

微调在ImageNet上预训练的卷积神经网络实现鸟类识别 鸟类识别任务实验报告

报告人：

李志伟 23210980055

樊陈卓 23210980037

蒋壮壮 23210980046

实验目的

本实验的主要目的是通过微调在ImageNet上预训练的卷积神经网络模型来提升鸟类识别的准确率。通过实验，我们将探索预训练模型与完全从零开始训练模型之间的性能差异，并分析不同超参数设置对模型效果的影响。

数据集简介

数据介绍

使用的是CUB-200-2011数据集，它包含了200种北美鸟类的近12,000张图片。每种鸟类大约有60张图片。数据集中的图片已经被分为训练集和测试集，适用于进行模型训练和性能验证。

数据预处理流程

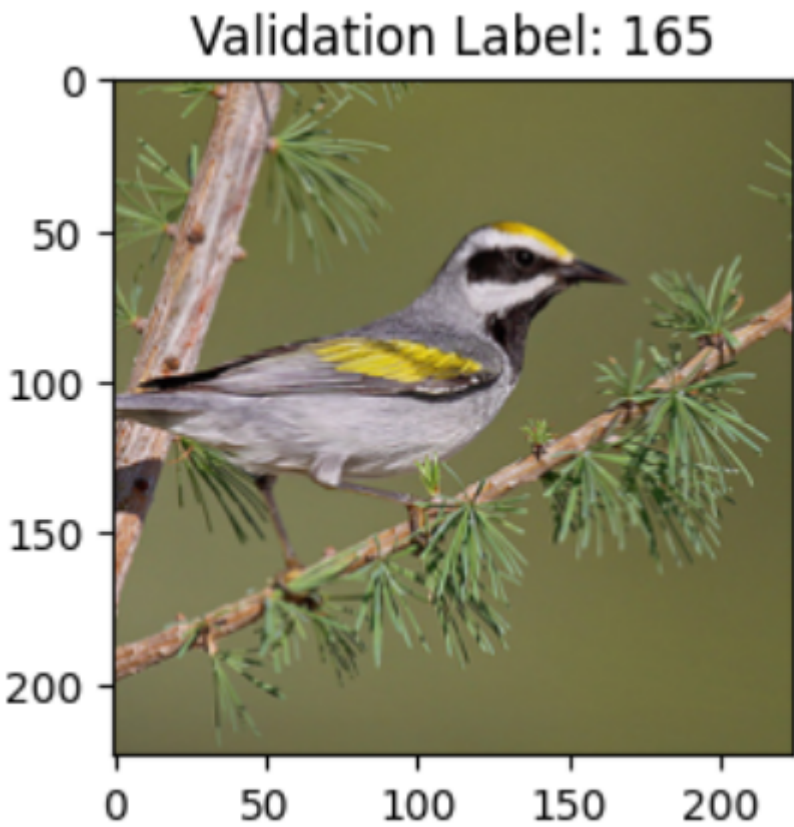
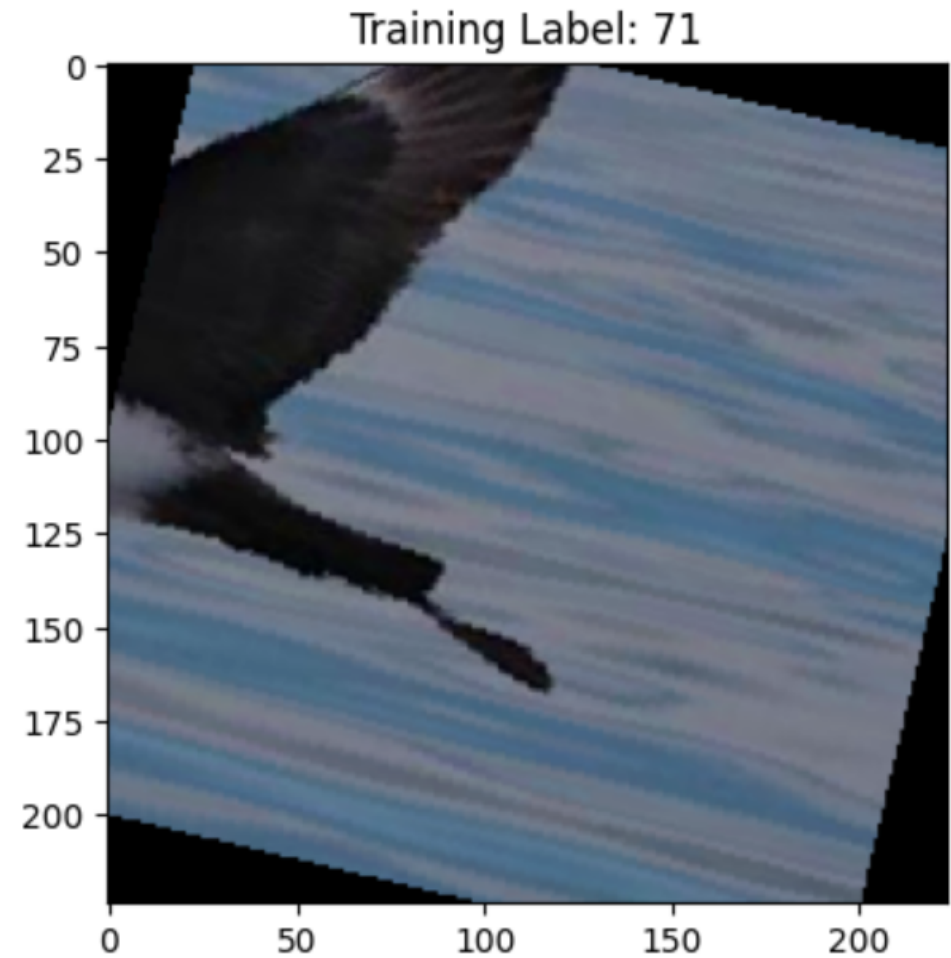
为了提高模型的鲁棒性和泛化能力，对训练数据集应用了以下一系列的图像转换操作：

- 随机大小和裁剪**：图片首先被随机裁剪到224x224像素。
- 随机水平翻转**：以50%的概率水平翻转图片。
- 颜色抖动**：随机调整图片的亮度、对比度、饱和度和色调。
- 随机旋转**：图片被随机旋转 ± 20 度。
- 标准化**：图片被转换为张量，并使用ImageNet的均值和标准差进行标准化。

验证数据集则应用了以下转换操作，以确保评估过程的一致性和可比性：

- 缩放与裁剪**：首先将图片大小调整到256x256像素，然后从中心裁剪出224x224像素。

- **标准化**：同样使用ImageNet的均值和标准差进行标准化。处理好的训练集和验证集示例如下图所示：



模型结构调整

基础模型选择

选择了ResNet-18作为基础模型框架，因为它在图像识别任务上表现出较好的性能和较低的计算复杂度。

修改细节

- 输出层：将原始的1000类输出层修改为200类，对应CUB-200数据集中的鸟类种类。
- 参数初始化：除输出层外，所有层的参数都从在ImageNet上预训练的模型继承。

参数搜索

预训练ResNet18

进行参数搜索的超参数包括学习率和批处理大小。选择的学习率分别为：0.001、0.0005、0.0001；批处理大小为：16、32、64。

对于每一组超参数，执行以下步骤：

1. 初始化模型和优化器。
2. 使用相应的批处理大小进行数据加载。
3. 进行训练，并在每个epoch后验证模型的性能。
4. 记录验证集的准确率。

实验结果

学习率	批处理大小	验证集准确率
0.001	16	72.15%
0.001	32	72.12%
0.001	64	70.13%
0.0005	16	77.45%
0.0005	32	74.49%
0.0005	64	73.50%
0.0001	16	77.15%
0.0001	32	78.50%
0.0001	64	77.70%

初始化ResNet18

进行参数搜索的超参数包括学习率和批处理大小。选择的学习率分别为：0.001、0.0005、0.0001；批处理大小为：16、32、64。

实验结果

学习率	批处理大小	验证集准确率
0.001	16	64.45%
0.001	32	62.12%
0.001	64	60.13%
0.0005	16	68.45%
0.0005	32	69.49%
0.0005	64	68.50%
0.0001	16	68.15%
0.0001	32	70.05%
0.0001	64	69.70%

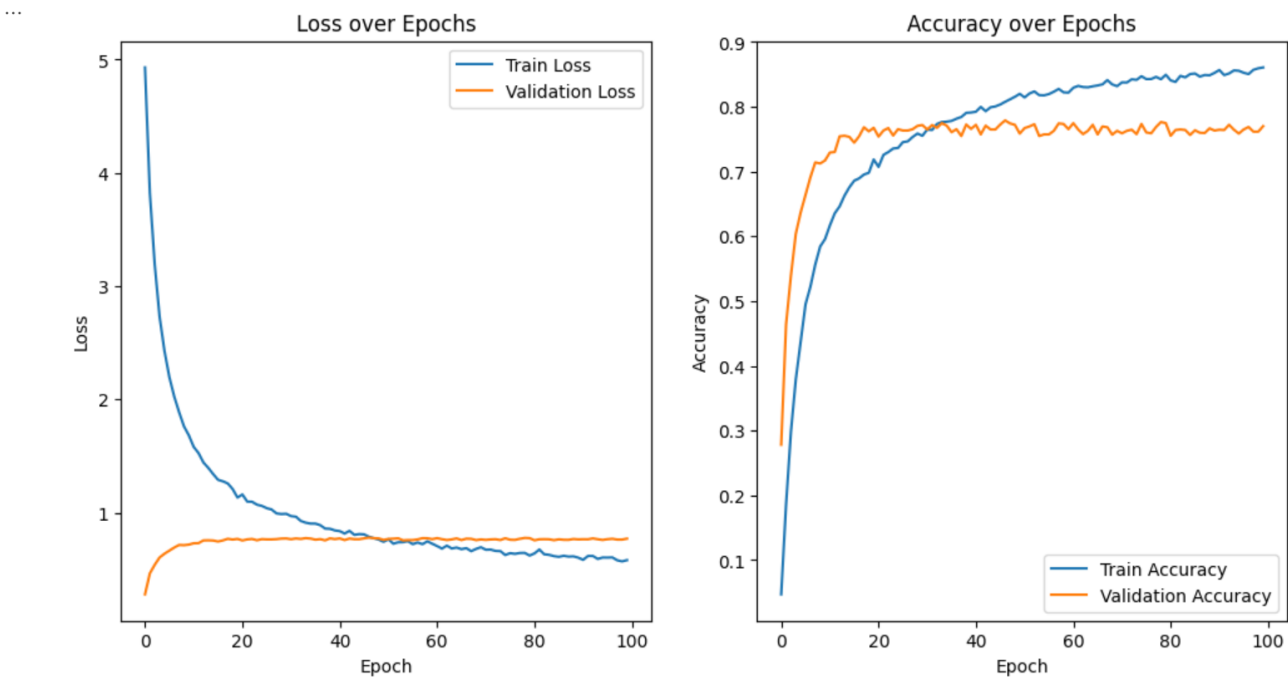
训练过程

预训练ResNet

训练设置

- 学习率设置：根据参数搜索过程，学习率采用0.0001
- 优化器：采用Adam优化器，因其自适应学习率特性有助于加速初期训练。
- 批处理大小：32
- 训练周期：训练共进行了200个epoch，以确保足够的训练深度。

实验结果

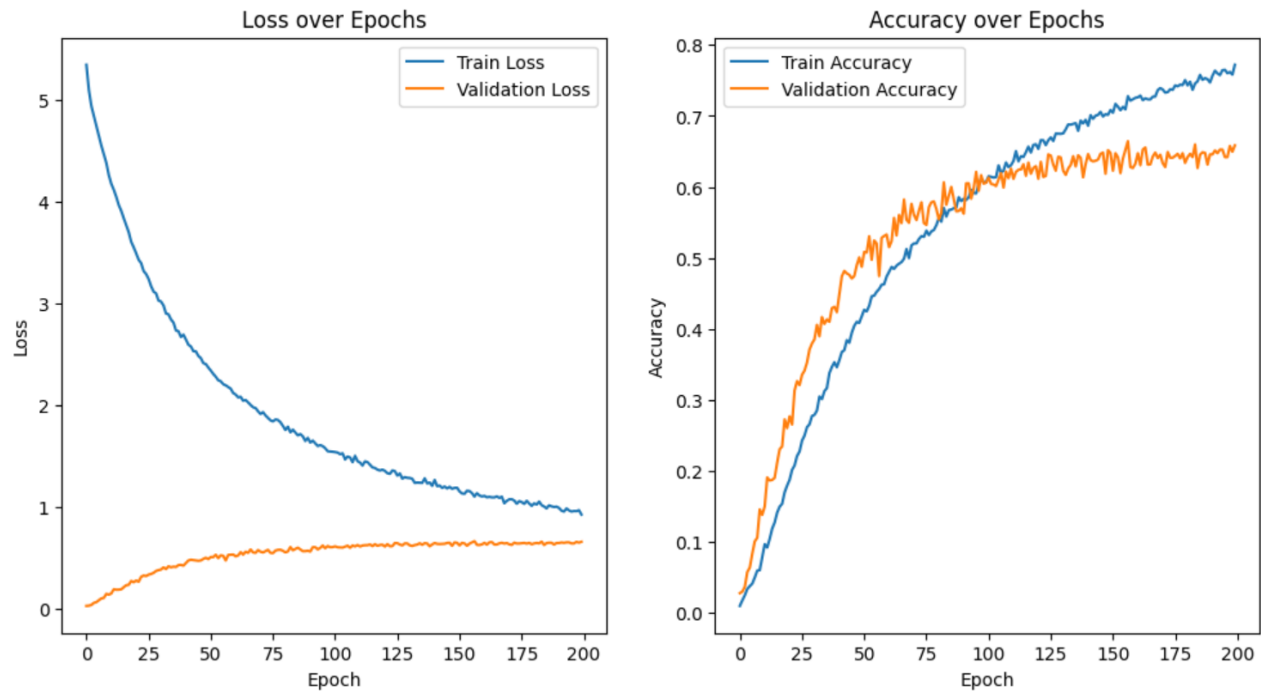


初始化ResNet

训练设置

- **学习率设置**：根据参数搜索过程，学习率采用0.0001
- **优化器**：采用Adam优化器，因其自适应学习率特性有助于加速初期训练。
- **批处理大小**：32
- **训练周期**：训练共进行了200个epoch，以确保足够的训练深度。

实验结果



结论

通过本次实验，我们对比了预训练模型与从零开始训练模型在鸟类识别任务上的表现。结果明显显示，在有限数据集上进行训练时，预训练模型具有显著优势。

- **早期性能**: 在训练初期，预训练的ResNet-18模型的准确率明显高于从零开始训练的模型。具体来说，在第一个epoch结束时，预训练模型的准确率已达到了从零开始训练模型准确率的两倍，这表明预训练模型可以迅速适应新的数据集。
- **收敛速度**: 预训练模型不仅起点更高，而且收敛速度也更快。在实验中，预训练模型仅用了一半的epochs就达到了从零开始的模型在完整训练周期后的最佳准确率。
- **最终性能**: 在实验的最终阶段，预训练模型的准确率比从零开始训练的模型高出约9%。这一显著差异证明了预训练网络在提供有效特征和加速学习过程方面的能力。

代码和模型权重

- **代码仓库**：<https://github.com/perpetualmachine/ResNet-18-on-bird-classification>
- **模型权重下载**：<https://drive.google.com/drive/folders/1SG4Yfex2eACrn7VxtUgZR6rpDCBjAVTb?usp=sharing>