

《模式识别》课程实验报告

# 实验二：字符识别

院 系： 机电工程学院自动化系

姓 名： 朱颖聪

学 号： 200320708

班 级： 自动化7班

日 期： 2023年12月30日

# 目录

[1 项目背景 2](#_Toc154858495)

[1.1 背景介绍 2](#_Toc154858496)

[1.2 数据集说明 2](#_Toc154858497)

[2 项目目标 2](#_Toc154858498)

[3 理论介绍 2](#_Toc154858499)

[4 实验过程及结果 3](#_Toc154858500)

[任务一：卷积神经网络 3](#_Toc154858501)

[任务二：K-L变换+MLP 4](#_Toc154858502)

[附录：关键代码注释 7](#_Toc154858503)

# 字符识别

## 1 项目背景

### 1.1 背景介绍

字符识别作为计算机视觉领域的重要任务，涵盖了许多实际应用，如自然场景中的车牌识别、手写文字识别等。本次实验旨在比较并探索基于K-L变换和多层感知机（MLP）的特征提取与分类方法，以及基于卷积神经网络（CNN）的字符识别。这将帮助我们深入了解不同方法在字符识别任务中的性能和适用性。

### 1.2 数据集说明

本次实验所用数据集为Chars74K系列数据集。Chars74K系列数据集是一个用于字符识别的综合数据集，包含英文大写字母和数字0~9的图像，图像大小为28x28像素，字符为白色，背景为黑色。该数据集由基础数据集（包括Chars74K数据集和网络数据集）以及经过旋转、缩小还原等操作的变换数据集构成。数据集分为训练集（Chars74K\_train）和测试集（Chars74K\_test）。训练集共计74700个图像，测试集则包含6294个独立的图像。

## 2 项目目标

**任务一：**

**基于K-L变换和MLP的字符识别：**

* + **特征提取：** 利用K-L变换从原始图像中提取高效的特征向量。
  + **分类方法：** 使用MLP进行字符分类。

**任务二：**

**基于卷积神经网络的字符识别：**

* + **特征提取：** 利用卷积神经网络（CNN）自动学习图像特征。
  + **分类方法：** 使用CNN进行字符分类。

## 3 理论介绍

**1. K-L变换：**

K-L变换是一种特征提取方法，通过线性变换将数据映射到一个新的坐标系，使得在新坐标系中数据的方差最大化。在字符识别中，K-L变换可用于将原始图像转换为更紧凑的特征表示，有助于提高后续分类器的性能，并且可以大大简化训练成本。

**2. 多层感知机（MLP）：**

多层感知机是一种前馈神经网络，由输入层、隐藏层和输出层组成。每个神经元都与上一层的所有神经元连接，具有权重和激活函数。

在字符识别中，MLP可用于学习从K-L变换后的特征到字符类别的映射。通过训练，模型能够自动学习字符的抽象表示。

**3. 卷积神经网络（CNN）：**

卷积神经网络是一类包含卷积计算且具有深度结构的前馈神经网络，是深度学习的代表算法之一。卷积神经网络具有表征学习能力，能够按其阶层结构对输入信息进行平移不变分类，因此也被称为“平移不变人工神经网络”。

## 4 实验过程及结果

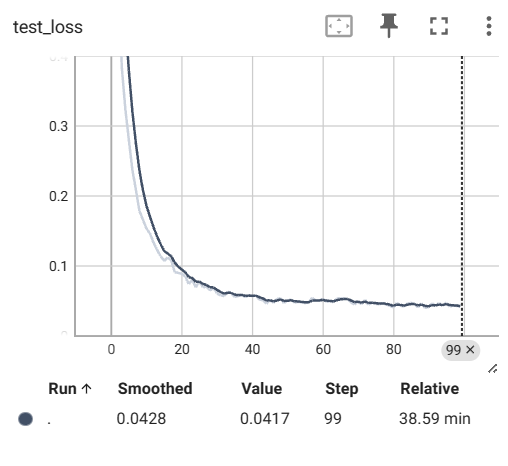
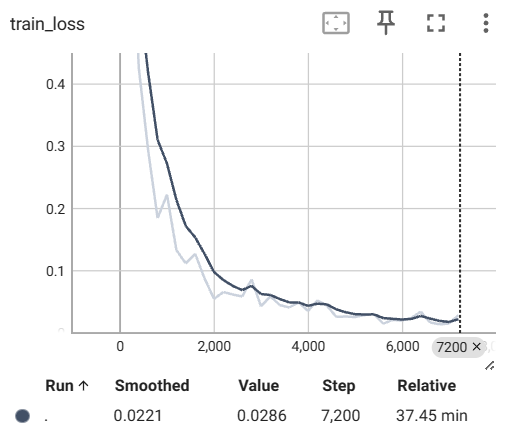
### 任务一：卷积神经网络

1. 网络结构：本次实验使用的卷积神经网络一共包含两个卷积层、一个池化层以及两个全连接层。

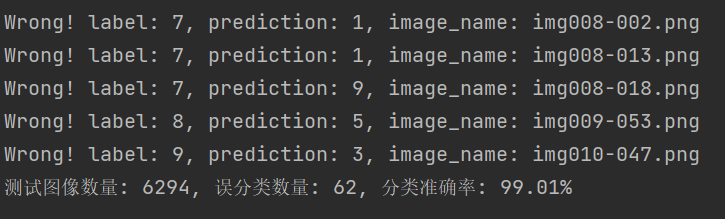


1. 选择优化器为Adadelta，损失函数为CrossEntropyLoss，设置batch\_size = 1024,训练100 epochs，载入训练数据，训练网络。
2. 根据训练结果选取最优网络，测试并评估结果。

以下为训练曲线：

训练损失和测试损失随着训练次数的增加不断下降，选取最后一个回合的模型进行测试。



准确率达到99.01%

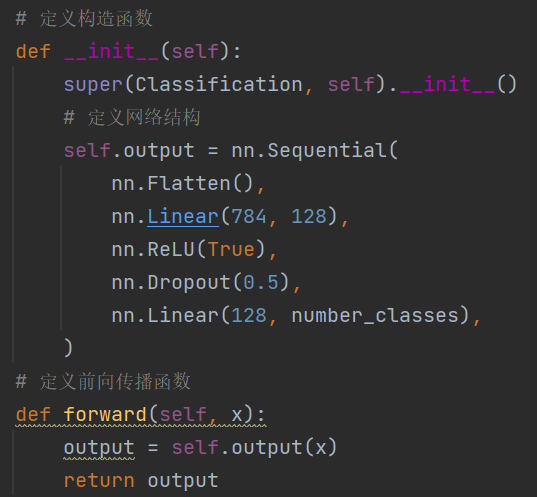
在该网络结构下，不需要额外的参数调整即可获得很好结果。

### 任务二：K-L变换+MLP

首先，对训练集的图像计算协方差矩阵，计算特征值和特征向量，根据累计贡献率选取前N大的特征向量，对图像进行降维处理。（代码见附录）

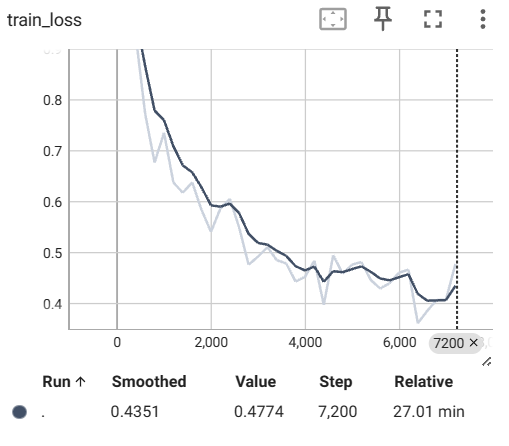
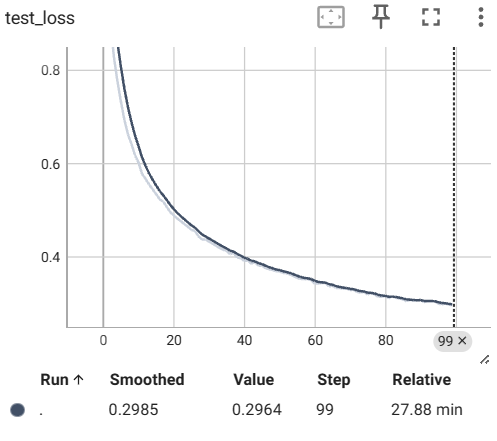
（1）选取累计贡献率e=1，即相当于不使用K-L转换,输入向量维度为784

设计网络结构为

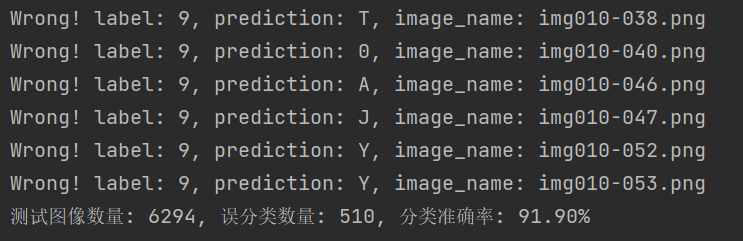


选择优化器为Adadelta，损失函数为CrossEntropyLoss，设置batch\_size = 1024,训练100 epochs。

训练过程损失函数如下所示：



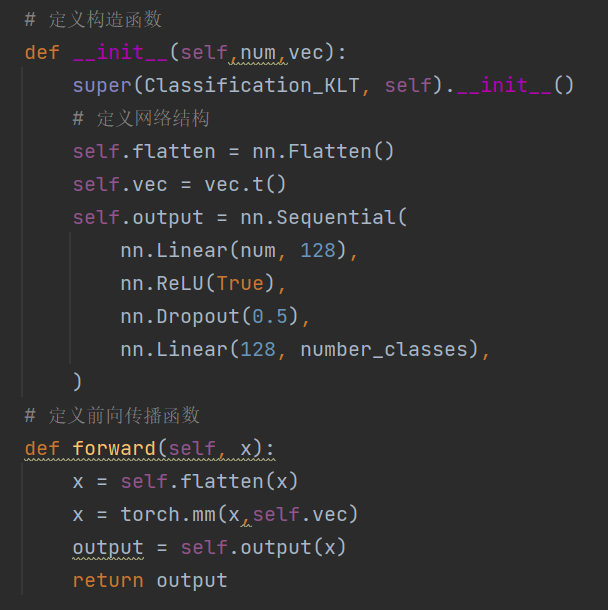
训练损失和测试损失随着训练次数的增加不断下降，选取最后一个回合的模型进行测试。



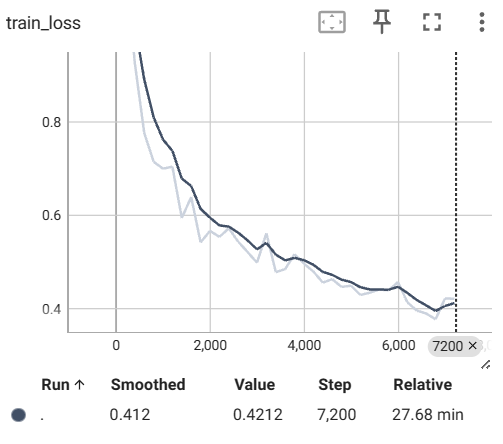
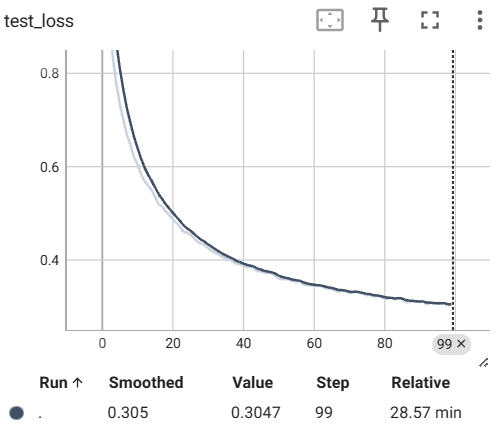
可以看到准确率仅91.9%，与CNN网络差距较大。

（2）选取累计贡献率e=0.99，降维后输入向量维数为170

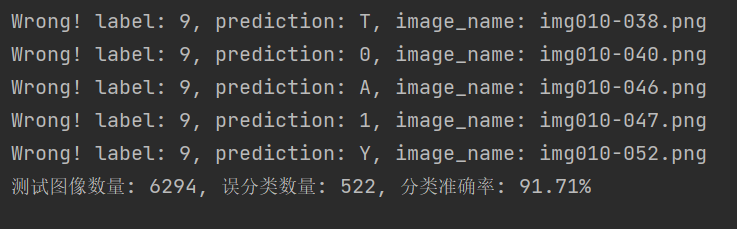
网络结构为



网络中的self.vec为特征向量，除输入外，其余网络结构与e=1时的基本一致。选择优化器为Adadelta，损失函数为CrossEntropyLoss，设置batch\_size = 1024,训练100 epochs。



训练损失和测试损失随着训练次数的增加不断下降，选取最后一个回合的模型进行测试。



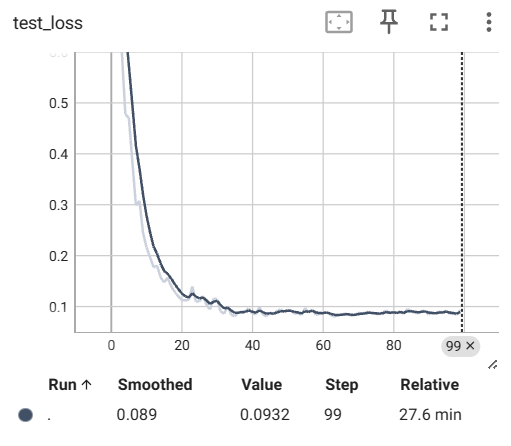
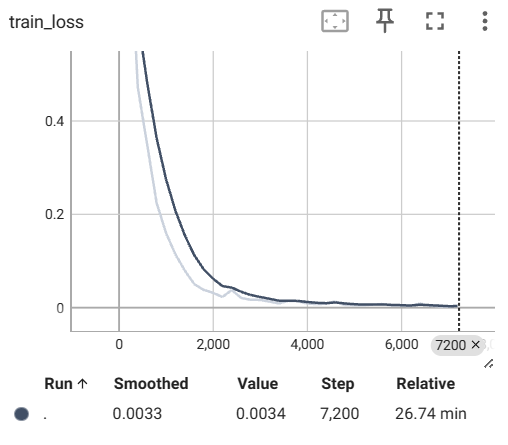
准确率为91.71%，与贡献率e=1时的准确率基本一致，可以看出K-L转换降低特征向量维度的同时基本保留特征向量的所有信息，这样可以提高训练的效率，节省时间。

（3）选取累计贡献率e=0.99，增加网络复杂度

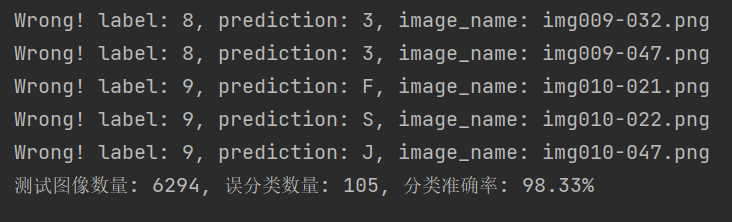
网络结构为



选择优化器为Adadelta，损失函数为CrossEntropyLoss，设置batch\_size = 1024,训练100 epochs。

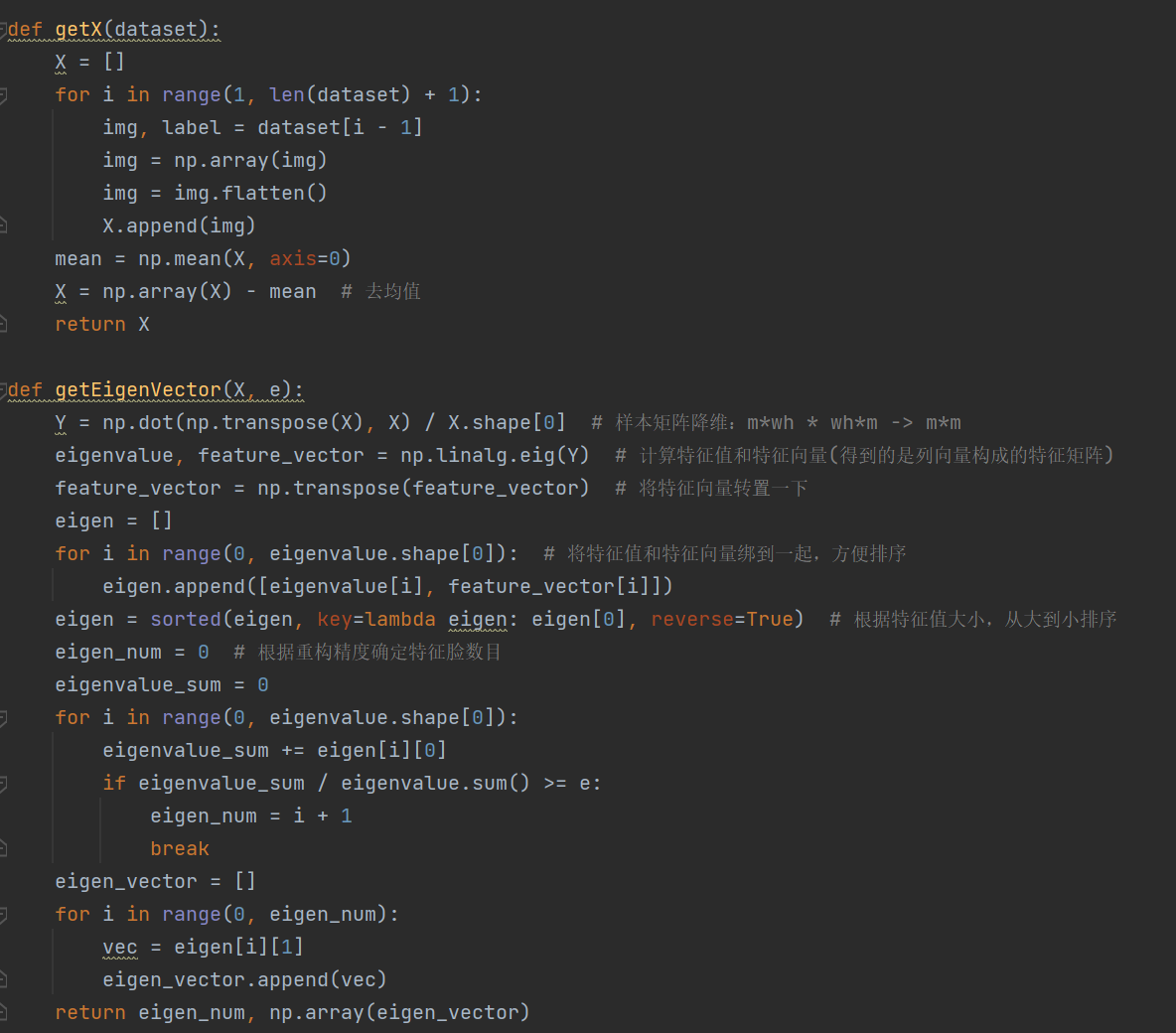
训练损失和测试损失随着训练次数的增加很快达到收敛，在第60回合中训练损失最低，故选择第60回合作为测试模型。



准确率为98.33%，可以看出提高模型复杂度可以大大提高准确率，与CNN网络的准确率相差不大。

## 附录：关键代码注释

K-L转换代码



主要训练过程

