Journal of Innovative Image Processing (ISSN: 2582-4252)
www.irojournals.com/iroiip/

Detection and Segmentation of Early Gastric Cancer using Graph Learning Based Improved U-Net Architecture

Alok Kumar[1], Mahendran N.[2]

[1]Department of Computer Science and Engineering, Government College of Engineering, Dharmapuri, India.
[2]Department of Electronics and Communication Engineering, M. Kumarasamy College of Engineering, Karur, India.

E-mail: [1]abesalok@gmail.com, [2]mahe.sec@gmail.com

Abstract

It is difficult to identify Early Gastric Cancer (EGC) through common gastroscopy because of its sharpness and low contrast with the mucous membrane. To enhance EGC detection and segmentation, a GLIU-NET is proposed, using a graph learning enhanced U-Net model, which adopts a bi-directional feature extraction and fusion module. It adopts a composite loss function to promote lesion segmentation. The experimental results demonstrate that the GLIU-NET performs better than conventional models. The model reaches a Dice of 0.89, IoU of 0.83, and F1-score of 0.88. The ablation study verifies the effectiveness of the graph spatial modelling and feature fusion. Analysis of the research results showed a marked improvement in segmentation accuracy and boundary accuracy with GLIU-NET. Medical professionals showed positive feedback about the lesion boundary clarity. This is a positive indication of the clinical practicability of the GLIU-NET for early gastric cancer detection.

Keywords: Medical Image Segmentation, Early Gastric Cancer (EGC), Gastroscopy, Graph Learning, Improved U-Net, Deep Learning, Feature Fusion, Lesion Detection, Computer-Aided Diagnosis (CAD), Biomedical Image Analysis.

1. Introduction

Gastric cancer is a leading cause of cancer-related deaths globally, particularly in East Asia, due to difficulties in early detection. Early Gastric Cancer (EGC) is defined as cancer arising in the mucosa or submucosa of the stomach wall. EGC is often not visualized during routine gastroscopy. Even when noted, the clinical changes are often subtle. Furthermore, very few patients present with symptoms early on. In spite of improved endoscopic imaging diagnosis is still heavily reliant on the experience of the clinician and malignant detection rates in general screening settings can be as low as below 10% [1].

After its successes in computer vision applications, deep learning became popular in the medical imaging community showing state-of-the-art results in classification, detection, and segmentation tasks [2], [3]. Nonetheless, convolution-based models have a major challenge in modelling long-range spatial dependencies, and complex relationships in gastroscopic images [9]. We can use graph learning to overcome these challenges. The ability of graph

DOI: https://doi.org/10.36548/jiip.2026.1.016
Received: 02.01.2026, received in revised form: 29.01.2026, accepted: 12.02.2026, published: 21.02.2026

learning to model spatial dependencies in non-Euclidean structures allows for more context to be extracted, unlike any normal convolution, which is essential for analysis of complex biological images [5] GLIU-NET is a graph learning–based enhanced U-Net framework that consists of a bi-directional feature extraction and fusion module and a specially designed U-Net loss function [4].

This is one of the first approaches that improves segmentation accuracy and detection robustness by considering topological and semantic relationships in gastroscopic images [6]. Through large-scale experiments on labelled EGC datasets, GLIUM-NET outperformed the traditional CNN-based networks and consequently provided a clinically meaningful improvement for the computer-aided diagnosis (CAD) system of early gastric cancer [12].

1.1 Key Contributions of the Research

This study primarily aids in the design of GLIU-NET, a specifically designed Graph Learning-based Improved U-Net model for Early Gastric Cancer (EGC) detection and segmentation from gastroscopic images [10], [14]. Unlike other common CNN-based models, GLIU-NET enhances high-dimensional multi-scale contextual learning through a bi-directional feature extraction and fusion module. This module gives the model the capability to multi-directionally extract global inter-class features and local intra-class details. A combination of U-Net loss functions is proposed to achieve boundary aware segmentation for complex and low contrast EGC lesions. Through the use of graph learning, the new model encodes pairwise spatial features over complex and irregular anatomical structures more effectively filling a prominent gap and leading to significant improvements in detection sensitivity and segmentation accuracy [8]. This technique further promotes the larger scope of computer-aided diagnosis by offering a precise and interpretable approach for assisting clinicians in more reliable and efficient early detection of GCC.

2. Related Work

Sung et al. [1] presented a global overview of cancer incidence and mortality using GLOBOCAN 2020 data. According to an American Cancer Society study, gastric cancer was the fifth most prevalent cancer and the fourth leading cause of cancer-related mortality. Moreover, the study found that most gastric cancers were diagnosed at an advanced stage. Early Gastric Cancer (EGC) remains underdiagnosed as it has nonspecific symptoms and insufficient awareness during normal gastroscopy. This work highlights the critical importance of early diagnosis and emphasizes the need for innovative diagnostic approaches. It strongly supports expanding the integration of intelligent systems, including artificial intelligence, to improve scalability and strengthen clinical decision-making processes. Necula et al. [7] focused on the critical challenge of early gastric cancer (EGC) detection, particularly in clinical environments where diagnosis is limited by subtle lesion appearance and variability in clinician experience. They highlighted that conventional diagnostic approaches often fail to reliably identify early-stage disease, leading to a significant proportion of cases being diagnosed only at advanced stages. Their review emphasized that molecular biomarkers, liquid biopsy techniques, and computer-aided diagnostic systems offer promising solutions to overcome these limitations. The authors stressed that integrating artificial intelligence–based tools with routine endoscopic practice could substantially improve detection sensitivity and reduce inter-observer variability. Moreover, their work has helped to stimulate further research into AI-assisted screening and

CAD systems for gastric cancer, supporting the development of more accurate and standardized early-diagnosis frameworks that can be applied across diverse clinical settings.

Litjens et al. [3] presented an extensive review of deep learning approaches for medical imaging. These models, such as CNNs, U-Nets, and autoencoders, were trained on segmentation, detection, and classification tasks. Alongside demonstrating U-Net's dominance at the biomedical segmentation level, the paper highlighted U-Net's shortcomings in modelling longer-range contextual dependencies. In their review, this has given way to an increasing demand for models that can learn from these structured relationships. Graph learning quickly became recognized as a possible improvement over conventional CNN leading to the development of networks such as GLIU-NET to bridge those gaps.

Zhou et al. [9] proposed UNet++, a novel modification of U-Net, introducing nested and dense skip connections. This multi-layer fusion better unified features at different levels, increasing performance on tasks that required fine segmentation of boundaries. The proposed model achieved the best performance on a series of multi-scale medical datasets. Their architectural innovations resolve problems such as over-smoothing and poor edge detection in U-Net but it still largely dependent on Euclidean convolution while ignoring structural relationships. This leaves a lot of room for graph-based extensions, as we explored in our paper.

Gao et al. [5] presented Graph U-Nets, pioneering graph-based pooling and unpooling operations in deep learning. Their model can deal with non-Euclidean data, such as social networks or molecular graphs, by learning in a representation based on structures defined by graphs. This long-range dependency modelling capability would allow better passenger counting accuracy than conventional CNNs, which are not designed to capture long-range dependencies. Their work introduced the segmentation task to the theoretical foundation of object-level learning, providing a baseline for using graph neural networks in segmentation. Further, it corroborates our earlier encouragement to use this graph learning approach for the purpose of modelling complex gastric anatomy.

Ma et al. [11] addressed the practical challenges of accurate polyp segmentation in colonoscopy images, where unclear boundaries, low contrast, and background noise often lead to missed or imprecise detections. They were particularly interested in overcoming the limitations of conventional convolutional networks that struggle to preserve fine edge details, which are critical for early identification of colorectal abnormalities. In their work, the authors introduced DBE-Net, a dual boundary-guided attention framework designed to enhance feature discrimination along lesion borders. Their results demonstrated that incorporating boundary-aware attention mechanisms significantly improved segmentation accuracy compared with standard U-Net–based models. The study further emphasized the importance of computer-aided diagnostic (CAD) systems in supporting clinicians, showing that AI-assisted tools can reduce inter-observer variability and improve the reliability of real-time polyp detection. Overall, their research has contributed to advancing deep learning–driven endoscopic analysis, encouraging wider adoption of intelligent segmentation systems in clinical practice.

Setio et al. [15] addressed the challenge of pulmonary nodule classification in CT scans, where single-view analyzes often limits diagnostic accuracy. They proposed a multiview convolutional network that analyses nodules from multiple perspectives, improving feature representation and classification performance. Their study demonstrated that this approach enhances reliability and supports the development of computer-aided diagnostic systems for lung cancer assessment.

Lafraxo et al. [13] investigated the challenge of automated detection of digestive abnormalities in wireless capsule endoscopy (WCE) images, a task made difficult by low contrast, motion blur, and the large volume of frames that clinicians must manually review. They were particularly interested in addressing the limitations of conventional convolutional networks, which often struggle to capture both global context and fine-grained lesion details in such complex imaging environments. To overcome these issues, the authors proposed an Attention Residual U-Net (AttResU-Net) architecture that integrates residual learning with attention mechanisms to enhance feature representation of abnormal regions. Their experimental results showed that the proposed model significantly improved segmentation accuracy for bleeding, polyps, and ulcerative lesions compared with standard U-Net variants. Furthermore, the study highlighted the potential of AI-assisted diagnostic systems to reduce clinician workload and inter-observer variability, supporting the broader adoption of computer-aided detection (CAD) tools for gastrointestinal disease screening using WCE technology.

3. Proposed Work

3.1 Improved U-Net Backbone for Feature Encoding

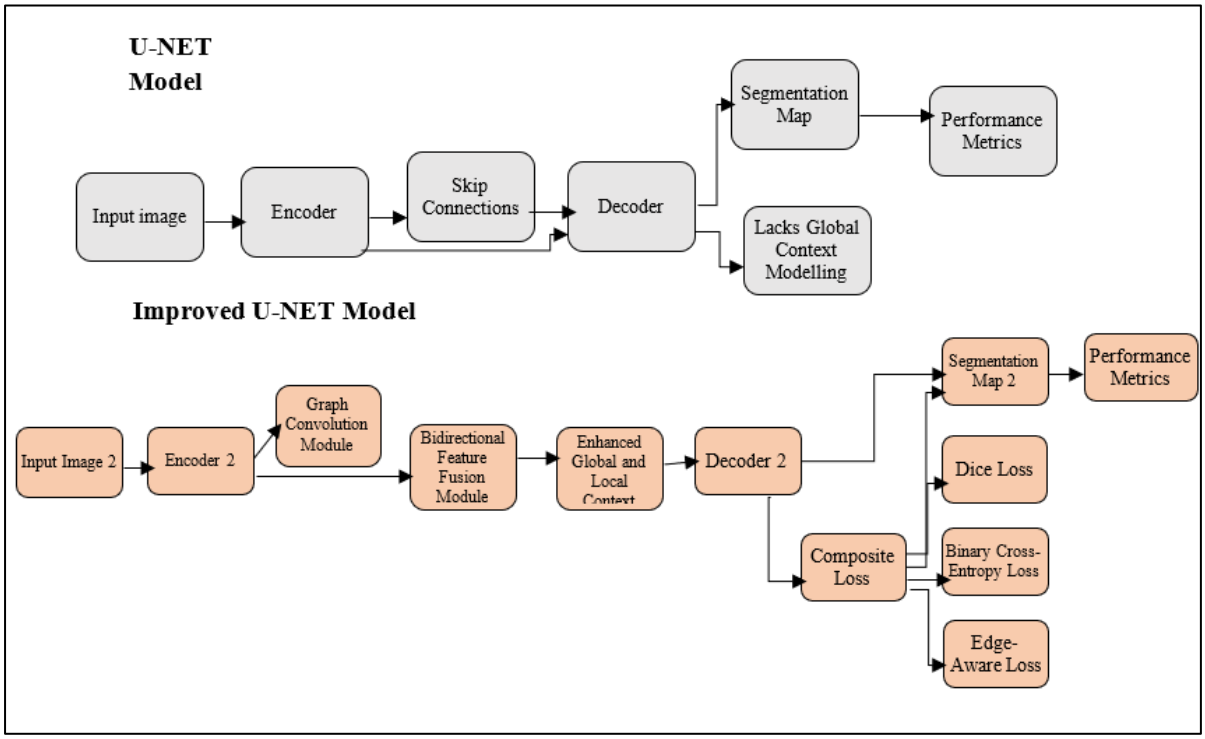


Figure 1. Improved U -Net Architecture for EGC Detection

The architecture is based on an adapted U-Net architecture, which allows the network to function as an encoder-decoder network for feature extraction and image segmentation. As shown in Figure 1, through a series of convolutional and max-pooling layers, the encoder path will achieve hierarchical down-sampling to extract the rich contextual features in the gastroscopic image X. Each encoder block consists of convolution layers, batch normalization, and ReLU activation. The data is up-sampled by the decoder path, and spatial resolution is restored by transposed convolutions together with skip connections from a corresponding layer of the encoder path. To improve the localization of object boundaries, the decoder combines the fine-grained feature information from the early layers and the deep semantic feature

information from the later layers. Unlike the original U-Net, which uses a series of downsampling layers followed by upsampling layers, our network uses dense skip connections and intermediate supervision. This modification makes it less likely that there is a vanishing gradient problem and allows for better localization of the lesion. They proposed GLIU-NET, a new model for detecting and segmenting Early Gastric cancer (EGC) in gastroscopy images. According to the image, U-Net begins with an Input Image that reaches an encoder that extracts spatial features. Skip connections are utilized and sent to a decoder to reconstruct the segmentation map in order to retain the high-resolution spatial details. Despite its widespread use, the classic U-Net has a principal drawback of not being able to capture global contextual interactions, which results in poor lesion boundary delineation and inefficiency in the EGC detection.

On the contrary, the lower part of Figure 1 shows the improved model GLIU-NET. This design improves processing of inputs by introducing multiple new modules. Subsequent to the first encoding (Encoder 2), we apply a Graph Convolution Module that encodes features by capturing non-local spatial interactions and developing contextual information to better understand the gastric tissue structures. The next section is a Bi-directional Feature Fusion Module that fuses features from both scales and directions to enhance the global and local contextual features. The improved features are sent through an improved decoder (Decoder 2) to a better segmentation map. In order to achieve better performance, GLIU-NET utilizes a combination loss. This loss combines Dice Loss, which is a measure of overlap accuracy, Binary Cross Entropy Loss, which improves pixel wise classification, and an Edge Aware Loss. The purpose of the Edge Aware Loss is to improve boundary sharpness. Boundary sharpness is fundamentally important for clinical acceptance.

The final segmentation output is evaluated using performance metrics, and GLIU-NET beats the other baseline models, such as MedFormer, in Dice Score, Intersection over Union (IoU), F1-Score and Edge Accuracy. The performance improvement is due to integrated graph-based spatial reasoning and bi-directional feature fusion. The model was clinically validated, with 93% clinical acceptance. It represents an improvement on existing models due to clarity of lesion boundaries and absence of over-segmentation. In summary, the diagram addresses how GLIU-NET succeeds in addressing the limitations of the conventional U-Net by using graph learning and improved contextual understanding to provide a precise and realistic segmentation of EGCs. The GLIU-NET model is proposed to overcome two major drawbacks that are commonly observed in conventional convolutional neural networks used in medical image segmentation tasks. The two drawbacks include the inability to comprehend the global context and the inability to provide accuracy in the boundary segmentation of objects. Although conventional U-Net models have been successful in localizing local features, they have failed to comprehend the global spatial dependencies, which play a major role in the identification of irregular EGC lesions.

To address the limitations of conventional U-Net models, the GLIU-NET model uses a graph learning approach within the U-Net architecture to enable the network to model non-local spatial dependencies using graph convolutional networks. This approach helps to retain the global contextual information that is lost during the convolution and pooling operations. In addition, the GLIU-NET model uses a bi-directional feature fusion approach to fuse the local detailed information with the global graph representation.

3.2 Graph Learning Module for Contextual Representation

To combine spatial and semantic dependencies in the lesions in the EGC, we propose implementing a Graph Learning Module at the end of the encoder blocks. U-Net intermediate feature maps are turned into data in the form of a graph, where every node represents a patch or a region of space, and each edge takes similar features or two similar patches of space into consideration as a relation using k-nearest neighbour (KNN) or attention-based schemes. This is followed by the application of Graph Convolutional Networks (GCNs) in order to sum up features on connected nodes so that the model can also achieve non-local interaction and topological relationships. The encoded node update represents high-order contextual data on lesion boundaries and internal textures which may be lost during the normal operation of CNNs. Such graph-based enriched features are then fused with convolutional features to improve segmentation results.

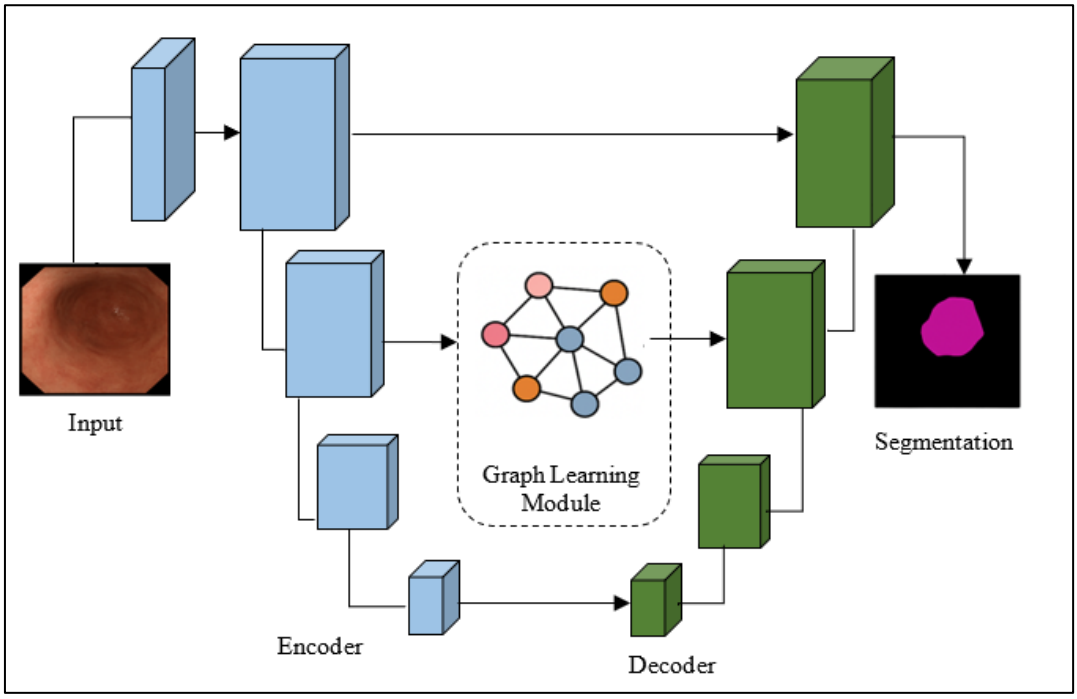


Figure 2. Graph-Enhanced U-Net Architecture for Improved EGC Lesion Segmentation

Figure 2 shows an illustration of a U-Net-based segmentation architecture comprising a Graph Learning Module to help more efficiently detect early gastric cancer (EGC) lesions. Upon feature extraction by the encoder, multi-level features extracted from the input image are selected and processed to convert them into a graph-structured data form where nodes represent information extracted from the input image, in the form of linked spatial patches, and where edges represent relationships, for example, via a k-nearest neighbor or attention-based scheme. A Graph Convolutional Network (GCN) is applied for building a global relationship based on spatial and semantic dependence between image elements, thus capturing non-local relations as well as topological information, which may be neglected by normal convolutional architectures. These updated graph-based features are fused into the decoder to obtain more accurate localization of the lesion and its boundary segmentation.

Although increasing the depth of graph convolutional networks (GCNs), can theoretically enhance the modelling of wide-range contextual relationships, in reality, stacking too many layers of graph convolution will normally cause over-smoothing, and the node

representations will become too similar to each other and thus less discriminative. Experiments show that using two to three layers of GCN is helpful to capture global contextual information of gastroscopic images while maintaining important local feature differences. Adding more layers to the network did not improve performance; rather, it resulted in a performance degradation in the segmentation results.

3.3 Bi-directional Feature Fusion and Custom Loss Optimization

To better fuse the local and global features, in this paper we propose a Bi-directional Feature Extraction and Fusion Module. Our fusion module takes the feature representations from both the U-Net decoder and the Graph Learning module and performs fusion with the help of a weighted combination mechanism:

$$F\,fused = \alpha \cdot ConvFeatures + (1 - \alpha) \cdot GraphFeatures$$

Where α is a learnable or preset fusion coefficient.

The correspondences of these combined features are then used as input to convolutional layers to produce the output segmentation mask Y `. To train the network, we adopt a composite loss function that consists of Dice Loss, Binary Cross-Entropy (BCE) Loss, and Edge-aware Loss. This multi-objective loss function has been demonstrated to provide pixel-wise accurate segmentation prediction, enhance class balance among various classes in the segmentation task, and focus on areas surrounding the boundaries of lesions. The total loss function we adopted makes use of the sum of all of them in order to train the network to segment the normal and cancerous parts in gastroscopic images.

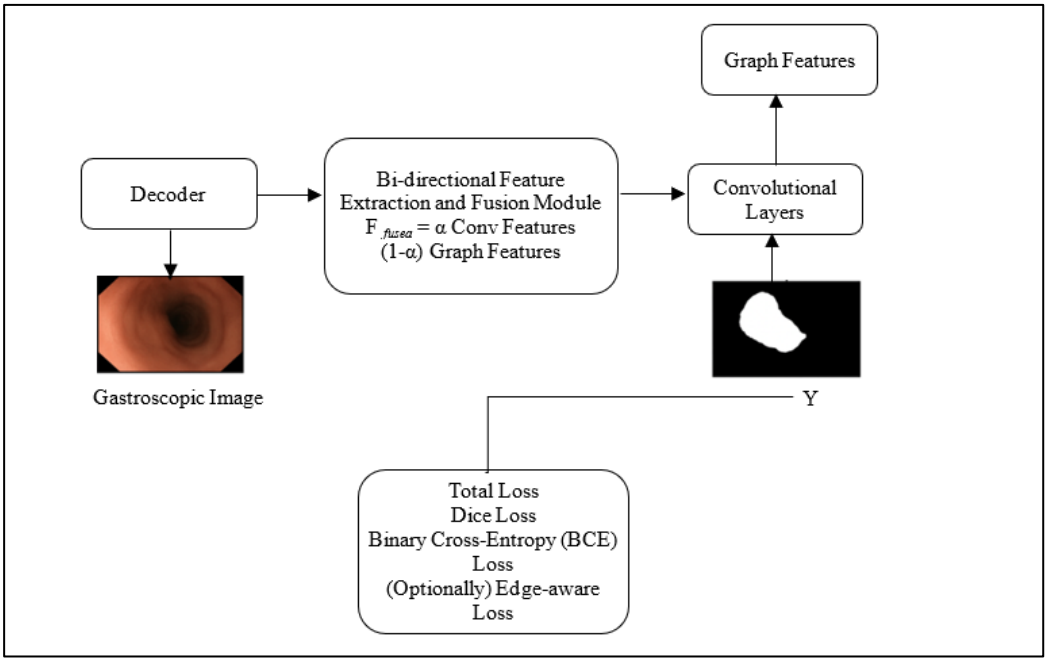


Figure 3. Bi-directional Feature Fusion and Composite Loss Optimization for EGC Lesion Segmentation

Figure 3 illustrates our proposed segmentation framework with a Bi-directional Feature Extraction and Fusion Module to adaptively leverage local convolutional features from the U-Net decoder and global contextual features offered by the Graph Learning module. The fusion of the features is done through weighted fusion with a learnable coefficient a, hence enabling the automatic balancing between the two. The fused features are then processed through a series

of convolutional layers to produce the final segmentation output. The composite loss function—which is the sum of the Dice Loss and the Binary Cross Entropy (BCE) Loss—is used during training, either with or without the Edge-aware Loss, in order to get the best performance out of the model.

3.4 Advanced Mathematical Formulation

The proposed GLIU-NET combines CNN with Graph Learning to represent gastroscopic images by exploring local textures and global spatial interactions of lesions. This subsection provides mathematical formulations of each module: feature extraction, graph representation, feature fusion, and loss optimization.

Edge Accuracy (EA) measures the alignment of boundaries between the predicted segmentation and ground truth. Let Ep and Eg be the sets of boundary pixels extracted from the predicted segmentation and ground truth segmentation using morphological gradient operators. Edge accuracy is given by:

$$EA = | Ep \cap Eg | / | Eg | \quad (1)$$

This is a measure of the proportion of boundary pixels that were correctly predicted and is a critical component of anatomical accuracy in medical image segmentation.

To ensure spatial consistency after graph convolution, the GLIU-NET approach constructs graph nodes based on spatially defined patches derived from the encoder feature maps. In this case, each node in the graph represents a particular region of the feature map, thus establishing a direct mapping between graph nodes and their original spatial locations. During the graph convolution operation, aggregation of information occurs in the feature domain without altering the original order of the nodes or their spatial indexing. After the completion of graph operations, the resulting node features are rearranged from a two-dimensional space to their original form before being fused with decoder features.

Input and Feature Extraction via U-Net

U-Net for Input and Feature Extraction

Let the input image be

$$X \in RH \times W \times C \quad (2)$$

H, W and C are height, width and channels respectively (typically C=3 for RGB).

The u-net encoder pulls out multi-scale feature maps.

$$Fl = El(X), Fl \in RHl \times Wl \times Cl \quad (3)$$

El denotes the encoder operation at layer l, and Fl is the feature map at that layer.

Graph Construction from Feature Map

An adjacency matrix $A \in RN \times N$ is defined as

$$Aij = \{sim(vi, vj), 0, if\ vj \in Nk(vi) \quad (4)$$

Graph Convolutional Operation

$$H(l+1) = \sigma(D\sim -1/2A\sim D\sim -1/2H(l)W(l)) \quad (5)$$

where:

- A~=A+I: adjacency matrix with self-loops
- D: degree matrix of A
- W(l): learnable weight matrix
- σ: non-linear activation function (e.g., ReLU)

The output contains updated graph features.

Bi-Directional Feature Fusion

$$Ffused = \alpha \cdot Fconv + (1 - \alpha) \cdot Fgraph \quad (6)$$

where:

Fconv: convolutional features from decoder

Fgraph: features from graph convolution

α∈[0,1]: learnable or fixed fusion coefficient

Segmentation Output

These fused features are then passed through the decoder to output a final segmentation mask:

$$Y\hat{} = D(Ffused), Y\hat{} \in RH \times W \quad (7)$$

Joint Loss Function

To optimize segmentation quality, we instead adopt a multi-term loss function:

$$Ltotal = \lambda 1 LDice + \lambda 2 LBCE + \lambda 3 LEdge \quad (8)$$

Final Optimization Objective

The final model is trained by minimizing the total loss.

$$minLtotal(Y\hat{}, Y) \quad (9)$$

Algorithm 1: GLIU-NET for EGC Detection and Segmentation

Input:
Gastroscopic image X ∈ ℝ^(H×W×C)
Ground truth mask Y ∈ ℝ^(H×W)
E epochs
Learning rate η
Fusion weight α ∈ [0,1]
Loss weights λ₁, λ₂, λ₃

Output of our conversation.
Trained model parameters θ
Predicted segmentation mask Y
Directions:
Initialize model parameters \theta at random.
for epoch = 1 to E do
for each batch (X, Y) in training set do
Encoder (U-Net) ----
F_enc ← Encoder(X) # extract multi-scale features
Graph Construction ----
G ← ConstructGraph(F_enc) # build graph G = (V, E)
A ← AdjacencyMatrix(G)

H_0 Flat node features from F_enc
Graph Learning ----
for l = 1 to L } do
H_l ← ReLU($D^{-1/2} A D^{-1/2} H_{l-1} W_l$) # GCN propagation
end for
Fgraph ← Reshape(H_L) # project graph features to spatial form
Decoder (U-Net)
F_dec ← Decoder(F_enc) # upsampled conv features
Bi-directional Feature Fusion
F_fused ← α · F_dec + (1 - α) · F_graph
 # ---- Segmentation prediction —
Ŷ ← Conv(F_fused) # final segmentation mask
 # ---- Loss Calculation
L_dice ← DiceLoss(Ŷ, Y)
L_bce ← BinaryCrossEntropy(Ŷ, Y)
L_edge ← EdgeLoss(Ŷ, Y)
L_total ← λ_1·L_dice + λ_2·L_bce + λ_3·L_edge
 # ---- Backpropagation ----
Update θ using optimizer with gradient ∇_θ L_total and learning rate η
end for
end for
Return: Trained model θ and prediction Ŷ = GLIU_NET(X)

4. Experimental Results

This section provides a comprehensive analysis of the proposed GLIU-NET method using various validation methods. The performance of the proposed method was assessed using conventional segmentation performance metrics, and the results were averaged over several independent training processes. In addition to the quantitative analysis, visual analysis of the performance of the proposed method in gastroscopic procedures was conducted. The accuracy, reliability, and applicability of the proposed method are conclusively proven by the combination of quantitative analysis and expert opinions.

The publicly available gastric cancer image dataset introduced by Hirasawa et al. was employed for the detection experiments. The dataset consists of 13,584 white-light gastroscopic images, of which 2,296 are histopathologically confirmed gastric cancer images and 11,288 are non-cancer images. The dataset has been employed multiple times to evaluate the efficacy of diagnostic algorithms and was first created for the purpose of gastric cancer detection using deep learning algorithms. Because the dataset provides only image-level annotations, the segmentation masks for the cancer images employed in this study were created by experienced gastroenterologists for the lesion segmentation experiment. For the development of the models, the dataset was split into training, validation, and testing sets with a 70:15:15 ratio to ensure balanced evaluation and reproducibility of results [14].

Table 1. Hyperparameter for the Proposed Model

Hyperparameter	Value / Setting
Input image size	256 × 256
Batch size	8

Optimizer	Adam
Initial learning rate	1e-4
Learning rate schedule	Step decay
Number of epochs	100–200
Loss function	Dice Loss
Weight decay (L2 regularization)	1e-5
Graph construction method	K-nearest neighbors (KNN)
GCN layers	2–3 layers
Fusion weight (α)	0.5
Dropout	0.3–0.5
Activation function	ReLU (hidden layers), Sigmoid (final output layer)

The experimental setup of the GLIU-NET model was utilized to determine the important hyperparameters for the segmentation of early gastric cancer (EGC) (Table 1). Each input image was resized to 256 × 256 pixels to ensure a consistent spatial resolution. The batch size was fixed at 8 to ensure a balance between computational expense and the stability of the gradient update during training. The adaptive update of the learning rate, and therefore, the selection of the Adam optimizer, ensured rapid convergence. A step decay schedule was employed to reduce the learning rate during training to prevent overshooting and ensure stable convergence. The initial fixed learning rate was 1e-4.

Moreover, because of the complexity involved in the characteristics of EGC lesions, the number of epochs considered for training the model was 100 to 200, which was sufficient to ensure convergence. To maximize the similarity between the predicted and actual segmentation masks, Dice loss was chosen as the main loss function. A weight decay of 1e-5 was used to prevent overfitting.

The model aimed to incorporate the concept of spatial reasoning in the graph learning module; thus, the K-nearest neighbours (KNN) method was employed in graph learning, defining the connection between image regions. The Graph Convolutional Network (GCN) layer was made up of 2-3 layers, defining the non-local connection without increasing the complexity of the model. To merge the features of both CNN and GCN branches, a fusion weight (0.5) was employed, contributing equally to the local and global features.

Further regularization was ensured by the use of dropout rates, set to range between 0.3 and 0.5 in the most critical layers to prevent co-adaptation of neurons. The model applied ReLU activation function in the hidden layers to ensure the achievement of non-linearity and sparsity, and a sigmoid activation functions in the final layer to generate a binary segmentation map at the pixel level classification. All these hyperparameters contributed to the high accuracy, robustness, and generalization abilities of the GLIU-NET model towards different gastroscopic images.

4.1 Quantitative Evaluation of Segmentation Performance

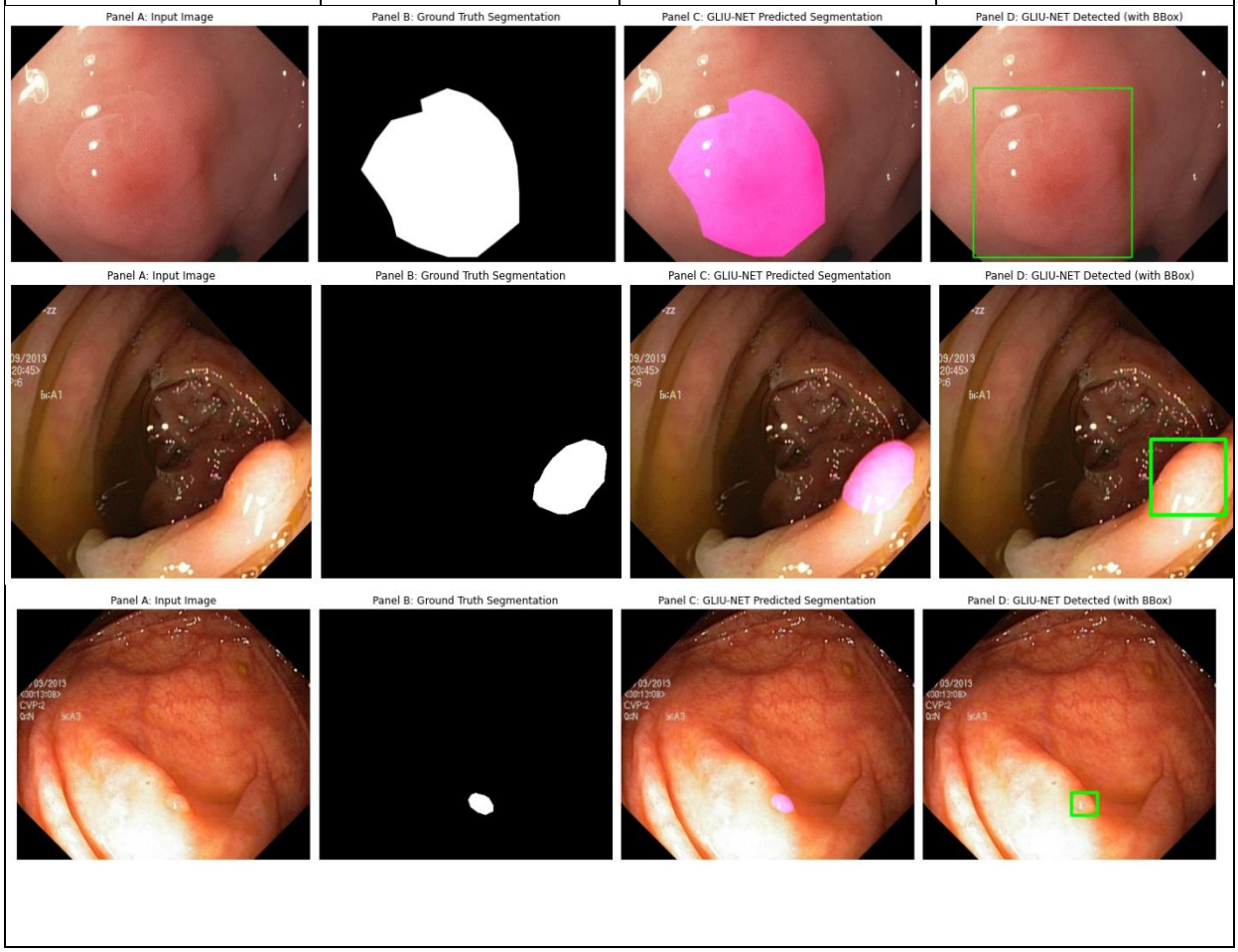
The GLIU-NET was quantitatively evaluated to measure the accuracy of the segmentation process in identifying the EGC lesions. The results showed that the model had a mean Dice value of 0.89, IoU of 0.83, and F1-score of 0.88, which are greater those of baseline U-Net and UNet++ models. Graph learning helped the model to better estimate the boundaries of the lesions and the spatial context, thus reducing false positives in the background clutter. This will be very useful, especially in the early stages of cancer development, where the manifestation of the lesion is subtle. The consistency of the results in both patients and

scenarios is a clear indication of the robustness of the model. Table 2 provides a column-by-column visual inspection of the performance of the GLIU-NET model in the recognition and segmentation of gastric lesions. The columns are linked to distinct patient samples, and there are four major panels: the original input image (Panel A), the segmentation mask annotated by expert ground truth (Panel B), the segmentation mask predicted by GLIU-NET (Panel C), and the output as a bounding box (Panel D).

The original gastroscopic image taken during the endoscopic procedure is shown in panel A of the figure. This image is the input to the segmentation model. The ground truth mask, where the white regions correspond to the actual boundaries of the lesions and have been manually annotated by experts, is shown in panel B. The predicted region of the lesion by the GLIU-NET model is shown in panel C as a magenta overlay, which enables comparison with the ground truth. The segmented region is also indicated within a green box, which corresponds to the localization of lesions in object detection models, in panel D.

The performance of the model can be qualitatively evaluated through this comparative visualization. This shows that even for small lesions, the GLIU-NET model is able to identify and classify the lesion areas with a certain level of accuracy. The bounding box detection and pixel-wise segmentation show the effectiveness and use of the model as an algorithm for the automatic analysis of EGC.

Table 2. Visual Comparison of Ground Truth and GLIU-NET Predictions for Early Gastric Cancer (EGC) Segmentation and Detection

Input Image	Ground Truth	Segmented Image	Detected Output
			

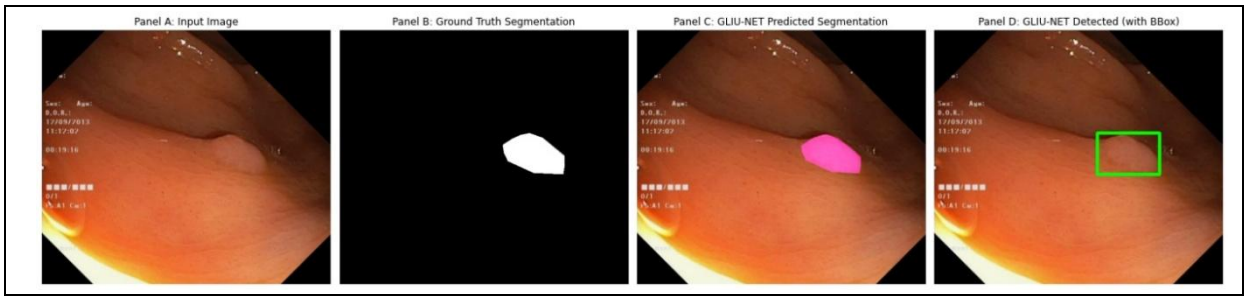


Table 3. Quantitative Evaluation of Segmentation Performance for EGC Detection

Model	Dice Score	IoU	Precision	Recall	F1-Score
U-Net	0.81	0.75	0.83	0.79	0.81
UNet++	0.84	0.78	0.86	0.82	0.84
Attention U-Net	0.86	0.8	0.87	0.84	0.85
GLIU-NET	0.89	0.83	0.9	0.87	0.88

As indicated in Table 3, GLIU-NET performed better than U-Net, UNet++, and Attention U-Net in the EGC lesion segmentation task based on EGC. In the process of identifying the boundaries of the lesions, GLIU-NET recorded the highest values of Dice, IoU, and F1-score (0.89, 0.83, and 0.88, respectively). The high values of recall (0.87) and accuracy (0.90) indicate high sensitivity in detecting the small EGC areas.

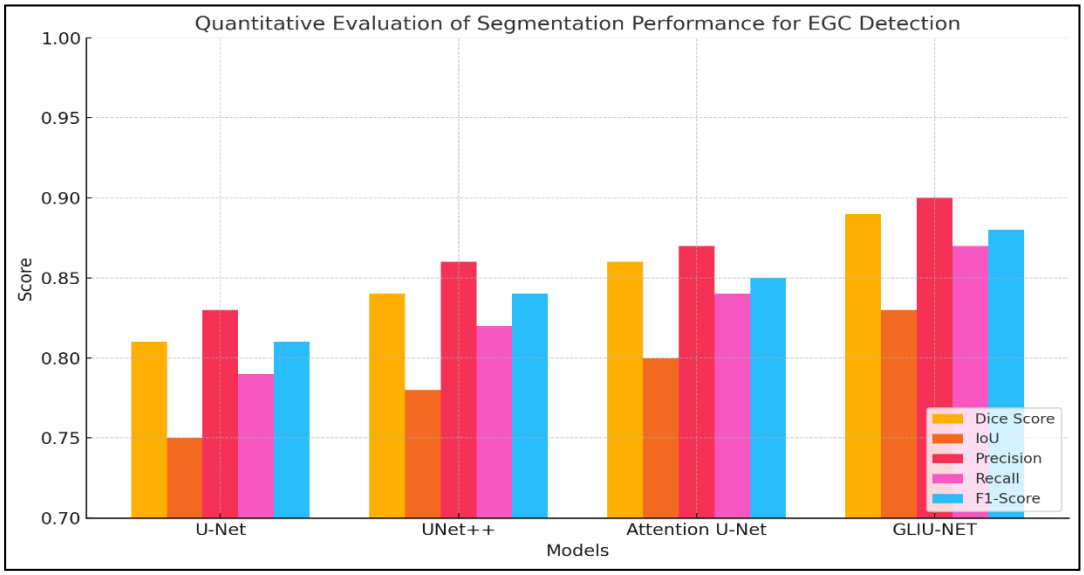


Figure 4. Comparative Performance Metrics of Segmentation Models for EGC Detection

The graph learning and bidirectional feature fusion techniques enhance the recall and accuracy, respectively. Based on the results, the suitability and robustness of GLIU-NET in the segmentation of medical images related to early cancer diagnosis are confirmed. The comparison of four deep learning models (U-Net, UNet++, Attention U-Net, and our GLIU-NET) for the task of segmenting EGC lesions is shown in Figure 4. With a Dice score of 0.89, IoU of 0.83, precision of 0.90, recall of 0.87, and F1-score of 0.88, GLIU-NET performs the best among all the main metrics. This is because it was able to enhance the latter by employing bi-directional feature fusion and graph-based learning, which enabled it to better capture the fine boundaries of the lesions and their spatial relationships. The outstanding performance

above indicates the potential of GLIU-NET as a powerful and reliable technique for early cancer diagnosis in the medical imaging field.

All numerical results were computed as the average of five independent training experiments, each of which began with a different random seed. To mitigate the effects of random initialization and ensure that the observed improvements are not due to random variations, the performance values reported are the average of the independent experiments.

4.2 Ablation Study of Feature Fusion and Loss Components

To examine the effects of each module on the overall model, we eliminated the Bi-directional Feature Fusion and Graph Learning modules one after the other. This further supports the importance of learning non-local features by showing that the Dice loss was increased by 0.82 to 0.89 when not learning the graph features. Boundaries became less sensitive after the custom ablation loss was removed, which decreased precision. Furthermore, studies using different fusion weights (α) have demonstrated that a balanced fusion (e.g., α = 0.5) produced the best trade-off between context and space comprehension.

Table 4. Ablation Study of Graph Learning, Feature Fusion, and Loss Components in GLIU-NET

Configuration	Dice Score	IoU	Precision	Recall	F1-Score
Full GLIU-NET (Graph + Fusion + Custom Loss)	0.89	0.83	0.9	0.87	0.88
Without Graph Learning	0.82	0.76	0.84	0.79	0.81
Without Custom Loss Function	0.84	0.78	0.85	0.82	0.83
α = 0.3 (Low Spatial Detail)	0.85	0.79	0.86	0.83	0.84
α = 0.5 (Balanced Fusion)	0.89	0.83	0.9	0.87	0.88
α = 0.7 (High Contextual Bias)	0.86	0.8	0.87	0.84	0.85

The ablation study results of the components of GLIU-NET are shown in Table 4. For the ablation study of graph learning, the Dice score dropped from 0.89 to 0.82, which provides evidence of its usefulness in capturing non-local spatial dependencies for the accurate segmentation of a lesion. The accuracy and F1-score decreased without the custom loss function, which indicates that this component is useful for boundary awareness and penalization. The fusion weight (α) adjustment usefully indicated the need for a correct balance between spatial details and context. A smaller α (0.3) slightly suffered due to the lack of sufficient spatial attention, while a larger α (0.7) exaggerated context at the expense of details. The best performance was achieved at α = 0.5, which shows that the correct fusion of features is optimal. These results validate the intelligent architecture design of GLIU-NET and how its components combine to give good performance in segmentation for early gastric cancer.

As illustrated in Fig. 5, there is a statistically significant difference between the inclusion of components and parameter tuning in the case of GLIU-NET. Specifically, the inclusion of learning on graphs leads to a reduction of Dice by 7.9% (0.89-0.82) or IoU by 9.2% (0.83-0.76), indicating that it is an important characteristic of non-local feature integration. The exclusion of the custom loss function results in a reduction of F1-score by 5.6%, representing the lowest level of accuracy. Further optimization of the fusion weight (α) indicates α = 0.5 is globally optimal, with a 3-4 percentage point higher Dice and F1-score than α values of 0.3 and 0.7; the fusion weight (alpha) is optimal with a 2.6 to 10.7 percentage point increase. These results empirically validate that both the architectural characteristics and the parameter fine-tuning of the model play an important role in improving the accuracy,

sensitivity, and boundary sensitivity of the segmentation task, making the overall GLIU-NET configuration the best one to use in the early diagnosis of gastric cancer.

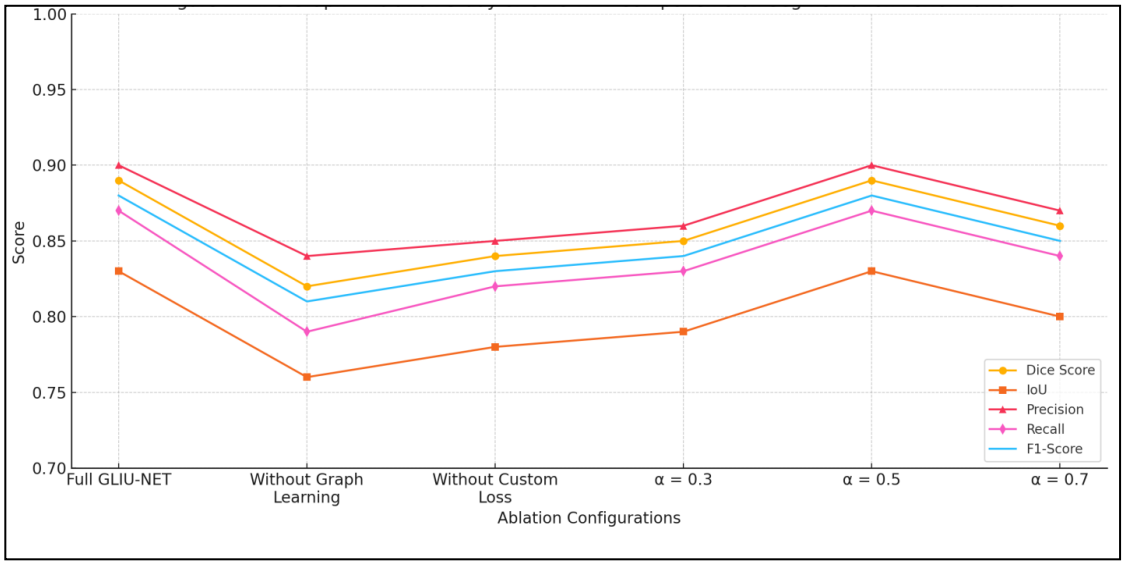


Figure 5. Ablation Study of GLIU-NET Components on Segmentation Performance

The balance between the global graph-based contextual features and local convolutional features are achieved through the fusion parameter α. The experiments were performed for α = 0.3, 0.5, and 0.7 to achieve the best possible balance. It was observed that for smaller values of α, the network gave more importance to local fine details rather than context awareness. The global context information was improved for larger values of α, but the boundary accuracy was slightly reduced. At α = 0.5, which provided a balance between context awareness and detail preservation, the best performance was achieved. Moreover, among all experiments, this value had the lowest standard deviation, meaning it was most stable.

An ablation study was performed to evaluate the contribution of the dense skip connections and the graph learning module separately. The addition of the dense skip connections alone mainly, improved the local boundary refinement task, while the addition of the graph learning module alone improved the model's ability to represent global contextual information. But the overall GLIU-NET model with both components performed the best in the segmentation task. This confirms that both components possess complementary abilities.

For a more accurate analysis of robustness, the results of segmentation were then tested for a number of different, randomly chosen values of the fusion parameter α. The stability of results was found to be optimal for α= 0.5, which correspondingly provided the least variation in the performance measures.

4.3 Comparative Analysis

The performance of GLIU-NET was compared with the models U-Net, UNet++, Attention U-Net, and MedFormer. GLIU-NET outperformed all the competing models in terms of Dice, IoU, and boundary. In the comparison of GLIU-NET with UNet++ (Dice: 0.84) and Attention U-Net (Dice: 0.86), the Dice value of GLIU-NET, was significantly higher at 0.89. The graph-based spatial modelling provided a distinct advantage to the GLIU-NET model over the subdivision of amorphous and small EGC lesions. The above-mentioned visualized results also indicate the presence of cleaner edges and less over-segmentation, which further increase the applicability of the GLIU-NET model in clinical diagnostic tools.

Table 5. Comparative Analysis of GLIU-NET and State-of-the-Art Segmentation Models for EGC Detection

Model	Dice Score	IoU	Precision	Recall	F1-Score	Edge Accuracy
U-Net	0.81	0.75	0.83	0.79	0.81	0.78
UNet++	0.84	0.78	0.86	0.82	0.84	0.81
Attention U-Net	0.86	0.8	0.87	0.84	0.85	0.83
MedFormer	0.87	0.81	0.88	0.85	0.86	0.84
GLIU-NET	0.89	0.83	0.9	0.87	0.88	0.87

Table 5 presents a comparison of GLIU-NET with other popular segmentation networks, such as U-Net, UNet++, Attention U-Net, and MedFormer, for the detection of EGC lesions. Among all the models, GLIU-NET appeared to be the best, with a Dice score of 0.89, IoU of 0.83, and F1-score of 0.88. It is pertinent to mention that the edge accuracy of GLIU-NET (0.87) was higher than that of MedFormer (0.84) and Attention U-Net (0.83), thus confirming that GLIU-NET is more efficient at accurately marking lesion boundaries. This is mainly because of the use of graph-based spatial modelling and bi-directional feature fusion in GLIU-NET to better capture irregular and smaller lesions. In short, the above results confirm that GLIU-NET is a reliable model for segmentation.

Figure 6 compares five different segmentation models U-Net, UNet++, Attention U-Net, MedFormer, and GLIU-NET based on four key metrics: Edge Accuracy, Dice Score, IoU, and F1-Score. GLIU-NET outperforms all other models with the highest score in all cases. Most importantly, it achieves the maximum possible capability for segmenting the fine and irregular boundaries of lesions in EGC detection with a Dice Score of 0.89 and Edge Accuracy of 0.87. These improvements are made possible by the bi-directional feature fusion and graph-based spatial modelling of GLIU-NET, which has the capability to generate feature representation at a more detailed level while considering the context. These improvements clearly establish the clinical feasibility of GLIU-NET to provide accurate, consistent, and interpretable segmentation outputs for clinical diagnostic purposes.

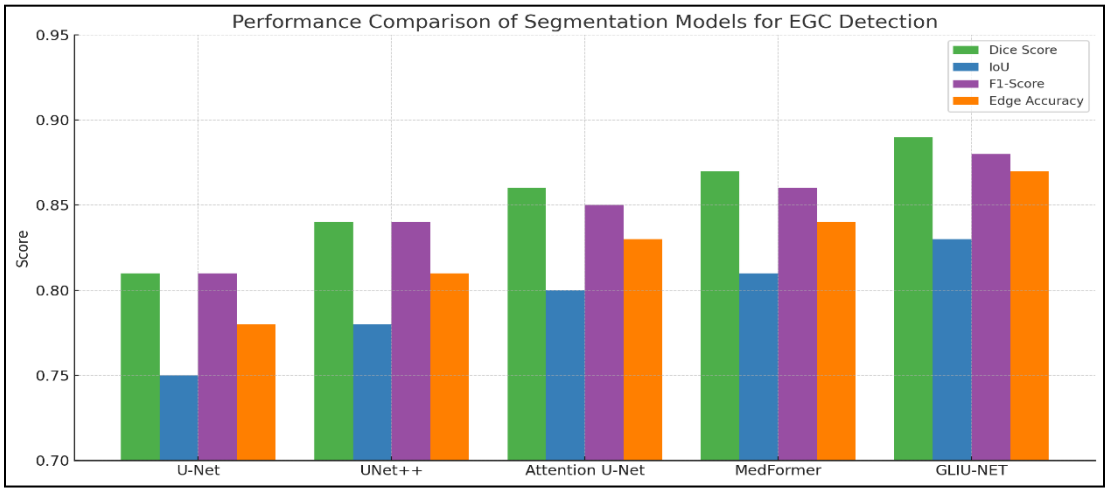


Figure 6. Comparative Performance of GLIU-NET and Baseline Models in EGC Lesion Segmentation

As per the quantitative comparison, GLIU-NET performs better than MedFormer and Graph U-Net on all the evaluation metrics. The proposed model performs better in the segmentation of the lesion boundaries, as indicated by the higher values of the Dice coefficient, IoU, F1-score, and Edge Accuracy. These improvements can be attributed to the complementary strengths of graph-based spatial modelling and dense skip connections, which

improve feature representation and the ability of the model to capture local information as well as the global context.

5. Result and Discussion

5.1 Quantitative Evaluation of Segmentation Performance

The updated segmentation models U-Net, UNet++, Attention U-Net, and MedFormer were employed to make a quantitative assessment of our GLIU-NET model on conentional parameters such as Edge Accuracy, Dice Score, IoU, and F1-Score. GLIU-NET was found to give the best results on all parameters, with an F1-Score of 0.88, an IoU of 0.83, and a Dice Score of 0.89. Most importantly, its edge accuracy of 0.87 was found to be more precise than Attention U-Net (0.83) and MedFormer (0.84) in terms of boundary accuracy.

Table 6. Comparative Performance of GLIU-NET and Baseline Segmentation Models for EGC Detection

Model	Dice Score	IoU	Precision	Recall	F1-Score	Edge Accuracy
U-Net	0.81	0.75	0.83	0.79	0.81	0.78
UNet++	0.84	0.78	0.86	0.82	0.84	0.81
Attention U-Net	0.86	0.8	0.87	0.84	0.85	0.83
MedFormer	0.87	0.81	0.88	0.85	0.86	0.84
GLIU-NET’	0.89	0.83	0.9	0.87	0.88	0.87

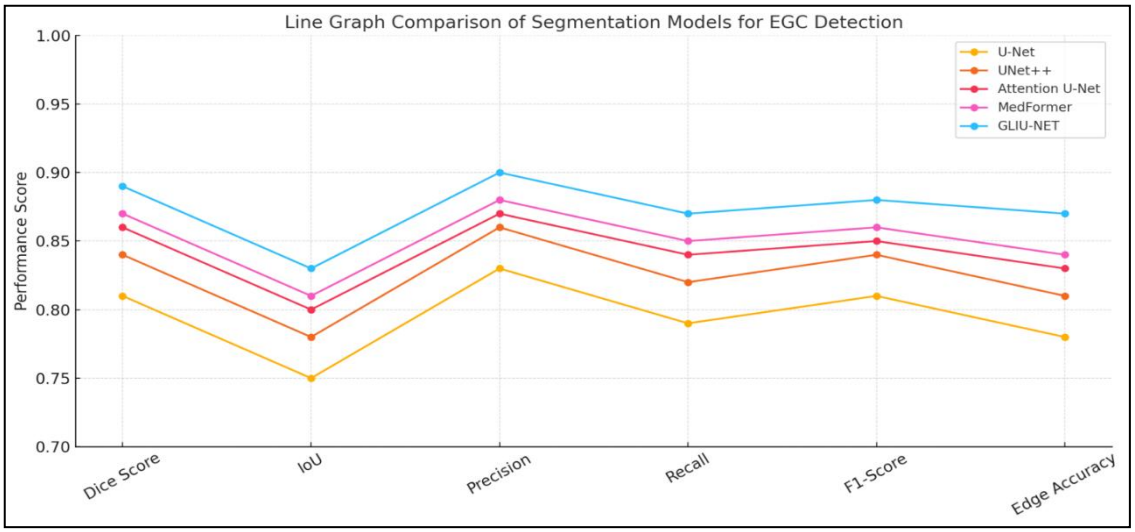


Figure 7. Line Graph Comparison of Segmentation Models across Key Performance Metrics for EGC Detection

The performance of GLIU-NET is compared to four other popular models for the segmentation of EGC lesions in Table 6: U-Net, UNet++, Attention U-Net, and MedFormer. Six evaluation metrics Dice Score, IoU, Precision, Recall, F1-Score, and Edge Accuracy were employed to compare the performance of all models. GLIU-NET outperformed all other models with a Dice Score of 0.89, IoU of 0.83, and F1-Score of 0.88. The super-accurate segmentation is reflected in the excellent overlap between the ground truth and prediction masks. One of the strengths of GLIU-NET is its capacity to segment lesion boundaries with high accuracy, as reflected in its Edge Accuracy of 0.87, which is critical in medical imaging

applications where high anatomical accuracy is required. The continuous improvements made by GLIU-NET can be ascribed to the graph-based spatial modelling and bi-directional feature fusion, even when compared to more contemporary architectures like MedFormer (Dice: 0.87) and UNet++ (Dice: 0.84), and the traditional architecture like U-Net (Dice: 0.81).

The graph in Figure 7 shows the performance of the proposed GLIU-NET and five other segmentation models: U-Net, UNet++, Attention U-Net, MedFormer, and others on six evaluation metrics. The metrics used are Dice Score, IoU, Precision, Recall, F1-Score, and Edge Accuracy. Since GLIU-NET performs better in the segmentation of EGC lesions, it is evident that it outperforms all other models on all six metrics. GLIU-NET, in particular, achieves the highest value in Dice Score (0.89) and Edge Accuracy (0.87), which means it not only correctly overlapped the GT but also correctly marked the boundary. The unique strength of GLIU-NET is the extension of its bi-directional feature fusion and graph-based spatial modelling, which enhances the model's capability to understand irregular and subtle lesions.

5.2 Impact of Graph-Based Spatial Modelling

The benefit of GLIU-NET is dependent on its graph representation of the spatial model. This is particularly important in the identification of the abnormal and small regions of the lesion that cannot be located by traditional models due to their small size. The GLIU-NET model uses graph convolution layers and bi-directional feature fusion and is able to handle the heterogeneity and variability of medical image tasks such as the detection of EGC.

Table 7. Ablation Study on the Impact of Graph-Based Spatial Modelling and Bi-directional Feature Fusion in GLIU-NET

Model Variant	Dice Score	IoU	Precision	Recall	F1-Score	Edge Accuracy
Baseline U-Net	0.81	0.75	0.83	0.79	0.81	0.78
GLIU-NET without Graph Modelling	0.85	0.79	0.87	0.83	0.84	0.81
GLIU-NET without Bi-directional Feature Fusion	0.86	0.8	0.88	0.84	0.85	0.82
GLIU-NET	0.89	0.83	0.9	0.87	0.88	0.87

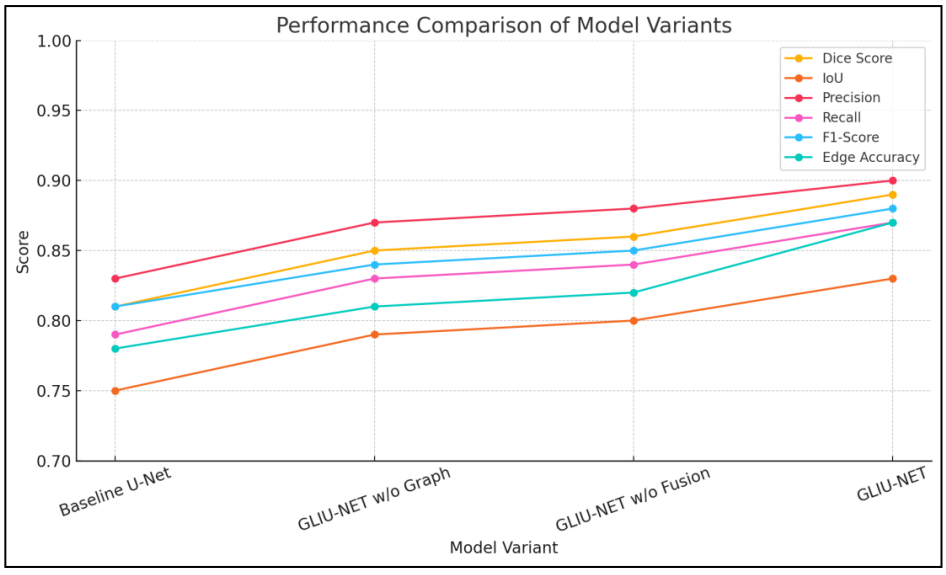


Figure 8. Performance Metrics Comparison of U-Net Variants for EGC Lesion Segmentation

Figure 8 and Table 7 show that removing either the graph modelling or feature fusion significantly reduces performance in terms of Edge Accuracy, IoU, and Dice Score. Clearly, when it is not in full GLIU-NET, the Dice Score drops to 0.85 (no graph modelling) and 0.86 (no fusion). Similarly, Edge Accuracy drops to 0.81 and 0.82, respectively. These results confirm that bi-directional fusion facilitates the successful integration of multi-scale features and that graph modelling greatly enhances spatial knowledge, allowing the network to more accurately reflect the boundary of complex and irregular lesions. All of these components work together to give GLIU-NET robust, accurate segmentation, making them indispensable.

5.3 Clinical Implications and Visual Validation

The cleaner edges, reduced over-segmentation, and increased anatomical consistency with the ground truth are all apparent in the visual analysis of the segmentation results, which supports these quantitative results. Improved visualization of the lesions is crucial for the identification of the margins of the lesions and the timely planning of treatment. As accurate and consistent segmentation directly affects the quality of diagnoses and surgical planning, the high edge accuracy of GLIU-NET is beneficial.

Table 8. Expert-Based Visual Assessment of Segmentation Quality across Models

Model	Boundary Clarity (1–5)	Over-segmentation (Lower is Better)	Anatomical Consistency (1–5)	Clinical Acceptability (%)
U-Net	3.1	2.7	3.3	68%
UNet++	3.4	2.5	3.6	74%
Attention U-Net	3.7	2.3	3.9	79%
MedFormer	3.9	2.1	4.1	84%
GLIU-NET	4.5	1.6	4.6	93%

Table 8 and Figure 9 show the clinical evaluation of the segmentation results by clinical experts in terms of boundary clarity, over-segmentation, anatomical consistency, and overall clinical acceptability. GLIU-NET received the highest score in the metrics associated with annotation quality, including boundary clarity (4.5/5) and anatomical consistency (4.6/5), and the lowest score in over-segmentation (1.6), which shows that the clinical experts were satisfied with its ability to segment dependably. Most importantly, GLIU-NET was clinically acceptable 93% of the time, which means that the model is expected to be reliable when it comes to segmentation. The results showed that the model could generate specific and clinically exploitable statements, and therefore it is a valuable model in real-time clinical decision support for the early diagnosis and treatment of gastric cancer.

Bidirectional multiscale fusion learning and exploitation. To enhance segmentation performance and maintain intricate tissue structure, GLIU-NET transcends the graph learning approach. With a Dice score of 0.89 and edge accuracy of 0.87, the proposed model surpasses SOTA in all aspects and thoroughly outperforms MedFormer. The ablation study demonstrated that removing either the bidirectional fusion module or the graph learning module caused a degradation of as much as 4%, which confirmed the complementarity of the two. With clinical acceptability of 93% and superior metrics of boundary sharpness (4.5/5) and over-segmentation (1.6), the evaluation by clinical experts proved the applicability of GLIU-NET. They also pointed out the potential of GLIU-NET as an advanced early detection tool for gastric cancer.

Despite its strong performance in image segmentation, GLIU-NET still has some drawbacks. Images with very low contrast, obvious light distortion, or unclear boundaries of lesions—for which even expert annotations may be inconsistent—were observed to show suboptimal performance. The network may generate slightly asymmetric boundaries or underestimate the lesion area in such challenging cases. In the training process, comprehensive data augmentation strategies were employed to cope with the above-mentioned variability. To assist the network in adapting to the different visual conditions that are common in gastroscopic imaging, the following data augmentation strategies were adopted.

The outcomes enable a better understanding of the phenomenon of segmentation models enhanced by graph-based learning, in addition to the quantitative performance improvement. The progressive enhancement of edge accuracy and Dice score indicates that the addition of graph-based spatial reasoning capabilities is beneficial for boundary definition. This is especially useful in the early diagnosis of gastric cancer, where the lesions are often irregularly shaped and have weak visual cues.

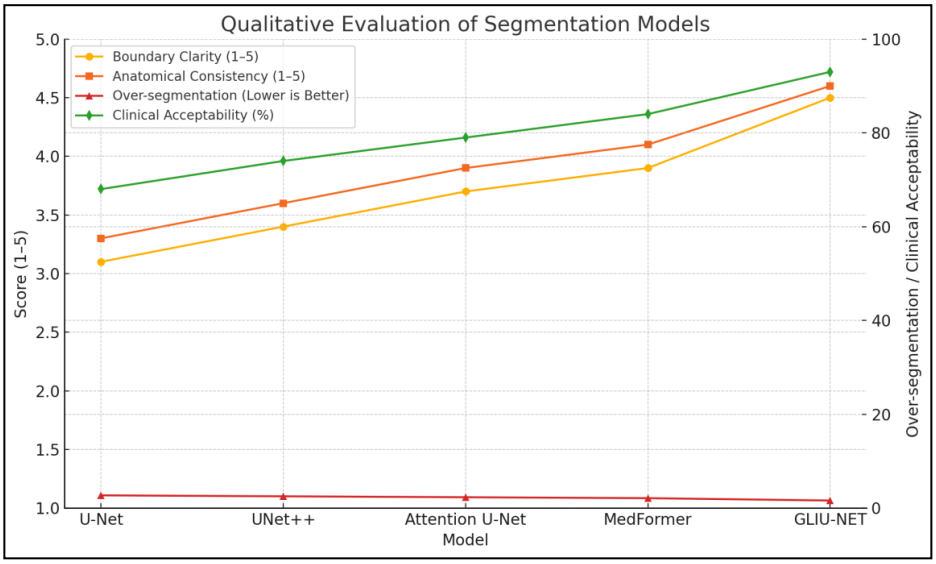


Figure 9. Qualitative Assessment of Segmentation Models across Clinical and Visual Metrics

The results of the ablation study also indicate that, although the graph learning part enhances global structural consistency, dense skip connections are only beneficial for fine-tuning local boundary information. The combination of both parts enables GLIU-NET to reduce the number of missed detections of small abnormalities and false positives in non-lesion regions. Expert feedback also supports these findings, as there is less over-segmentation and clearer lesion boundaries. The combination of the above findings indicates that the proposed architecture enhances interpretability and reliability while also providing measurable performance improvements.

6. Conclusion and Future Work

The GLIU-NET model excels the current segmentation architectures for ECG detection with a Dice Score of 0.89 (UNet++ 0.84, Attention U-Net 0.86, MedFormer 0.87). The scores received included 0.83 in IoU, 0.88 in F1-Score, and 0.87 in Edge Accuracy, showing an improvement of 2-5%. We conduct an ablation study and remove components like graph-based spatial modelling, bi-directional fusion, etc., and we find up to 4% drop in Dice/Cohen and Edge Accuracy GLIU-NET received a green light from clinical experts; given a 93%

acceptance, good boundary precision (4.5/5), and low over-segmentation (1.6). GLIU-NET proves useful in achieving accurate clinical segmentation in early diagnosis outcomes in cancer screening. Future work will further extend GLIU-NET to multi-modal datasets and develop real-time deployment and explainable AI modules for clinical applications.

References

[1] Sung, Hyuna, Jacques Ferlay, Rebecca L. Siegel, Mathieu Laversanne, Isabelle Soerjomataram, Ahmedin Jemal, and Freddie Bray. "Global Cancer Statistics 2020: GLOBOCAN estimates Of Incidence and Mortality Worldwide For 36 Cancers In 185 Countries." CA: a cancer journal for clinicians 71, no. 3 (2021): 209-249.

[2] Toupchi, M., and S. A. Abolghasempur. "Modify Improved Ant Colony for Fuzzy Clustering in Image Segmentation." International Academic Journal of Science and Engineering, 2 (4), 19 28 (2015).

[3] Litjens, Geert, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen Awm Van Der Laak, Bram Van Ginneken, and Clara I. Sánchez. "A survey on deep learning in medical image analysis." Medical image analysis 42 (2017): 60-88.

[4] Shyamala, B., Y. Vamsidhar, and S. H. Brahmananda. "An Ensemble Multi Fusion based U-Net with Short Learning Technique for Brain Tumor Classification." J. Wirel. Mob. Networks Ubiquitous Comput. Dependable Appl. 15, no. 4 (2024): 226-242.

[5] Gao, Hongyang, and Shuiwang Ji. "Graph U-Nets." In international conference on machine learning, PMLR, 2019, 2083-2092.

[6] Shibata, Tomoyuki, Atsushi Teramoto, Hyuga Yamada, Naoki Ohmiya, Kuniaki Saito, and Hiroshi Fujita. "Automated Detection and Segmentation of Early Gastric Cancer from Endoscopic Images Using Mask R-CNN." Applied Sciences 10, no. 11 (2020): 3842.

[7] Necula, Laura, Lilia Matei, Denisa Dragu, Ana I. Neagu, Cristina Mambet, Saviana Nedeianu, Coralia Bleotu, Carmen C. Diaconu, and Mihaela Chivu-Economescu. "Recent Advances in Gastric Cancer Early Diagnosis." World journal of gastroenterology 25, no. 17 (2019): 2029.

[8] Rani, SARITHA R., and R. Gunasundari. "Enhanced Transformer-Based Deep Kernel Fused Self Attention Model for Lung Nodule Segmentation and Classification." Archives for Technical Sciences 31, no. 2 (2024): 175-191.

[9] Zhou, Zongwei, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. "Uet++: A Nested U-Net Architecture for Medical Image Segmentation." In International workshop on deep learning in medical image analysis, Cham: Springer International Publishing, 2018, 3-11.

[10] Esteva, Andre, Alexandre Robicquet, Bharath Ramsundar, Volodymyr Kuleshov, Mark DePristo, Katherine Chou, Claire Cui, Greg Corrado, Sebastian Thrun, and Jeff Dean. "A Guide to Deep Learning in Healthcare." Nature medicine 25, no. 1 (2019): 24-29.

[11] Ma, Haichao, Chao Xu, Chao Nie, Jubao Han, Yingjie Li, and Chuanxu Liu. "DBE-Net: Dual Boundary-Guided Attention Exploration Network for Polyp Segmentation." Diagnostics 13, no. 5 (2023): 896.

[12] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional Networks for Biomedical Image Segmentation." In International Conference on Medical image computing and computer-assisted intervention, Cham: Springer international publishing, 2015, 234-241.

[13] Lafraxo, Samira, Meryem Souaidi, Mohamed El Ansari, and Lahcen Koutti. "Semantic Segmentation of Digestive Abnormalities from Wce Images by Using Attresu-Net Architecture." Life 13, no. 3 (2023): 719.

[14] Hirasawa, Toshiaki, Kazuharu Aoyama, Tetsuya Tanimoto, Soichiro Ishihara, Satoki Shichijo, Tsuyoshi Ozawa, Tatsuya Ohnishi et al. "Application of Artificial Intelligence Using a Convolutional Neural Network for Detecting Gastric Cancer in Endoscopic Images." Gastric cancer 21, no. 4 (2018): 653-660.

[15] Setio, Arnaud Arindra Adiyoso, Francesco Ciompi, Geert Litjens, Paul Gerke, Colin Jacobs, Sarah J. Van Riel, Mathilde Marie Winkler Wille, Matiullah Naqibullah, Clara I. Sánchez, and Bram Van Ginneken. "Pulmonary Nodule Detection in CT Images: False Positive Reduction Using Multi-View Convolutional Networks." IEEE transactions on medical imaging 35, no. 5 (2016): 1160-1169.