3D Reconstruction of Colonic Polyps: A Methodology for Improved Visualization and Analysis

By

Ishtiaque Ahmed
212-15-4217

Md. Saimim Islam Khan Hamim
212-15-4219

FINAL YEAR DESIGN PROJECT REPORT

This Report Presented in Partial Fulfillment of the Requirements for the

Degree of Bachelor of Science in Computer Science and Engineering

Supervised by

Dr. Md Zahid Hasan
Associate Professor
Department of Computer Science and Engineering
Daffodil International University

Co-Supervised by

Arpita Ghose Tusi
Lecturer
Department of Computer Science and Engineering
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY
Dhaka, Bangladesh

May 14, 2025

APPROVAL

This Project titled **"3D Reconstruction of Colonic Polyps: A Methodology for Improved Visualization and Analysis,"** submitted by **Ishtiaque Ahmed**, ID: **212-15-4217** and **Md. Saimim Islam Khan Hamim**, ID: **212-15-4219** to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on **14-05-2025**.

BOARD OF EXAMINERS

Dr. Sheak Rashed Haider Noori (SRH)
Board Chairman
Professor & Head, Department of CSE, FSIT
Daffodil International University

Mohammad Monirul Islam (MMI)
Internal Examiner
Assistant Professor, Department of CSE, FSIT
Daffodil International University

Bushra

Ms. Tasfia Anika Bushra (TAB)
Internal Examiner
Lecturer, Department of CSE, FSIT
Daffodil International University

Dr. Md. Zulfiker Mahmud (ZM)
External Examiner
Professor, Department of CSE
Jagannath University

DECLARATION

We hereby declare that this project has been done by us under the supervision of **Dr. Md Zahid Hasan, Associate Professor**, Department of Computer Science and Engineering, Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:

Dr. Md Zahid Hasan
Associate Professor
Department of Computer Science and Engineering
Daffodil International University

Co-Supervised by:

Arpita 14.05.25

Ms. Arpita Ghose Tusi
Lecturer
Department of Computer Science and Engineering
Daffodil International University

Submitted by:

Ishtiaque 13-05-2025

Ishtiaque Ahmed
Student ID: 212-15-4217
Department of Computer Science and Engineering
Daffodil International University

13.05.25

Md. Saimim Islam Khan Hamim
Student ID: 212-15-4219
Department of Computer Science and Engineering
Daffodil International University

ACKNOWLEDGEMENTS

This work would not have been possible without the support and contributions of many individuals over the past two semesters. We are deeply grateful to everyone who has assisted us in one way or another.

First, we express our heartfelt thanks and gratefulness to the almighty for His divine blessing making it possible for us to complete the **Final Year Design Project (FYDP)** successfully.

We are grateful and wish our profound indebtedness to **Dr. Md Zahid Hasan**, **Associate Professor**, Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh. Deep knowledge and keen interest of our supervisor in the field of **Computer Vision** to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

We would like to express our heartfelt gratitude to the Head of the Department of Computer Science and Engineering, for his kind help in finishing our project and also to other faculty members and the staff of the Department of Computer Science and Engineering, Daffodil International University.

We would like to thank our entire course mates at Daffodil International University, who took part in this discussion while completing the coursework.

Finally, we must acknowledge with due respect the constant support and patience of our parents.

ABSTRACT

Colorectal cancer is one of the most common and life-threatening diseases worldwide, and colonoscopy remains the most effective method for early detection and removal of precancerous polyps from the human colon. However, a major limitation of conventional colonoscopy is the restricted field of view of the endoscope, which often prevents the complete visualization of a polyp's surface. Due to complex polyp shapes or difficult camera angles, certain regions of the polyp may remain unobserved during the procedure, increasing the risk of misdiagnosis or incomplete removal. In this research, a 3D reconstruction-based approach is proposed to improve the visualization and analysis of colonic polyps using 2D endoscopic images. The overall framework begins with organizing the dataset based on lesion and video annotations, followed by automatic frame selection using mask-based conditions, image contrast, feature points, and depth quality scores to select the most informative frames for reconstruction. To improve image quality, reflection removal is applied using the EndoSRR framework. Depth maps are then estimated using ZoeDepth, a state-of-the-art monocular depth estimation model. Using these depth maps, dense point clouds are generated, and Region of Interest (ROI) extraction is performed for accurate reconstruction of the polyp surface. Clean 3D meshes are constructed using the Ball Pivoting Algorithm (BPA), along with refinement and hole-filling techniques. Additionally, the reconstructed 3D models are analyzed through various geometric and structural features, including shape descriptors and curvature. Finally, silhouette projection and alignment validation are performed to compare the 3D reconstructed model with the original 2D mask, ensuring accuracy and reliability of the reconstruction. This 3D reconstruction pipeline provides enhanced visualization of colonic polyps, enabling better inspection of previously unobserved regions and supporting further quantitative analysis for future computer-aided diagnosis systems.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

Colorectal cancer is the second most commonly occurring cancer in women and the third most commonly occurring cancer in men all over the world. Colonoscopy is considered as the gold-standard method to detect changes and remove precancerous polyps in the large intestine. Medical imaging has revolutionized modern healthcare by enabling non-invasive visualization of internal body structures, facilitating accurate diagnosis, treatment planning, and monitoring of various diseases. Techniques such as X-rays, computed tomography (CT), magnetic resonance imaging (MRI), ultrasound, and endoscopic imaging have been instrumental in providing two-dimensional (2D) representations of anatomical structures. While these 2D images offer valuable insights, they often fall short in conveying the complex spatial relationships inherent in human anatomy [1]. Despite advancements in medical imaging technology, a significant portion of clinical diagnosis still relies on 2D images. These images possess inherent limitations when representing complex anatomical structures, especially in capturing depth, volume, and spatial relationships. The limitations may lead to inadequate evaluation of essential areas or impede precise diagnosis. This discrepancy has prompted the creation of three-dimensional reconstruction techniques aimed at converting two-dimensional medical images into detailed three-dimensional representations. These models enhance depth perception and improve the analysis of morphological features, facilitating superior visualization of anatomical structures. In addition to enhancing diagnosis, 3D models are essential for surgical planning, the advancement of computer-aided diagnosis (CAD) systems, and medical education. Early detection of colorectal cancer through gastrointestinal endoscopy relies on precise identification and characterization of colonic polyps. Colonoscopy is regarded as the most precise technique for identifying and excising polyps; however, it has a restricted field of vision. The two-dimensional perspective of endoscopy may hinder the endoscopist's ability to accurately evaluate the dimensions, shape, and surface characteristics of a polyp obscured by colonic folds or situated in difficult locations [2]. Moreover, during colonoscopy, it is common that certain regions of a polyp may remain unobserved due to occlusions or restricted angles of visualization. This can lead to missed diagnoses or incomplete treatment of potentially cancerous lesions. Therefore, a system that can reconstruct the 3D structure of a polyp using the available 2D endoscopic frames would greatly assist clinicians in visualizing the entire surface of the polyp, both during the procedure and in offline analysis.

In this context, various advantages come from 3D reconstruction of polyp using endoscopic pictures. First it makes it possible to create a full-surface textured map of the polyp, therefore facilitating the detection of uninspected areas. Second, it improves morphological analysis, therefore enabling doctors to more precisely examine surface roughness, topology, and volume. Integrating automated feature extraction and quantitative measurements is beneficial in computer-aided diagnosis systems for colorectal cancer screening. Reconstructing a 3D model from monocular 2D endoscopic images presents challenges related to specular reflections, variations in illumination, occlusions, and the presence of textureless surfaces. These components can distort visual content, complicate depth estimation, and hinder the development of point clouds and meshes.

This study introduces an automated and efficient framework for the three-dimensional reconstruction of colonic polyps using monocular two-dimensional endoscopic images. The reconstruction of an accurate three-dimensional structure from limited and variable-quality two-dimensional images is complex, especially in the medical imaging domain where precision and reliability are essential. This process is complicated by several practical challenges, including varying illumination conditions within the colon, specular reflections on wet polyp surfaces, non-rigid tissue movement, occlusions caused by folds or fluids, and the restricted field of view characteristic of endoscopic procedures. This research proposes a methodology that integrates multiple advanced and robust techniques, which operate sequentially and collaboratively to address the identified obstacles. The process initiates with the automated selection of frames from colonoscopy video streams, guided by intelligent mask analysis criteria[3]. The selection is influenced by the presence of the lesion as well as the image quality, which includes contrast, ORB feature richness, and depth map quality score. This guarantees the selection of only the most informative and clinically relevant frames for reconstruction. The reflection removal stage utilizes EndoSRR-based methods to remove undesirable specular highlights that can distort depth estimation processes [4]. This research utilizes ZoeDepth [5], an advanced deep learning model, for precise depth estimation, enabling the generation of detailed depth maps from single monocular images. Subsequently, precise lesion localization is achieved through the use of Region of Interest (ROI), thereby eliminating extraneous background noise. Point cloud creation, adhering to geometric projection guidelines, results from calculated depth maps and is subsequently enhanced through noise removal and outlier filtering techniques. The Ball Pivoting Algorithm (BPA) is recognized for its efficacy in generating meshes from scattered point clouds and is employed to construct the surface model of the polyp. To produce a high-quality, clinically interpretable 3D model, the resulting mesh undergoes additional cleaning, hole-filling, and polishing processes. This research emphasizes quantitative analysis of the polyp model alongside surface reconstruction. Geometric features are extracted from the point cloud and mesh, including global shape descriptors, critical curvature. The computation of curvature and slope along the geodesic path between maximum and minimum elevation points of the polyp offers insights into the surface complexity and morphology of the lesion, which are essential for clinical classification, risk assessment, and surgical planning [3].

The implementation of a comprehensive and automated 3D reconstruction pipeline provides various advantages in practical clinical settings. This methodology offers endoscopists an enhanced visualization environment, allowing for comprehensive viewing, manipulation, and analysis of the polyp's entire surface topology, surpassing the constraints of conventional 2D imaging. Regions previously challenging or unfeasible to inspect during live procedures because of occlusions or blind spots can now be reconstructed and assessed comprehensively [2]. Secondly, the reconstructed 3D information is essential for improving Computer-Aided Diagnosis (CAD) systems by offering precise surface and volume measurements, shape analysis, and detection of structural abnormalities, all of which are crucial for assessing the malignancy potential of a polyp. The utilization of accurate 3D visual content significantly enhances comprehension compared to static 2D images. This system supports lesion documentation, enables follow-up comparisons, and aids in surgical training and medical education. Advanced feature extraction methods, such as curvature and slope analysis, enhance the development of intelligent classification systems capable of distinguishing between benign and malignant polyps based on surface characteristics, thereby significantly advancing automated colorectal cancer screening tools. This study may enhance clinical decision-making, reduce the likelihood of overlooked lesions, assist in pre-operative planning, facilitate patient monitoring over time, and contribute to the development of future artificial intelligence-based diagnostic tools. This study establishes a solid foundation for future medical imaging research, encompassing automated polyp classification, early cancer detection systems, polyp growth prediction models, and real-time 3D visualization in live colonoscopy procedures. Contemporary medical imaging and colorectal cancer diagnosis significantly rely on a 3D reconstruction approach that integrates advanced computer vision techniques with essential healthcare requirements [1].

1.2 Motivation

The limits of conventional colonoscopy in precisely viewing the entire surface structure of colonic polyps are addressed in this work. Conventional colonoscopy is the endoscopist using two-dimensional (2D) images, usually impacted by limited angles, blockages, poor lighting, or the complex colonic polyp placement. The recognized difficulties could cause essential polyp areas to be neglected, so influencing the diagnosis accuracy and the way therapeutic treatments are carried out, so causing mistakes. This work aims to build a three-dimensional (3D) model of polyps using 2D endoscopic pictures in order to enhance polyp surface view. In addition to aiding in the visualization of obstacles, three-dimensional representation provides critical morphological information, including shape, structure, size, and surface roughness, which are challenging to acquire from traditional two-dimensional photographs. This 3D reconstruction framework utilizes geometric features and surface characteristics to enhance the assessment of polyp malignancy potential. This additional layer of study would greatly enhance the ability of endoscopists to make correct and confident clinical decisions. Research in this area begins with the establishment of a powerful 3D reconstruction framework. This platform allows researchers to design intelligent virtual reality simulations for medical education and training, automated polyp classification systems based on artificial intelligence, and detection systems. By improving surgeons' knowledge of size, depth, and colonal position, a detailed 3D model of the polyp greatly helps pre-operative planning. This could improve polyp removal techniques' efficiency, safety, and simplicity of use. By means of clinical practices and future advancements in artificial intelligence-driven medical diagnosis, automated 3D analysis systems, and enhanced surgical planning tools in colorectal cancer treatment, this paper closes the gap between computer vision technology and healthcare requirements.

1.3 Objectives

The purpose of this thesis is to generate three-dimensional (3D) models of colonic polyps from two-dimensional (2D) endoscopic pictures by means of an automated and efficient method. This method aims to surpass the boundaries of traditional colonoscopy methods and thus let doctors engaged in treatment planning by raising the accuracy of polyp detection and diagnosis.

- To develop a fully automated framework that converts 2D endoscopic images into accurate 3D models of colonic polyps to improve diagnostic and clinical decision-making.
- To evaluate the performance and accuracy of the developed framework by comparing the reconstructed 3D models with ground truth data and assessing the clinical potential of the 3D polyp models for improved diagnosis and treatment planning.
- To design and implement an automated pipeline for 3D reconstruction of colonic polyps, leveraging advanced techniques such as frame selection, depth map generation, and reflection removal to enhance the quality of the reconstructed models.
- To enhance the visualization and analysis of polyps by utilizing advanced techniques like depth map generation, frame selection, and geometric feature extraction.

1.4 Methodology

This work aims to provide a completely automated and efficient framework for rebuilding three-dimensional (3D) models of colonic polyps from two-dimensional (2D) endoscopic pictures. By improving the visualization, analysis, and interpretation of polyps, this suggested approach seeks to surpass the restrictions of conventional colonoscopy and so help doctors in improved diagnosis and treatment planning. Starting with data preparation using the publicly accessible PICCOLLO dataset, the approach. Along with their matching mask annotations, this dataset has a large variety of endoscopic polyp images. Carefully arranged and split into training and testing subsets, the dataset guarantees heterogeneity in terms of polyp size, shape, and texture for more generalised model development. Following data preparation, the most instructive frames for reconstruction are chosen automatically using a frame selection algorithm. Mask conditions (like area, center, and boundary assessment), image contrast evaluation for better visibility, ORB (Oriented FAST and Rotated BRIEF) feature richness analysis for selecting frames with more texture detail, and depth quality estimation for ensuring that only frames with meaningful depth information are considered. Reflection removal methods as the EndoSRR framework are utilized to reduce visual artifacts caused by illumination and fluid interference inside the colon, therefore guaranteeing a better input for the next reconstruction procedure. Following the selection and enhancement of high-quality frames, ZoeDepth, a state-of-the-art monocular depth estimation algorithm based on modern deep learning, is used. This model produces consistent depth maps from the 2D frames, therefore offering essential spatial information needed for 3D reconstruction. Then bitwise operations and mask morphology help to extract the Region of Interest (ROI). This technique guarantees that only the polyp area is identified and taken into account for 3D reconstruction, therefore removing pointless background data that could compromise accuracy. Dense point clouds expressing the 3D structure of the polyp are produced with the depth maps generated and ROI extracted. These point clouds are produced from 3D coordinates generated from depth information then using geometric approaches to improve their quality. The following stage uses the Ball Pivoting Algorithm (BPA) to generate mesh from the point clouds. This method precisely and smoothly creates a 3D mesh model of the polyp. Any anomalies are eliminated using further cleaning and hole-filling methods, therefore guaranteeing that the mesh is fit for clinical study. Several geometric characteristics are extracted in order to acquire a better knowledge of the rebuilt polyp model. This covers worldwide geometric features that give an overall description of the structure, shape distribution histograms that statistically reflect the shape characteristics, and curvature and slope analysis using geodesic paths and differential geometry techniques to evaluate surface variations and complexity. Projection-based evaluation helps to validate the reconstructed 3D model's accuracy at last. Reconstructed 3D mesh is projected onto 2D planes and compared with original annotated masks from PICCOLLO dataset. Ground truth alignment accuracy between the reconstructed model and the ground truth is quantitatively measured using evaluation criteria including Dice Coefficient and Intersection over Union (IoU). This thorough effort guarantees correct reconstruction of colonic polyps and offers necessary geometric insights that can be very important for surgical planning, future AI-based diagnostic instruments, and medical research developments in gastrointestinal imaging.

1.5 Project Outcome

Many noteworthy and significant findings resulting from this research project help the field of medical imaging as well as useful therapeutic applications. This method has clearly demonstrated how especially in the framework of colorectal cancer screening and diagnosis, 2D endoscopic images can increase the understanding, visualization, and analysis of colonic polyps. The following lists the primary findings of this study:

1. Design of a robust, totally automated basis for 3D reconstruction from monocular 2D endoscopic pictures. By guaranteeing consistency, economy, and accuracy in reconstruction of complex polyp structures, our automated method minimizes manual participation and guarantees a great step forward for including advanced computer vision into clinical practice.

2. The effective development of improved visualization technologies enabling doctors to have a whole perspective of polyp shape. This covers a thorough examination of polyp surface structure, size, and form even in areas challenging for conventional colonoscopy. These improved visualizations are quite crucial in lowering the possibility of missing significant lesions and provide doctors more confidence during diagnosis and therapy planning.

3. The extraction of several significant geometric features from the reconstructed 3D models: form descriptors, curvature, slope, and global geometry features. These retrieved features are quite useful since they provide fresh opportunities for classifying polyps more precisely, differentiating between benign and malignant formations, and allowing additional analysis in next artificial intelligence-based research projects.

4. The assessment of the reconstructed 3D polyp models through Dice Coefficient and intersection over Union (IoU) to ensure reliability and accuracy. The validation methods reinforce the credibility of the proposed framework by ensuring structural coherence between the reconstructed 3D models and the original 2D images.

5. The building of a robust basis for the future integration of this 3D reconstruction framework into computer-aided diagnostic (CAD) systems for colorectal cancer detection and screening. This approach creates new possibilities for sophisticated diagnostics tools able to offer surgeons planning guidance, polyp risk assessment, and automated analysis. Furthermore, the framework gives next researchers a scalable and flexible platform that might be developed or used to manage different medical imaging issues.

1.6 Organization of the Report

This report is organized into six comprehensive chapters to provide a clear and systematic presentation of the entire research study. Each chapter is carefully structured to reflect the progressive development of this work, starting from conceptual understanding to practical implementation and evaluation.

The first chapter introduces the research topic, highlighting the significance of 3D reconstruction of colonic polyps from 2D endoscopic images. It presents the background, research motivation, objectives, adopted methodology, expected project outcomes, and overall report organization. This chapter forms the basis for understanding the aim and scope of the research.

Chapter 2 presents the Literature Review, where existing research works, methods, and technologies related to 3D reconstruction in medical imaging, particularly in the field of endoscopy, are discussed. It focuses on similar applications, related research works, and provides a gap analysis to identify the necessity and novelty of this research.

Chapter 3 describes the Research Methodology and Design Specification adopted in this study. It elaborates the complete working procedure of the proposed system, covering dataset preparation, frame selection, reflection removal, depth estimation, ROI extraction, point cloud generation, mesh construction, and feature extraction processes. In addition, it outlines the system requirements, project plan, task allocation, and design diagrams used in the project.

Chapter 4 focuses on Implementation and Results, where the practical execution of the proposed methodology is explained. It describes the environment setup, testing and evaluation procedures, performance analysis, and presents the experimental results. Both qualitative and quantitative evaluations are provided based on the reconstructed 3D polyp models.

Chapter 5 discusses Engineering Standards and Design Challenges along with the Impact on Society, Environment, and Sustainability. It addresses the ethical aspects, compliance with hardware and software standards, and the real-world implications of this research in terms of medical advancements and future sustainability.

Finally, Chapter 6 presents the Conclusion and Future Work, summarizing the overall research findings, contributions, and limitations. It also outlines possible future research directions and emphasizes how the current work can pave the way for advanced studies in 3D reconstruction, automated diagnosis systems, and improved colorectal cancer screening technique.

Chapter 2

Background

2.1 Introduction

Colorectal cancer (CRC) remains a significant global health concern, being one of the leading causes of cancer-related deaths in both men and women. Early detection and precise treatment of colonic polyps, the primary precursors to CRC, are crucial for reducing mortality rates. Colonoscopy is considered the gold-standard procedure for detecting and removing these polyps, but conventional colonoscopy techniques are largely dependent on two-dimensional (2D) imaging, which limits the visualization of the complete structure of the polyps. Medical imaging technologies like CT, MRI, and endoscopy have significantly improved clinical diagnostics [5]. However, the inherent limitations of 2D endoscopic images, especially in capturing depth, surface irregularities, and spatial structure, pose challenges during polyp detection, evaluation, and treatment planning [6]. These challenges have led to the development of three-dimensional (3D) reconstruction techniques to generate detailed and spatially accurate models from 2D images. 3D reconstruction provides clinicians with enhanced visualization of polyp morphology, allowing for better assessment of size, shape, and surface characteristics. It also creates prospects for offline inspection, surgical planning, and integration into computer-aided diagnostic (CAD) systems [7].

Such reconstruction methods are becoming more and more important in gastrointestinal imaging since they can help to promote correct diagnosis and enhance patient outcomes [8]. Reconstruction of reliable 3D models from monocular endoscopic data remains a difficult chore despite recent developments. Low texture areas, specular reflection, and occlusions that hamper depth estimation and 3D modeling procedures abound in endoscopic pictures [9]. This work proposes an automated method employing the PICCOLLO dataset for the 3D reconstruction of colonic polyps, therefore addressing these difficulties. Using cutting-edge methods including frame selection based on contrast, ORB feature richness, reflection removal using EndoSRR, depth estimate using ZoeDepth [10], ROI extraction, and mesh generation via the Ball Pivoting Algorithm (BPA), the framework In order to enable future study in AI-driven polyp classification systems, the rebuilt 3D models help not only with better clinical vision but also the extraction of geometric data including curvature, roughness, and form descriptors. This work provides a basis for further research in real-time 3D reconstruc-tion, AI-based polyp classification, and enhanced colorectal cancer detection technologies [6] as well as helps to close the gap between 2D imaging limits and the demands of modern clinical diagnostics..

2.2 Literature Review

The advancements in computer-aided diagnosis (CAD) systems for colorectal polyp detection and classification have gained significant attention due to the increasing need for early detection of colorectal cancer. Several existing research works have focused on polyp segmentation, classification, depth estimation, and 3D reconstruction techniques using various machine learning and deep learning approaches. However, most of the studies have certain limitations when compared to the comprehensive framework proposed in this research, which incorporates an end-to-end pipeline starting from Paris classification-based dataset splitting to validation through silhouette and projection analysis. Yengec-Tasdemir et al.[5] introduced a supervised contrastive learning-based framework for polyp classification using histopathological images but lacked any 3D reconstruction or depth analysis. Hui Chen et al.[6] proposed 3D point cloud reconstruction from a single image through image retrieval, yet their method is not polyp-specific and lacked mesh generation, feature extraction, or depth analysis techniques related to medical images. Nur-A-Alam et al. [7] developed an automatic colorectal polyp detection model using hybrid fused features but did not consider any 3D reconstruction or point cloud generation in their framework.

Huang et al. [8] integrated SimCLR-based contrastive learning for polyp classification, but their work also remained limited to 2D analysis without any depth estimation or 3D modeling. Hossain et al. [9] proposed DeepPoly for polyp segmentation and classification but excluded any 3D reconstruction pipeline. Rock et al. [11] focused on 3D shape completion from depth images using thin-plate spline deformation but did not include polyp-based analysis or point cloud feature extraction. Chowa et al. [10] classified breast tumors by transforming ROI into 3D meshes using Point-e and applied GAT for classification, whereas their methodology is not applicable for colonoscopy-based polyp reconstruction. Younas et al. [12] applied deep ensemble learning for polyp classification but missed depth, point cloud, or mesh generation analysis. Krenzer et al. [13] focused on polyp classification using Paris and NICE schemes but ignored any 3D analysis. Cao et al. [14] developed an adaptive learning model for polyp classification using CTC images but lacked real-world colonoscopy depth estimation or mesh generation.

Ahmad et al. [23] proposed shape-from-shading (SFS) and shape-from-focus (SFF)-based gastrointestinal 3D reconstruction but did not perform any Paris classification, best frame

Table 2.1: Summary of Literature Reviewed

Author(s)	**Year**	**Title**	**Methodology**	**Key Findings**
Yengec-Tasdemir et al. [5]	2024	Effective Colorectal Polyp Classification Using Supervised Contrastive Learning	Deep Learning (BiT Model with Supervised Contrastive Learning)	Achieved 86.2% accuracy on custom dataset and 70.1% on UniToPatho dataset for polyp classification.
Hui Chen et al. [6]	2020	3D Point Cloud Generation Reconstruction from Single Image Based on Image Retrieval	Image Retrieval + Deep Learning	Generated accurate 3D point clouds from single image using improved reconstruction techniques.
Nur-A-Alam et al. [7]	2022	An Automatic System to Detect Colorectal Polyp Using Hybrid Fused Method	Hybrid Method (EMD, Wavelet, CNN)	Achieved 99.45% accuracy for polyp detection using hybrid fused method.
Huang et al. [8]	2023	Polyp Classification in Endoscopic Images Using Self-Supervised Structured Learning	Self-Supervised Learning	Improved classification accuracy using SimCLR and Look-into-Object modules with limited labeled data.
Hossain et al. [9]	2023	DeepPoly: Deep Learning-Based Polyps Segmentation and Classification	Deep Learning (DoubleU-Net & ViT)	Achieved 99% accuracy and 0.956 dice coefficient in polyp segmentation and classification.
Rock et al. [11]	2015	Completing 3D Object from Single Depth View	Shape Completion (Exemplar-Based)	Robust shape completion using 3D model retrieval and deformation.
Chowa et al. [10]	2023	Tumor Classification by 3D Mesh Representation	3D Reconstruction + GAT	Achieved 99.34% accuracy using 3D mesh of tumor and Graph Attention Networks.
Younas et al. [12]	2023	Deep Ensemble Learning for Colorectal Polyp Classification	Ensemble CNN	Achieved 96.3% accuracy on UCI dataset and 81.2% on PICCOLO dataset.

Table 2.2: Summary of Literature Reviewed

Author(s)	**Year**	**Title**	**Methodology**	**Key Findings**
Krenzer et al. [13]	2023	Dual Classification of Polyps in Endoscopy	Deep Learning	Achieved 89.35% accuracy for Paris classification and 81.34% for NICE classification.
Cao et al. [14]	2022	Adaptive Learning for Polyp Classification in CT Colonography	Machine Learning	Achieved 0.925 AUC using multi-scale texture features and hierarchical classifier.
Zhang et al. [15]	2023	ColDE: Depth Estimation from Colonoscopy Video	Deep Learning (ColDE Framework)	High-quality depth map generation using geometric consistency losses.
Yousuf et al. [16]	2022	Polyp Detection Using 3D Geometric Features	CNN + Geometric Features	Improved detection using curvature, depth, and normal features.
Duran et al. [17]	2021	LiDAR Point Cloud Classification Using Machine Learning	Machine Learning	Achieved 96% accuracy using geometric feature extraction (linearity, planarity, sphericity).
Deka et al. [18]	2023	Dense 3D Reconstruction of Endoscopic Polyp	ORB Matching + Depth Estimation	Generated dense 3D reconstruction using ORB features and photometric optimization.
Ahmad et al. [19]	2017	Endoscopic 3D Surface Reconstruction from Single Images	SFS + Anisotropic Diffusion	Improved depth estimation and surface reconstruction from single images.
Wesp et al. [20]	2021	Premalignant Polyp Classification in CT Colonography	CNN	Achieved 0.83 AUC in polyp classification using CT images.

Table 2.3: Summary of Literature Reviewed

Author(s)	**Year**	**Title**	**Methodology**	**Key Findings**
Nogueira-Rodríguez et al. [21]	2022	Shape Completion using 3D Encoder-Predictor CNNs	Deep Learning (Shape Completion)	Completed 3D shape from partial voxel inputs using 3D CNN and shape synthesis.
Widya et al. [22]	2022	Whole Stomach 3D Reconstruction from Monocular Endoscopy	SfM + Poisson Mesh Generation	Reconstructed 3D stomach with frame localization and texture mapping.
Ahmad et al. [23]	2018	3D Reconstruction of Gastrointestinal Regions Using SFS & SFF	Shape from Shading & Shape from Focus	Reconstructed endoscopic surfaces for enhanced visualization.
Nadeem et al. [24]	2020	Depth Estimation and Polyp Detection in Optical Colonoscopy	Depth Dictionary & Machine Learning	Predicted depth from RGB images and improved polyp detection accuracy.

selection, or feature extraction from point clouds. Nadeem et al. [24] generated depth maps for polyp detection from colonoscopy images but did not consider mesh generation or complete 3D modeling. Zhang et al. [15] proposed ColDE for self-supervised depth estimation and mesh generation but did not include point cloud feature extraction or ROI extraction. Yousuf et al. [16] worked on 3D geometric feature-based polyp detection but focused on CT images, not real colonoscopy videos, and lacked ROI extraction and validation using silhouette projection. Duran et al. [17] proposed point cloud classification using geometric features but targeted LiDAR and photogrammetric data, not medical images. Deka et al. [18] performed dense 3D reconstruction using monocular video but did not consider Paris classification or point cloud feature extraction. Ahmad et al. [19] used SFS methods for single-image 3D reconstruction but excluded mesh generation, curvature, or slope analysis. Wesp et al. [20] differentiated adenomas from hyperplastic polyps using deep learning in CT colonography but did not perform depth estimation or 3D reconstruction from colonoscopy images. Nogueira-Rodriguez et al. [21] used YOLOv3-based real-time polyp detection but without any 3D modeling or depth analysis. Widya et al. [22] performed whole stomach 3D reconstruction using Structure-from-Motion (SfM) but did not apply Paris classification, best frame selection, reflection removal, or detailed point cloud analysis.

In contrast to the above studies, the proposed research presents a complete 3D reconstruction framework for colorectal polyp analysis. The framework uniquely applies Paris classification-based dataset splitting, mask condition-based best frame selection, reflection removal using EndoSRR, depth map generation using ZoeDepth, 3D point cloud generation, mesh generation using BPA and Poisson, detailed point cloud feature extraction, curvature and slope analysis using Dijkstra path, ROI extraction using void region mask, and final validation through silhouette and projection analysis. No existing literature combines all these techniques together in a single pipeline specifically targeted at colorectal polyp reconstruction and analysis from real-world colonoscopy images.

2.2.1 Similar Applications

Several existing research studies, methodological contributions, and frameworks have been proposed for colorectal polyp detection, classification, and 3D reconstruction using computer vision and deep learning techniques. However, these applications are mostly limited to specific stages of polyp analysis, such as 2D classification, segmentation, or depth estimation, without integrating a complete 3D reconstruction pipeline. Yengec-Tasdemir et al. [5] introduced a supervised contrastive learning-based classification framework for colorectal polyps using histopathological images. Similarly, Younas et al. [12] developed an ensemble learning-based classification method for polyps, while Krenzer et al. [13] applied Paris and NICE classification systems using Vision Transformer models. Although these methods provided effective classification results, none of them performed depth estimation, 3D reconstruction, or analysis on colonoscopy video data. Hui Chen et al. [6] and Rock et al. [11] proposed 3D point cloud generation and shape completion techniques using image retrieval and depth image matching, respectively. Their methodologies were effective for general object reconstruction but were not specifically designed for polyp reconstruction or medical imaging datasets. Zhang et al. [15] introduced ColDE, a depth estimation framework for colonoscopy videos, generating high-quality depth maps and mesh reconstruction. However, their system lacked point cloud feature extraction, curvature analysis, or silhouette-based validation steps. Ahmad et al. [23], [19] applied shape-from-shading (SFS) and shape-from-focus (SFF) techniques for gastrointestinal 3D reconstruction but did not perform polyp classification, best frame selection, or point cloud-based analysis. Widya et al. [22] focused on whole stomach 3D reconstruction from monocular endoscopy video using Structure-from-Motion (SfM) and Poisson surface mesh generation, while Deka et al. [18] implemented dense 3D reconstruction of polyps from monocular video using photometric optimization techniques. However, these works lacked essential processes such as Paris classification, reflection removal, and detailed geometric feature analysis from the point cloud. On the other hand, some methods like Nadeem et al. [24] and Yousuf et al. [16] worked on depth estimation and 3D geometric feature extraction for polyp detection, mainly targeting CT colonography or virtual colonoscopy data rather than real colonoscopy videos. Apart from academic research, various mobile and web-based applications exist for polyp detection, such as AI-based colonoscopy assistants integrated into endoscopy devices. These tools mainly focus on real-time polyp localization and classification, using 2D detection frameworks like YOLOv3 or SSD, but lack depth generation, 3D modeling, or analytical features like curvature analysis.

Overall, although several studies have contributed to different components of polyp detection or 3D reconstruction, no existing work provides a complete end-to-end pipeline integrating Paris classification, best frame selection, reflection removal, depth map generation, 3D point cloud generation, mesh generation, point cloud feature extraction, curvature and slope analysis, ROI extraction, and validation through silhouette and projection analysis. The proposed framework in this research uniquely integrates all these components to create a comprehensive polyp reconstruction and analysis system from real colonoscopy images and videos.

2.2.2 Related Research

The growing integration of computer vision (CV) and deep learning (DL) techniques in medical imaging has greatly influenced the research and development of advanced systems for the detection, classification, and analysis of colorectal polyps. Numerous studies have focused on different stages of polyp analysis such as classification, segmentation, depth estimation, and partial 3D reconstruction. However, most of these works are limited to specific tasks and fail to offer a complete pipeline covering the entire 3D analysis of polyps.

Polyp Classification Approaches

Most of the early studies concentrated on polyp classification based on 2D colonoscopy or histopathological images. Yengec-Tasdemir et al. [5] applied a supervised contrastive learning strategy for polyp classification using BiT models. Similarly, Younas et al. [12] proposed an ensemble learning-based approach, while Krenzer et al. [13] incorporated both Paris and NICE classification systems for enhanced diagnosis. However, these works focused only on classification tasks without integrating depth estimation or 3D modeling.

3D Point Cloud and Mesh Generation Methods

To address structural analysis, Hui Chen et al. [6] introduced a 3D point cloud reconstruction framework using image retrieval. Rock et al. [11] developed shape completion techniques using depth images for general objects, while Chowa et al. [10] generated 3D meshes of breast tumors for classification. Though effective, these methods were not specifically developed for colonoscopy images or polyp reconstruction.

Depth Map Estimation and 3D Reconstruction

In terms of depth estimation, several frameworks have emerged targeting colonoscopy data. Zhang et al. [15] proposed ColDE for depth map generation from RGB colonoscopy videos, while Nadeem et al. [24] used an RGB-Depth dictionary-based approach for depth estimation. Ahmad et al. [23], [19] applied SFS and SFF methods for gastrointestinal surface reconstruction from a single image, but these works lacked point cloud feature analysis and validation techniques.

Polyp Detection with 3D Geometric Features

Other studies, such as Yousuf et al. [16], focused on combining 3D geometric features like curvature, depth, and normal maps with CNNs for polyp detection in virtual colonoscopy images. Duran et al. [17] utilized geometric features for classifying LiDAR point clouds but did not apply their method to medical images.

Dense 3D Reconstruction from Endoscopic Video

More recent research aimed to reconstruct dense 3D models from monocular endoscopy videos. Deka et al. [18] implemented dense 3D reconstruction using ORB feature matching and depth estimation. Widya et al. [22] developed a whole stomach 3D reconstruction pipeline using Structure-from-Motion (SfM) and mesh generation techniques, mainly for visualization purposes.

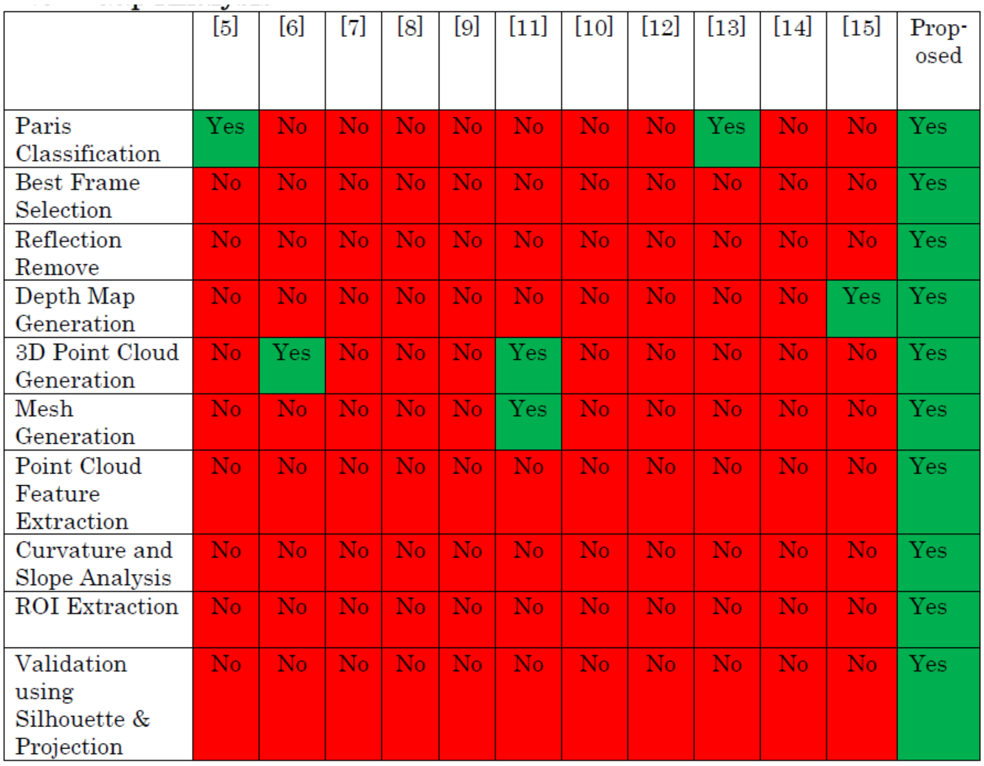
Real-Time Detection and Deep Learning Models

In the domain of real-time detection, Nogueira-Rodríguez et al. [21] proposed a YOLOv3-based framework integrated with object tracking for real-time polyp detection. Wesp et al. [20] utilized 3D CNNs for classifying premalignant polyps in CT colonography without considering depth estimation or 3D modeling.

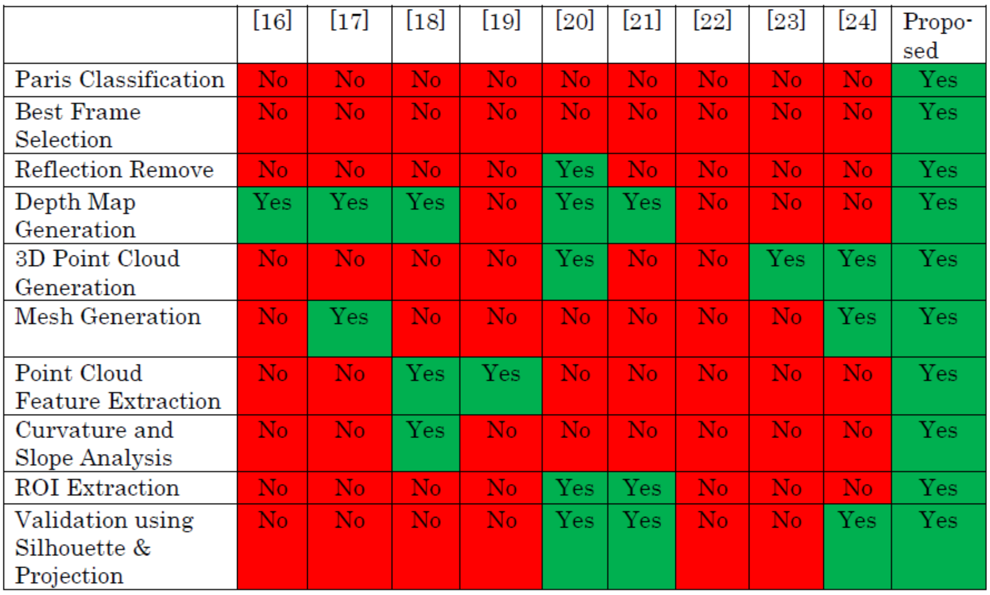
2.3 Gap Analysis

The gap analysis highlights that most previous studies only cover individual parts of the 3D reconstruction process. While some focus on depth estimation or mesh generation, they often overlook key steps like selecting the best frames, removing reflections, or analyzing shape features. These missing elements are important for producing high-quality 3D visuals and reliable structural insights. Our proposed method brings these components together into one complete pipeline. By doing so, it not only improves visual reconstruction but also supports more accurate analysis, offering a more practical and clinically useful solution.

Table 2.4: Gap analysis table based on comparison across existing studies, highlighting the presence or absence of key components such as Paris classification, 3D reconstruction steps, and feature extraction



	[5]	[6]	[7]	[8]	[9]	[11]	[10]	[12]	[13]	[14]	[15]	Proposed
Paris Classification	Yes	No	No	No	No	No	No	No	Yes	No	No	Yes
Best Frame Selection	No	No	No	No	No	No	No	No	No	No	No	Yes
Reflection Remove	No	No	No	No	No	No	No	No	No	No	No	Yes
Depth Map Generation	No	No	No	No	No	No	No	No	No	No	Yes	Yes
3D Point Cloud Generation	No	Yes	No	No	No	Yes	No	No	No	No	No	Yes
Mesh Generation	No	No	No	No	No	Yes	No	No	No	No	No	Yes
Point Cloud Feature Extraction	No	No	No	No	No	No	No	No	No	No	No	Yes
Curvature and Slope Analysis	No	No	No	No	No	No	No	No	No	No	No	Yes
ROI Extraction	No	No	No	No	No	No	No	No	No	No	No	Yes
Validation using Silhouette & Projection	No	No	No	No	No	No	No	No	No	No	No	Yes



	[16]	[17]	[18]	[19]	[20]	[21]	[22]	[23]	[24]	Proposed
Paris Classification	No	No	No	No	No	No	No	No	No	Yes
Best Frame Selection	No	No	No	No	No	No	No	No	No	Yes
Reflection Remove	No	No	No	No	Yes	No	No	No	No	Yes
Depth Map Generation	Yes	Yes	Yes	No	Yes	Yes	No	No	No	Yes
3D Point Cloud Generation	No	No	No	No	Yes	No	No	Yes	Yes	Yes
Mesh Generation	No	Yes	No	No	No	No	No	No	Yes	Yes
Point Cloud Feature Extraction	No	No	Yes	Yes	No	No	No	No	No	Yes
Curvature and Slope Analysis	No	No	Yes	No	No	No	No	No	No	Yes
ROI Extraction	No	No	No	No	Yes	Yes	No	No	No	Yes
Validation using Silhouette & Projection	No	No	No	No	Yes	Yes	No	No	Yes	Yes

2.4 Summary

This chapter has comprehensively explored and analyzed the existing literature related to colorectal polyp detection, classification, depth estimation, and 3D reconstruction techniques using various computer vision and deep learning methods. Several studies were found to contribute to individual segments of polyp analysis, such as 2D classification, depth map generation, or partial 3D modeling. However, most of these works were found to be limited to specific tasks without offering a complete end-to-end 3D reconstruction framework. A detailed summary of the reviewed literature was provided, highlighting their methodologies and key findings. The research gap analysis clearly revealed that none of the existing works have integrated all the essential components together into a single pipeline. These missing components include Paris classification-based dataset splitting, mask condition-based best frame selection, reflection removal using EndoSRR, depth map generation using ZoeDepth, 3D point cloud generation, mesh generation using BPA & Poisson, detailed point cloud feature extraction, curvature and slope analysis, ROI extraction using void region masks, and validation through silhouette and projection analysis. The proposed work in this research is therefore novel and unique in its comprehensive integration of all these techniques, creating a robust and fully automated framework specifically designed for colorectal polyp 3D reconstruction and analysis from real-world colonoscopy images and videos.

Chapter 3

Research Methodology

3.1 Methodology

3.1.1 Overview

This chapter outlines the methodology used in this research to enhance the detection and analysis of colonic polyps through advanced 3D reconstruction techniques. The aim is to transform standard 2D colonoscopy images into detailed 3D models, providing more accurate insights into polyp morphology for better diagnosis and treatment planning. Combining multiple advanced techniques, Paris classification-based dataset splitting to improve data organization, and a mask condition-based frame selection procedure employing contrast, ORB features, and depth map quality, the approach yields Using the EndoSRR technique, reflection reduction is done to solve issues including reflection in colonoscopy pictures; ZoeDepth is utilized to generate accurate depth maps. Moreover, the work focuses on generating 3D point clouds and meshes using cutting-edge techniques such Poisson surface reconstruction and Ball Pivoting Algorithm (BPA), which enable the geometric feature extraction including slope and curvature. Improving diagnosis accuracy depends critically on these characteristics. This work examines the curvature and slope of the polyp surface through geodesic path calculation, following validation with shadow and project analysis. From data preparation to model training and evaluation, every stage of the process, including data preprocessing, is integrated into a cohesive pipeline designed to exceed the capabilities of traditional 2D imaging.This approach not only offers a greater visual awareness of polyps but also makes these approaches accessible and scalable in clinical environments, so enabling more efficient and effective colorectal cancer screening.

3.1.2 Dataset Collection

High-quality colonoscopy images from the PICCOLLO dataset provide the dataset used in this work for the 3D reconstruction and analysis of colorectal polyps. Particularly designed for polyp detection and classification activities is this dataset. It comprises pictures taken with conventional colonoscopy techniques, which constitute the basis for 3D models from

2D photos. Polyp locations and their matching Paris classification are noted on the photos. Training, validation, and test sets separate the dataset to guarantee that the model may be trained and evaluated successfully. Corresponding masks complement the images in the collection and are utilized for segmentation activities during reconstruction. Correct polyp recognition and noise reduction from the photos depend on these masks. Development of the pipeline spanning frame selection, depth estimation, 3D point cloud creation, and mesh reconstruction depends on this dataset. It enables a strong assessment of the suggested approaches in this work, therefore giving the 3D reconstruction process clinical relevance.

Dataset Overview

Specifically selected for polyp detection and 3D reconstruction tasks, this study employs 3,433 pictures acquired from clinical colonoscopy procedures. Comprising both narrow band imaging (NBI) images and white light, the PICCOLO Widefield collection contains this dataset. These images were taken during colonoscopy exams on 48 patients including 76 distinct lesions overall. Clinical evaluation guides each lesion into one of three categories: Class 0 (hyperplastic polyps), Class 1 (adenomatous polyps), Class 2 (sessile serrated lesions). These class designations are applied all through model training and evaluation and show clinically significant differences. Additionally included in the dataset are clinical metadata linked to every lesion, which supports supervised learning methods and provides necessary background for diagnosis interpretation.Three independent subsets training, testing, and validation have the images set to ensure a strong, objective evaluation of the proposed approach. Segmentation activities determine accurate polyp localization and the production of 3D reconstructions; so, every image is matched with a corresponding binary mask to assist with them.

Table 3.1: Dataset Distribution

Dataset Split	**Images**	**Masks**	**Lesions**	**Patients**
Train	2,203	2,203	76	48
Test	333	333	-	-
Validation	897	897	-	-
Total	3,433	3,433	-	-

Sample Representation

Three primary categories define the dataset used in this work: White Light Images, Narrow Band Imaging (NBI) Light Images, and Void images. These groups offer several and complimentary perspectives of the colon. The White Light Images help to provide a natural, broad perspective of the colon's surface, therefore facilitating basic tissue evaluation. The Narrow Band Imaging (NBI) Light Images improve mucosal architecture and vascular patterns, therefore increasing the visibility of tiny features vital for a correct diagnosis.

At last, the void images aid to highlight areas without relevant features, therefore enabling the differentiation between various kinds of tissue and lesions. Images taken during clinical colonoscopy operations from the PICCOLO dataset abound in every one of these categories. This visual diversity is vital for creating a comprehensive model that can analyze and detect colonic abnormalities, thereby enhancing the diagnostic process in medical imaging, as illustrated in 3.1.

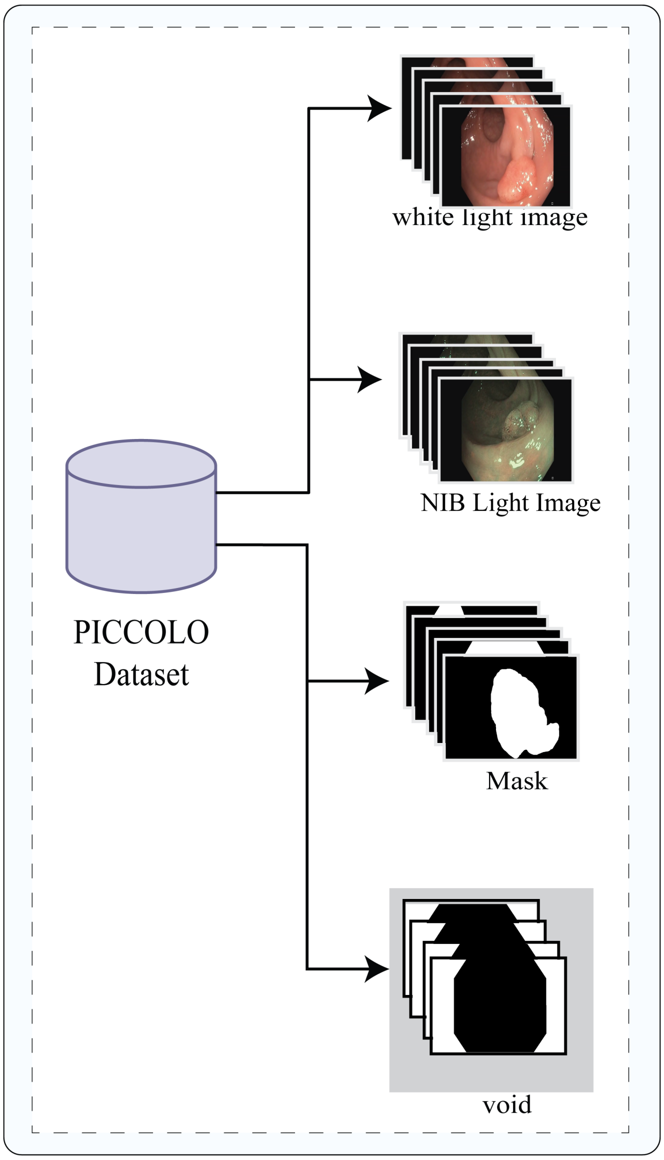


Figure 3.1: Sample images from the PICCOLO dataset, including white light, NBI light images, and void areas

3.1.3 Proposed Methodology

This section presents the methodology adopted for 3D reconstruction and classification of colorectal polyps using a fully automated and clinically oriented pipeline.

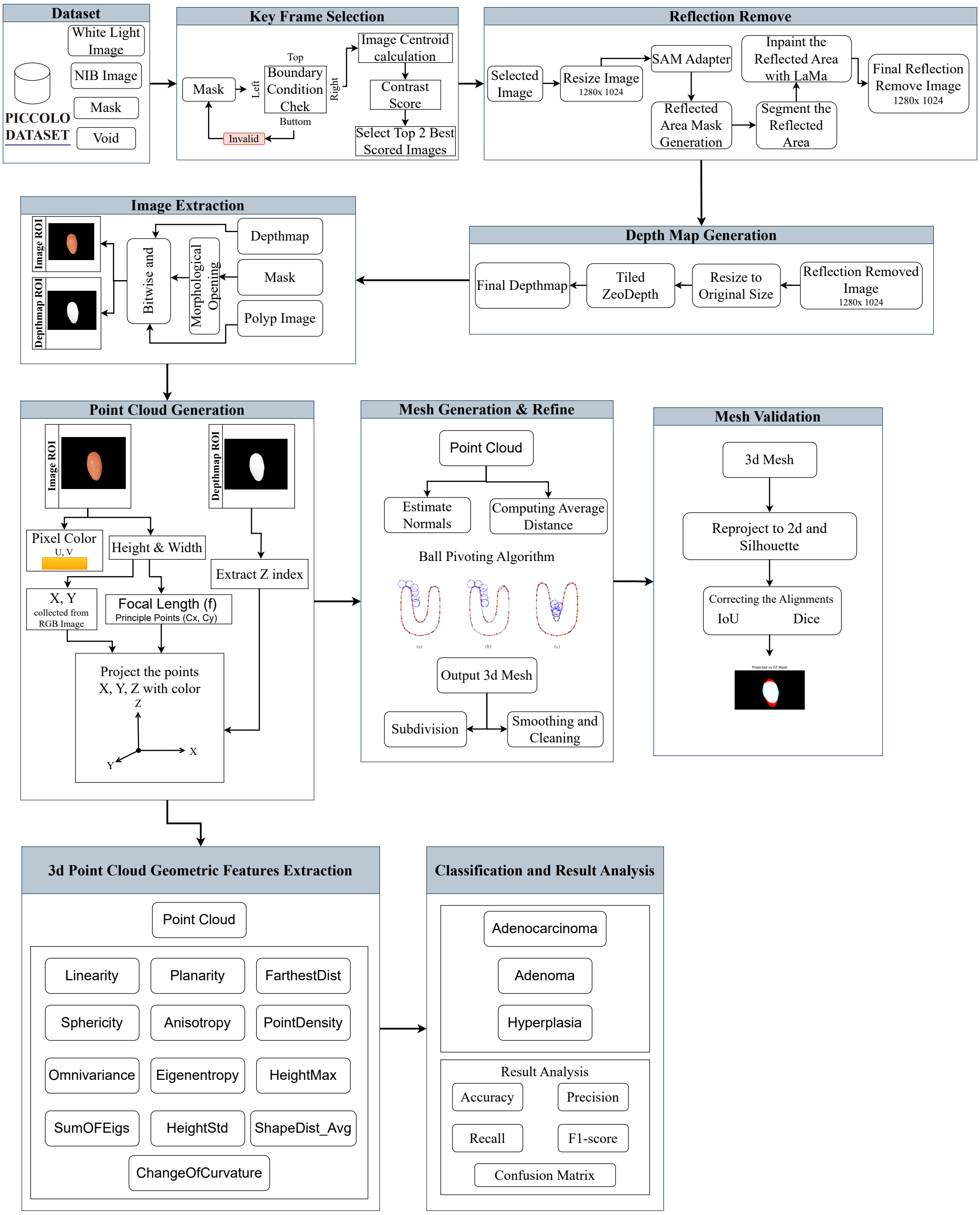


Figure 3.2: Proposed methodology for 3D reconstruction and classification of colorectal polyps based on 3D model features

The proposed system is designed to transform 2D colonoscopy video frames into 3D mesh representations and extract meaningful geometric features for classification. To ensure clinical relevance, we focused on enhancing diagnostic accuracy while maintaining real-time applicability. Figure 3.2 shows the architecture of the proposed methodology.

3.1.3.1 Frame Selection

In our pipeline, each lesion's sequence of 10–20 frames are initially grouped, and any frame with a binary polyp mask that intersects the image borders is deleted to ensure near complete segmentation. From the remaining frames, we calculate a normalized mask centrality score, which is based on the distance of the mask centroid from the image center, and analyze each frame's global quality through RMS contrast in its grayscale intensities. The two metrics are min-max normalized and equally weighted to generate a composite score for each candidate frame. Finally, the frames are ranked based on this composite score, with the top two from each sequence selected for subsequent 3D reconstruction, as shown in Figure 3.3. This approach aligns with representative frame extraction in photogrammetric workflows, automatic frame selection techniques for regularized 3D endoscopic mapping, and key-frame selection in endoscopic video analysis.

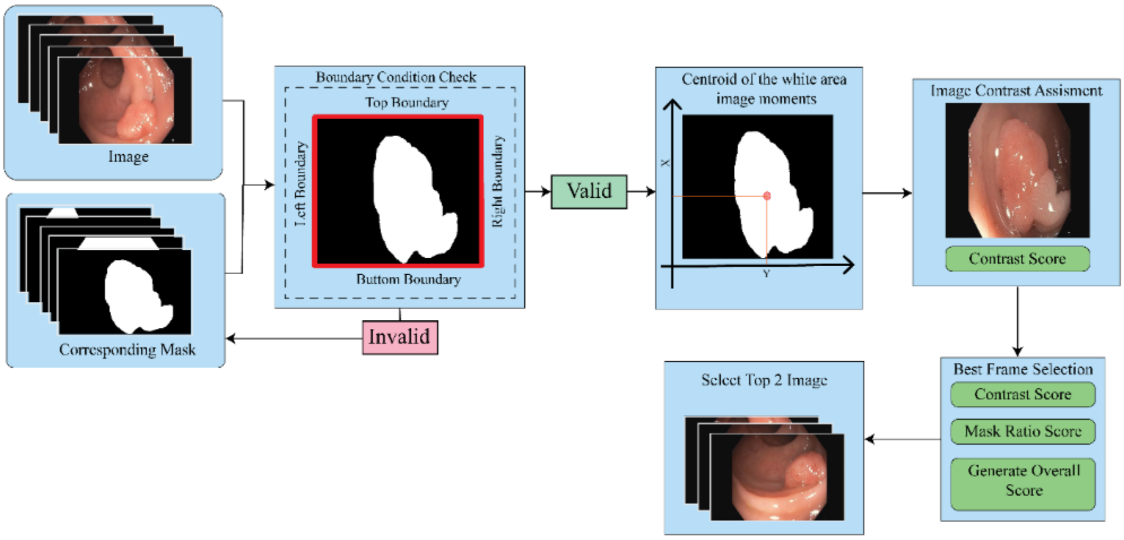


Figure 3.3: Frame selection based on border exclusion, mask centrality, and RMS contrast for 3D polyp reconstruction

3.1.3.2 Specular Reflection Removal

Specular reflections in endoscopic images, often caused by moist and highly reflective biological tissues, create bright artifacts that obscure surface details and limit computer vision algorithms. To address this, the EndoSRR algorithm is implemented for systematic reflection removal. EndoSRR uses a two-stage pipeline: first, it fine-tunes a Segment Anything Model (SAM) to generate precise reflection masks, and then applies a LaMa-based Fourier convolution inpainting strategy [25] to restore uniform illumination and continuity of underlying textures [26]. By removing high-intensity highlights during preprocessing, the algorithm ensures that subsequent stages of depth estimation and feature extraction operate on reliable surface information, significantly enhancing the accuracy of geometric interpretations in 3D reconstruction. If specular reflections are not removed, they can result in anomalous outliers or distortions in 3D point clouds and mesh models, causing spurious vertices and topological errors. Therefore, their removal is crucial for achieving accurate morphological representations. Integrating EndoSRR into endoscopic imaging workflows, as depicted in Figure 3.4, considerably improves the accuracy of subsequent 3D modeling and quantitative anatomical analyses.

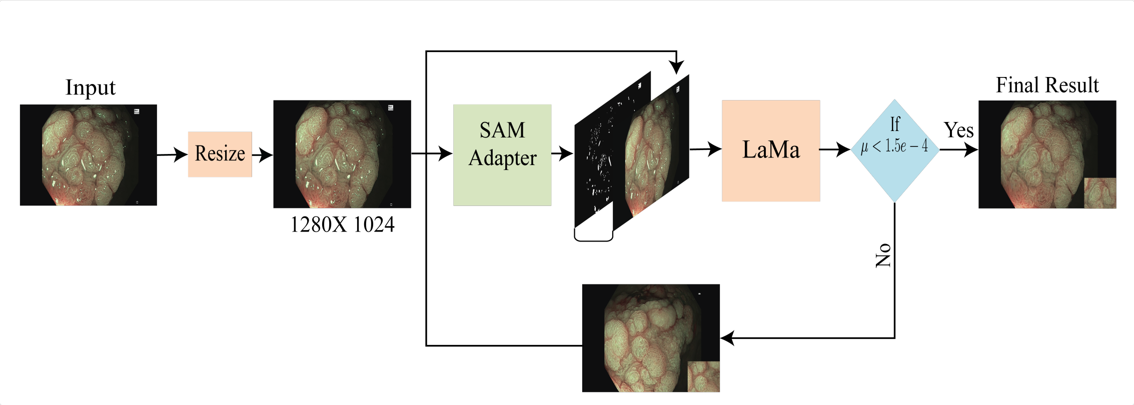


Figure 3.4: Reflection removal using Endo-SRR with SAM segmentation and LaMa inpainting for improved 3D reconstruction

3.1.3.3 Depth Map Generation

We employ the Tiled ZoeDepth algorithm [27] to improve the resolution and precision of depth maps in endoscopic imaging, using a tiling method for the efficient processing of high-resolution images. This method involves dividing the original image into overlapping tiles, generating depth maps for each tile, and then reconstructing these into a complete depth map through the application of gradient masks and average weighting based on the initial depth map. A low-resolution depth map for the entire image is initially generated using the ZoeDepth model [28]. The image is subsequently divided into overlapping segments, and depth maps are generated for each individual segment. Minimization of edge artifacts and assurance of tile transitions are achieved through the use of gradient-based filters. Utilizing the gradient information from surrounding regions, these filters aim to reduce discontinuities at the edges of neighboring tiles.

The reassembly process involves estimating average depth values from the overlapping regions of adjacent tiles and applying weights derived from gradient information to select the more reliable depth estimates. This method effectively reduces artifacts and enhances the overall quality of the depth map.

The individual depth maps are reassembled into a comprehensive depth map using gradient-based masks and averaging techniques, enabling seamless transitions between tiles. This method retains intricate details and ensures spatial consistency throughout the image, which is essential for accurate 3D reconstruction of polyp areas. The full process is illustrated in Figure 3.5.

The final depth value at each pixel (x, y) is computed using a weighted average over all overlapping tiles, defined as:

$$D_{\text{final}}(x, y) = \sum_{i,j} \left(\frac{W_{i,j}(x, y) \cdot D_{i,j}(x, y)}{\sum_{i,j} W_{i,j}(x, y)} \right)$$

where $D_{\text{final}}(x, y)$ is the final depth value at pixel (x, y), $D_{i,j}(x, y)$ is the depth value from the (i, j)-th tile, and $W_{i,j}(x, y)$ is the weight assigned to that tile's depth value based on its proximity to the pixel. This weighted averaging ensures that depth information from all tiles contributes appropriately to the final depth map, enhancing the overall accuracy of the depth estimation.

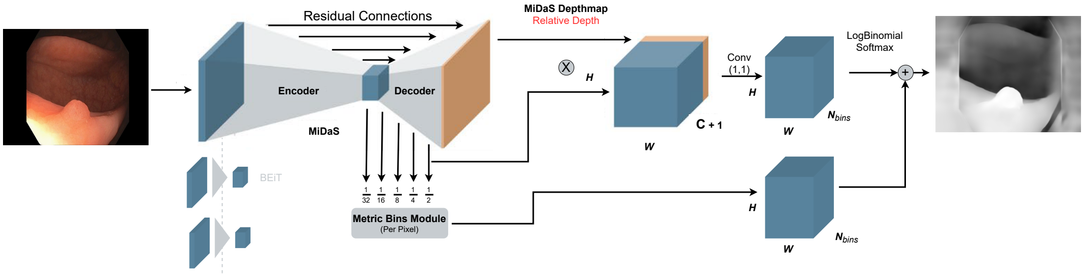


Figure 3.5: Depth map generation using ZoeDepth with tiled inference for high-resolution monocular 3D reconstruction

3.3.1.4 Region of Interest (ROI) Extraction

The Region of Interest (ROI) extraction method under development is meant to separate the polyp areas with great spatial accuracy so that only pertinent anatomical features are applied in further 3D reconstruction projects. This stage starts with using the binary segmentation masks produced from the images that have been eliminated for reflections.

A morphological opening operation on the mask helps eliminate minor artifacts and noise while maintaining the polyp region's form, hence improving the ROI. The improved mask then bitwise ANDs with the matching RGB picture, so isolating the polyp and removing pointless background areas. This targeted approach ensures that only the essential features of the polyp are retained for depth estimation, point cloud generation, and mesh reconstruction, thereby improving both computational efficiency and reconstruction accuracy. The overall ROI extraction procedure is illustrated in Figure 3.6.

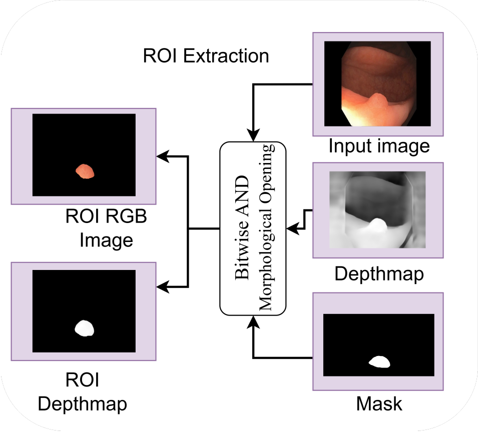


Figure 3.6: ROI extraction using polyp and void masks

3.3.1.5 Point Cloud Generation

In order to create a 3D point cloud from depth and RGB images, we apply the camera's intrinsic parameters, particularly the focal length and principal point, to project the depth value of each pixel into 3D space.The first stage is the construction of a camera matrix K reflecting these inherent properties [29]. Each pixel in the depth map has its 3D coordinates computed by back-projecting its position using the inverse of K and scaling in line with the depth value. The output offers the 3D coordinates inside the camera frame. One uses a rotation matrix R and a translation vector t to align the point cloud with the world coordinate system. The 3D points then follow an adjustment into the global coordinate frame. These three-dimensional points are commonly connected with RGB values obtained from the appropriate color image to enable simpler presentation and later analysis. This produces a colored point cloud. This method is extensively applied for the reconstruction of spatial environments utilizing depth and color information in many applications in computer vision and robotics. The overall point cloud generation workflow is illustrated in Figure 3.7.

$$f = \frac{0.5 \times W}{\tan\left(0.5 \times FOV \times \frac{\pi}{180}\right)}$$

$$K = \begin{bmatrix} f & 0 & 0.5 \times W \\ 0 & f & 0.5 \times H \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = K^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \times d(u, v)$$

Each 3D point is then assigned the RGB color from the corresponding pixel in the color image to create a visually instructive representation, hence producing a dense, colored point cloud that faithfully records both geometric and photometric data. From environment re-

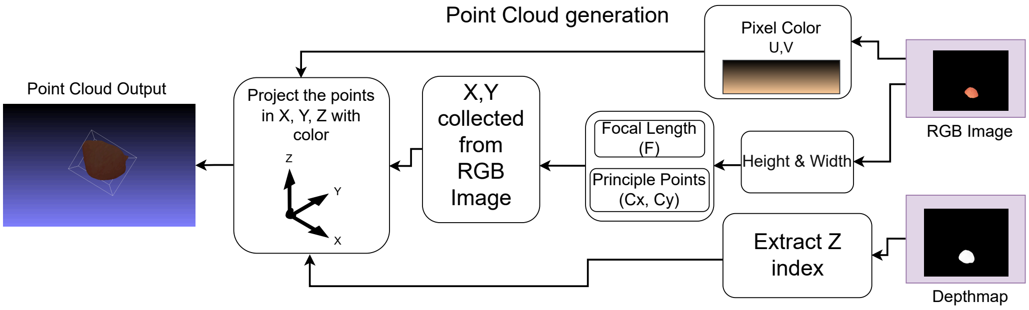


Figure 3.7: Point cloud generation from depth maps using camera intrinsics for 3D reconstruction

construction to object detection and navigation, this simple projection-and-color-mapping process underpins a broad spectrum of uses in computer vision and robotics.

3.3.1.6 Mesh Generation

A fundamental operation in 3D computer vision and graphics, mesh reconstruction from point clouds turns individual 3D points into continuous surface representations. Applications ranging from virtual reality to robotics to medical imaging depend on this metamorphosis. Many algorithms have been created to handle this work; each has special benefits and method of approach.

One of the foundational methods is the Ball Pivoting Algorithm (BPA) [30], which simulates a ball rolling over the point cloud to form triangles, effectively creating a mesh. Another prominent technique is Poisson Surface Reconstruction, which solves a Poisson equation to generate a smooth surface that best fits the input points. Additionally, the Alpha Shapes method constructs a surface by creating a family of shapes parameterized by a radius α, capturing the underlying geometry of the point cloud. The mesh reconstruction process, including the BPA method applied in this work, is illustrated in Figure 3.8.
Our implementation utilizes the Ball Pivoting Algorithm (BPA) for polyp modeling, applying its ability to accurately trace the intrinsic geometry of point clouds while avoiding the creation of spurious boundary artifacts. BPA generates a virtual sphere of a specified radius that traverses adjacent triplets of points, pivoting around edges to create triangles only in areas with sufficient local point density, resulting in a mesh that accurately reflects the underlying polyp topology. In contrast, our experiments with Poisson reconstruction under Neumann boundary conditions generated extraneous surface patches around the mesh, complicating the isolation of the polyp surface. By implementing BPA, we eliminate these unwanted extensions and create a clean, anatomically accurate 3D mesh.

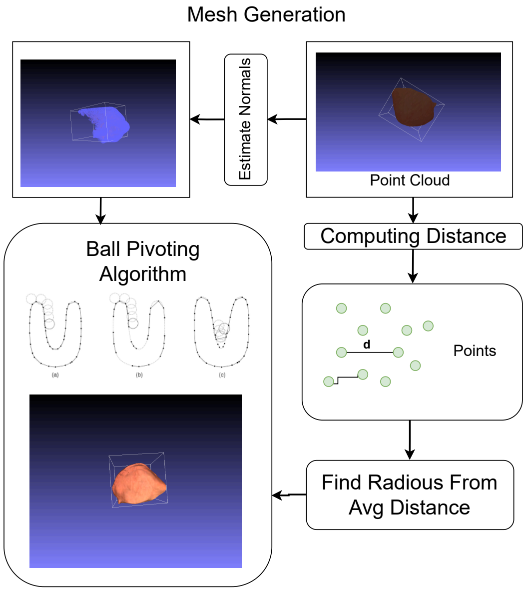


Figure 3.8: Surface mesh generation from point clouds using the Ball Pivoting Algorithm (BPA)

3.3.1.7 Validation through Silhouette and Projection

To ensure the accuracy of the reconstructed 3D mesh and its consistency with the original image and segmentation mask, the pipeline includes a validation step based on silhouette and projection analysis. This process is designed to verify the geometric fidelity of the reconstruction and is carried out after mesh generation. In this step, the reconstructed 3D point cloud is first reprojected onto the 2D image plane using the camera intrinsic matrix. This projection simulates the camera's view of the reconstructed surface. The silhouette of the projected points is then computed and compared directly with the binary polyp mask originally used during the frame selection phase.

The comparison is visualized by overlaying the reprojected 2D silhouette on top of the ground truth segmentation mask. A high degree of alignment indicates that the 3D reconstruction has maintained accurate spatial correspondence with the original polyp region. Minor mismatches help identify local inconsistencies in depth estimation or geometric artifacts introduced during meshing. This validation technique leverages both geometric re-projection and binary mask matching to provide visual confirmation of mesh accuracy. It is particularly effective in confirming that the 3D shape aligns with the clinical region of interest and ensures that further feature extraction from the mesh is performed on spatially reliable structures.

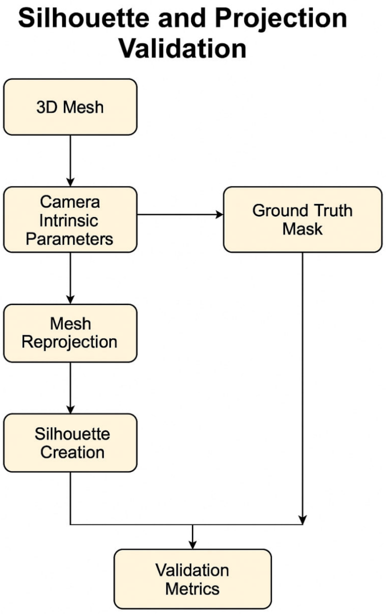


Figure 3.9: Mesh Validation

Proposed Algorithmic Framework

The proposed algorithmic framework aims to reconstruct the three-dimensional structure of polyp regions from colonoscopy video sequences. It integrates traditional computer vision methods with state-of-the-art deep learning architectures—including transformer-based and generative models—to achieve high spatial accuracy, texture preservation, and computational efficiency.

1. Frame Scoring and Selection Algorithm

To ensure only high-quality frames are used for 3D reconstruction, a scoring mechanism is applied on a per-sequence basis. Three key metrics are computed for each frame:

- Mask Centrality Score: Measures how centrally the polyp mask is positioned within the image.
- RMS Contrast: Evaluates the global contrast to identify visually rich frames.
- ORB Feature Count: Quantifies texture detail by detecting ORB keypoints.

Each metric is normalized to the range $[0, 1]$, and a weighted average is used to compute the final frame score. The two top-scoring frames per lesion are selected to ensure both diversity and quality.

2. Reflection Removal with EndoSRR

To handle specular reflections that impair segmentation and depth accuracy, the EndoSRR pipeline is employed. It consists of:

- SAM-ViT (Segment Anything Model - Vision Transformer): Performs zero-shot segmentation of polyp regions.
- big-LaMa (Large Mask Inpainting): Removes reflections using generative inpainting for realistic restoration.

This two-stage process significantly improves input quality and mask accuracy.

3. ROI Extraction with Morphological Refinement

Following segmentation, morphological opening removes small artifacts. The refined binary mask is applied via a bitwise AND operation with the RGB image to isolate the Region of Interest (ROI). This ensures that depth estimation targets only the polyp, reducing noise and computational cost.

4. Depth Estimation using ZoeDepth V3 (Tiled Inference)

A modified ZoeDepth V3 model is used with tiled inference to handle high-resolution inputs. The process includes:

- Dividing the ROI into overlapping tiles.
- Predicting depth for each tile.
- Merging tiles using gradient-based filtering and averaging to reduce edge artifacts.

ZoeDepth's transformer architecture provides reliable depth gradients, even in texture-sparse endoscopic imagery.

5. Point Cloud Construction

3D coordinates are computed for each pixel using inverse projection with the camera intrinsics and the depth map. Each point is colored with its corresponding RGB value, producing a dense and textured point cloud.

6. Mesh Generation using Ball Pivoting Algorithm (BPA)

To convert the point cloud into a surface mesh, the Ball Pivoting Algorithm (BPA) is applied. It forms triangles by pivoting a virtual ball around triplets of nearby points. Unlike Poisson reconstruction, BPA maintains local surface integrity and avoids overfitting to sparse regions. This leads to anatomically coherent meshes, ideal for preserving polyp topology.

3.1.4 Workflow

The proposed methodology integrates a robust frame selection mechanism, a deep-learning-based reflection removal process, depth estimation, and geometric mesh generation to produce accurate 3D polyp reconstructions from colonoscopy videos. The workflow is designed to address common challenges in endoscopic reconstruction such as poor lighting, low texture, specular highlights, and anatomical deformation, ultimately enabling high-fidelity 3D modeling of polyps for clinical analysis. The entire process is modular and optimized for both visual quality and downstream analytical usability.

The workflow begins with frame grouping and cleaning, where each video sequence containing a polyp is divided into 10 to 20 individual frames. These frames are first filtered based on binary polyp segmentation masks. Any frame where the mask touches the image borders is eliminated. This step ensures that the selected frames capture polyps that are fully visible and centered, avoiding distortions in later stages of reconstruction.

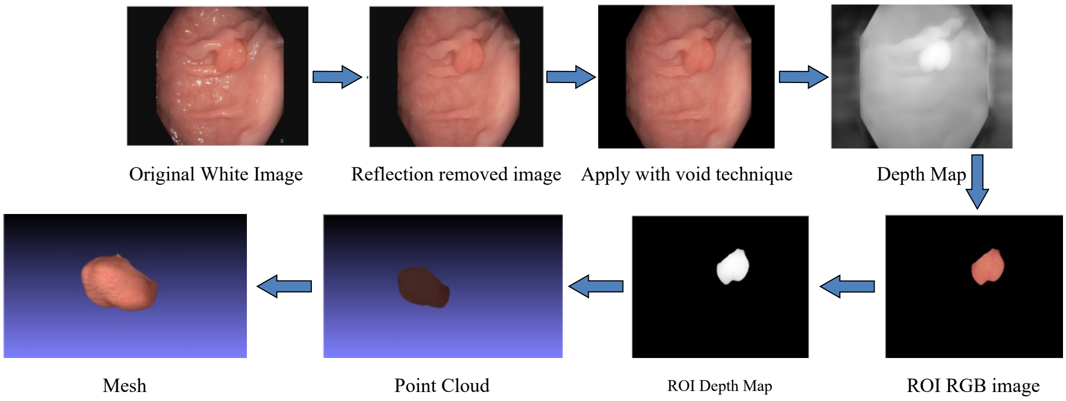


Figure 3.10: Workflow of the proposed 3D reconstruction and classification pipeline for colorectal polyps

Next, a scoring system is applied to all remaining frames to identify the most representative samples for 3D modeling. Each frame is evaluated using three distinct metrics: mask centrality, RMS contrast, and ORB feature count. The mask centrality measures how close the polyp lies to the image center; RMS contrast captures the overall visual quality of the frame; and ORB feature count quantifies the amount of texture and keypoint information in the image. These min-max normalized values are aggregated with equal weighting to get a composite score at last. This score guides the selection of the top two frames from every video sequence, therefore guaranteeing the best possible frame quality and anatomical relevance.

The picked frames go through reflection removal with the EndoSRR process following the phase of frame choice. This two-stage technique comprises inpainting and segmentation. First high-quality, zero-shot polyp masks are extracted using the Segment Anything Model (SAM-ViT). These masks then are sent to the big-LaMa model, which removes specular highlights and overexposed areas from the photos using deep learning-based inpainting. The end effect is a visually consistent, artifact-free polyp image more fit for depth measurement.

Removing reflections isolates the Region of Interest (ROI) for depth prediction. Opening and dilating the polyp mask helps to eliminate noise by means of morphological procedures.

To get a clean ROI, this mask is then bitwise AND applied to the inpainted image. This stage reduces computing work and stops meaningless background areas from affecting the depth estimation output.

The Zoe Depth V3 depth estimating model then runs over the cleaned ROI. We use a tiling technique to manage GPU memory limits and high-resolution input. Each of the overlapping tiles in the ROI image receives unique application using ZoeDepth. To create a seamless and realistic full-image depth map, the tiles are then rejoined back together with smoothing in the overlapping areas. Crucially for the accuracy of the 3D reconstruction, this depth map catches the polyp's fine-grained geometry.

After the depth map is created comes generation of point clouds. By means of depth values and camera intrinsic features, the pinhole camera model projects every pixel into 3D space. The resultant 3D coordinates blended with RGB values from the original image create dense colored point clouds. This point cloud fairly captures the polyp's surface morphology and texture.

Then the Ball Pivoting Algorithm (BPA) transforms the point cloud into a 3D surface mesh. BPA creates triangles whereby the ball contacts all three by rolling a simulated ball over triplets of points. This approach captures local curvature and helps especially to preserve surface continuity. In contrast to Poisson surface reconstruction, which may result in unnecessary smoothing or inaccurate shapes, BPA generates meshes that accurately represent the true surface contours of the polyp, avoiding the pitfalls of excessive smoothing or distortion associated with Poisson surface reconstruction. Ideal for clinical assessment, the resulting mesh exhibits anatomical precision and structural integrity.

Finally, we arrive at the stages of visualization and evaluation. With the application of 3D rendering technology, medical professionals or analysts have the capability to rotate, zoom, and examine the polyp within a fully immersive 3D environment, allowing for a comprehensive visualization of the reconstructed mesh. For further investigation one can additionally calculate geometric properties such surface variation, planarity, and curvature. Quantitative examination about Hausdorff Distance and mesh quality ratings ensures the accuracy and consistency of the reconstructions.

To meet the special difficulties in processing endoscopic pictures, our whole approach combines modern neural networks with conventional image analysis techniques. By use of geometric surface modeling, monocular depth computing, intelligent frame selection, and deep reflection reduction, our approach defines a new benchmark in 3D polyp reconstruction.It is scalable, privacy-preserving, versatile for both research-based and diagnostic applications.

3.2 Detailed Methodology and Design

From endoscopic video sequences, the proposed approach presents a complete and flexible pathway for 3D polyp reconstruction. This structure is especially designed to solve the frequent medical imaging problems including limited vision, reflecting surfaces, and different polyp appearances. The method is organized into precisely defined phases covering frame choice, elimination of reflections, region of interest extraction, depth estimation, point cloud generation, mesh reconstruction, and final evaluation. Every element of the pipeline combines modern computer vision and deep learning methods to guarantee scalability and resilience in several clinical environments. Frame preparation and filtering starts the process; each polyp video sequence is arranged into 10–20 separate frames. First passed by a binary segmentation mask filter derived from pretrained polyp segmentation networks, are these frames If the polyp mask crosses any of the image borders, frames are thrown away since these probably reflect partial or obstructed perspectives of the polyp. This stage guarantees that retained frames for additional processing only feature fully visible polyps. Then, from every sequence, the two most representative frames are found using a quality-based frame selection technique. Three complementary metrics direct the choice: RMS contrast, which measures global image sharpness and clarity; mask centrality, which gauges the polyp's central location within the frame; and the number of ORB (Oriented FAST and Rotated BRIEF) features, which records textural richness and detail. Min-max scaling helps each statistic to be standardized thereby guaranteeing equitable contribution. Equally weighting the three normalized indicators produces an aggregate score; the top two frames with the best composite scores are chosen for reconstruction. Minimizing redundancy and optimizing anatomical representation and data quality depend on this choosing phase.

The selected frames then undergo reflection removal using a two-stage deep learning framework based on the EndoSRR pipeline. Initially, the Segment Anything Model (SAM-ViT) is used to automatically generate precise object masks of the polyp, operating in a zero-shot manner without prior task-specific training. These masks are passed to the LaMa model, a large generative inpainting model that removes specular highlights and other bright artifacts by intelligently hallucinating plausible textures in corrupted regions. This step significantly improves visual consistency and prepares the image for accurate downstream analysis. To refine the area of interest for depth estimation, morphological operations such as opening and dilation are applied to the segmentation mask to eliminate noise and small artifacts. The cleaned mask is then used to extract the region of interest (ROI) from the reflection-removed image using a bitwise AND operation. This step ensures that only the relevant polyp area is passed to the depth estimator, improving the efficiency and accuracy of subsequent 3D reconstruction.

For depth estimation, the reflection-cleaned ROI is passed through the ZoeDepth V3 model, a state-of-the-art monocular depth estimation framework. To handle high-resolution inputs and GPU memory limitations, a tiling strategy is implemented, wherein the ROI is split into overlapping tiles. Each tile is individually processed by ZoeDepth, and the outputs are then stitched together using blending in overlapping regions to form a seamless depth map. This step results in a high-fidelity depth map that captures the nuanced topography of the polyp.

Dense 3D point clouds are produced by means of known camera intrinsic characteristics and depth map. The pinhole camera model projects every pixel in the depth map into 3D space along with matching RGB values from the original frame linked to each 3D coordinate. Acting as an intermediary representation for mesh construction, the output is a colored point cloud encoding both spatial and appearance information of the polyp surface.

The three-dimensional point cloud is subsequently produced using the Ball Pivoting Algorithm (BPA). This method generates triangles at every instance where the ball intersects three neighboring points by simulating a virtual ball traversing the point cloud. BPA preserves surface continuity and prevents the introduction of synthetic geometry, making it particularly suitable for smooth anatomical features such as polyps. Increased control over mesh grain and a reduced likelihood of over-smoothing render BPA less susceptible compared to Poisson-based reconstruction.

After the mesh is built, it is displayed using 3D rendering software that facilitates examination from various perspectives. Healthcare professionals and investigators can analyze the shape, curvature, and various geometric characteristics of polyps. Furthermore, features derived from mesh data, including planarity, omnivariance, and surface variation, can be calculated to enhance diagnostic insights or facilitate subsequent machine learning applications.

Chamfer Distance, Hausdorff Distance, and Mesh Structural Similarity are among the quantitative measurements used in evaluation together with qualitative visualization. These criteria enable one to evaluate the accuracy of the reconstruction toward actual anatomical forms. Moreover, timing and memory benchmarks are tracked to make sure the pipeline is useful for clinical settings with restricted processing capacity.

To produce a quite realistic and interpretable 3D reconstruction of polyps, the proposed method efficiently blends contemporary models with specialized approaches. The pipeline addresses major gastrointestinal imaging challenges by means of careful frame selection, advanced artifact removal using deep learning, accurate depth estimation, and simplified surface modeling, so opening the path for future uses in diagnosis support, surgical planning, and disease progression monitoring. Designed with flexibility in mind, the entire system allows plug-in replacement of components as well as simple interaction with current clinical or research infrastructure.

3.3 Summary

This chapter discussed the entire architecture and technique of a modular pipeline for three-dimensional colonic polyps from endoscopic video sequences. The method was meant with clinical usefulness in mind by comprising frame selection, reflection reduction, depth estimate, point cloud generation, and mesh reconstruction. Starting with segmentation-based frame filtering to reject incomplete views, the pipeline moves using a scoring system integrating mask centrality, RMS contrast, and ORB feature count to select the most representative images. First using LaMa for inpainting and SAM-ViT for object masking,

the EndoSRR pipeline removes reflections thereby guaranteeing artifact-free images for depth analysis.
By means of a tiled form of ZoeDepth V3, the depth estimation stage preserves spatial consistency under high-resolution inputs. The Ball Pivoting Algorithm then turns dense point clouds generated from this depth data into finely detailed surface models. BPA was selected since it kept geometric continuity better than standard Poisson reconstruction and more effectively managed complex anatomical elements. Moreover stressing excellent preprocessing, efficient management of computational resources, and precision in spatial representation helps to increase diagnosis reliability.

Considering all the factors, the technology offers 3D polyp modeling as a scalable and cohesive method improving exact morphological analysis and visualizing. The foundation for next chapters exploring the specifics of implementation, experimental validation, and the performance criteria of the suggested pipeline is laid here.

Chapter 4

Implementation and Results

4.1 Experimental Setup

These experiments were conducted using Google Colab and Kaggle Notebook with Python 3.12 running an NVIDIA Tesla T4 GPU backend. Both systems offer a cloud-based framework that fits very well with machine learning models such TensorFlow, Keras, and PyTorch. Data preparation, analysis, and visualization relied on basic libraries including Pandas, NumPy, Matplotlib, Seaborn, and Scikit-Learning. The systematically arranged specified system cleared the path for development in 3D reconstruction of a polyp cancer by enabling the efficient implementation, optimization, and validation of the proposed model.

4.2 Evaluation And Performance Analysis

4.2.1 Evaluation Metrics

To assess the performance of our classification models in polyp reconstruction and categorization, we utilized several standard evaluation metrics. These include precision, recall, F1-score, Intersection over Union (IoU), and the Dice Similarity Coefficient (DSC). Each of these metrics offers unique insights into the model's effectiveness in predicting the correct class labels, especially in the presence of class imbalance and overlapping structures.
A confusion matrix provides a summary of the prediction results on a classification problem. It displays the number of true and false predictions broken down by each class. The components of the confusion matrix are defined as follows:

- True Positive (TP): The model correctly predicts the positive class.
- True Negative (TN): The model correctly predicts the negative class.
- False Positive (FP): The model incorrectly predicts the positive class.
- False Negative (FN): The model incorrectly predicts the negative class.

These values form the basis for computing the core performance metrics:

- Recall: Recall, also known as sensitivity or true positive rate, measures the proportion of actual positives that are correctly identified by the model. It is computed as:
$$\text{Recall} = \frac{TP}{TP + FN}$$
- Precision: Precision quantifies the proportion of predicted positive instances that are truly positive. It is defined as:
$$\text{Precision} = \frac{TP}{TP + FP}$$
- F1-score: The F1-score is the harmonic mean of precision and recall. It balances the two metrics and is computed as:
$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$
- Intersection over Union (IoU): Also known as the Jaccard Index, IoU measures the overlap between the predicted and ground truth regions. Using set theory, it is defined as:
$$IoU(A, B) = \frac{|A \cap B|}{|A \cup B|}$$
where A is the set of predicted pixels and B is the set of ground truth pixels.
- Dice Similarity Coefficient (DSC): This metric measures the overlap between the predicted and ground truth regions, placing more emphasis on the area of agreement. It is calculated using the following formula:
$$\text{Dice}(A, B) = \frac{2 \cdot |A \cap B|}{|A| + |B|}$$
where A is the predicted set and B is the ground truth set.

4.2.2 Evaluation of model performance on 3D features

To characterize the 127 unique 3D point-cloud polyp models, we computed fourteen local geometric descriptors: Linearity, Planarity, Sphericity, Omnivariance, Anisotropy, Eigenentropy, SumOFEigs, Change of Curvature, Farthest Distance, Point Density, HeightMax, HeightStd, and ShapeDist_Avg. Each descriptor was normalized to a scale of 1.0 to ensure comparability. These features represent diverse structural properties of the point cloud, ranging from shape elongation and flatness to volumetric dispersion and local curvature variation.

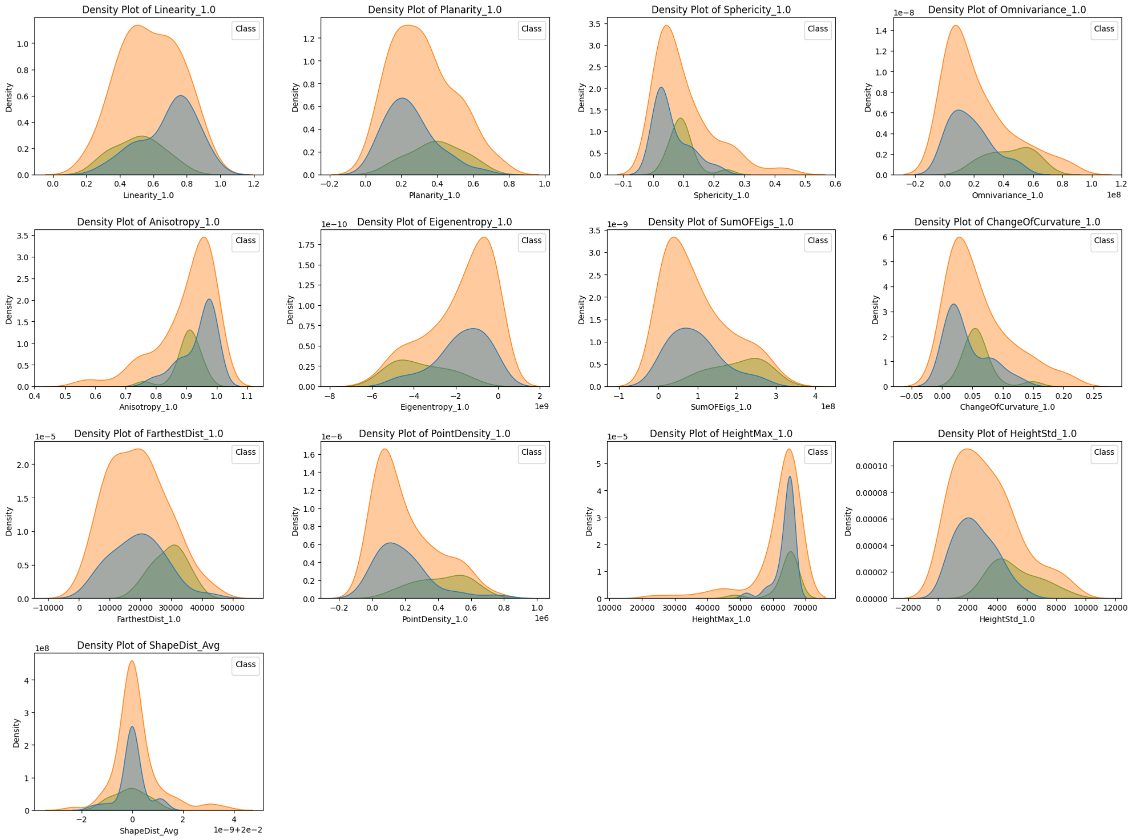


Figure 4.1: Point cloud density plot

To assess the distribution and potential discriminative power of these features, we generated kernel density plots for each descriptor across the dataset. The figure 4.1 highlights the density distributions of key point-cloud features, linearity, planarity, sphericity, omnivariance, anisotropy, eigenentropy, and the sum of eigenvalues, for three distinct polyp classes. Notably, features such as linearity and planarity display visible separation among Class 1, Class 2, and Class 3, suggesting possible class-specific patterns. However, overlapping distributions in several feature plots imply that accurate classification may remain challenging without further validation or feature selection, as feature space boundaries between classes are not always clearly defined.

Table 4.1: Accuracy of individual classifiers

Model	Accuracy (%)
XGBoost	80.77
LightGBM	76.92
MLP	76.92
KNN	73.08

For the initial evaluation, we applied a standard 80/20 train-test split without cross-validation. Classification results, including model accuracies, are presented in Table 4.1

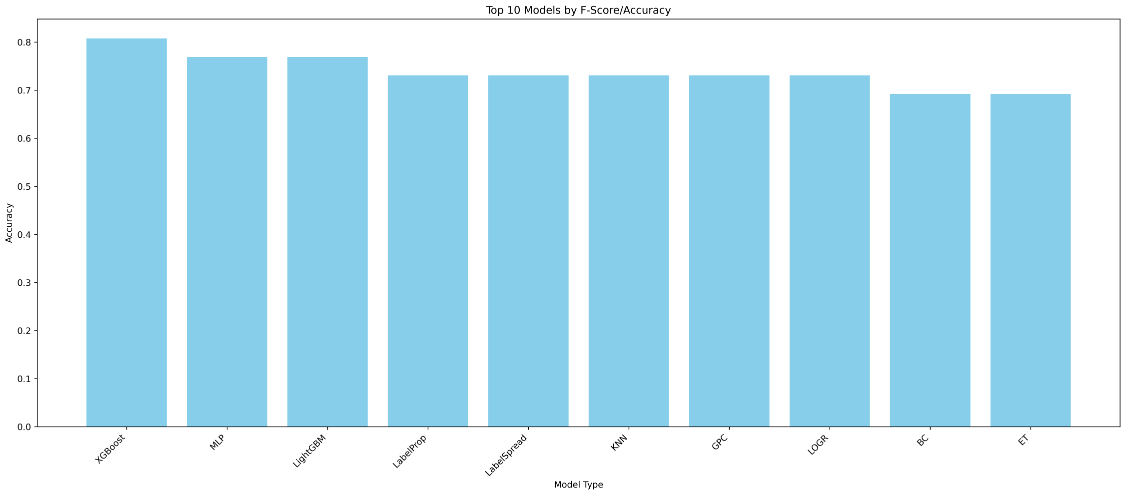


Figure 4.2: Accuracy for top 10 models

and visualized in the hold-out accuracy bar chart (Figure 4.2). Among the models tested, XGBoost achieved the highest test accuracy of 88.77%, followed by LightGBM and Multi-Layer Perceptron (MLP), each at 76.92%, and k-Nearest Neighbors (KNN) at 73.08%. In addition to the individual classifiers, we developed a hard-voting ensemble model that combines predictions from XGBoost, MLP, KNN, and Label Propagation. Using 5-fold cross-validation, this ensemble achieved fold-wise accuracies of 0.6154, 0.7308, 0.6000, 0.6000, and 0.7200. The ensemble's mean accuracy was 65.32% with a standard deviation of 5.93%. While this ensemble did not outperform the best individual model (XGBoost), its relatively stable performance across folds (Figure 4.3) indicates that combining diverse learners may help reduce overfitting, especially in small or imbalanced datasets.

4.3 Result And Discussion

In this section, we present a comprehensive analysis of the proposed 3D polyp reconstruction and classification pipeline. The evaluation encompasses frame selection, depth estimation, point cloud generation, mesh reconstruction, and the extraction of geometric features such as linearity, planarity, and curvature. These features were used to train and evaluate multiple machine learning classifiers to assess their effectiveness in polyp type classification. We further analyze the performance of each classifier using precision, recall, F1-score, and accuracy metrics. Additionally, we present the confusion matrices and evaluate model performance on a hold-out test set to validate the generalizability and robustness of our proposed approach.

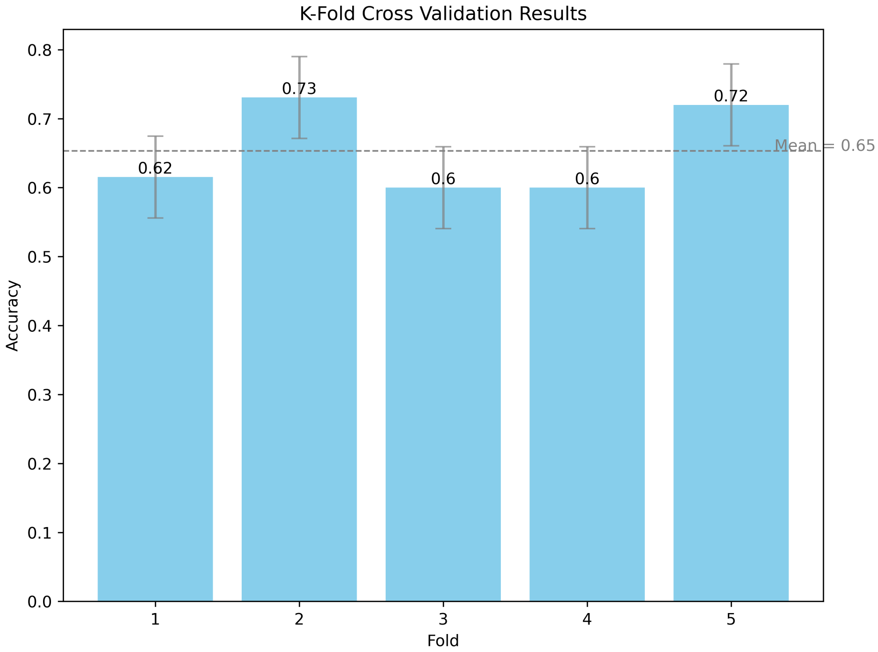


Figure 4.3: K-fold cross validation on 5-fold split.

4.3.1 Classification Results and Evaluation

This study implements a comprehensive pipeline for the 3D reconstruction of polyp images, leveraging point cloud generation and mesh reconstruction techniques. The goal is to accurately represent polyp structures in three dimensions using key geometric features. Features such as linearity, planarity, and curvature-based descriptors are extracted from the generated meshes and serve as input to machine learning classifiers for categorization. To evaluate the classification performance, four machine learning models—XGBoost, Multi-Layer Perceptron (MLP), LightGBM, and K-Nearest Neighbors (KNN)—were applied to the extracted 3D features. Their effectiveness was measured using precision, recall, F1-score, and accuracy across three polyp classes. The results on the hold-out test set (26 samples) are summarized in the following tables:

Table 4.2: Classification Performance Comparison Across Models

Metric	**XGBoost**	**MLP**	**LightGBM**	**KNN**
Accuracy	0.81	0.77	0.77	0.73
Macro Avg Precision	0.73	0.65	0.62	0.68
Macro Avg Recall	0.85	0.71	0.58	0.62
Macro Avg F1-Score	0.76	0.64	0.59	0.64
Weighted Avg Precision	0.86	0.82	0.76	0.79
Weighted Avg Recall	0.81	0.77	0.77	0.73
Weighted Avg F1-Score	0.82	0.78	0.76	0.75

The table 4.3 below presents a breakdown of classification metrics (accuracy, precision, recall, and F1-score) for each class across four different machine learning models (XGBoost,

Table 4.3: Class-wise Evaluation Metrics for Constituent Models

Model	**Class**	**Accuracy (%)**	**Precision**	**Recall**	**F1-Score**	**Support**
XGBoost	Adenocarcinoma	81.00	0.50	1.00	0.67	3
	Adenoma	81.00	0.94	0.79	0.86	19
	Hyperplasia	81.00	0.75	0.75	0.75	4
MLP	Adenocarcinoma	77.00	0.50	0.33	0.40	3
	Adenoma	77.00	0.94	0.79	0.86	19
	Hyperplasia	77.00	0.50	1.00	0.67	4
LightGBM	Adenocarcinoma	77.00	0.50	0.33	0.40	3
	Adenoma	77.00	0.85	0.89	0.87	19
	Hyperplasia	77.00	0.50	0.50	0.50	4
KNN	Adenocarcinoma	73.00	0.20	0.33	0.25	3
	Adenoma	73.00	0.83	0.79	0.81	19
	Hyperplasia	73.00	1.00	0.75	0.86	4

MLP, LightGBM, and KNN). This per-class evaluation offers insight into how each model performs across classes with different levels of support, enabling identification of strengths and weaknesses in the classification of underrepresented classes (e.g., Class 0 and Class 2).

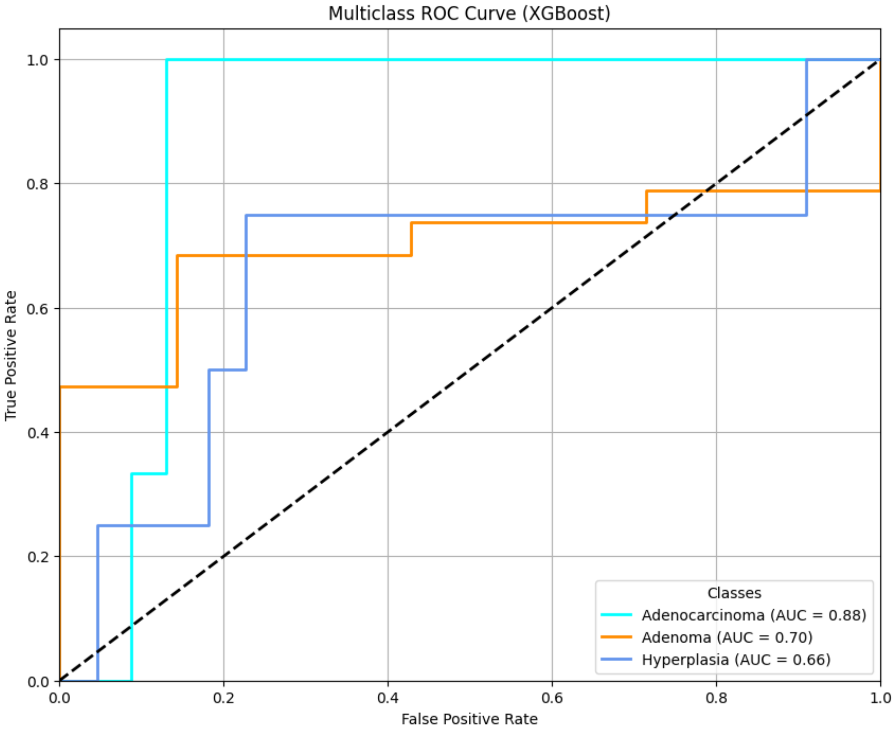


Figure 4.4: Confusion matrix for XGBoost classifier

Figure 4.4 presents the multiclass ROC (Receiver Operating Characteristic) curve for the XGBoost classifier, illustrating its performance across three classes. The ROC curve plots the true positive rate (sensitivity) against the false positive rate, allowing for a threshold-independent evaluation of classifier effectiveness. The dashed diagonal line represents random guessing.

For Class 0, the area under the curve (AUC) reaches 0.88, indicating a high level of discriminative capability and suggesting that the classifier reliably distinguishes this class from the others. In contrast, Class 1 and Class 2 achieve lower AUC scores of 0.70 and 0.66, respectively. These values indicate moderate to low performance in distinguishing those classes, likely due to feature overlap or class imbalance within the dataset.

The disparity in AUC values across the three classes suggests that while XGBoost performs well for at least one class (Class 0), its ability to generalize across all categories is limited. This observation is consistent with prior findings regarding class overlap in the point cloud feature space. Further improvement may be achieved by refining feature selection, augmenting the training data for underperforming classes, or applying class-weighted training strategies.

4.4 Confusion Matrix Analysis

The confusion matrix offers a granular view of the performance of each classifier by comparing predicted class labels against the actual class distributions. It serves as a crucial diagnostic tool to evaluate how effectively the models distinguish between the three polyp categories used in this study — Hyperplastic (Class 0), Adenomatous (Class 1), and Serrated (Class 2). Understanding the patterns of true positives, false positives, and false negatives is essential to assess the reliability and robustness of the classification stage, particularly in a sensitive application such as polyp type prediction based on 3D geometry.

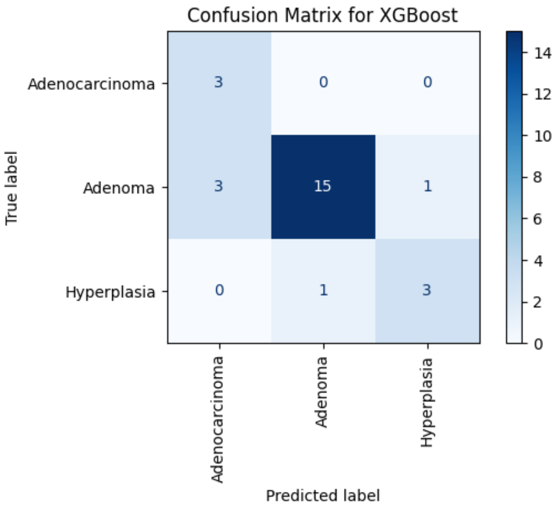


Figure 4.5: Confusion matrix for XGBoost classifier

XGBoost's classification accuracy was supported by minimal false positives and false negatives, reflecting its strong discriminatory capabilities. This is particularly important for clinical applications, where missing rare but potentially significant polyps can lead to diagnostic errors. The confusion matrix, illustrated in Figure 4.5, confirms its reliability by showing near-perfect predictions in the most challenging class distributions. ROC curve analysis further validates its superior performance, with an AUC close to 1, reinforcing its suitability for real-world deployment.

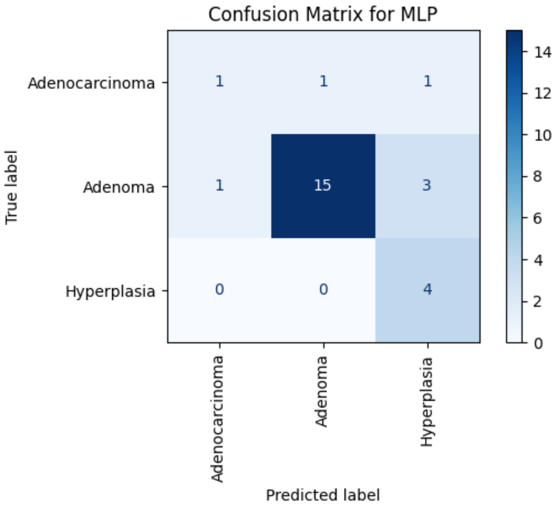


Figure 4.6: Confusion matrix for MLP

The MLP classifier also delivered competitive performance but showed noticeable weaknesses in handling the minority group, leading to reduced recall. Nevertheless, it performed well in distinguishing the more distinct class boundaries, suggesting sensitivity to class representation and a need for further optimization to generalize across all categories. Its confusion matrix (Figure 4.6) reflects these tendencies with a skewed distribution of misclassification.

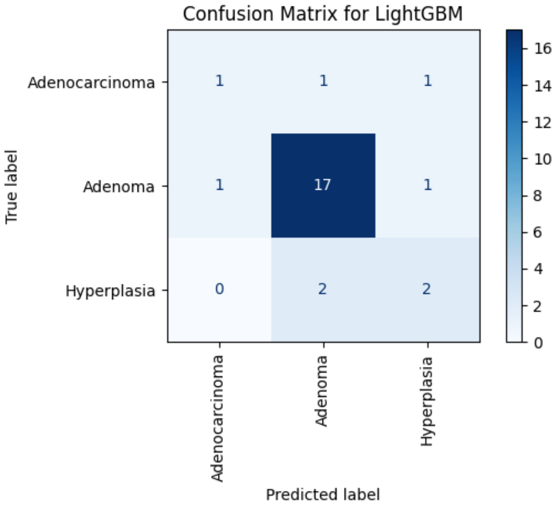


Figure 4.7: Confusion matrix for LightGBM classifier

LightGBM produced results similar to MLP, with overall stable classification for the majority but moderate confusion in minority categories. The model demonstrated better consistency than MLP in certain predictions, though its confusion matrix (Figure 4.7) revealed some misclassifications likely due to overlapping features between geometrically similar polyp types. These findings underline the importance of carefully engineered features when working with subtle anatomical differences.

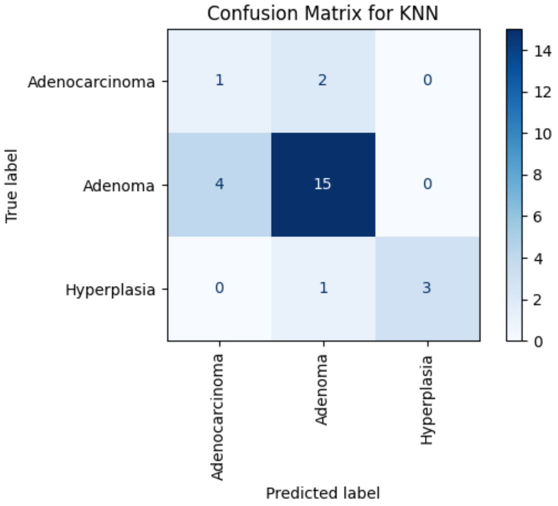


Figure 4.8: Confusion matrix for KNN

KNN, when evaluated under the same conditions, showed the weakest performance among the models. The confusion matrix (Figure 4.8) reveals multiple misclassifications, particularly with samples from the first two classes, indicating limited capacity in capturing the nuanced geometry encoded in the feature space. While it performed correctly in a few instances, its overall accuracy and recall were significantly lower, suggesting that KNN may not scale well in this context due to its sensitivity to local neighborhood noise and class imbalance.

4.4.1 Discussion

The outcomes of this work show how well the machine learning classifiers (XG-Boost, MLP, LightGBM, and KNN) classify polyps depending on 3D geometry and point cloud parameters. XGBoost shown to be the most dependable and accurate classifier among the studied models, showing better performance over all kinds of polyp. The classifier performed especially well in identifying the minority class, which usually presents more difficulty because of its poor frequency in the dataset. XGBoost proved strong even in the presence of class imbalance by attaining great precision and recall in all categories.
The XGBoost classifier's performance was particularly impressive for the minority polyp type, where it demonstrated the ability to effectively capture the subtle geometric features crucial for accurate classification. Its low false positive and false negative rates highlight even more its possible use in the real world since misclassifications may result in important diagnostic mistakes. XGBoost therefore makes a good model for application in healthcare environments where early and dependable polyp detection is essential.

In comparison, the MLP classifier showed comparable overall accuracy but faced challenges with classifying the minority polyp type. Despite these difficulties, MLP performed well for the other types, demonstrating its capability to learn and distinguish their distinctive geometric features. This suggests that MLP may benefit from further optimization, particularly in addressing class imbalance by better utilizing the feature patterns of underrepresented classes.

LightGBM also faced challenges with the minority class but performed better than MLP in correctly identifying the majority polyp type. Its confusion matrix analysis revealed some misclassifications between the minority class and the majority class, indicating overlapping feature characteristics between certain polyp types. This highlights the importance of feature selection and the need for refinement in extracting the most discriminative features to improve classification accuracy.

The KNN classifier, while simpler, also delivered decent results, with its performance being relatively consistent across the polyp categories. KNN struggled slightly with the minority class but still managed to produce accurate predictions for the majority class. Its performance suggests that although KNN may not be as powerful as tree-based models like XGBoost, it remains a useful model when paired with effective feature extraction methods.

The confusion matrix analysis for each classifier, particularly XGBoost, provides valuable insights into the model's performance. It highlights the ability of XGBoost to accurately classify rare polyp types, which is crucial in clinical diagnostics where timely and precise detection can significantly impact patient outcomes. The ROC curve analysis further confirmed the strong discriminatory power of XGBoost, with its area under the curve (AUC) indicating a high level of performance across different operating thresholds.

While XGBoost outperformed the other models in terms of classification accuracy, all models demonstrated a strong ability to classify the majority polyp type with high precision. This is expected, as the majority class was better represented in the dataset. However, the true measure of a model's robustness lies in its ability to handle minority classes effectively, where misclassifications could lead to missed diagnoses or delayed treatments. In this respect, XGBoost clearly stood out, with minimal false negatives and high recall for the minority polyp type.

The overall performance of the classifiers in this study underscores the effectiveness of using advanced point cloud geometry and depth map features for polyp classification. These features allowed the models to capture intricate patterns and enhance their ability to distinguish between different polyp types. Additionally, the results suggest that future improvements in feature extraction, along with the integration of additional data sources (e.g., endoscopic video), could further enhance model performance and broaden its generalization across various polyp types.

4.5 Summary

This chapter outlines the implementation, evaluation, and results of the machine learning classifiers (XGBoost, MLP, LightGBM, and KNN) applied to polyp classification based on 3D geometry and point cloud features. The experiments were conducted using cloud-based platforms, leveraging high-performance computing resources to ensure efficient training and testing. Industry-standard libraries such as Scikit-Learn and custom implementations for feature extraction and mesh-based processing were utilized for streamlined model development.

The proposed models demonstrated strong performance, with XGBoost emerging as the most accurate and reliable classifier for polyp detection. It achieved high classification accuracy, particularly for the minority polyp type, with low false positives and false negatives. The models effectively handled the class imbalance in the dataset, showcasing the capability of machine learning classifiers to work with complex 3D geometry and point cloud data for polyp classification.

Evaluation metrics such as precision, recall, F1 score, and confusion matrix analysis confirmed the robustness of the XGBoost model. The results highlight its potential for deployment in clinical settings where accurate and early detection of polyps is crucial for diagnostic purposes. Additional evaluations, including ROC curve analysis, further validated the discriminatory power of XGBoost, making it a promising tool for medical applications.

Overall, this work provides a valuable contribution to the development of automated systems for polyp classification, offering significant potential for improving early detection, diagnosis, and treatment planning in clinical practice.

Chapter 5

Engineering Standards and Design Challenges

5.1 Compliance with the Standards

5.1.1 Software Standards

This study ensures that scalability, compatibility, and efficiency are largely achieved through the utilization of specific tools and software. Python stands out as the leading programming language in areas such as data management, image processing, and model implementation, largely due to its extensive library ecosystem that includes Pandas, OpenCV, and NumPy. This allows for seamless integration of advanced learning frameworks such as TensorFlow and PyTorch, which are crucial for the training and refinement of models applicable to tasks including segmentation, classification, and 3D reconstruction. For the processing requirements of extensive datasets such as the PICCOLLO dataset, these models utilize GPU acceleration.

Deep learning activities and model training find Google Colab to be perfect because of its availability to free GPUs and Python integration, which enable it as the working platform. On GitHub, version control also ensures well-documented, repeatable, ordered code. Google Drive and Kaggle provide excellent cloud storage for datasets as well as basic access and faultless Python script interaction. Stressing code modularity, documentation, and maintainability helps to ensure that the work follows IEEE 730-2014 Standard for Software Quality Assurance Processes, therefore meeting industry standards for software engineering.

5.1.2 Hardware Standards

Selected to strike a compromise between performance and cost-effectiveness, the hardware configuration for this project combines local development system with cloud-based tools. We depend on Google Colab, which offers access to an NVIDIA Tesla T4 GPU, for much of the computing activity. This GPU is fundamentally essential for teaching the deep learning models used, as it facilitates mixed-precision training by utilizing both FP16 and FP32 data types, thereby accelerating the process. This capability is crucial for tasks such as 3D point cloud generation and depth map estimation, where efficient handling of substantial data sets is essential.

Apart from the cloud environment, I use a local development setup with a PC running Intel Core i7 CPU, 16GB of RAM, and SSD. Even without GPU capabilities, this system is still quite useful for data preparation, model diagnosis, and smaller test execution. This approach is particularly beneficial for activities such as preliminary model training and validation assessments, which do not require the extensive capabilities of cloud infrastructure.

Integrating Google Colab's robust cloud features with a minimal local configuration ensures that the project maintains the necessary processing power when it matters most, while also providing flexibility and cost-effectiveness. A strong combination of effectiveness and usability will facilitate the project's progression at every stage.

5.1.3 Communication Standards

From data gathering to model building, this project relies on precise and clear communication throughout the data collection process to ensure everything functions as intended. The integration of digital technologies with personal interactions ensures alignment and facilitates ongoing communication. We hold frequent virtual meetings via Google Meet or Zoom to keep everyone informed and aligned.. These meetings give a forum for addressing development, problem-solving, and future directions. They enable real-time teaming and help us to keep in line on goals. Google Docs make sharing of resources and documentation simple for us to do. It enables the team run naturally and rapidly change and suggest. Version control and code sharing also rely on GitHub, which allows us to manage the project and properly track changes. Regarding formal reports and technical documentation, we ensure correct and professional formatting by means of LaTeX more notably, Overleaf which guarantees Email is also quite vital for communication even if we make sure it's direct and succinct. This helps us to avoid useless back-and-forth and guarantees that all questions and updates are responded fast. Taken together, these communication rules help us to be a coherent team that guarantees our objective attainment and efficient forward direction of the project.

5.2 Impact on Society, Environment and Sustainability

The planned research on 3D reconstruction of colorectal polyps and their integration into AI-driven diagnostic systems is expected to have a major impact on many spheres of life, the surroundings, ethical issues, and sustainability.

5.2.1 Impact on Life

Still one of the most common and fatal malignancies globally, colorectal cancer (CRC) depends on early discovery to lower death rates. This work directly solves this problem by using modern 3D reconstruction methods to increase polyp identification accuracy and efficiency. The suggested approach allows more accurate and early polyp diagnosis by generating extremely detailed and accurate 3D models from regular colonoscopy pictures, therefore greatly lowering the chance of colonoscopy development. Especially in areas with restricted access to specialist medical professionals, the deployment of automated systems driven by artificial intelligence guarantees that more patients can gain from quicker diagnosis. Furthermore, this project guarantees that patients get quick and efficient therapy by means of bettering the diagnostic workflow and lowering human mistake. By enabling sophisticated diagnostic tools available in resource- constrained environments, where medical imaging equipment and professional radiologists may be rare, the approach also has the potential to democratize healthcare. Furthermore, the suggested approaches help to solve privacy issues related to the usage of patient data in medical research by guaranteeing the protection of delicate health information. By maintaining patient data privacy, the study creates a safer atmosphere for cooperative research and may thus help to overcome any obstacle to advance in the field of medical artificial intelligence.

5.2.2 Impact on Society and Environment

This project seeks to greatly increase access to healthcare services and sustainability. The combination of 3D reconstruction in colonoscopy tests opens the path for creative diagnosis approaches to be used in fields usually lacking expensive, high-tech medical instruments. Because colonoscopy allows the generation of 3D models from standard 2D colonoscopy images, therefore reducing the total costs connected with these procedures, colonoscopy is often more accessible and cost-effective in developing nations.

The automation of polyp detection is another major benefit of this research. By automating the process, healthcare professionals will have less of a workload when it comes to analyzing colonoscopy images. This, in turn, allows them to dedicate more time and focus to other aspects of patient care, improving the overall efficiency of medical practices. This is particularly important in a healthcare environment where time and resources are often limited, as it helps ensure that medical staff can provide the best possible care to their patients.

An other main advantage of this work is the automation of polyp detection. Automating the process will help medical experts to analyze colonoscopy images with less effort. This therefore enables them to provide other facets of patient care more time and attention, so increasing the general effectiveness of medical procedures. In a healthcare setting when time and resources are sometimes limited, this is especially crucial since it guarantees that medical staff members may treat their patients with the best available quality.

5.2.3 Ethical Aspects

Especially concerning patient privacy and data security, ethical considerations hold significant importance for this study. The implementation of robust encryption and stringent access controls ensures that the system adheres to high standards, such as GDPR and HIPAA, safeguarding sensitive patient information. Clarity is crucial; the system enables healthcare providers to understand the basis of projections, thereby fostering trust in the technology through Explainable AI (XAI) methods. Furthermore, the structure seeks to be fair among several patient groups, so guaranteeing the objective and fair performance of the artificial intelligence, which is essential for fostering system faith.

5.2.4 Sustainability Plan

Scalability made possible by the architecture helps to incorporate forthcoming developments in healthcare technologies. Pre-trained models and datasets are among open-access materials that encourage cooperation and inspire medical artificial intelligence researchers. The system will be guaranteed to be strong and flexible by regular updates informed by user insights and new technology developments. Moreover, cooperation with governmental agencies and healthcare institutions will help to implement and maintain it in clinical settings, so ensuring its ongoing relevance and availability for next years.

5.3 Project Management and Financial Analysis

Project Management

Good project management determines both the timely achievement of objectives for this research and the efficient use of resources. The program followed a rigorous approach with three phases: planning, implementing, and observing. First focus in the planning phase was ensuring key resources—including hardware improvements, computational tools, and specialized training via online courses. Making use of an NVIDIA Tesla T4 GPU's processing capacity, the project created models and trained on Google Colab during the implementation stage. One local PC was also upgraded to assist with activities including testing and debugging.

The project timeline was sequential, beginning with data preprocessing, followed by model development, training, and evaluation. Risk management strategies were implemented, including secure data backup, effective utilization of computational resources, and access to premium tools to address potential bottlenecks. Regular monitoring and adjustments ensured the project remained on track and milestones were met effectively.

Financial Analysis

The financial management of this project followed a cost-effective approach, ensuring that necessary resources were acquired without exceeding the allocated budget. Below is a breakdown of the financial requirements for the project:

Table 5.1: Cost Breakdown for Learning and Development

Item	**Cost (BDT)**	**Purpose**
Deep Learning and ML Course	8,000	Acquiring advanced skills and knowledge
Hardware Upgrade (Core i5, SSD)	10,000	Enhancing computational capacity for local debugging and processing
Premium Application Subscription	4,000	Accessing advanced features for efficient model development
Total	**22,000**	

5.4 Complex Engineering Problem

5.4.1 Complex Problem Solving

Over here, We are going to map complex problem solve according to this study-

Table 5.2: Mapping with complex problem solving.

EP1 Dept of Knowledge	EP2 Range of Conflicting Requirements	EP3 Depth of Analysis	EP4 Familiarity of Issues	EP5 Extent of Applicable Codes	EP6 Extent of Stakeholder Involvement	EP7 Interdependence
√		√	√			√

EP1 – Depth of Knowledge

The project combines medical imaging, computer vision, deep learning, and computational geometry among other disciplines of knowledge. Building a 3D reconstruction from 2D colonoscopy pictures calls for a strong knowledge of mesh generation, depth estimation, and image processing. The combination of machine learning for segmentation and polyp categorization adds still another level of complexity and calls for knowledge of neural network topologies and ethical issues with artificial intelligence. This project is a complete challenge since it synthesizes information from many spheres.

EP3 – Depth of Analysis

The project necessitates in-depth evaluations that demand a comprehensive understanding of various disciplines. In detail, it encompasses the examination and reproduction of techniques such as depth estimation, mesh construction, 3D point cloud generation, and feature extraction from 2D endoscopic images. Furthermore applied for appropriate classification and validation are machine learning techniques. Comprehensive analysis ensures the dependability and robustness of the system by way of performance testing using many criteria including accuracy, recall, precision, and F1-score. Apart from traditional image processing, the level of analysis necessary investigates innovative approaches for model validation.

EP4 – Familiarity of Issues

Investigating a spectrum of fast changing and understudied topics in medical imaging, our work focuses on depth estimation combining endoscopic frames and monocular images for three-dimensional reconstruction. Dealing with these difficulties requires a strong awareness of computer vision ideas in addition to knowledge of the constraints placed on real clinical settings including noise, occlusion, and low-quality images. Furthermore adding complication is the research navigating the ethical and legal elements of artificial intelligence in healthcare.

EP7 - Interdependence

The interaction among several elements of the project is absolutely important. The framework makes use of 3D reconstruction, depth estimate, mesh generation, artificial intelligence for classification. Every approach has to cooperate perfectly with others to create a logical system. For example, whilst the polyp classification depends on precise segmentation and 3D modeling of the polyp regions, accurate mesh generation depends on proper depth map estimate. This interdependence calls for careful coordination and data source optimization among the methods.

Mapping with Knowledge Profile for EP1

Table 5.3 is designed to map the EP1 to the Knowledge Profile.

Table 5.3: Mapping with knowledge Profile.

K3 Engineering Fundamentals	K4 Specialist Knowledge	K5 Engineering Design	K6 Engineering Practice	K8 Research Literature
√	√	√	√	√

K3 - Engineering Fundamentals

The project applies key engineering fundamentals related to machine learning, specifically image processing, neural networks and deep learning architectures like Convolutional Neural Networks (CNNs). Additionally, it leverages advanced image processing techniques for depth estimation, mesh generation, and point cloud generation, which are core elements of engineering design in computational geometry. The understanding of these fundamentals supports the core methodology of transforming 2D colonoscopy images into 3D models, crucial for colorectal polyp detection.

K4 – Specialist Knowledge

The research requires specialized knowledge in integration of 3D reconstruction with deep learning models. Expertise in medical imaging, specifically colonoscopy images, is necessary to address challenges like noise removal, occlusions, and depth estimation. Knowledge of specialized algorithms such as Ball Pivoting Algorithm (BPA) and techniques for reflection removal and feature extraction from point clouds are essential for achieving accurate 3D models and improving clinical diagnostics.

K5 – Engineering Design

The project necessitates thoughtful engineering design to develop a comprehensive pipeline for 3D reconstruction and analysis. This involves the careful selection of machine learning models, depth estimation techniques, and point cloud feature extraction methods. The integration of various components, such as Paris classification-based dataset splitting, best frame selection, and reflection removal using EndoSRR, showcases the intricate design necessary to ensure a seamless, scalable, and efficient system.

K6 – Engineering Practice

Practical application of engineering tools such as Python and machine learning frameworks like TensorFlow and PyTorch are integral to the project. Additionally, utilizing cloud-based environments like Google Colab, along with computational resources like GPUs, ensures the scalability and computational efficiency required for deep learning tasks. Moreover, adherence to healthcare privacy standards (e.g., GDPR and HIPAA) ensures that the project complies with legal and ethical requirements, demonstrating the practical aspect of engineering in medical applications.

K8 - Research Literature

The proposed framework builds upon the existing body of research, contributing new insights and methodologies to the field of colorectal cancer detection through 3D modeling.

The work also reflects an ongoing engagement with academic literature to ensure the proposed techniques are state-of-the-art and relevant to current trends in medical AI research.

5.4.2 Engineering Activities

Here this study goes through by this Engineering activities

Table 5.4: Mapping with complex engineering activities.

EA1 Range of resources	EA2 Level of Interaction	EA3 Innovation	EA4 Consequences for society and environment	EA5 Familiarity
√	√	√	√	√

EA1: Range of Resources

The project effectively utilizes a combination of cloud-based and local resources for processing, training, and validating deep learning models. Cloud platforms like Google Colab provide powerful GPUs for training large models, while local hardware aids in preprocessing and debugging tasks. Additionally, advanced software frameworks such as PyTorch and TensorFlow are employed for developing and deploying models. The use of these diverse resources ensures the smooth operation of the system while also balancing performance and cost-efficiency.

EA2: Level of Interaction

The interaction among several elements of the project is absolutely important. The framework makes use of 3D reconstruction, depth estimate, mesh generation, artificial intelligence for classification. Every approach has to cooperate perfectly with others to create a logical system. For example, whilst the polyp classification depends on precise segmentation and 3D modeling of the polyp regions, accurate mesh generation depends on proper depth map estimate. This interdependence calls for careful coordination and data source optimization among the methods.

EA3: Innovation

Combining several advanced technologies into a single pipeline—including 3D point cloud production, mesh generation, reflection reduction, and depth map estimation—this project is unique in that it allows The study addresses colorectal polyp detection including occlusions, noise, and depth estimates by proposing a new approach integrating modern artificial intelligence technology for enhanced diagnosis.

EA4: Consequences for Society and Environment

This study has great possible impact for society as well as the surroundings. By means of advanced 3D reconstructions of colonoscopy images, the project aims to save lives and reduce the burden on healthcare systems by early colorectal cancer detection. Especially in places with limited access to high-end medical equipment, the recommended approach makes advanced tests more readily available. Moreover, by reducing the need on costly and energy-consuming specialist hardware, the study helps to reduce the environmental effect of medical imaging equipment.

EA5: Familiarity

This project emphasizes recognized techniques in machine learning and computer vision, such as 3D reconstruction, while ensuring a foundation in reliable and established technology. While these techniques are recognized across the AI community, their use in 3D reconstruction and colorectal polyp detection introduces a novel challenge. The project combines various established technologies into a coherent sequence, grounded in recognized engineering practices and innovative concepts.

5.5 Summary

This chapter underlines the primary engineering aspects, design challenges, and standards followed during the development of this project. The application of generally accepted benchmarks for hardware, software, and communication is discussed in this work thereby guaranteeing the effective execution and expansion potential of the system. Leveraging cloud resources such Google Colab with powerful GPUs, the software framework uses Python along with deep learning frameworks such as PyTorch and TensorFlow to manage the challenging needs of 3D reconstruction, depth estimate, and polyp classification. Combining local processing capacity like an Intel Core i5 PC with 16GB RAM along with cloud-based GPUs guarantees the efficient performance of training and validation tasks.

The initiative also covers important ethical issues, especially with relation to privacy and data security. Furthermore included in the framework are modern 3D reconstruction techniques and approaches for minimizing reflections, improving polyp visibility, and giving doctors high-quality images that support better diagnosis and treatment planning. Especially in regions with limited access to advanced diagnostic methods, this finding could greatly impact the identification of colorectal cancer. The project aims to utilize routine colonoscopy images for 3D reconstruction to reduce the costs associated with identifying colorectal polyps, ultimately enhancing the affordability and accessibility of this technology with a positive impact on social and environmental factors. The decreased dependence on costly imaging technologies and the transition to cloud-based solutions contribute to lowering environmental impact, thus supporting sustainability in healthcare.

The parts on financial analysis and project management provide a smart strategy to achieve outstanding performance while preserving economy. Resources were set aside for training, computer tools, and necessary hardware improvements, therefore ensuring the viability of the project and preserving fiscal discipline. The financial study shows that the project can generate notable results and operate within the allocated budget.

Additionally, the engineering efforts involved in this study align with the systematic approach required to address complex issues in colorectal cancer diagnosis. In conjunction with sophisticated feature extraction and analysis, the processes of 3D reconstruction, depth estimation, and mesh generation illustrate the project's alignment with engineering best practices. This investigation offers a comprehensive and adaptable basis for advancing colorectal cancer detection by showcasing the application of advanced engineering concepts to healthcare challenges.

Chapter 6

Conclusion

6.1 Summary

This paper presents a significant progress in the identification and diagnosis of colorectal polyps by means of modern 3D reconstruction techniques. We have merged several strong methods like Paris classification-based dataset splitting, reflection removal with EndoSRR, and depth estimate via ZoeDepth using a comprehensive pipeline to increase the accuracy and quality of polyp identification from standard 2D colonoscopy images. By creating 3D point clouds and models made feasible by this work, polyp shape can be considerably more precisely analyzed including crucial aspects like curvature and slope. Apart from diagnosis, these geometric features enable doctors to make better decisions regarding treatment planning. This work is special because of combining these innovative technologies into a perfect, automated process with 2D colonoscopy images. This approach generates fresh chances for more exact and consistent polyp detection over the limits of standard 2D imaging. Thanks to its cloud-based design, the method also ensures that these advanced features can be utilized even on sites with restricted means. This raises the technological scalability and accessibility, therefore improves the quality of treatment in many clinical environments.

Better knowledge of polyp size, form, and structure helps one to show how much 3D reconstruction can raise diagnosis accuracy. Although this method is grounded on solid engineering, it benefits doctors by improving the accuracy of colorectal cancer diagnosis and so facilitating better treatment planning. Since it follows privacy laws and provides a way to strike between ethical obligation and creativity, the technology also guarantees respect of patient privacy. This work lays a basis for future developments and helps the expanding field of artificial intelligence-driven medical diagnostics. Modern 3D imaging combined with deep learning techniques will enable early cancer detection and provide clinicians with the tools they need to make faster, better decisions. Reducing colorectal cancer death rates is eventually the goal; hence, improving patient outcomes and allowing a wider range of healthcare providers to access these life-saving medicines.

6.2 Limitation

Limitations and Challenges

Although the method described in this paper has significant potential to improve colorectal polyp diagnosis, various restrictions and issues should be taken into account.

1. **Complexity of 3D Reconstruction:** While the 3D reconstruction of polyps offers more data, the quality of the 2D images affects point cloud generation and depth estimate accuracy. Inconsistent image quality that is, from poor illumination or obscurities could produce less reliable 3D reconstructions.
2. **Segmentation Errors:** Although powerful, the segmentation stage still causes problems in particularly complicated or small polyp forms. Small segmentation errors or noise in the point cloud data could impair the integrity of the 3D model, therefore affecting the validity of geometric feature analysis.
3. **Computational Demands:** The proposed method relies on cloud-based computational resources, which despite their scalabilitymay induce delays or reliance on internet connectivity particularly in settings with limited resources. Moreover computationally difficult is the development of 3D models from 2D images, which requires high-performance technology.
4. **Data Availability and Diversity:** The availability of a varied and high-quality dataset determines mostly the performance of the pipeline. Should datasets be small or uneven, the model's accuracy could be undermined. Generalizing the performance of the model over other types of polyps and patient demographics depends on a bigger and more varied dataset.
5. **Integration with Clinical Workflows:** The intricacy and demand for extra education among medical professionals suggest that adding this 3D reconstruction technology into effective therapeutic procedures may cause problems. If the system is to be most efficient in hectic clinical settings, it must be simple and easy for usage.
6. **Interpretability of 3D Models:** Although the 3D models offer improved vision, for doctors not versed with 3D imaging technologies reading them can be difficult. More work is required to give medical professionals sufficient instruction and to simplify the interface.

6.3 Future Work

We want to increase the accuracy of our machine learning models in future by means of improved extraction and use of 3D information from polyp structures. Among other major limitations, the lack of class annotations in our dataset reduced our ability for complete polyp categorization. To address this, future studies should incorporate expert-labeled data that clearly identifies polyp types, making the system more clinically relevant. We also plan to explore better mesh reconstruction techniques and experiment with more advanced classifiers to enhance prediction reliability. As the ultimate goal is to support doctors in diagnosis and treatment planning, improving model performance and moving closer to real-time analysis will be key steps forward.

References

[1] Marjan Laal. Innovation process in medical imaging. *Procedia-Social and Behavioral Sciences*, 81:60–64, 2013.

[2] Perry J Pickhardt, B Dustin Pooler, David H Kim, Cesare Hassan, Kristina A Matkowskyj, and Richard B Halberg. The natural history of colorectal polyps: overview of predictive static and dynamic features. *Gastroenterology Clinics of North America*, 47(3):515, 2018.

[3] R Johnson et al. A review of three-dimensional medical image visualization. *Health Data Science*, (1):83–101, 2022.

[4] Rammah Yousef, Gaurav Gupta, Nabhan Yousef, and Manju Khari. A holistic overview of deep learning approach in medical imaging. *Multimedia Systems*, 28(3):881–914, 2022.

[5] Sena Busra Yengec-Tasdemir, Zafer Aydin, Ebru Akay, Serkan Dogan, and Bulent Yilmaz. An effective colorectal polyp classification for histopathological images based on supervised contrastive learning. *Computers in Biology and Medicine*, 172:108267, 2024.

[6] Hui Chen, Yipeng Zuo, Yong Tong, and Li Zhu. 3d point cloud generation reconstruction from single image based on image retrieval. *Results in Optics*, 5:100124, 2021.

[7] Md Nur-A-Alam, Khandaker Mohammad Mohi Uddin, MMR Manu, Md Mahbubur Rahman, and Mostofa Kamal Nasir. An automatic system to detect colorectal polyp using hybrid fused method from colonoscopy images. *Intelligent Systems with Applications*, 22:200342, 2024.

[8] Qi-Xian Huang, Guo-Shiang Lin, and Hung-Min Sun. Classification of polyps in endoscopic images using self-supervised structured learning. *IEEE Access*, 11:50025–50037, 2023.

[9] Md Shakhawat Hossain, Md Mahmudur Rahman, M Mahbubul Syeed, Mohammad Faisal Uddin, Mahady Hasan, Md Aulad Hossain, Amel Ksibi, Mona M Jamjoom, Zahid Ullah, and Md Abdus Samad. Deeppoly: deep learning-based polyps

segmentation and classification for autonomous colonoscopy examination. *IEEE Access*, 11:95889–95902, 2023.

[10] Sadia Sultana Chowa, Sami Azam, Sidratul Montaha, Md Rahad Islam Bhuiyan, and Mirjam Jonkman. Improving the automated diagnosis of breast cancer with mesh reconstruction of ultrasound images incorporating 3d mesh features and a graph attention network. *Journal of Imaging Informatics in Medicine*, 37(3):1067–1085, 2024.

[11] Jason Rock, Tanmay Gupta, Justin Thorsen, JunYoung Gwak, Daeyun Shin, and Derek Hoiem. Completing 3d object shape from one depth image. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2484–2493, 2015.

[12] Farah Younas, Muhammad Usman, and Wei Qi Yan. A deep ensemble learning method for colorectal polyp classification with optimized network parameters. *Applied Intelligence*, 53(2):2410–2433, 2023.

[13] Adrian Krenzer, Stefan Heil, Daniel Fitting, Safa Matti, Wolfram G Zoller, Alexander Hann, and Frank Puppe. Automated classification of polyps using deep learning architectures and few-shot learning. *BMC Medical Imaging*, 23(1):59, 2023.

[14] Weiguo Cao, Marc J Pomeroy, Shu Zhang, Jiaxing Tan, Zhengrong Liang, Yongfeng Gao, Almas F Abbasi, and Perry J Pickhardt. An adaptive learning model for multiscale texture features in polyp classification via computed tomographic colonography. *Sensors*, 22(3):907, 2022.

[15] Yubo Zhang, Jan-Michael Frahm, Samuel Ehrenstein, Sarah K McGill, Julian G Rosenman, Shuxian Wang, and Stephen M Pizer. Colde: a depth estimation framework for colonoscopy reconstruction. *arXiv preprint arXiv:2111.10371*, 2021.

[16] Mohamed Yousuf, Samir Harb, Islam Alkabbany, Asem Ali, Salwa Elshazley, and Aly Farag. Colorectal polyps detection in virtual colonoscopy using 3d geometric features and deep learning. In *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 1–4. IEEE, 2024.

[17] Zaide Duran, Kubra Ozcan, and Muhammed Enes Atik. Classification of photogrammetric and airborne lidar point clouds using machine learning algorithms. *Drones*, 5(4):104, 2021.

[18] Ankur Deka, Yuji Iwahori, Manas Kamal Bhuyan, Pradipta Sasmal, and Kunio Kasugai. Dense 3d reconstruction of endoscopic polyp. *BIOIMAGING*, 2018:159–166, 2018.

[19] Bilal Ahmad, Pål Anders Floor, Ivar Farup, and Casper Find Andersen. Single-image-based 3d reconstruction of endoscopic images. *Journal of Imaging*, 10(4):82, 2024.

[20] Philipp Wesp, Sergio Grosu, Anno Graser, Stefan Maurus, Christian Schulz, Thomas Knösel, Matthias P Fabritius, Balthasar Schachtner, Benjamin M Yeh, Clemens C Cyran, et al. Deep learning in ct colonography: differentiating premalignant from benign colorectal polyps. *European Radiology*, 32(7):4749–4759, 2022.

[21] Alba Nogueira-Rodríguez, Rubén Domínguez-Carbajales, Fernando Campos-Tato, Jesus Herrero, Manuel Puga, David Remedios, Laura Rivas, Eloy Sánchez, Agueda Iglesias, Joaquín Cubiella, et al. Real-time polyp detection model using convolutional neural networks. *Neural Computing and Applications*, 34(13):10375–10396, 2022.

[22] Aji Resindra Widya, Yusuke Monno, Masatoshi Okutomi, Sho Suzuki, Takuji Gotoda, and Kenji Miki. Whole stomach 3d reconstruction and frame localization from monocular endoscope video. *IEEE Journal of Translational Engineering in Health and Medicine*, 7:1–10, 2019.

[23] Bilal Ahmad, Pål Anders Floor, Ivar Farup, and Øistein Hovde. 3d reconstruction of gastrointestinal regions using single-view methods. *Ieee Access*, 11:61103–61117, 2023.

[24] Saad Nadeem and Arie Kaufman. Depth reconstruction and computer-aided polyp detection in optical colonoscopy video frames. *arXiv preprint arXiv:1609.01329*, 2016.

[25] Bingyan Wang, Jinjia Peng, Hui Li, and Huibing Wang. Generating high-quality texture via panoramic feature aggregation for large mask inpainting. *Knowledge-Based Systems*, 286:111382, 2024.

[26] Junde Wu, Ziyue Wang, Mingxuan Hong, Wei Ji, Huazhu Fu, Yanwu Xu, Min Xu, and Yueming Jin. Medical sam adapter: Adapting segment anything model for medical image segmentation. *Medical image analysis*, 102:103547, 2025.

[27] Sijie Zhao, Wenbo Hu, Xiaodong Cun, Yong Zhang, Xiaoyu Li, Zhe Kong, Xiangjun Gao, Muyao Niu, and Ying Shan. Stereocrafter: Diffusion-based generation of long and high-fidelity stereoscopic 3d from monocular videos. *arXiv preprint arXiv:2409.07447*, 2024.

[28] Ashkan Ganj, Hang Su, and Tian Guo. Hybriddepth: Robust metric depth fusion by leveraging depth from focus and single-image priors. In *2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 973–982. IEEE, 2025.

[29] Zewei Wu, Cui Wang, Wei Zhang, Guodong Sun, Wei Ke, and Zhang Xiong. Online 3d behavioral tracking of aquatic model organism with a dual-camera system. *Advanced Engineering Informatics*, 61:102481, 2024.

[30] Zelin Huang, Yanjie Zhu, Wen Xiong, and CS Cai. Bridge scour morphology identification and reconstruction using 3d sonar point cloud data. *Automation in Construction*, 175:106205, 2025.

212-15-4217

ORIGINALITY REPORT

15%	10%	9%	7%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	4%
2	arxiv.org Internet Source	1%
3	Submitted to United International University Student Paper	1%
4	Submitted to Malta College of Arts,Science and Technology Student Paper	1%
5	dblp.dagstuhl.de Internet Source	<1%
6	dblp.org Internet Source	<1%
7	Submitted to University of New South Wales Student Paper	<1%
8	opus.lib.uts.edu.au Internet Source	<1%
9	Tasneem Ahmed, Shrish Bajpai, Mohammad Faisal, Suman Lata Tripathi. "Advances in Science, Engineering and Technology: A Path to the Future - Proceedings of the International Conference on Advances in Science, Engineering and Technology (ICASET - 2024), Organized by Department of	<1%