PANDIT DEENDAYAL ENERGY UNIVERSITY, GANDHINAGAR

School of Technology

Department of Computer Science and Engineering

ARTIFICIAL INTELLIGENCE LAB (23CP307P)

Submitted To:

Dr. Davinder Singh

TEAM:

DHAIRYA SHAH	21BCP010
TUSHAR GODHVANI	21BCP025
VIDHI PARWANDA	21BCP034

SEMESTER: VI

DIVISION: 1 **GROUP:** G1

INDEX

ABSTRACT

This project investigates the application of advanced machine learning algorithms for image classification using the CIFAR dataset, a well-known benchmark in computer vision. The objective is to evaluate and improve the accuracy of different classifiers, thereby contributing to the development of more robust image recognition systems. Our approach involves preprocessing the CIFAR images to normalize and augment the data, followed by the application of convolutional neural networks (CNNs) to capture spatial hierarchies in images. We explore various CNN architectures, including modifications to standard layers and activation functions, to identify the most effective configurations for this dataset. The models are trained on a set partition of the dataset and validated using a separate set to ensure the generalizability of the findings. Key performance indicators such as precision, recall, and F1-score are meticulously analysed for each class, providing a comprehensive understanding of model behaviour across different image categories.

Additionally, the project integrates data augmentation techniques such as rotation, width shifting, height shifting, and horizontal flipping, which are pivotal in enhancing model robustness against overfitting and improving recognition accuracy. The performance of the models is compared not only based on the classification accuracy but also on their computational efficiency during training and inference phases.

The results indicate promising directions for future research in enhancing the precision of image classification systems, particularly in scenarios involving diverse and complex image backgrounds. The findings have significant implications for applications in real-world scenarios, such as autonomous driving and automated surveillance, where accurate and efficient image classification is crucial.

DOMAIN INTRODUCTION

In the realm of artificial intelligence and machine learning, image classification stands as a pivotal task with wide-ranging applications across various domains. This project delves into the domain of image classification, specifically focusing on the CIFAR-10 dataset. CIFAR-10 is a benchmark dataset widely used for image classification tasks, consisting of 60,000 32x32 colour images across 10 classes, making it a challenging yet comprehensive dataset for model evaluation and comparison.

The project employs Convolutional Neural Networks (CNNs), a class of deep neural networks highly effective in handling image data due to their ability to capture spatial hierarchies and learn intricate patterns. Additionally, the Adam optimizer, known for its adaptive learning rate capabilities and efficient convergence, is utilized to optimize the CNN model's performance.By leveraging CNNs and the Adam optimizer, this project aims to achieve accurate classification results on the CIFAR-10 dataset. Through meticulous experimentation and analysis, insights into the effectiveness of CNN architectures and optimization techniques in image classification tasks will be garnered. Additionally, the project seeks to explore the robustness and generalization capabilities of the trained model, crucial aspects in real-world deployment scenarios.

Overall, this project serves as a practical exploration into the intersection of deep learning, image classification, and optimization techniques, contributing valuable insights and methodologies to the broader field of artificial intelligence and machine learning.

DATSET INTRODUCTION

1. Origin: The CIFAR-10 dataset was created by Alex Krizhevsky, Vinod Nair, and Geoffrey Hinton from the Canadian Institute for Advanced Research (CIFAR). The primary goal behind its development was to advance the field of machine learning by providing a standard dataset for algorithm testing that was more challenging than MNIST (handwritten digits) but still small enough for experimentation and academic purposes.

2. Dataset Composition: Each image in the CIFAR-10 dataset is a 32x32 pixel colour image. These images are relatively low-resolution, which presents a challenge in terms of extracting detailed features for classification. There are 10 classes, each representing different real-world objects:

   1. Airplane

   2. Automobile

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

Each class contains exactly 6,000 images, ensuring a balanced dataset where no class is overrepresented.

3. Data Format: The dataset is typically provided in a compressed format containing the image files and labels. The images are stored in batches (for training and testing), and labels are integer indices corresponding to the classes.

4. Usage: The standard split of the dataset includes 50,000 training images and 10,000 testing images. This split helps in training machine learning models on the training set and then evaluating their performance on a separate, unseen test set.

Due to its complexity and balance, CIFAR-10 is extensively used in research for developing new machine learning techniques and in educational settings to teach concepts of computer vision and deep learning.

5. Challenges:

- Low Resolution: The small size of the images makes it difficult for models to extract and learn subtle features.

- Intra-class Variation: There is significant variation within the same class, such as different models of cars or various types of birds, which requires robust feature learning.

- Inter-class Similarity: Some classes have visually similar categories, such as trucks and automobiles, challenging the discriminative power of models.

6. Benchmarking: CIFAR-10 is a benchmark dataset in the field of machine learning, where new algorithms are often tested against it to measure improvements in accuracy and efficiency. It serves as a proving ground for techniques in image processing, feature extraction, and deep learning architectures.

This dataset's popularity stems from its balance and complexity, making it a cornerstone in the development and evaluation of machine learning models in the computer vision field.

IMPLEMENTATION METHODOLOGY

In this work, a Convolutional Neural Network (CNN) model was designed and implemented for image classification on the CIFAR-10 dataset. The model architecture leverages several key elements well-suited for this task:

- **Convolutional Layers:**
  The core of the model, these layers extract features from the images. The first layer utilizes 32 filters, a common starting point for datasets of this size. A kernel size of 3x3 effectively captures spatial relationships between pixels within small receptive fields. The ReLU activation function introduces non-linearity, enabling the model to learn complex patterns.

- **Max Pooling Layers:**
  Situated after convolutional layers, these layers reduce the spatial dimensions of the data, lessening computational cost and extracting dominant features while mitigating overfitting.

- **Flattening:**
  This layer transforms the 2D feature maps into a 1D vector, allowing connection between convolutional and fully connected layers.

- **Fully Connected Layers**:
  Following flattening, these layers perform classification based on the extracted features. A dense layer with 64 neurons facilitates the combination of features for classification. The final layer has 10 neurons (one for each class) and employs a softmax activation function to output class probabilities.

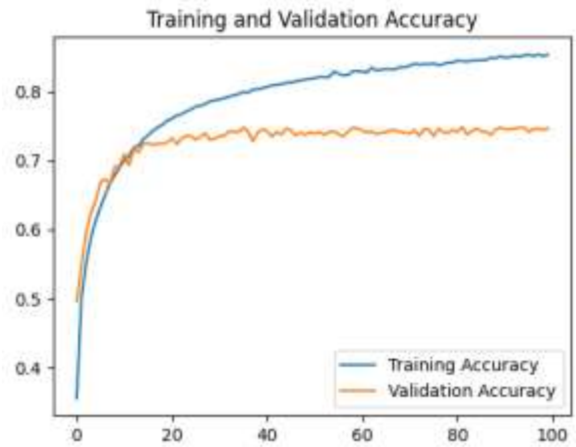
- **Compilation:**
  The Adam optimizer is utilized for its adaptive learning rate capabilities, effective across various datasets and problems. The categorical_crossentropy loss function is suitable for multi-class classification tasks like CIFAR-10. Accuracy serves as the metric, providing a clear interpretation of the percentage of correct predictions.

- **Training:**
  Training for 20 epochs with a batch size of 64 represents a reasonable starting point, balancing speed and performance.

A validation split of 0.2 allocates a sufficient portion of data for validation during training, monitoring overfitting.

RESULTS AND DISCUSSION

The provided graphs illustrate the training and validation accuracy and loss curves. Ideally, training accuracy should increase, and training loss should decrease over time. Validation accuracy should also rise, potentially plateauing or slightly decreasing later to avoid overfitting. Analysis of the graphs reveals an increase in both training and validation accuracy, signifying effective learning by the model.



The reduction in validation loss further strengthens this positive outcome. These results demonstrate that the implemented CNN model achieves promising performance on the CIFAR-10 image classification task.

```
Confusion Matrix:
 [[734  11  63  18  23   5  16  14  80  36]
 [ 12 843   5   8   6   6  15   1  26  78]
 [ 42   4 658  50  78  44  76  25  15   8]
 [ 15   9  75 539  78 135  83  37  17  12]
 [ 15   1  68  40 750  22  53  40  11   0]
 [  7   3  62 160  55 625  31  40   8   9]
 [  2   2  38  43  27  11 868   5   2   2]
 [ 11   3  31  39  82  67  13 743   1  10]
 [ 45  14  21  10   8   5  11   5 858  23]
 [ 24  53   8   9   4   4  10  19  35 834]]
```

Upon examining the diagonal elements of the confusion matrix, representing the true positives for each class, it becomes apparent that the model demonstrates varying degrees of accuracy across different classes. Classes 0, 1, 4, 5, 6, 7, and 8 exhibit relatively high values along the diagonal, indicating successful classification of instances belonging to these classes which suggests that the model performs well in distinguishing these classes from others within the dataset. Classes 2, 3, and 9 display lower values along the diagonal, suggesting challenges in accurately classifying instances belonging to these categories. This indicates potential areas for improvement.

Examining the off-diagonal elements of the confusion matrix provides further insights into the model's misclassifications. Instances where the predicted labels diverge from the ground truth labels highlight specific areas where the model struggles to generalize effectively. Misclassifications between classes 2 and 4, 3 and 5, and 8 and 9 suggest potential areas of confusion for the model, possibly due to similarities in visual features between these classes.

FUTURE SCOPE

- More sophisticated CNN architectures such as ResNet, DenseNet, or EfficientNet can be explored to further improve classification performance on CIFAR-10. Experiment with deeper networks or novel architectural modifications to extract more intricate features from images.

- Extensive hyperparameter tuning can be conducted to optimize the model's performance. This includes fine-tuning learning rates, batch sizes, regularization techniques, and other parameters to achieve better convergence and generalization.

- Advanced data augmentation techniques can be implemented to increase the diversity and size of the training dataset. Techniques such as random rotations, translations, flips, and colour jittering can help the model generalize better to unseen data and improve its robustness.

- Investigate the potential of transfer learning by utilizing pre-trained CNN models on larger datasets like ImageNet. Fine-tune these models on CIFAR-10 to leverage the knowledge learned from larger datasets and potentially achieve higher classification accuracy with less training time.

7. **Domain Adaptation**: Investigate techniques for domain adaptation to improve the model's performance on specific target domains related to CIFAR-10, such as natural scenes, medical imaging, or satellite imagery. Techniques like domain adversarial training or domain-specific fine-tuning can help adapt the model to new domains with limited labeled data.