Data Collection and Preprocessing Phase

Date	15 July 2024
Team ID	SWTID1720418653
Project Title	Crystal Clear Vision: Revolutionizing Cataract Prediction through Transfer Learning Mastery
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 is taken from kagale. It consists images of eyes, specifically focusing on those with cataract conditions. Each image is labeled to indicate the presence or absence of cataracts, making it suitable for training machine learning models for classification tasks. The images are organized in folders with separate directories for cataract and non-cataract images. This dataset is primarily used for developing automated cataract detection systems to assist in medical diagnosis.
Resizing	In the project, images are resized to 224x224 pixels using the 'ImageDataGenerator' class from Keras.
Normalization	The project normalizes pixel values to the range [0, 1] using the 'rescale' parameter in 'ImageDataGenerator'.
Data Augmentation	The project uses data augmentation techniques like shear range, zoom range, and horizontal flip.
Denoising	-

Edge Detection	-
Color Space Conversion	-
Image Cropping	-.
Batch Normalization	-
Data Preprocessing Code Screenshots	

Loading Data	`!chmod 600 ~/.kaggle/kaggle.json`
	`!kaggle datasets download -d hemooredaoo/cataract`
	
	`chmod: cannot access '/root/.kaggle/kaggle.json': No such file or directory`
	`Dataset URL: https://www.kaggle.com/datasets/hemooredaoo/cataract`
	`License(s): unknown`
	`Downloading cataract.zip to /content`
	`100% 551M/552M [00:08<00:00, 91.8MB/s]`
	`100% 552M/552M [00:08<00:00, 71.1MB/s]`
	
	`!unzip '/content/cataract.zip'`
	
	`inflating: cataract-image-dataset/processed_images/train/normal/image_47.png`
	`inflating: cataract-image-dataset/processed_images/train/normal/image_48.png`
	`inflating: cataract-image-dataset/processed_images/train/normal/image_49.png`

Resizing	`training_set = train_datagen.flow_from_directory(`
	`directory=train_directory,`
	`target_size = (224, 224),`
	`batch_size = 32,`
	`class_mode='binary'`
	`)`
	
	`test_set = test_datagen.flow_from_directory(`
	`directory=test_directory,`
	`target_size = (224, 224),`
	`batch_size = 32,`
	`class_mode='binary'`
	`)`

Normalization	`train_datagen=ImageDataGenerator(rescale=1./255, zoom_range=0.2, rotation_range=48, fill_mode='nearest', horizontal_flip=True, shear_range=0.2)`
	
	`test_datagen=ImageDataGenerator(rescale=1./255)`

Data Augmentation	`train_datagen=ImageDataGenerator(rescale=1./255, zoom_range=0.2, rotation_range=48, fill_mode='nearest', horizontal_flip=True, shear_range=0.2)`
Denoising	-
Edge Detection	-
Color Space Conversion	-
Image Cropping	-
Batch Normalization	-