

# 神经网络是什么？如何构成的？

---

同学们，大家好！

今天给大家带来的小科普是：神经网络是什么，及其是如何构成的。

想必大家已经了解过一些深度学习的知识，深度学习是基于神经网络的高级算法，那么神经网络是什么呢？

## 神经网络的定义

猜一猜，下图中的动物是什么？



尽管图中的动物胖得出奇，你也一眼就能猜到它是一只长颈鹿。人类的大脑拥有强大的辨识能力，它是一个由大约 800 亿个神经元组成的复杂网络。即使某物体并非我们熟知的模样，我们也能够轻松地识别。

大脑神经元彼此协同工作，它们把输入信号（比如长颈鹿的图片）转换成相应的输出标签（比如“长颈鹿”）。神经网络技术的诞生正是受到人脑神经网络的启发。

神经网络是自动图像识别的基础，近年来，神经网络技术大热，主要有以下 3 个原因：

- 数据存储和共享技术取得进步。这为训练神经网络提供海量数据，有助于改善网络的性能。
- 计算能力越来越强大。GPU ( graphics processing unit, 图形处理器 ) 的运行速度最快能达到 CPU ( central processing unit, 中央处理器 ) 的 150 倍。之前，GPU 主要用来在游戏中显示高品质图像。后来，人们发现它能为在大数据集上训练神经网络提供强大的支持。
- 算法结构的改进。虽然目前神经网络在性能上还很难与人脑媲美，但是已经有一些能大幅改善其性能的技术。本文会介绍其中一些技术。

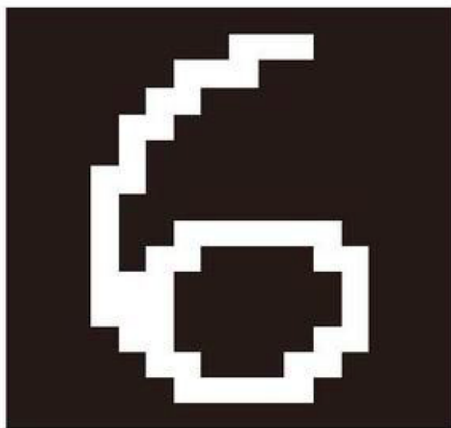
图像识别是神经网络技术的有力例证，它被应用于许多领域，包括视觉监控、汽车自主导航，还出现在智能手机中，来识别手写文字。下面来看看如何训练能识别手写数字的神经网络。

## 神经网络示例：识别手写数字

下面示例中使用的手写数字来自于 MNIST 数据库，如下图所示。



为了让计算机读取图像，黑色像素用 0 表示，白色像素用 1 表示，如下图所示。如果图像是彩色的，则可以使用三原色的色相值来表示。



```

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 1 1 1 0 0 0 0
0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 0
0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0
0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0
0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0
0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0
0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 1 0 0 1 1 1 1 1 1 0 0 0 0
0 0 0 1 0 1 1 0 0 0 0 1 1 0 0 0
0 0 0 1 1 1 0 0 0 0 0 0 1 0 0 0
0 0 0 1 1 1 0 0 0 0 0 0 1 0 0 0
0 0 0 0 1 1 0 0 0 0 0 0 1 1 0 0
0 0 0 0 1 1 0 0 0 0 0 1 1 0 0 0
0 0 0 0 0 1 1 0 0 0 1 1 0 0 0 0
0 0 0 0 0 0 1 1 1 1 1 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

```

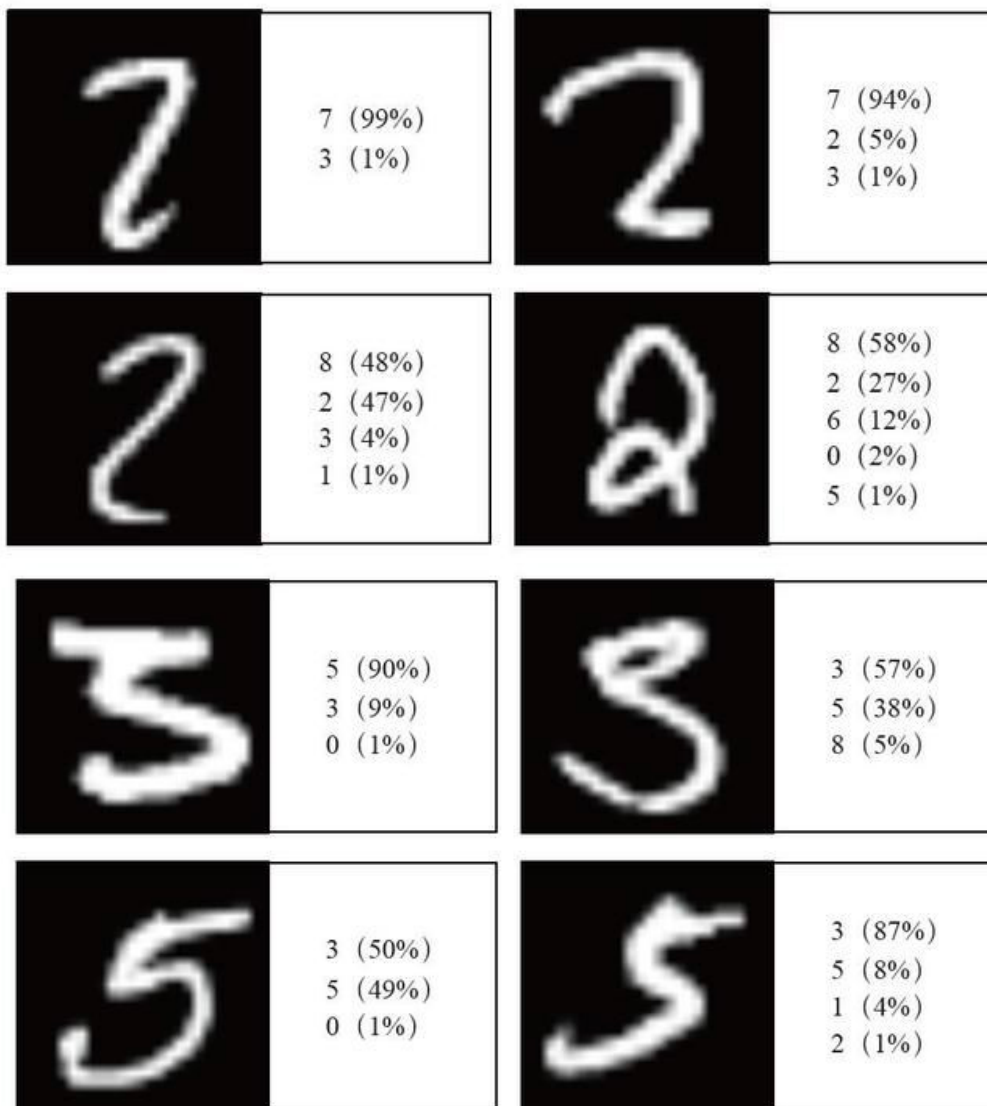
一旦图像完成像素化，就可以把得到的值交给神经网络。在本例中，神经网络总共得到 10000 个手写数字以及它们实际所表示的数字。在神经网络学过手写数字及其对应标签的联系之后，我们拿 1000 个新的手写数字（不带标签）来测试它，看看它是否能够全部识别出来。测试发现，神经网络从 1000 个新的手写数字中正确识别出了 922 个，即正确率达到了 92.2%。下图是一张列联表，可以用它来检查神经网络的识别情况。

		预测的数字										总计	%
		0	1	2	3	4	5	6	7	8	9		
实际的数字	0	84	0	0	0	0	0	1	0	0	0	85	99
	1	0	125	0	0	0	0	1	0	0	0	126	99
	2	1	0	105	0	0	0	0	4	5	1	116	91
	3	0	0	3	96	0	6	0	1	0	1	107	90
	4	0	0	2	0	99	0	2	0	2	5	110	90
	5	2	0	0	5	0	77	1	0	1	1	87	89
	6	3	0	1	0	1	2	80	0	0	0	87	92
	7	0	3	3	0	1	0	0	90	0	2	99	91
	8	1	0	1	3	1	0	0	2	81	0	89	91
	9	0	0	0	0	1	0	0	6	2	85	94	90
总计		91	128	115	104	103	85	85	103	91	95	1000	92

上图列联表总结了神经网络的表现：第一行指出，共有 85 个“0”，神经网络正确识别出 84 个，最后一个“0”被错误地识别为“6”。最后一列是识别准确率。

从上图可以看到，“0”和“1”的手写图像几乎全部被正确识别出来了，而“5”的手写图像最难识别。接下来详细看看那些被识别错的数字。

“2”被错误识别成“7”或“8”的情况大约占8%。虽然人能够轻松识别出图中这些数字，神经网络却可能被某些特征给难住，比如“2”的小尾巴。有趣的是，神经网络对“3”和“5”也比较困惑，识别错误的情况约占10%。

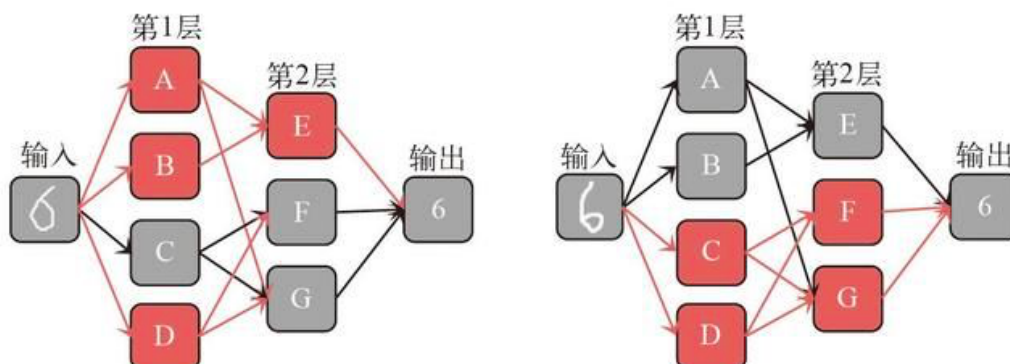


尽管出现了这些错误，但是神经网络的识别速度远快于人类，并且从总体上看，神经网络的识别准确率很高。

通过上面手写数字识别案例，大家对神经网络有了更深认识，那么神经网络究竟如何构成的呢？

### 神经网络的构成

为了识别手写数字，神经网络使用多层神经元来处理输入图像，以便进行预测。下图为双层神经网络示意图。



输入不同，但是输出相同，其中红色表示被激活的神经元。

在上图双层神经网络中，虽然输入是“6”的两幅不同形态的图像，但输出是一样的，并且该神经网络使用不同的神经元激活路径。尽管每一个神经元组合产生的预测是唯一的，但是每一个预测结果都可以由多个神经元组合实现。

神经网络通常由如下几部分组成：

- 输入层

该层接收输入图像的每个像素。

- 隐藏层

在像素进入神经网络之后，它们经过层层转换，不断提高和那些标签已知的图像的整体相似度。标签已知是指神经网络以前见过这些图像。虽然转换得越多，预测准确度就会越高，但是处理时间会明显增加。一般来说，几个隐藏层就足够了。

- 输出层

该层产生最终预测结果。

- 损失层

虽然上图并未显示损失层，但是在神经网络的训练过程中，损失层是存在的。该层通常位于最后，并提供有关输入是否识别正确的反馈；如果不正确，则给出误差量。

在训练神经网络的过程中，损失层至关重要。若预测正确，来自于损失层的反馈会强化产生该预测结果的激活路径；若预测错误，则错误会沿着路径逆向返回，这条路径上的神经元的激活条件就会被重新调整，以减少错误。这个过程称为反向传播。

通过不断重复这个训练过程，神经网络会学习输入信号和正确输出标签之间的联系，并且把这些联系作为激活规则编入每个神经元。因此，为了提高神经网络的预测准确度，需要调整管理激活规则的部件。