Advanced Machine Learning: Project 3

Problem

The goal of this assignment was to implement and evaluate three different neural network architectures on the CIFAR-10 image classification dataset. The target was to explore how architectural choices—such as convolutional layers, skip connections, and dense connectivity—affect classification performance. Specifically, we aimed to:

1. Design and train a standard Convolutional Neural Network (CNN).
2. Implement a Residual Network (ResNet) with skip connections.
3. Train a Dense Convolutional Network (DenseNet) with densely connected layers.
4. Compare their performance and analyze their training behavior.

Solution

To solve the problem, I used TensorFlow and TensorFlow Datasets libraries, which I installed inside a virtual environment named `myenv`. The CIFAR-10 dataset was loaded and split into training and testing sets, with supervised learning enabled to return image-label pairs.

Model 1: Convolutional Neural Network (CNN)

The first architecture was a CNN, commonly used for image classification tasks. My implementation included:

1. Three convolutional layers, each followed by ReLU activation and max-pooling to reduce spatial dimensions.
2. A flattened output layer connected to four dense layers.
3. A final dense output layer with 10 units and softmax activation for multi-class classification.
4. Data augmentation (random flips, rotations, zoom, crop, and contrast adjustment) using the Keras `Sequential` API.
5. Training using the Adam optimizer and sparse categorical cross-entropy loss function.

Initially, training with 10 epochs yielded very low accuracy. After increasing to 74 epochs, the model reached 73% accuracy, and eventually, with 100 epochs, achieved a test accuracy of 74.03%. I believe that extending training beyond 500 epochs could potentially reach the target of 95%.

Model 2: Residual Network (ResNet)

The second architecture was a ResNet, designed to combat the vanishing gradient problem by introducing shortcut (residual) connections. This architecture featured:

1. Several residual blocks, each with two or more convolutional layers and identity shortcut paths.
2. Deeper architecture compared to the CNN model, allowing for richer hierarchical feature extraction.
3. Final fully connected layers and softmax classification.

This model achieved significantly better results, with a test accuracy of 92.37% after training for 100 epochs.

Model 3: Dense Convolutional Network (DenseNet)

The third architecture was a DenseNet, in which each layer receives inputs from all preceding layers. Its structure included:

1. Dense blocks, each containing multiple convolutional layers connected in a feed-forward manner.
2. Bottleneck layers to control the number of parameters.
3. A final classification layer with softmax activation.

Despite its theoretical advantages, this model was difficult to train under the current conditions and showed signs of overfitting. After training for 68 epochs, the best test accuracy achieved was 66.16%.

Performance Summary

1. **CNN (100 epochs)**: Test Accuracy = **74.03%**
2. **ResNet (100 epochs)**: Test Accuracy = **92.37%**
3. **DenseNet (68 epochs)**: Test Accuracy = **66.16%**

Conclusion

Among the three architectures, ResNet performed the best on the CIFAR-10 dataset, thanks to its ability to train deeper networks effectively using residual connections. The CNN model, although simpler, achieved moderate performance but lacked the depth to capture complex features. The DenseNet model struggled to generalize under the current training setup, despite its potential for high performance. This project highlights how architectural design, regularization, and training duration significantly influence model accuracy in deep learning.