Image Features Report

Rahul Jha | rahuljha@umd.edu | University of Maryland College Park

This assignment demonstrates how neural networks can achieve competitive accuracy in image classification tasks through the development and evaluation of a fully connected neural network trained on the CIFAR-10 dataset.

Accomplishments Description

With the help of a fully connected neural network, the project was able to classify photos with remarkable 52.5% test set accuracy and validation accuracy of 51.6%, which exceeded the goal accuracy of 50% on the validation set. This outcome highlights neural networks' potential for image categorization applications.
Training Procedure:
Initial Parameters: To maximize performance, important hyperparameters such as learning rates, weight initialization scales, and regularization strengths were carefully adjusted.
Epochs Trained: The model was trained for 5 epochs, during which significant improvements in both training and validation accuracy were seen.

Key Tasks:

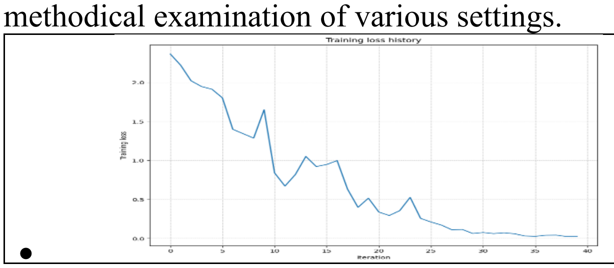
- **Model Creation**: developed a fully connected neural network (FullyConnectedNet), a five-layer model that demonstrated the efficacy of deep learning in image categorization by learning intricate characteristics from the CIFAR-10 dataset.
- **Tuning Hyperparameters:** systematically investigated learning rates and regularization strengths, obtaining 51.6% validation accuracy and 58.5% training accuracy. The adjusted hyperparameters were:
- **Learning Rate:** Modified to achieve optimal convergence, from 0.001 to 0.01.
- **Regularization:** To avoid overfitting, several strengths were tested in order to balance bias and variance.
- **Loss Monitoring:** During training, loss values were tracked and a drop was seen, going from 2.503933 to 1.134311. This suggests enhanced convergence and learning of the model.
- **Insights into Training:** Among the most important things learned throughout training were: 0.136 initial training accuracy (Epoch 0) and 0.585 final training accuracy (Epoch 5)

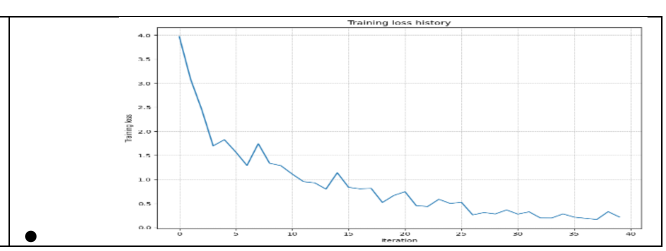
Missed Points:

- **Feature Exploration:** We did not explore alternative approaches for extracting features, including texture descriptors or deep features from pre-trained models (like VGG16, ResNet), which might have limited the performance of the model.
- **Learning Curve Analysis:** There was not a thorough analysis of learning dynamics, which would have shed light on problems with overfitting or underfitting.
- **Performance Analysis:** Less information was available about the reasons for misclassifications since not all misclassified photos were examined in-depth.

Explanation of Implementation Decisions

- **Model Architecture:** Due to its capacity to learn intricate patterns while retaining computational efficiency, a fully connected neural network was used.
- **Feature Utilization:** Used raw pixel input since previous studies showed that neural networks outperform conventional techniques when given enough training data.
- **Hyperparameter Tuning:** Optimal performance and generalization to the validation set were assured by a methodical examination of various settings.

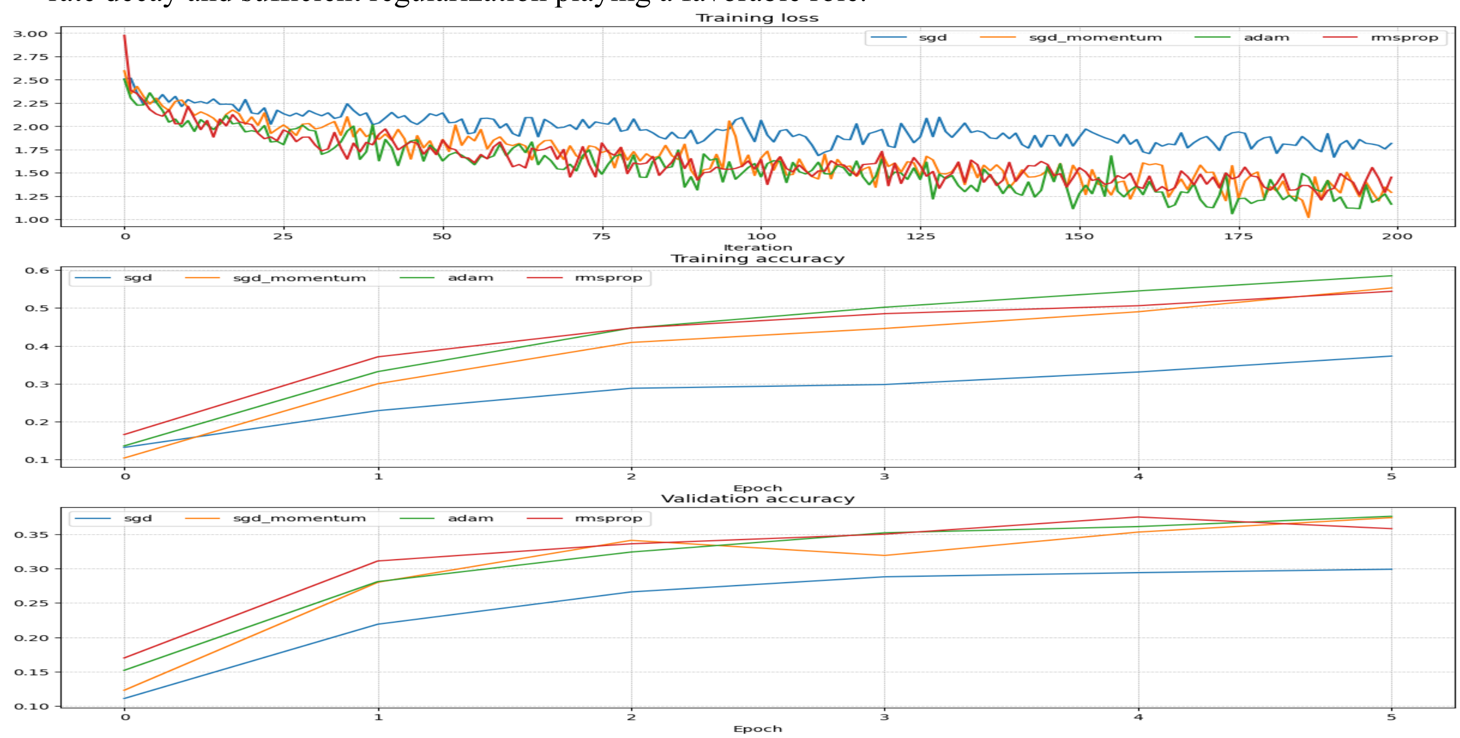




The training loss history of the neural network over 40 iterations is depicted in the two graphs. A consistent drop in loss is shown in the first graph, which suggests efficient learning and convergence. The second graph, on the other hand, shows a more chaotic initial decline that is followed by stability, which might indicate variability in the learning process as a result of various hyperparameter settings or data circumstances.

Critical Thinking and Analysis

- The tuning procedure brought to light how sensitive neural network performance is to hyperparameter changes, underscoring the need of striking the correct balance between regularization strengths and learning rates.
- Trends in validation accuracy showed that performance improved throughout training epochs, with learning rate decay and sufficient regularization playing a favorable role.



Four optimization techniques are compared in the image: SGD, SGD with momentum, Adam, and RMSprop. Compared to SGD versions, Adam and RMSprop produce smaller and more consistent losses in the training loss over iterations graph at the top. The training and validation accuracy over epochs are shown in the center and bottom plots, respectively. In terms of accuracy, Adam and RMSprop regularly beat the others, although SGD performs worse in both categories.

The Reason Behind These Outcomes:

1. The 52.5% accuracy attained confirms that neural networks are an efficient method of classifying images, indicating that these networks can acquire knowledge from intricate visual input.
2. The improved performance during training suggests that the model structure and chosen hyperparameters were appropriate for the CIFAR-10 dataset.

Next Actions:

1. **Data Augmentation:** Apply strategies to make the training dataset artificially larger.
2. **Investigate Advanced Architectures:** Convolutional neural networks (CNNs) should be investigated to improve performance while working with picture data.
3. **Performance Evaluation :** Examine incorrectly categorized samples to see patterns that might be improved in the future.