计算机视觉实践实验报告（二）

目录

[一． 实验目的 2](#_Toc132657328)

[二． 实验原理 2](#_Toc132657329)

[2.1MNIST数据集 2](#_Toc132657330)

[2.2 LeNet-5模型 3](#_Toc132657331)

[2.3 卷积层 5](#_Toc132657332)

[2.4池化层 6](#_Toc132657333)

[2.5全连接层 8](#_Toc132657334)

[2.6激活函数 9](#_Toc132657335)

[2.7损失函数 10](#_Toc132657336)

[2.8优化器 10](#_Toc132657337)

[三． 实验步骤 12](#_Toc132657338)

[四． 数据集 12](#_Toc132657339)

[五． 程序代码 12](#_Toc132657340)

[六． 实验结果 16](#_Toc132657341)

[6.1 训练损失 16](#_Toc132657342)

[6.2 测试结果 16](#_Toc132657343)

[七．实验分析与总结 17](#_Toc132657344)

# 一． 实验目的

* 熟悉卷积神经网络的基本结构，包括卷积层，池化层，激活函数及最后的全连接层等。
* 学习LeNet-5模型。
* 下载MNIST数据集。
* 实现 LeNet-5 在MNIST 数据集上的训练和测试。

# 二． 实验原理

主要包括MNIST数据集，LeNet-5模型，卷积层，池化层，全连接层，激活函数，损失函数和优化器 8个部分。

# 2.1MNIST数据集

MNIST数据集是一个手写数字图像数据集，由70,000张图像组成，每张图像大小为28x28像素。这些图像被分为60,000张用于训练和10,000张用于测试。每个图像都标记为一个数字，从0到9之间的整数，表示图像中的手写数字。

MNIST数据集是机器学习领域中最常用的数据集之一。它被广泛用于训练和测试各种图像分类算法和模型，例如卷积神经网络（Convolutional Neural Networks, CNNs）。这是因为MNIST数据集规模较小，易于使用，同时其具有代表性，可以很好地代表手写数字图像。

通过使用MNIST数据集，研究人员可以测试新的算法和模型的性能，并与其他算法和模型进行比较。此外，该数据集还可用于演示如何构建和训练图像分类模型，从而帮助初学者快速入门深度学习和机器学习领域

形状, 箭头

描述已自动生成

# 2.2 LeNet-5模型

LeNet-5是一种经典的卷积神经网络模型，由Yann LeCun等人于1998年提出。它是第一个成功应用于手写数字识别任务的卷积神经网络，并在此后的几十年里被广泛使用于图像识别任务中。LeNet-5模型的结构相对简单，仅包含两个卷积层和三个全连接层，但其卓越的性能证明了卷积神经网络在图像识别任务中的优越性。

LeNet-5模型的结构如下：

输入层：图像的像素值

卷积层1：使用6个大小为5x5的卷积核进行卷积操作，输出6张特征图

池化层1：对卷积后的特征图进行2x2的最大池化操作，输出6张池化特征图

卷积层2：使用16个大小为5x5的卷积核进行卷积操作，输出16张特征图

池化层2：对卷积后的特征图进行2x2的最大池化操作，输出16张池化特征图

全连接层1：将池化后的特征图展平后输入到包含120个神经元的全连接层中

全连接层2：包含84个神经元的全连接层

输出层：包含10个神经元，对应于10个数字的概率输出值

LeNet-5模型的主要用途是手写数字识别任务，可以在MNIST数据集上达到较高的准确率。此外，LeNet-5模型的结构对于其他图像分类任务也有很好的迁移性，可以用于人脸识别、车辆识别、文字识别等应用领域。它的成功也为后来更加复杂的深度卷积神经网络模型的发展奠定了基础。

图示

描述已自动生成

# 2.3 卷积层

卷积层（Convolutional Layer）是卷积神经网络（Convolutional Neural Network, CNN）中的核心组件之一。卷积层可以有效地处理图像和其他类型的数据，对于图像分类、目标检测、语音识别等任务都有着广泛的应用。

卷积层的主要作用是从输入数据中提取特征。在图像处理中，卷积层将图像划分为不同的区域，并对每个区域进行卷积操作。卷积操作是指将卷积核（Convolution Kernel）应用于图像的每个像素，并将结果存储到输出特征图（Feature Map）中。

卷积层的用处主要有以下几个方面：

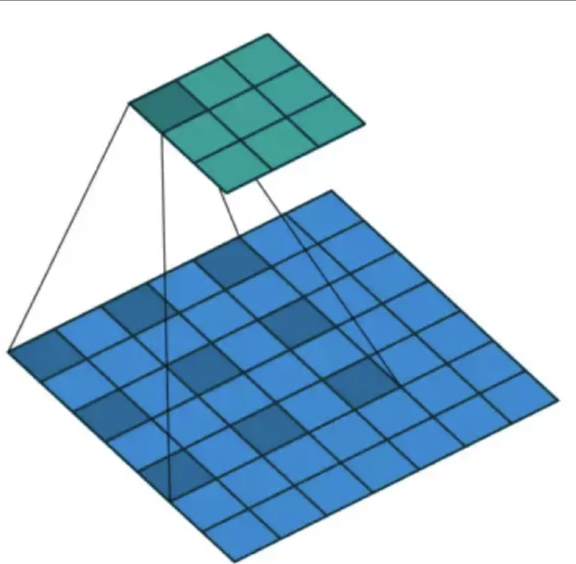
1.特征提取：卷积层通过卷积操作从输入数据中提取出了特征，使得模型可以更好地理解和识别图像、语音等数据。

2.参数共享：卷积层中的卷积核可以共享参数，这意味着在处理不同的区域时，使用相同的卷积核，从而大大减少了模型的参数数量，降低了过拟合的风险。

3.空间不变性：卷积层在卷积操作时只考虑了局部区域的信息，因此对于平移、旋转、缩放等图像变换具有一定的不变性，使得模型更加健壮和鲁棒。

4.多通道处理：卷积层可以对多通道的输入数据进行卷积操作，从而更好地捕捉图像的多层次信息，提高模型的表现能力。

卷积层是深度学习中非常重要的一种神经网络层，它能够自动学习输入数据的特征，从而实现高效的图像分类、目标检测等任务。



# 2.4池化层

池化层（Pooling Layer）是卷积神经网络（Convolutional Neural Network, CNN）中的一种常用层，用于减小特征图的空间大小，减少参数数量，从而提高模型的效率和鲁棒性。在卷积层之后通常会接一个池化层，两者共同构成了CNN中的特征提取部分。

池化层的作用是通过对输入数据的局部区域进行聚合操作，生成新的输出特征图。常用的池化操作有最大池化（Max Pooling）和平均池化（Average Pooling）。

池化层的用处主要有以下几个方面：

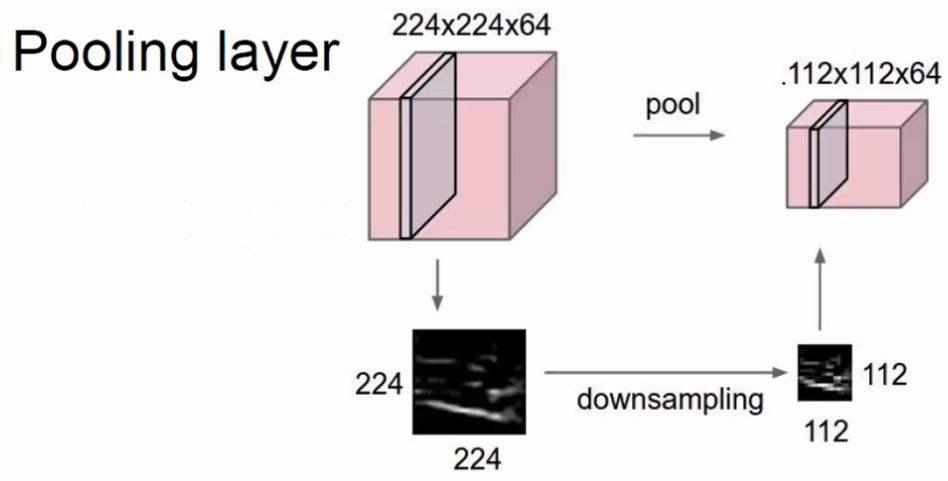
1.降维：池化层可以减小特征图的空间大小，从而降低计算量和内存占用，加速模型的训练和推理过程。

2.不变性：池化层可以使得模型对于输入数据的微小变化具有一定的不变性，例如平移、旋转、缩放等操作。

3.提取主要特征：池化层可以对输入数据的局部区域进行聚合操作，从而提取出主要的特征，减少噪声对模型的影响。

4.防止过拟合：池化层可以减少特征图中的参数数量，从而降低模型的复杂度和过拟合的风险。

总之，池化层是卷积神经网络中不可或缺的一部分，可以提高模型的效率、鲁棒性和泛化能力，从而实现更好的图像分类、目标检测等任务。



# 2.5全连接层

全连接层（Fully Connected Layer）是神经网络中的一种常见层，也称为密集连接层（Dense Layer）。全连接层将前一层的所有节点都连接到后一层的所有节点，是神经网络中最基本的层之一。

全连接层的作用是对前一层的特征进行线性组合，从而生成新的特征。在卷积神经网络（Convolutional Neural Network, CNN）中，全连接层通常在卷积层和池化层之后，用于将高维的特征图转换为一维向量，并进行分类、回归等任务。

全连接层的用处主要有以下几个方面：

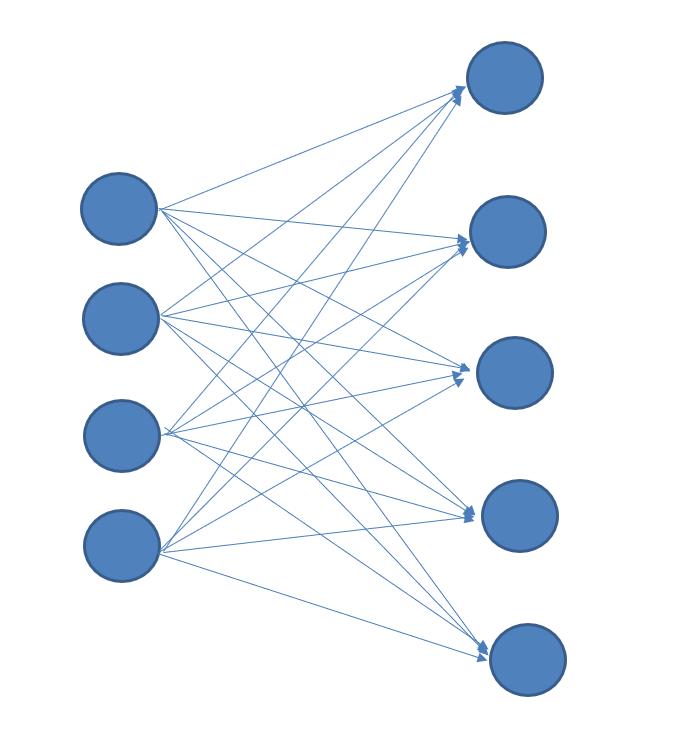
1.特征提取：全连接层可以对前一层的特征进行线性组合，提取出更高层次的特征，从而更好地捕捉数据的复杂性。

2.分类、回归等任务：全连接层可以将高维的特征图转换为一维向量，并进行分类、回归等任务。

3.参数优化：全连接层的参数数量较多，可以使得模型更加灵活和可调节，从而优化模型的性能。

4.提高表现能力：全连接层可以将前一层的特征进行非线性组合，从而增加模型的表现能力。

总之，全连接层是神经网络中重要的一种层，可以对输入数据进行高维特征提取和非线性组合，从而实现分类、回归等任务。在卷积神经网络中，全连接层通常与卷积层和池化层组合使用，构成了神经网络的完整架构。



# 2.6激活函数

激活函数主要用于对输入数据进行非线性变换，使神经网络能够拟合更加复杂的函数关系。在神经网络中，每个神经元都有一个激活函数，其作用是对输入信号进行加权求和后进行非线性变换，从而生成输出信号。常见的激活函数包括sigmoid函数、ReLU函数、tanh函数等。

激活函数的作用主要有以下几个方面：

1.引入非线性：激活函数可以引入非线性，从而使神经网络可以拟合更加复杂的函数关系。

2.增强模型表现力：激活函数可以增强模型的表现力，从而提高模型的性能。

3.避免梯度消失：某些激活函数可以避免梯度消失问题，从而加速模型的训练过程。

# 2.7损失函数

损失函数主要用于衡量模型预测结果与真实结果之间的差异，从而指导模型参数的优化。在神经网络中，损失函数通常用于计算模型的误差，并通过反向传播算法更新模型参数。常见的损失函数包括均方误差（Mean Squared Error, MSE）、交叉熵（Cross-entropy）等。

损失函数的作用主要有以下几个方面：

1.评估模型性能：损失函数可以用于评估模型的性能，从而指导模型的训练过程。

2.引导参数优化：损失函数可以引导参数优化，使得模型在训练过程中逐渐接近最优状态。

3.避免过拟合：某些损失函数可以避免过拟合问题，从而提高模型的泛化能力。

# 2.8优化器

在机器学习中，优化器是一种用于更新模型参数的算法。在训练过程中，优化器根据当前的损失函数和参数状态，计算出一个更新方向和步长，以使得损失函数最小化。常见的优化器包括梯度下降法、Adam、Adagrad、RMSprop等。

优化器的作用是在训练模型时，通过自适应地调整参数的学习率和步长，使得模型更快地收敛到最优解。在使用优化器时，通常需要根据具体的任务和模型选择适合的优化器。例如，Adam优化器对于处理稀疏梯度和噪声的情况效果较好，而Adagrad适用于处理具有不同特征稀疏度的数据。

优化器在深度学习中具有重要的作用，不仅可以提高模型的训练效率和准确率，还可以防止模型陷入局部最优解。在实践中，通常需要对不同的优化器进行试验和比较，以得出最佳的优化效果。

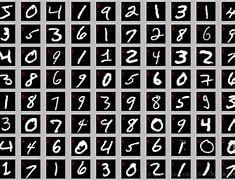
# 三． 实验步骤

图示

描述已自动生成

# 四． 数据集

本次实验使用的是MNIST数据集，一个手写数字图像数据集，由70,000张图像组成，每张图像大小为28x28像素。这些图像被分为60,000张用于训练和10,000张用于测试。每个图像都标记为一个数字，从0到9之间的整数，表示图像中的手写数字。



# 五． 程序代码

1.导入需要的包

图片包含 文本

描述已自动生成

2. 加载MINIST数据集，并且划分测试训练集

电脑屏幕截图

描述已自动生成

3.定义模型类图片包含 文本

描述已自动生成

4.定义超参数

文本

描述已自动生成

5.开始训练

文本

描述已自动生成

6.进行测试并打印



# 六． 实验结果

# 6.1 训练损失

图形用户界面, 文本

描述已自动生成

# 6.2 测试结果

文本

描述已自动生成

# 七．实验分析与总结

* **MINIST数据集：**

MINIST数据集是一个常用的手写数字识别数据集，它由60000张训练图片和10000张测试图片组成。每张图片都是28x28像素的灰度图像，代表了0-9十个数字中的一个。

在使用MINIST数据集进行机器学习实验时，可以得到以下几点结论：

1.MINIST数据集是一个相对简单的数据集，由于数字是相对规整的，且图像大小较小，因此相对容易进行分类。

2.由于MINIST数据集已经是广泛使用的数据集，因此有很多优秀的算法和模型可以直接使用或参考，例如SVM、CNN、LeNet-5等。

3.在使用MINIST数据集进行实验时，需要注意过拟合问题，特别是对于一些较为复杂的模型，如卷积神经网络。可以通过数据扩增等方法来减少过拟合问题。

4.MINIST数据集可以用于学习和实验不同的机器学习算法和模型，如分类、聚类、降维等。同时也可以用于测试一些新的算法和模型的效果。

由于MINIST数据集是相对简单的数据集，因此其实验结果并不一定能够代表在更复杂的数据集上的表现，需要结合其他数据集进行评估。

* **LeNet-5 模型：**

LeNet-5是一种经典的卷积神经网络模型，用于手写数字的识别，也是第一个被广泛应用的卷积神经网络模型之一。它由卷积层、池化层、全连接层等组成，采用了多层卷积与池化，并使用了ReLU激活函数，从而有效地降低了过拟合问题，提高了模型的准确性。

在使用LeNet-5模型进行机器学习实验时，可以得到以下几点结论：

1.LeNet-5是一种较为简单但有效的卷积神经网络模型，适用于手写数字的识别等图像分类问题。

2.LeNet-5模型在设计中采用了多层卷积与池化，可以提取图像的特征并降低维度，从而进一步降低了模型的复杂度和计算量。

3.LeNet-5模型采用了ReLU激活函数，可以在不引入额外参数的情况下，提高模型的非线性表达能力。

4.LeNet-5模型在训练时需要注意对超参数的调节，如学习率、批量大小等，这些参数的调节会对模型的训练速度和准确率产生影响。

LeNet-5模型是一种基于手工设计的模型，需要人工设计和调整网络结构和参数，这在一定程度上限制了其应用范围和效果。随着深度学习的发展，人们提出了很多自适应网络结构的模型，如VGG、ResNet等。

* **实验结果分析：**

在训练过程中，LeNet-5 模型能够快速收敛，测试集平均准确率达到了 97.03%。同时，损失函数在训练过程中不断下降，表明模型的拟合效果较好。总体来说，LeNet-5 在 MNIST 数据集上表现良好，达到了较高的准确率和较低的损失函数。