

Fashion-MNIST Image Classification – Final Report

Generated: 2025-11-04 09:16:14

Overview: This report summarizes training and evaluation of two architectures (MLP & CNN) on the Fashion-MNIST dataset using TensorFlow/Keras. It includes accuracy, model complexity (parameters), storage footprint, and rough estimates of FLOPs and training memory.

Model	Test Accuracy	Trainable Parameters	Saved Model Size (MB)	FLOPs (Training)	FLOPs (Inference)
MLP	0.8687	235,146	2.72 MB	0.46M	0.23M
CNN	0.8796	56,714	0.69 MB	2.00M	1.00M

Conclusion:

- The CNN achieved higher test accuracy, thanks to spatial feature extraction via convolution and weight sharing, while keeping parameter count and saved size lower than the MLP.
- The MLP is simpler and has fewer FLOPs per inference in this setup, but it discards spatial structure by flattening, which typically limits image classification performance.
- For image tasks, CNNs are generally superior due to learning hierarchical, translation-aware features with fewer parameters.

Notes: Reported FLOPs are rough academic estimates for this specific architecture. Actual runtime cost depends on hardware, libraries, batch size, and kernel implementations. Values marked "N/A" indicate the session lacked those variables/files at generation time.