# 实验四 基于卷积神经网络的图像分类算法

## 一、实验目的

1. 了解卷积神经网络的工作原理、搭建流程。

2. 了解基于MATLAB的卷积神经网络搭建流程，能够通过有规则地将不同网络层搭配在一起，从而实现图像分类任务。

3. 了解一般分类卷积神经网络中各种超参数的作用，能够通过调整这些超参数提升神经网络的性能。

## 二、实验平台

本节分别介绍本实验的基本平台、必要组件及可选组件。其中，基本平台建议使用西安交通大学提供的正版Matlab 2023a，非正版可能导致无法正常下载必备组件，若低于2023a版本则可能导致编程过程与2023a版本存在细微差异。（Matlab 2022a版本测试过也没有任何问题）

* **基本平台**：Matlab 2023a
* **必备组件**：

（1）Deep Learning Toolbox：实现Matlab深度学习编程的基础；

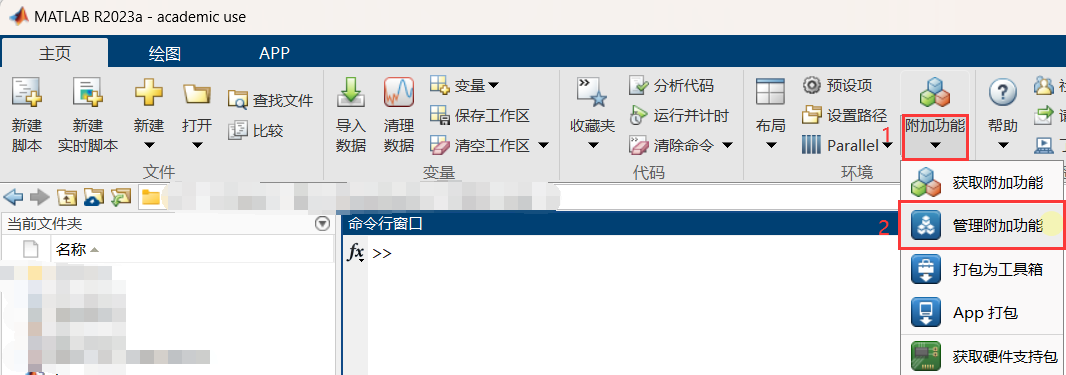
（2）Parallel Computing Toolbox：可实现Matlab程序的多CPU并行计算，大大减少神经网络的计算时间；

* **可选组件**：

（1）CUDA (Compute Unified Device Architecture)：英伟达(Nvidia)推出的运算平台，通过调用显卡GPU算力，极大幅度提高神经网络训练速度。无须在MATLAB平台下载，只需在CUDA官网找到与自己显卡型号相对应的CUDA版本进行下载和安装即可，MATLAB会自动进行适配。

可通过以下方法自查是否已安装上述必备组件，如图1所示，首先点开主页菜单栏，点击“附加功能”，然后点击“管理附加功能”，即可打开附加功能管理器，从“已安装”一栏中检查是否已安装必备组件。

若尚未安装上述必备组件，则点击“获取附加功能”，在出现的界面搜索栏中搜索相应组件名称，找到符合要求的组件下载即可。



**图1 打开MATLAB附加功能管理器，检查是否已安装必备组件**

对于加速运算组件Parallel Computing Toolbox和CUDA，两者只选其一即可。由于CUDA并非本实验必备组件，且安装步骤繁琐，其安装过程可参阅网络资料，本文不再赘述。以下给出三种神经网络计算方式在同一个分类任务上的运算时间，如表1所示：1）不调用Parallel Computing Toolbox (表1中为单CPU)；2）调用Parallel Computing Toolbox (表1中为多CPU并行)；以及3）调用CUDA（表一表示为单GPU）。

**表1 神经网络的三种运算方式的计算时间对比，使用的CPU型号为AMD Ryzen9 5900X(Intel 12th i7-1260P, 2.10 GHz, 精度64%的训练时间14秒，精度98%以上的训练时间29秒 ), 显卡型号为 NVIDIA RTX 3060**

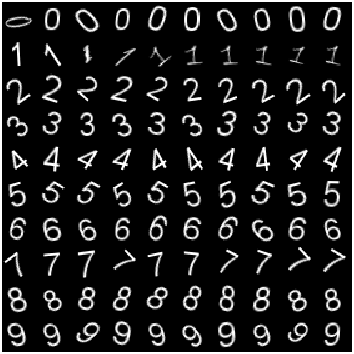
|  |  |  |  |
| --- | --- | --- | --- |
| 计算方式 | 单CPU | 多CPU并行 | 单GPU |
| 运行时间（秒） | 34.16(24) | 18.21(14) | 10.39 |

## 三、实验任务

给定MNIST手写数字数据集，样图如图2所示。原MNIST数据集包含70000张手写数字灰度图像，分别率为，涵盖10个类别，对应从“0”到“9”的十个数字。考虑到实验设备算力有限，本实验给出的MNIST数据集已经过精简化处理，从原始的70000张图像缩减为10000张图像，其中每个类别有1000张图像。

现按照0.75和0.25的比例划分训练集和测试集，将MNIST数据集中的10000张图像按类别均分为7500张图像的训练集和2500张图像的测试集，即每个类别有750张训练图像和250张测试图像。

现搭建一个分类卷积神经网络，其网络架构及相应参数如表2所示，分别包含输入层、卷积层、Relu层、全连接层和Softmax层。其中，输入层维度与输入图像维度一致，即[28 28 1]，这里的第三个参数表示图像通道。卷积层通过卷积运算对输入图像进行特征提取，其卷积核的维度为，通道数为20，步幅为1（函数中不设置此参数时默认为1）。Relu层则对上一卷积层提取的特征进行非线性映射。池化层对Relu层输出的特征进行最大池化操作，池化核维度为，步幅为2。全连接层对前序输出的特征作线性映射，得到维度为10的特征，与MNIST手写数字数据集中的10个类别相对应。然后，将这10维的特征输入到softmax层，得到每个类别的预测分数。将这些预测分数输入到分类层，分数高者即为预测的类别，分类层的作用即输出这些类别。



**图2 MNIST手写数字数据集样图**

**表2 范例代码的神经网络架构及设置**

|  |  |
| --- | --- |
| **模块（按顺序）** | **设置** |
| 输入层 | [28 28 1] |
| 卷积层 | 5@, 步幅: 1 |
| Relu层 | - |
| 最大池化层 | , 步幅: 2 |
| 全连接层 | 输出维度：10 |
| Softmax层 | - |
| 分类层 | - |

**现要求完成以下任务：**

搭建一个卷积神经网络，在给定数据集上进行图像分类任务，并使分类识别率达到99%以上。首先，按指定要求把数据集中的10000张图像按类别均分为7500张图像的训练集和2500张图像的测试集。同学们可参考现有的一切教材，网络结构和样式不限，只须达到指定精度即可。最后，通过绘图或列表格的方式描述你的网络架构及参数，格式可参考表2。

## 四、注意事项

1.实验代码基本框架部分的代码已给出，同学们只需在网络结构模块以及超参数模块进行编程即可。

2.本实验使用的MNIST数据集由的单通道灰度图像组成，在搭建网络时请确保数据输入层的参数和数据集图像尺寸的通道数相对应。

3.训练神经网络前，可通过trainingOptions()调整参数和硬件设备。训练神经网络前请确保‘ExecutionEnvironment’后面参数为’Parallel’或’gpu’，否则MATLAB可能只使用单CPU训练网络，导致训练速度大幅下降。

4.若训练神经网络时不收敛，首先考虑减小超参数中的学习率。

5.可以通过“help 函数”指令调阅该函数的官方文档，从而学习如何使用该函数。例如在命令行窗口键入：

>>>help trainingOptions

即可调阅trainingOptions()函数的官方文档，可从中参考如何设置神经网络超参数；

## 四、实验报告模板（见下页）

|  |  |  |
| --- | --- | --- |
| **姓名：** | **班级：** | **学号：** |

# 实验报告：实验四

基于**卷积**神经网络的图像分类算法

**任务：**搭建一个卷积神经网络，在给定数据集上进行图像分类任务，并使分类识别率达到99%以上。首先，按指定要求把数据集中的10000张图像按类别均分为7500张图像的训练集和2500张图像的测试集。同学们可参考现有的一切教材，网络结构和样式不限，只须达到指定精度即可。最后，通过绘图或列表格的方式描述你的网络架构及参数。

**搭建的卷积神经网络结构：**

|  |  |
| --- | --- |
| **模块（按顺序）** | **设置** |
| 输入层 | [28 28 1] |
| 卷积层 | 512@, 步幅: 1 |
| Relu层 | - |
| 最大池化层 | , 步幅: 2 |
| 卷积层 | 1024@, 步幅: 1 |
| Relu层 | - |
| 最大池化层 | , 步幅: 2 |
| 全连接层 | 输出维度：10 |
| Softmax层 | - |

**学习率：0.0005 Epoch：10 Batch Size：128 准确率：0.9968**

卷积层的参数大虽然训练效果好，但会导致占用的显存高、模型的训练速度慢，可以继续增加卷积层的个数或全连接层的个数。

|  |  |
| --- | --- |
| **模块（按顺序）** | **设置** |
| 输入层 | [28 28 1] |
| 卷积层 | 20@, 步幅: 1 |
| Relu层 | - |
| 最大池化层 | , 步幅: 2 |
| 卷积层 | 40@, 步幅: 1 |
| Relu层 | - |
| 最大池化层 | , 步幅: 2 |
| 全连接层 | 输出维度：540 |
| Relu层 | - |
| 全连接层 | 输出维度：190 |
| Relu层 | - |
| 全连接层 | 输出维度：10 |
| Softmax层 | - |

**学习率：0.001 Epoch：10 Batch Size：128 准确率：0.9948**

|  |  |
| --- | --- |
| **模块（按顺序）** | **设置** |
| 输入层 | [28 28 1] |
| 卷积层 | 35@, 步幅: 1 |
| Relu层 | - |
| 最大池化层 | , 步幅: 2 |
| 卷积层 | 80@, 步幅: 1 |
| Relu层 | - |
| 最大池化层 | , 步幅: 2 |
| 卷积层 | 150@, 步幅: 1 |
| Relu层 | - |
| 最大池化层 | , 步幅: 2 |
| 全连接层 | 输出维度：1000 |
| Relu层 | - |
| 全连接层 | 输出维度：500 |
| Relu层 | - |
| 全连接层 | 输出维度：150 |
| Relu层 | - |
| 全连接层 | 输出维度：10 |
| Softmax层 | - |

**学习率：0.001 Epoch：10 Batch Size：128 准确率：0.9964**