Executive Summary

This study presents an end-to-end implementation of artificial neural networks (ANN) and convolutional neural networks (CNN) for image classification using the Fashion-MNIST dataset. The goal was to evaluate how architecture influences predictive accuracy and generalisation performance on low-resolution image data. The dataset contains 70,000 grayscale images across ten categories such as T-shirt/top, trouser, pullover, dress, coat, sandal, shirt, sneaker, bag, and ankle boot. The experiment employed TensorFlow/Keras, drawing methodological insights from Preda (2021) on Kaggle.

Methodology Overview

The dataset was normalised to [0,1], with 50,000 training, 10,000 validation, and 10,000 test samples. The ANN flattened inputs into 784 features, using dense layers of 512, 256, and 128 neurons (ReLU activations, dropout=0.3). The CNN preserved spatial structure with convolutional layers (32, 64, 128 filters, 3×3 kernels) and max pooling, followed by batch normalisation, a dense layer (128 neurons), dropout (0.4), and a 10-class softmax output. Both used Adam optimisation (lr=0.001) with early stopping.

Results and Key Findings

The CNN achieved 91.0% test accuracy, outperforming the ANN's 88.9%. Validation curves show faster convergence and reduced loss variance for the CNN. The ANN required 22 epochs to stabilise, while the CNN achieved optimal performance after 9 epochs.

Interpretation and Recommendations

The CNN demonstrated superior spatial feature extraction, particularly in distinguishing similar garments such as shirts, coats, and pullovers. Both models exhibited strong performance on simpler classes (e.g., trousers, sandals, bags). The confusion matrices reveal persistent misclassifications in upper-body garments, highlighting dataset complexity. Future improvements should include image augmentation, learning-rate scheduling, and deeper networks to reach state-of-the-art accuracy.

Conclusion

The CNN achieved better generalisation and reduced validation error, confirming the strength of convolutional inductive bias. While both models achieved robust performance, the CNN's accuracy advantage and stability make it preferable for real-world classification tasks. This aligns with findings by Preda (2021) and LeCun et al. (1998).

References

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278–2324. https://doi.org/10.1109/5.726791

Preda, G. (2021). Fashion-MNIST classification using Keras CNN. Kaggle. https://www.kaggle.com/code/gpreda/fashion-mnist-with-keras-cnn

Xiao, H., Rasul, K., & Vollgraf, R. (2017). Fashion-MNIST: A novel image dataset for benchmarking machine learning algorithms. arXiv preprint arXiv:1708.07747.