Final Evaluation Report: Histopathological Breast Cancer Detection

Introduction

This report summarizes the performance of two distinct deep learning models—a Simple Convolutional Neural Network (CNN) and a Transfer Learning model using VGG16—for the binary classification of cancer in histopathological images. The project involved a critical improvement by increasing the training dataset size to combat overfitting and improve model generalization.

Methodology

The project was divided into two main phases:

1. **Initial Training and Evaluation**: The models were first trained on a smaller dataset consisting of 1,000 training images (500 for each class). The performance was then evaluated on a separate 1,000-image test set.
2. **Improved Training with More Data**: The training dataset was significantly expanded to over 9,000 images, and both models were re-trained from scratch. This was done to mitigate the signs of overfitting observed during the initial runs.

Results

Phase 1: Small Dataset (1,000 Training Images)

- **Simple CNN Test Accuracy**: **82%**
- **Transfer Learning (VGG16) Test Accuracy**: **74%**

Analysis: Initially, the Simple CNN outperformed the Transfer Learning model. The low validation accuracy during training indicated that both models were overfitting to the small training set, but the simple CNN's architecture happened to generalize better to the specific images in the test set.

Phase 2: Expanded Dataset (9,000+ Training Images)

- **Simple CNN Test Accuracy**: **70%**
- **Transfer Learning (VGG16) Test Accuracy**: **81%**

Analysis: This phase demonstrated the critical importance of data size. The Simple CNN's test accuracy unexpectedly decreased, highlighting its sensitivity to the new, larger dataset. In contrast, the Transfer Learning model's test accuracy **increased by 7%**, from 74% to a final score of **81%**. This shows that the VGG16 model, with its robust, pre-trained feature extraction

capabilities, was able to leverage the additional data more effectively to learn better, more general patterns.

Conclusion

Based on the final results, the **Transfer Learning model using VGG16 is the superior choice for this task**. Although the simple CNN showed a good result on the initial small dataset, its performance proved to be unstable and ultimately decreased with more data. The Transfer Learning approach, by contrast, demonstrated a clear and significant performance gain by being exposed to more data. This outcome confirms a core principle of deep learning: providing a pre-trained model with more diverse data is an extremely effective strategy for achieving robust and high-performing results.