

作者先后供职于奇虎360人工智能研究院与陌陌深度学习实验室  
凝聚作者六年多的深度学习研究与开发心得，业内四位大咖鼎力推荐

言有三  
◎著

# 深度学习之 图像识别

## 核心技术与案例实战

Image Recognition by Deep Learning: Core Technologies and Practices

涵盖深度学习核心理论、数据与模型可视化、模型压缩与优化等内容  
详解图像分类、图像分割和目标检测的三大核心技术和八大经典案例

鲁华祥

中科院半导体研究所研究员

陈强

360人工智能研究院技术副总监

邱学侃

作业盒子AI算法科学家

杨世才

浙江大学计算机系博士后

共同  
推荐



机械工业出版社  
China Machine Press

# 深度学习之 图像识别

## 核心技术与案例实战

Image Recognition by Deep Learning: Core Technologies and Practices

言有三◎著



机械工业出版社  
China Machine Press

## 图书在版编目 ( CIP ) 数据

深度学习之图像识别：核心技术与案例实战 / 言有三著. —北京：机械工业出版社，2019.4

ISBN 978-7-111-62472-1

I. 深… II. 言… III. 人工智能—算法—应用—图象识别 IV. TP391.413

中国版本图书馆CIP数据核字 ( 2019 ) 第065997号

本书全面介绍了深度学习在图像处理领域中的核心技术与应用。书中不但重视基础理论的讲解，而且从第4章开始的每章都提供了一到两个不同难度的案例供读者实践，读者可以在已有代码的基础上进行修改和改进，从而加深对所学知识的理解。

本书共10章，首先从深度学习的基础概念开始，介绍了神经网络的基础知识和深度学习中的优化技术；然后系统地介绍了深度学习中与数据相关的知识，包括经典数据集的设计、数据集的增强以及数据的获取与整理；接着重点针对图像开发领域，用3章内容系统地介绍了深度学习在图像分类、图像分割和目标检测3个领域的核心技术与应用，这些内容的讲解均结合实战案例展开；另外，还对深度学习中损失函数的发展、数据和模型的可视化以及模型的压缩和优化进行了详细介绍，为读者设计和训练更加实用的模型提供了指导；最后以微信小程序平台为依托，介绍了微信小程序前后端开发技术，完成了深度学习的模型部署，让本书的内容形成了一个完整的闭环。

本书理论与实践结合，深度与广度兼具，特别适合深度学习领域的相关技术人员与爱好者阅读，尤其适合基于深度学习的图像从业人员阅读，以全方位了解深度学习在图像领域中的技术全貌。另外，本书还适合作为相关培训机构的深度学习教材使用。

## 深度学习之图像识别：核心技术与案例实战

出版发行：机械工业出版社（北京市西城区百万庄大街22号 邮政编码：100037）

责任编辑：欧振旭 李华君

责任校对：姚志娟

印刷：中国电影出版社印刷厂

版次：2019年5月第1版第1次印刷

开本：186mm×240mm 1/16

印张：17.5

书号：ISBN 978-7-111-62472-1

定价：79.00元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

客服热线：(010) 88379426 88361066

投稿热线：(010) 88379604

购书热线：(010) 68326294

读者信箱：hzit@hzbook.com

版权所有·侵权必究

封底无防伪标均为盗版

本书法律顾问：北京大成律师事务所 韩光/邹晓东

如何使用好数据？

深度学习的核心优化技术有哪些？

如何做好图像分类任务？

如何做好图像分割任务？

如何做好目标检测任务？

如何对数据和模型进行可视化分析？

如何压缩和优化一个工业级深度学习模型？

如何理解分类和回归等任务的损失函数？

如何在微信小程序上部署模型？

.....

通过阅读本书，你将了解这些复杂问题背后的原理，甚至你都可以自己解决这些问题。

## 本书核心知识

---

### 理论基础

- 神经网络与卷积神经网络基础
- 卷积神经网络核心理论技术
- 开源框架简介

### 图像识别核心知识

- 数据集的发展与使用
- 数据和模型可视化
- 模型压缩与优化
- 损失函数
- 图像分类
- 图像分割
- 目标检测

### 八大经典案例

- 人脸表情分类
- 鸟类细粒度分类
- 人脸属性分割
- 人像分割
- 猫脸检测
- TensorFlow可视化应用
- MobileNet语义分割模型压缩
- 微信小程序部署



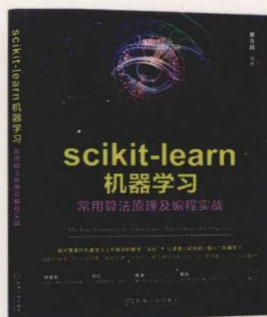
## 作者简介



**言有三** 真名龙鹏。2012年本科毕业于华中科技大学，后保研至中国科学院并于2015年毕业。先后在奇虎360人工智能研究院和陌陌深度学习实验室从事与计算机视觉相关的工作，积累了丰富的传统图像处理算法和深度学习项目

实战经验。运营微信公众号《有三AI》，内容覆盖深度学习的理论、实战经验、开源框架、模型架构，以及深度学习在各应用领域的技术分析，还提供国内外AI研究院的核心技术报导及AI工程师成长路线的完整规划。在知乎上开设专栏《有三AI学院》和其他子方向专栏。在GitChat和网易云课堂上开设若干图文和视频课程。

## 推荐阅读



欢迎IT领域的各位技术高手洽谈出书事宜。如果您有写书或投稿意向，请加QQ或微信与编辑具体商谈。

QQ: 627173439

微信: oyzx\_xp

试读结束 需要全本请在线购买:

[www.ertongbook.com](http://www.ertongbook.com)

# 前言

机器学习、深度学习、人工智能，这些关键词在最近几年“声名鹊起”。以深度学习为代表的无监督机器学习技术在图像处理、语音识别和自然语言处理等领域里频频取得新的突破。但深度学习其实并不是一门全新的学科，其历史可以追溯到 20 世纪 40 年代。

深度学习背后的核心技术包括神经网络的结构设计和最优化方法等，其理论体系虽然有一定进展但是尚不完备。可以说，当前的主流深度学习技术是一门应用性极强的工程技术，这种尚不完备的理论加上具有较高门槛的应用工程特点，对于初学者来说具有一定的困难。如何系统性地了解理论知识，又能够紧随理论进行全面的实践，成为一名合格的图像处理领域的深度学习算法工程师，这是本书所要解决的问题。

笔者有超过 6 年的图像行业背景，最近几年也多以深度学习技术为基础进行相关项目的开发，在多年的知识积累和项目实践中，总结出了大量的经验，浓缩成了这本书。本书从深度学习的背景和基础理论开始讲起，然后介绍了深度学习中的数据及图像处理中的几大重要方向，并介绍了神经网络的可视化、优化目标、模型的优化和模型的线上部署。

本书内容由浅入深，讲解图文并茂，紧随工业界和学术界的最新发展，理论和实践紧密结合，给出了大量的图表与案例分析。本书抛开了过多的数学理论，完整地剖析了深度学习在图像处理领域中各个维度的重要技术，而不是只停留于理论的阐述和简单的结果展示，更是从夯实理论到完成实战一气呵成。相信读者跟随着本书进行学习，将会对深度学习领域的图像处理技术和其在实际开发中的应用有更深入的理解。

## 本书特色

### 1. 内容全面，紧跟最新技术发展

本书内容涵盖了深度学习的理论知识、数据获取与增强，以及深度学习在图像分类、分割和检测这三大基础研究领域中的发展、数据与模型的可视化、优化目标、模型压缩与部署等相关知识，基本上囊括了深度学习在图像开发中所必须要掌握的大部分基础知识。

### 2. 深度与广度兼具

本书在讲解每个知识点时力求详尽，而且紧密结合了学术界与工业界相关技术的最新

发展。这样的安排既注重知识的广度，也兼具知识的深度，可以为图像处理领域中的从业者提供系统性的学习指导。

### 3. 理论与实践案例紧密结合

本书不仅对理论知识进行了阐述，而且还给出了大量的实践内容，以帮助读者提高实际的动手能力。除了第 1、2 章主要介绍了深度学习的基础理论外，后续章节则大多采用了先系统介绍该章涉及知识的发展现状，然后有针对性地设计了一到两个实践案例带领读者学习，有较好的学习效果。

### 4. 参考了不同层次学习者的意见

本书若干内容的简化版本已在笔者运营的公众号平台上接受了不同层次读者的反馈，力求知识的完备性和准确性；另外，本书有多位编写者参与，他们或理论见长，或善于动手，让本书从不同层面得到了广泛的意见，可以满足不同人群的学习需求。

## 本书内容

第 1 章神经网络基础，首先介绍了神经网络的生物基础与数学模型，然后介绍了卷积神经网络的基础知识，这也是当前深度学习模型的基础。

第 2 章深度学习优化基础，首先介绍了深度学习主流开源框架，如 Caffe、TensorFlow、Pytorch 和 Theano 等，并对其特点与性能做了对比；然后介绍了网络优化参数，包括激活函数、正则化方法和归一化方法等。本章旨在让大家对深度卷积神经网络有一个较为全面的认识，给后续章节的学习打好基础。

第 3 章深度学习中的数据，首先介绍了深度学习发展过程中的几个数据集，给读者展示了数据集对深度学习的重要性；然后介绍了几大重要发展方向中的数据集；接着讲述了数据增强的方法；最后讲述了数据的收集、整理及标注等相关问题。

第 4 章图像分类，首先介绍了图像分类的基础、基于深度学习的图像分类研究方法及发展现状，以及图像分类任务中的挑战；然后以一个移动端的基准模型为例，展示了图像分类任务的实践流程；最后介绍了一个细粒度级别的图像分类任务，以一个较高的基准模型，展示了较难的图像分类任务训练参数的调试技巧。

第 5 章图像分割，首先介绍了从阈值法到活动轮廓模型的传统图像分割方法；然后介绍了基于深度学习的图像分割方法的基本原理与核心技术；接着讲述了一个移动端的实时图像分割任务，该任务以 MobileNet 为基准模型，展示了图像硬分割任务实践的完整流程；最后介绍了一个更加复杂的肖像换背景的任务，展示了图像软分割任务的基本流程和应用场景。

第 6 章目标检测，首先介绍了目标检测的基础和基本流程，并讲述了一个经典的 V-J 目标检测框架；然后介绍了基于深度学习的目标检测任务的研究方法与发展现状，并分析了目标检测中的核心技术；最后给出了一个目标检测任务实例，通过分析 faster rcnn 的源代码，使用该框架自带的 VGG CNN 1024 网络完成训练和测试，并总结目标检测中的难点。

第 7 章数据与模型可视化，首先对包括低维与高维数据的可视化做了简单介绍；然后对深度学习中的模型可视化做了详细介绍，包括模型的结构和权重可视化；最后介绍了一个基于 Tensorflow 和 Tensorboard 的完整案例。

第 8 章模型压缩，首先详细介绍了模型压缩的方法，然后以一个典型的模型压缩实战案例来阐述项目中的模型压缩上线。

第 9 章损失函数，首先介绍了分类任务的损失函数；然后介绍了回归任务的损失函数；最后介绍了这些损失函数在几大经典图像任务中的使用。

第 10 章模型部署与上线，依托微信小程序平台从 3 个方面介绍了模型部署的问题。首先介绍了微信小程序的前端开发，然后介绍了微信小程序的服务端开发，最后介绍了 Caffe 的环境配置。

## 本书配套资源获取方式

本书涉及的源代码文件及其他资料需要读者自行下载。请登录华章公司网站 [www.hzbook.com](http://www.hzbook.com)，在该网站上搜索到本书，然后单击“资料下载”按钮即可在页面上找到“配书资源”下载链接。

## 本书读者对象

- 图像处理技术人员；
- 深度学习技术人员；
- 深度学习技术爱好者；
- 深度学习技术研究人员；
- 相关院校的学生和老师；
- 相关培训机构的学生和老师。

## 本书作者

本书作者龙鹏，笔名言有三，运营微信公众号《有三 AI》。本书的第 1 章和第 2 章

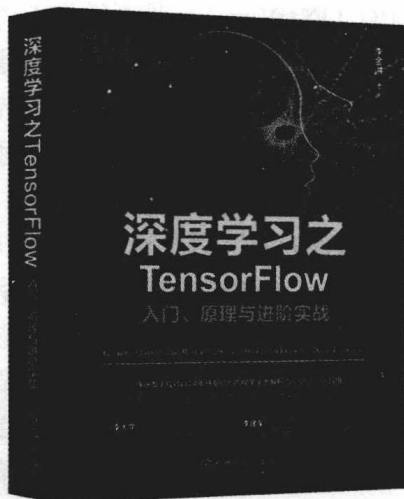
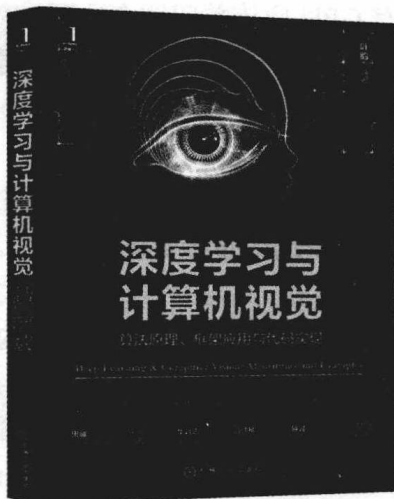


的深度学习理论知识，由华中科技大学的在读博士生徐冰璐主笔，展示出了作者扎实的深度学习理论知识功底和细致的写作水平；鲍琦琦、陶玉龙及杨皓博参与了第 3、6、10 章部分内容的写作与校对工作；另外，在本书成书前，从中国科学技术大学毕业 3 年的硕士研究生胡郡郡参与了全书的阅读和反馈。再次对他们表示感谢！

由于作者水平所限，加之写作时间较为仓促，书中可能还存在一些疏漏和不足之处，敬请各位读者批评指正。联系我们请发电子邮件到 [hzbook2017@163.com](mailto:hzbook2017@163.com)。

作者

## 推荐阅读



### 深度学习与计算机视觉：算法原理、框架应用与代码实现

作者：叶韵 书号：978-7-111-57367-8 定价：79.00元

**全面、深入剖析深度学习和计算机视觉算法，西门子高级研究员田疆博士作序力荐！**

**Google软件工程师吕佳楠、英伟达高级工程师华远志、理光软件研究院研究员钟诚博士力荐！**

本书全面介绍了深度学习及计算机视觉中的基础知识，并结合常见的应用场景和大量实例带领读者进入丰富多彩的计算机视觉领域。作为一本“原理+实践”教程，本书在讲解原理的基础上，通过有趣的实例带领读者一步步亲自动手，不断提高动手能力，而不是枯燥和深奥原理的堆砌。

本书适合对人工智能、机器学习、深度学习和计算机视觉感兴趣的读者阅读。阅读本书要求读者具备一定的数学基础和基本的编程能力，并需要读者了解Linux的基本使用。

### 深度学习之TensorFlow：入门、原理与进阶实战

作者：李金洪 书号：978-7-111-59005-7 定价：99.00元

**磁云科技创始人/京东终身荣誉技术顾问李大学、创客总部/创客共赢基金合伙人李建军共同推荐**

**一线研发工程师以14年开发经验的视角全面解析TensorFlow应用**

**涵盖数值、语音、语义、图像等多个领域的96个深度学习应用实战案例！**

本书采用“理论+实践”的形式编写，通过大量的实例（共96个），全面而深入地讲解了深度学习神经网络原理和TensorFlow使用方法两方面的内容。书中的实例具有很强的实用性，如对图片分类、制作一个简单的聊天机器人、进行图像识别等。书中每章都配有一段教学视频，视频和图书的重点内容对应，能帮助读者快速地掌握该章的重点内容。本书还免费提供了所有实例的源代码及数据样本，这不仅方便了读者学习，而且也能为读者以后的工作提供便利。

本书特别适合TensorFlow深度学习的初学者和进阶读者作为自学教程阅读。另外，本书也适合作为相关培训学校的教材，以及各大院校相关专业的教学参考书。

# 目录

前言	
第 1 章 神经网络基础	1
1.1 神经网络的生物基础与数学模型	1
1.1.1 神经元	1
1.1.2 感知机	2
1.1.3 BP 算法	6
1.2 卷积神经网络基础	10
1.2.1 卷积神经网络的基本结构	11
1.2.2 卷积与权值共享	13
1.2.3 感受野与池化	14
第 2 章 深度学习优化基础	17
2.1 深度学习主流开源框架	17
2.1.1 Caffe 简介	18
2.1.2 TensorFlow 简介	18
2.1.3 PyTorch 简介	19
2.1.4 Theano 简介	20
2.1.5 Keras 简介	20
2.1.6 MXNet 简介	21
2.1.7 Chainer 简介	21
2.2 网络优化参数	22
2.2.1 常用激活函数	22
2.2.2 参数初始化方法	26
2.2.3 最优化方法	27
2.2.4 归一化方法	31
2.2.5 正则化方法	33
第 3 章 深度学习中的数据	36
3.1 深度学习通用数据集的发展	36
3.1.1 MNIST 数据集	36
3.1.2 CIFAR10 和 CIFAR100 数据集	37

3.1.3	PASCAL 数据集	38
3.1.4	ImageNet 数据集	38
3.1.5	Microsoft COCO 数据集	39
3.2	常见的计算机视觉任务数据集	40
3.2.1	人脸数据集	40
3.2.2	自动驾驶数据集	52
3.2.3	医学数据集	55
3.3	数据增强	59
3.3.1	有监督数据增强	60
3.3.2	无监督数据增强	63
3.4	数据的收集与标注	65
3.4.1	数据收集	65
3.4.2	数据标注	67
3.4.3	数据清洗与整理	68
<b>第 4 章</b>	<b>图像分类</b>	<b>70</b>
4.1	图像分类基础	70
4.1.1	图像分类问题	70
4.1.2	深度学习图像分类发展简史	72
4.1.3	评测指标与优化目标	75
4.1.4	图像分类的挑战	76
4.2	移动端实时表情分类实战	77
4.2.1	项目背景	78
4.2.2	数据预处理	80
4.2.3	项目方案	82
4.2.4	模型训练与测试	84
4.2.5	项目总结	88
4.3	细粒度图像分类实战	89
4.3.1	项目背景	89
4.3.2	项目方案	90
4.3.3	模型训练与测试	92
4.3.4	参数调试	97
4.3.5	项目总结	102
<b>第 5 章</b>	<b>图像分割</b>	<b>103</b>
5.1	传统图像分割方法	103
5.1.1	阈值法	103
5.1.2	区域生长法与超像素	105
5.1.3	图切割	105



5.1.4	活动轮廓模型	106
5.2	深度学习图像分割	109
5.2.1	基本流程	110
5.2.2	反卷积	110
5.2.3	多尺度与感受野	112
5.2.4	CRF 方法	113
5.2.5	Image Matting 与图像融合	114
5.3	移动端实时图像分割项目	115
5.3.1	项目背景	115
5.3.2	项目方案	116
5.3.3	模型训练与总结	126
5.4	一个实时肖像换背景项目	127
5.4.1	项目背景	127
5.4.2	项目方案	128
5.4.3	模型训练与测试	134
5.4.4	项目总结	138
<b>第 6 章</b>	<b>目标检测</b>	<b>139</b>
6.1	目标检测基础	139
6.1.1	检测窗口选择	140
6.1.2	特征提取	141
6.1.3	分类器	142
6.1.4	V-J 人脸检测算法	143
6.2	深度学习目标检测方法	145
6.2.1	Selective search 与 R-CNN	146
6.2.2	RoI Pooling 与 SPPNet	147
6.2.3	Fast R-CNN 与 Faster R-CNN	149
6.2.4	YOLO 方法	152
6.2.5	SSD 方法	154
6.2.6	目标检测中的关键技术	155
6.3	实战 Faster-R-CNN 目标检测	157
6.3.1	项目背景	157
6.3.2	py-faster-rcnn 框架解读	157
6.3.3	模型定义与分析	170
6.3.4	模型训练与测试	180
6.3.5	项目总结	183
<b>第 7 章</b>	<b>数据与模型可视化</b>	<b>185</b>
7.1	数据可视化	185
7.1.1	低维数据可视化	185

7.1.2	高维数据可视化	187
7.2	模型可视化	190
7.2.1	模型结构可视化	190
7.2.2	模型权重可视化	198
7.2.3	特征图可视化	201
7.3	可视化案例	202
7.3.1	项目背景	202
7.3.2	数据接口定义	204
7.3.3	网络结构定义	205
7.3.4	可视化代码添加	207
7.3.5	可视化训练指标	207
<b>第 8 章</b>	<b>模型压缩</b>	<b>209</b>
8.1	模型压缩方法	209
8.1.1	模型设计压缩	210
8.1.2	网络剪枝与量化	213
8.1.3	张量分解	216
8.1.4	模型蒸馏与迁移学习	216
8.2	模型压缩实战	218
8.2.1	网络分析	220
8.2.2	输入尺度和第一层卷积设计	224
8.2.3	网络宽度与深度压缩	226
8.2.4	弥补通道损失	228
8.2.5	总结	230
<b>第 9 章</b>	<b>损失函数</b>	<b>231</b>
9.1	分类任务损失	231
9.1.1	什么是 0-1 loss	231
9.1.2	熵与交叉熵 loss	231
9.1.3	softmax loss 及其变种	232
9.1.4	KL 散度	237
9.1.5	Hinge loss 简介	237
9.1.6	Exponential loss 与 Logistic loss	237
9.1.7	多标签分类任务 loss	238
9.2	回归任务损失	238
9.2.1	L1 loss 与 L2 loss	238
9.2.2	L1 loss 与 L2 loss 的改进	239
9.3	常见图像任务与 loss 使用	240
9.3.1	图像基础任务	240
9.3.2	风格化与图像复原, 超分辨重建	240

---

9.3.3 生成对抗网络 .....	241
9.3.4 总结 .....	245
<b>第 10 章 模型部署与上线 .....</b>	<b>246</b>
10.1 微信小程序前端开发 .....	246
10.1.1 小程序的技术特点与定位 .....	246
10.1.2 Web 前端基础 .....	248
10.1.3 小程序开发工具 .....	251
10.1.4 小程序前端目录 .....	252
10.1.5 小程序前端开发 .....	254
10.2 微信小程序服务端开发 .....	260
10.2.1 域名注册与管理 .....	260
10.2.2 服务端框架简介 .....	261
10.2.3 算法搭建与实现 .....	262
10.3 Caffe 环境配置 .....	264
10.3.1 依赖库安装 .....	264
10.3.2 Caffe 编译安装 .....	266

# 第 1 章 神经网络基础

自计算机发展以来，人们逐渐使用计算机来代替繁杂冗余的计算及任务。在传统的编程方法中，我们需要告诉计算机如何去做，人为地将大问题划分为许多小问题，精确定义任务以便计算机执行。这些方法在解决图像识别、语音识别和自然语言处理等领域的疑难问题时，往往效果不佳；而使用神经网络处理问题时，不需要我们告诉计算机如何分解问题，而是由神经网络自发地从观测数据中进行学习，并找到解决方案。

深度学习作为目前最热门的技术研究方向之一，为许多通过传统方法解决不了的问题提供了另一种思路。本书面对的读者需要具备基本的数学知识和计算机知识。本节将对神经网络的基础做简单的回顾。

本章将从以下两个方面进行介绍。

- 1.1 节将从神经元到感知机再到 BP 算法，对神经网络的基础做概述。
- 1.2 节将对卷积神经网络的结构、核心概念进行简单介绍，这是当前深度学习模型的基础。

## 1.1 神经网络的生物基础与数学模型

深度学习并不是近几年才诞生的全新技术，而是基于传统浅层神经网络发展起来的深层神经网络的别称。本节将从神经网络的生物学基础到它的优化算法，对神经网络的基础做概述。

### 1.1.1 神经元

人工神经网络（Artificial Neural Network）简称神经网络（Neural Network, NN），是人类模拟生物神经网络的结构和功能提出的数学模型，广泛应用于计算机视觉等领域。

人工神经网络与生物神经网络有大量的相似之处，例如两者最基础的单元都是神经元。神经元又称神经细胞，是生物神经网络的基本组成，其外观和大小在神经系统中的差异很大，但都具有相同的结构体、胞体、树突和轴突。

胞体又称为核周体，由内质网、微管、游离核糖体、神经丝和核组成。轴突和树突是



神经元的突起，在神经元间传递电信号。神经元的功能是接受信号并对其做出反应、传导兴奋、处理并储存信息以及发生细胞之间的联结等，有这些功能，动物才能迅速对环境的变化做出整合性的反应。

神经元之间相互连接，当某一神经元处于“兴奋”状态时，其相连神经元的电位将发生改变，若神经元电位改变量超过了一定的数值（也称为阈值），则相连的神经元被激活并处于“兴奋状态”，向下一级连接的神经元继续传递电位改变信息。信息从一个神经元以电传导的方式跨过细胞之间的联结（即突触），传给另一个神经元，最终使肌肉收缩或腺体分泌。

神经元可以处理信息，也可以以某种目前还未知的方式存储信息。神经元通过突触的连接使数目众多的神经元组成比其他系统复杂得多的神经系统。从神经元的结构特性和生物功能可以得出结论：神经元是一个多输入、单输出的信息处理单元，并且对信息的处理是非线性的。

基于上述情形，1943年 McCulloch 和 Pitts 提出了 MP 模型，这是一种基于阈值逻辑算法的神经网络计算模型，由固定的结构和权重组成。

在 MP 模型中，某个神经元接受来自其余多个神经元的传递信号，多个输入与对应连接权重相乘后输入该神经元进行求和，再与神经元预设的阈值进行比较，最后通过激活函数产生神经元输出。每一个神经元具有空间整合特性和阈值特性。

MP 模型把神经元抽象为一个简单的数学模型，模拟生物神经元形式，成功证明了神经元能够执行逻辑功能，开创了人工神经网络研究的时代。从此神经网络的研究演变为两种不同的研究思路：一种是继续进行生物学原理方面的探究，着重关注大脑中信息传递与处理的生物学过程；另一种则演变为计算机学科，即神经网络在人工智能里的实际应用。后者的研究是模仿前者的原理机制。

## 1.1.2 感知机

感知机 (Perceptron) 是 Frank Rosenblatt 在 1957 年提出的概念，其结构与 MP 模型类似，一般被视为最简单的人工神经网络，也作为二元线性分类器被广泛使用。

通常情况下感知机指单层人工神经网络，以区别于多层感知机 (Multilayer Perceptron)。尽管感知机结构简单，但能够学习并解决较复杂的问题，其结构如图 1.1 所示。

假设我们有一个  $n$  维输入的单层感知机， $x_1$  至  $x_n$  为  $n$  维输入向量的各个分量， $w_1$  至  $w_n$  为各个输入分量连接到感知机的权量（或称权值）， $\theta$  为阈值， $f$  为激活函数（又称为激励函数或传递函数）， $y$  为标量输出。理想的激活函数  $f(\cdot)$  通常为阶跃函数或者 sigmoid 函数。感知机的输出是输入向量  $\mathbf{X}$  与权重向量  $\mathbf{W}$  求得内积后，经激活函数  $f$  所得到的标量：

$$y = f\left(\sum_{i=1}^n w_i x_i - \theta\right) \quad (1.1)$$