Image classification with CNN

Prepared by Endework Abera

1. **Introduction**

   The project aimed to develop a deep learning model for image classification using the CIFAR-10 dataset. This dataset is a collection of 60,000 32x32 color images, evenly distributed across 10 different classes. The objective was to train a model that could accurately classify these images into their respective categories. The project utilized convolutional neural networks (CNNs), a type of deep learning model particularly effective for image analysis tasks due to their ability to capture spatial hierarchies of features.

2. **Data Exploration**

   The CIFAR-10 dataset was thoroughly explored to understand its structure and characteristics. The dataset comprises 10 classes: airplane, automobile, bird, cat, deer, dog, frog, horse, ship, and truck. Each image is 32x32 pixels in size and has three color channels (RGB). Summary statistics were computed, visualizations were created, and the distribution of classes was analyzed to gain insights into the dataset. Challenges observed included the relatively low resolution of images, which could make it harder for the model to distinguish between subtle features, and the presence of diverse object poses and backgrounds, which could add to the complexity of the classification task.

3. **Data Preprocessing**

   Preprocessing steps were applied to prepare the data for model training. The images were resized to ensure consistent input dimensions across all images. Normalization was performed to scale pixel values to a range suitable for neural network training, typically 0-1. These preprocessing steps aimed to enhance model convergence and performance by providing it with standardized and appropriately scaled input data.

4. **Model Architecture**

   The deep learning model was built using a convolutional neural network (CNN) architecture. The model consisted of multiple convolutional layers for feature extraction, interspersed with max-pooling layers for spatial down-sampling. This was followed by dense layers for classification. Rectified Linear Unit (ReLU) activation functions were used in the convolutional and dense layers to introduce non-linearity without the vanishing gradient problem. A softmax activation function was used in the output layer to facilitate multi-class classification, providing a probability distribution over the 10 classes.

5. **Model Training**

The model was trained on the CIFAR-10 training set using stochastic gradient descent (SGD) as the optimizer. The number of epochs, batch size, and learning rate were carefully chosen to balance training speed and model convergence. During training, accuracy and loss metrics were monitored on both the training and validation sets to assess model performance, guide hyperparameter tuning, and detect any signs of overfitting or underfitting.

6. **Results**

The trained model demonstrated promising results on the test set, achieving a high accuracy rate. Evaluation metrics, including precision, recall, F1-score, and confusion matrices, were computed to provide a detailed analysis of the model's performance across different classes. Sample predictions were visualized, showcasing instances where the model successfully identified objects as well as instances where it faced challenges. For example, the model had difficulty distinguishing between some classes that share similar features, such as horse and deer.

7. **Conclusion**

In conclusion, the image classification project successfully achieved its objectives. The developed model demonstrated competence in accurately classifying images from the CIFAR-10 dataset. The use of a CNN architecture, combined with appropriate preprocessing techniques and careful model training, contributed to the model's effectiveness.

8. **Future Work**

To further enhance the project, future work could involve experimenting with more complex CNN architectures, such as ResNet or Inception, to improve feature extraction capabilities. Advanced data augmentation techniques could be explored to increase the diversity of the training data and potentially improve model generalization. Transfer learning approaches could also be investigated, leveraging pre-trained models on larger datasets to improve performance. Additionally, addressing specific challenges identified during model evaluation, such as confusion between certain classes, could lead to improvements in overall performance.