CS89BD Deep Learning, Fall 2024
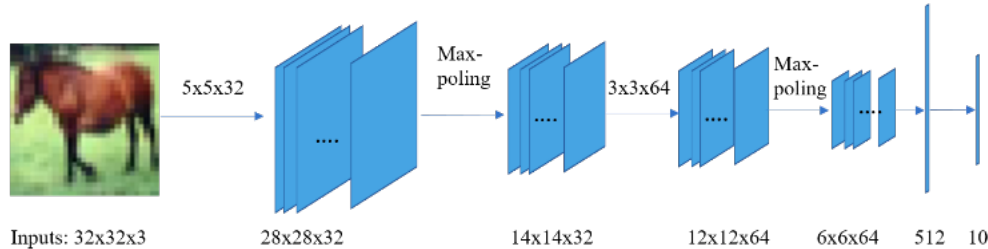
Assignment 2
Due: September 26, 2024
Total Points: 100

Question 1 (20 Points): Design a convolutional neural network (CNN) model considering the following criteria:

    a. Design a CNN model with the following configuration

Inputs: 32x32x3   28x28x32   14x14x32   12x12x64   6x6x64   512   10

    b. Use activation function: ReLU (in hidden layer)
    c. Batch Normalization
    d. Number of epochs: 100 and batch size: 32
    e. Optimizer: SGD with a learning rate of 0.1 and momentum.
    f. Use Cosine Scheduler to reduce learning rate until final learning rate of 0.001

Evaluate your model on CIFAR-10 dataset.

You need to report training logs, testing errors and accuracy, and computational time. Please use graphs for training logs (i.e., training loss, validation loss)

Question 2 (80 Points): The aim of this task is to develop a Deep Convolutional Neural Networks (DCNN) model and get better accuracy for scene understanding or classification tasks. (Limit your training to 500 epochs. You can set a smaller number of epochs as well. But please make sure your model provides a decent performance.)
Dataset: 15-Scene_Image_Dataset

Task 1: Prepare your data (30)
    a. Prepare 3 different sets (Training set, Validation Set, and Testing Set) from the original data. You must create a validation set from the training set taking 10% of images per class randomly.
    b. The training set should use 80% of images per class, remaining 20% should be kept for testing your model.
    c. If data are class-wise imbalanced, please make it balanced.

An example of images from the 15-scene dataset [1]

Task 2: Implement a DCNN model to classify the images. Please consider the following characteristics while designing your model. (Note: you can consult any implementation that uses 15-scene datasets. Please make sure to include references) (50)

1. Data Augmentation
2. Better Initialization methods (i.e., Xavier or He)
3. Activation Function: ELU
4. Dropout (Regularization)
5. Batch Normalization

Compare your model performance for different optimization functions:

I. SGD
II. Adam
III. RMSProp

Report training log (i.e., training loss, validation loss, training accuracy, validation accuracy) and computation time for each optimizer.

Write a report that represents a comparison of two models.

The report outlines: The report must contain:

- The title page includes course title, course number, your name, WSU ID, and assignment number
- Introduction
- Methodology
- Deep Learning Architecture
- Experiment and Results (graphs)

- o A graph that represents training error (y-axis) and training time(x-axis)
- o A graph that represents errors (i.e., training error and testing error on the y-axis) and training time (x-axis)
- o A graph that represents loss (training loss and testing loss) and number of epochs (x-axis)
- o Report **state-of-the-art** accuracy for this dataset
- Conclusion
- Reference

Reference:

[1] Lazebnik, S., Schmid, C., & Ponce, J. (2006, June). Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In 2006 IEEE computer society conference on computer vision and pattern recognition (CVPR'06) (Vol. 2, pp. 2169-2178). IEEE.