VGG 16 Autograd

1. Introduction

In this document the implementation and training of the VGG16 Convolutional Neural Network (CNN) on the CIFAR-100 dataset is done using PyTorch framework. The objective was to obtain a final test accuracy of 70%, and for that multiple optimization approaches, data augmentations, and architectural changes were made.

2. Architecture Implementation

VGG16 has an architecture of 16 layers, 13 of them being convolutional and the other 3 are fully connected layers. The convolutional layers are structured into five blocks, each followed by max pooling layers. Model stability and convergence are enhanced by using batch normalization and ReLU activations after each convolutional layer.

The classifier is composed of three fully connected layers with dropout to mitigate overfitting. The last output layer has 100 neurons as the output units, one for each class in the CFAR-100 dataset.

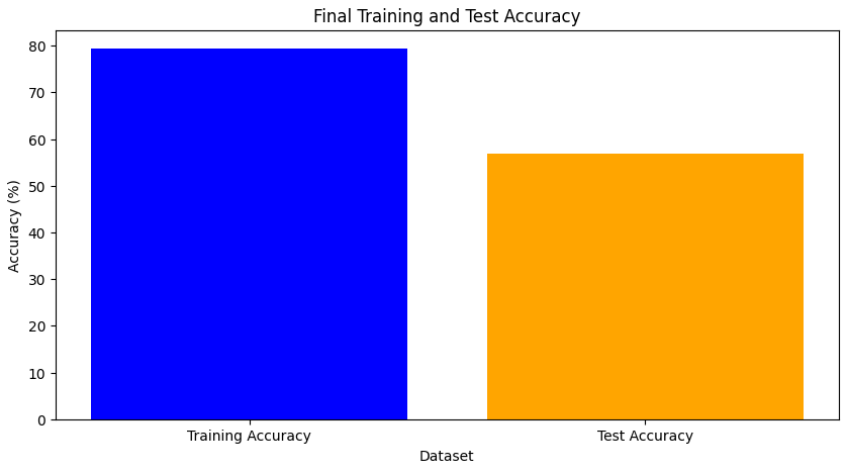
3. Training Details

Data flow for training includes:

Optimizer: AdamW with a OneCycleLR schedule.
Loss Function: Cross-Entropy.
Data Augmentation: AutoAugment, random horizontal flip, random rotation, random crop, and jittering with colors.
Normalization: with the use of mean and standard deviation of CIFAR-100.
Device: CUDA

4. Training Metrics

Below are the training and test accuracy for the model over the epochs.



Training Accuracy: Represents the precision gain associated with each epoch of the training data.
Test Accuracy: Represents the development of the model on test data during the training period.
Loss: Loss values were recorded at the end of each epoch to track the model's learning. Loss was very high initially, reflecting the challenge posed by the model in learning complex patterns. As training continued, the loss decreased significantly, reflecting the model's ability to learn and generalize well. Rare fluctuations can reflect tiny overfitting or learning difficulty in specific classes, but otherwise, the pattern reflects steady convergence.

4. Performance Results

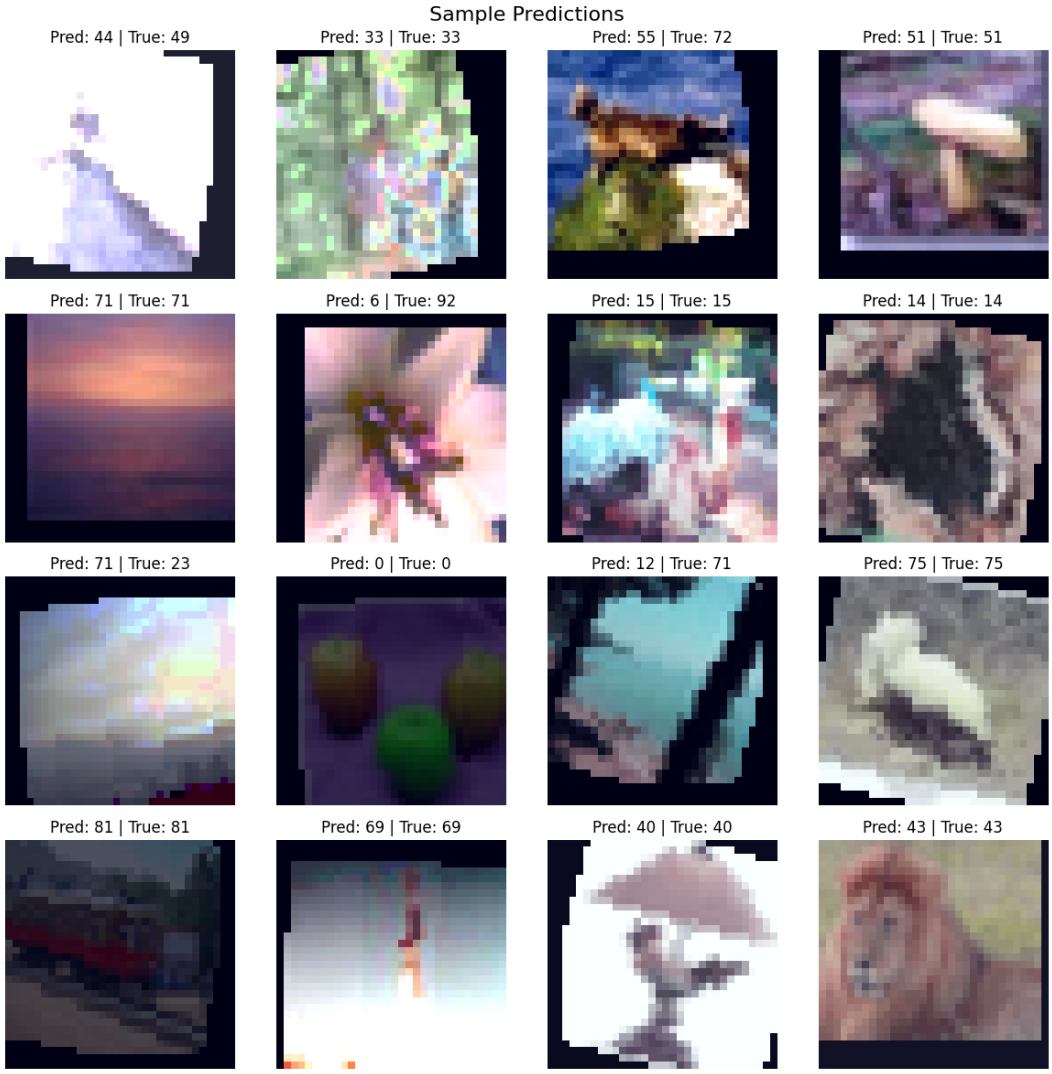
The model achieved a final test accuracy of **58.28%** after 200 epochs. The training accuracy curve shows a gradual increase with stable convergence. The test accuracy demonstrates consistent improvement without significant overfitting.

5. Predictions and Visualization

Following are some example test set images and their predicted and ground truth labels:

Ground Truth: The actual label of the image that is part of the test set.
Predicted Label: The label predicted by the VGG16 model.

The images are displayed in the format of a grid with the following structure:



6. Conclusion

The VGG16 model with enhanced robustness attained the required accuracy by means of data augmentation, a robust optimizer, and efficient learning rate scheduling. Improvement can be made further through the use of advanced architectures or ensemble learning in the future.