黑色的仙人掌

中度可信度描述已自动生成

**计算机与信息学院**

**《人工智能》课程作业**

**2024年秋季学期**

|  |  |
| --- | --- |
| 课程类型： | 专业核心课 |
| 学 号： | 202210120518 |
| 姓 名： | 胡国昌 |
| 专 业： | 计算机科学与技术 |
| 授课教师： | 臧兆祥 |

完成日期：2024年 10月 20日

**手写体数字识别**

# 案例内容与要求

## 实验内容与任务

MNIST数据集是一个手写体数字的图像数据集，训练集包括60000张图片，测试集包括10000张图片，每张图片是一个8位的灰度图片，尺寸为28×28,训练集的前20张图片如图1.1所示。现要求训练一个卷积神经网络，用于识别数字图片。

## 实验过程及要求

1. 实验环境要求：Windows/Linux操作系统，Python编译环境，numpy、keras、matplotlib等程序库。
2. 学习理解神经网络、卷积网络层、图像处理等知识。
3. 下载数据集，构建卷积神经网络，进行网络训练与评估。
4. 调整网络超参数，记录网络训练的过程。
5. 撰写实验报告。

# 原理论述及解决方法

## 原理概述

神经网络是一种运算模型，由大量的节点（或称神经元）之间相互联接构成。每个节点代表一种特定的激活函数，每两个节点间的连接都代表一个对于通过该连接信号的加权值，网络的输出是连接方式、权重值和激活函数的综合作用。通过调整神经网络的结构、规模、参数，它可以逼近任何函数，因此神经网络目前成为机器学习的重要工具。图像处理中可以通过卷积计算获得图像特征，卷积神经网络可以进行图像特征提取，以实现图像分类、目标检测等任务。2012年，在大规模机器视觉识别竞赛(ILSVRC)上，卷积神经网络AlexNet超出其他学习方法，取得了最好结果。从此，卷积神经网络技术在图像处理领域得到广泛的应用。

## 解决方法

### 神经网络与层

各种信息处理过程可以看作是一个函数处理 ，然而 的形式往往是未知并且复杂的。神经网络使用简单的线性处理或非线性处理进行连续复合的方式，来逼近或者模拟 。用 表示输入数据，用 , , 表示第 次复合处理的输入、简单处理函数和输出，则神经网络处理过程为

…

其中，是输入层，，是个隐含层，称为输出层。在神经网络中，通常定义为的线性函数或非线性函数的组合，例如常见的全连接层的定义为

这些简单函数的复合可以逼近复杂函数，而每个网络层的,等参数一起构成了神经网络的参数，的每个元素是一个神经元。神经网络学习的任务是通过大量的训练数据来找到各网络层的参数。

### 激活函数

网络层要加入非线性成分，因为多个线性处理的复合还是线性的，只有线性处理的神经网络的函数模拟能力有限。可以如同公式1.2一样，在线性处理后加入一个非线性的成分，称为激活函数。当然也可以把激活函数定义为一个单独的层。常用的激活函数有sigmoid、tanh、relu函数等。因为神经网络是使用梯度下降法完成参数训练，sigmoid函数在多数区域梯度为0，会影响训练的速度，因此常常使用relu函数作为激活函数。

### 损失函数与优化计算

给定一批训练数据。为了计算网络参数，定义一个损失函数来衡量网络预测与真实数据之间的差异，通过在训练集上最小化损失函数来获得对参数的估计。假设神经网络用函数表示，用神经网络进行回归处理时，常用均方误差(MSE)作为损失函数，

当用神经网络进行分类处理时，输出一个概率向量，用负的对数似然作为损失函数，也称为交叉熵损失函数，

其中用one-hot形式表示。

由于神经网络的线性和非线性函数都是基本光滑的，因此损失函数也是光滑的。神经网络中用梯度下降法来最小化损失函数，从而获得最佳的参数。虽然主要应用梯度下降，但在具体的计算中，嵌入了不同的处理手段，目前常用的优化算子有随机梯度下降(SGD)，自适应学习率优化算子AdaGrad, RMSProp等。

### 图像处理与卷积层

用一个2值图片1(a)为例，其中有一个“7”字。图1(b)是一个尺寸较小的模板，称为卷积核。将模板在图片1(a)上滑动，并与覆盖的区域作乘法，

其中是输出特征，是输入的图片，是卷积核，和是卷积核的尺寸。这个操作称为卷积(跟一般信号处理的严格定义有差别)。卷积的结果按位置列在图1(c)中。可以发现，在位置（2,3）处出现最大的卷积值5。这个示意图说明通过卷积操作，能大致判断在图片1(a)的(2,3)处，可能有一个“7”字存在。卷积核可以有不同的尺寸，我们还可以用两个小一点的模板组合起来处理，如模板1(d)通过卷积判断有一个水平线，模板1(e)通过卷积判断有一个垂直线，则很可能有一个“7”存在。组合的方法可以避免使用复杂的卷积核。通常情况下，图像处理中卷积核是方形的。在进行卷积操作时，除了卷积核尺寸可以不同外，在图片上滑动的步幅也可以不同。卷积核的尺寸和卷积的步幅，可以影响到输出的尺寸。

手机屏幕截图

描述已自动生成

图1 卷积操作示意图

卷积处理能获取空间关系的特征，从而特别适用于图像处理。神经网络中的卷积层，由一批不同值的卷积核构成，它们进行卷积操作后的输出，相当于提取了一次图片特征。下一个网路层还可以继续是卷积层，提取更高级的特征。

### 池化层

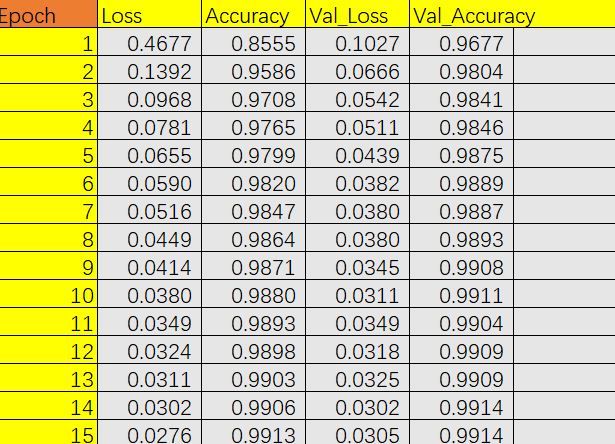
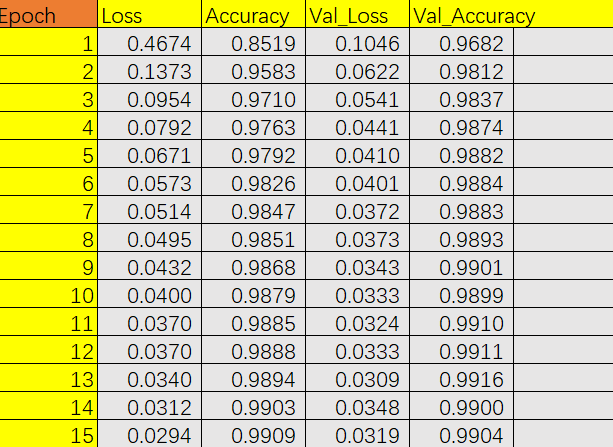
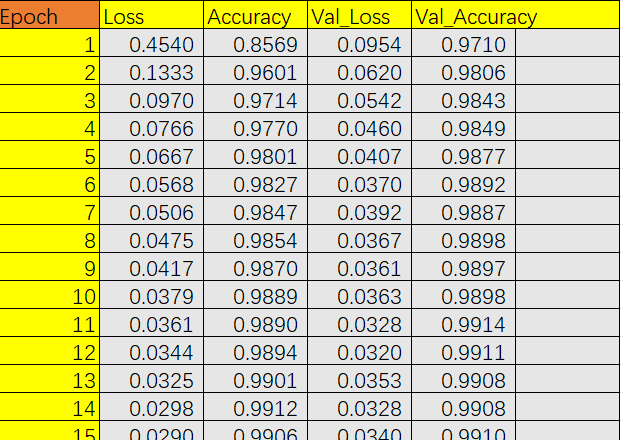
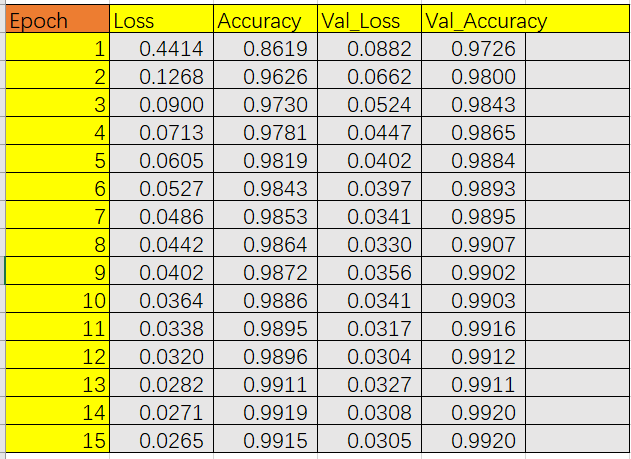
神经网络的一个网络层输入的数据量越多，需要配置的参数就越多，计算耗费的时间就越大，训练也就越困难。而图片作为一个空间信息的载体，大部分邻近的信息是相同的，因此通过采样的方式，扔掉一部分数据，但图片基本特征还能保持。池化层完成的就是这种采样来缩小输入的工作。跟卷积操作类似，用一个池化窗口在图片上滑动，池化结果是窗口内数据取平均值或者最大值，称为平均池化或者最大池化。数据缩小的程度与窗口尺寸和步幅相关。

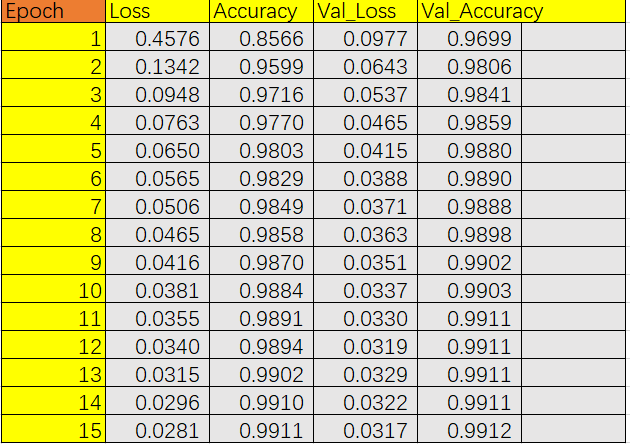
### 丢弃处理Dropout

神经网络具有大量的参数，有很强的数据拟合能力。但是拟合能力过强，在用来预测时并不一定有好的效果，因为训练数据本身可能会有误差，实际模型也可能并没有那么复杂。避免过拟合的方法之一是训练过程中，随机丢弃一些神经元，使得网络结构变小，从而能部分抑制过拟合。另外，丢弃一些神经元，可以迫使网络其他的神经元能学到一些更一般的特征。

# 计算结果与讨论

## 计算结果





在四次实验的平均数据中（图1-4为三次实验数据，图5为其平均值），模型在15个训练轮次内表现出显著的性能提升，具体结果如下：

1. 训练集损失（Loss）：从第1轮的0.4576逐渐降低至第15轮的0.0281，表明模型在训练过程中逐渐收敛。
2. 训练集准确率（Accuracy）：从第1轮的85.66%提升至第15轮的99.11%，显示了模型在手写数字识别任务上的学习效果。
3. 验证集损失（Val\_Loss）：验证集损失从第1轮的0.0977降低至第15轮的0.0317，表明模型在验证集上的泛化能力有所提高。
4. 验证集准确率（Val\_Accuracy）：验证集准确率从第1轮的96.99%提升至第15轮的99.12%。

这些结果表明，随着训练轮次的增加，模型的训练和验证性能均得到了显著的改善。

## 实验讨论

从实验数据可以看出，模型在初期轮次内的损失值和准确率变化较为显著。在前5轮训练中，损失值迅速下降，准确率则快速上升。这说明模型在初期阶段能够快速学习到MNIST数据集的特征。随着训练的进行，损失值下降和准确率上升的速度逐渐放缓，并在最后几轮趋于稳定，这表明模型逐渐收敛。

在验证集上的表现也类似，验证集的损失和准确率从初始轮次到最后轮次逐步改善，说明模型不仅在训练集上表现良好，而且具备良好的泛化能力。

此外，模型在验证集上的损失值和准确率趋于稳定，表明没有出现明显的过拟合现象。可以认为，此次训练过程中使用的模型结构和参数设置在MNIST手写数字识别任务中较为适合，达到了预期的识别效果。

# 作业总结

在本次实验中，我进行了基于MNIST数据集的手写体数字识别实验。本实验的核心任务是利用卷积神经网络（CNN）对手写数字进行分类识别，实验步骤包括数据集下载、模型构建、训练、评估等，环境为Python编程，并使用了NumPy、Keras、Matplotlib等库。

实验过程与心得：

1. 数据处理与模型搭建：通过对MNIST数据集的预处理，将28×28的灰度图像输入CNN模型。模型的主要结构包含卷积层、池化层以及全连接层，通过多次实验调整了网络的超参数以优化识别效果。
2. 训练与评估：在训练过程中，通过不断优化损失函数，模型的训练集和验证集的准确率均达到了99%以上，展现了较好的性能。损失值在15轮次后收敛，准确率也逐渐趋于稳定，说明模型适合该数据集。
3. 结果分析：实验表明，通过适当的卷积核尺寸、池化方式以及正则化手段，模型能够有效避免过拟合，同时具备良好的泛化能力。模型在训练集和验证集上的表现均较为稳定。

本次实验加深了我对神经网络和卷积神经网络的理解，尤其是在图像处理方面的应用。通过本实验，我基本了解了图像特征提取、模型调参等技能，也对机器学习的应用有了更深的体会。未来，希望进一步探索深度学习在更复杂数据集上的应用，提高模型的识别精度和泛化能力。

# 参考文献

[1] 徐义春. 人工智能案例与实验[M]. 2024年5月第1版. 清华大学出版社, 2024.

# 实验代码