Assignment 5

Name: Mohamed-Obay Alshaer
Student Number: 300170489

Submission Date: March 21, 2025

Professor: Daniel Shpairo

SEG4300 - Applied Machine Learning

1 Deep Learning for Image Classification on CIFAR-10 Dataset

1.1 Dataset Selection and Overview

The CIFAR-10 dataset was selected for this deep learning classification task. This dataset consists of 60,000 32×32 color images across 10 balanced classes (6,000 images per class): airplane, automobile, bird, cat, deer, dog, frog, horse, ship, and truck. The dataset is split into 50,000 training and 10,000 test images.

CIFAR-10 is ideal for this assignment because it offers a meaningful challenge for deep learning models while remaining computationally manageable. The variety of object classes with different shapes, textures, and contexts makes it suitable for demonstrating convolutional neural network capabilities.

1.2 Model Architecture and Training

A Convolutional Neural Network (CNN) architecture was implemented using PyTorch for this classification task. The model consisted of:

- Multiple convolutional layers with increasing filters (32, 64, 128)
- MaxPooling layers to reduce spatial dimensions
- Batch normalization for training stability
- Dropout layers (p=0.25) for regularization
- Fully connected layers for final classification

The model was trained using cross-entropy loss and the Adam optimizer with a learning rate of 0.001 for 50 epochs. Data augmentation techniques including random horizontal flips and slight rotations were applied to enhance model generalization.

1.3 Results and Evaluation

The trained CNN achieved 83.03% accuracy on the test set, which is respectable for the CIFAR-10 dataset using a relatively simple architecture. Performance varied across classes:

- Best performing classes: automobile (94.7% recall), truck (93.3% recall), and ship (93.0% recall)
- Most challenging classes: cat (58.9% recall) and bird (68.9% recall)

Analysis of misclassifications revealed consistent patterns of confusion between visually similar classes:

- Cat/dog confusion (229 cats misclassified as dogs)
- Bird/dog confusion (77 instances)
- Deer/horse confusion (73 instances)
- Transportation-related confusions (airplane/ship, automobile/truck)

1.4 Key Insights and Analysis

The learning curves demonstrate that the model began to plateau after approximately 40 epochs. Despite this, the validation accuracy continued to slightly improve while maintaining a reasonable gap with training accuracy, indicating the effectiveness of our regularization strategies.

The model performed better on classes with distinctive shapes and more uniform backgrounds (vehicles) compared to natural categories with higher visual variability (animals). This aligns with expectations, as the low resolution (32×32) of CIFAR-10 images makes subtle distinctions challenging.

Most misclassifications occurred between semantically and visually similar categories. For example:

- Cats and dogs share many visual features and pose variations
- Birds and planes both appear in sky contexts with similar profiles
- Automobiles and trucks have similar structural components

1.5 Limitations and Future Improvements

While the current model achieved solid performance, several enhancements could further improve results:

- Implement advanced architectures such as ResNet or DenseNet to allow deeper networks without vanishing gradient issues
- Apply transfer learning by utilizing models pre-trained on larger datasets
- Conduct systematic hyperparameter tuning to optimize learning rates and regularization strengths
- Implement attention mechanisms to help the model focus on discriminative image regions
- Explore ensemble methods to combine predictions from multiple models

1.6 Conclusion

This project demonstrated a complete deep learning workflow using PyTorch for image classification. The CNN model achieved 83.03% accuracy on CIFAR-10, showing the effectiveness of convolutional architectures for computer vision tasks even with limited image resolution. The analysis of model performance across different classes provided valuable insights into the strengths and limitations of the approach, while identifying clear paths for future improvements.