Tensorflow$_{M}odel_{C}NN_{R}eport$

Kashif ur Rahman

August 2025

Project: Image Classification and Analysis on CIFAR-10 Dataset Report Documentation

0.1 Introduction

The CIFAR-10 dataset is a well-known benchmark in computer vision, consisting of 60,000 colored images (32×32 pixels) across 10 classes: airplane, automobile, bird, cat, deer, dog, frog, horse, ship, and truck. The main objective of this project is to apply classical machine learning models and deep learning methods to classify images and compare their performances.

0.2 Dataset Exploration

- Training set: 50,000 images
- Test set: 10,000 images
- Balanced dataset: each class has 6,000 images
- Labels were mapped to human-readable names (e.g., 0 = airplane, 1 = automobile, etc.).

0.3 Data Visualization & Preprocessing

- Visualized 20 random images with labels.
- Plotted histogram of class distribution (balanced).
- Applied OpenCV for preprocessing: grayscale conversion, edge detection (Canny), rotation, flipping, and noise.
- Normalized pixel values from [0–255] to [0–1].
- One-hot encoded the labels.
- Optionally resized images to 64×64 for experimentation.

0.4 Machine Learning Baseline

- Flattened each image into a vector of 3,072 features ($32 \times 32 \times 3$).
- Trained a **Logistic Regression** model (on subset for efficiency).
- Achieved accuracy of ~30–40%.
- **Observation:** Traditional ML struggles with raw pixel data because spatial relationships between pixels are lost.

0.5 Deep Learning Approach (CNN)

- Built a CNN with 3 convolutional layers, pooling layers, dropout, and fully connected layers.
- Trained for 10 epochs using Adam optimizer.
- Achieved ~70–75% accuracy on the test set.
- CNN successfully captured spatial and visual features that ML models missed.
- Accuracy/Loss curves showed good convergence.

0.6 Advanced Analysis

- **Misclassifications:** Common between visually similar classes (e.g., cat vs. dog, truck vs. automobile).
- **Confusion Matrix:** Confirmed overlaps among similar categories.
- **Data Augmentation:** Applied rotation, shift, zoom, and flipping using `ImageDataGenerator`.
- Improved CNN accuracy to **~78–80%**, showing better generalization and reduced overfitting.

0.7 Key Findings & Conclusion

1. Machine Learning methods are not effective for image data due to loss of spatial information.
2. Deep Learning (CNNs) significantly outperformed ML, showing its suitability for vision tasks.
3. Data augmentation improved robustness and boosted performance.
4. CIFAR-10 remains a challenging dataset due to small image size and visual similarity between classes.

Final Note: CNNs with augmentation provide a strong baseline, but more advanced architectures (ResNet, DenseNet, etc.) could achieve even higher accuracy.

Deliverables:

- Google Colab Notebook with full implementation.
- This report summarizing methodology, results, and conclusions.