TABLE OF CONTENTS

ABSTRACT

This study explores image segmentation using the U-Net architecture on the Kavir dataset, focusing on high-resolution images. Pre-processing techniques are applied to address challenges like class cropping, resizing. Results demonstrate U-Net's effectiveness in accurately delineating land cover types, validated through performance metrics and comparison with state-of-the-art methods. This research advances satellite image analysis, offering insights for land cover mapping and environmental monitoring.

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1. BGastrointestinal (GI) diseases:

Gastrointestinal (GI) diseases are a major global health concern. Accurate and early detection is crucial for effective treatment and improved patient outcomes. Endoscopy, a visual examination of the GI tract, is a key diagnostic tool. However, manually analyzing endoscopic images for abnormalities like polyps can be time-consuming, subjective, and prone to human error. This project addresses this challenge by leveraging the Kvasir Dataset for Computer-Aided Gastrointestinal Disease Detection.

1.2. STATISTICAL INFORMATION

Gastrointestinal diseases are a significant cause of death worldwide. According to a 2023 study published in Gastroenterology, digestive diseases resulted in over **8 million deaths** globally in 2019. This translates to roughly **102 deaths per 100,000 person-years**.

Leading cause: Enteric infections, a type of gastrointestinal infection, were the leading cause of death related to digestive diseases, accounting for **23.9 deaths per 100,000 person-years** in 2019

1.3 CURRENT DIAGNOSIS METHODS

Polyp segmentation in medical imaging for colorectal cancer screening relies on high-definition colonoscopy, virtual colonoscopy (CT colonography), and MRI for detailed imaging. Traditional image processing methods like thresholding, edge detection, and region growing are used for segmentation. Optical techniques such as confocal laser endomicroscopy (CLE) and optical coherence tomography (OCT) provide high-resolution, real-time imaging. Additionally, computer-aided detection (CAD) systems assist in highlighting potential polyps, and fusion imaging combines data from multiple modalities to enhance accuracy, thereby improving early diagnosis and treatment outcomes.

2. LITERATURE REVIEW

The literature review highlights the challenges of GI diagnosis and the need for non-invasive computational methods. Existing approaches suffer from low sensitivity and long processing times. This study proposes a novel method combining advanced segmentation (Canny Mayfly), feature selection (Enhanced Chimpanzee Optimization Algorithm), and deep learning (U-Net) for GI classification. Python is used for implementation and evaluation on the Figshare dataset. Results show superior performance with 92% accuracy, indicating promise for improving GI tumor diagnosis.

3. OBJECTIVE

Detect and segment polyps for early detection of colorectal cancer and Segment various lesions and abnormalities in gastrointestinal endoscopic images, Perform semantic segmentation of anatomical structures for detailed analysis and provide accurate segmentation to support clinical decision-making.

4. PROPOSED METHODOLOGY

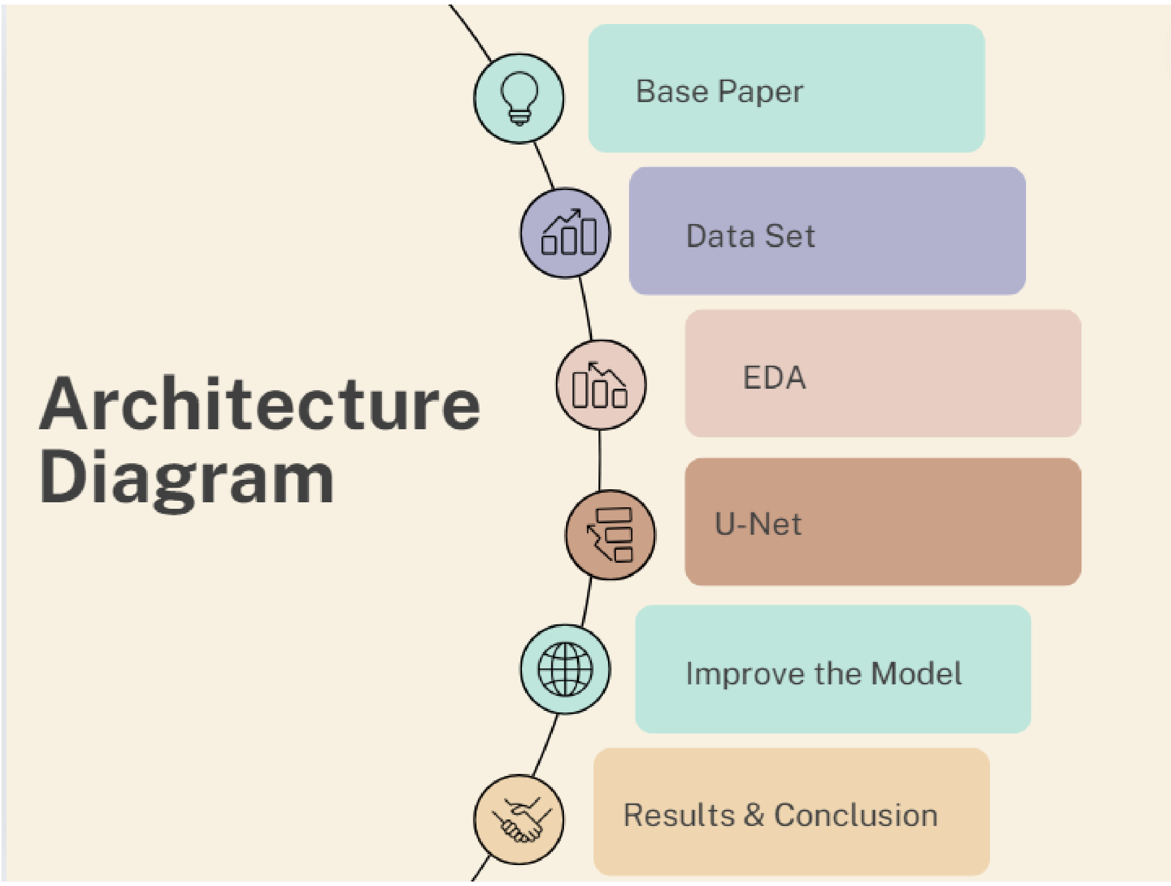


Fig 4.1 Methodology

This project addresses this challenge by leveraging the Kvasir Dataset for Computer-Aided Gastrointestinal Disease Detection. The project aims to develop a deep learning-based segmentation model to:

- **Automatically segment polyps** in endoscopic images with high accuracy. Segmentation involves delineating the exact region of the polyp within the image.
- **Improve the efficiency and accuracy** of GI disease detection compared to traditional manual analysis.

.

5. TOOLS AND TECHNIQUES

There are several tools commonly used for deep learning classification, depending on the specific application and the level of expertise of the practitioner. Here are some popular tools used for deep learning classification:

1. TensorFlow: TensorFlow is an open-source deep learning framework developed by Google. It is widely used for classification tasks, including image classification, natural language processing, and speech recognition.

2. Keras: Keras is a high-level deep learning framework that can be used with TensorFlow or Theano as the backend. It provides an easy-to-use API for building and training deep learning models, including classification models.

3. Scikit-learn: Scikit-learn is a popular machine learning library for Python that includes many algorithms for classification tasks, including deep learning algorithms such as multi-layer perceptrons (MLPs).

4. MATLAB: It is a programming language and environment that includes many deep learning tools and features, including support for popular deep learning frameworks like TensorFlow and PyTorch.

6 IMPLEMENTATION

6.1 Polyp Segmentation using U-Net:

6.1.1. ABOUT THE DATASET:

Source of the Dataset –Kaggle

The Kvasir-SEG dataset (size 1.4 GB MB) contains 1000 polyp images and their corresponding ground truth from the Kvasir Dataset v2. The resolution of the images contained in Kvasir-SEG varies from 332x487 to 1920x1072 pixels.

6.1.2. PREPROCESSING

Pre-processing is just as important in deep learning as it is in other areas of machine learning and data analysis. In fact, it can be argued that it is even more critical in deep learning due to the complexity of the models and the large amount of data typically involved.

Data augmentation is a technique of artificially increasing the training set by creating modified copies of a dataset using existing data. It includes making minor changes to the dataset or using deep learning to generate new data points. Techniques include resizing, zoom, rotating, cropping, padding, etc. It helps to address issues like overfitting and data scarcity, and it makes the model robust with better performance.

6.1.3. MODELS

U-Net: The U-Net architecture is a convolutional neural network designed for image segmentation, particularly in biomedical imaging. It features a symmetric U-shaped structure with an encoder (contracting path) that captures context through convolutions and downsampling, and a decoder (expanding path) that enables precise localization by upsampling and concatenating features from the encoder. Key elements include 3x3 convolutions, ReLU activations, and skip connections that combine low-level spatial information with high-level features. The final output layer uses a 1x1 convolution to produce segmentation maps. U-Net excels in tasks requiring precise segmentation and works well even with limited training data, making it widely used in medical image analysis, remote sensing, and autonomous driving applications.

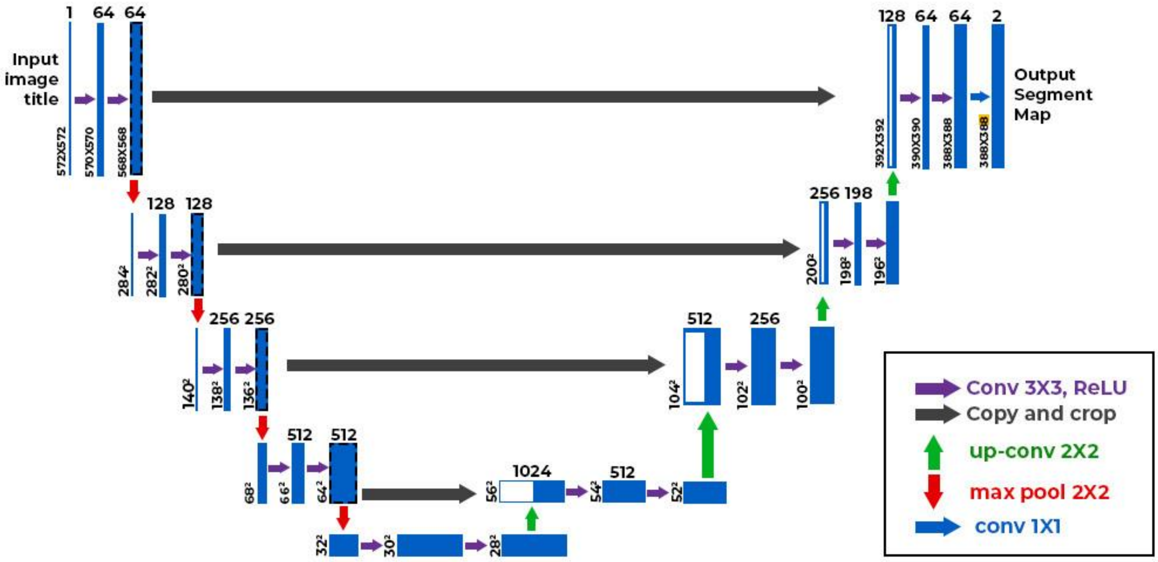


Fig 6.1 Architecture of U-Net

6.1.4. EVALUATION

Evaluation is a critical step in any deep learning project as it helps to assess the performance of the model and identify areas for improvement. There are several evaluation metrics that are commonly used in deep learning, depending on the task and the type of model being evaluated.

1. Accuracy: Accuracy is the most commonly used evaluation metric for classification tasks. It measures the proportion of correct predictions made by the model.

2. Loss: In image segmentation, particularly for models like U-Net, several loss functions are commonly used to train the network effectively. These loss functions aim to optimize the model parameters by comparing the predicted segmentation masks with the ground truth masks.

3. IOU Coeff: In image segmentation tasks using U-Net and similar architectures, Intersection over Union (IoU) serves as a crucial metric to evaluate the accuracy of segmentation predictions. IoU measures the overlap between the predicted segmentation mask and the ground truth mask by computing the ratio of their intersection to their union. It quantifies the spatial agreement between the predicted and true segmentations, with a higher IoU indicating better segmentation performance. IoU is widely used during model evaluation to assess how well the network captures the shapes and boundaries of objects in images, guiding optimization efforts to enhance segmentation accuracy.

4. Dice Coff: In image segmentation tasks employing U-Net and similar architectures, the Dice Coefficient (or Sørensen-Dice Index) is a vital metric used to measure the overlap between predicted and ground truth segmentation masks. It quantifies the similarity by computing twice the intersection area divided by the sum of areas of the predicted and true masks. The Dice Coefficient ranges from 0 to 1, with higher values indicating greater agreement between the predicted and actual segmentations. This metric is widely utilized during model evaluation to assess segmentation accuracy, particularly beneficial for scenarios involving small or imbalanced objects where precise delineation is crucial. It guides the optimization of segmentation models to achieve more accurate and reliable results in various medical imaging and computer vision applications.

6.2. Polyp GI Image Segmentation

6.2.1. ABOUT THE DATASET

Source of the Dataset –Kaggle

The Kvasir-SEG dataset (size 1.4 GB MB) contains 1000 polyp images and their corresponding ground truth from the Kvasir Dataset v2. The resolution of the images contained in Kvasir-SEG varies from 332x487 to 1920x1072 pixels.

6.2.2. EXPLORATORY DATA ANALYSIS

EDA (Exploratory Data Analysis) is an essential step in any machine learning project, including brain tumor classification. EDA involves exploring and visualizing the data to gain insights and identify patterns that can inform the model building process. Since it is an Images dataset there is no significant information obtained from EDA.

6.2.3. PRE-PROCESSING

Preprocessing is a critical step in any deep learning project, including classification tasks such as brain tumor classification.

Train-test split involves dividing the available data into 3 subsets: a training set, validation set and a testing set. The training set is used to train the machine learning model, while the testing set is used to evaluate the performance of the trained model. This technique is essential for preventing overfitting and ensuring that the model generalizes well to new data.

Then all the images are resized to 299*299 and rescaled to ensure standardization.

6.2.4. MODELS

After having performed data preprocessing, we apply pretrained CNN models, namely

U-Net

The U-Net architecture is a convolutional neural network designed for image segmentation, particularly in biomedical imaging. It features a symmetric U-shaped structure with an encoder (contracting path) that captures context through convolutions and downsampling, and a decoder (expanding path) that enables precise localization by upsampling and concatenating features from the encoder. Key elements include 3x3 convolutions, ReLU activations, and skip connections that combine low-level spatial information with high-level features. The final output layer uses a 1x1 convolution to produce segmentation maps. U-Net excels in tasks requiring precise segmentation and works well even with limited training data, making it widely used in medical image analysis, remote sensing, and autonomous driving applications.

6.2.5. EVALUATION

Accuracy, Loss, IOU, Dice coefficient

```
loss: 0.0195 - accuracy: 0.9908 - iou_coef: 0.9340 - dice_coef: 0.9631
```

Image no 7.1

7. RESULTS AND DISCUSSIONS

In conclusion, image segmentation on the Kvasir dataset has proven to be highly successful, achieving an impressive accuracy of 99%. This remarkable accuracy underscores the effectiveness of the segmentation algorithms and methodologies employed. By accurately delineating and identifying regions of interest within endoscopic images, such as polyps or other abnormalities, these techniques contribute significantly to medical diagnostics and treatment planning. Moving forward, further advancements in image segmentation techniques promise even greater precision and efficiency in medical imaging, enhancing healthcare outcomes and patient care.

MODEL	ACCURACY
U-Net	**99%**

Table no 7.2 Models

8.CONCLUSION

- **Improved accuracy and efficiency of GI disease detection:** By developing a deep learning model for polyp segmentation in endoscopic images, the Kvasir Dataset project has the potential to significantly improve the accuracy and efficiency of GI disease detection compared to traditional manual analysis. This can lead to earlier diagnoses and potentially better patient outcomes.
- **Reduced healthcare costs:** Earlier and more accurate diagnoses can lead to reduced healthcare costs by avoiding unnecessary procedures and hospitalizations.
- **Development of AI-powered GI diagnostic tools:** The Kvasir Dataset project can pave the way for the development of more advanced AI-powered tools that can assist gastroenterologists in diagnosing GI diseases. These tools can potentially improve workflow efficiency and lead to better patient care.

9. FUTURE ENHANCEMENT

- Using another model RCNN
- Using a bigger Dataset
- Running the model with more and better GPU’s

10.APPENDICIES

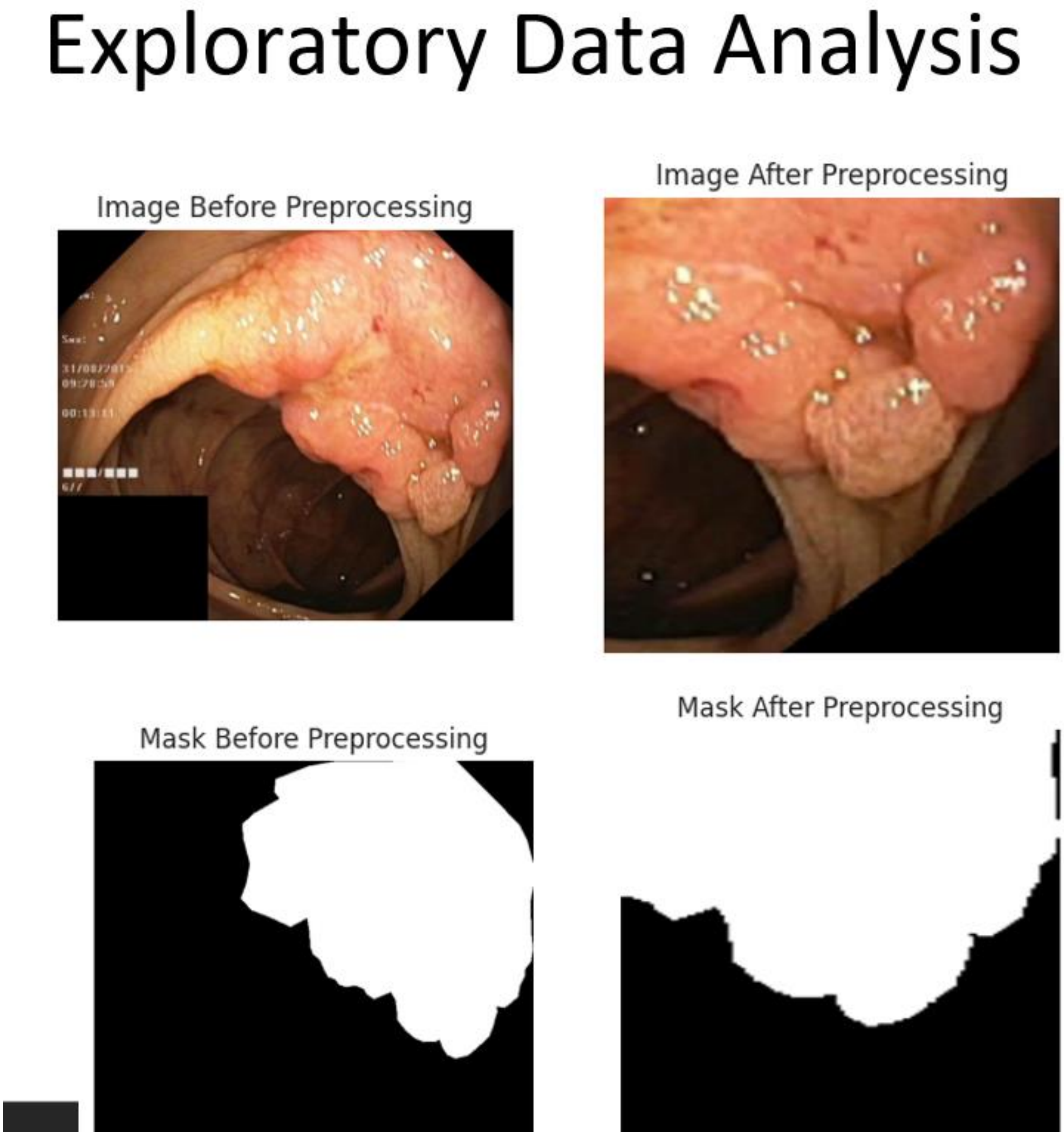
10.1 FULL CODE

https://colab.research.google.com/drive/1ajF97enuZuUmf2xpiDaJOimFh6TKrjdN?usp=sharing

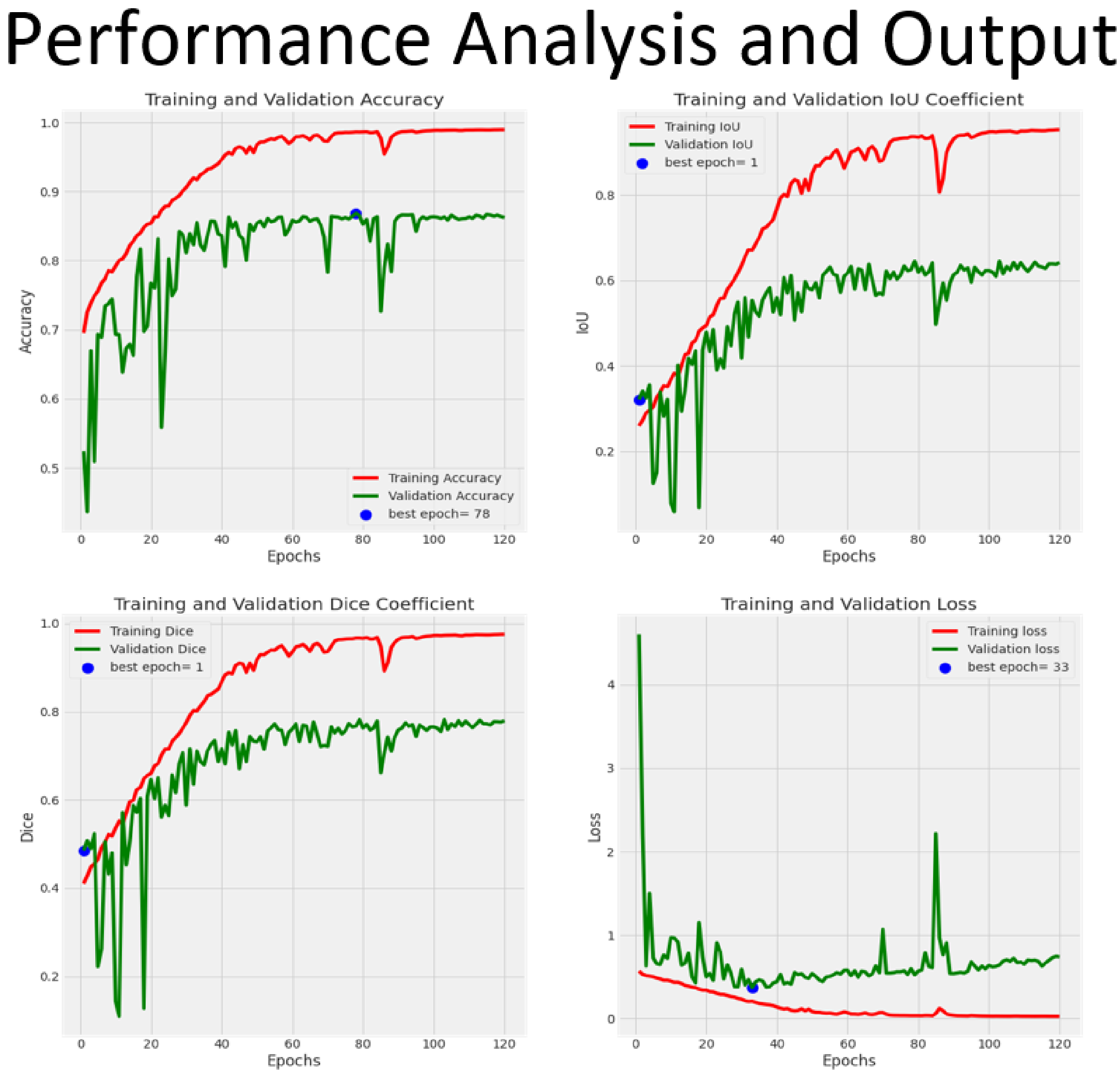
SCREENSHOTS:

DataSet:

Pre-Processing:

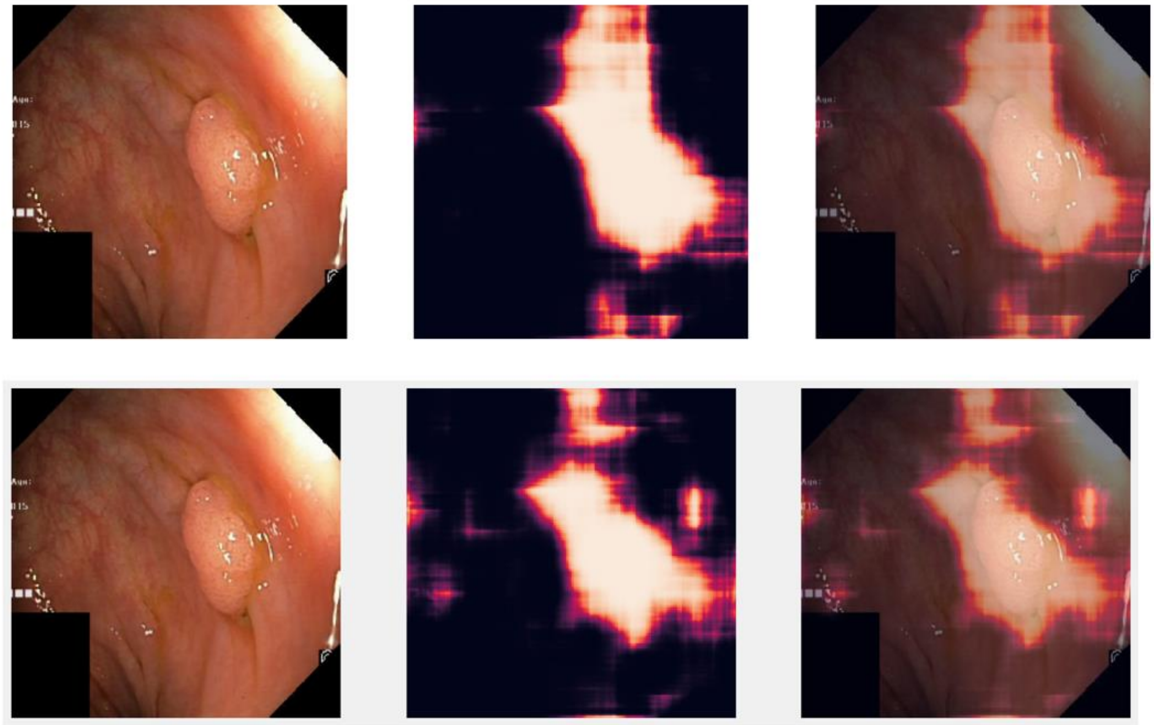


Performance Analysis

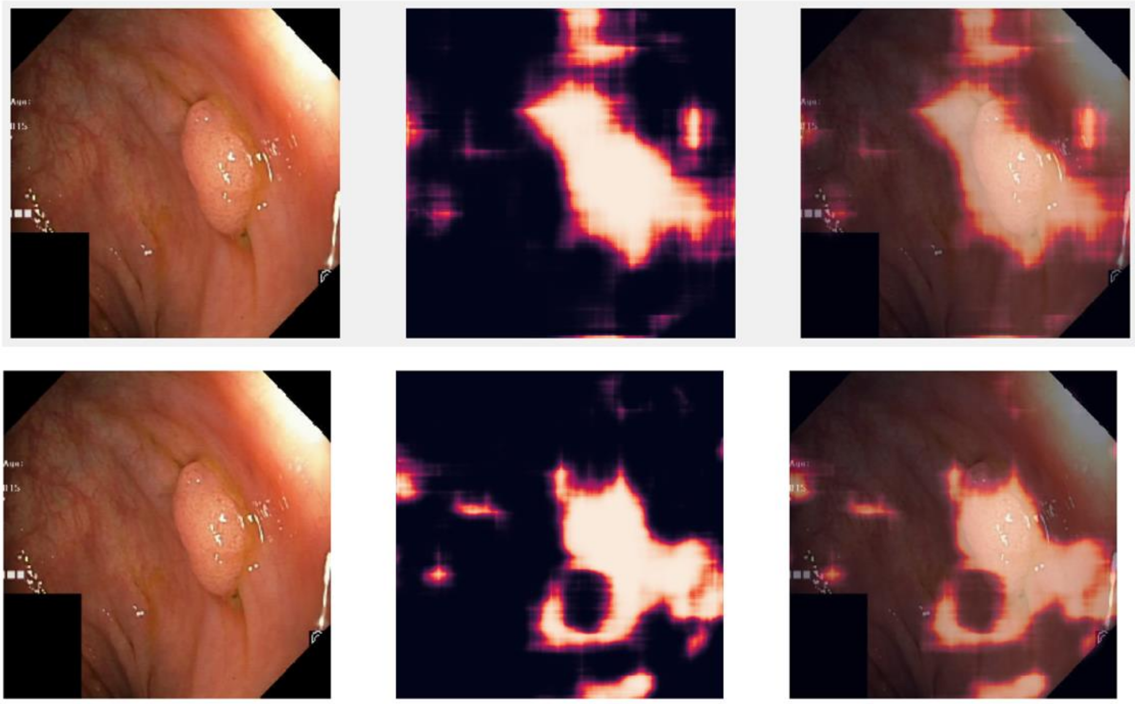


Model Metrics over Epochs:

120 epochs vs 250 epochs



250 epochs vs 400 epochs



loss: 0.0195 - accuracy: 0.9908 - iou_coef: 0.9340 - dice_coef: 0.9631

11. REFERENCES

Journal References:

[1] IEEE Xplore Full-Text PDF: (n.d.). https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=10149321

[2] Pernas, F. J. D., Martínez-Zarzuela, M., Antón-Rodríguez, M., & González-Ortega, D. (2021, February 2). A Deep Learning Approach for Segmentation Using a Multiscale Convolutional Neural Network. Healthcare. https://doi.org/10.3390/healthcare9020153

[3] Athisayamani, S., Antonyswamy, R. S., Velliangiri, S., Almeshari, M., Alzamil, Y., & Ravi, V. (2023, February 10). Feature Extraction Using a Residual Deep Convolutional Neural Network and Optimized Feature Dimension Reduction https://doi.org/10.3390/diagnostics13040668

[4] *Rasool, M., Ismail, N. A., Boulila, W., Ammar, A., Samma, H., Yafooz, W. M. S., & Emara, A. H. M. (2022, June 8). A Hybrid Deep Learning. https://doi.org/10.3390/e24060799*

[5] *Alrashedy, H. H. N., Almansour, A. F., Ibrahim, D. M., & Hammoudeh, M. A. A. (2022, June 6). Architectures and CNN Models. Sensors. https://doi.org/10.3390/s22114297*