

Data Collection and Preprocessing Phase

Date	11 November 2024
Team ID	team-739757
Project Title	Tomato Plant Disease Detection From Leaf Images 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 consists of images of tomato plant leaves affected by various diseases, such as bacterial spots, late blight, and healthy leaves. These images vary in resolution, lighting conditions, and orientations. The data is collected from publicly available Kaggle datasets (e.g., Tomato leaf disease detection)
Resizing	All images will be resized to a uniform target size of $224 \times 224 \times 224$ pixels to maintain consistency and compatibility with the Resnet architecture.
Normalization	Pixel values will be normalized to a range of 0 to 1 or scaled using mean subtraction and standard deviation to improve training convergence and stability.
Data Augmentation	Apply augmentation techniques such as flipping, rotation, shifting, zooming, or shearing.
Denoising	Apply denoising filters like Gaussian or median filters will be applied to reduce noise in images, improving the clarity of disease-related patterns.

Edge Detection	Apply edge detection algorithms to highlight prominent edges in the images.
Color Space Conversion	Convert images from one color space to another.
Image Cropping	Crop images to focus on the regions containing objects of interest.
Batch Normalization	Batch normalization will be applied during training to normalize the activations within layers, improving convergence and generalization of the model.

Data Preprocessing Code Screenshots



Loading Data

```
import cv2
import glob
import matplotlib.pyplot as plt

# Load images from a directory
image_paths = glob.glob('/content/drive/MyDrive/dataset/tomato/train/Tomato_Early_blight/0034a')
images = [cv2.imread(img_path) for img_path in image_paths]

# Display first loaded image as a sample
if images:
    plt.imshow(cv2.cvtColor(images[0], cv2.COLOR_BGR2RGB))
    plt.axis('off')
    plt.show()
else:
    print("No images found in the specified directory.")
```



<p>Resizing</p>	<pre> resized_images = [cv2.resize(img, (640, 640)) for img in images] # Display a resized image as a sample plt.imshow(cv2.cvtColor(resized_images[0], cv2.COLOR_BGR2RGB)) plt.axis('off') plt.show() </pre> 
<p>Normalization</p>	<pre> normalized_images = [img / 255.0 for img in resized_images] # Display a normalized image as a sample plt.imshow(normalized_images[0]) plt.axis('off') plt.show() </pre> 

Data Augmentation

```
augmented_images = []
for img in resized_images:
    flipped_img = cv2.flip(img, 1) # Horizontal flip
    rotated_img = cv2.rotate(img, cv2.ROTATE_90_CLOCKWISE) # 90-degree rotation
    augmented_images.extend([flipped_img, rotated_img])
```

```
# Display an augmented image as a sample
plt.imshow(cv2.cvtColor(augmented_images[0], cv2.COLOR_BGR2RGB))
plt.axis('off')
plt.show()
```



Denoising

```
denoised_images = [cv2.GaussianBlur(img, (5, 5), 0) for img in resized_images]
```

```
# Display a denoised image as a sample
plt.imshow(cv2.cvtColor(denoised_images[0], cv2.COLOR_BGR2RGB))
plt.axis('off')
plt.show()
```



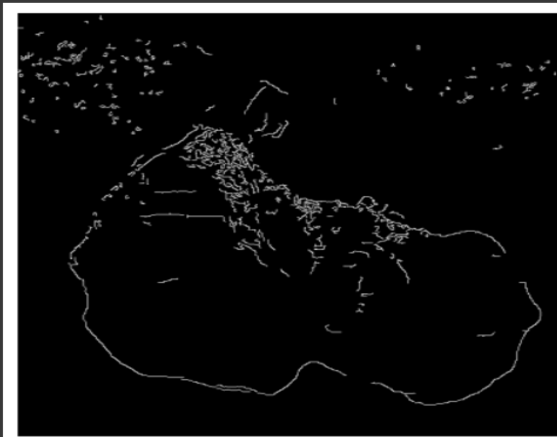
Edge Detection

```

edge_detected_images = [cv2.Canny(img, 100, 200) for img in resized_images]

# Display an edge-detected image as a sample
plt.imshow(edge_detected_images[0], cmap='gray')
plt.axis('off')
plt.show()

```



Color Space Conversion

```

grayscale_images = [cv2.cvtColor(img, cv2.COLOR_BGR2GRAY) for img in resized_images]

# Display a grayscale image as a sample
plt.imshow(grayscale_images[0], cmap='gray')
plt.axis('off')
plt.show()

```

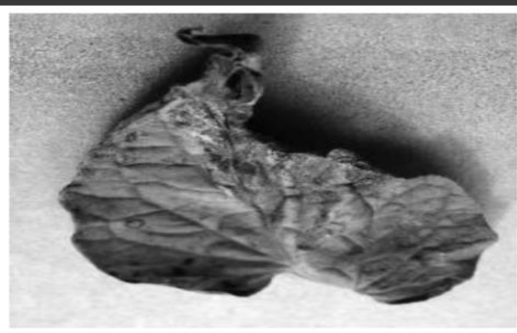


Image Cropping

```

cropped_images = [img[100:540, 100:540] for img in resized_images]

# Display a cropped image as a sample
plt.imshow(cv2.cvtColor(cropped_images[0], cv2.COLOR_BGR2RGB))
plt.axis('off')
plt.show()

```



Batch Normalization

```
import torch
import torch.nn as nn

# Example batch of images in tensor form
image_batch = torch.tensor([img.transpose(2, 0, 1) for img in normalized_images]).float()

# Apply batch normalization
batch_norm = nn.BatchNorm2d(3)
normalized_batch = batch_norm(image_batch) # Convert back to numpy for visualization
sample_image = normalized_batch[0].detach().numpy().transpose(1, 2, 0)

# Display a batch-normalized image
plt.imshow(sample_image)
plt.axis('off')
plt.show()
```

WARNING:matplotlib.image:Clipping input data to the valid range for imshow with RGB

