# 基于混合神经网络的MINST识别算法

## 功能介绍

本算法旨在通过用利用神经网络的优势，高效的对手写数字进行识别，即将手写在纸上或设备上的数字的图片信息高正确率地转换为纸上或设备上的数字。

## 整体架构介绍

为了充分发挥各个神经网络的优点，使其既有卷积神经网络高效，速度快，对噪点的容错性高的优点，又有全联接的前导神经网络准确性高的特点，本算法使用将卷积神经网络与前导神经网络抽样后全联接的方法来高效的完成对手写数字图像的识别。

>整体网络图 figure\_all

本神经网络的主体分为输入数据处理，神经网络运算和输出数据处理。

其中输入数据处理（IN\_PRO）是将原始数据处理成神经网络容易识别的且经过标准化的数据集。

神经网络(在运算上分为训练阶段与识别阶段。在训练阶段神经网络通过所给的训练数据与答案修改与更新自己的参数，从而使自己通过输入数据所计算出来的答案越来越接近真实正确的答案。识别阶段即使用通过大量数据训练过的神经网络，让测试数据通过神经网络进行运算从而得出预测的结果。

在这里神经网络部分（NN）分为输入层（IN）卷积层（C1），次抽样层（S1），隐藏层（H1），输出层（OUT）

输出处理则将大量的原始，为优化结果而标准化的数据处理成可以利用，统计的直观数据。在这里输出处理只有一部分（OUT\_PRO）。

## 详细介绍

下面将对本算法的各个层次的原理，设计与实现进行详尽的介绍。

### 输入数据处理部分

>输入处理图 in1

本部分旨在对原始数据的处理，其中原始数据有训练数据（32000张PNG图片，其中每个图片的分辨率为28\*28，图片为灰度图片）；训练数据的标签，即这32000张图片所显示的具体数字的答案；测试数据（10000张图片，图片格式与训练数据相同）以及测试数据的标签。

对训练与测试数据的处理：

训练数据原始为图片，先用python的Image库将整个图片加载成矩阵，由于输入的图片均为灰阶图片，所以根据Image库的算法，输出的矩阵为输入图片对应像素的灰阶值。

然后对这个矩阵进行标准化，标准化的意思是，通过线性或者非线性转换，将矩阵的最大值-最小值范围控制在0-1之间且让最大值和最小值尽量分别接近0与1。对于本文针对的灰阶矩阵，其最大值为255，最小值为0，所以我们可以通过把每个数字除以255的方式将其标准化。

用此方法，我们成功将输入矩阵进行了标准化。对于标准化的作用，本文将在后面介绍线性神经元与其梯度下降算法的部分详细介绍。

对训练与测试标签数据的处理：

训练数据与测试数据通过神经网络的计算会得到其预测的结果的特征，在这里这个特征的值就决定了本神经网络对输入图像的识别结果。区别于数字本身的连续性，对于神经网络的分类任务来说，其特征应该是互不相干的，例如：1与9之间的相关性应和1与2之间的相关性完全一样。所以需要用编码的方式排除数字本身的相关性，本算法使用的编码方式是独热编码（One-Hot Encoding）

独热编码：

独热编码直观来说就是用总状态数的比特数，其中一个比特为0其他比特为1的方法来表示状态的一种码制。

在本文的算法中，总状态数即为10。下面是编码表：

|  |  |  |
| --- | --- | --- |
| 原始数据 | 二进制 | 独热编码 |
| 0 | 0000 | 1000000000 |
| 1 | 0001 | 0100000000 |
| 2 | 0010 | 0010000000 |
| 3 | 0011 | 0001000000 |
| 4 | 0100 | 0000100000 |
| 5 | 0101 | 0000010000 |
| 6 | 0110 | 0000001000 |
| 7 | 0111 | 0000000100 |
| 8 | 1000 | 0000000010 |
| 9 | 1001 | 0000000001 |

表格 1-a

经过编码后的标签由于其向量积均为0，所以互不相关。且在总状态数较少的情况下独热编码更简单高效，另一方面又增加了状态的空间，使算法更加准确。

### 卷积层（C1）

卷积神经网络(Convolutional Neural Network, CNN)：

卷积神经网络是深度学习与计算机视觉中极其重要的网络结构之一，许多及其成功的模型都是基于CNN的，特别是基于ImageNet等的各种图像物体识别算法。卷积神经网络算法相对于传统的图像处理算法有很多优点，其中一个就是其巧妙的避免了复杂耗时的预处理工作（如人工特征的提取），原始图像数据经过CNN可以直接进入下一步的运算，不需要人工的参与。

如前文所说，在图像识别的过程中，图像往往会被看成一个或多个的二维矩阵，本文用到的MINST图像库中，图像库中每一个图片就可以看成是28 \* 28 的二维矩阵，因为它是黑白图片，只有一个颜色通道。如果是彩色图片就会有RGB三个颜色通道，就会被表示为三个二维矩阵。传统的神经网络算法使用的是全链接的方式，这种连接方式是将每一个颜色通道的每一个像素都与下一层的神经元·所连接，这样会使参数的数量指数增长，导致参数的数量巨大，使整个神经网络的训练及其耗时甚至时间数量级巨大，无法训练。而卷积神经网络通过局部链接，权值共享两个主要方法，在几乎不消耗准确率的情况下大幅减少变量数，从而避免这一困难的发生。

（图片）

直观的看，卷积层的参数可以看作是一系列的可以训练和学习的过滤器（卷积核）。在前向学习的过程中，每次我们将图片的每一个小部分通过一个过滤器，点乘后再组成新的二维数据，每个过滤器计算成的二维数据再组成三维数据。我可以理解成每个过滤器只关心一部分图片的特征，再综合图片的各个特征进行下一步的训练。

局部连接：

局部连接是卷积神经网络独特的原因之一，上文提到在图片中，使用全连接的神经网络在工程上是不可行的，而局部链接的意思就是卷积神经网络的神经元只会与上层的神经元的局部区域进行连接。可以把上层数据想象成一张画，画的上面有一个滑动的窗口，每次运算时只有窗口内的数据会与下层进行关联。这就引出了卷积神经网络的一个重要的变量：窗口大小，一般窗口的长与宽是相等的，窗口的长\*宽叫做感受野（receptive field），在整个运算中，这个窗口会逐格滑动，一步步近似覆盖图片的每一个小区域。

在本算法中，我们的MINST图片库输入的图片数据大小为[1\*28\*28]，我们设定感受野（窗口）大小为[5\*5]，为了提取不同的特征，我们有6个不同的卷积核。

权值共享：

权值共享是卷积神经网络中很巧妙的处理方方式，