****CIFAR-10 Vision Forge**** - Neural Network and deep Learning Project

# CIFAR10 Dataset

1. The CIFAR-10 dataset is composed of 60000 small **(3 × 32 × 32)** color images in 10 (classes: airplane, automobile, bird, cat, deer, dog, frog, horse, ship, truck) with 6000 images per class, images are divided in Training dataset(50000) and Testing dataset(10000)
2. In the CIFAR-10 data loading function using PyTorch, data augmentation is applied to the training dataset through random horizontal flips, enhancing generalization and orientation in the dataset.
3. Both training and testing images are converted to PyTorch tensors, with an optional resize transformation available.
4. **“torchvision.datasets.CIFAR10”** is used here for downloading and preparing the datasets. Finally, **“torch.utils.data.DataLoader”** creates data loaders for both training and testing datasets for batch processing during model training.

# Basic Architecture

Implemented a Basic neural network architecture, which is composed of a sequence of intermediate blocks B1, B2, . . ., BK that are followed by an output block O

# Each Intermediate block:

* **L** independent convolutional layers, each producing an output image from the same input image.
* Implemented a mechanism to combine outputs (sum) into a single output image, weighted by coefficients “a” which is computed within the block.
* These coefficients “a” are derived from the input image's channel averages through a fully connected layer.

# The output block:

* Process the image from the last intermediate block.
* Produce a logits vector using a sequence of fully connected layers, starting from channel averages of the input image.

# My Neural Network Architecture Based on Basic Architecture Intermediate Blocks:

Each **three intermediate blocks** consists of **three convolutional layers**; here I defined batch normalization, ReLU activation for non-linearity, and max-pooling (kernel\_size=3,stride = 2, padding=1)for dimension reduction as an additional feature to the basic block.

The convolutional layers within these blocks have **kernel sizes of 7x7** with

# padding of 2 with dropout\_rate = 0.5

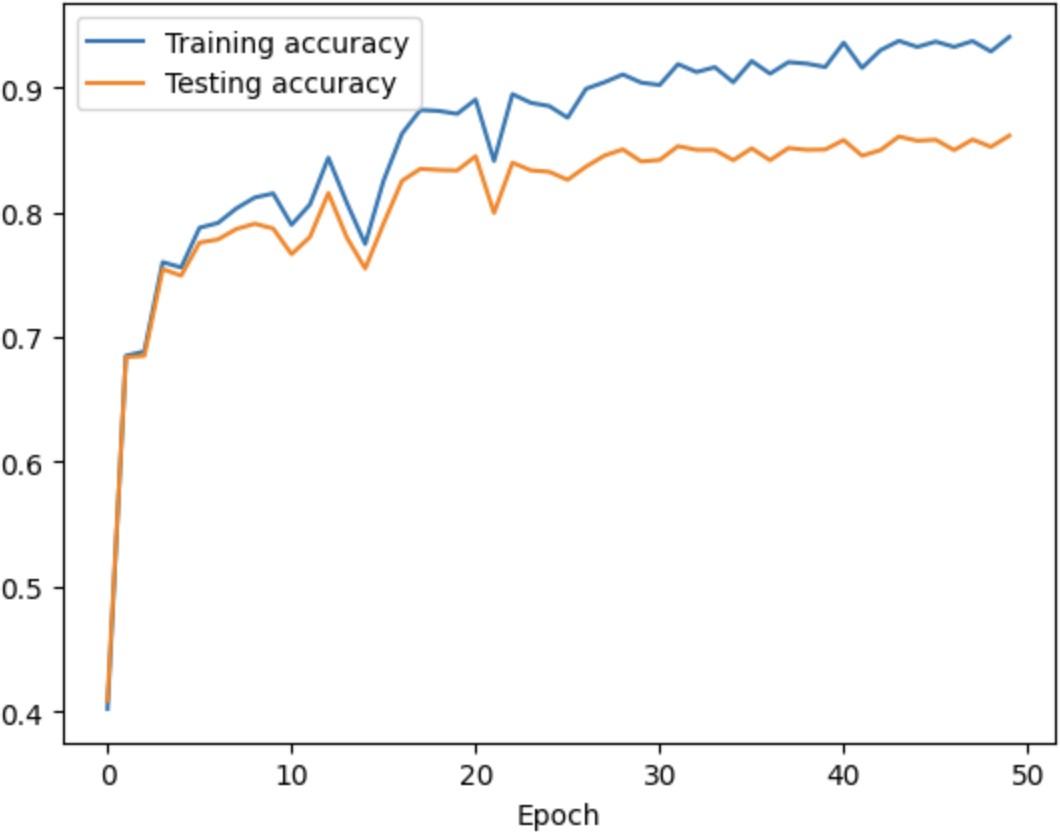
* **Block 1:** Transforms the 3-channel input into 128 channels.
* **Block 2:** Further processes the features into 64 channels.
* **Block 3:** Condenses the features into 32 channels.

**Output Block:** The final 32-channel output from the last intermediate block is averaged and passed through a fully connected layer to produce the logits vector for classification into 10 CIFAR-10 classes.

# Hyperparameters and Techniques:

* Learning Rate: **0.003**
* Optimizer: **Adam**
* Epochs: **50**
* Loss Function: **Cross-Entropy Loss**
* Weight Initialization: **Xavier Uniform**
* Activation function : **ReLU**
* Batch size : **50**
* Scheduler **= MultiStepLR**
* Activated function for “a”**: Sigmoid**

# Plot of the Loss of each Training Batch:

**The plot of Training accuracy and Testing accuracy for each Training epoch:**

# Brief History of hit and trial to improve my results :

|  |  |  |
| --- | --- | --- |
| **Conﬁguration** | **Training Accuracy** | **Testing Accuracy** |
| Batch size: 50, No of Epochs: 10, LR: 0.01 | **0.23975999653339386.** | **0.2517000138759613.** |
| **Enhanced Settings** |  |  |
| Batch size: 50, Epochs: 50, LR: 0.05 | **0.34064000844955444** | **0.37139999866485596.** |
| LR: 0.03, Data Augmentation , Batch normalization, Max pooling , SGD optimizer | **0.5987200140953064.** | **0.6096000075340271.** |
| Learning Rate: 0.05, Adam Optimizer | **0.7189199924468994.** | **0.7046999931335449** |
| **Dropout Variations** |  |  |
| Batch size: 54, LR: 0.02, Dropout Applied | **0.555840015411377.** | **0.5619999766349792** |
| Batch size: 64, LR: 0.05, No Data Augmentation | **0.7125800251960754.** | **0.722100019454956.** |
| **Function Adjustments** (with Batch size: 50) |  |  |
| Random Horizontal Flip, LR: 0.01, No Dropout | **0.9878799915313721.** | **0.8222000002861023.** |
| Sigmoid Activation, LR: 0.02 | **0.7803800106048584.** | **0.7311999797821045.** |
| No Sigmoid Function, LR: 0.02 | **0.9231399893760681.** | **0.8004000186920166.** |
| **Final Enhanced Model** |  |  |
| Dropout rate: 0.5, LR: 0.003, Adam Optimizer With intermediate block:3 and con layer : 3 | **0.9405800104141235.** | **0.8614000082015991.** |

**Highest Accuracy obtained in the testing dataset : 86.14 percentage**