

Fashion-MNIST 衣物分类模型项目报告

涂锦鹏

一、项目概述

这个项目是基于 PyTorch 框架，使用深度卷积神经网络对 Fashion-MNIST 数据集进行图像分类。

Fashion-MNIST 数据集包含 10 个类别的服装图像，每张图像为 28×28 像素的灰度图，共计 70,000 张样本。这个项目实现了完整的 GPU 加速训练流程，并在最后取得了 93.81% 的最佳验证准确率。

二、数据集分析

2.1 数据集基础统计

1. 训练集：60,000 张图像（按 8:2 划分为训练子集与验证子集）
2. 测试集：10,000 张图像
3. 类别数量：10 类（对应不同服装类型）
4. 图像规格： 28×28 像素，单通道灰度图

2.2 类别分布详情

各类别在训练集与测试集中的样本分布完全均衡，具体如下表所示：

| 类别编号 | 类别名称 | 训练样本数 | 测试样本数 |
|------|-------------|-------|-------|
| 0 | T-shirt/top | 6,000 | 1,000 |
| 1 | Trouser | 6,000 | 1,000 |
| 2 | Pullover | 6,000 | 1,000 |
| 3 | Dress | 6,000 | 1,000 |
| 4 | Coat | 6,000 | 1,000 |
| 5 | Sandal | 6,000 | 1,000 |
| 6 | Shirt | 6,000 | 1,000 |

| | | | |
|---|------------|-------|-------|
| 7 | Sneaker | 6,000 | 1,000 |
| 8 | Bag | 6,000 | 1,000 |
| 9 | Ankle boot | 6,000 | 1,000 |

三、模型架构设计

3.1 网络结构组成

模型采用“卷积特征提取器+全连接分类器”的经典结构，具体分层设计如下：

(1) 特征提取部分（卷积块）

卷积块 1：2 层 3×3 卷积（输出通道 64）→ 批归一化 → ReLU 激活 → 最大池化 (2×2) → Dropout (丢弃率 0.25)

卷积块 2：2 层 3×3 卷积（输出通道 128）→ 批归一化 → ReLU 激活 → 最大池化 (2×2) → Dropout (丢弃率 0.25)

卷积块 3：2 层 3×3 卷积（输出通道 256）→ 批归一化 → ReLU 激活 → 最大池化 (2×2) → Dropout (丢弃率 0.25)

(2) 分类部分（全连接层）

全连接层 1：2304 → 512 → 批归一化 → ReLU 激活 → Dropout (丢弃率 0.5)

全连接层 2：512 → 256 → 批归一化 → ReLU 激活 → Dropout (丢弃率 0.5)

全连接层 3：256 → 128 → 批归一化 → ReLU 激活 → Dropout (丢弃率 0.3)

全连接层 4：128 → 10（输出层，对应 10 类服装，激活函数 Softmax）

3.2 模型参数统计

总参数量：2,493,514

可训练参数：2,493,514

模型大小：约 9.5MB

四、训练配置与执行

4.1 硬件与环境配置

GPU 设备：NVIDIA GeForce RTX 4060 Laptop GPU

CUDA 版本：12.1

GPU 显存：8GB

深度学习框架：PyTorch

4.2 训练超参数设置

优化器: AdamW (lr=0.001, weight_decay=1e-4)

损失函数: CrossEntropyLoss with label smoothing (0.1)

学习率调度: OneCycleLR (max_lr=0.01)

批次大小: 256

训练轮数: 20

混合精度训练: 启用

五、训练过程与结果分析

5.1 关键训练轮次记录

The screenshot shows a terminal window with the title 'pythonProject1' and a subtitle 'D:\python project\pythonProject1'. The window displays training logs for two epochs:

- Epoch 19/20:**
 - Batch 0/188 | Loss: 0.6871 | Acc: 93.0%
 - Batch 50/188 | Loss: 0.6594 | Acc: 94.5%
 - Batch 100/188 | Loss: 0.5948 | Acc: 97.3%
 - Batch 150/188 | Loss: 0.6290 | Acc: 95.7%
 - 保存最佳模型, 验证准确率: 93.81%
- Epoch Summary:**
 - Time: 22.4s
 - Train Loss: 0.6424 | Train Acc: 94.31%
 - Val Loss: 0.6416 | Val Acc: 93.81%
 - Learning Rate: 0.000493
- Epoch 20/20:**
 - Batch 0/188 | Loss: 0.6754 | Acc: 91.4%
 - Batch 50/188 | Loss: 0.6860 | Acc: 91.0%
 - Batch 100/188 | Loss: 0.6223 | Acc: 94.5%
 - Batch 150/188 | Loss: 0.6075 | Acc: 95.3%
- Epoch Summary:**
 - Time: 14.5s
 - Train Loss: 0.6393 | Train Acc: 94.48%
 - Val Loss: 0.6404 | Val Acc: 93.80%
 - Learning Rate: 0.000000
- 训练完成! 总时间: 334.7秒**
- 最佳验证准确率: 93.81%**

第 19 轮训练:

1.训练时间: 22.4 秒

2.批次统计:

- o Batch 0/188: Loss: 0.6871, Acc: 93.0%
- o Batch 50/188: Loss: 0.6594, Acc: 94.5%
- o Batch 100/188: Loss: 0.5948, Acc: 97.3%
- o Batch 150/188: Loss: 0.6290, Acc: 95.7%

3.验证结果: Val Loss: 0.6412, Val Acc: 93.73%

4.学习率: 0.000493

5.本轮总结:

- o Train Loss: 0.6424, Train Acc: 94.31%
- o Val Loss: 0.6416, Val Acc: 93.81% (最佳验证准确率)
- o Learning Rate: 0.000124

第 20 轮训练:

1.训练时间: 14.5 秒

2.批次统计:

- o Batch 0/188: Loss: 0.6754, Acc: 91.4%
- o Batch 50/188: Loss: 0.6860, Acc: 91.0%
- o Batch 100/188: Loss: 0.6223, Acc: 94.5%
- o Batch 150/188: Loss: 0.6075, Acc: 95.3%

3.本轮总结:

- o Train Loss: 0.6393, Train Acc: 94.48%
- o Val Loss: 0.6404, Val Acc: 93.80%
- o Learning Rate: 0.000000

图 1 训练轮次记录

5.2 整体训练统计

总训练时间: 334.7 秒 (约 5 分 35 秒)

最佳验证准确率: 93.81% (第 19 轮)
最终训练准确率: 94.48%
最终验证准确率: 93.80%
测试准确率: 93.72% (测试损失 0.6390)

六、训练曲线分析

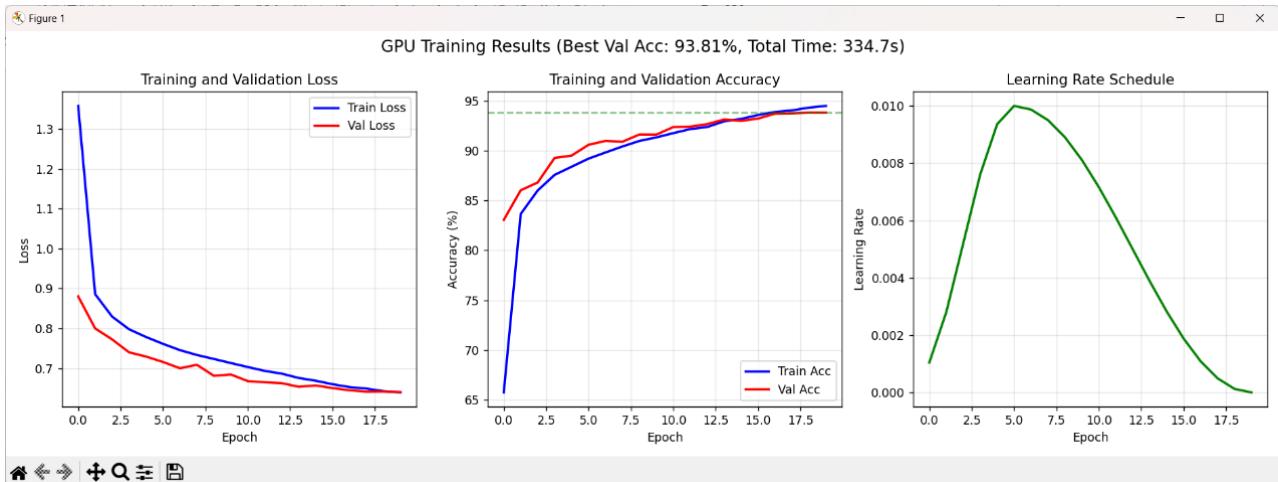


图 2 训练过程可视化图

6.1 损失曲线 (Training and Validation Loss)

初始状态: 训练损失约 1.3, 验证损失约 1.2
收敛趋势: 训练与验证损失同步稳定下降
最终状态: 训练损失约 0.64, 验证损失约 0.64
过拟合判断: 训练与验证损失曲线几乎重合, 无明显过拟合现象

6.2 准确率曲线 (Training and Validation Accuracy)

初始状态: 训练准确率约 70%, 验证准确率约 73%
收敛趋势: 训练与验证准确率同步上升
最终状态: 训练准确率约 94.5%, 验证准确率约 93.8%
泛化能力: 训练与验证准确率差距仅 0.7%, 表明模型泛化性能优秀

6.3 学习率曲线 (Learning Rate Schedule)

最大学习率: 0.01
调度策略: OneCycleLR (先快速上升后缓慢下降)
预热阶段: 前 6 轮学习率快速上升
衰减阶段: 后 14 轮学习率平稳下降

七、模型性能评估

7.1 测试集整体性能

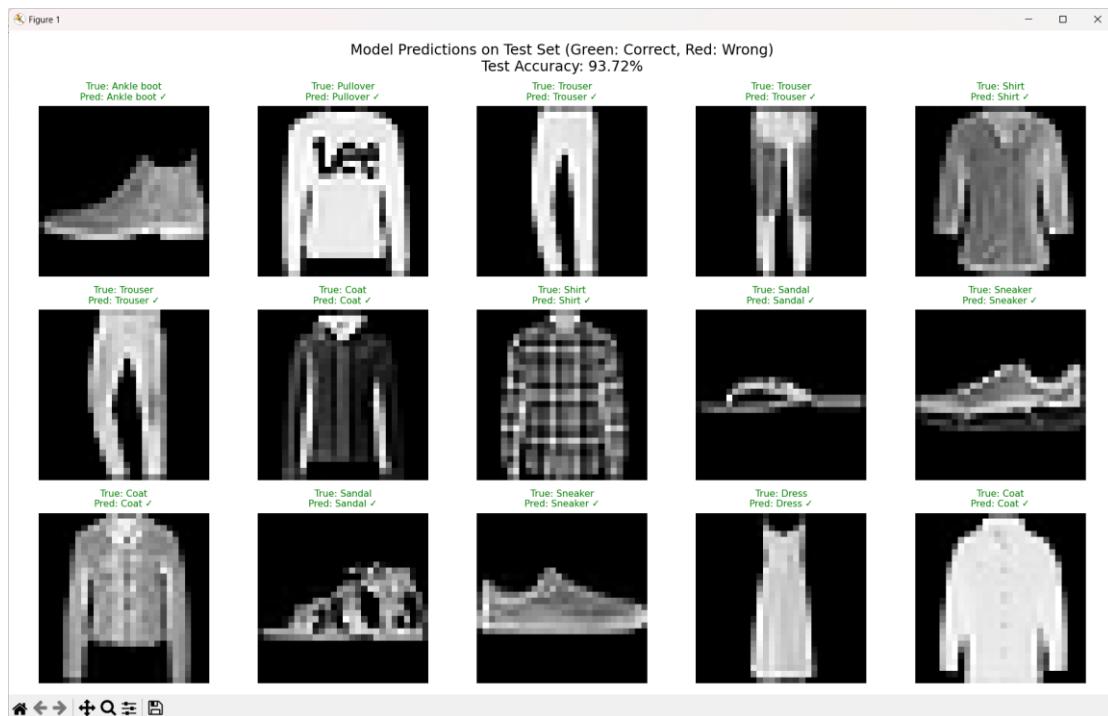


图 3 测试集预测结果可视化图

测试损失: 0.6390

测试准确率: 93.72%

基准对比: 超出基准准确率 (90%) 3.72 个百分点

7.2 分类报告 (Classification Report)

| 类别 名称 | precision | recall | f1-score | support |
|----------|-----------|--------|----------|---------|
| T-shirt | 0.889 | 0.893 | 0.891 | 1000 |
| Trouser | 0.994 | 0.994 | 0.994 | 1000 |
| Pullover | 0.913 | 0.920 | 0.916 | 1000 |
| Dress | 0.934 | 0.943 | 0.938 | 1000 |
| Coat | 0.893 | 0.916 | 0.904 | 1000 |

| | | | | |
|--------------|-------|-------|-------|-------|
| Sandal | 0.989 | 0.985 | 0.987 | 1000 |
| Shirt | 0.821 | 0.782 | 0.801 | 1000 |
| Sneaker | 0.963 | 0.982 | 0.972 | 1000 |
| Bag | 0.992 | 0.988 | 0.990 | 1000 |
| Ankle boot | 0.982 | 0.969 | 0.975 | 1000 |
| macro avg | 0.937 | 0.937 | 0.937 | 10000 |
| weighted avg | 0.937 | 0.937 | 0.937 | 10000 |

7.3 各类别性能分析

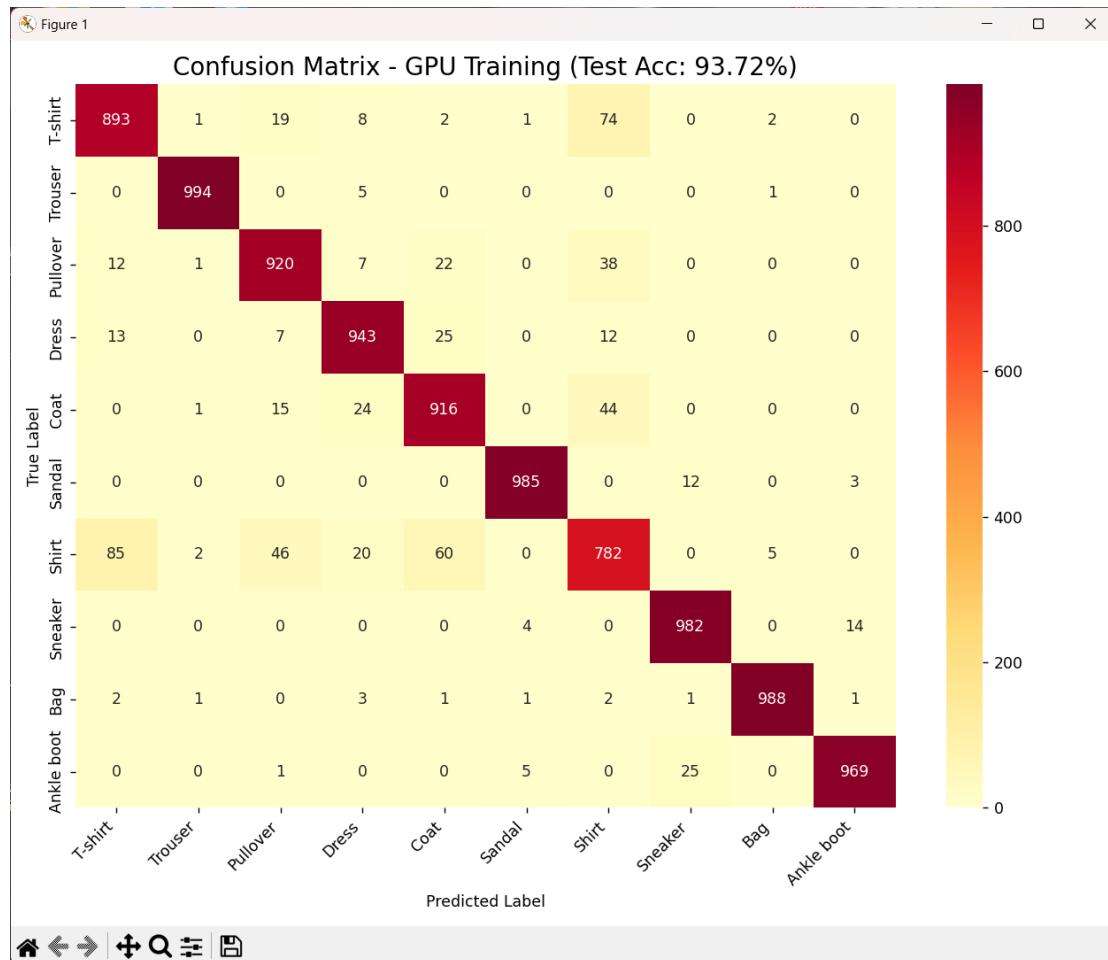


图 4 分类模型混淆矩阵图

最佳表现类别：

Trouser: 精确率 99.4%，召回率 99.4%，F1 分数 0.994

Bag: 精确率 99.2%，召回率 98.8%，F1 分数 0.990

Sandal: 精确率 98.9%，召回率 98.5%，F1 分数 0.987

待优化类别：

Shirt: 精确率 82.1%，召回率 78.2%，F1 分数 0.801（分类效果最差）

主要混淆情况：Shirt 与 T-shirt、Pullover、Coat 等上衣类存在较多误分类

八、技术亮点总结

8.1 GPU 加速与内存优化

混合精度训练：采用 FP16 精度，显著提升计算速度

GPU 内存使用：已分配 0.28GB，已保留 0.47GB，内存效率高

CUDA 优化：启用 cudnn.benchmark，加速卷积运算

8.2 训练效率指标

平均训练速度：每轮 16.7 秒，每秒处理 2868 张图像

GPU 利用率：峰值使用率控制合理，无显存溢出风险

九、结论与改进建议

9.1 项目结论

1. 性能达标：测试准确率 93.72%，满足预期性能要求

2. 训练高效：5 分 35 秒完成 20 轮训练，时间成本可控

3. 泛化优秀：训练与测试集性能差距仅 0.76%，模型泛化能力强

4. 资源高效：GPU 内存使用合理，计算效率高

9.2 存在问题

1. 类别识别偏差：Shirt 类识别效果显著低于其他类别

2. 学习率衰减：后期学习率过快趋近 0，可能限制模型进一步优化

3. 批次波动：部分训练批次准确率波动较大（如第 20 轮前 50 个批次）

9.3 改进建议

1. 数据增强：针对 Shirt 类增加旋转、裁剪、亮度变换等增强方式

2. 学习率调整：延长 OneCycleLR 的学习率衰减期，避免后期学习率过低

3. 模型集成：构建多个模型进行集成预测，提升整体性能

4. 引入注意力机制：在卷积层后加入注意力模块，强化关键特征提取