

神经网络和深度学习

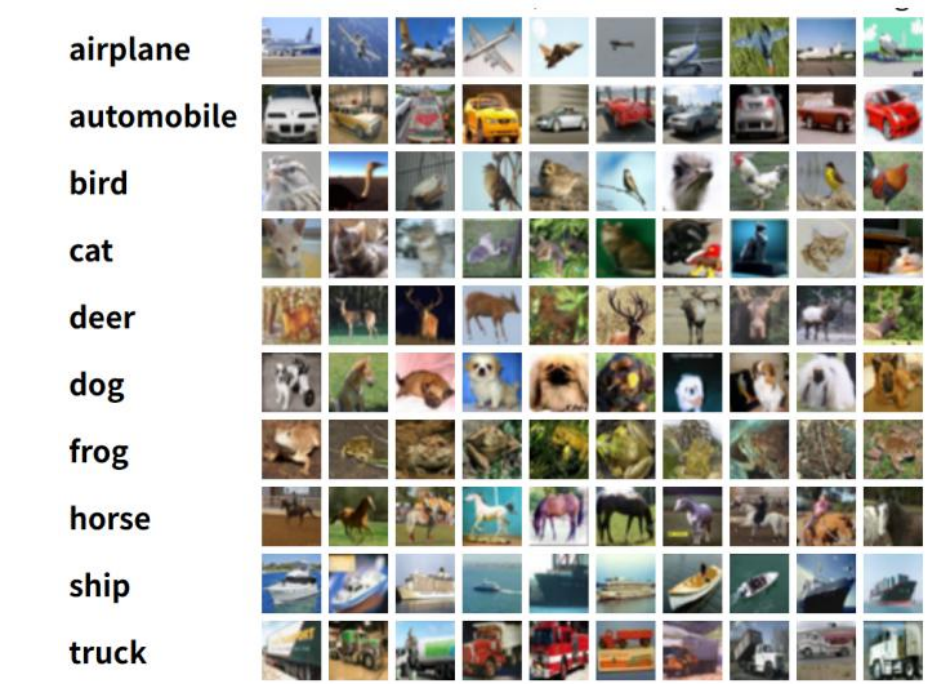
陈俊辉(24210980138)

本实验通过从零开始构建三层全连接神经网络分类器，在 CIFAR-10 数据集上实现了图像分类任务。基于 NumPy 手动实现了反向传播算法，设计支持自定义隐藏层大小与激活函数的模块化模型架构，并结合随机梯度下降（SGD）优化器、交叉熵损失函数、L2 正则化及学习率下降策略完成模型训练。通过调节学习率、隐藏层维度、正则化强度等超参数，系统探究了其对模型性能的影响，并基于验证集结果自动保存最优权重。实验结果表明，最佳配置下模型在测试集上达到 **50.82%** 的分类准确率，训练过程中损失曲线收敛平稳，验证集准确率随迭代逐步提升。

数据集 CIFAR 10

1.1CIFAR-10 数据集介绍

CIFAR-10 (Canadian Institute for Advanced Research 10-Class Dataset) 是计算机视觉领域广泛使用的基准图像分类数据集，由 Alex Krizhevsky、Vinod Nair 和 Geoffrey Hinton 整理发布。该数据集包含 10 个类别的彩色图像，分别为飞机 (airplane)、汽车 (automobile)、鸟类 (bird)、猫 (cat)、鹿 (deer)、狗 (dog)、青蛙 (frog)、马 (horse)、船 (ship) 和卡车 (truck)，每个类别包含 6000 张图像，总计 60,000 张。数据集按标准划分为 50,000 张训练图像和 10,000 张测试图像，所有图像的分辨率为 32×32 像素，采用 RGB 三通道格式。



1.2 数据集处理

在实验初始阶段，本研究对 CIFAR10 数据集实施了三个关键性预处理操作，具体技术流程如下：

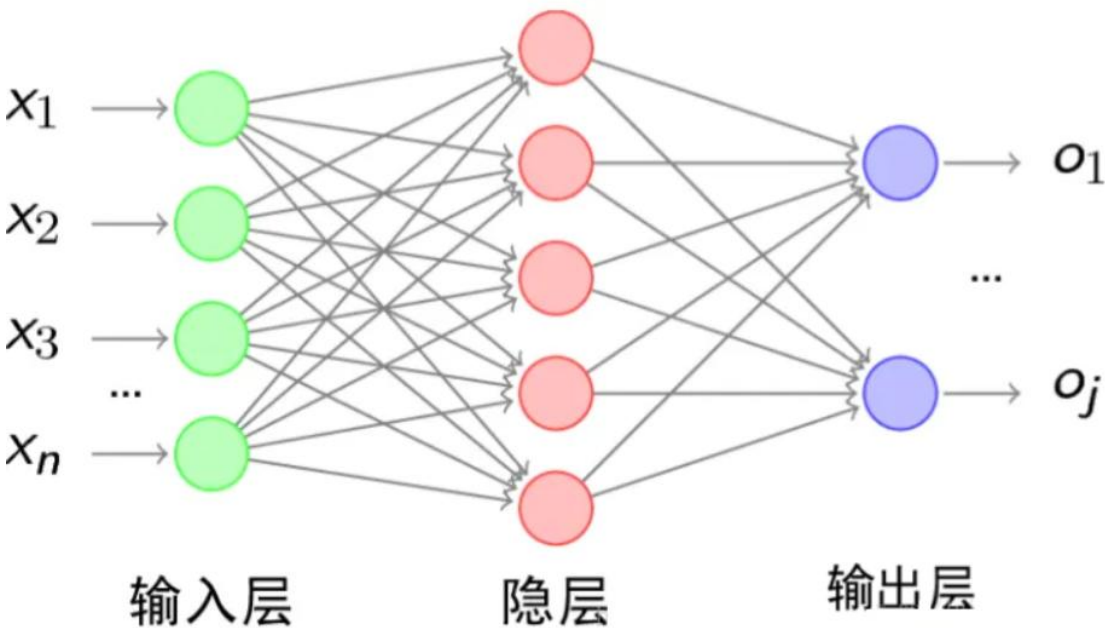
数据标准化：对图像 RGB 通道数值实施线性变换，通过将原始 0-255 整数范围的像素强度值统一除以标量系数 255，实现特征值域向[0,1]区间的映射，该操作有效提升了梯度计算的数值稳定性。

特征重构：将原始图像的三维空间结构（ $3 \times 32 \times 32$ ）通过维度重组操作转化为单一维度特征表达，生成 3072 维的输入向量，这种数据重塑方式为后续全连接网络提供了适配的输入格式。

概率化数据分割：采用蒙特卡洛随机抽样方法，按照 9:1 的样本配比动态划分训练样本，建立包含验证机制的数据集结构。具体而言，每个训练样本以伯努利分布概率 $p=0.9$ 归入训练子集，余下 $p=0.1$ 则划入验证子集。

全连接神经网络

全连接神经网络 **FCNN**（**Fully Connected Neural Network**），又称多层感知机（**Multilayer Perceptron, MLP**），是一种基础且广泛应用的深度学习模型。其核心特征在于网络中的每一层神经元均与相邻层的所有神经元通过权重参数全连接，能够通过非线性变换逐步提取输入数据的抽象特征，最终完成分类或回归任务。



2.1 网络结构

输入层: 接收原始数据（如 CIFAR-10 的 $32 \times 32 \times 3$ 像素图像，输入维度为 3072）。

隐藏层: 由若干全连接层构成，每层通过线性变换（权重矩阵乘法与偏置相加）和非线性激活函数（如 ReLU、Sigmoid）实现特征变换。

输出层: 根据任务类型设计，分类任务中通常使用 Softmax 函数输出类别概率分布。

2.2 前向传播与反向传播

前向传播: 数据从输入层逐层传递至输出层，每层的计算为：

$$z^{(l)} = W^{(l)}a^{(l-1)} + b^{(l)}, a^{(l)} = f(z^{(l)})$$

其中， f 为激活函数， $a^{(l)}$ 为第 l 层的输出。

反向传播: 基于链式法则计算损失函数对权重和偏置的梯度，通过优化器（如 SGD）更新参数，实现损失最小化。

2.3 激活函数的作用

ReLU: 缓解梯度消失，加速收敛（常用隐藏层）。

Softmax: 输出层归一化为概率分布（多分类任务）。

Sigmoid/Tanh: 早期网络常用，易导致梯度饱和，现多被 ReLU 替代。

实验

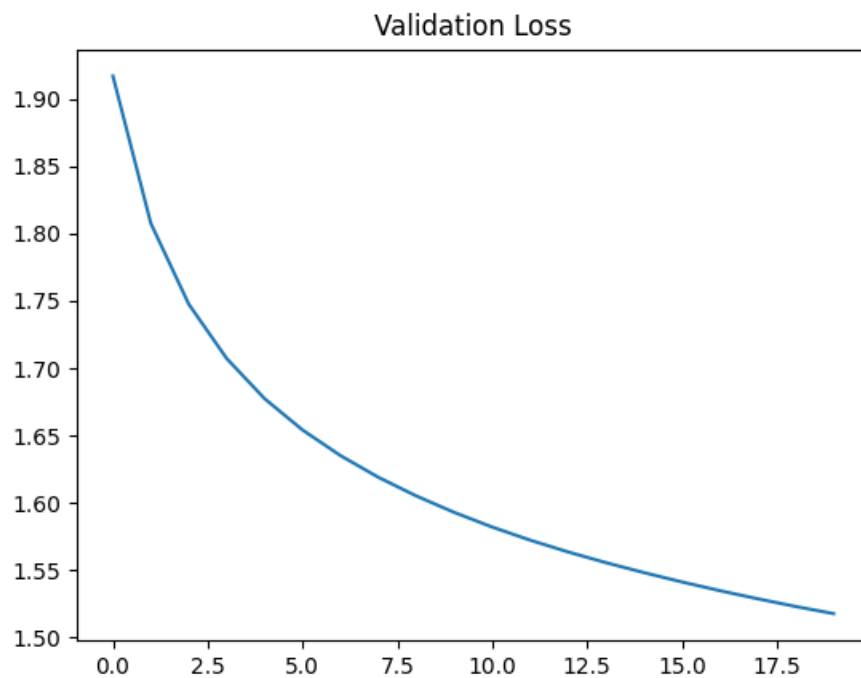
超参数选择

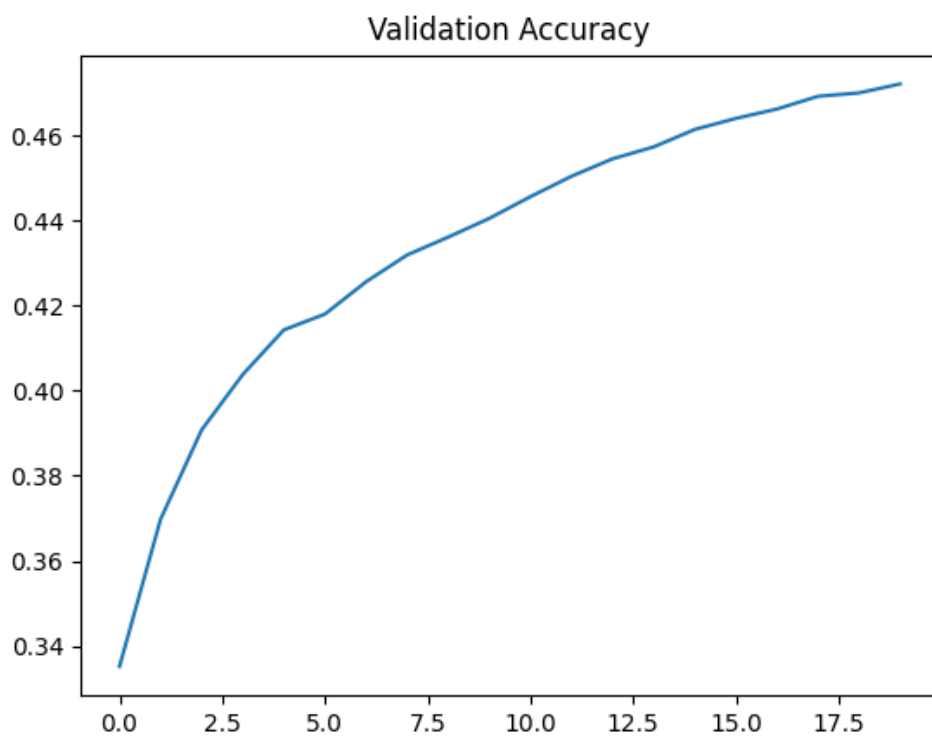
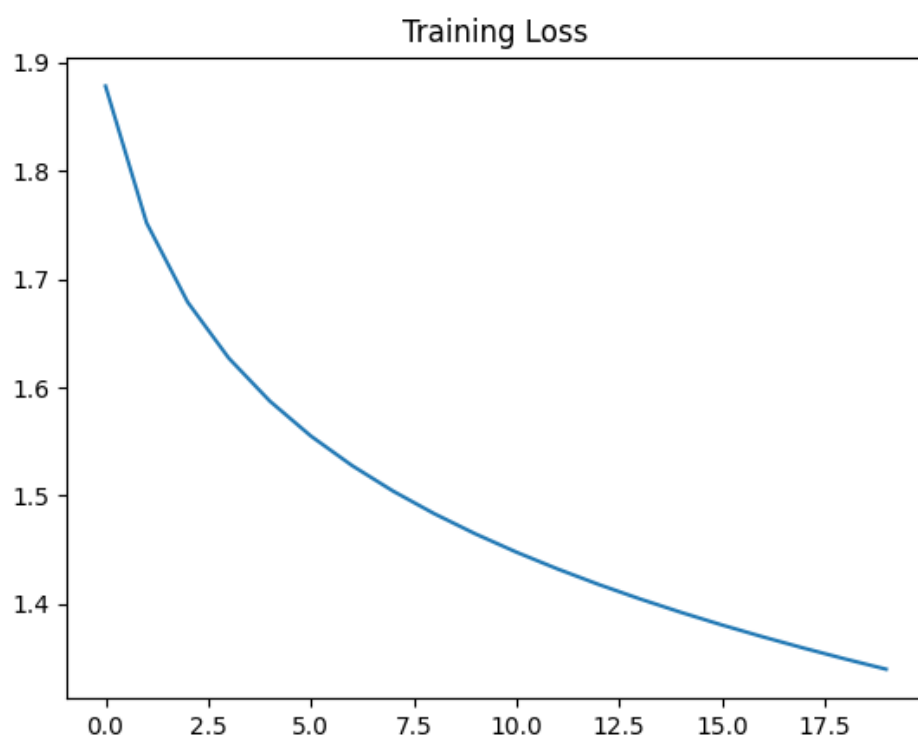
在本研究中，我们构建了超参数优化系统并实施参数寻优策略。实验基准配置维持恒定参数：批量样本量（**batch size**）固定为 64，迭代轮次（**epoch**）预设为 50 周期，网络激活函数统一设定为 **RELU** 非线性单元。优化的超参为：**隐藏层维度（hidden size）**、**初始学习率（learning rate）** 以及 **L2 正则化系数（regularization parameter）**。

```
{
  "best_val_acc": 0.502,
  "config": {
    "hidden_size": 256,
    "learning_rate": 0.001,
    "reg_strength": 0.001,
    "activation": "relu",
    "batch_size": 64
  }
},
{
  "best_val_acc": 0.5082,
  "config": {
    "hidden_size": 256,
    "learning_rate": 0.001,
    "reg_strength": 0.01,
    "activation": "relu",
    "batch_size": 64
  }
}
```

最优化模型

经过超参优化之后, 确定了最优参数为隐藏层维度设为 256, 初始学习率设置为 0.001, L2 正则化系数取 0.01。此外, 我们令 batch size=64, epoch=50, 并选 RELU 函数作为激活函数。模型的训练结果如下:





总结

本次作业, 手工搭建三层神经网络分类器, 在数据集 CIFAR-10 上进行训练以实现图像分类. 实验表明采用交叉熵损失函数、L2 正则化 (系数 0.01) 和阶梯型学习率衰减的 SGD 优化算法, 结合隐藏层维度 256 和初始学习率 0.001 的并选 RELU 函数作为激活函数的超参数组合, 最终模型在测试集上达到 50.82% 的分类准确率.

代码链接

GitHub: [Raymond1227/FCNN](https://github.com/Raymond1227/FCNN)

通过网盘分享的文件: best_model.pkl

链接: <https://pan.baidu.com/s/1ZERSqwNakkGPKmhKag0nLg> 提取码: gb2r