Executive Summary: Fashion-MNIST Classification Performance Analysis

Sharuk Shaik
P2904355

1. Overview: The present summary gives a glimpse of the major results brought about by the application of deep-learning algorithms on the Fashion-MNIST dataset (10 clothing types): Original ANN (Multi-Layer Perceptron), Original CNN, and their variants. The main purpose was to determine the architecture and methods producing the highest classification accuracy. The Improved CNN achieved a peak accuracy of 93.62%, thus confirming the superiority of convolutional layers combined with best-in-class regularization.

2. Models Used:

Adam optimizer and Cross-Entropy Loss were employed for training all models. The dataset was divided into 48000 training, 12000 validation, and 10000 testing samples.

Original Models (10 Epochs):

ANN (MLP): This model served as a benchmark and consisted of three Fully Connected (FC) layers (784→128→64→10).

CNN: A typical architecture employing two convolutional layers, Max Pooling layers, and two FC layers was used.

Improved Models (20 Epochs):

Improved ANN: It received a makeover of deeper FC layers (4 total), BatchNorm1d, and Dropout (0.5).

Improved CNN: It received a makeover with BatchNorm2d after convolutions, Dropout (0.25) in FC layers, and trained with Data Augmentation.

3. Core Results:

Performance Comparison:

Model	Architecture Type	Best Validation Accuracy	Final Test Accuracy
Original ANN	Fully Connected	88.64%	87.43%
Original CNN	Convolutional	91.17%	90.89%
Improved ANN	FC + Regularization	88.53%	87.35%
Improved CNN	CNN + Regularization/Augmentation	93.62%	91.55%

Key Findings:

CNN Superiority: The Original CNN outperformed the Original ANN by +3.46 percentage points in test accuracy (90.89% vs. 87.43%).

Optimal Performance: The Improved CNN obtained the highest accuracy, increasing the base CNN by +2.45 percentage points.

Diminishing Returns for ANN: Regularization methods could not make up for the basic shortcomings of the ANN, which could not cope with the spatial quality of the image data.

4. Key Innovations (Improved CNN):

The peak performance of the Enhanced CNN is linked with the strategic combination of modern techniques:

Data Augmentation: Randomly Horizontally Flipping and very little Rotating were done to the training images, which resulted in a dataset that was much larger and the model's ability to generalize was greatly improved.

Batch Normalization (BN): BatchNorm2d was used to normalize the signals of the network's layers so that training would be more stable and faster, thus making it easier and quicker to get close to the solution.

Dropout: A Dropout (0.25) ratio in the fully connected layers was considered as a strong regularizer, which prevented feature co-adaptation and thereby controlled overfitting.

Learning Rate (LR) Scheduling: A StepLR scheduler steadily decreased the learning rate by a factor of 0.1 every 7 epochs, ensuring that the weights were optimally fine-tuned towards the end of the training process.

5. Conclusion and Future Steps:

The Improved CNN has been evaluated as the best model for the Fashion-MNIST classification task through its effective and high performance.

Conclusion: Convolutional Networks are the basis for this visual task. 93.62% accuracy can be achieved only with the combination of these networks and Data Augmentation and Batch Normalization.

Insight: The most difficult classification remaining problems tend to be between items that look alike, for example, 'Shirt' and 'T-shirt/top' or 'Pullover' and 'Coat.'

Recommendation: The next step is to try the deeper state-of-the-art CNNs or transfer learning that could lead to more than 94% accuracy and less confusion between the most similar clothing items.