

# 华中科技大学

## 2023

### 计算机视觉实验

### · 实验报告 ·

专    业：          计算机科学与技术

班    级：          CS2106

学    号：          U202115514

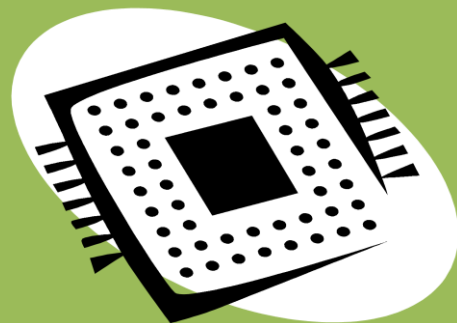
姓    名：          杨明欣

电    话：          13390396012

邮    件：          ymx@hust.edu.cn

指导老师：          刘康

完成日期：          2023-12-25



计算机科学与技术学院

## 目录

1	实验内容 .....	2
1.1	实验要求 .....	2
1.2	实验内容 .....	2
1.2.1	卷积神经网络 .....	2
1.2.2	残差神经网络 .....	2
1.3	实验方法 .....	3
1.3.1	构建数据集 .....	3
1.3.2	网络架构设计 .....	4
1.4	实验设计 .....	5
1.4.1	数据集 .....	5
1.4.2	实验环境 .....	5
1.5	实验分析与结果 .....	5
1.5.1	数据集评估 .....	5
1.5.2	分类性能 .....	6
1.6	实验总结 .....	7
2	实验心得 .....	8

## 1 实验内容

### 1.1 实验要求

设计一个卷积神经网络，输入为两张 MNIST 手写体数字图片，如果两张图片为同一个数字（注意，非同一张图片），输出为 1，否则为 0。

具体要求：从 MNIST 数据集的训练集中选取 10% 作为本实验的训练图片，从 MNIST 数据集的测试集中选取 10% 作为本实验的测试图片。请将该部分图片经过适当处理形成一定数量的用于本次实验的训练集和测试集。

### 1.2 实验内容

#### 1.2.1 卷积神经网络

卷积神经网络（Convolutional Neural Networks，简称 CNN）是一种具有局部连接、权值共享等特点的深层前馈神经网络（Feedforward Neural Networks），是深度学习（deep learning）的代表算法之一，擅长处理图像特别是图像识别等相关机器学习问题，比如图像分类、目标检测、图像分割等各种视觉任务中都有显著的提升效果，是目前应用最广泛的模型之一。

卷积神经网络（CNN）具有表征学习能力，能够按其层次结构对输入信息进行平移不变分类。它适用于监督学习和非监督学习，并通过卷积核参数共享和层间连接的稀疏性，以较小的计算量对像素和音频等格点化特征进行学习。CNN 能够稳定地实现分类任务，无需额外的特征工程，因此在计算机视觉、自然语言处理等领域得到广泛应用。

#### 1.2.2 残差神经网络

深度残差网络（Deep Residual Networks，ResNet）自从 2015 年首次提出以来，就在深度学习领域产生了深远影响。通过一种创新的“残差学习”机制，ResNet 成功地训练了比以往模型更深的神经网络，从而显著提高了多个任务的性能。深度残差网络

# 华中科技大学课程实验报告

---

通过引入残差学习和特殊的网络结构，解决了传统深度神经网络中的梯度消失问题，并实现了高效、可扩展的深层模型。

传统的深度神经网络试图学习目标函数 ( $H(x)$ )，但是在 ResNet 中，每个网络层实际上学习的是一个残差函数 ( $F(x) = H(x) - x$ )。然后，这个残差结果与输入 ( $x$ ) 相加，形成 ( $H(x) = F(x) + x$ )。这一机制使得网络更容易学习身份映射，进而缓解了梯度消失问题。

## 1.3 实验方法

本实验需要设计一个卷积神经网络，输入为两张 MNIST 手写体数字图片，如果两张图片为同一个数字（注意，非同一张图片），输出为 1，否则为 0。总结来说就是针对新构成的两张图片作为输入的二分类数据集进行分类。为了实现这一任务，本实验在此分成两个部分，第一部分合理构建数据集，第二部分则为正确设计网络针对数据集完成分类任务。

### 1.3.1 构建数据集

一个高质量的数据集往往能够提高模型训练的质量和预测的准确率。本实验中由于没有现成的数据集，因此需要在 MNIST 数据集的基础上构建新的数据集。在这个部分，我们尝试了两种构架数据集的方法，第一种方法是简单地构建标签为 0 和 1 的数据各一半，其中两个数字为从训练数据（测试数据）随机选取，经实验发现效果很差，猜测为标签均衡并不代表数据均衡，因此接下来着重介绍第二种构建数据集的理由、具体方法和结果。

第二种构建数据集的理由是为了在让标签均衡的情况下，让数据更加均衡。首先从 MNIST 数据集的训练集中随机选取 10% 作为本实验的训练图片，从 MNIST 数据集的测试集中随机选取 10% 作为本实验的测试图片。进一步以训练数据为例，具体方法就是第一步取出每个数字类别的全部数字，第二步构建两张图片的标签相同的数据集（包括 0|0, 1|1, ..., 9|9 共 10 种情况），数据量为 500000，每个数字的数据量为 50000，具体数字为在每个数字类别中随机选取。第三步构建两张图片的标签不同的数据集（包括 0|1, 0|2, ..., 8|9 共 45 种情况），数据量为 450000，每种情况的数据量为 10000，

# 华中科技大学课程实验报告

同样根据 45 种数字组合，具体数字为在每个数字类别中随机选取。总体上标签为 0 和标签为 1 的数据比接近 1:1，同时每个数字出现的次数相同。测试集进行相同的操作。

最终构建数据集的结果如表 1-1 所示。

表 1-1 构建数据集的结果

指标	训练集	测试集
数量	950000	95000
标签为 0 数量	450000	45000
标签为 1 数量	500000	50000
每个数字出现 的数量	190000	19000

## 1.3.2 网络架构设计

针对前一部分构建的数据集，我们设计实现了一个模型架构实现对于数据的分类任务。模型架构如图 1-1 所示。

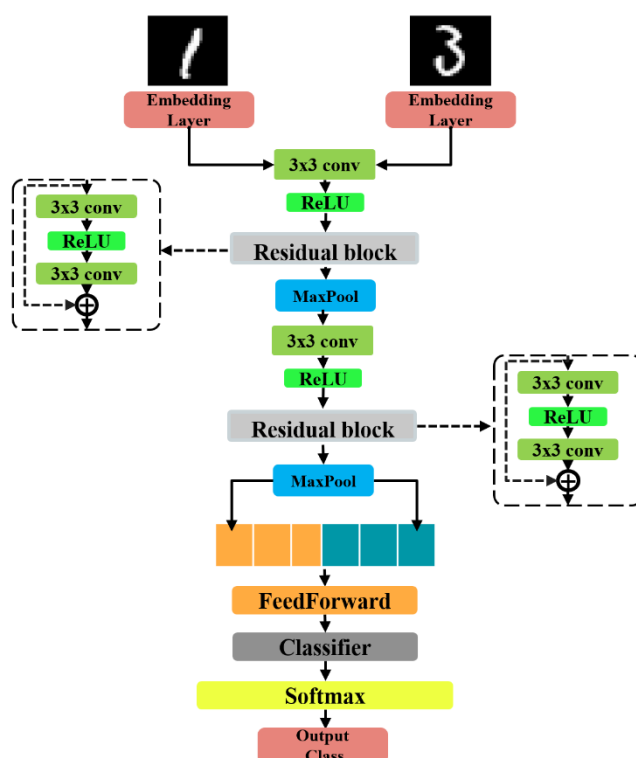


图 1-1 模型架构示意图

# 华中科技大学课程实验报告

---

根据图 1-1 所示，模型架构主要包括两个部分，第一个部分为特征提取部分，由两张图片共享参数，主要包括卷积层和残差块，实现对于图像特征的提取，最终获得图像的特征向量，第二个部分为标签分类部分，将两张图像的拼接在一起，通过一层全连接层和线性分类层获得最终的标签结果。

## 1.4 实验设计

### 1.4.1 数据集

本实验采用自行构建的数据集，数据总量为 1045000。

接下来首先将数据集进行打乱（shuffle），训练集：测试集 = 10:1，以备模型训练和测试。

### 1.4.2 实验环境

系统版本：Windows 11 专业版

处理器：13th Gen Intel(R) Core(TM) i9-13900HX 2.20 GHz

机带 RAM：16.0 GB (15.7 GB 可用)

系统类型：64 位操作系统，基于 x64 的处理器

显卡版本：NVIDIA GeForce RTX 4060 Laptop GPU

显存：8188MiB

CUDA 版本：12.0

Python 版本：3.9.18

torch 版本：2.1.1+cu118

## 1.5 实验分析与结果

### 1.5.1 数据集评估

查询资料可得一个高质量的数据集应该具备以下特点：

1. 代表性：数据集应能代表所研究的对象或问题，也就是说，数据的分布应该广泛，包括各种不同的情况和特征。

2. 完整性：数据集应包含足够的数据样本，以保证分析和模型的准确性。数据样本数量越多，模型的预测能力和稳定性通常会更好。
3. 可解释性：数据集中的特征和目标变量应该有明确的含义，使得研究人员或机器学习算法能够理解和使用这些数据。
4. 平衡性：对于分类问题，数据集中的各类样本数量应该大致相等。这样可以避免模型在训练和测试时出现类别不平衡的问题。

分析构建的数据集可知，我们的数据集充分保证了数据平衡和标签平衡，构建数据集的方式和结果具有可解释性，数据量可观，能够代表要研究的问题的大致情况，数据完成，对于模型的训练来说质量较高。

## 1.5.2 分类性能

对于 1.3.2 中的网络架构进行训练，使用交叉熵损失函数（Cross\_entropy），使用 SGD 优化器，epoch 为 4 轮，batchsize 为 128，学习率设置为 0.0001，权重衰减设置为 0.0001，训练过程中训练集的损失和准确率如图 1-2 所示。

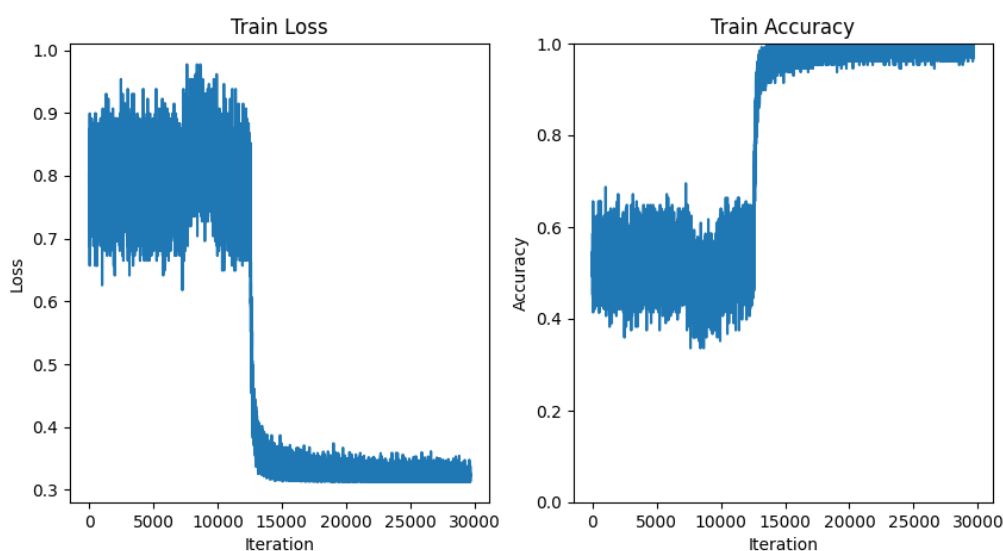


图 1-2 训练过程中训练集的损失和准确率

最终在训练集上进行测试可得整体的准确率为 99.27%，其中标签为 0 的准确率为 99.19%，标签为 1 的准确率为 99.34%。

训练过程中测试集的损失和准确率如图 1-3 所示。

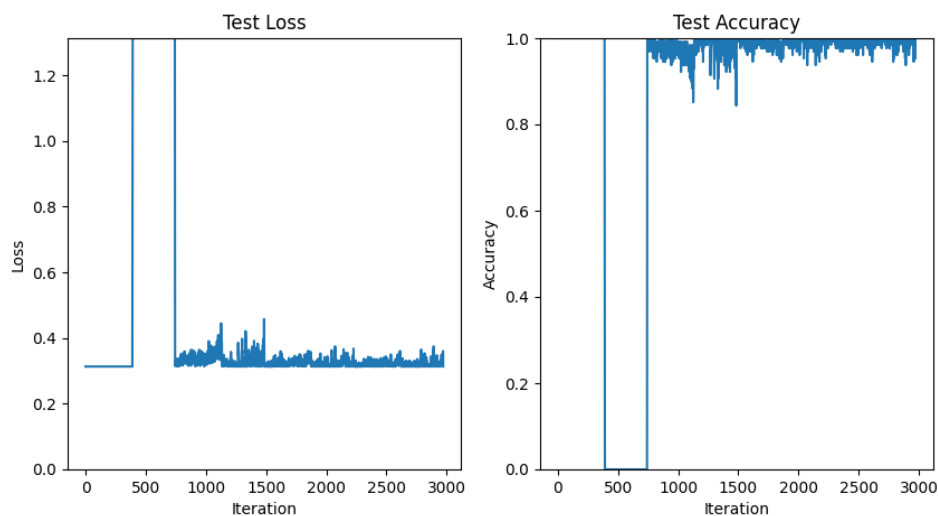


图 1-3 训练过程中测试集的损失和准确率

最终在测试集上进行测试可得整体的准确率为 99.13%，其中标签为 0 的准确率为 99.12%，标签为 1 的准确率为 99.14%。

最终可以得出结论我们的模型最终收敛到了一个比较好的效果，能够很好地完成二分类任务。

## 1.6 实验总结

本次实验主要完成了如下几点工作：

- 1) 在 MNIST 数据集的基础上继续构建新的数据集，输入为两张 MNIST 手写体数字图片，如果两张图片为同一个数字（注意，非同一张图片），输出为 1，否则为 0。
- 2) 设计实现了一个网络架构，对上述构建好的数据实现分类任务。



## 2 实验心得

在完成基于新构建的二分类手写数字识别数据集实现分类任务的实验过程中，我获得了许多宝贵的经验。

- 1) 高质量数据集构建对于实验成功的重要性。机器学习算法通过从大量的数据中学习模式和规律，从而实现对未知数据的预测和分类。一个好的数据集可以提供丰富多样的样本和标签，帮助算法更好地理解问题的本质。通过不断调整数据分布，我逐渐找到了适合当前模型的最佳数据集构建方式。这个过程需要耐心和细心。
- 2) 学会如何根据构建好的数据集设计神经网络，同时要多进行尝试，比如在本实验设计的模型架构中，我惊人地发现的当模型中多加一层全连接层时，模型的收敛速度和准确率都有了显著提高。

总的来说，这次实验让我认识到了构建高质量数据集的重要意义。我不仅学会了如何构建和训练神经网络模型，还掌握了如何构建高质量的数据，收获很大。

## • 指导教师评定意见 •

---

### 一、原创性声明

本人郑重声明本报告内容，是由作者本人独立完成的。有关观点、方法、数据和文献等的引用已在文中指出。除文中已注明引用的内容外，本报告不包含任何其他个人或集体已经公开发表的作品成果，不存在剽窃、抄袭行为。

特此声明！

作者签字：杨明欣



### 二、对课程实验的学术评语（教师填写）

### 三、对课程实验的评分（教师填写）

评分项目 (分值)	课程目标 1 工具应用 (10 分)	课程目标 2 设计实现 (70 分)	课程目标 3 验收与报告 (20 分)	最终评定 (100 分)
得分				

指导教师签字：\_\_\_\_\_