Neural Networks Coursework (Fashion-MNIST)

Executive Summary

This project compares two neural network architectures, an Artificial Neural Network (ANN) and a Convolutional Neural Network (CNN) for a image classification on the Fashion-MNIST dataset. This analysis evaluates their performance in terms of loss, accuracy and model efficiency. The CNN outperformed the ANN by approximately 3 to 4%, demonstrating ace feature extraction through convolutional operations while using fewer parameters.

Context and Purpose

Fashion-MNIST dataset is a good benchmark for image classification task it contains a grayscale images (28×28 pixels) of 10 types of clothing items.The goal of this coursework is simple.

- Implement and train both an ANN and a CNN on the same data-set.
- Compare their accuracy, loss, confusion matrices, and class wise performance.
- Understand how the convolutional layers improve model efficiency and feature learning for image data.

Methodology

- **Dataset.** In this I normalized images with the dataset mean/std, used a 90/10 train-val split from the 60k training set, and then kept the 10k test set untouched for a evaluation.

- **Training Setup.** I used Adam optimizer (lr=0.001), Cross Entropy loss, 18 epochs, batch size 128, and a ReduceLROnPlateau scheduler to pushed the learning rate down when validation accuracy delayed.

- **ANN Architecture.** Three hidden layers (512→256→128) with ReLU, batch norm, and dropout (0.3). It treats the image as a flat 784 dimensional vector fast to implement but oblivious to spatial layout.
- **CNN Architecture.** Three convolutional stages (1→32→64→128) with batch norm, max-pooling, and light dropout, followed by a small fully connected head. This setup preserves and draw on local structure in the image.

Key Findings

Metric	ANN	CNN
Train Accuracy	89%	94%
Validation Accuracy	88%	92%
Test Accuracy	88 – 89%	92%
Parameter Count	592,000	438,000
Improvement	—	+3 – 4%

- CNN consistently outperformed the ANN on validation and test sets despite having fewer parameters, while ANN struggled.

- Confusion matrix of CNN shows improved classification for “Sneakers,” “Trousers,” and “Bags.”
- Most confused classes were, ANN: *Shirt* ↔ *T-shirt* and CNN: *Shirt* ↔ *Coat / Pullover*

Discussion

The ANN throws away the exact thing that make images special the spatial relationships. While the CNN keeps those patterns intact and stacks them into high level cues. That is why it wins here, even with a fewer parameters. Still some classes are inherently ambiguous in 28×28 grayscale. A shirt with certain sleeves can look a lot like a T-shirt at this resolution, so errors there are not shocking.

Recommendations

- **Architectural Tweaks.** Trying a slightly deeper CNN or a small residual block. Even a compact ResNet-style module can improve gradients without inflating parameters.
- **Training Improvements.** Consider trying early stopping, a cyclical schedule, and a brief warm up phase. These often give smoother convergence.
- **Analysis & Explainability.** Visualize feature maps or Grad CAM on mistakes it’s useful to see what the network looked at when it got things wrong.

Conclusion

This study shows that the CNN outperform ANN on image classification tasks due to their spatial feature extraction capability. CNN perform better with fewer parameters. While ANN remains a simpler but less powerful measure for visual data.