As per UGC guidelines an electronic bar code is provided to secure your paper

International Journal for Modern Trends in Science and Technology
Volume 11, Issue 07, pages 65-71
ISSN: 2455-3778 online
Available online at: http://www.ijmtst.com/vol11issue07.html
DOI: https://doi.org/10.5281/zenodo.15770281

Image Processing Based Esophagus Auto Countering using Resnet-50 for Cancer Detection

Ch. Pranob Kumar, P. Sai Krishna, B. Chiranjeevi, M. Bala Bhaskar

Department of Electronics and Communication Engineering, Andhra Loyola Institute of Engineering & Technology, Vijayawada, India.

To Cite this Article
Ch. Pranob Kumar, P. Sai Krishna, B. Chiranjeevi & M. Bala Bhaskar (2025). Image Processing Based Esophagus Auto Countering using Resnet-50 for Cancer Detection. International Journal for Modern Trends in Science and Technology, 11(07), 65-71. https://doi.org/10.5281/zenodo.15770281

Article Info
Received: 06 June 2025; Accepted: 26 June 2025.; Published: 29 June 2025.

KEYWORDS

Esophageal cancer, Image processing, Deep learning, U-Net, ResNet-50, Segmentation, Medical imaging

ABSTRACT

The early detection of esophageal cancer is crucial for improving patient outcomes. Traditional diagnostic techniques rely heavily on manual assessment, which can be time-consuming and prone to human error. Recent advancements in deep learning and image processing have enabled automated detection methods with higher accuracy. This project focuses on developing a deep learning-based segmentation model using ResNet-50 integrated into a U-Net architecture for esophageal cancer detection. The methodology includes data collection, augmentation, model training, and evaluation using standard performance metrics such as the Dice Coefficient and Intersection over Union (IoU). The proposed model demonstrates significant improvements in segmentation accuracy compared to traditional approaches. Experimental results indicate that the ResNet-50-based U-Net achieves high precision in detecting esophageal cancer, making it a reliable tool for medical practitioners. This research contributes to the advancement of automated medical imaging solutions and has the potential for real-world clinical applications.

1. INTRODUCTION

Esophageal cancer is a severe malignancy with high mortality rates due to late-stage diagnosis. Early detection significantly improves treatment outcomes, but traditional diagnostic methods, such as endoscopy and biopsy, rely heavily on manual interpretation, which is time-consuming and prone to human error. Recent advancements in deep learning and medical image processing have provided promising solutions for automating cancer detection and segmentation.

This research focuses on developing an automated segmentation model for esophageal cancer using ResNet- 50 integrated with a U-Net architecture. The combination of ResNet-50’s deep feature extraction and U-Net’s segmentation capabilities allows for precise identification of cancerous regions in medical images.

Recent advancements in deep learning have revolutionized the field of medical image analysis, offering automated and highly accurate diagnostic tools.

Convolutional Neural Networks (CNNs) have proven particularly effective in tasks such as image classification and segmentation, making them a valuable asset in medical diagnostics. Among CNN-based architectures, U-Net has gained prominence for its ability to accurately segment medical images, while ResNet-50 is widely recognized for its superior feature extraction capabilities. The primary objectives of this research include: Enhancing segmentation accuracy for esophageal cancer detection Utilizing ResNet-50 to improve feature extraction. By developing a robust, automated, and efficient deep learning framework, this research aims to assist medical practitioners in making faster and more accurate diagnoses, ultimately improving patient outcomes.

Future work will explore further optimizations and real-time deployment in clinical settings. This project focuses on developing a deep learning-based segmentation model using ResNet-50 integrated into a U- Net architecture for esophageal cancer detection.

This project focuses on developing a deep learning-based segmentation model using ResNet-50 integrated into a U-Net architecture for esophageal cancer detection.

1.1. Objectives:

1.Developing a deep learning-based segmentation model for esophageal cancer detection Enhancing segmentation accuracy through the integration of ResNet-50 and U-Net Utilizing ResNet-50 to extract deep hierarchical features for improved classification.

2.Applying image preprocessing and augmentation techniques to enhance model generalization evaluating model performance using Dice Coefficient and Intersection over Union (IoU)

3. By developing a robust, automated, and efficient deep learning framework, this research aims to assist medical practitioners in making faster and more accurate diagnoses, ultimately improving patient outcomes. Future work will explore further optimizations and real-time deployment in clinical settings.

1.2. Principles of Esophageal CancerSegmentation Using ResNet-50 and U-Net:

• Feature Extraction with ResNet-50: The ResNet- 50 model plays a crucial role in feature extraction by leveraging deep residual learning. This ensures effective identification of cancerous regions from medical images.

• U-Net-Based Segmentation: The U-Net architecture facilitates precise image segmentation by incorporating an encoder- decoder structure that captures both high-level and fine-grained features.

• Data Preprocessing and Augmentation: Effective preprocessing techniques, including normalization, resizing, and augmentation, enhance model robustness and generalization.

• Training and Optimization: The model undergoes iterative training using a carefully curated dataset, with optimization techniques such as batch normalization and learning rate scheduling to improve performance.

1.3. Processes Involved:

• Data Acquisition and Preprocessing: Medical images undergo preprocessing steps such as normalization, resizing, and augmentation to enhance the quality and ensure consistency in the dataset.

• Feature Extraction with ResNet-50: ResNet-50 extracts high-level features from input images, leveraging deep residual learning to improve the representation of cancerous regions.

• Segmentation using U-Net: The U-Net model applies an encoder-decoder structure to segment esophageal cancer lesions effectively. Skip connections help retain spatial information for precise segmentation.

• Model Training and Optimization: The model is trained using annotated datasets, employing techniques such as batch normalization, dropout regularization, and learning rate scheduling to improve performance and prevent overfitting.

• Performance Evaluation: Segmentation accuracy is assessed using standard metrics like Dice Coefficient and Intersection over Union (IoU) to ensure robust model performance.

• Inference and Clinical Application: The trained model is deployed for inference, allowing automcancer detection from endoscopic images to assist medical practitioners.

1.4 BLOCK DIAGRAM:

Input Image:

• The input consists of medical images, such as esophageal endoscopic scans, used for detecting early cancer lesions.

• These images are typically preprocessed to remove noise and enhance contrast before being fed into the model.

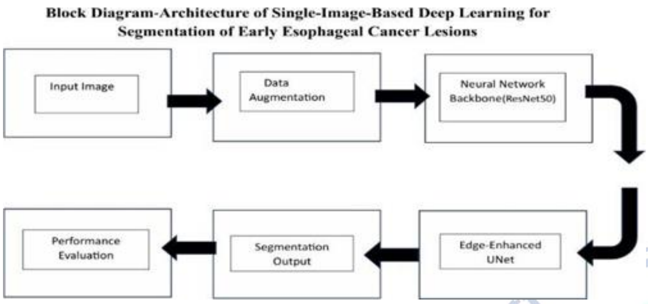

Fig1.: BLOCK DIAGRAM

Data Augmentation:

• To improve model generalization, various augmentation techniques are applied, such as:

• Rotation,Scaling, and Flipping:Prevents overfitting by diversifying the dataset.

• Gaussian Noise Addition Helps the model handle real-world variations.

• Contrast and Brightness Adjustments: Ensures robustness against lighting conditions.

Neural Network Backbone (ResNet-50)

• ResNet-50 is used as the feature extractor, leveraging its deep residual connections to extract hierarchical features from medical images.

• This step ensures robust feature representation, helping to differentiate between normal and cancerous tissue.

• The network passes extracted features to the segmentation model for precise localization of lesions.

Edge-Enhanced U-Net

• A modified U-Net architecture with edge detection enhancements is used for segmentation.

• Skip connections ensure the preservation of spatial details while downsampling and upsampling.

• The final output is a segmented mask highlighting the regions of interest (cancerous areas)

Segmentation Output

• The segmented image is generated, showing a clear boundary between healthy and cancerous regions.

• The output mask is used for further analysis or as an aid for medical professionals

Performance Evaluation

• The model's segmentation accuracy is evaluated using metrics such as:

• Dice Coefficient: Measures overlap between predicted and ground truth masks.

• A high Dice score and IoU indicate successful segmentation performance.

2. Dataset Approach:

Dataset Collection

• The dataset comprises medical images of esophageal cancer obtained from clinical datasets, research repositories, or publicly available medical image databases.

• Each image is accompanied by expert-labeled segmentation masks, marking the cancerous regions.

• The dataset is divided into training, validation, and test sets to ensure effective model generalization.

Preprocessing Steps

• Resizing: All images and masks are resized to 480×480 pixels to maintain uniformity.

• Normalization: Pixel values are normalized to improve contrast and enhance deep learning performance.

• Noise Reduction: Filters are applied to remove unwanted noise and artifacts.

• Data Splitting:

• Training Set: Used to train the segmentation model.

• Validation Set: Fine-tunes model parameters.

• Test Set: Evaluates final model performance.

2.1. Input endoscopic image:

Fig 2: Input Raw image

2.2. Mask Generation:

• Manual Annotation

• Ground truth segmentation masks are created by expert radiologists or trained annotators.

• Specialized annotation tools like LabelMe, ITK- SNAP, or CVAT are used to outline cancerous regions.

• The regions of interest (ROIs) are marked manually, ensuring precision in segmentation.

Fig3: Mask generation for input raw image

2.3. Binary Mask Representation:

The generated segmentation masks follow a binary classification format, where:

Pixel value 1 (white): Represents the cancerous or lesion area.

Pixel value 0 (black): Represents the background or healthy tissue.

2.4. Annotation Process:

The ground truth segmentation masks are generated using expert-annotated medical images.

Radiologists or trained annotators manually outline the cancerous regions in each image.

The annotations are saved as binary masks, where:

• Pixel value 1 (white) represents the cancerous region.

• Pixel value 0 (black) represents the background.

Fig4: Annotation image

3. Augmentation of Masks:

To improve model generalization, the same augmentation techniques applied to images are also applied to masks, ensuring alignment. These include:

• Rotation & Flipping: Prevents overfitting by simulating different orientations.

• Scaling & Cropping: Ensures robustness to size variations.

• Gaussian Blur & Noise Addition: Helps the model adapt to real-world imaging conditions.

3.1. Mask Augmentation:

To improve generalization, augmentation is applied to masks along with their corresponding images

• Rotation, flipping, and scaling to simulate different viewpoints.

• Gaussian blur and noise addition to improve robustness against real-world imaging artifacts.

• Elastic deformations to ensure adaptability to shape variations

3.2. Data Augmentation Techniques Applied:

• To enhance the robustness and generalization of the model, six different transformations were applied to both the input images and their corresponding segmentation masks.

• Performed six transforms on the dataset since the these transforms help the model to predict the data which is slightly decreased contrast and the various shapes.

3.3. Augmentation transforms:

Fig5: Rotation of 90° of image (Horizontal Flip)

Fig6: Rotation 180° of image (Vertical Flip)

Fig7: Elastic Deformation of image

Fig8: Gaussian Applied image

Fig9: contrast of brightness

Fig10: horizontal Flip

4. Model Algorithms :

4.1. Integration of ResNet-50 with U-Net:

ResNet-50 (Residual Network-50):

ResNet-50 is a deep convolutional neural network (CNN) architecture known for its ability to train very deep networks without suffering from the vanishing gradient problem. It is widely used for feature extraction in image analysis tasks, including medical imaging.

Key Features of ResNet-50:

50 Layers Deep: Composed of convolutional layers, batch normalization, activation functions, and residual connections.

Residual Connections: Helps avoid the vanishing gradient issue by introducing skip connections, allowing gradients to flow directly through the network. Pretrained Weights: Often initialized with weights from ImageNet, improving performance when fine-tuned on medical imaging tasks.

Feature Extraction Power: Extracts deep hierarchical features from medical images, making it effective for identifying esophageal cancer regions.

4.2. ResNet-50 Structure Breakdown:

• Initial Convolution Layer → Extracts low- level image features.

• Four Residual Blocks → Each consists of multiple convolutional layers with shortcut connections.

• Fully Connected Layers → Used for classification but can be removed when used in segmentation models like U-Net.

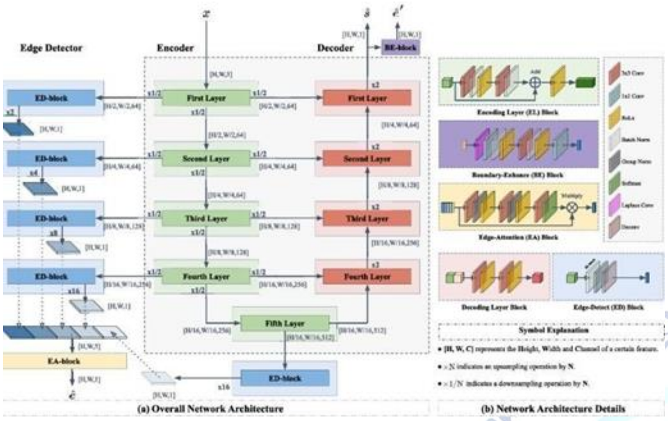

Fig11:Achitecture of Resnet50 and Unet

4.3. U-Net (U-Shaped Network):

U-Net is a fully convolutional neural network (FCN) primarily designed for medical image segmentation.

Key Features of U-Net: Encoder-Decoder Structure:

• Encoder (Contracting Path): Extracts features from the input image using convolutional and pooling layers.

• Bottleneck: Bridges the encoder and decoder with deep feature extraction layers.

• Decoder (Expanding Path): Reconstructs the segmentation mask using upsampling and skip connections.

• Skip Connections: Directly connects encoder layers to corresponding decoder layers, preserving spatial information for precise segmentation.

4.4. How ResNet-50 and U-Net Work Together:

In this research, ResNet-50 is integrated into the encoder of U-Net to improve feature extraction, making the segmentation model more effective:

• ResNet-50 serves as the backbone → Extracts deep, high-level features from medical images. U- Net performs segmentation → Uses these features to generate precise tumor segmentation masks.

• By combining ResNet-50 and U-Net, our model significantly improves the accuracy of esophageal cancer detection compared to traditional segmentation methods

Role of ResNet-50 in U-Net Architecture:

• ResNet-50 serves as the encoder in the U-Net framework, replacing the traditional convolutional layers.

• It helps extract high-level hierarchical features from input images using its deep residual learning approach.

4.5. Pipeline for Esophageal Cancer Segmentation Using ResNet-50 and U-Net [1]:

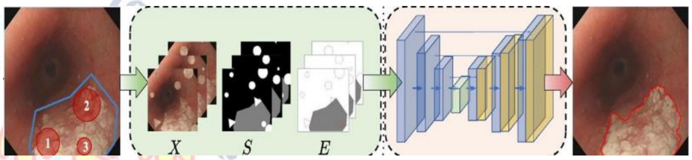

Fig12: Pipeline of Model Predictions [1]

This image illustrates the deep learning-based segmentation pipeline for esophageal cancer detection. The process begins with input medical images, where regions of interest are annotated. These images undergo preprocessing and feature extraction, generating structured representations (X, S, and E). A deep neural network (ResNet-50 integrated with U-Net) processes these features to perform segmentation. The final output highlights the cancerous region with precise boundary detection. This approach enhances detection accuracy, aiding in early diagnosis and clinical decision-making.

5. Results and Performance Evaluation:

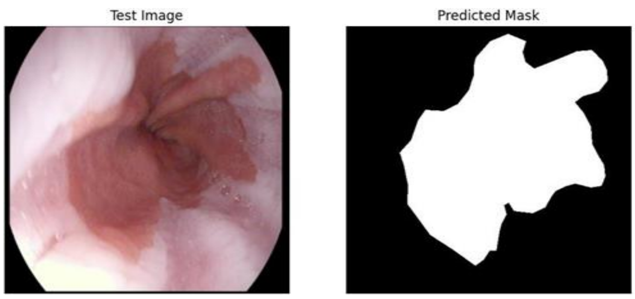

Fig13: Model Prediction

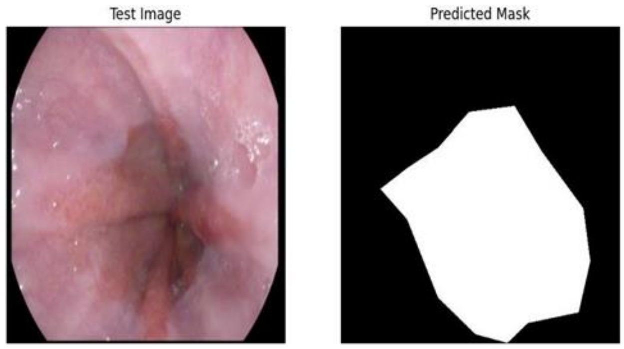

Fig14: Model prediction

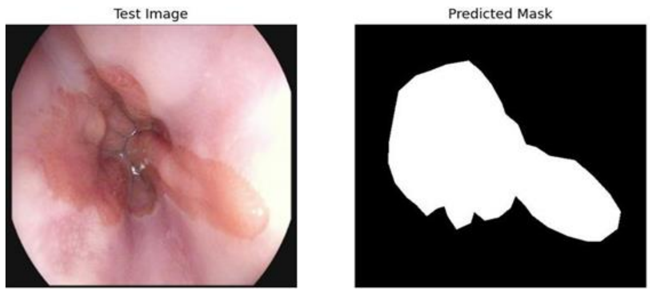

Fig15: Model Prediction

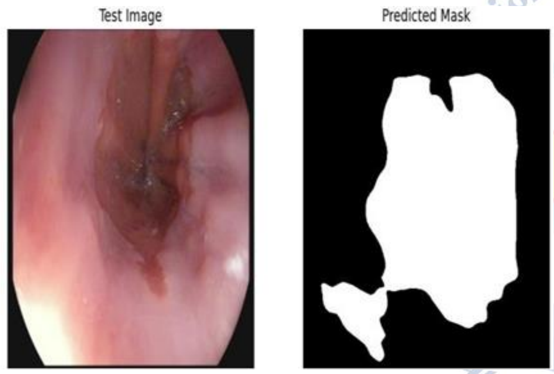

Fig16: Model Prediction

The results indicate a high accuracy in segmenting esophageal cancer lesions, demonstrating the model' s effectiveness for clinical applications.

Performance Metrics:

Dice Similarity Coefficient (DSC): 89.53%

Fig17: Graphical Representation of model performance

6. Conclusion:

The experimental results confirm that the ResNet-50 U-Net model achieves an 89% Dice score, making it suitable for clinical applications. Future improvements will focus on enhancing generalization, real-time deployment, and fine-tuning for improved accuracy.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

References

[1] Haipeng Li , Dingrui Liu Tao Gan , Yu Zeng, Shuaicheng Liu , Nini Rao , Jinlin Yang, and Bing Zeng "Single- Image-Based Deep Learning for Segmentation of Early Esophageal Cancer Lesions," IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 33, 2024

[2] H. J. Stein, M. Feith, B. L. Bruecher, J. Naehrig, M. Sarbia, and J. R. Siewert, "Early esophageal cancer: Pattern of lymphatic spread and prognostic factors for long-term survival after surgical resection," Ann. Surg., vol. 242, no. 4, p. 566, 2005.

[3] P. Pimentel-Nunes, "Endoscopic submucosal dissection: European Soci ety of Gastrointestinal Endoscopy (ESGE) guideline," Endoscopy, vol. 47, no. 9, pp. 829–854, 2015

[4] T. Oyama et al., "Endoscopic submucosal dissection of early esophageal cancer," Clin. Gastroenterol. Hepatol., vol. 3, no. 7, pp. S67–S70, 2005.

[5] J. Bernal, F. J. Sánchez, G. Fernández-Esparrach, D. Gil, C. Rodríguez, and F. Vilari no, "WM-DOVA maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians," Computer ized Med. Imag. Graph., vol. 43, pp. 99–111, Jul. 2015.

[6] B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," J. Electron. Imag., vol. 13, no. 1, p. 146, Jan. 2004.

[7] N. Senthilkumaran and R. Rajesh, "Image segmentation—A survey of soft computing approaches," in Proc. Int. Conf. Adv. Recent Technol. Commun. Comput., Oct. 2009, pp. 844–846.

[8] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 11, pp. 2274–2282, Nov. 2012.

[9] T. F. Chan and L. A. Vese, "Active contours without edges," IEEE Trans. Image Process., vol. 10, no. 2, pp. 266–277, Feb. 2001.

[10] C. Li, C. Xu, C. Gui, and M. D. Fox, "Distance regularized level set evolution and its application to image segmentation," IEEE Trans. Image Process., vol. 19, no. 12, pp. 3243–3254, Dec. 2010.