Cancer Detection Using Histopathological Images

Objective:
Learn the basics of image data preprocessing and training CNNs for cancer detection using histopathological images.

Dataset:
We used a subset of lung cancer histopathological images consisting of 250 Lung Cancer and 250 Normal images. Images were resized to 128x128 for CNN and 224x224 for VGG16. Dataset split: 80% training, 20% validation.

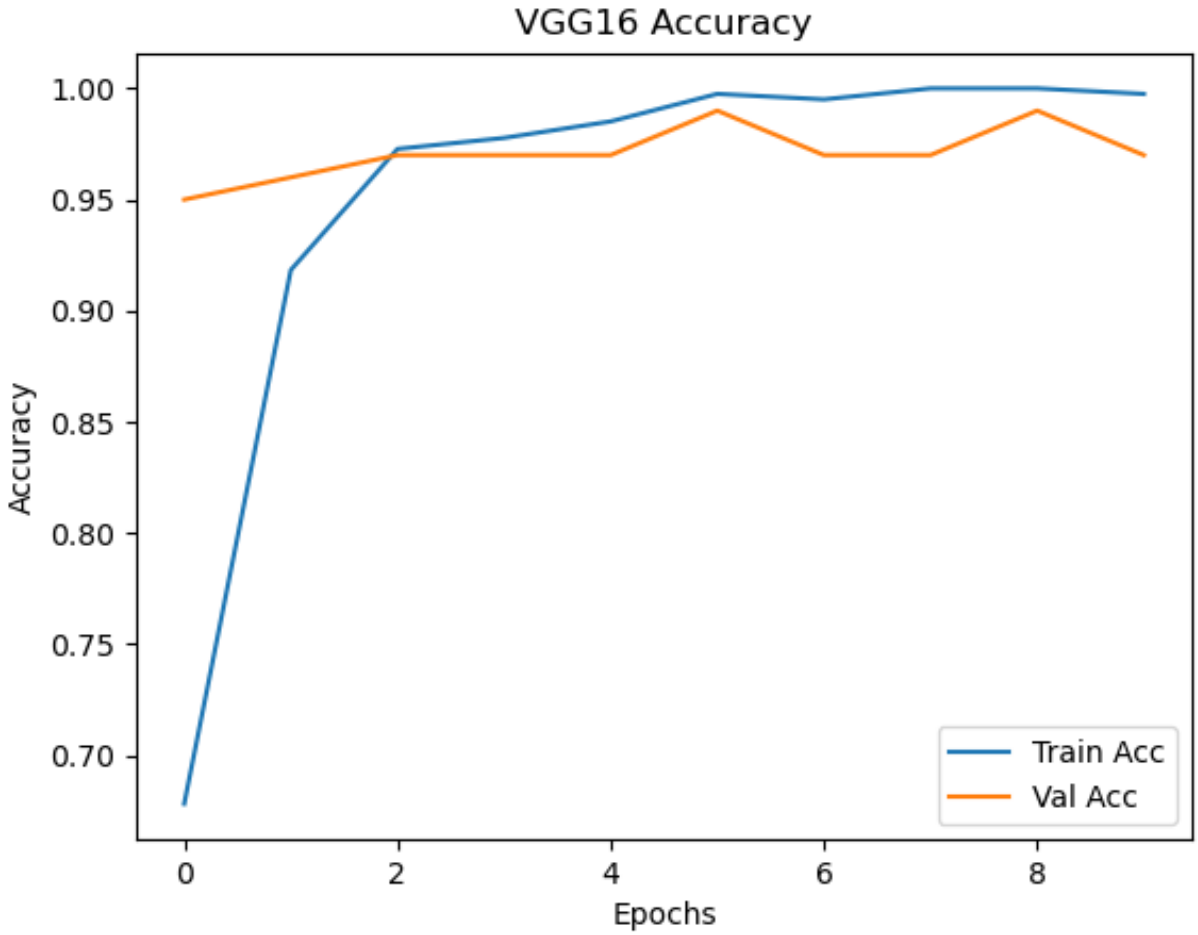
Data Preprocessing:
- Resized images to 128x128 (CNN) and 224x224 (VGG16)
- Normalized pixel values to [0,1]
- Automatically split dataset into training and validation sets (80/20)

Model Training:
- CNN: 3 convolutional layers, 10 epochs
- VGG16: Transfer learning with pre-trained ImageNet weights, 10 epochs, fully connected layers added on top

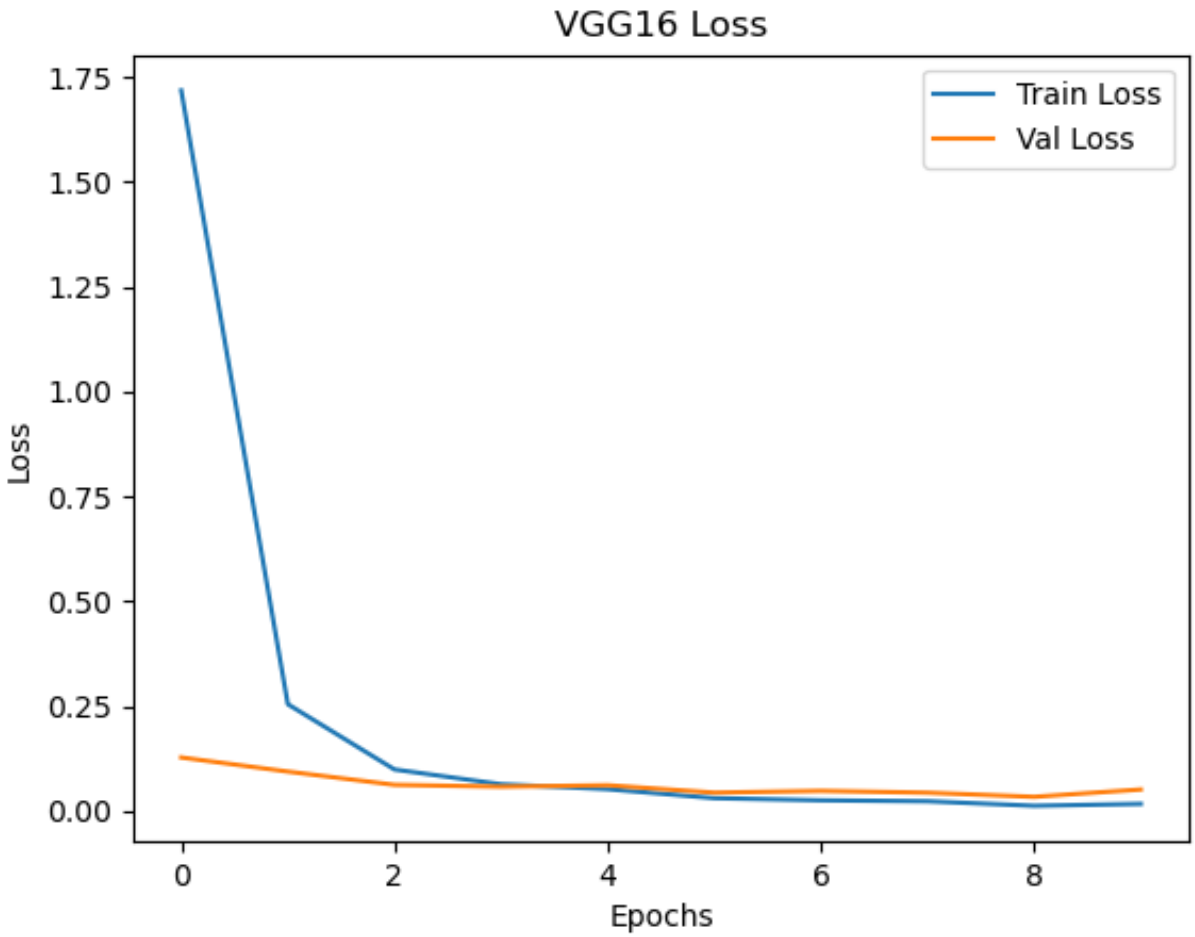
Evaluation:
- CNN Test Accuracy: 72.28%
- VGG16 Validation Accuracy: 99.00%

Graphs

VGG16 Accuracy



VGG16 Loss



Conclusion:

- CNN provides a basic model for cancer detection with moderate accuracy.
- VGG16 Transfer Learning significantly improves accuracy using pre-trained ImageNet features.
- Both models and graphs are included for comparison and evaluation purposes.