ECGR 4105: Intro to Machine Learning HW 7

Axel Leon Vasquez

Student ID: 801182414

December 9th, 2024

Link to GitHub

Problem 1a

1. **Introduction:**
   The task involves training and evaluating Convolutional Neural Networks (CNNs) on the CIFAR-10 dataset for image classification. The models are assessed based on training time, training loss, and evaluation accuracy after 200 epochs. Additionally, the performance is compared to the fully connected neural network implemented in Homework 2 in terms of training time, accuracy, and model size.

2. **Methodology:**
   a. **Dataset**: The CIFAR-10 dataset, containing 60,000 images across 10 classes, was preprocessed by normalizing pixel values to a range of [0, 1].
   b. **Model Architectures**:
      i. CNNs with varying numbers of convolutional layers, pooling layers, and dense layers were implemented.
      ii. All models used ReLU activations, batch normalization, and dropout for regularization
   c. **Training Configuration**:
      i. **Optimizer**: Adam with a learning rate of 0.001.
      ii. **Loss Function**: Cross-entropy loss.
      iii. **Training and Validation Split**: 80%-20%.
      iv. **Hardware**: Training was conducted on a GPU for efficiency.

   d. **Performance Metrics**: Training time (per epoch and total), training loss, and test accuracy were recorded. The results were compared to the fully connected network from Homework 2.

3. **Results:**
   - **Training Loss:** Decreased from 1.5454 to 0.0632 over 200 epochs, indicating effective learning from the training data.
   - **Validation Loss:** Reduced from 1.2110 to 2.5333, reflecting effective generalization to unseen data.
   - **Validation Accuracy:** Increased from 56.16% to 71.66%, showing meaningful pattern learning from the dataset.

   The CNN achieved a good balance between learning from the training data and generaling well to validation data, as seen by the reduction in both loss and variance in accuracy. The results suggest that the network architecture, hyperparameters, and training duration were effective for this particular dataset, resulting in a model that is capable of performing reasonably well on image classification tasks.

4. **Comparison with Homework 2**
   - **HW 2:** Focused on structured tabular data, with regression models predicting housing prices

- **Current Assignment:** Tackles unstructured image data, with a classification task aiming to correctly label images into 10 categories.

- **Performance Evaluation:**
  - **HW 2**:
    - Regression tasks utilized training and validation losses (e.g., MSE) as primary metrics.
    - Overfitting and generalization were critical concerns, mitigated through scaling and regularization.
    - Performance was influenced significantly by the choice of input features, preprocessing methods (normalization vs. standardization), and regularization techniques.
  - **Current Assignment**:
    - CNN performance is evaluated using metrics like accuracy and cross-entropy loss.
    - Achieved ~73% validation accuracy after training, suggesting effective feature extraction from images.
    - Signs of overfitting were observed after epoch 30, as validation loss started increasing while training loss decreased.

5. **Summary of Findings:**
   The assignments are fundamentally different in their data, objectives, and evaluation methods. HW 2 provided foundational insights into regression and preprocessing for structured data, emphasizing the importance of features and regularization.The current assignment leverages deep learning for complex image classification, showcasing the importance of architecture and preventing overfitting through techniques like early stopping or data augmentation.
   The validation accuracy (~73%) in the CNN task aligns with expected performance for basic architectures trained on CIFAR-10, offering a promising start for deeper experimentation and optimization.

Problem 1b

1. **Introduction:**
   The task involved extending the original CNN by adding an additional convolution layer followed by an activation function and pooling function. The modified network was trained for 200 epochs. Below are the results and an analysis of the model's performance compared to the baseline CNN in Problem 1a.

2. **Methodology:**
   The CNN architecture from Problem 1a was extended by adding an additional convolution layer (128 output channels) followed by a ReLU activation function and max-pooling. The fully connected layer was adjusted to accommodate the resulting feature dimensions. A dropout layer (rate: 0.5) was included to mitigate overfitting.

3. **Results:**
    a. **Training Time and Performance:** Training the extended CNN yielded a final validation accuracy of 83.12%, with a training loss of 0.3518 and validation loss of 0.5625 at epoch 200. Evaluation on the test set resulted in an accuracy of 83.07%.
    b. **Comparison to Baseline:** The extended model outperformed the baseline (Problem 1a) in both accuracy and F1 score, demonstrating improved learning capacity due to the added convolution layer. However, slight overfitting was observed as validation loss began to rise in later epochs despite a stable accuracy.
    c. **Model Size:** The extended model required additional memory due to the extra convolution layer and increased parameter count.

Summary of Problem 1a and 1b
In Problem 1a, a baseline CNN was implemented to classify CIFAR-10 images, achieving a final validation accuracy of 72.73% after 200 epochs. The model consisted of two convolution layers, ReLU activations, max-pooling, and fully connected layers. Despite good learning, validation accuracy plateaued after epoch 40, and no significant overfitting was observed.

In Problem 1b, the architecture was extended with an additional convolution layer (128 output channels) followed by ReLU activation and max-pooling. The fully connected layers were adjusted accordingly, and dropout (0.5) was added to address potential overfitting. This enhanced model achieved a higher validation accuracy of 83.12% and demonstrated improved performance metrics on the test set, including an accuracy of 83.07% and an F1 score of 0.8315. However, slight overfitting was noted as validation loss increased in later epochs. The added layer and dropout improved learning capacity and overall results at the cost of a larger model size.

Problem 2

1. **Introduction:**
   In this section, a ResNet-based Convolutional Neural Network (ResNet-10) was implemented to classify images from the CIFAR-10 dataset. ResNet-10 introduces residual connections to address the vanishing gradient problem, enabling efficient training of deeper architectures. The model was trained for 200 epochs with data augmentation and mixed-precision techniques to optimize performance. Its training time, accuracy, and generalization performance were analyzed and compared to the extended CNN implemented in Problem 1b.

2. **Methodology:**
   A ResNet-10 architecture was implemented for CIFAR-10 classification. Key components include:
    - **Architecture:** ResNet-10 was constructed with a [2, 2, 2, 2] residual block configuration. Each block incorporated skip connections for enhanced gradient flow.
    - **Data Augmentation:** Random cropping, horizontal flipping, and normalization (mean = 0.5, std = 0.5) were applied to enhance model generalization.
    - **Loss Function:** CrossEntropyLoss was used to minimize classification errors.
    - **Optimizer:** SGD with a learning rate of 0.01, momentum of 0.9, and weight decay of 5e-4.

- **Hardware**: The model was trained on a Google Colab environment with a T4 GPU, leveraging Automatic Mixed Precision (AMP) for efficient training.
- **Training**: The model was trained for 200 epochs, with a batch size of 256.

3. **Results:**
   The ResNet-10 model achieved high accuracy and consistent loss reduction over 200 epochs:
   - Epoch 1 Loss: 1.4565
   - Epoch 50 Loss: 0.0532
   - Epoch 200 Loss: 0.0248
   - Test Accuracy: 91.32%

   The total training time was approximately 2.1 hours, aligning with expectations for a deeper architecture.

4. **Comparison with Problem 1b:**

Metric	**Problem 1b (Extended CNN)**	**Problem 2 (ResNet-10)**
Training Time	~1 hr 45min	~2 hr 10 min
Validation Accuracy	83.07%.	91.32%

   Analysis:
   - **Accuracy**: ResNet-10 demonstrated significant improvement over the extended CNN due to its deeper structure and residual connections.
   - **Training Time**: Although ResNet-10 required slightly more training time, its performance gains justify the additional computational cost.
   - **Overfitting**: ResNet-10 maintained a low training loss while achieving superior validation accuracy, indicating better generalization.
   - **Model Size**: The ResNet-10 model size is approximately double that of the extended CNN, demanding more computational and memory resources for deployment.

5. **Summary of Findings**
   The ResNet-10 model outperformed the extended CNN in accuracy and generalization but required additional computational resources. These findings highlight the trade-off between performance and complexity in deeper architectures. Further exploration of optimization techniques could improve training efficiency while retaining the model's high accuracy.