Executive Summary

Traditional fully connected network designs are commonly pitted against the spatially aware capabilities of Convolutional Neural Networks within the context of this field of visual classification. This work speaks to that dichotomy by performing an extensive benchmarking of both ANNs and CNNs on the challenging Fashion MNIST dataset, specifically isolating 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 objective of this coursework was to explore capacity scaling and generalization. In this regard, both architectures were enhanced with best-practice techniques: Batch Normalization was introduced to stabilize training gradients and accelerate convergence, while Dropout was implemented to avoid overfitting and enhance the model's ability to generalize when faced with previously unseen data.

These modifications led to clear results. The Improved ANN performance ticked up to 89.6%, indicating better stability. Most importantly, the Improved CNN leaped massively, establishing itself as the best classifier with a peak test accuracy of 92.5% and achieving the minimum test loss of 0.216.

This performance difference conclusively verifies the fact that, though ANNs can process flattened image data, the mechanism behind CNNs in feature mapping, especially the ability 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 that the refined CNN architecture provides the required balance between deep learning capability and high generalization for robust image classification, thus setting the validated standard for visual recognition tasks.