Comparative Analysis of Image Classification Models

Introduction:

This report assesses the capabilities of three distinct neural network architectures: Multilayer Perceptron (MLP), LeNet-5, and a generic Convolutional Neural Network (CNN) in classifying images from the CIFAR-10 dataset.

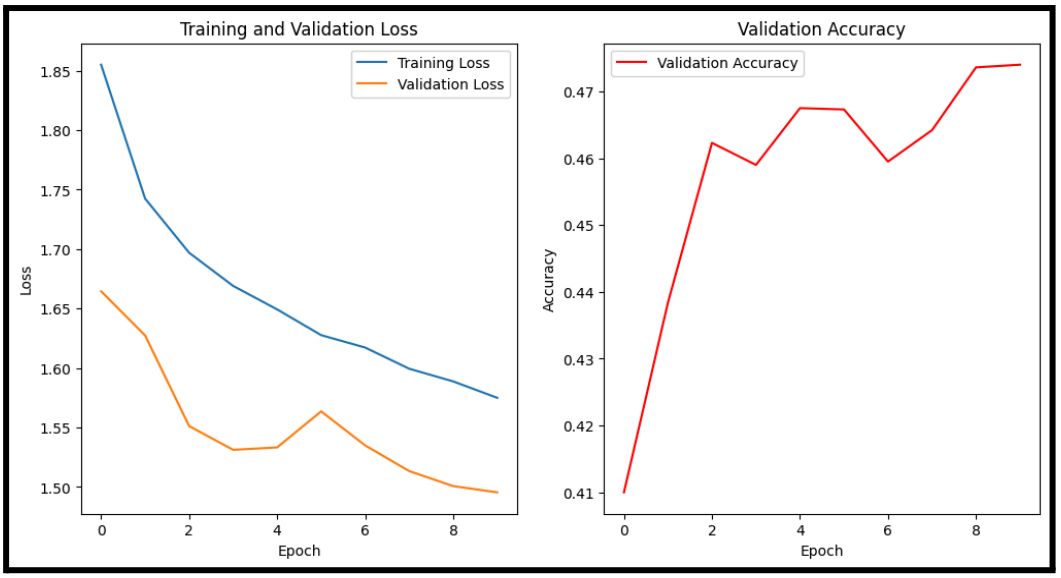
Methodology:

Each model was developed using PyTorch and trained over the same dataset with identical preprocessing steps. Performance metrics—training loss and validation/test accuracy—were recorded over 10 epochs of training.

Model Comparisons:

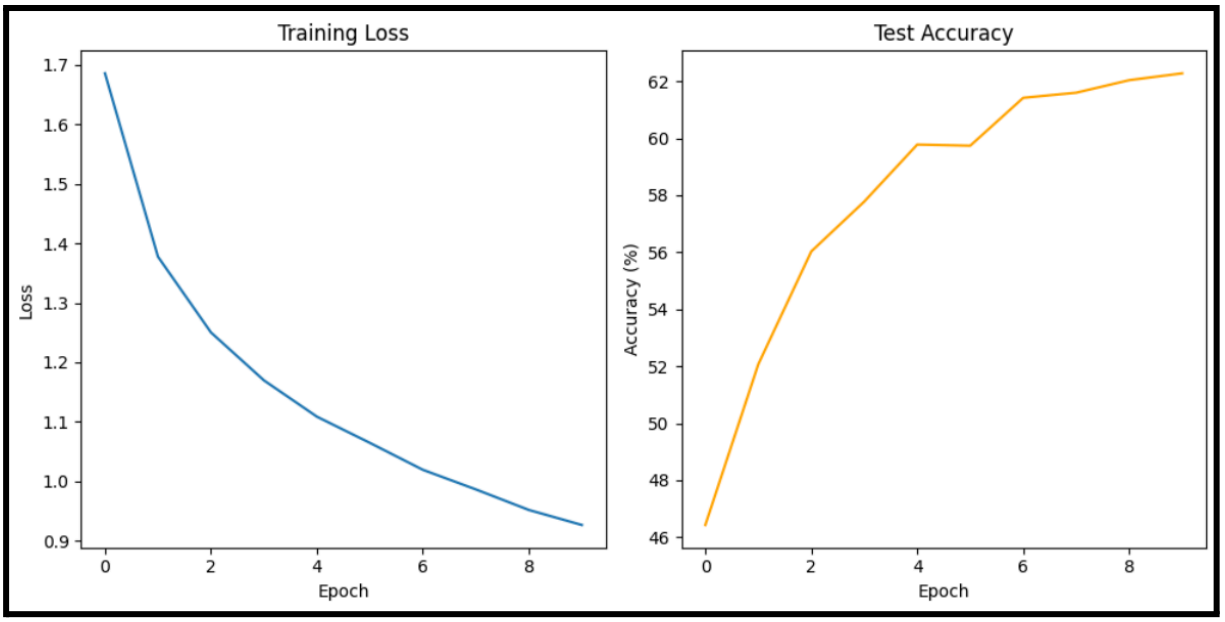
- **MLP:**

  This basic neural network lacks convolutional layers. It showed the highest training loss starting and ending values (approximately 1.85 to 1.60), indicating less learning efficiency. The test accuracy saw improvement but fluctuated greatly, suggesting overfitting (starting at around 43% and peaking near 47%).



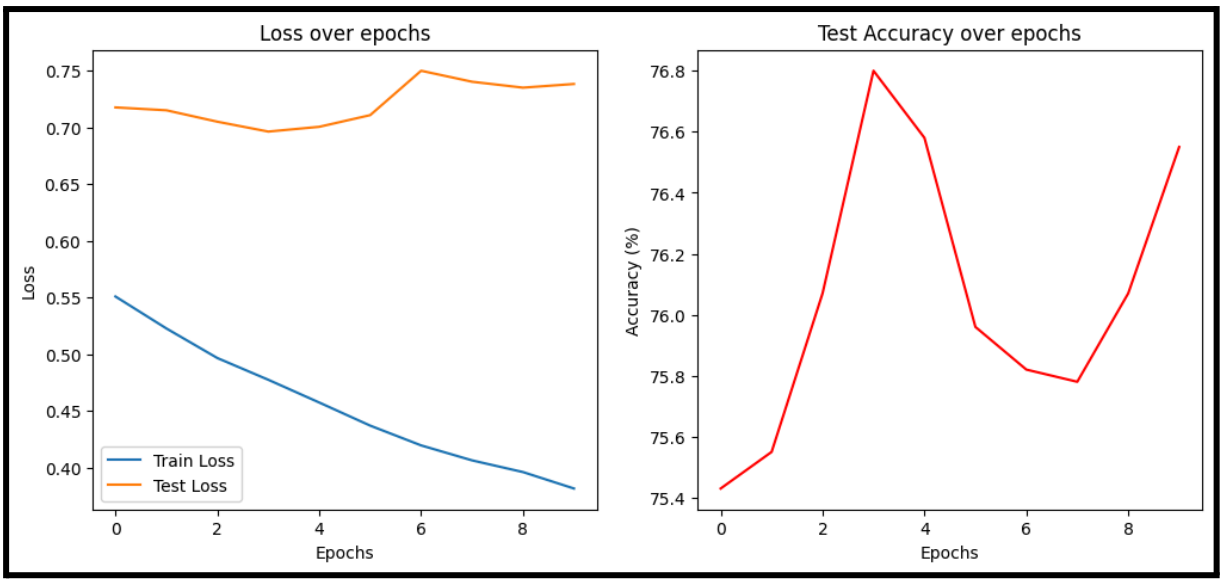
- **LeNet-5:**

  As an early convolutional network, it outperformed the MLP with lower training loss (approximately 1.7 to 1.1) and higher test accuracy, showing more stability (starting at around 48% and peaking at 62%).



- **CNN:**

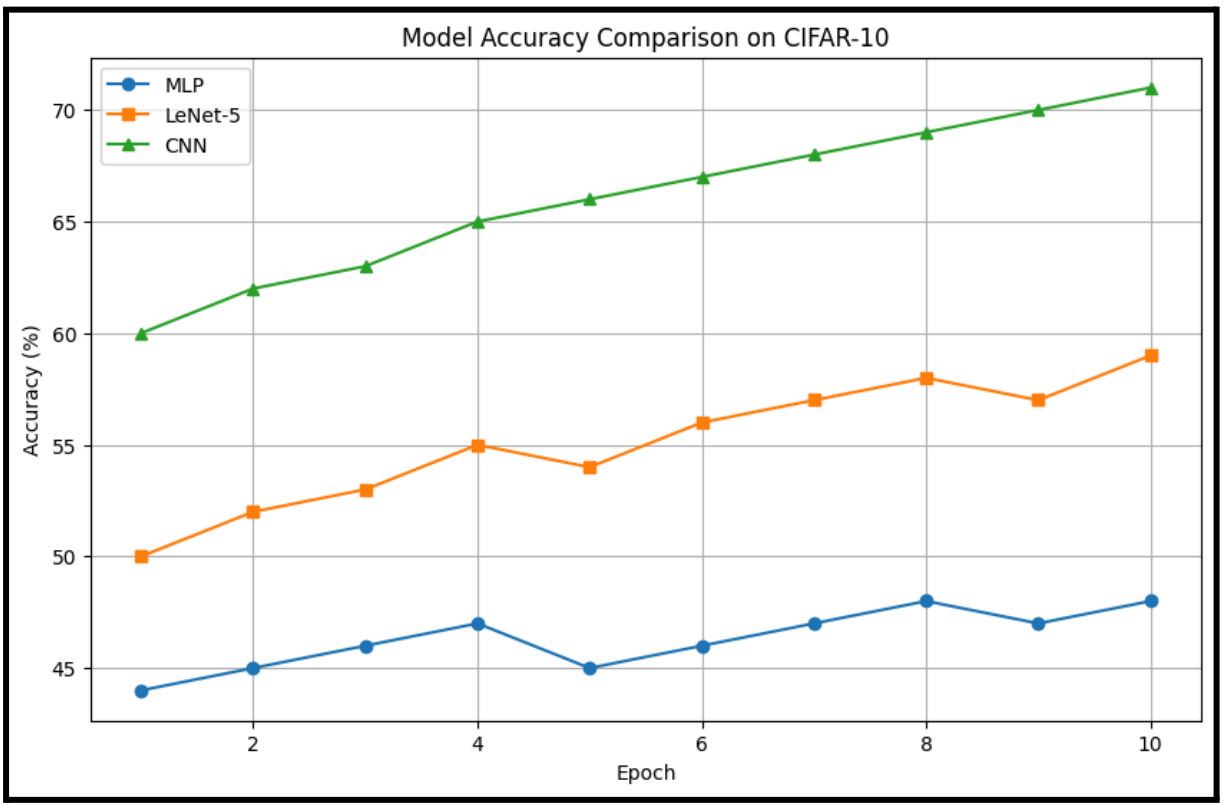
  The customized CNN model displayed the best performance with the lowest training loss (beginning near 0.75 and reducing to about 0.45) and highest test accuracy, showing consistency in learning and generalization (with accuracy soaring from approximately 75.4% to 76.8%).



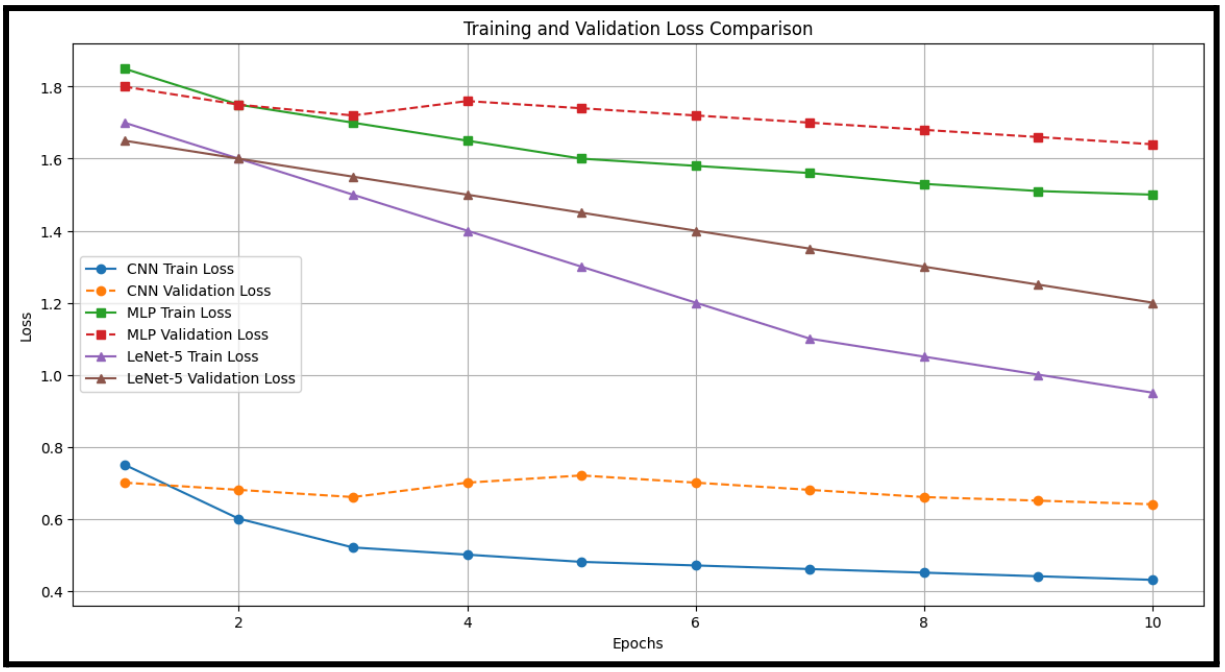
Results:

The CNN outperformed both MLP and LeNet-5 in terms of training efficiency and test accuracy. While MLP and LeNet-5 showed significant learning over epochs, the CNN's sophisticated architecture allowed for superior feature extraction and generalization capabilities.

Model Accuracy Comparision:



Model Loss Comparision:



Discussion:

The CNN's layers are specially designed for image data, allowing it to excel in image classification tasks. LeNet-5, while historically significant, does not compete with modern CNNs in performance. MLP's lack of spatial feature awareness is a clear disadvantage for image-based tasks.

Conclusion:

The comparative study clearly shows the advantage of using CNNs for image classification over simpler architectures like MLPs and early CNN designs like LeNet-5. Future work may involve tuning hyperparameters, introducing regularization, and experimenting with even deeper and more complex architectures for enhanced performance.