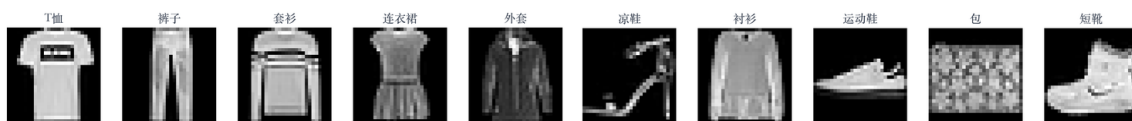


2026 年秋季学习《深度学习与计算机视觉》课程实验

从线性分类器到多层感知机（对应课程 3-4 讲）

廖振宇



Fashion-MNIST 十个类别的图像示例，来源：[官方数据仓库](#)

1 实验内容

1. **数据与任务**：在 Fashion-MNIST 数据集上完成十分类任务。将官方训练集固定划分为 50 000 个训练样本和 10 000 个验证样本，官方测试集只用于评价。
2. **线性基线**：根据课上描述的内容，训练线性 Softmax 分类器，记录训练与验证曲线、测试准确率、混淆矩阵、逐类召回率和典型错误可视化。
3. **反向传播**：仅使用 NumPy 在固定 mini-batch 上实现单隐藏层 MLP 的 forward 与 backward，并用中心差分完成梯度检查。（无需使用 NumPy 训练完整数据集。）
4. **受控比较**：使用 PyTorch 比较线性 Softmax 分类器、两个线性层但中间没有激活函数，以及加入 ReLU 的 MLP。后两个模型使用相同隐藏层宽度，其余训练条件尽量保持一致。
5. **训练诊断**：选择一种实际出现的训练问题，先提出原因判断，再只改变一个关键因素进行验证，例如学习率、初始化、正则化或 Early Stopping。

2 实验要求

- 实验由每位学生独立完成，不组队，不共享代码、训练结果或报告
- 必须提交：梯度检查结果、三模型性能表、训练与验证曲线、混淆矩阵、逐类召回率和至少一组典型错误
- 提交内容包括可运行源代码、README 说明文件（说明环境、运行命令、配置、随机种子等）与 PDF 实验报告；截止时间与提交入口以课程平台通知为准
- 评价重点是实现正确性、比较是否受控、证据是否完整以及分析是否合理

3 相关链接与参考文献

1. Fashion-MNIST 原始论文：Xiao, Rasul, Vollgraf, “[Fashion-MNIST: a Novel Image Dataset for Benchmarking Machine Learning Algorithms](#),” 2017
2. [官方数据与说明](#)与[Torchvision 数据接口](#)