

### 1 问题背景

图像识别作为计算机视觉领域的一个核心问题，对于自动驾驶、视频监控、图像检索等应用至关重要。特别是在自动驾驶领域，准确识别车辆与非车辆对象是确保行车安全的基本前提。车辆/非车辆的二分类问题，即将图像中的对象划分为车辆和非车辆两大类，对于理解交通场景、决策制定具有显著的作用。在本次作业实践中，将深入理解 CNN 在图像分类任务中的应用，并熟悉理解 PyTorch 框架和相关环境的构建以及训练神经网络模型的技能。

### 2 模型原理

卷积神经网络（CNN）已成为图像识别任务的主流模型之一，特别适合处理具有网格结构的数据，如图像（二维网格）。在提供的代码中，CNN 通过一系列卷积层和池化层自动学习图像特征。卷积层使用滤波器在图像上滑动，捕捉局部特征，而池化层则降低特征的空间维度，同时增加对图像位移的不变性。随后，展平层将卷积层的输出转换为一维特征向量，供全连接层进一步处理。全连接层对特征进行分类，其中最后一个全连接层的输出通过 sigmoid 激活函数转换为概率值，用于二分类决策。在训练过程中，CNN 通过反向传播算法和 Adam 优化器不断更新网络权重，以最小化损失函数，从而提高分类准确性。这种端到端的学习方式使得 CNN 在图像分类任务中表现出色，能够自动提取复杂的图像特征并进行有效分类。

### 3 数据集

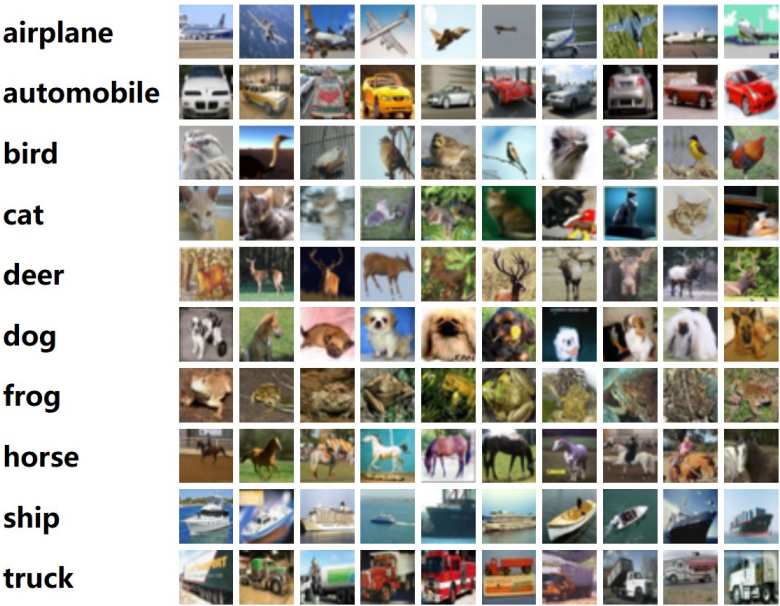


图 1 CIFAR-10 数据集

CIFAR-10 数据集是图像分类研究中的标准基准，它包含 10 个类别的 60000 张 32x32 图像，其中包括汽车这一类别。对于车辆分类问题，CIFAR-10 提供了一个理想的平台，因为它不仅包括了多样化的车辆图像，还涵盖了其他类别的图像，可以用来训练和测试模型对车辆的识别能力。此外，由于数据集中每个类别的图像数量均衡，这有助于防止模型对某些类别产生偏见，确保了分类的公平性。CIFAR-10 的易用性和广泛认可，加上其对深度学习模型构成的挑战，

使其成为研究车辆二分类问题（如车辆与非车辆识别）的优选数据集。在本作业中，我们使用 CNN 对 CIFAR-10 数据集中的图像进行车辆与非车辆的二分类。其中，训练集包含 50000 张图像，每个类别 5000 张，测试集包含 10000 张图像，每个类别 1000 张。

#### 4 模型结构

本作业采用的神经网络模型是一个简单的卷积神经网络，具体结构如下：

- 1.第一层卷积：使用 3x3 的卷积核，从输入图像中提取初步特征，其中输入通道数为 3（RGB 三通道），输出通道数为 6。
- 2.最大池化层：采用 2x2 的最大池化窗口，减少特征的空间维度，同时增加对图像位移的不变性。
- 3.第二层卷积：再次使用 3x3 的卷积核，增加模型的深度，以提取更复杂的特征，输出通道数为 16。
- 4.展平层：将卷积层的输出展平为一维向量，以适配全连接层的输入。
- 5.全连接层：包括两个全连接层，第一个全连接层有 120 个神经元，第二个有 84 个神经元，用于在特征向量上学习分类任务的高级表示。
- 6.输出层：单个神经元使用 sigmoid 激活函数输出概率，用于二分类决策。

#### 5 网络参数设置

在本作业中，网络的训练参数如下：

- 1.学习率：设置为 0.001，这是优化算法在每一步更新模型权重时的步长。
- 2.批量大小：设置为 64，即每次迭代使用 64 张图像来更新模型权重。
- 3.训练周期：设置为 8、10、15，即整个数据集将分别被遍历 8、10、15 次。
- 4.损失函数：使用二元交叉熵损失函数（BCELoss），适用于二分类问题，衡量
- 5.模型预测概率与实际标签之间的差异。
- 6.优化器：使用 Adam 优化器，它结合了动量和 RMSProp 的特点，通常在训练深度学习模型时表现良好。

#### 6 实验结果及分析

结合以上网络结构和参数设置，分别进行 8 次、10 次、15 次 epoch 的网络训练。得到以下图 2、3、4 所示的结果。

1) 8 次 epoch，损失迭代过程：

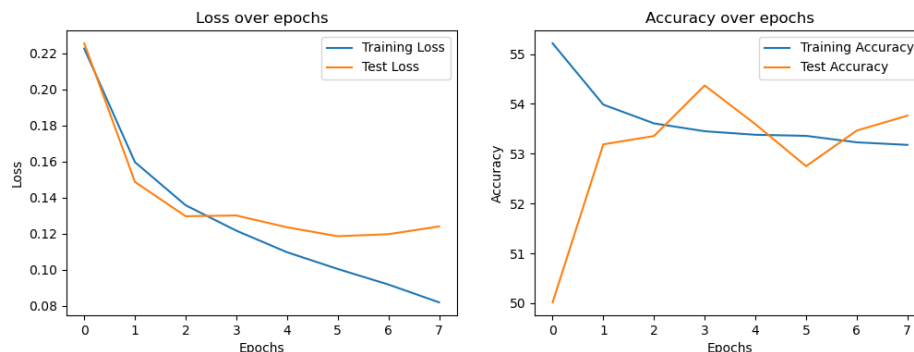


图 2 8 次 epoch 迭代结果

2) 10 次 epoch，损失迭代过程：

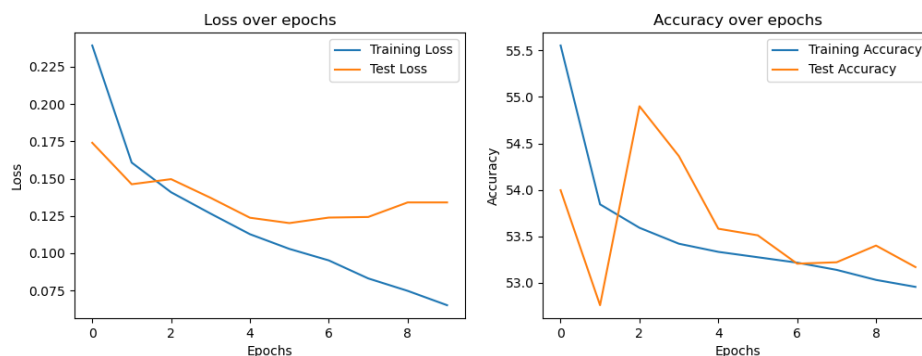


图 3 10 次 epoch 迭代结果

3) 15 次 epoch，损失迭代过程：

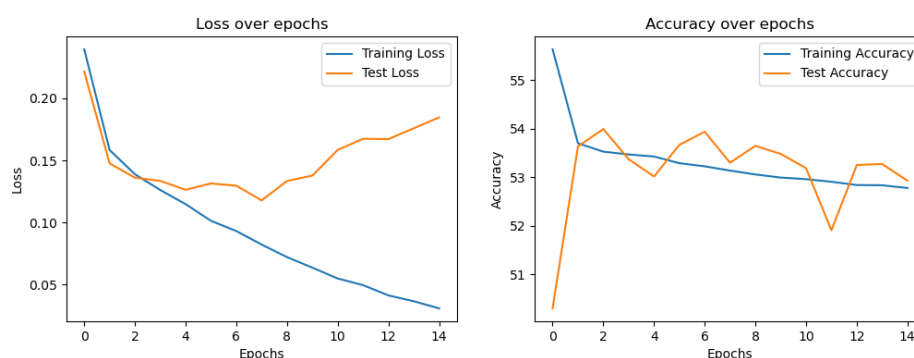


图 4 15 次 epoch 迭代结果

**Epoch 8 的结果：**训练损失和测试损失都较低，测试损失略低于训练损失，这表明模型在测试集上的表现相对较好。训练精度和测试精度都比 Epoch 10 的结果要高，但仍然有改进空间。

**Epoch 10 的结果：**训练损失和测试损失都较高，且测试损失没有明显低于训练损失，这表明模型可能尚未完全训练或存在过拟合。训练精度和测试精度都较低，模型在训练集和测试集上的性能都不是很好。

**Epoch 15 的结果：**训练损失和测试损失随着 epoch 的增加而降低，测试损失略低于训练损失，表明模型在测试集上的表现逐渐提升。测试精度有所提升，但仍然不高。即模型虽然在训练集上表现良好，但在测试集上的泛化能力还有待提高。

## 7 总结与展望

从测试精度来看，三个 epoch 中只有 Epoch 8 的结果相对更好一些，因为其具有较低的测试损失和较高的测试精度。然而，即使是最高的测试精度（Epoch 8 的结果），也远未达到理想的水平。这可能意味着模型的复杂度不够，或者当前的网络结构和参数设置不是最优的。后续工作可以从以下方面进行改进从而提升分类效果：

- (1) 调整网络结构：增加卷积层的深度或宽度，添加更多的全连接层，或使用不同类型的激活函数。
- (2) 调整超参数：改变学习率、批量大小或优化器设置。

- (3) 数据增强：应用更多的数据增强技术，如旋转、缩放、裁剪等，以提高模型的泛化能力。
- (4) 正则化技术：使用 **dropout**、L2 正则化等方法来减少过拟合。
- (5) 学习率调度：实施学习率衰减策略，例如学习率随着 **epoch** 的增加而降低。
- (6) 早停法：在验证集上实施早停法，以避免过拟合。
- (7) 阈值调整：根据需要调整分类阈值，以优化精度。
- (8) 模型评估：使用混淆矩阵、ROC 曲线和 AUC 等指标对模型进行全面评估。
- (9) 更复杂的模型：如果简单的 CNN 不足以解决问题，可以考虑使用更复杂的模型，如残差网络（ResNet）或卷积循环网络。