Fashion-MNIST Image Classification Report

1. Executive Summary

This paper examines and compares the results of two deep learning models: Artificial Neural Network (ANN) and Convolutional Neural Network (CNN) for image classification of fashion items from the Fashion-MNIST dataset images. The task was to develop the model to classify grayscale images automatically representing fashion products like t-shirts, footwear, dresses, and bags. Each model was developed with the help of PyTorch.

Although the ANN is learning with the flat pixel information and processing the image as a one-dimensional vector, the CNN is able to extract spatial information like edges, textures, and shapes before finally proceeding to generate predictions. The results obtained have shown that CNNs perform significantly better than ANNs because of their capability to extract hierarchical features. CNN's test accuracy stood at 88.49%, while that of ANN was 83.81%.

This report also covers data preparation for datasets, model structure, training parameters, and comparison results. The paper concludes with recommendations on how to improve results, such as data augmentation methods, learning rate control, and transfer learning.

2. Technical Implementation

2.1 Dataset and Preprocessing

It includes 70,000 grayscale images with dimensions 28×28 pixels. It is categorized into 10 fashion types. These include T-shirt/top, trouser, pullover, dress, coat, sandal, shirt, sneaker, bag, and ankle boots. It is normalized to the range 0-1 and is then divided into training, validation, and testing sets.

Dataset	Samples	Purpose
Training	60,000	Model learning
Validation	12,000	Hyperparameter tuning
Testing	10,000	Final performance evaluation

2.2 Model Designs

Two models have been developed and contrasted:

• ANN - feed-forward network that takes all the image's pixel information in as a vector.

• CNN - Convolutional Neural Network - extracts local spatial information before classification

Also, the CNN architecture involves convolutional layers and pooling to acquire hierarchical patterns. Thus, it reinforces learning during training.

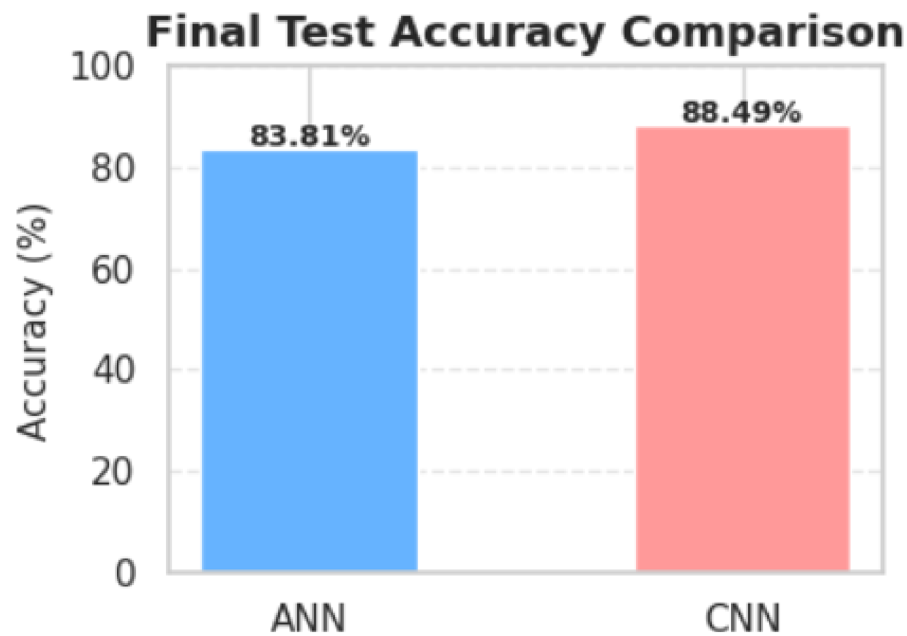
2.3 Training Setup

Parameter	Value
Optimizer	Adam
Loss Function	Cross Entropy
Learning Rate	0.001
Batch Size	64
Epochs	15
Hardware	CPU / GPU
Framework	PyTorch

2.4 Results and Evaluation

Model	Test Accuracy	Remarks
ANN	83.81%	Slower, less accurate, limited spatial understanding
CNN	88.49%	Faster convergence, better image feature learning

```
=== Summary ===
ANN Test Accuracy: 83.81%
CNN Test Accuracy: 88.49%
```



The CNN performed better in the validation and testing tasks with improved learning and generalization. The networks had instances where they misclassified items with similar patterns like shirts and coats. The CNN successfully minimized errors in such cases.

2.5 Comparison of Models

Feature	ANN	CNN
Accuracy	83.81%	88.49%
Speed	Medium	Fast
Feature Extraction	Weak	Strong
Overfitting	Slight	Minimal
Visual Understanding	Low	High

Thus, in conclusion, CNNs perform better in image-based tasks than ANNs because they retain spatial information and hierarchical features.

2.6 Possible Improvements

• • Implement new data augmentation techniques such as random rotations, flips, and changes in brightness.

• Dropout layers and batch normalization for better regularization.

• Adaptive learning rate schedulers is implemented.

• Pre-trained models such as VGG16, ResNet, or EfficientNet to improve performance the transfer learning is applied.

3. Conclusion

In this experiment it confirms that the CNNs are more effective than ANNs for image Classification task on the Fashion-MNIST dataset. By leveraging spatial feature extraction and hierarchical learning, CNNs achieve higher accuracy and faster training.