# 《人工智能基础》

# 实验报告

项目名称： 手写体数字识别

项目地址链接：（项目需要设置为公开）

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 成员 | 组长 | 组员1 | 组员2 | 组员3 | 组员4 |
| 学号 | 20301164 | 20301155 | 20301133 |  |  |
| 姓名 | 柳萱莹 | 邹佳琪 | 李昊晴 |  |  |

1. 实验数据集
   1. 数据预处理
2. 说明对数据进行了哪些预处理工作，比如：裁剪、改变通道、改变格式、人脸检测等。说明处理的目的以及处理的结果（没有可以不写）。如果自己加入了新的数据，需要介绍新加入的数据（给出数量、尺寸、特点、如何采集的），以及加入新数据的目的。
3. 介绍最后所用数据集的规模、训练集中的样本数、测试集中的样本数、输入图像的尺寸。

答：（1）

预处理：

1. 读出所有训练数据和测试数据到train\_reader和test\_reader列表，将它们转换为数组形式分为图片数组和标签数组，数据类型转为double型和int型；
2. 通过给训练数据和测试数据的图片数组整体加负号，即将原黑底白字转化为白底黑字（新数据，即数量、尺寸与原数据集相同，特点就是变为白底黑字的了。原训练数据集数量60000，测试集数量10000）。接着将黑底白字图片数组与白底黑字图片数组拼接起来生成新数组，随机以相同状态打乱图片数组和标签数组，使它们打乱后依然对应；

处理的目的：

将数据集转化为numpy数组便于从测试数据和训练数据中分出图片和标签，也便于将黑底白字图转为白底黑字图并把二者做拼接并打乱作为新的训练数据和测试数据。

处理的结果：

得到混合有白底黑字和黑底白字图以及对应标签的四个数组all\_train\_image, all\_test\_image, all\_train\_label, all\_test\_label

（2）最后所用数据集规模：140000

训练集中样本数：120000

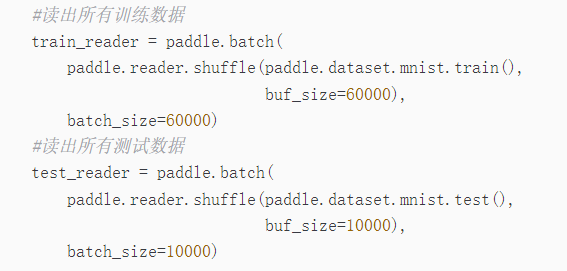
测试集中样本数：20000

输入图像的尺寸：[-1,1]（归一化处理后）

* 1. 数据加载

描述如何将训练数据和测试数据加载进入PaddlePaddle框架中的。说明加载的具体细节，比如：使用了PaddlePaddle里面哪种数据加载器，参数是如何设置的，这样设置的理由是什么。

答：（1）通过paddle框架获取mnist官方数据集的训练集和测试集



读出训练数据部分：

使用paddle.reader.shuffle()修改数据集大小，第一个参数是mnist训练数据集的creator，它返回一个reader creator，reader中的每个样本的图像像素范围是[-1,1]，标签范围是[0,9]；第二个参数是需要数据集的大小，这里设置为整个训练集的大小60000，先读出所有训练数据，为的是方便作后续的转白底黑字图的操作。

使用paddle.batch()将训练数据打包，每个batch的大小为60000，这里是先读出所有训练数据，为的是方便作后续的转白底黑字图的操作。

读出测试数据部分：

与读训练数据同理。

1. 处理成批处理数据应用于后面的模型

定义一个**reader\_creator**(image\_array,label\_array)其中有一个带有yield的reader（）函数，返回一个关于image的数组和此image数组第一个元素图片对应的label生成器，用于后面训练数据的提供器，每次从缓存中随机读取批次大小的数据。





使用paddle.reader.shuffle()修改数据集大小，第二个参数是需要数据集的大小，每次从**reader\_creator**中随机读取512大小的数据集，将指定大小为512的数据集按照batch大小为128打包成批处理数据，每批处理128条数据.这样有利于提高算法的执行效率，以小单位来处理。

1. 网络结构
2. **详细介绍所采用的网络结构，给出网络结构图，说明：**
3. 每个卷积层所采用的滤波器个数、卷积尺寸、滤波器深度、激活函数

VGG一共使用了5个卷积层

第1个卷积层：滤波器64\*2=128个，卷积尺寸：28\*28\*1，滤波器深度：1，激活函数：Relu

第2个卷积层：滤波器128\*2=256个，卷积尺寸：14\*14\*128，滤波器深度：128，激活函数：Relu

第3个卷积层：滤波器256\*3=768个，卷积尺寸：7\*7\*256，滤波器深度：256，激活函数：Relu

第4个卷积层：滤波器512\*3=1536个，卷积尺寸：3\*3\*768，滤波器深度：768，激活函数：Relu

第5个卷积层：滤波器512\*3=1536个，卷积尺寸：1\*1\*1536，滤波器深度：1536，激活函数：Relu

1. 池化层的个数、位置、池化窗口的尺寸及深度

VGG池化层个数：5

位置：在每个卷积层后

池化窗口的尺寸：2\*2

深度：第1个池化层：1；第2个池化层：128；第3个池化层：256；第4个池化层：768；第5个池化层：1536

池化类型：max

1. 全连接层的个数、每个全连接层的滤波器个数、尺寸、深度

VGG全连接层的个数：3个

第1个全连接层的滤波器个数：512，尺寸：1\*1，深度：1536

第2个全连接层的滤波器个数：512，尺寸：1\*1，深度：152

第3个全连接层的滤波器个数：10，尺寸：1\*1，深度：152

4）输出层的神经元个数、激活函数

神经元个数：10

激活函数：softmax

1. **简述选择该网络结构的理由。**

VGG通过堆叠2个3\*3的卷积核来代替1个5\*5的卷积核，

通过堆叠3个3\*3的卷积核来代替1个7\*7的卷积核，

可以减少参数量。

**（3）说明尝试了几种网络模型，分别是什么模型。**

尝试使用了两种网络模型，Lenet和VGG神经网络模型。

1. LeNet

1998年的诞生的LeNet(LeCunetal.)可谓是现在各种卷积神经网络的始祖了，其网络结构虽然只有5层，却包含了卷积神经网络的基本组件(卷积层，池化展，全连接层)。

使用的问题：

（1）训练难。图像的数据由一个个的像素组成，如果将每个像素都输入到一个的单独的神经元中，那神经网络的尺寸，权值参数将是非常

的大的。这么多的参数，需要消耗非常多的计算资源，以及需要非常大的数据集。

（2）无法应对图像内容位置的变化。对干手写体的数字图像，其每个字符的大小，倾斜度，书写的位置都会导致特征位置的变化。如果将

每个像素输入到一个单独的神经元中，则需要大量的不同位置的训练数据输入到神经元中。

（3）忽略了图像的局部相关性。图像的相邻像素有很强的相关性。单独一个像素输入到神经元中则丢失了这种相关性。

1. VGG

如上文2.1和2.2介绍。

1. 损失函数及学习率

（1）给出网络所使用的损失函数公式，说明选择该损失函数的理由，说明反向传播是如何作用在权重上的（可以简写或不写）以及其他相关的内容。

（2）说明尝试了几种损失函数，分别是什么函数。

（3）学习率是多少？简述确定该学习率的过程。

1. 优化器

（1）介绍所采用的优化器的原理、特点，及选择它的原因。

（2）说明尝试了几种优化器，分别是什么优化器。

1. 实验结果
2. 用表格或图的形式，对比各种尝试（使用不同的网络模型、损失函数、优化器）的实验结果（recall，precision，accuracy，error-rate）；
3. 分析、解释实验过程中出现的一些现象。
4. 总结与收获（必写）

谈谈学习这门课的心得体会。每个小组成员分别写自己的心得（不得少于100字），汇总于此。

学生1 柳萱莹：

通过这学期的学习，对人工智能有了更进一步的了解和认识。

个人觉得研究人工智能的目的，一方面是要创造出具有智能的机器，另方面是要弄清人类智能的本质，因此，人工智能既属于工程的范畴，又属于科学的范畴。通过研究和开发人工智能，可以辅助，可以辅助，部分替代甚至拓宽人类的智能，使计算机更好的造福人类。人工智能研究的近期目标是使现有的计算机不仅能做般的数值计算及非数值信息的数据处理，而且能运用知识处理问题，能模拟人类的部分智能行为。按照这一目标，根据先行的计算机的特点研究实现智能的的有关理论、技术和方法，建立相应的智能系统。例如目前研究开发的专家系统，机器翻译系统、机器人等。随着社会的发展，技术的进步，人工智能的发展是任何人都无法想象的。

在这门课后，将会自主的对人工智能进行更多的探索和学习，感谢这门课程为我打开了人工智能的窗户，让我有幻想和愿望去探索更加丰富的人工智能世界。

学生2姓名：心得。。。。

学生3姓名：李昊晴：

学生4姓名：心得。。。。

学生5姓名：心得。。。。