

## Data Collection and Preprocessing Phase

Date	9 July 2024
Team ID	SWTID1720043892
Project Title	WCE Curated Colon Disease Using Deep Learning
Maximum Marks	6 Marks

### Preprocessing Template

The images will be preprocessed by resizing, normalizing, augmenting, denoising, adjusting contrast, detecting edges, converting color space, cropping, batch normalizing, and whitening data. These steps will enhance data quality, promote model generalization, and improve convergence during neural network training, ensuring robust and efficient performance across various computer vision tasks.

Section	Description
Data Overview	The dataset used in this project is the Curated Colon Disease Dataset for Deep Learning, which consists of 5000 colonoscopy images divided into four classes: Adenoma, Hyperplastic, Inflammatory, and Normal. The images are in JPEG format and have varying sizes.
Resizing	Resize images to a target size of 224x224 pixels to ensure consistent input sizes for the neural network.
Normalization	Normalize pixel values to the range [0, 1].
Data Augmentation	<p>Normalize pixel values to a specific range.</p> <p><b>Data Augmentation</b> Apply augmentation techniques such as flipping, rotation, shifting, zooming, or shearing. Ex: <u>Image before augmenting:</u></p>

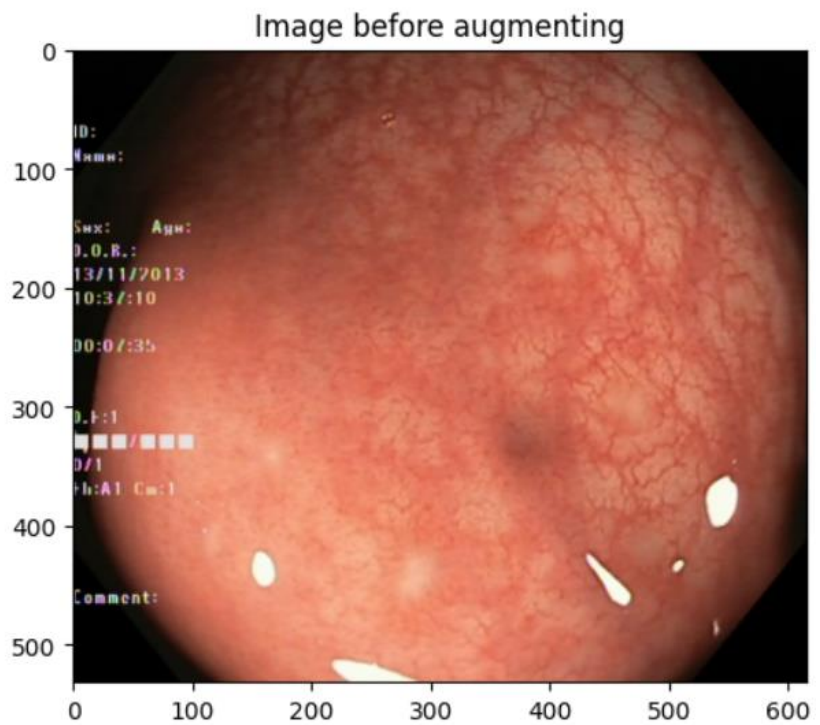
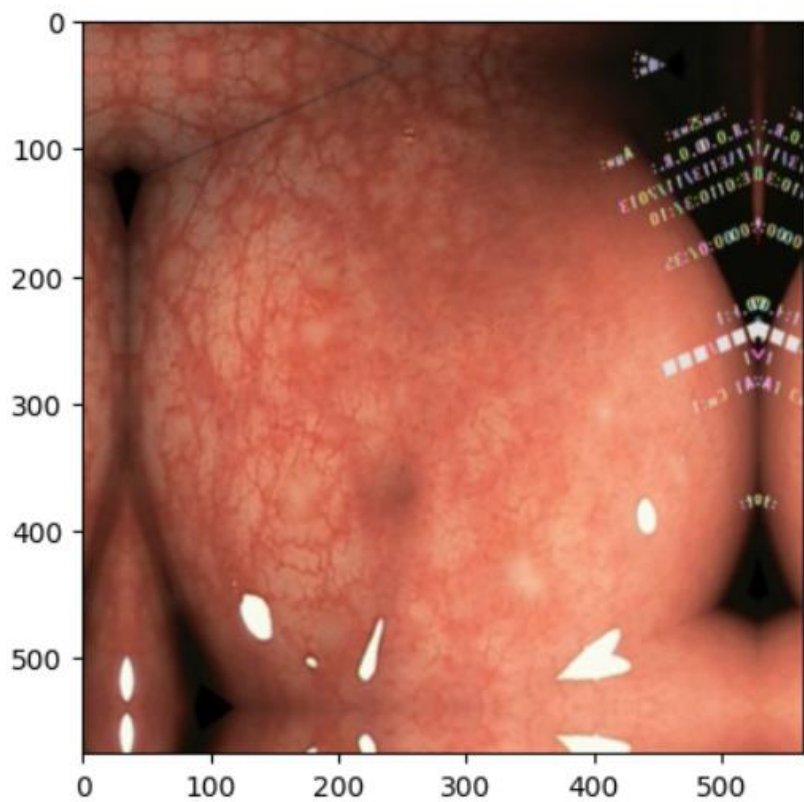
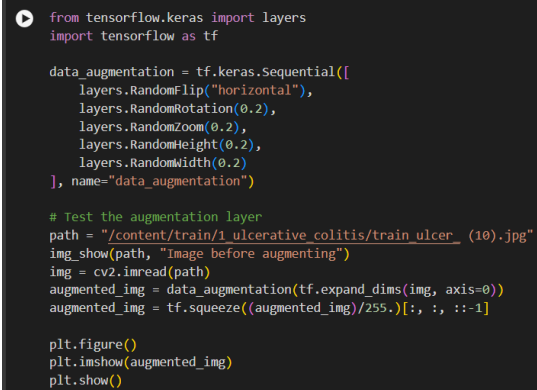


Image after augmenting:



Denoising	<p>Apply denoising filters to reduce noise in the images.</p> <p>N/A</p>
Edge Detection	<p>Apply edge detection algorithms to highlight prominent edges in the images.</p> <p>N/A</p>
Color Space Conversion	<p>Convert images from one color space to another.</p> <p>N/A (Images are already in RGB format)</p>
Image Cropping	<p>Crop images to focus on the regions containing objects of interest.</p> <p>N/A (Images in the dataset have already been cropped).</p>
Batch Normalization	<p>Apply batch normalization to the input of each layer in the neural network.</p>
<b>Data Preprocessing Code Screenshots</b>	
Loading Data	<p>Downloading from Kaggle:</p> <div data-bbox="599 1150 1421 1476"> <p>✓ Set Up Kaggle API and Download Dataset</p> <pre># Create the .kaggle directory and copy the kaggle.json file !mkdir -p ~/.kaggle !cp kaggle.json ~/.kaggle/  # Download the cataract dataset from Kaggle !kaggle datasets download -d francismon/curated-colon-dataset-for-deep-learning  # Unzip the downloaded dataset !unzip curated-colon-dataset-for-deep-learning.zip -d /content/</pre> </div> <p>Loading data from train ,test and validation directories:</p> <div data-bbox="599 1583 1206 1717"> <pre>import tensorflow as tf from tensorflow.keras.preprocessing.image import ImageDataGenerator train_dir = '/content/train' test_dir = '/content/test' val_dir = '/content/val'</pre> </div>
Resizing	<div data-bbox="599 1753 1421 1837"> <pre>train = train_datagen.flow_from_directory(trainPath,target_size =(224,224),batch_size=16,class_mode='categorical') test = test_datagen.flow_from_directory(testPath,target_size =(224,224),batch_size=16,class_mode='categorical') val = val_datagen.flow_from_directory(valpath,target_size =(224,224),batch_size=16,class_mode='categorical')</pre> </div>

Normalization	Give the code snippet as an image (copy and paste the picture in this block).
Data Augmentation	<div><div>▼ Data Augmentation</div><div><pre>from tensorflow.keras import layers import tensorflow as tf  data_augmentation = tf.keras.Sequential([     layers.RandomFlip("horizontal"),     layers.RandomRotation(0.2),     layers.RandomZoom(0.2),     layers.RandomHeight(0.2),     layers.RandomWidth(0.2) ], name="data_augmentation")  # Test the augmentation layer path = "/content/train/1 ulcerative colitis/train ulcer_ (10).jpg" img_show(path, "Image before augmenting") img = cv2.imread(path) augmented_img = data_augmentation(tf.expand_dims(img, axis=0)) augmented_img = tf.squeeze((augmented_img)/255.)[: , :, :-1]  plt.figure() plt.imshow(augmented_img) plt.show()</pre></div></div>
Denoising	Give the code snippet as an image (copy and paste the picture in this block).  N/A
Edge Detection	Give the code snippet as an image (copy and paste the picture in this block).  N/A
Color Space Conversion	Give the code snippet as an image (copy and paste the picture in this block).  N/A
Image Cropping	Give the code snippet as an image (copy and paste the picture in this block).  N/A
Batch Normalization	Give the code snippet as an image (copy and paste the picture in this block).