Multiclass Classification and **Semantic Segmentation** of **Colorectal Cancer Cells** from **Histopathology Images**

Advanced Image Analysis | Machine Learning | Deep Learning
2023/2024

By
Mahdi Islam
Musarrat Tabassum
Sirada Kittipaisarnkul
Sarita Mourya
Tuhinangshu Gangopadhyay

Dataset

- Publicly available EBHI-Seg Dataset containing 2,228 colorectal histopathology images and their corresponding images.
- This dataset has the following classes:
  - Normal: 76 images
  - Polyp: 474 images
  - Low-grade IN: 637 images
  - High-grade IN: 186 images
  - Adenocarcinoma: 795 images
  - Serrated Adenoma: 58 images

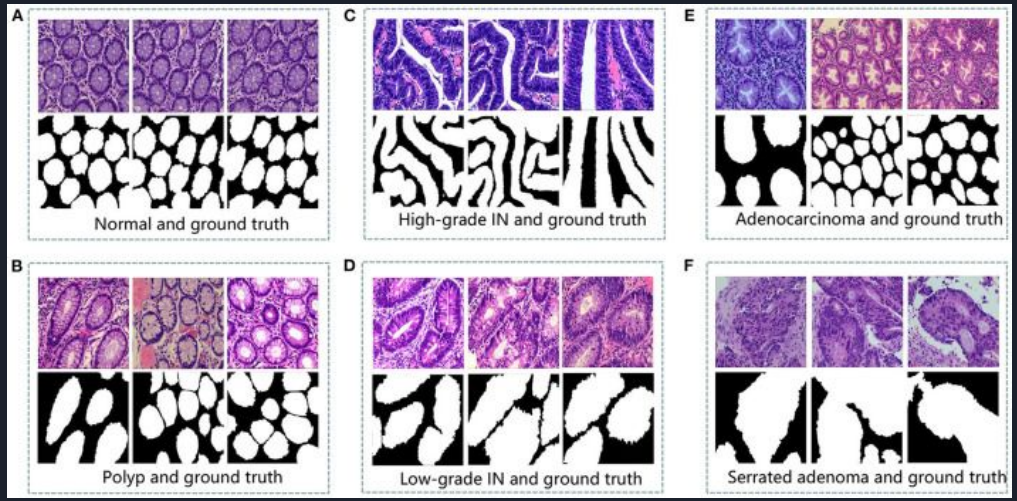


Image Processing based Segmentation

The general image Processing based Segmentation pipeline has the following steps involved:

- Image Preparation Pipeline:
  - Conversion to Gray-scale
  - Inversion of Gray-scale image
  - Grayscale Morphology (Erosion)
  - Smoothing

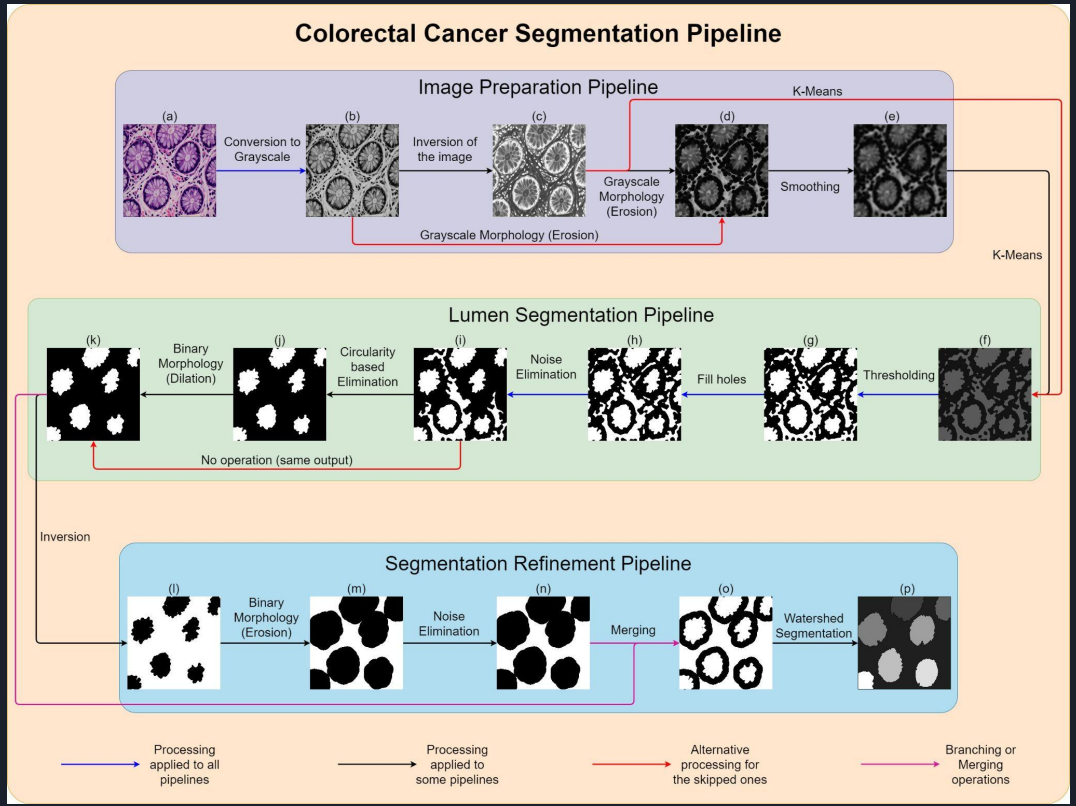
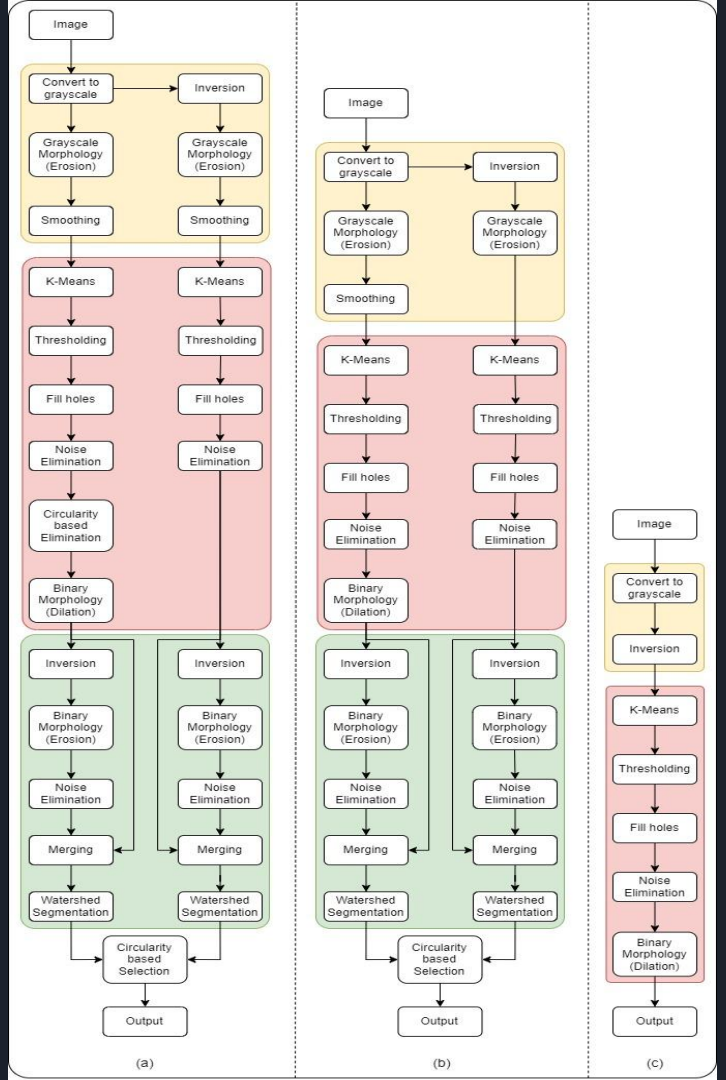


Image Processing based Segmentation (continued)

- Lumen Segmentation Pipeline:
  - K-Means
  - Thresholding
  - Filling Holes
  - Noise Elimination
  - Circularity based Elimination
  - Binary Morphology (Dilation)
- Segmentation Refinement Pipeline
  - Inversion
  - Binary Morphology (Erosion)
  - Noise Elimination
  - Merging
  - Watershed Segmentation

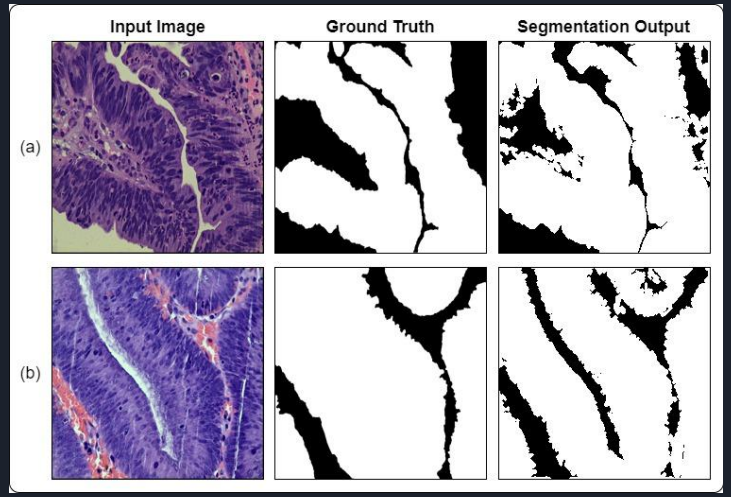
This general pipeline was further updated class wise to capture the specific patterns observed in the image.

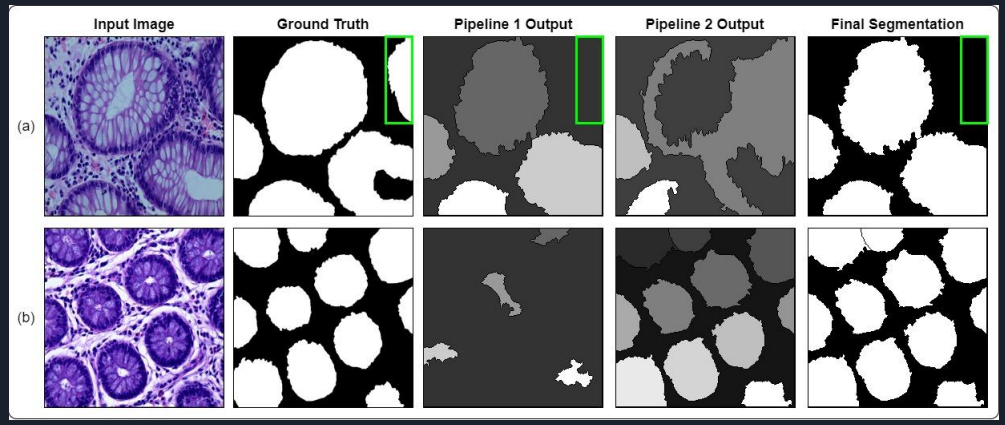


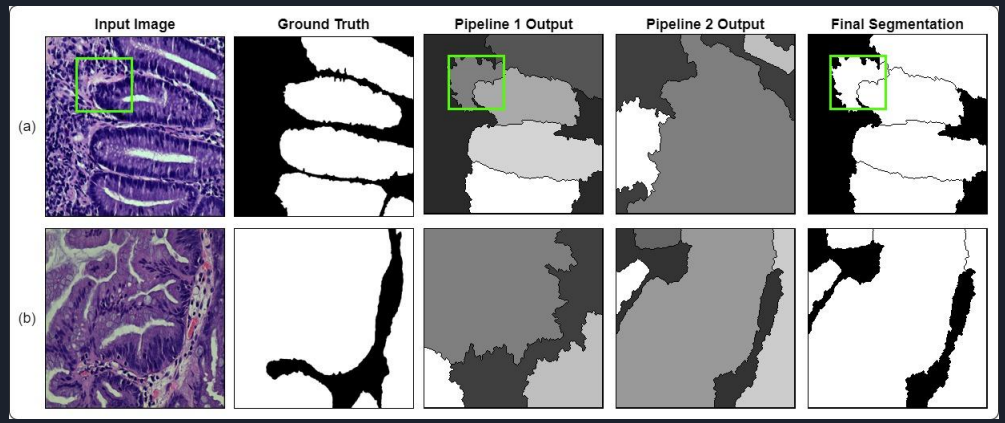
Segmentation Result

	Precision (P)	Recall (R)	Jaccard Similarity (J)	Dice Score (D)	Accuracy (A)
Normal	0.90	0.78	0.73	0.83	0.81
Polyp	0.89	0.78	0.72	0.83	0.82
Low-grade IN	0.86	0.71	0.63	0.76	0.72
High-grade IN	0.85	0.93	0.80	0.88	0.83
Adenocarcinoma	0.90	0.82	0.74	0.84	0.79
Serrated Adenoma	0.78	0.85	0.67	0.80	0.73
Combined Performance over all classes	**0.86**	**0.81**	**0.71**	**0.82**	**0.78**

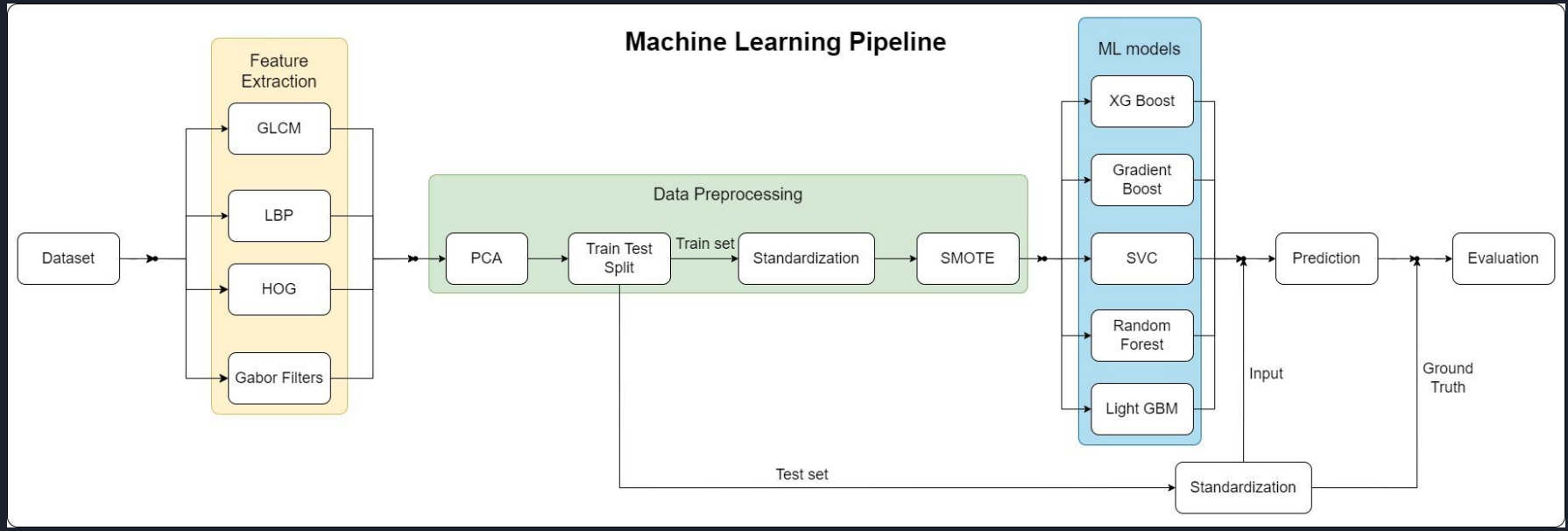
Segmentation Result







Machine Learning for Classification



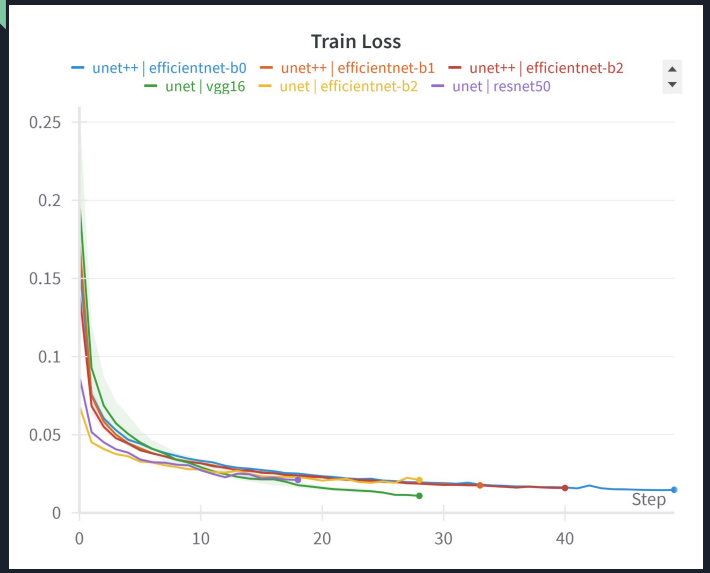
Classification Result

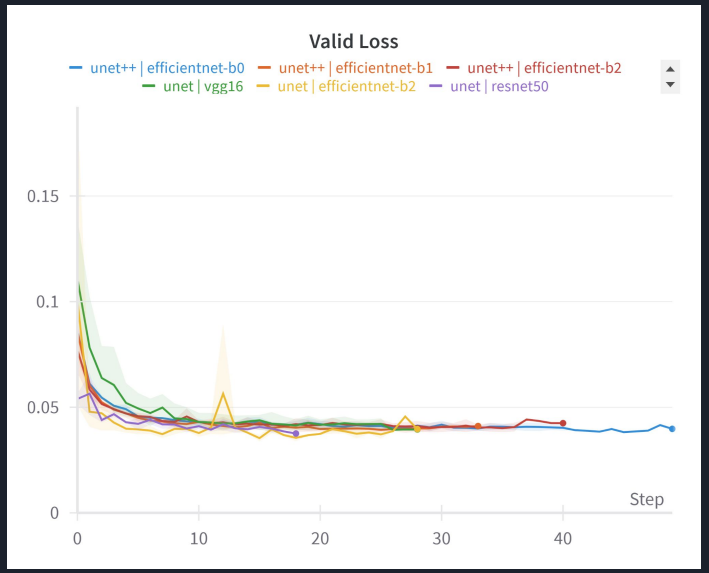
Features	Number of Features	Models	Precision (P)	Recall (R)	Accuracy (A)
LBP1 + LBP2 + LBP3	36	XGB GB SVC RF LGBM	0.73 0.72 0.76 0.70 0.75	0.74 0.73 0.77 0.71 0.75	0.74 0.73 0.77 0.71 0.75
LBP1 + LBP2 + LBP3 + HOG1 + PCA (n=400)	436	XGB GB SVC RF LGBM	0.67 0.69 0.56 0.63 0.70	0.71 0.72 0.64 0.64 0.71	0.71 0.72 0.64 0.64 0.71
LBP1 + LBP2 + LBP3 + HOG2 + PCA (n=400)	436	XGB GB SVC RF LGBM	0.70 0.69 0.60 0.61 0.71	0.73 0.72 0.64 0.64 0.73	0.73 0.72 0.64 0.64 0.73
LBP1 + LBP2 + LBP3 + GLCM + Gabor filters	**252**	**XGB GB SVC RF LGBM**	**0.83 0.79 0.82 0.79 0.81**	**0.83 0.80 0.81 0.79 0.82**	**0.83 0.80 0.81 0.79 0.82**
LBP1 + LBP2 + LBP3 + GLCM + Gabor filters + SMOTE	252	XGB GB SVC RF LGBM	0.82 0.82 0.82 0.78 0.81	0.81 0.82 0.82 0.78 0.81	0.81 0.82 0.82 0.78 0.81

Deep Learning for segmentation: Pipeline

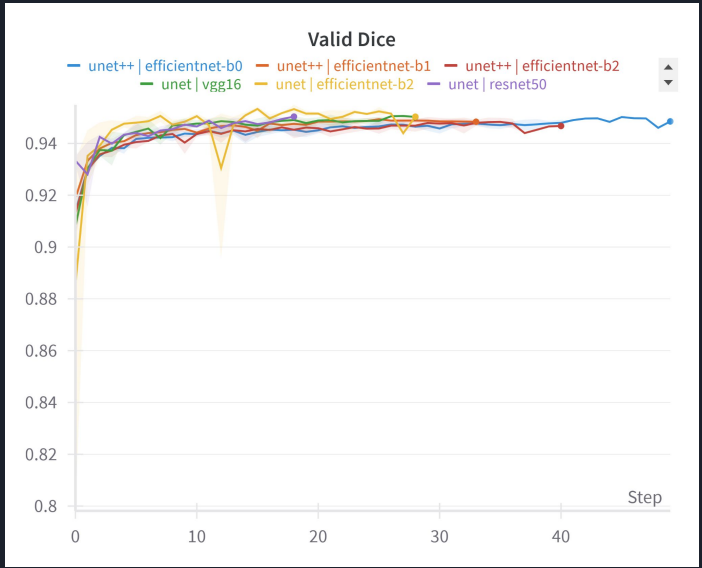
- Dataset Preparation Pipeline:
  - Image Normalization
  - Mask Thresholding
  - Image Augmentation (not used during inference)
- Segmentations Architectures
  - UNet: Backbones-
    - VGG16
    - Resnet50
    - EfficientNet-B2
  - UNet++: Backbones-
    - EfficientNet-B0
    - EfficientNet-B1
    - EfficientNet-B2
- Experiments:
  - Detailed Performance of DL models
  - Class wise performance for each model
  - Sample Image Visualization for checking output masks

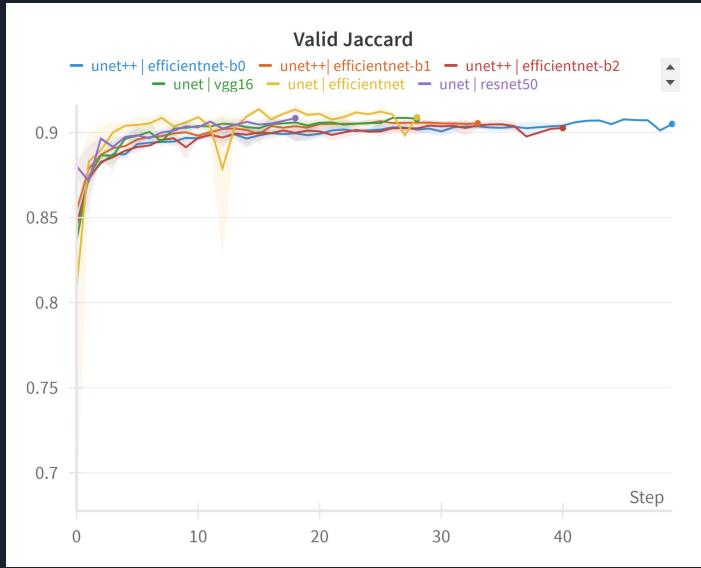
Deep Learning for segmentation:
Training Phase





Deep Learning for segmentation:
Training Phase cont.





Deep Learning for segmentation: Inference Phase

Detailed Performance of DL models

Models	Dice Score	Jaccard Similarity	Precision (P)	Recall (R)	Accuracy (A)
Baseline Unet Backbone: VGG16	0.9337	0.8921	0.9305	0.9397	0.9402
UNet Backbone: ResNet50	0.9353	0.8952	0.9308	0.9429	0.9421
UNet **Backbone: EfficientNet-B2**	**0.9372**	**0.8986**	**0.9323**	**0.9448**	**0.9448**
UNet++ Attention: scSE Backbone: EfficientNet-B0	0.9328	0.8909	0.9315	0.9377	0.9387
UNet++ Attention: scSE Backbone: EfficientNet-B1	0.9337	0.8926	0.9305	0.9408	0.9403
UNet++ Attention: scSE Backbone: EfficientNet-B2	0.9289	0.8846	0.9211	0.9417	0.9330

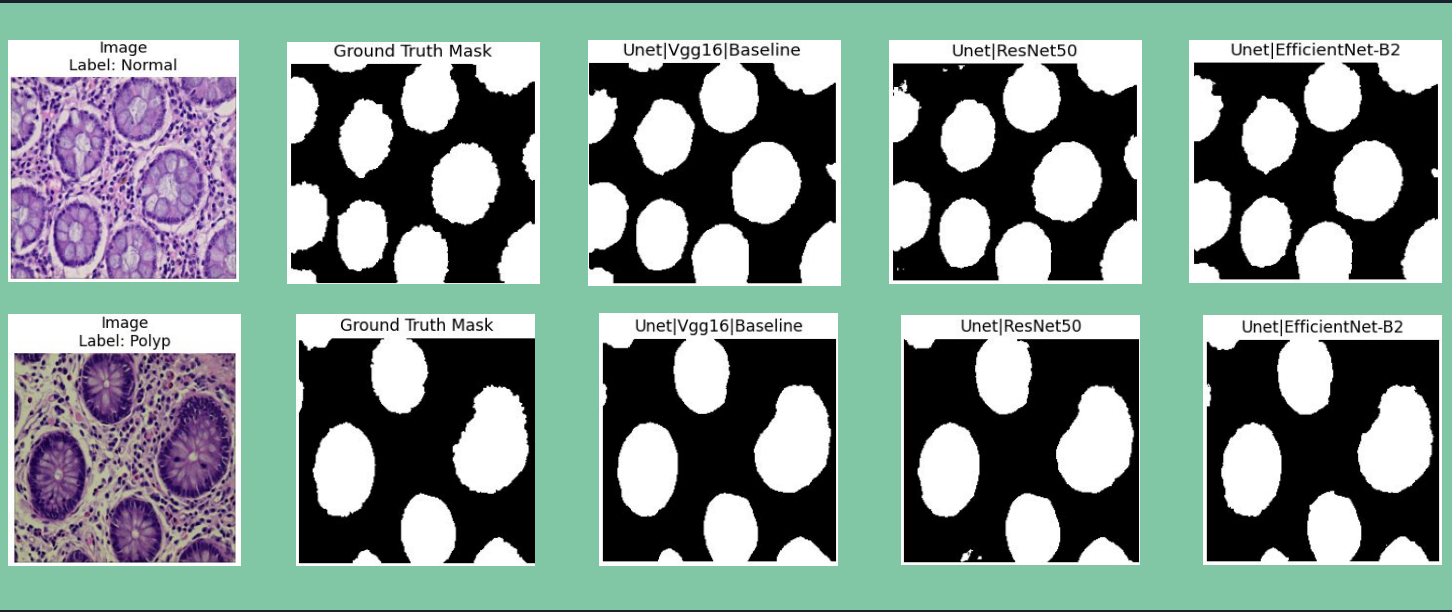
Deep Learning for segmentation: Inference Phase

Class wise performance of best model

	Dice Score	Jaccard Similarity	Precision (P)	Recall (R)	Accuracy (A)
Normal	0.9337	0.8921	0.9305	0.9397	0.9402
Polyp	0.9353	0.8952	0.9308	0.9429	0.9421
Low-grade IN	0.9372	0.8986	0.9323	0.9448	0.9448
High-grade IN	0.9328	0.8909	0.9315	0.9377	0.9387
Adenocarcinoma	0.9337	0.8926	0.9305	0.9408	0.9403
Serrated Adenoma	0.9289	0.8846	0.9211	0.9417	0.9330

Deep Learning for segmentation: Inference Phase

Sample Segmentation Outputs



THANK YOU