**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 28 November 2024 |
| Team ID | 739996 |
| Project Title | Deep fruit veg: Automated Fruit And Veg  Identification |
| Maximum Marks | 6 Marks |

**Preprocessing Template**

The dataset consists of 30,000 images representing 30 classes of fruits and vegetables. Images are split into training (60%), validation (20%), and test (20%) sets. This diverse dataset ensures comprehensive learning, enabling accurate classification and robust generalization across various conditions.

|  |  |
| --- | --- |
| **Section** | **Description** |
| Data Overview | The dataset consists of 30,000 images representing 30 classes of fruits and vegetables. Images are split into training (60%), validation (20%), and test (20%) sets. This diverse dataset ensures comprehensive learning, enabling accurate classification and robust generalization across various conditions. |
| Resizing | * Resize images to a uniform size (e.g., 224x224) to standardize input for the neural network. * **Code Snippet (Screenshot Placeholder):** Include Python code using cv2.resize or TensorFlow's image. Resize. |
| Normalization | Normalize pixel values to the range [0, 1] to ensure numerical stability during model training.  **Code Snippet (Screenshot Placeholder):** Include code using libraries like NumPy or TensorFlow. |
| Data Augmentation | Apply augmentation techniques such as flipping, rotation, shifting, zooming, or shearing. |
| Denoising | Remove unwanted noise from images, which can improve feature extraction.  **Implementation:** Apply filters like Gaussian Blur or Bilateral Filter using OpenCV (cv2.GaussianBlur orcv2.bilateralFilter). |
| Edge Detection | Enhance object outlines to improve feature detection for classification.  Use algorithms like Canny Edge Detection (cv2.Canny) or Sobel filters. |
| Color Space Conversion | Convert images to grayscale or alternative color spaces (e.g., HSV) to highlight specific features or reduce complexity.  Use OpenCV (cv2.cvtColor) to convert between color spaces. |
| Image Cropping | Focus on the region of interest (fruits/vegetables) by removing unnecessary background information.  Automate cropping using bounding box detection or manually crop key regions. |
| Batch Normalization | Normalize input activations for each layer in the neural network, improving convergence and model stability.  Include batch normalization layers in the model architecture using Keras (Batch Normalization). |
| **Data Preprocessing Code Screenshots** | |
| Loading Data | sdir=r'/content/drive/MyDrive/ttv\_plants'  files=glob.glob(sdir+'/\*\*/\*.jpg',recursive=True)  print(len(files)) |
| Resizing |  |
| Normalization |  |
| Data Augmentation |  |
| Denoising |  |
| Edge Detection |  |
| Color Space Conversion |  |
| Image Cropping |  |
| Batch Normalization |  |