Report: Training a CNN Model with CIFAR-10

By Ignacio Cabrera Lugo, Roberto J. González Hernández, Jesús Hernández

CNN Architecture Description

The proposed model is a Convolutional Neural Network (CNN) designed to classify the 10 classes of the CIFAR-10 dataset. The architecture consists of three main blocks of convolutional layers followed by pooling layers, ending with a fully connected layer and a softmax output layer. The main components are:

1. **Convolutional Layers**: Each block contains two convolutional layers with (3, 3) filters and ReLU activations, combined with Batch Normalization to stabilize learning.
2. **Pooling**: MaxPooling2D is used to reduce spatial dimensionality after each convolutional block.
3. **Regularization**: Includes L2 regularization in the convolutional layers and Dropout (0.5) in the fully connected layer to prevent overfitting.
4. **Global Average Pooling Layer**: Reduces dimensions before the fully connected layers.
5. **Output Layer**: A Dense layer with softmax activation to classify the 10 classes of the dataset.

Preprocessing Steps

1. **Normalization**: Pixel values were scaled to the range [0, 1] by dividing by 255.
2. **Data Splitting**: The training and test sets provided by CIFAR-10 were used.
3. **Data Visualization**: 10 random images per class were selected to visually inspect the data distribution.

Training Process Details

- **Optimizer**: Adam with a learning rate of 0.0005.
- **Loss Function**: SparseCategoricalCrossentropy with from_logits=True.
- **Metrics**: Accuracy.
- **Batch Size**: 64.
- **Number of Epochs**: 50 (with Early Stopping).
- **Callbacks**: EarlyStopping with a patience of 10 epochs to stop training if validation loss does not improve.

Results and Model Performance Analysis

Training and Validation Metrics

- **Final Training Accuracy**: 94.49%
- **Final Validation Accuracy**: 73.42%
- **Test Accuracy**: 83.32%
- **Final Test Loss**: 0.8888

Confusion Matrix Analysis

The confusion matrix shows that "automobile" and "truck" classes have the highest prediction rates, while "cat" and "frog" have lower performance due to visual similarity with other classes.

Classification Report

Class	**Precision**	**Recall**	**F1-Score**
Airplane	87%	83%	85%
Automobile	93%	91%	92%
Bird	83%	72%	77%
Cat	64%	79%	71%
Deer	93%	69%	79%
Dog	80%	78%	79%
Frog	77%	93%	85%
Horse	83%	91%	87%
Ship	92%	87%	89%
Truck	90%	90%	90%
Average	**84%**	**83%**	**83%**

Confusion Matrix Visualization

The confusion matrix confirms key patterns and areas for improvement, such as confusion between "cat" and "dog."

Best Model and Reasons

The best model was obtained by restoring weights from the last epoch with the lowest validation loss, thanks to the EarlyStopping callback. This model is suitable due to its balance between training and validation accuracy, showing good generalization to unseen data.

Insights from the Experimentation Process

1. **Batch Normalization**: Significantly helped stabilize learning, especially with a relatively low learning rate.
2. **Regularization**: The combination of L2 and Dropout prevented overfitting.
3. **EarlyStopping Strategy**: Saved training time and prevented overtraining.
4. **Deep Structure**: The depth of three convolutional blocks was sufficient to capture complex patterns without vanishing gradient issues.

Visualizations and Charts

Key visualizations include:

Confusion Matrix: To analyze correct and incorrect predictions by class.