Data Collection and Preprocessing Phase

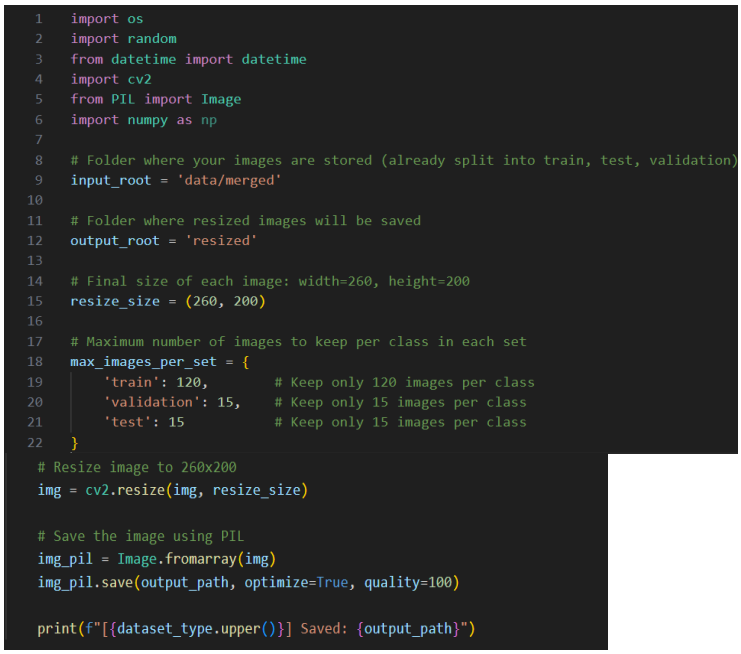
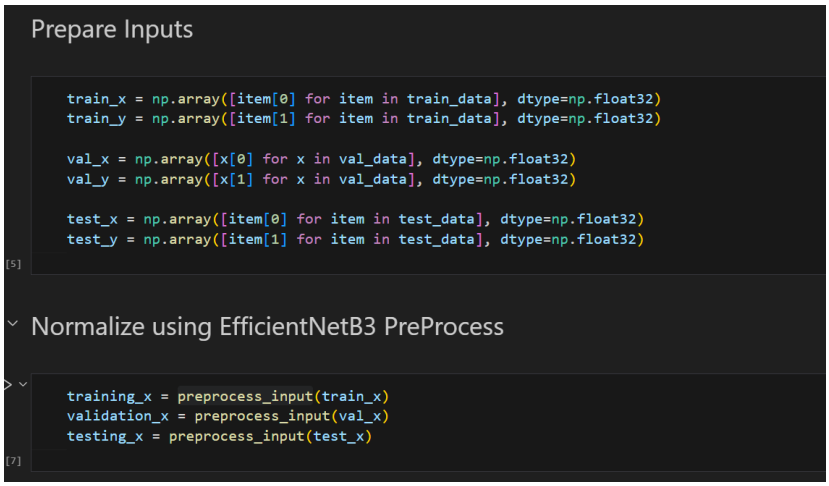
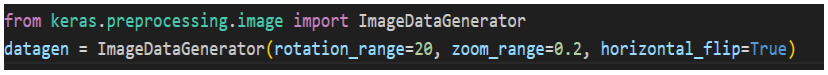
Date	19 June 2025
Team ID	SWTID1750170729
Project Title	**Deepfruitveg: Automated Fruit And Veg Identification**
Maximum Marks	6 Marks

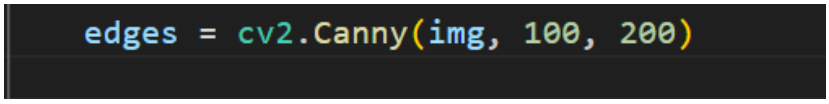
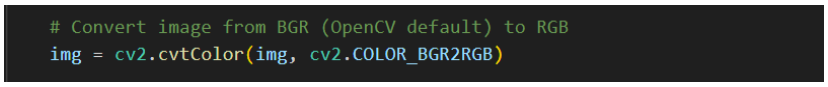
Preprocessing Template

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	This dataset contains images of fruits and vegetables across 30 categories. Each class has more than 150 images for training, and 100 images each for validation and testing. These images are resized and preprocessed before training a EfficientNetB3-based CNN model.
Resizing	Resize images to a fixed target size of (260, 200) or (224, 224) pixels depending on the model's input requirement.
Normalization	Normalize image pixel values using EfficientNetB3's preprocess_input() to scale them into the range expected by the pre-trained model.
Data Augmentation	Using Keras ImageDataGenerator to apply transformations like rotation, flipping, and zoom to increase data variability.
Denoising	Appling OpenCV's Non-Local Means denoising algorithm to reduce image noise before feeding into the model.

Edge Detection	Appling Canny edge detection to highlight object boundaries — useful for feature extraction or visualizations.
Color Space Conversion	Convert image from BGR (OpenCV default) to RGB for compatibility with TensorFlow/Keras.
Image Cropping	Crop images to focus on the regions containing objects of interest.
Batch Normalization	Apply batch normalization to the input of each layer in the neural network.
Data Preprocessing Code Screenshots	
Loading Data	[TRAIN] Saved: resized\train\aloevera\20250619-085057-069361.jpg [TRAIN] Saved: resized\train\aloevera\20250619-085057-351944.jpg [TRAIN] Saved: resized\train\aloevera\20250619-085057-650062.jpg [TRAIN] Saved: resized\train\aloevera\20250619-085058-407130.jpg [TRAIN] Saved: resized\train\aloevera\20250619-085059-081043.jpg [TRAIN] Saved: resized\train\aloevera\20250619-085059-365170.jpg [TRAIN] Saved: resized\train\aloevera\20250619-085059-657034.jpg [TRAIN] Saved: resized\train\aloevera\20250619-085059-941451.jpg [TRAIN] Saved: resized\train\aloevera\20250619-085100-254267.jpg [TRAIN] Saved: resized\train\aloevera\20250619-085100-546014.jpg [TRAIN] Saved: resized\train\aloevera\20250619-085101-449667.jpg [TRAIN] Saved: resized\train\aloevera\20250619-085102-177268.jpg [TRAIN] Saved: resized\train\aloevera\20250619-085102-454637.jpg [TRAIN] Saved: resized\train\aloevera\20250619-085102-738072.jpg [TRAIN] Saved: resized\train\aloevera\20250619-085103-031117.jpg [TRAIN] Saved: resized\train\aloevera\20250619-085103-321347.jpg [TRAIN] Saved: resized\train\aloevera\20250619-085103-579228.jpg [TRAIN] Saved: resized\train\aloevera\20250619-085103-863240.jpg [TRAIN] Saved: resized\train\aloevera\20250619-085104-127757.jpg [TRAIN] Saved: resized\train\aloevera\20250619-085104-410383.jpg [TRAIN] Saved: resized\train\aloevera\20250619-085105-098785.jpg [TRAIN] Saved: resized\train\aloevera\20250619-085105-798903.jpg [TRAIN] Saved: resized\train\aloevera\20250619-085106-071716.jpg [TRAIN] Saved: resized\train\aloevera\20250619-085106-352922.jpg [TRAIN] Saved: resized\train\aloevera\20250619-085107-542332.jpg ... [TEST] Saved: resized\test\watermelon\20250619-095102-235309.jpg [TEST] Saved: resized\test\watermelon\20250619-095102-831708.jpg [TEST] Saved: resized\test\watermelon\20250619-095103-101143.jpg [TEST] Saved: resized\test\watermelon\20250619-095103-364991.jpg

Resizing	
Normalization	
Data Augmentation	

Denoising	
Edge Detection	
Color Space Conversion	
Image Cropping	Not Needed As resized already Done
Batch Normalization	Not used explicitly, but EfficientNetB3 uses it internally.