Executive Summary

The spatially aware capabilities of Convolutional Neural Networks are usually pitted against traditional fully connected network designs within the context of this field of visual classification. This paper addresses that dichotomy through a far-reaching benchmarking of the ANNs and CNNs on the challenging Fashion MNIST dataset, isolating specifically the impact of critical architectural refinements on model efficacy.

The first step was to establish two baseline models on the 70,000 grayscale images: a foundational ANN composed simply of dense, fully-connected layers and a minimal CNN containing basic convolutional and pooling layers. Curiously, the less complex Baseline ANN achieved a test accuracy of 88.4%, improving on the performance of the Baseline CNN with 87.4%. This outcome arguably indicates that for fairly simple feature sets, any immediate advantage derived from convolutions is not especially evident at this early stage, and a well-configured dense network can perform in a competitive manner.

However, the core purpose of this coursework was to investigate capacity scaling and generalization. To that end, both architectures were augmented with best-practice techniques: Batch Normalization to stabilize training gradients and speed up convergence, while Dropout aimed at preventing overfitting and improving the ability of the model to generalize when faced with previously unseen data.

These changes brought forth clear results. The Improved ANN performance ticked up to 89.6%, showing better stability. Most importantly, the Improved CNN jumped massively, establishing itself as the best classifier with a peak test accuracy of 92.5% and attaining the minimum test loss of 0.216.

This difference in performance is a certain verification of the fact that, while ANNs can process flattened image data, the mechanism behind CNNs in feature mapping-especially the capability to learn hierarchical spatial features such as edges, textures, and ultimately object shapes-is demonstrably superior. When this mechanism is coupled with strong regularization, the true potential of the network is unleashed. The paper concludes by noting that the refined CNN architecture provides the required balance between deep learning capability and high generalization for robust image classification, setting the validated standard for visual recognition tasks.