CIFAR-10 Dataset

Understanding and Utilizing a Classic Image Classification Benchmark

By Suman Rai

Contents

1

Introduction

The CIFAR-10 Dataset is created by Canadian Institute for Advanced Research (CIFAR).The CIFAR-10 dataset is primarily used for image classification tasks and testing deep learning models. The CIFAR-10 dataset is a widely used collection of labeled images for machine learning and computer vision tasks It contains 60,000 color images of size 32x32 pixels, divided into 10 distinct classes. There are 50,000 training images and 10,000 test images for evaluation.

Objectives

- To build a model that can classify images from the CIFAR-10 Dataset.
- To undderstand the role of Convolution Neural Networks (CNN) in image recognition.

Dataset Composition

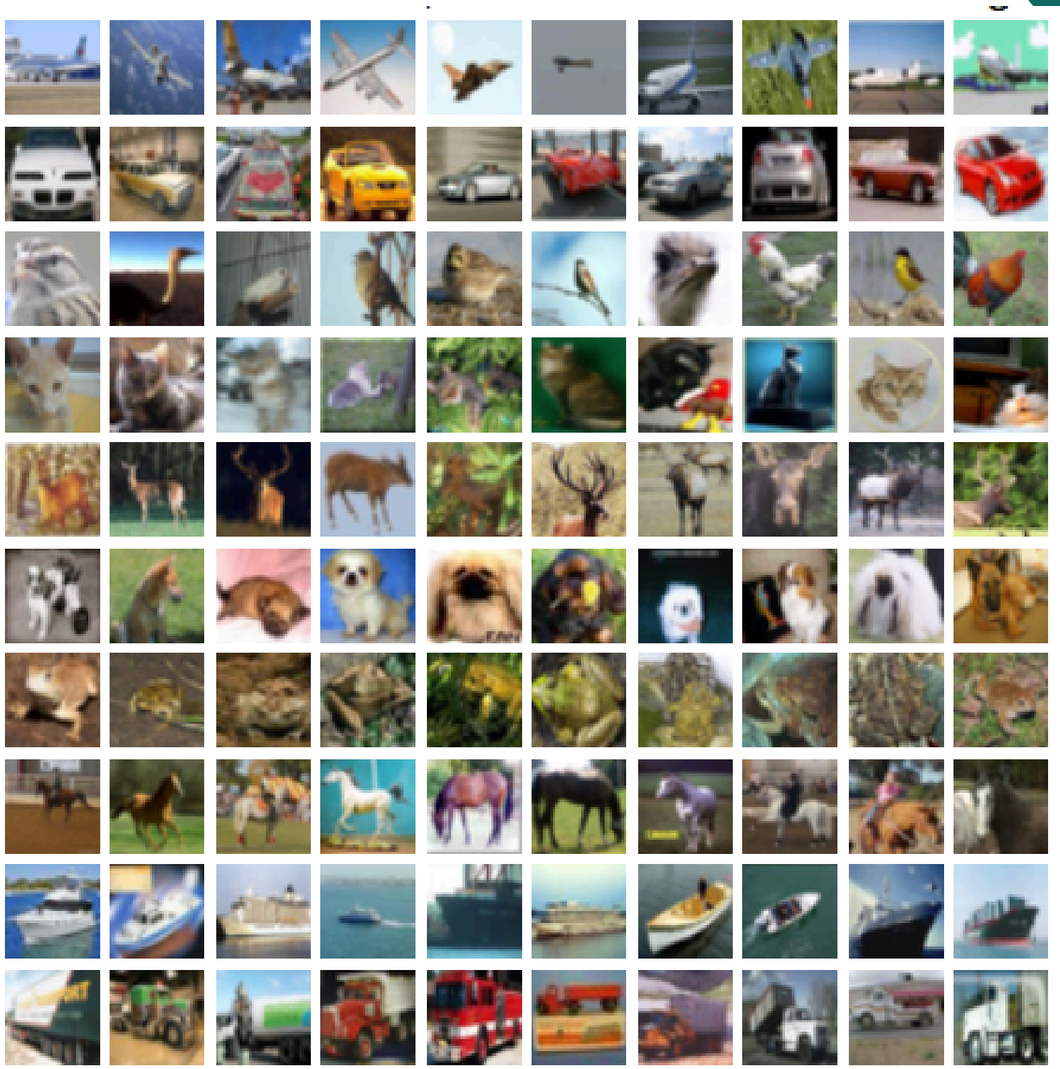
3

- Contains 60,000 color images (32x32 pixels).
- Divided into 10 classes, with 6,000 images per class.
- 50,000 images for training, 10,000 images for testing.
- Classes: airplane, automobile, bird, cat, deer, dog, frog, horse, ship, truck.
- Images are low-resolution, making feature extraction challenging.

CLASSES

This Dataset has 10 categories:

1. Plane
2. Car
3. Bird
4. Cat
5. Deer
6. Dog
7. Frog
8. Horse
9. Ship
10. Truck



Approaches and Algorithms

CNNs are designed for image data and can automatically extract hierarchical features.

Convolutional Neural Network (CNN)

A Convolutional Neural Network (CNN) is a type of deep learning model mainly used for image and video recognition. It works by automatically detecting important features (like edges, shapes, and texture) using convolution layers. CNNs also use pooling layers to reduce data size and fully connected layers to make final predictions. They are highly effective because they learn directly from raw data without needing manual feature extraction. CNNs are widely applied in areas like computer vision, medical imaging, self-driving cars, and facial recognition. They consist of three main types of layers:

1. Convolution Layer
2. Pooling Layer
3. Fully Connected Layer

CNN Architecture

Convolution Layer

A convolution layer is the core building block of a CNN that extracts features from input data. It uses filters(kernels) that slide over the input to detect the local patterns like edges, textures, or shapes. The layers apply operations like dot products and summation between the filter and the input region It helps in the learning of complex patterns.

Pooling Layer

A pooling layer is used in CNNs to reduce the spatial size of feature maps. Pooling makes the model more robust to small shifts and distortions in the input image. It helps to minimize overfitting problem.

Fully Connected Layer

A fully connected (FC) layer connects every neuron in one layer to every neuron in the next layer. It combines the features learned by convolution and pooling layers to make final predictions.

Model Training

- Loss function: CrossEntropyLoss
- Optimizer: SGD
- Trained for 8 epochs to minimize loss and improve accuracy
- Test accuracy: 76.18%
- Train accuracy: 82.28%

Conclusion

- Using the CIFAR-10 dataset helps in understanding image classification and realizing the importance of CNN architecture.
- With the help of convolution, pooling, and fully connected layers, the model can learn features from images and make class predictions.
- Provides a foundation for understanding more complex computer vision like CIFAR-100 and ImageNet.
- CIFAR-10 remains a valuable dataset for learning and experimenting with image classification.
- CIFAR-10 is a simple yet powerful dataset for beginners and researchers.
- Helps in understanding and building image recognition models.
- A stepping stone towards more complex datasets

Thank You