

**中国地质大学（武汉）**

**模式识别与机器学习实验报告**

**学 院： 学院**

**课 程：**

**指导老师：**

**学 号：**

**班 级：**

**姓 名：**

**2022年12月27日**

**目录**

[第一章 引言 1](#_Toc87533944)

[1.1实验目的 1](#_Toc87533945)

[1.2 实验意义 1](#_Toc87533946)

[1.3 国内外研究现状 1](#_Toc87533947)

[第二章 实验内容 1](#_Toc87533948)

[2.1所选题目 2](#_Toc87533949)

[2.2 报告的内容及要求 2](#_Toc87533950)

[第三章 实验方法 3](#_Toc87533951)

[3.1实验原理 3](#_Toc87533952)

[3.2 实验流程 4](#_Toc87533953)

[第四章 实验结果与分析 5](#_Toc87533954)

[4.1实验结果 5](#_Toc87533955)

[4.2 分析与评价 6](#_Toc87533956)

[第五章 总结 7](#_Toc87533957)

[5.1收获与体会 7](#_Toc87533958)

[5.2 遇到的问题及解决方案 7](#_Toc87533959)

[5.3 参考文献 7](#_Toc87533960)

# 第一章 引言

1.1实验目的

1、使学生进一步巩固模式识别基本的概念。

2、掌握相关实验的工具箱，能熟练使用工具箱中的各种功能。

3、增强学生应用Python编写程序及分析、解决实际问题的能力。

1.2 实验意义

实验引导学生重温模式识别方法的基本概念和原理，并在计算机上进行验证实验，实际操作和观察图像模式识别的过程和运行结果，通过实验使学生掌握应用模式识别基本原理和方法解决实际问题的基本技巧，培养学生应用理论知识解决实际问题的能力。

## 1.3 国内外研究现状

近年来，手写数字的识别是模式识别和人工智能领域的研究热点，是一种利用计算机自动识别手写数字的技术．该技术被广泛应用到邮政编码、财务报表、税务系统数据统计、银行票据等手写数字自动识别录入中。由于不同的手写者字 体不同，对大量的手写数字实现完全正确识别不是一件容易的事情．随着信息化技术的不断发展，人们对自动识别的要求越来越高．因此，研究一种高效且准 确 的手写数字识别方法有着重要的意义。

对于国外来说，从最开始的LeNet，到后来深度更深、宽度更宽的AlexNet、GoogLeNet，再到ResNet、MobileNet等深度残差卷积神经网络，手写数字识别的准确率越来越高。

# 第二章 实验内容

2.1所选题目

手写数字识别

2.2 报告的内容及要求

报告的格式内容如下：

1. 实验意义

这部分说明基于本实验在生活有什么实际意义。

1. 国内外研究现状

调研人工智能与模式识别国内外研究背景，发展现状。

1. 实验原理和实验流程

解释实验的原理、程序代码、注释、程序框图。

1. 实验结果及分析

实验结果截图及对结果的分析。

1. 总结

实验中遇到的问题及解决方案，学到的东西和取得的经验总结，心得体会。

报告要求：

1. 必须按照以上要求完成实验报告。
2. 注意实验字体，整个实验报告的排版。
3. 独立完成实验报告。

# 第三章 实验方法

3.1实验原理

MNIST数据集是从NIST的两个手写数字数据集：Special Database 3 和Special Database 1中分别取出部分图像，并经过一些图像处理后得到的。MNIST数据集共有70000张图像，其中训练集60000张，测试集10000张。所有图像都是28×28的灰度图像，每张图像包含一个手写数字。

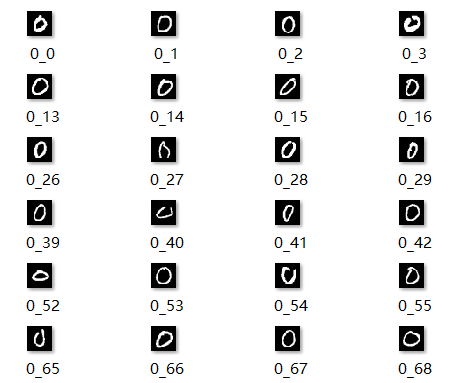


图 3.1 MNIST样张

LeNet[1]、AlexNet[2]、VGG[3]是经典的卷积神经网络，其原理到现如今还在被广泛应用，通过阅读相关论文可以了解这类神经网络的结构及具体的参数，使用PyTorch框架搭建神经网络对MNIST数据集进行识别。



图 3.2 LeNet（左）、AlexNet（右）

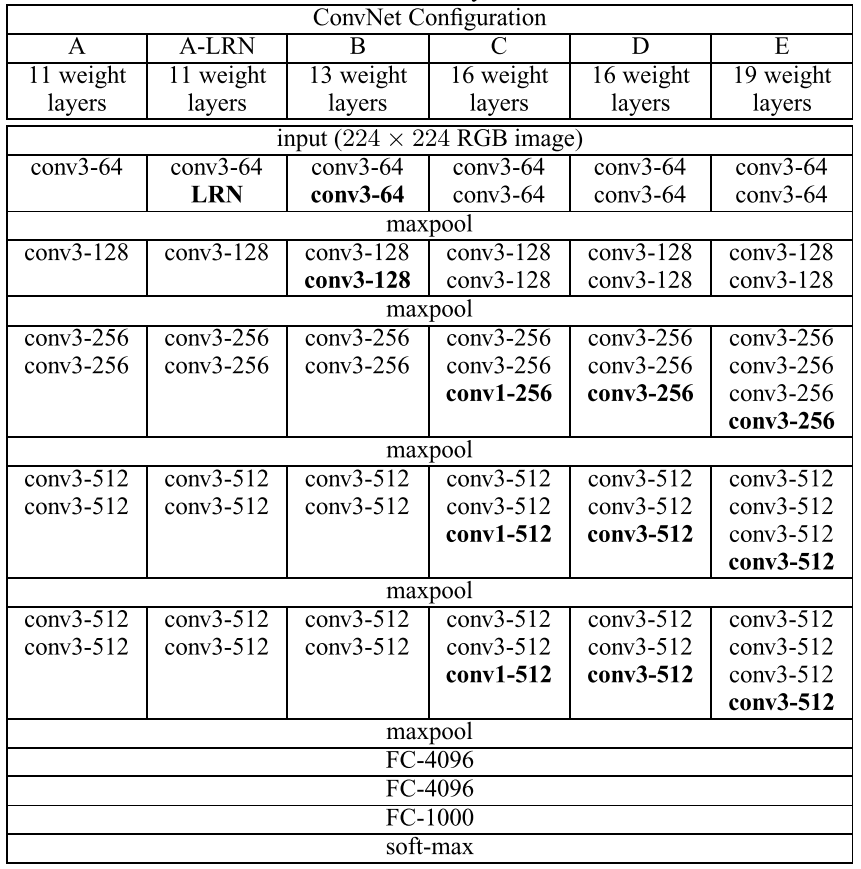


图 3.3 VGG各层参数

3.2 实验流程

首先确定使用Python作为编程语言，使用PyTorch作为时间平台框架，分别搭建并使用LeNet、ALexNet、VGG13、VGG16、VGG19等深度卷积神经网络模型在MNIST数据集上进行训练，并比较不同模型的识别率。

训练过程中，先对图片进行预处理，例如随机裁剪、随机水平翻转、归一化，并将图片转化为Tensor类型。将数据集分为训练集和测试集，分别设计train函数和test函数进行模型的训练和测试。最后编写可视化代码，观察Less曲线和acc曲线，并对比不同模型的效果。

第四章 实验结果与分析

4.1实验结果

实验结果如表4.1所示，其中识别率排名前两位的VGG13和VGG16的Loss曲线和acc曲线如图4.1和图4.2所示。

表 4.1 实验结果

|  |  |
| --- | --- |
| 模型 | 测试集准确率 |
| LeNet | 99.49% |
| AlexNet | 99.37% |
| VGG13 | 99.79% |
| VGG16 | 99.71% |
| VGG19 | 99.70% |

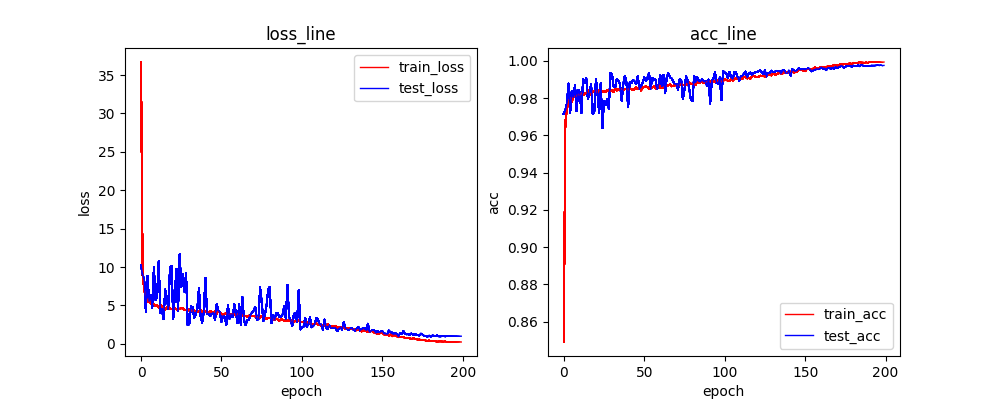
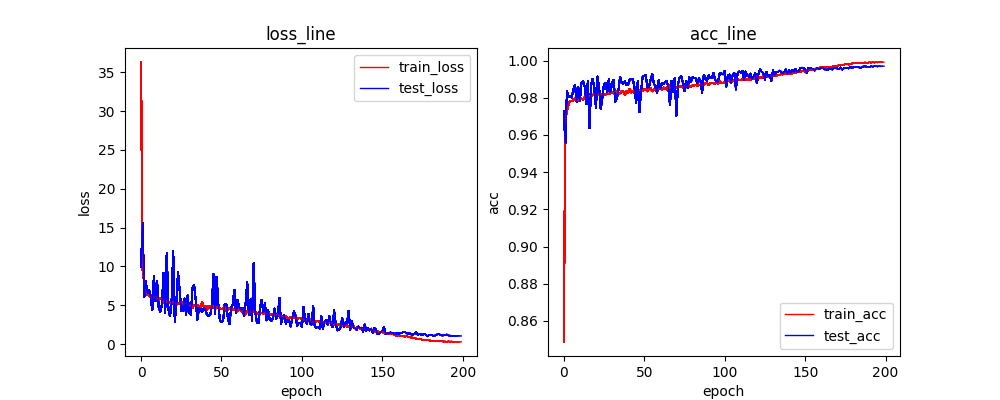


图 4.1 VGG13的Loss曲线（左）、准确率曲线（右）

图 4.2 VGG16的Loss曲线（左）、准确率曲线（右）

4.2 分析与评价

根据实验结果（表4.1），我们可以看出，所有的模型准确率均在99%以上，整体的识别率还是比较高的，效果比较好，可能是因为MNIST数据集的图片是灰度图只有1个通道，并且测试集的数据和训练集的数据长的比较像，可能存在一定的过拟合。其中VGG13的准确率最高。同时我发现，AlexNet相比于LeNet深度更深、宽度更宽，但是准确率却没有LeNet高；深度更深的VGG19和VGG16的准确率也没有VGG13高。原因可能是更深的卷积神经网络可能有更大的识别空间（如图4.3），但是识别空间并没有真正的向我们的最优解靠拢，使用ResNet可能可以解决这一问题。

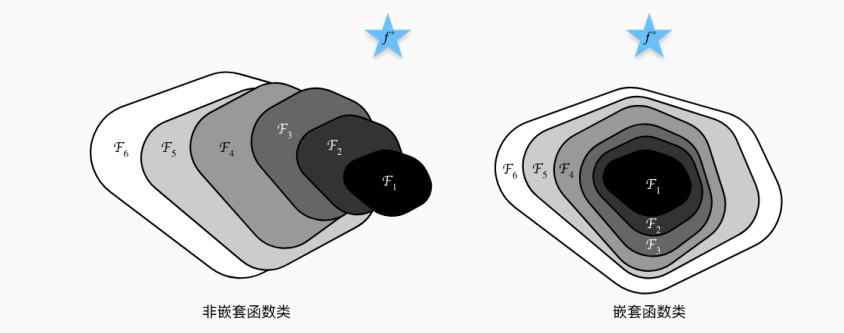


图 4.3 嵌套函数类和非嵌套函数类

第五章 总结

5.1收获与体会

通过手写数字实验，我对深度学习中一些经典的网络结构以及典型深度学习训练技术（Dropout、BN等）有了初步的了解。在使用PyTorch搭建神经网络的过程中，参考了一些网上的代码，但是发现网上部分代码最终的效果并不好（猜测是因为搭建的模型与论文中有较大的差异），于是对着论文中的模型结构一步步查看代码是否有误，将错误修改后，发现部分模型的识别率相较于之前确实有较大的提升。在这过程中我也慢慢锻炼了阅读文献和在文献中提取重点的能力。

5.2 遇到的问题及解决方案

在使用VGG对MNIST数据集进行识别的过程中发现程序报错。原因是MNIST数据集中图像的大小为28×28，但是在VGG中进行了5次下采样，所以在第五次下采样的过程中会出现参数不够的问题。

解决方案：在对图片进行预处理的过程中，将图片Resize成32×32的大小。

## 5.3 参考文献

[1] Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.

[2] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.

[3] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition[M]. arXiv, 2015.