
Lab 8: Image Classification

1 Intro

本次实验旨在通过图像分类任务，实战学习一些经典的计算机视觉网络结构(e.g., AlexNet and ResNet)及其变体，和一些优化训练方式(e.g., Dropout, BatchNorm, Optimizer)，强化对基础视觉概念的理解和掌握，在经典分类数据上积累的实验经验和sense对之后更宽更深的知识学习过程中也会很有帮助。

CIFAR-10 Dataset CIFAR-10数据集是Tiny Images数据集的一个子集，由60000张32x32的彩色图像组成。每个图像被分类为10个互斥的类别之一：airplane, automobile (but not truck or pickup truck), bird, cat, deer, dog, frog, horse, ship, and truck (but not pickup truck)。每类有6000张图像，每类有5000张训练图像和1000张测试图像。

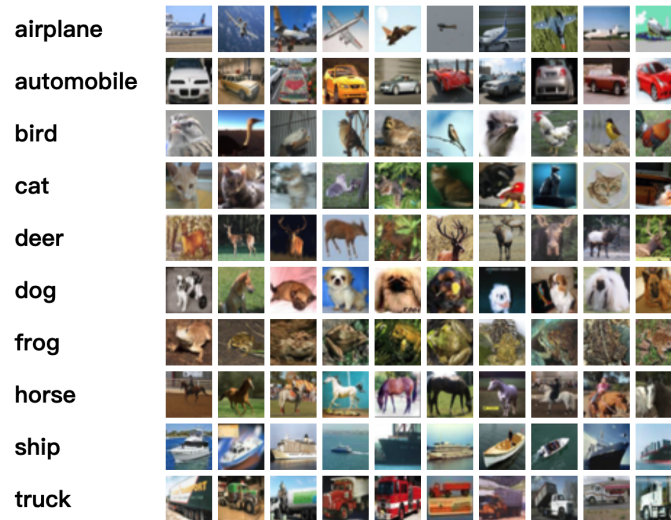


Figure 1: CIFAR-10示例

2 Requirements

numpy, torch, torchvision

3 TODO

在提供的code base基础上，设计自己的神经网络，完成图像分类任务(测试集准确率至少60%)，要求尝试的网络结构至少有VGG和ResNet(modified版本即可模仿原设计大致的结构和设计思想不要求完全复刻)

- accuracy > 60%

CIFAR-10: <https://www.cs.toronto.edu/~kriz/cifar.html>

- VGGNet (modified) and ResNet (modified)

4 TODO

- 关于运行
提供的代码是可以完整运行、且跑出一个比较base的结果的(accuracy 55)
- 关于网络设计
 - 提供的BaselineNet只是简单示例网络，可以在此基础上修改、也可另行设计
 - 自行设计过程中请不要调用第三方模型库，需要用torch.nn里的卷积层、线性层、池化层等自己搭建网络，目的是对不同经典网络设计思想和自己修改历程有所体会
 - 参照资料可以找VGGResNet原论文，一些博客和GitHub公开CV项目，或者一些tutorial: pytorch tutorial MS tutorial
- 任务量
最后提交的代码和报告里，写或者介绍的模型数请不要超过3个，不要在这个数量上内卷；准确率不是最终评价的唯一指标，不用过度追求过高的准确率，有一定尝试和提升即可，截止2022.11.22，该数据集SOTA方案efficient adaptive ensembling准确率为99.612，我们的实验目的是学习掌握一些设计思想特别是经典网络及其变体的思想、和自己一些修改的尝试，不必一味追求过高的准确率、复现SOTA方案

Submit

- 2023xxxxxx_xiaoming_lab8.zip (./code ./report.pdf)
- 提交至 <https://k.ruc.edu.cn>, DDL 2025.12.19 23:59

VGG: <https://arxiv.org/abs/1409.1556>
 ResNet: <https://arxiv.org/abs/1512.03385>
 pytorch tutorial: https://pytorch.org/tutorials/beginner/blitz/cifar10_tutorial.html
 MS tutorial: <https://learn.microsoft.com/en-us/windows/ai/windows-ml/tutorials/pytorch-train-model>
 efficient adaptive ensembling: <https://arxiv.org/abs/2206.07394>