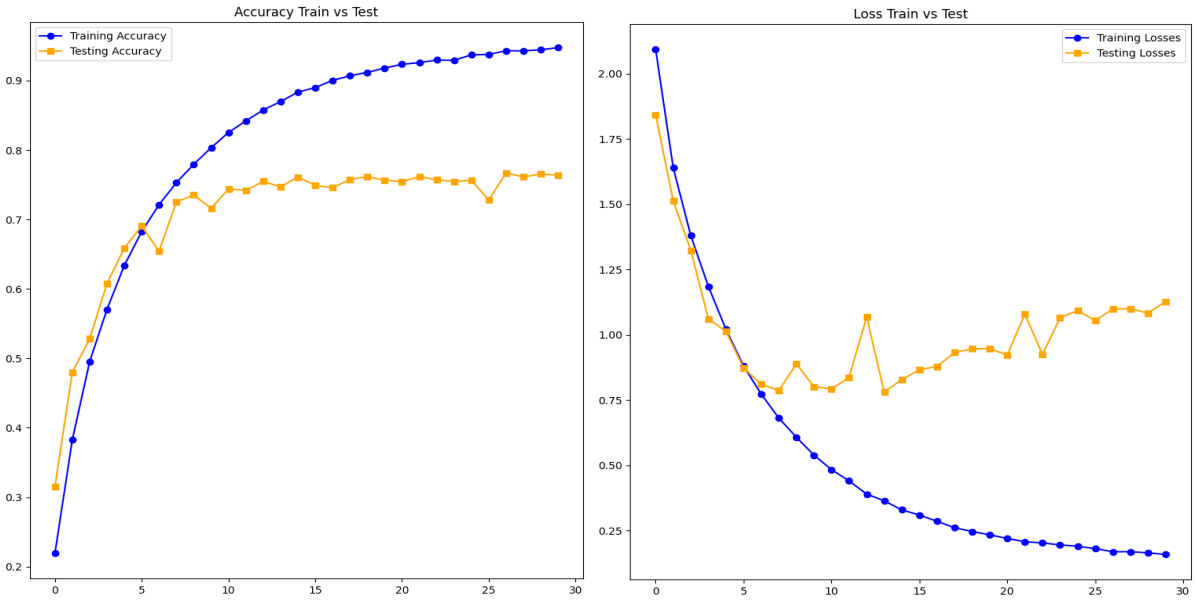
COMPUTER VISION (CAP5415)-PA-4 REPORT
IMAGE CLASSIFICATION

Objective:

To classify RGB color images in the CIFAR-10 dataset, build a customized Convolutional Neural Network (CNN) structure with multiple convolutional layers and fully connected layers. Enhancing feature extraction and classification precision for all the classes the dataset represents is the goal.

Model 1:

The model is constructed with two hidden layers in the fully connected network and four layers in the convolutional network by using the ReLU activation function and Cross Entropy Loss. Using a mini-batch size of 10 and a learning rate of 0.01 over 30 epochs, dropout regularization techniques are included in the training process of stochastic gradient descent.

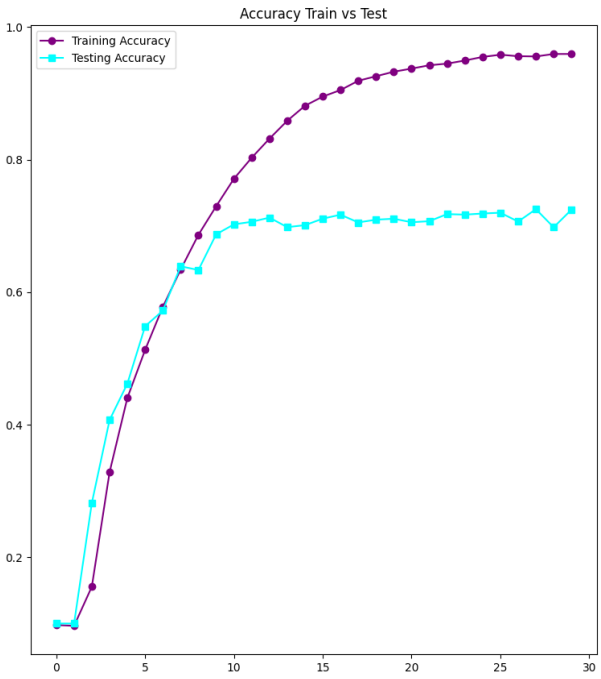


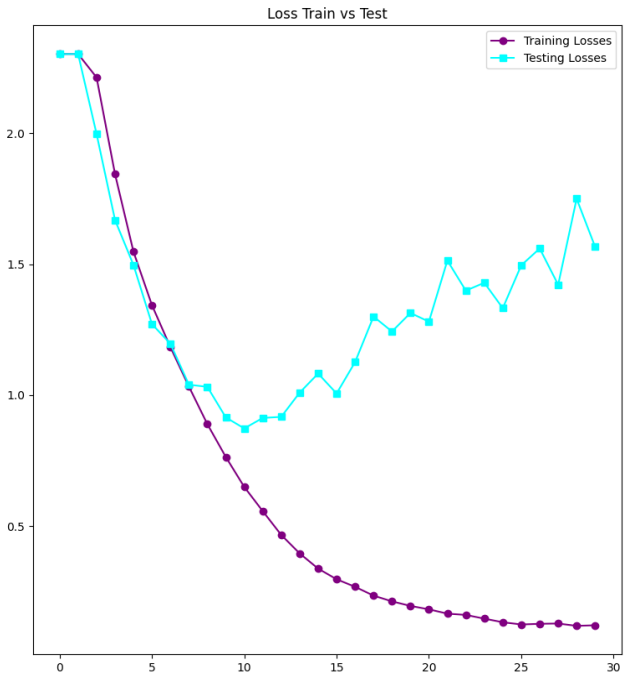
Results:

Train Accuracy: 95%
Test Accuracy: 76.73%

Model 2:

The fully connected network consists of two hidden layers plus two extra convolutional layers, all of which use the ReLU activation function and Cross Entropy Loss in training and design. Training is done using stochastic gradient descent with a learning rate of 0.01 over 30 epochs, a mini-batch size of 10, and dropout integration for regularization.





Results:

Train Accuracy: 96%
Test Accuracy: 72.51%

Observations:

When experimenting with hyperparameters, it's clear that boosting the number of convolutional layers doesn't necessarily enhance accuracy. Proper adjustment of this parameter is essential, considering the dataset characteristics and input features.

Additionally, elevating the learning rate leads to overfitting. This is particularly noticeable in smaller datasets like CIFAR-10, where the model quickly becomes overly tailored to the training data, causing a significant gap between training and test accuracies.