**一. 实验介绍**

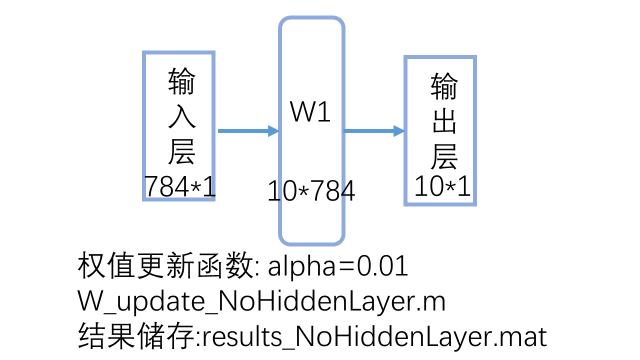
实验中总共尝试了三种结构的网络模型：无隐层神经网络，简单的带隐层的神经网络和带一层卷积池化的简单CNN，在普通神经网络中，更改隐层层数，在CNN中，又尝试了加入dropout算法和更改判别网络的隐藏层层数，看最后的结果有何不同。

**二. 训练方法**

在权值更新函数中统一采用SGD算法，使用60000个训练数据更新一轮权值。

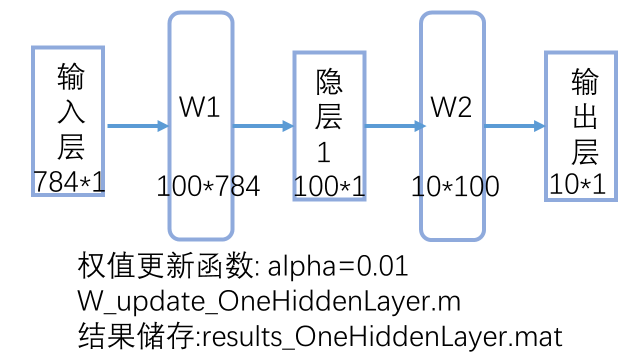
**三. 网络设计**

**1.无隐层网络**

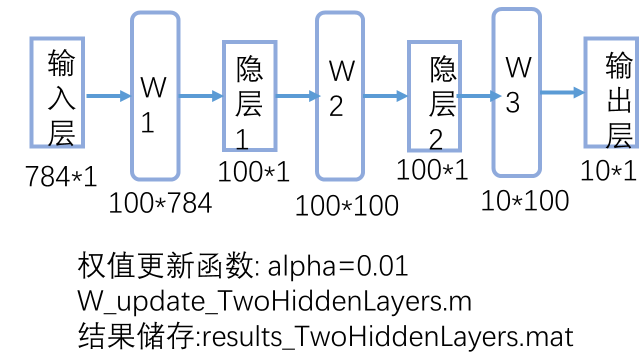


**2.带隐层简单神经网络**

**2.1 单隐层**

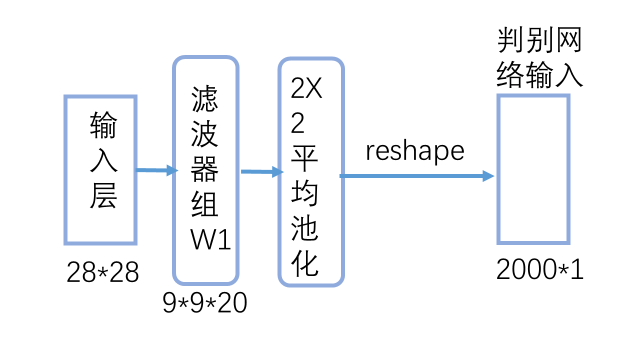


**2.2 双隐层**

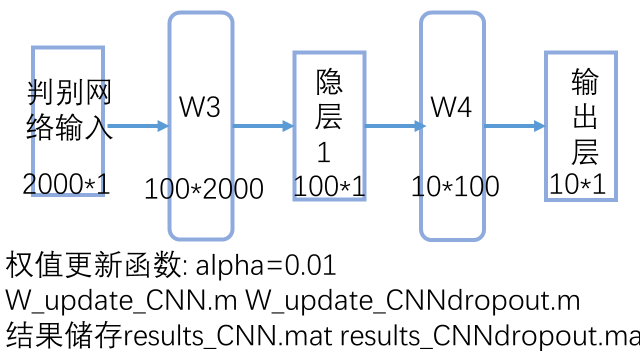


**3. CNN**

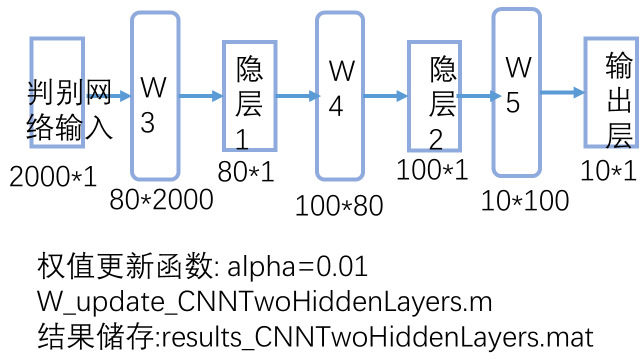
特征提取网络都一致：



**3.1 判别网络单隐层,区分使用dropout和不使用dropout方法**



**3.2 判别网络双隐层,未使用dropout**



**四. 实验结果**

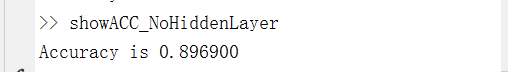
三类网络的实验结果如下：

**1. 无隐层网络**

运行时间：



准确率：



**2.普通神经网络**

**2.1 单隐层**

运行时间：



准确率：



**2.2 双隐层**

运行时间：



准确率：



**3.CNN**

**3.1 判别网络单隐层**

运行时间：



准确率：



**3.2 判别网络单隐层+dropout**

运行时间：



准确率：



**3.3 判别网络双隐层**

运行时间：



准确率：



**五. 结果分析比较**

令人惊奇的是，使用普通的神经网络就能获得比较高的正确率（最高92%左右）。

就简单的无隐层，单双隐层网络而言。单隐层网络的正确率（92%）高于无隐层（89.6%）的网络，这是复合预期的。但到了双隐层，正确率（90%）就有一定的下降，判断是因为参数设置的不合理，随着层数的增加，出现了过拟合。不过，在实验中加入dropout算法后，结果也没有变好。应该是隐层节点数设置的不合理。

在网络中加入卷积和池化层组成的特征提取网络之后，正确率最高可以达到（97.63%）。在单隐层中应用dropout算法之后，正确率（97.3%）有一点下降。考虑到下降的幅度并不大，以及初始权值的变化会导致结果出现波动的原因，推测dropout方法并未明显的影响最终的正确率。应该是因为只有一个隐层，dropout效果不明显。

将判别网络变为双隐层后，正确率的结果（97.55%）同在简单神经网络中一样，也没有明显的变化。推测还是由于隐层的节点数设置的不合理。

**六. 实验结论**

最终，在经过各种尝试后，还是按照老师给出的网络结构得出的结果最好，达到97.63%，训练一轮耗时将近7分钟。尝试将数据多训练几轮，正确率有一定提高，如下表：

|  |  |  |
| --- | --- | --- |
| **训练轮数** | **正确率** | **耗时** |
| **2** | **0.982400** | **685.032 s（11min）** |
| **3** | **0.984000** | **1021.610 s（17min）** |
|  | **0.986200** | **1403.957 s（23min）** |

**七. 总结与反思**

1.dropout算法只加在层数不多的判别网络中似乎没有什么效果，在网上查了一些资料后发现可以在卷积池化层中也加入dropout算法。但也有人说这样做没有必要，效果不是很显著。

2.可以考虑加深特征提取网络的层数，没准会有更好的效果。