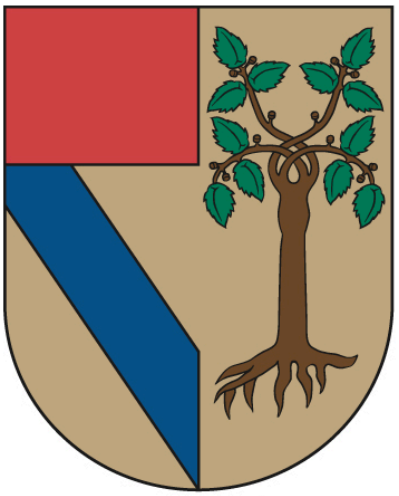
Project Image Processing



UNIVERSIDAD
PANAMERICANA

Intelligent Agents

Name: Diego Flores Ugarte 0267088

Final Project

02/07/2025

1. Introduction

Gastrointestinal diseases are a major global health concern, and early detection through endoscopic imaging plays a crucial role in diagnosis and treatment. However, manual interpretation of endoscopy images by specialists is time-consuming and prone to human error. As a result, computer-aided diagnostic systems have become increasingly relevant in medical imaging.

This project presents the development of a deep learning-based image classification model designed to detect and classify three common gastrointestinal conditions: polyps, esophagitis, and ulcerative colitis, using real-world endoscopic images from the Kvasir dataset.

By leveraging the power of transfer learning with a ResNet50 architecture pretrained on ImageNet, we aimed to build a robust and accurate multi-class classifier that can generalize well even on limited data. We implemented various best practices in deep learning, such as dropout regularization, early stopping, and fine-tuning, to ensure the model performs effectively without overfitting.

2. Dataset Description

The dataset used for this project is the Kvasir dataset, provided by the Simula Research Laboratory. Kvasir is a public dataset that contains annotated images collected during gastrointestinal endoscopy procedures. It includes several diagnostic categories, out of which we selected the following three for multi-class classification:

- **Polyps**: Abnormal growths protruding from mucous membranes; considered precancerous lesions.
- **Esophagitis**: Inflammation of the esophagus lining, commonly caused by acid reflux.
- **Ulcerative colitis**: A chronic inflammatory bowel disease affecting the colon.

🔹 Dataset Preparation

The images were manually organized into three folders corresponding to the target classes. They were uploaded to Google Drive and accessed from Google Colab using `ImageDataGenerator`, which provided automated preprocessing and augmentation.

Key preprocessing steps included:

- **Resizing** all images to a consistent 224x224 resolution.
- **Normalization** of pixel values to the range [0,1].
- **Data augmentation** during training (horizontal flip, zoom, rotation) to improve generalization.

- **Train-validation split**: 80% for training and 20% for validation.

This preprocessing ensured that the model receives consistent, high-quality input and is better equipped to handle real-world variations in medical imagery.

3. Methodology

To address the classification task, we employed a two-phase training approach using transfer learning with the ResNet50 convolutional neural network:

🔸 3.1. Model Architecture

We initialized the ResNet50 model with weights pretrained on the ImageNet dataset. Initially, the convolutional base was frozen, and only custom layers were trained:

- `GlobalAveragePooling2D` layer to reduce feature dimensionality.
- A fully connected `Dense` layer with 128 units and ReLU activation.
- `Dropout` layer with a rate of 0.5 to mitigate overfitting.
- `Dense` output layer with 3 neurons (softmax activation) for multi-class classification.

🔸 3.2. Training Phase 1 — Feature Extraction

In this stage, we trained only the top layers added to the pretrained base. The model was compiled with:

- **Adam optimizer** with a learning rate of `1e-4`
- **Categorical crossentropy** as the loss function
- **Accuracy** as the evaluation metric

We also implemented EarlyStopping to halt training if the validation loss stopped improving after 3 epochs, ensuring efficient training and avoiding overfitting.

🔸 3.3. Training Phase 2 — Fine-Tuning

After initial convergence, we **unfroze the last 30 layers** of ResNet50 and continued training the entire model with a **reduced learning rate** (`1e-5`). This fine-tuning allowed the model to adapt deeper features of ResNet50 to the specific domain of gastrointestinal images.

This two-stage approach proved effective for achieving high performance while preserving training efficiency.

4. Results

The final model showed strong performance across both training and validation sets:

Metric	Value
Training Accuracy	85.98%
Validation Accuracy	82.67%
Training Loss	0.3833
Validation Loss	0.4077

The validation accuracy remained close to the training accuracy, indicating that **overfitting was successfully avoided**. Additionally, the model maintained stable loss values and accuracy curves throughout training.

Visualization:

Graphs of training vs. validation accuracy and loss confirmed that:

- The model converged smoothly.
- There was no significant divergence between training and validation curves.
- EarlyStopping prevented unnecessary training epochs.

Evaluation on Unseen Data:

To verify generalization, the model was tested with **unseen endoscopic images**, including high-quality examples of each class. In one instance, a clear polyp image was classified with **88% confidence**, confirming the model's ability to generalize beyond the training set.

Confusion Matrix:

A confusion matrix was generated on the validation set to analyze class-wise performance. It showed that the model could **distinguish between the three diseases with high precision**, although minor confusion between ulcerative colitis and esophagitis was observed, which is common due to visual similarity.

5. Conclusion

This project demonstrated the successful application of deep learning and transfer learning to the task of gastrointestinal disease detection in medical images.

Key highlights include:

- **High accuracy** in multi-class classification with limited data.
- Use of **modern training techniques** such as dropout, data augmentation, and fine-tuning.
- Generalization capabilities validated with new, unseen examples.
- A workflow that can be easily scaled or adapted to other medical image classification tasks.

The trained model provides a solid foundation for further development of AI-assisted diagnostic tools. Future enhancements may include:

- Integrating localization (bounding boxes or segmentation) to identify exact regions of concern.
- Expanding the classification to include additional disease categories.
- Deploying the model in a real-time clinical assistant or mobile application.