图标

描述已自动生成

**计算机视觉**

**上机实验二报告**

|  |  |  |  |
| --- | --- | --- | --- |
| **专业** | **：** | **计算机电脑** |  |
| **班级** | **：** | **CS2002** |  |
| **学号** | **：** | **I201920029** |  |
| **姓名** | **：** | **冯就康** |  |
| **电话** | **：** | **15623031879** |  |
| **邮箱** | **：** | **sokhorng526@gmail.com** |  |

Contents

[I. 简介 3](#_Toc154479538)

[1. 卷积神经网络的简介 3](#_Toc154479539)

[2. ResNet模块的简介 4](#_Toc154479540)

[II. 网络的架构 5](#_Toc154479541)

[1. 模型的结构 5](#_Toc154479542)

[2. 残差块的结构 5](#_Toc154479543)

[3. 模型的权重和偏差 5](#_Toc154479544)

[III. 模型训练过程 6](#_Toc154479545)

[1. 数据准备： 6](#_Toc154479546)

[2. 损失函数和优化器： 6](#_Toc154479547)

[3. 训练循环： 6](#_Toc154479548)

[IV. 实验结果以及分析 7](#_Toc154479549)

[1. 实验结果 7](#_Toc154479550)

[2. 实验分析 7](#_Toc154479551)

[V. 实验总结 8](#_Toc154479552)

[1. 实验总结： 8](#_Toc154479553)

# 简介

## 卷积神经网络的简介

**卷积神经网络**（**Convolutional** **Neural** **Networks**，简称**CNN**）是一种深度学习算法，主要用于图像分类和计算机视觉任务。它由卷积层、池化层和全连接层组成。

* **卷积层**---是卷积神经网络的核心构建块，负责执行大部分计算。
* **池化层**---执行降维操作，旨在减少输入中参数的数量。
* **全连接层**---根据通过先前层及其不同的过滤器提取的特征，执行分类任务。

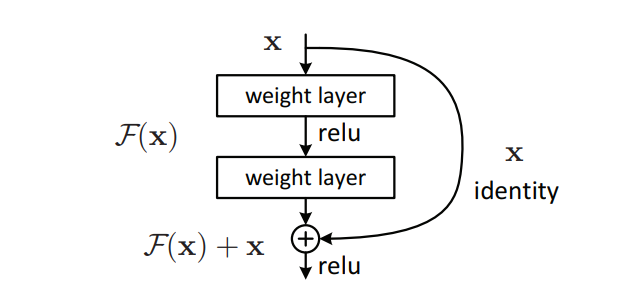
A diagram of a diagram of a process

Description automatically generated with medium confidence**卷积神经网络**利用线性代数原理（特别是矩阵乘法）来识别图像内的图案。**卷积神经网络**的结构是一个分层结构，随着层级的递进，**卷积神经网络**的复杂性也逐步增加，能够识别图像的更多部分。**卷积神经网络**的出现，提供了更加可扩展的方法来执行图像分类和对象识别任务。

## ResNet模块的简介

残差网络(**ResNet**)是一种深度神经网络架构。**ResNet**旨在解决当我们增加深度神经网络的层数时出现的梯度消失问题。梯度消失问题导致梯度变得太大或为零，从而导致训练和测试错误率增加。

**ResNet**引入了**Residual** **Blocks**的概念来解决梯度消失问题。在此架构中，我们使用一种称为“跳过连接”的技术，该技术通过跳过中间的某些层将层的激活连接到其他层。这形成了一个残差块，**ResNets** 就是通过将这些残差块堆叠在一起而形成的。该网络背后的方法是让网络适应残差映射，而不是学习底层映射的层。

****残差网络用于各种应用，例如图像分类、对象检测和语义分割。

# 网络的架构

## 模型的结构

整个模型由一系列卷积层、残差块和全连接层组成。

* **nn.Conv2d** 函数用于定义第一个卷积层，该层接收大小为 28x28 像素的灰度图像。 **nn.BatchNorm2d**函数用于对卷积层的输出进行归一化。
* **self.\_make\_layer** 函数用于创建残差块。每个残差块由两个卷积层和绕过一个或多个卷积层的跳跃连接组成。
* **stride**参数用于指定每个残差块中第一个卷积层的步长。
* **nn.function.avg\_pool2d** 函数用于对最后一个残差块的输出执行平均池化。
* **out.view**函数用于压平平均池化层的输出。
* 最后，**nn.Linear**函数用于定义全连接层，该层产生模型的最终输出。

## 残差块的结构

残差块由两个卷积层和绕过一个或多个卷积层的跳跃连接组成。

* **nn.Conv2d** 函数用于定义第一个卷积层，该层接受具有 **in\_channels** 个通道的输入张量。 **nn.BatchNorm2d**函数用于对第一个卷积层的输出进行归一化。
* **self.shortcut** 函数用于定义跳过连接。如果第一个卷积层的步长不等于1或者输入通道数不等于输出通道数，则跳过连接由卷积层和批量归一化层组成。
* **forward**函数用于定义**Residual Block**的前向传播。输入张量经过第一个卷积层、批量归一化层和**ReLU**激活函数。
* 然后，第一个卷积层的输出通过第二个卷积层和批量归一化层。第二个卷积层的输出被添加到跳跃连接的输出中。
* 最后，**Residual Block** 的输出通过 **ReLU** 激活函数。

## 模型的权重和偏差

**ResNet** 模型由多个卷积层和残差块组成，每个层都有自己的一组权重和偏差。权重和偏差是在训练过程中使用反向传播学习的。

**ResNet**模型中的权重用于调整网络中神经元之间的连接强度。偏差用于将每个神经元的输出移动一定量。

在**ResNet**模型的前向传递过程中，输入图像经过卷积层和残差块，最终残差块的输出经过全连接层，产生模型的最终输出。卷积层和残差块的权重和偏差用于从输入图像中提取特征，而全连接层的权重和偏差用于对输入图像进行分类。

# 模型训练过程

## 数据准备：

我们要使用的数据集是 **MNIST** 数据集。 **MNIST** 数据集是一个大型手写数字数据库，通常用于训练各种图像处理系统。它还广泛用于机器学习领域的训练和测试。该数据集包含 60,000 张训练图像和 10,000 张测试图像。图像为灰度图像，分辨率为 28x28 像素。我们可以从**torch**.**vision**下载**MNIST**数据集，并用**torchvison**.**transform**.**ToTensor**()把数据变成**Tensor**。

## 损失函数和优化器：

* **nn.CrossEntropyLoss** 函数用于定义损失函数，用于计算预测标签和实际标签之间的差异。
* **optim.Adam**函数用于定义优化器，用于在训练期间更新模型的权重。

## 训练循环：

**MNIST** 数据集上的 **ResNet** 模型的训练循环。

由于数据集是图像，因此很大。所以我们首先指定 **device** 变量来指定是使用 **GPU** 还是 **CPU** 进行训练。 **ResNet**函数用于定义**ResNet**模型，然后使用**to**方法将其移动到指定设备。

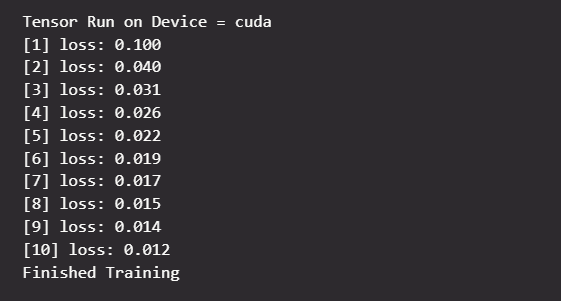
训练循环由两个嵌套循环组成。外循环迭代指定数量的 **epoch**，而内循环则分批迭代训练数据。 **enumerate** 函数用于批量迭代训练数据。

使用 **to** 方法将输入图像和标签移动到指定设备。优化器使用 **zero**\_**grad** 方法进行归零。模型的输出是使用输入图像计算的，损失是使用预测标签和实际标签计算的。使用后向方法计算梯度，并使用步进方法更新模型的权重。

运行损失在内循环的每次迭代时更新，并在每个时期结束时打印平均运行损失。训练循环重复指定的时期数。

# 实验结果以及分析

## 实验结果

经过**10个Epochs**训练的训练集的损失值：

A screenshot of a black screen

Description automatically generated测试集每个数字的准确率：

## 实验分析

该模型的总体准确率达到了99.43%。

该模型对数字0、3、8和9实现了最高准确率，准确率分别为99.8%、99.8%、99.69%和99.31%。该模型对数字 4 的准确率最低，准确率为 98.98%。

总体而言，**ResNet**模型在 **MNIST** 测试数据集上的确能取得一个了非常高的准确率，这表明它在分类手写数字方面的效率非常好。

# 实验总结

## 实验总结：

**ResNet**架构在实验中的准确率达到了99.43%。结果表明 **ResNet** 架构可以有效地对手写数字进行分类。然而，值得注意的是，该模型在其他数据集或实际应用中可能表现不佳。

在训练过程中，面临的挑战之一是过度拟合，当模型在训练数据上表现良好但在测试数据上表现不佳时就会发生这种情况。为了应对这一挑战，使用了 **dropout** 和数据增强等技术来防止模型记忆训练数据并提高其泛化性能。

训练过程中面临的另一个挑战是梯度消失问题，当梯度变得太小而无法用于更新模型权重时就会发生这种情况。为了应对这一挑战，**ResNet** 架构引入了残差块的概念，它使用跳跃连​​接来允许梯度更轻松地流过网络。

总体而言，实验获得的结果证明了 **ResNet** 架构对于 **MNIST** 数据集分类任务的有效性。然而，需要进一步的实验来确定模型在其他数据集和实际应用中的性能。