Atzin Cruz

Classification of Gastrointestinal Endoscopy

A little introduction

Why is this important?

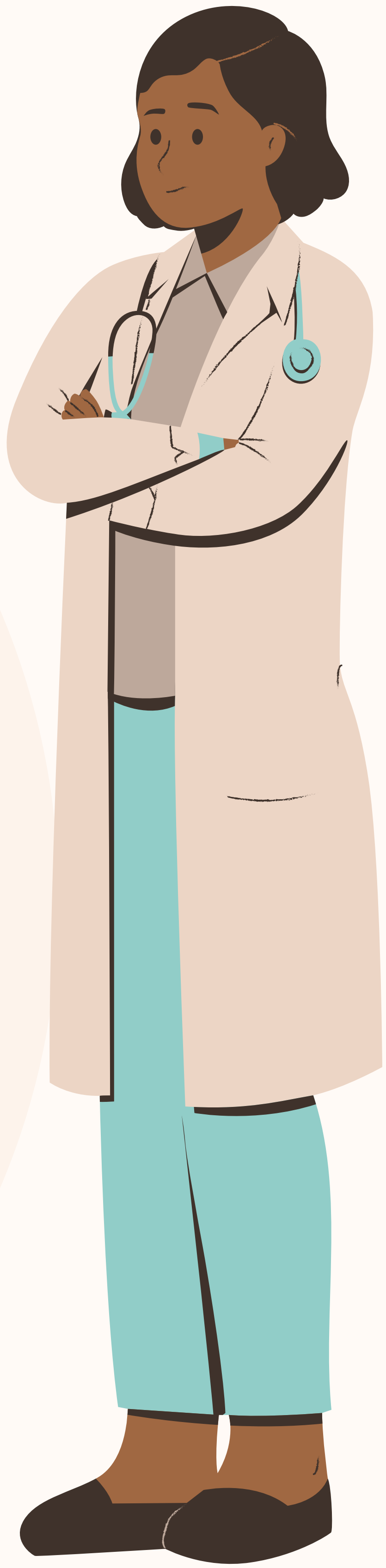
Gastrointestinal (GI) endoscopy is critical for diagnosing GI conditions

The dataset

Objective: Classify endoscopy images using transfer learning with ResNet50. Dataset: Kvasir v2, a multi-class dataset for GI disease detection.

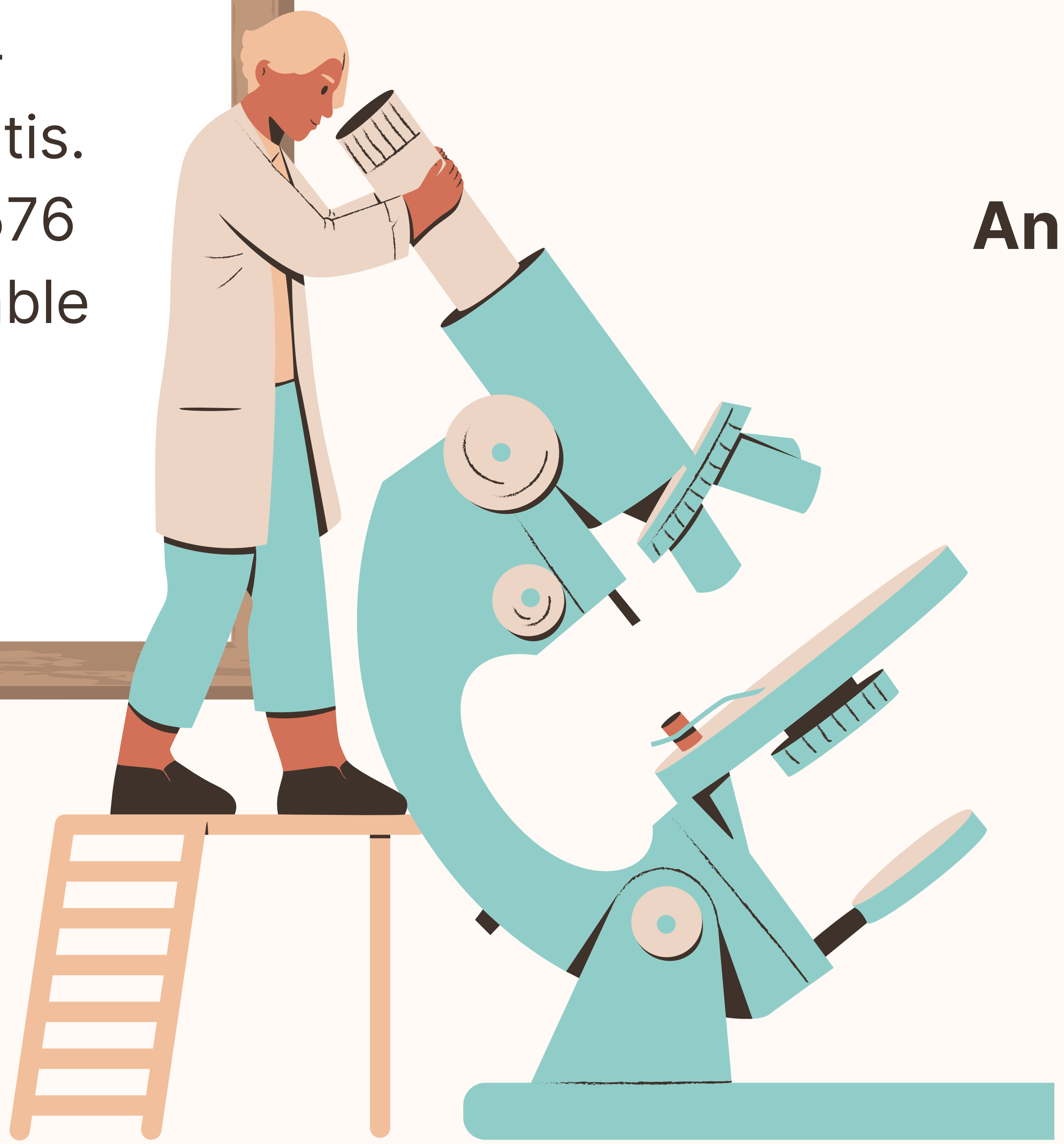
Goal

Achieve high accuracy for clinical applicability.



THE DATASET

- Contains 8,000 high-resolution RGB endoscopy images.
- Eight classes: dyed-lifted-polyps, dyed-resection-margins, esophagitis, normal-cecum, normal-pylorus, normal-z-line, polyps, ulcerative-colitis.
- Approximately 1,000 images per class, 720x576 pixels. Annotated by medical experts for reliable ground truth



An exploration of innovations in medicine

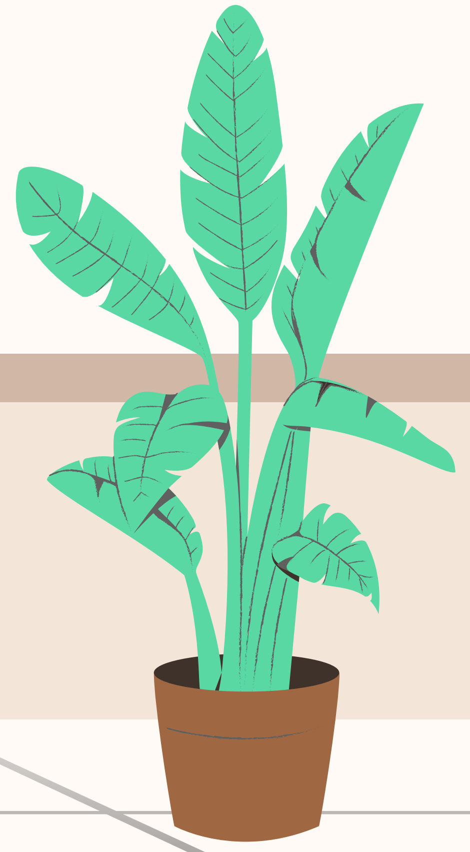
The methodology

Model and evaluation metrics

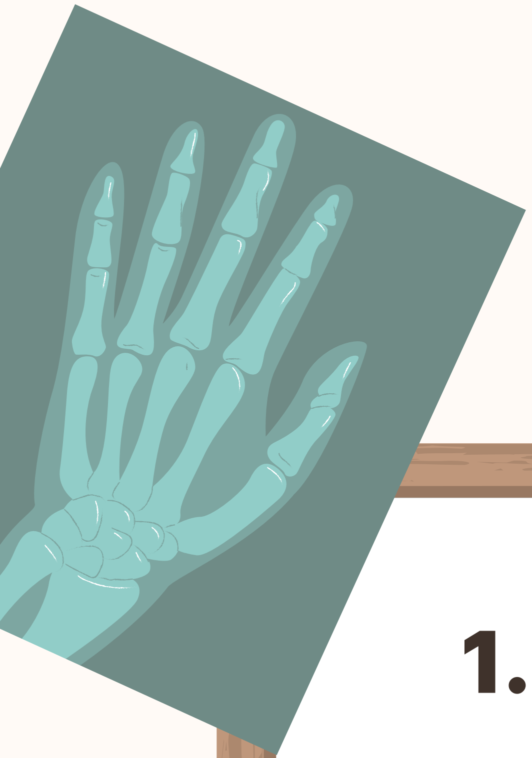
- ResNet50, pre-trained on ImageNet, fine-tuned on Kvasir v2.
- Accuracy, precision, recall, F1-score, specificity, ROC-AUC.

Training procedure

- Two-stage training: Initial feature extraction, then fine-tuning.
- Advanced data augmentation (rotation, flips, color jitter).
- Optimizer: Adam; Loss: Categorical cross-entropy.



Why resnet50?



1. Residual Learning Solves Vanishing Gradients

- Deep networks often suffer from vanishing/exploding gradients.
- ResNet50 uses residual connections (skip connections):
- $F(x)=H(x)+xF(x) = H(x) + xF(x)=H(x)+x$
- This allows gradients to flow directly through identity mappings, making training more stable, even in very deep networks

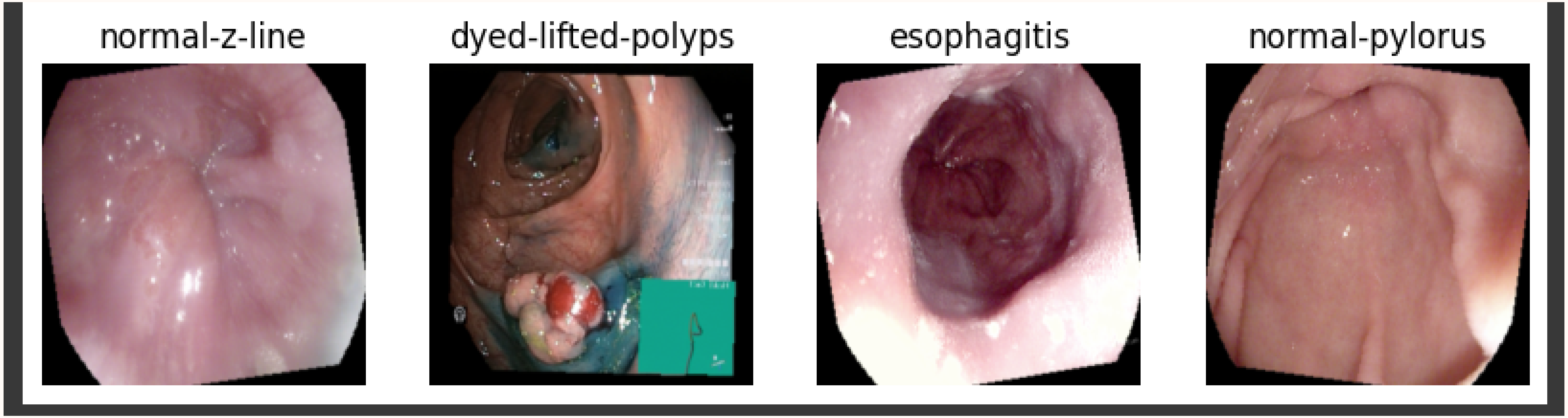
2. Depth + Efficiency Balance

- ResNet50 has 50 layers, providing enough depth to extract high-level and low-level features.
- It is deep enough to capture complex medical image patterns, but not as computationally heavy as deeper versions like ResNet101 or ResNet152.

Data augmentation

- Random rotation (±30°)
- Horizontal & vertical flips (probability 0.5)

- Brightness & contrast adjustments (±0.2)
- Random cropping and resizing


normal-z-line
dyed-lifted-polyps
esophagitis
normal-pylorus

MODEL ARCHITECTURE

- ResNet50 (pre-trained on ImageNet)
- Replaces final layer with 8-class softmax output
- Leverages residual connections to prevent vanishing gradients:
- $F(x)=H(x)+x$

Training Procedure

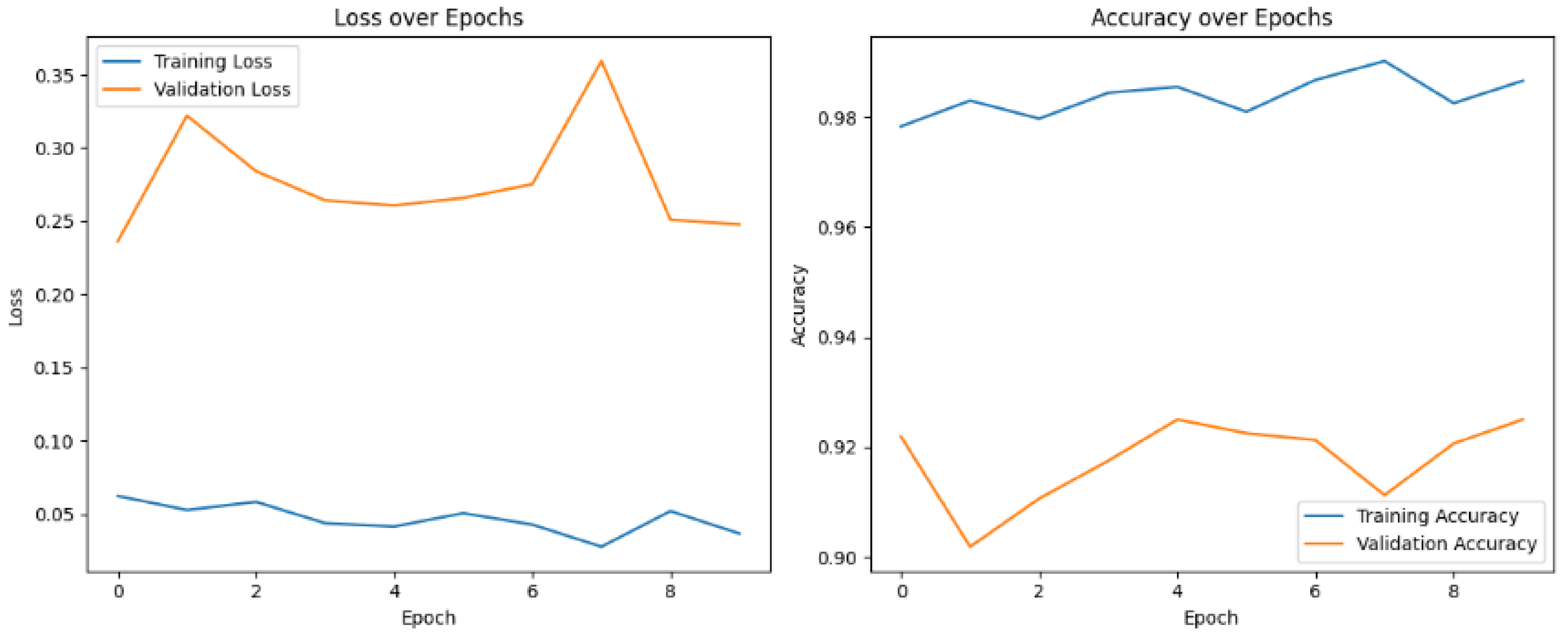
1. **Stage 1:** Freeze all weights → Train for 5 epochs
2. **Stage 2:** Unfreeze last 10 layers → Fine-tune for 5 epochs

- **Optimizer:** Adam, learning rate = 1e-4
- **Loss:** Categorical cross-entropy
- **Batch size**: 32
- Tracked with **TensorBoard**

RESULTS

- **Validation Accuracy**: 92.63%.
- **Per-Class Performance:**
  - **Strong**: normal-pylorus (F1-score: 0.8973, recall: 1.0000), normal-cecum (F1-score: 0.9858).
  - **Weak:** esophagitis (F1-score: 0.8130), normal-z-line (F1-score: 0.8244).
- **ROC-AUC:** Macro-average of 0.98; normal-pylorus: 1.00, esophagitis: 0.92.

Class	Precision	Recall	F1-Score	Specificity	Support
Dyed-lifted-polyps	0.9299	0.9803	0.9544	0.9893	203
Dyed-resection-margins	0.9836	0.9231	0.9524	0.9979	195
Esophagitis	0.9317	0.7212	0.8130	0.9921	208
Normal-cecum	0.9615	0.9756	0.9685	0.9943	205
Normal-pylorus	0.9760	1.0000	0.9878	0.9964	203
Normal-z-line	0.7412	0.9389	0.8284	0.9585	180
Polyps	0.9797	0.9234	0.9507	0.9971	209
Ulcerative-colitis	0.9353	0.9543	0.9447	0.9907	197
Macro Avg	0.9300	0.9269	0.9250	0.9895	1600
Weighted Avg	0.9308	0.9263	0.9250	–	1600

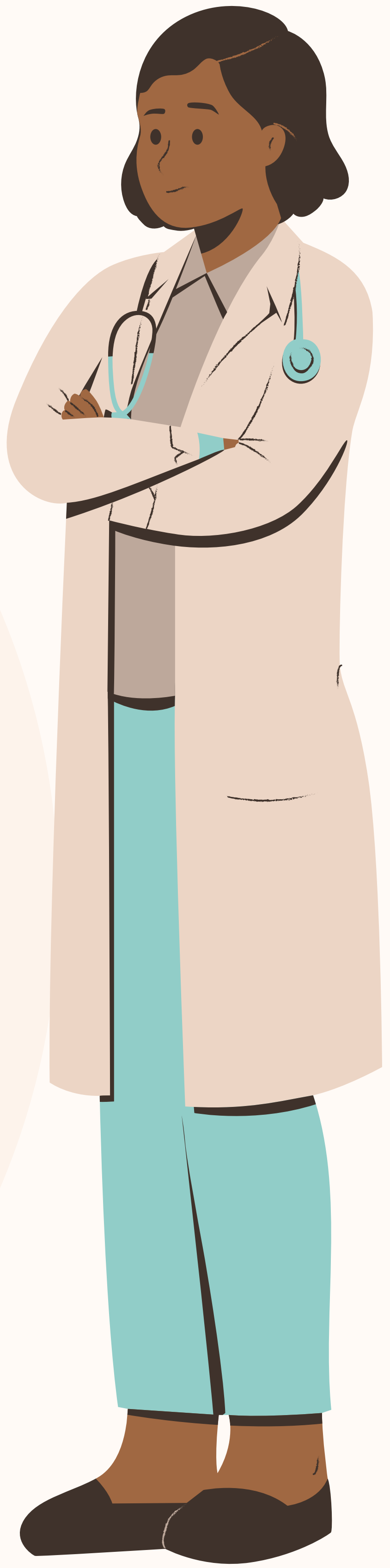

Loss over Epochs
Training Loss
Validation Loss
Loss
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0
2
4
6
8
Epoch
Accuracy over Epochs
Accuracy
0.98
0.96
0.94
0.92
0.90
Training Accuracy
Validation Accuracy
0
2
4
6
8
Epoch

	DLP	DRM	ESO	NC	NP	NZL	POL	UC
DLP	199	3	0	0	0	0	0	1
DRM	15	180	0	0	0	0	0	0
ESO	0	0	150	0	0	58	0	0
NC	0	0	0	200	0	0	2	3
NP	0	0	0	0	203	0	0	0
NZL	0	0	11	0	0	169	0	0
POL	0	0	0	3	3	1	193	9
UC	0	0	0	5	2	0	2	188

Legend: DLP: Dyed-lifted-polyps, DRM: Dyed-resection-margins, ESO: Esophagitis, NC: Normal-cecum, NP: Normal-pylorus, NZL: Normal-z-line, POL: Polyps, UC: Ulcerative-colitis.

CONCLUSION

- ResNet50 achieves 92.63% accuracy on Kvasir v2, suitable for clinical use.
- Challenges: Esophagitis and normal-z-line need better feature discrimination.
- Visualizations and ROC-AUC enhance model interpretability.
- Future Work:
  - Advanced augmentations (CutMix, MixUp).
  - Hybrid CNN-ViT architectures.
  - Ensemble methods and clinical validation.



Thank you for your attention