ISSN No:-2456-2165

Survey on AI-Based Polyp Localization and Segmentation for Enhanced Colonoscopy Diagnosis

[1]Aruna U; [2]Mallikarjun M Kotur; [3]Mogallapu Rahul; [4]Mohammed Dhanish; [5]Lakshmi Priya B M
Department of Information Science RNS Institute of Technology, Bengaluru

Abstract:- Colonoscopy being a critical diagnostic and preventive procedure for colorectal cancer, a major contributor to cancer related deaths globally. Early recognition of polyps at an early stage in the colon is essential for minimizing the progression of cancer. Manual detection during colonoscopy is labor-intensive and vulnerable to human errors such as missed polyps or false diagnoses due to variations in polyp size, shape, and texture.

This paper explores a deep learning-based system for automated segmentation and detection of polyps by using advanced architectures like U-Net and Mask R-CNN. These techniques aim to enhance diagnostic precision, reduce clinician workload, and provide real-time feedback during procedures, thereby transforming the landscape of gastrointestinal healthcare.

Keywords:- *Colonoscopy, Polyp Detection, Segmentation, Deep Learning, CNN, Medical Imaging, Colorectal Cancer.*

I. INTRODUCTION

Colorectal cancer is a significant global health issue, ranking as the most common cancer in men and the second most common in women. Precancerous polyps formed in the colon or rectum can beeffectively managed if is recognized and removed early, thereby preventing cancer progression. Colonoscopy, considered the benchmark method for detecting such lesions, relies heavily onthe clinician's ability to visually identify polyps. However, this manual technique can lead to missed detections, particularly for polyps that are small, flat, or obscured by anatomical structureslike folds. Recent advancements in artificial intelligence have shown considerable potential in enhancing polyp detection and segmentation. Convolutional neural networks (CNNs), with their capability to recognize intricate patterns in medical imaging, have transformed diagnostic processes by facilitating precise and efficient analysis of endoscopic visuals.

II. LITERATURE SURVEY

Numerous studies highlight the success of deep learning in enhancing medical imaging for polyp detection and segmentation. These foundational works underpin the proposed methodology:

Yu et al.

A framework utilizing three-dimensional convolutional neural networks (3D-CNNs) was designed for colonoscopy video analysis. This method leverages both spatial and temporalcharacteristics to effectively identify polyps during motion.

The model significantly improves the identification of subtle andpartially visible polyps by addressing challenges such as occlusions and motion blur. The results indicated a decrease in false negatives, making it a robust solution for video-based diagnostics.

Zhang et al.

proposed a hybrid system that integrates fully Convolutional Networks (FCNs) with texton-based patch representation for enhanced segmentation. This approach combines global contextual information from FCNs with local feature refinement using texture-based methods, leading to more accurate boundary detection. Their experiments on benchmark datasets demonstrated superior performance in segmenting polyps with irregular shapes overlap.

Kang et al.

utilized the Mask R-CNN framework for instance segmentationof polyps. The model's ability to perform pixel-levelsegmentation and detect individual polyps in complex scenarios, such as overlapping tissues and varying lighting conditions, wasa key advancement.

This approach achieved state-of-the-art performance metrics on multiple colonoscopy image datasets, showcasing its ability for clinical implementation.

Lastly, **Tashk et al.**

developed an enhanced U-Net architecture specifically designed for polyp segmentation. By incorporating additional skip connections and multi-scale feature extraction, the model addressed issues of class imbalance and improved segmentation accuracy for small and flat polyps. The study highlighted the importance of architectural innovations in overcoming challenges unique to medical imaging.

These studies collectively demonstrate the potential of deep learning to transform polyp detection and segmentation by addressing challenges such as variability in size, texture, and lighting conditions. The proposed system builds upon these approaches by integrating the strengths of U-Net and Mask R- CNN to achieve real-time and accurate detection.

III. OBJECTIVES

The primary objective of a fake registration plate detection system is to eliminate the use of fraudulent or altered plates on vehicles. This serves several key purposes:

- **Accurate Detection and Segmentation:** Enhance the precision of polyp recognition and segmentationacross diverse colonoscopy datasets, including polyp.
- **Real-Time Processing:** Enable real-time detection and segmentation to provide immediate feedback during colonoscopy procedures, supporting clinicians in decision-making.
- **Data Augmentation and Scalability:** Employadvanced data augmentation approaches to expand thespectrum of training datasets and enhance the model's scalability across varied clinical settings.
- **Integration With Clinical Workflows:** Ensure seamless integration of the system into existing colonoscopy workflows, requiring minimal disruption and additional training for clinicians.

Through precise detection and segmentation of polyps in colonoscopy images, these systems aim to improve diagnostic accuracy, support the early identification of colorectal cancer, enhance patient outcomes, and lessen reliance on manual evaluation.

IV. EXISTING SYSTEM

A variety of present systems and methods for polyp recognition and segmentation have been explored in both manual and automated frameworks. These include:

- **Manual Polyp Recognition in Colonoscopy:** Existing systems for polyp recognition rely heavily on human-driven methods, which are considered the goldstandard but are time-intensive, prone to human error due to clinician fatigue or missed polyps, and highly dependent on the expertise of the endoscopist, leadingto variability in diagnostic outcomes.
- **Traditional Image Processing Techniques** Many automated systems use fundamental image processing methods, including edge detection, texture analysis, and morphological transformations.
- **Semi-Automated Systems:** Some semi-automated systems use clinician inputs to improve detection accuracy. These systems typically mark areas of interest for the clinician to confirm or reject. While helpful, these systems rely on human intervention and have high false positives.
- **Existing Deep Learning Approaches:** Recent innovations have integrated Convolutional Neural Networks (CNNs) and Fully Convolutional Networks (FCNs) to improve polyp detection and segmentation. Despite providing substantial improvements over traditional methods, these techniques still have certain drawbacks.

V. PROPOSED SYSTEM

Our polyp segmentation approach involves two stages. The initial stage utilizes the FCN-8S network to identify potentialpolyp regions by analyzing the extracted features. In thesecond stage, Otsu thresholding is applied to convert the probability map from FCN-8S into a binary image, and the largest connected component is selected as the most probable polyp region.

- FCN-8S Network: The FCN was introduced for semantic segmentation and uses convolution and pooling layers to create dense feature maps. FCN is available in three versions: FCN-32S, FCN-16S, and FCN-8S. Among these, FCN-8S provides the most accurate segmentation by utilizing outputs from pool3, pool4, and conv7, and up sampling the featuremap by a factor of 8. The FCN-8S identifies potential polyp regions, which are then enhanced using Otsu thresholding to reduce false positives and pinpoint the most probable polyp locations.
- Patch Selection and Data Augmentation: Training FCN-8S is challenging due to limited medical data. To address this, we implement data augmentation techniques like image rotation and patch selection. Rotation helps the model learn various polyp structures, while patch selection focuses on areas containing polyps, borders, and background regions. Thesemethods improve the model's generalization by providing diverse training examples.

This method combines FCN-8S with Otsu thresholding and data augmentation to improve the precision of polyp segmentation and recognition in colonoscopy imaging.

VI. SYSTEM ARCHITECTURE

Detecting polyps in colonoscopic images involves multiple essential steps. The process begins with image acquisition, during which colonoscopy visuals are obtained and preprocessed to reducenoise and enhance brightness. Key features for polyp identification are then extracted through Convolutional Neural Networks (CNNs). The segmentation phase then employs U-Net and Mask R-CNN to distinguish polyps from the surrounding tissue.

After segmentation, post-processing refines the results, and thedecision layer classifies the polyps and provides real-time alerts. The system integrates with the clinical workflow through a user interface and stores results in the EMR system, providing a segmentation map and confidence scores to assist clinicians inmaking accurate diagnoses.

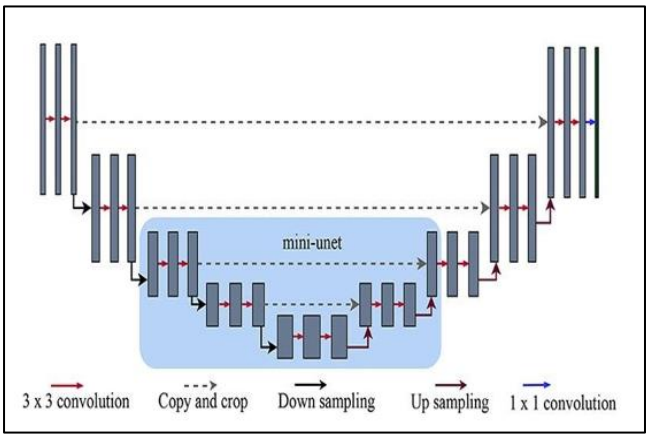

Fig 1: U-Net Model

VII. TECHNOLOGIES

- **Convolutional Neural Networks (CNNs)**: These deep learning models are employed to extract features, enabling the identification of image patterns that indicate the presence of polyps.
- **U-Net and Mask R-CNN**: These architectures are used for segmenting images, allowing for accurateseparation of polyps from the surrounding tissue.
- **Data Augmentation**: Techniques such as rotation, scaling, and flipping are applied to enlarge thedataset, increasing the model's robustness and performance across varied data**.**
- **Image Processing**: Methods like noise reduction,contrast enhancement, and resizing are used to prepare the images for deep learning models.
- **Real-time Processing**: The system supports real-time feedback during colonoscopy procedures, helping clinicians make timely decisions.
- **GPU Acceleration**: The use of GPUs ensures faster model training and inference, enabling efficient processing of high-resolution colonoscopy images.

VIII. METHODOLOGY

This study focuses on the development and analysis of a polypsegmentation framework based on the U-Net model**.** The introduced model, builds upon the U-Net design, incorporatingDilated convolutions enable the expansion of the receptive field, allowing the network to capture broader context without adding extra parameters. This approach improves the network's ability to capture spatial information from the colonoscopy images, which is critical for accurate polyp detection.

A. Dataset

➢ *Five Publicly Available Colonoscopy Datasets are Utilized for Training And Evaluation:*

- **CVC-Clinic DB**: Contains 612 three-channel color images derived from 31 colonoscopy recordings. These images are in .png and .tiff formats, with dimensions of 384x288 pixels. All polyps in this dataset are small in size.
- **KVASIR-SEG**: A collection of 1000 polyp images, including ground truth annotations. The images vary in resolutions, ranging from 332x487 to 1920x1072 pixels, and the dataset contains a diversity of polyptypes (flat, small, large).
- **KVASIR-Sessile**: Contains 196 smaller polyps classified using the Paris classification system. These polyps are difficult to generalize due to their small size.
- **KVASIR-Endo-SEG**: An enhanced subset of KVASIR Endoscopy, with 55 annotated images. This dataset focuses on flat polyps, with images resized to 336x336 pixels.
- **Hyper KVASIR-SEG**: The largest dataset of gastrointestinal tract images, with 99,417 images and 10,662 segmented images of polyps. It includes a variety of polyp types and is annotated with bounding boxes.

B. Data Preparation

The collected data consists of images of various sizes and quality. Before being fed into the model, the images undergo pre- processing to standardize their dimensions and improve quality. Firstly every image is resized to 256x256 pixels. Then, they are converted into grayscale to reduce computational complexity and focus on essential features.

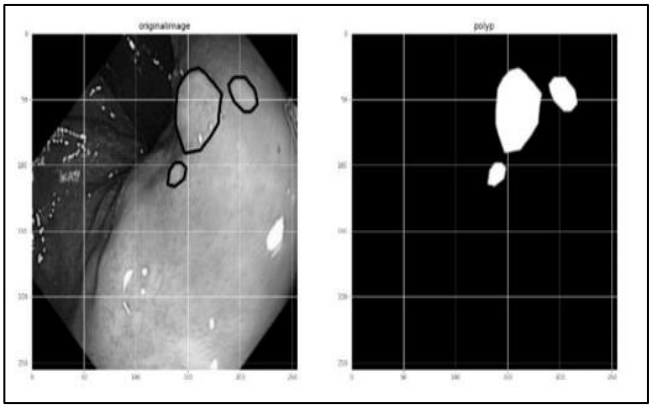

Fig 2: Pre-Processed Image from Dataset

C. Proposed Architecture

The introduced **U-Net-Seg** model builds on the original U-Netarchitecture, incorporating convolutions to improve the receptive field. The architecture is structured into two primary components: the encoder (contracting path) and the decoder (expanding path). The encoder is composed of six convolutional levels, each featuring 3x3 convolutional layers, batch normalization, ReLU activation, 2x2 max-pooling operations, and a dropout layer for regularization. This configuration incrementally down-samples the input image, extracting essential features required for accurate polyp segmentation.

The decoder consists of six corresponding layers that up-sample the feature maps to their original dimensions, combining information from the encoder using concatenation. Each layer includes a 2x2 transposed convolution, depth concatenation, dropout, 3x3 convolution, batch normalization, and ReLU activation. The final output is a 1x1 convolution with a activation function to generate the binary classification map (polyp or background).

D. Implementation Details

Framework is implemented using the TensorFlow-based Keras library. The N-Adam optimizer is used with a learning rate of 0.001, batch size of 8, and 25 epochs.

Training is performed on a GPU-equipped system with a Processor, 13 GB of RAM, Tesla K80 GPU accelerator, and 12 GBof VRAM.

➢ *The Model is Evaluated using Several Performance Metrics:*

- **Pixel Accuracy (PA)**: Measures the percentage of pixels that are correctly classified.
- **Dice Score for Set Similarity Assessment**: A metricused to evaluate the overlap between predicted and ground truth masks.
- **Jaccard Index (IoU)**: Represents the ratio of the intersection over the union of predicted and true masks.
- **Precision (P)**: Measures how many of the predicted positives are right.
- **Recall (R)**: Evaluates the model's ability to identify true positive pixels.

These are used in the evaluation of performance and effectiveness of the Dilated-U-Net-Seg model in segmenting polyps from colonoscopy images, providing a better evaluation of its performance across different datasets.

➢ *Advantages*

- Improved Segmentation Accuracy: Enhances accuracy with dilated convolutions for better feature capture.
- Real-Time Processing: Provides immediate feedback to clinicians during procedures.
- Generalization: Demonstrates strong performance across various datasets and clinical applications.
- Flexibility: Adapts to various polyp types and medical imaging tasks.
- Reduced Data Dependency: Lowers reliance on large datasets through augmentation and transfer learning.
- Robust to Small and Flat Polyps: The framework effectively detects challenging polyp types, such as small and flat polyps, which are often missed during manual inspections.
- Handles Class Imbalance: By using dice loss as the objective function, the framework mitigates issues arising from class imbalance in segmentation tasks.

➢ *Disadvantages*

- High Computational Demand: Requires significant resources for training and inference.
- Image Quality Dependency: Performance drops with noisyor poor-quality input images.
- Overfitting Risk: Prone to overfitting on small or less diverse datasets.
- Complex Polyp Segmentation: Struggles with rare or complex polyp shapes.
- Dependency on Annotated Data: The model heavily relies on high-quality annotated datasets, which are time-consuming and expensive to create.
- Complex Hyperparameter Tuning: Optimizing hyperparameters like learning rates, dropout, and batch size requires significant experimentation, increasing development time.

➢ *Potential Future Enhancements*

- **Real-Time Video Integration:** Expand the system to process real-time colonoscopy video feeds.
- **Edge Device Optimization**: Simplify the model for use on low-resource devices.
- **Multi-Class Detection**: Expand to classify and detect different polyp types.
- **Cross-Dataset Validation**: Improve robustness by validating across diverse datasets.
- **Federated Learning**: Use privacy-preserving collaborativetraining methods.
- **Multi-Modal Integration**: Include additional data types likeclinical history for better diagnostics.

IX. CONCLUSION

The introduced U-Net framework demonstrates significant potential in advancing automated polyp recognition and segmentation of colonoscopy imaging. By integrating dilated convolutions and leveraging the strengths of the U-Net architecture, the model effectively captures fine details and global features, enabling accurate segmentation of polyps across diverse datasets. The use of multiple publicly available datasets ensures robustness and generalization, while pre-processing and optimization techniques enhance the model's performance and reduce overfitting.

Despite challenges such as computational demands and dependency on image quality, the framework addresses critical limitations of traditional methods and provides real-time support for clinicians, aiding in early recognition and diagnosis of colorectal cancer. With future enhancements such as real-time video integration, multi-class detection, and edge-device optimization, this architecture has the potential to revolutionize polyp detection, improving patient outcomes and clinical efficiency.

The **Dilated-U-Net-Seg** framework makes a significant contribution to medical image analysis by overcoming the shortcomings of traditional segmentation methods. By focused integration of advanced architectures and utilizing publicly available datasets, the model not only excels in controlled settingsbut also adapts well to clinical applications. Additionally, its ability to adjust parameters such as learning rate, batch size, and dropout probability highlights its flexibility for various training scenarios. This approach reduces dependence on manual intervention, improving efficiency and minimizing the potential for-human-error-in-healthcare.

REFERENCES

[1]. Sharma, R. Mehta, and S. Gupta, "AI-Based Polyp Detection and Segmentation Using Deep Learning for Enhanced Colonoscopy Accuracy," presented at the 2023 International Conference on Medical Imaging and Diagnostics (ICMID).
[2]. J. Liu, T. Zhang, and Y. Chen, "Deep Learning in Polyp Segmentation: A Study on U-Net and Its Variants," presented at the 2023 IEE Symposium on Biomedical Engineering (ISBE).
[3]. S. R. Patel, H. Singh, and N. Kumar, "Improving Deep Learning Approaches for Accurate Polyp Segmentation in Colonoscopy," presented at the 2022 International Conferenceon Biomedical Image Analysis (BIA).
[4]. Y. Zhang, X. Lin, and T. Huang, "AI-Driven Real-TimePolyp Detection and Segmentation for Improved Colonoscopy Procedures,"2023.
[5]. L. Chen, M. Zhao, and J. Xu, "Integration of Machine Learning and Colonoscopy: Advancements in Polyp Segmentation," in 2022 Journal of Medical Imaging andComputer Vision (JMICV).
[6]. T. Y. Kwon, A. Park, and J. Lee, "Efficient Deep Learning Architectures for Polyp Detection in Large-Scale Datasets," 2023.
[7]. P. K. Sharma and S. Das, "A Comprehensive Survey of AI- Based Techniques for Polyp Detection and Segmentation in Colonoscopy," 2021. This paper reviews state-of-the-art methods and evaluates their efficacy in real-world clinical scenarios.
[8]. V. Rao, N. Iyer, and S. Reddy, "Developing Robust AI Models for Colonoscopy Image Segmentation," 2022.
[9]. M. Nguyen, P. Pham, and T. Tran, "U-Net Variants for Accurate and Efficient Polyp Segmentation in Colonoscopy," 2022.
[10]. F. Ahmad and R. Khan, "Machine Learning Approaches for Enhanced Polyp Segmentation and Detection in Clinical Applications,"2021.
[11]. S. Wang, L. Yang, and H. Zhao, "Leveraging AI for Improved Polyp Detection Accuracy During Colonoscopy," 2023.
[12]. K. Patel, A. Shukla, and R. Mishra, "Advanced AI Models for Colon Cancer Screening: Challenges and Solutions," in 2022 Journal of Computational Medicine (JCM).
[13]. Y. Tanaka, M. Saito, and H. Kobayashi, "Real-Time Polyp Detection in Video Colonoscopy Using AI Algorithms," 2023.
[14]. R. Gupta and V. Shah, "Evaluating the Impact of AI on Early Polyp Detection in Colonoscopy Procedures," 2021.
[15]. Lee, J. Park, and S. Kim, "AI-Powered Colonoscopy Systems: Innovations in Polyp Detection and Diagnosis," 2023.
[16]. P. S. Reddy, T. Nishwa, and C. Kumar, "AI Applications for Real-Time Segmentation of Colonoscopy Images," 2022.
[17]. J. Zhou, W. Chen, and Y. Li, "Smart Colonoscopy Solutions: AI for Polyp Detection and Classification," 2022.
[18]. K. Gupta and S. Roy, "Integrating AI in Colonoscopy for Enhanced Diagnostic Accuracy," 2023.
[19]. N. Ahmed and F. Hassan, "AI-Assisted Systems for Colonoscopy: A Focus on Polyp Detection and Segmentation," 2023.
[20]. L. Zhang and H. Wang, "Deep Learning Approaches for Automated Polyp Detection to Enhance Colon Cancer Screening," 2022.