Executive Summary Neural Systems

This study presents a comparative analysis of Artificial Neural Networks (ANN) and Convolutional Neural Networks (CNN) using the Fashion MNIST dataset, focusing on both baseline models and improved models. The objective was to evaluate model performance on image classification tasks and to investigate how architectural enhancements influence accuracy and generalization.

The Fashion MNIST dataset, including 70,000 grayscale images across ten apparel categories, was preprocessed through normalization and reshaping to suit both dense and convolutional network structures. The baseline ANN utilized fully connected layers with ReLU activations, while the CNN incorporated convolutional and pooling layers to exploit spatial hierarchies. Both CNN models have used Data Augmentation for the dataset input.

For the improved variants, dropout regularization and batch normalization were introduced to reduce overfitting and stabilize training. Activation functions are also the same as the baseline. Experimenting with different activation functions such as LeakyReLU did not result in the models improving. Adamax and SGD with momentum have also been tested, both failing to improve the models further however Nadam made an improvement on the ANN model. Callbacks such as Early Stopping and ReduceLROnPlateau were used to improve the models training time and efficiency, affecting the models positively. Hyperparameter tuning also improved the model further (especifically batch size). Finally, adding more layers of Convolution and Pooling made a noticeable positive impact on the models.

Empirical evaluation revealed clear performance trends. The baseline ANN achieved a test accuracy of ~88%, slightly outperforming the baseline CNN's ~87%, suggesting that for relatively simple architectures, dense networks can perform competitively. However, after applying architectural improvements, both models demonstrated measurable gains. The improved ANN reached ~90% test accuracy, whereas the improved CNN achieved the highest performance with ~93% accuracy and the lowest test loss of ~0.216. Validation accuracies followed similar patterns, confirming robust generalization for the enhanced CNN.

These findings underscore the efficiency of convolutional layers in capturing spatial features when combined with regularization techniques. While ANNs remain efficient for lower-complexity tasks, CNNs, particularly when refined with dropout and batch normalization, offer superior accuracy and stability. Overall, the improved CNN architecture provides the optimal balance between learning capacity and generalization, affirming its suitability for modern image classification applications.

P2669504 Periklis Lazarides