

第二次仿真实验

冯海桐-522031910557

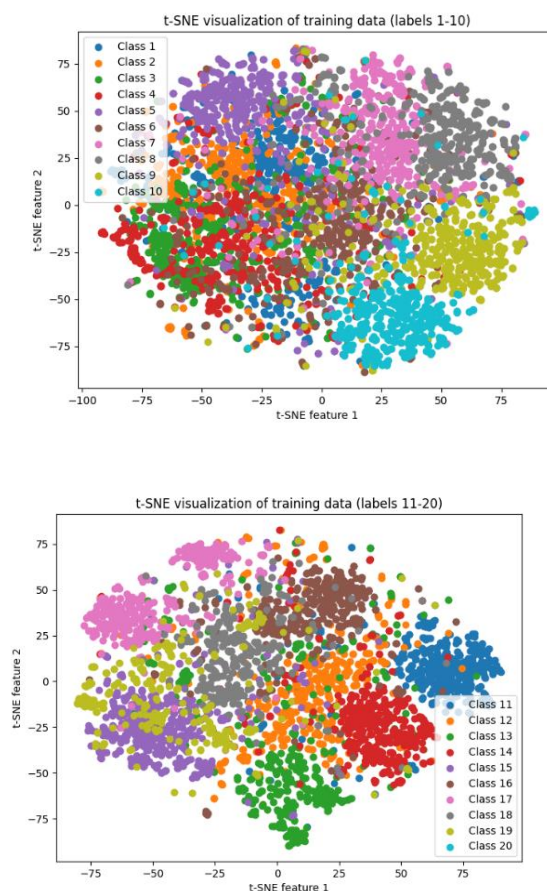
实验简述

编写数据可视化方法 t-SNE 的实现程序，并利用该程序可视化课程设计中训练数据的二维散点图。

t-SNE (t-Distributed Stochastic Neighbor Embedding) 是一种非线性降维技术，主要用于高维数据的可视化。它能够将高维数据嵌入到二维或三维空间中，同时尽可能地保留数据的局部结构，使得在低维空间中相似的数据点仍然相邻。t-SNE 的主要步骤是：对于每对数据点，先计算它们在高维空间和低维空间中的相似度，再通过梯度下降法，最小化高维空间和低维空间中相似度分布的差异。

实验结果

由于 20 个种类还是太多了，可视化结果会乱作一团。因此，我将 label 为 1~10 和 11~20 分别进行可视化，得到下面两张图：



可以看到，第二张图比较泾渭分明，而第一张图有几类混在一起，差别不明显。