Project Name

<u>CIFAR-10 Object Recognition Using ResNet50</u>

Submitted by:
Tanishq(102217227)

BE Third Year

CSE

Submitted to:

Dr. Anjula Mehto

Assistant Professor

THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

Computer Science and Engineering Department

Thapar Institute of Engineering and Technology, Patiala

November 2024

TABLE OF CONTENTS

Introduction or Project Overview

Object recognition is a critical task in computer vision with numerous applications in fields like autonomous driving, surveillance, healthcare, and robotics. This project focuses on object recognition using deep learning, specifically by leveraging the ResNet50 architecture to classify images from the CIFAR-10 dataset.

The ResNet50 model, known for its deep architecture and efficient training capabilities, utilizes residual learning, where "skip connections" mitigate the vanishing gradient problem by allowing gradients to flow through layers more effectively. This enables the network to have significantly more layers without sacrificing training quality, making it a preferred choice for complex image classification tasks.

The goal of this project is to accurately classify images from the CIFAR-10 dataset, which comprises 10 distinct object categories, demonstrating the effectiveness of ResNet50 for visual recognition tasks in resource-constrained settings.

Problem Statement

In this project, the primary problem is to design and implement a convolutional neural network (CNN)-based system to accurately classify images within the CIFAR-10 dataset. The following objectives were established:

1. **To apply data preprocessing techniques that prepare the CIFAR-10 dataset for optimal use in deep learning.**
2. **To fine-tune a pre-trained ResNet50 model specifically for CIFAR-10's 10-class categorization task.**
3. **To achieve high accuracy and minimal overfitting on the test set by applying regularization techniques and hyperparameter optimization.**
4. **To assess model performance using detailed metrics such as accuracy, precision, recall, and F1-score.**

This project aims to strike a balance between computational efficiency and classification accuracy, providing insights into the practical deployment of deep learning models for real-world object recognition.

Overview of the Dataset used

The **CIFAR-10 dataset** is a well-known dataset in computer vision, consisting of 60,000 32x32 color images divided into 10 classes, each representing a common object type: airplanes, automobiles, birds, cats, deer, dogs, frogs, horses, ships, and trucks. Each class has 6,000 images, with 50,000 images used for training and 10,000 for testing.

Key features of CIFAR-10 include:

- **Resolution:** Images are low-resolution (32x32 pixels), requiring models that can learn effectively from limited visual details.
- **Class Diversity:** The dataset contains balanced classes, with each class representing common everyday objects.
- **Challenge:** Due to the small size and diversity of the dataset, training deep models like ResNet50 requires effective preprocessing and data augmentation strategies to enhance generalization and avoid overfitting.

Data preprocessing involved normalizing images to a standard scale, which helps in speeding up convergence during model training and improves accuracy. Furthermore, data augmentation techniques (such as random cropping, horizontal flipping, and brightness adjustments) were applied to artificially expand the training data, improving the model's robustness.

Project Workflow

The project workflow includes several steps, each critical to achieving optimal model performance:

Step 1: Data Preprocessing

- **Normalization: CIFAR-10 images were normalized to improve model training stability. This involved scaling pixel values to a [0,1] range or applying standard normalization to zero-mean and unit-variance values.**
- **Data Augmentation: Techniques like rotation, flipping, and cropping were applied to increase the dataset size artificially, reduce overfitting, and help the model generalize better on unseen data.**

Step 2: Model Architecture (ResNet50)

- **Transfer Learning with ResNet50: ResNet50, pre-trained on the larger ImageNet dataset, was fine-tuned to work effectively on CIFAR-10. Transfer learning allows us to leverage features learned on large datasets, reducing the amount of data and training time needed for our task.**
- **Network Modifications: The final fully connected layer of ResNet50 was modified to output 10 classes instead of the original 1000 classes of ImageNet, aligning with CIFAR-10's class structure.**

Step 3: Training Process

- **Hyperparameter Tuning: The model was fine-tuned by experimenting with different learning rates, optimizers**

(Adam and SGD), and batch sizes to identify the configuration that yielded the best performance. Regularization Techniques: Regularization methods like dropout and L2 regularization were used to prevent overfitting. Dropout randomly deactivates neurons during training, while L2 regularization penalizes large weights.

- **Optimizer and Loss Function: The Adam optimizer was used for its efficient gradient-based optimization, while categorical cross-entropy served as the loss function to measure model performance on the multi-class classification task.**

Step 4: Evaluation Metrics

- **Accuracy: Primary metric to evaluate overall model performance.**
- **Confusion Matrix: A confusion matrix was generated to visually analyze the model's predictions across all classes, helping to pinpoint specific areas for improvement.**

Results

The ResNet50 model, trained and fine-tuned for CIFAR-10, yielded the following results:

- **Training Accuracy:** 93.86%
- **Confusion Matrix Observations:** The confusion matrix revealed that some classes, such as ships and trucks, were easier for the model to classify accurately. However, classes like cats and dogs showed lower accuracy, likely due to visual similarities and limited detail in 32x32 images.

These results demonstrate that ResNet50 is highly capable of learning distinguishing features across most classes within CIFAR-10, though certain classes with overlapping features present more of a challenge.

Conclusion

The project demonstrated the effectiveness of ResNet50 in tackling object classification challenges, especially with small datasets like CIFAR-10. By leveraging transfer learning, data augmentation, and hyperparameter optimization, the model achieved impressive classification performance.

Future Recommendations:

- **Larger and Higher-Resolution Datasets: Using higher-resolution images or larger datasets, such as CIFAR-100 or ImageNet, could further improve accuracy and generalization.**
- **Advanced Model Architectures: Exploring newer architectures, such as EfficientNet or Vision Transformers (ViTs), might yield better results with the same or fewer computational resources.**
- **Additional Regularization: Techniques like batch normalization or MixUp (a data augmentation method) could help in further improving model robustness.**

In summary, ResNet50, with appropriate preprocessing and tuning, proves to be an effective model for CIFAR-10, with promising applications in real-world object recognition tasks.