Executive Summary

Student ID: P2926259

Student name: Muhammad Jahanzaib Awan

Fashion-MNIST Classification: Comparative Analysis of ANN and CNN Architectures

1. Overview

In this study, I have implemented and compared two neural network structures for Fashion Mnist dataset. It involves Artificial Neural Network (ANN) and a Convolutional Neural Network (CNN). The dataset includes 70,000 greyscale images of clothes (60,000 training and 10,000 testing) which is divided across 10 different categories. It represents a very difficult classification task due to high intra class variance and inter class similarities.

ANN Architecture: This architecture is a 4 layer feedforward network with 567,434 parameters, featuring hidden layers of 512, 256 and 128 neurons. It uses ReLU activation function and 30% dropout regularization to prevent overfitting.

CNN Architecture: This architecture is a hierarchical feature extraction network with 422,474 parameters which utilizes three convolutional blocks (32, 64, and 128 filters) with max pooling, batch normalization and fully connected classification layers. Now even though this architecture has 25% less parameters, it still shows superiority over ANN through its spatial feature learning.

2. Major Findings

Performance Comparison:

CNN achieved %92.88% test accuracy compared to ANN's %89.34 which represents a %3.54 improvement. This performance gap shows the superiority of convolutional architectures for spatial pattern recognition tasks.

- **Training Efficiency:** CNN converged by epoch 6 versus epoch 7 for ANN, which demonstrated faster learning through hierarchical feature extraction.
- **Generalization:** CNN also showed %3.70 train validation gap versus ANN's %2.20 indicating slightly higher model complexity while maintaining excellent generalization.
- **Class-Specific Performance:** Both of these models excelled at geometrically distinct items like (Bags: 98-99%, Sneakers: 96-98%, Trousers: 97-99%) but the struggled with visually similar garments.
- **Critical Challenge:** The most problematic classification was Shirt (ANN: %68, CNN: %78) with %12.7 of shirts misclassified as T-Shirts in ANN versus %8.2 in CNN.

3. Notable Innovations

Regularization Strategy: Implemented multi faceted overfitting prevention through %30 dropout, batch normalization in CNN layers and strategic 80-20 train validation split. This combination yielded robust models with minimal overfitting.

Optimizer Configuration: Utilizing Adam optimizer (lr=0.001) with StepLR scheduler (gamma=0.1 and epoch 10), resulting in %1.32 validation accuracy improvements for ANN and %0.77 for CNN post decay which demonstrates effective learning rates management.

Advanced Visualization Techniques: Developed comprehensive model interpretability through CNN filter visualization, feature map analysis across three convolutional layers, interactive Plotly training curves and confusion matrix heatmaps. Feature maps revealed progressive abstraction from edge detection (Conv1) to complex pattern recognition (Conv3).

Data Augmentation Framework: Designed scalable augmentation pipeline with light (rotation +-10 degrees horizontal flip) and heavy (rotation +-15 degrees, translation +-10%, random erasing) strategies, providing foundation for futute performance enhancement (currently demonstrated but not applied to maintain baseline comparison.

4. Analysis and Insights

Why CNN Outperforms ANN:

The CNN's superiority over ANN stems from three fundamental advantages e.g Spatial Hierarchies: convolutional layers capture local patterns (edges, textures) and progressively combine them into a complex representations (garment shapes, styles) Translation: invariance, shared filter weights detect features regardless of spatial positionwhich is crucial for varied garment placements, Parameter Efficieny: local connectivity and weight sharing reduce parameters while increasing representational capacity which resulted in %25 fewer parameters yet %4 better performance.

Confusion Pattern Analysis:

Inter-class confusion clustering reveals systematic challenges e.g Upper body garments (shirt to t-shirt to coat to pullover) from a confusion cluster with 6 to 13% misclassification rates due to similar silhouettes and neckline variations, Footwear (sneaker, ankle boot, sandal) demonstrates minimal confusion (<%2) due to distinctive solve patterns, Accessories (bag) achieve a near perfect classification (98-99%) due to unique structural geometry.

Training Dynamics:

Both models exhibited healthy learning curves with rapid early improvement (epoch 1-6) followed by fine tuning phase. The %0.52 validation test gap for CNN indicates excellent generalization without validation set overfitting. Learning rate decay at epoch 10 provided beneficial fine tuning, particularly for ANN(%1.32 improvement) which demonstrated proper hyper scheduling effectiveness.

5. Nutshell

This comparative study conclusively demonstrates CNN architectural superiority for image classification tasks, achieving %92.88 accuracy versus ANN's %89.34 while using %25 fewer parameters. The success validates fundamental deep learning principles e.g hierarchical feature learning, spatial inductive biases, and parameter sharing mechanisms are essential for visual recognition tasks.

Key Technical Contributions: Comprehensive architectural comparison with rigorous evaluation across accuracy, confusion patterns and training dynamics; Advanced regularization combining dropout, batch normalization and learning rate scheduling; Model interpretability through feature visualization revealing hierarchical pattern extraction.

Recommended Improvements: Data augmentation integration (projected +1-2% accuracy); Ensemble methods combining multiple CNN architectures; Attention mechanisms for difficult classes(shirt, coat) Transfer learning from larger fashion dataset; Advanced architectures (ResNet style skip combinations) for deeper networks.

Learning Outcomes Achieved: Successfully implemented end to end deep learning pipeline including dataset preprocesing, model architecture design, training optimization, performance evaluation and result visualization. Gained practical experience in PyTorch framework, regularization techniques, hyperparameter tuning and critical analysis of competing architectures. Developed strong understanding of why CNNs excel at computer vision through empirical validation and theoretical analysis.