



深度学习

21天实战Caffe

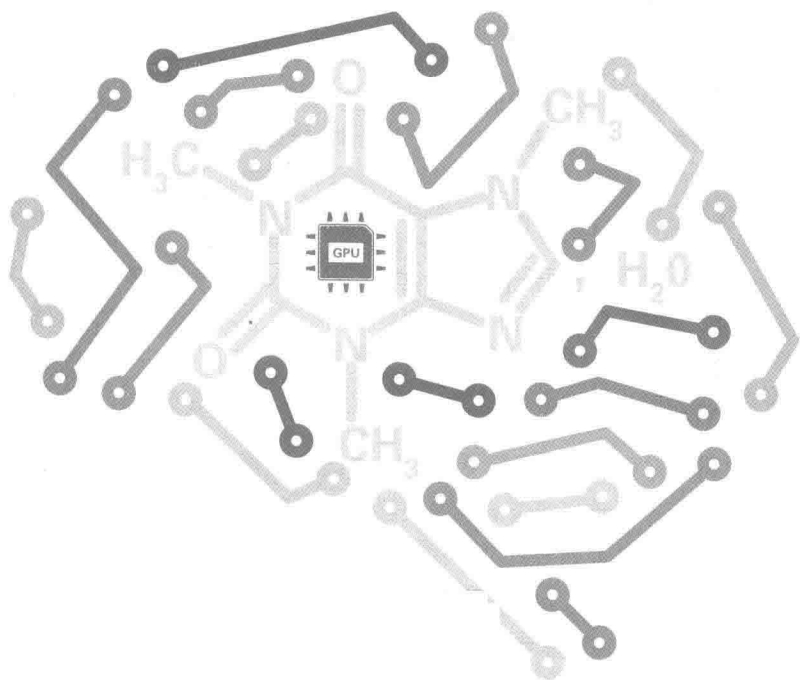
赵永科 著



中国工信出版集团



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
<http://www.phei.com.cn>



深度学习

21天实战Caffe

赵永科 著

电子工业出版社
Publishing House of Electronics Industry
北京·BEIJING

内 容 提 要

本书是一本深度学习入门读物。以目前已经大量用于线上系统的深度学习框架 Caffe 为例，由浅入深，从 Caffe 的配置、部署、使用开始学习，通过阅读 Caffe 源码理解其精髓，加强对深度学习理论的理解，最终达到熟练运用 Caffe 解决实际问题的目的。和国外机器学习、深度学习大部头著作相比，本书偏重动手实践，将难以捉摸的枯燥理论用浅显易懂的形式表达出来，透过代码揭开其神秘面纱，更多地贴近实际应用。

本书非常适合：对人工智能、机器学习感兴趣的读者；希望用深度学习完成设计的计算机或电子信息专业学生；准备开设机器学习、深度学习实践课的授课老师；学习过 C++，希望进一步提升编程水平的开发者；刚入坑的机器学习、语音、机器视觉、智能机器人研发或算法工程师。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有，侵权必究。

图书在版编目 (CIP) 数据

深度学习：21 天实战 Caffe / 赵永科著. —北京：电子工业出版社，2016.7
ISBN 978-7-121-29115-9

I. ①深… II. ①赵… III. ①学习系统 IV. ①TP273

中国版本图书馆 CIP 数据核字 (2016) 第 137508 号

策划编辑：张春雨

责任编辑：葛 娜

印 刷：三河市双峰印刷装订有限公司

装 订：三河市双峰印刷装订有限公司

出版发行：电子工业出版社

北京市海淀区万寿路 173 信箱 邮编：100036

开 本：787×980 1/16 印张：24.5 字数：530 千字

版 次：2016 年 7 月第 1 版

印 次：2016 年 7 月第 2 次印刷

定 价：79.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：(010) 88254888, 88258888。

质量投诉请发邮件至 zltz@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式：(010) 51260888-819, faq@phei.com.cn。

推荐序一

近年来，几乎整个智能学科的研究者们都注意到一个技术名词——深度学习（Deep Learning）。这个略带神秘色彩的名字和其代表的前沿性技术被著名的《MIT 技术评论》评选为 2013 年世界 10 大突破性技术之首。而在此之前，包括 Google、Microsoft、Facebook 等公司在内的诸多信息科技巨头都已争相在此技术上投入了前所未有的重视力度和战略资源，继而高调宣布布局智能应用领域。学术界和工业界不遗余力地抢占相关研究和技术的制高点，人们并没有感到奇怪，因为所有人都明白：这也许是人类在探索人工智能的伟大旅程和漫漫征途上的重要一刻。

关于神经网络的研究可以追溯到 20 世纪 40 年代。在其漫长的历史上经历了数次戏剧性的波折。然而近年来，随着大量数据的获得、先进理论的发现，以及高性能并行计算技术的发展，以深度神经网络为载体的特征学习技术相继在语音、视觉、语言等诸多研究领域中取得了突破性的成果，并且正以不可阻挡之势“入侵”传统技术占领的各个领域。

随着深度学习技术在学术界和工业界得到广泛认可，越来越多的人开始参与到深度学习的相关研究和实践中来。然而，由于存在一定的技术门槛，快速入手深度学习的研究并不是一件容易的事情。其中的一个重要原因是，深度学习中的许多问题非常依赖于实践。然而长期以来，学术界和工业界缺少一款专门为深度学习而设计的，兼具性能、灵活性和扩展性等诸多优势于一身的开源框架。这使得无论是快速实现算法，还是复现他人的结论，都存在着实践上的困难。研究人员和工程师们迫切需要一套通用而高效的深度学习开源框架。

2013 年，一款叫作“Caffe”的深度学习框架由加州大学伯克利分校的@贾扬清博士在 Github 上发布。发布伊始，Caffe 框架就得到了广泛的关注。Caffe 框架以“层”为单位对神经网络的结构进行了高度的抽象，通过一些精巧的设计显著优化了执行效率，并且在保持高效实现的基础上不失灵活性。无论在结构、性能上，还是代码质量上，Caffe 都是一款十分出色的开源框架。更重要的是，它将深度学习的每一个细节都原原本本地展现出来，供人们学习和实践。可以说，Caffe 框架的发布极大地降低了深度学习研究和开发的难度。

正是由于上述的诸多优势，Caffe 框架迅速流行起来，并且逐步形成了强大的用户社区。经过两年多的版本迭代，Caffe 框架已经在学术界和工业界得到了广泛的认可。在学术界，目前每天都有以 Caffe 框架作为底层实现的研究成果发布；而在工业界，已经有许多产品使用 Caffe 作为其深度学习算法实现的内核。从学术界到工业界，大家可以共享同一套底层代码，基于同一

套平台进行研究、交流和生产，这是一件令人愉悦的事情。可以说，Caffe 开源项目对于促进整个深度学习研究领域的快速发展具有不可磨灭的贡献。

对于刚刚接触深度学习的朋友们来说，通过结合 Caffe 的代码来加深对理论的理解，也许是一种事半功倍的方法。Caffe 框架天然的清晰层次和良好的代码可读性，为入手深度学习的朋友们提供了教科书般的实践指导。然而，由于 Caffe 中有大量技术细节是在论文中无法找到的，想要快速理解 Caffe 框架内部的种种精要往往需要费一番周折。幸好，有一些先行者为大家分享了相关的知识。

几天前，我有幸接到好友@卜居的邀请，为他的《深度学习：21 天实战 Caffe》新书做序。这本书是国内第一本在代码级别上全面剖析 Caffe 框架的指导书，同时也是一本真正的实战手册。本书涉及深度学习的基本理论、Caffe 的设计思想、Caffe 中各模块的具体实现，以及各种实例等内容。书中对 Caffe 框架的分析非常细致，涵盖的内容也颇为丰富，可以说是一本入手 Caffe 实践的技术手册，因此特别适合于 Caffe 的初学者阅读。相信本书可以帮助朋友们少走许多弯路。有关 Caffe 的诸多奥秘，@卜居将会在书中为您一一呈现。

感谢创立和推动深度学习研究的科学家们，感谢 Caffe 框架的作者贾扬清博士，感谢本书的作者@卜居，以及所有为深度学习技术的发展而奋斗的朋友们。我们的征途是星辰与大海，让我们一起努力，向着实现人工智能的伟大目标前进！

辛淼

Caffe 中国用户社区 (caffecn.cn) 创始人

推荐序二

在形形色色的科技创新报道中，人工智能（AI）成果引发的社会性冲击无疑前所未有的。一些科技术语如“深度学习”，对信息技术类学生而言变得不再陌生。虽然国际科技界对相关领域的研究已有数十年的历史，而跳跃式进步还只是最近若干年的事。正是这个原因，关于机器智能和“深度学习”的学习类书籍大多偏重理论，或散见于外刊上发表的研究论文、各个研究发展机构的研究报告、开源资料等，鲜有从工程实践出发系统地介绍深度学习的书籍。

国外研究机构设置的开源社区的繁荣发展，从工程实现方面补充了理论研究的不足。然而，对开源代码的阅读、理解、应用对于非机器学习专业人士有较大的挑战性。阿里云计算有限公司的赵永科工程师（博客昵称：卜居）在研发实践的基础上，对深度学习从基础理论到编程实践进行了系统的整理，形成了《深度学习：21 天实战 Caffe》一书。这是一个有技术深度、处于国际技术竞争中的领域；而本书是一个研发亲历者对技术深入理解后的总结，十分难得。

本书的写作风格是引导性的，围绕深度学习基础，通过代码导读方式，循序渐进，揭开了深度学习的神秘面纱，让深度学习技术，包括理论和工程实现，贴近所有 AI 爱好者。相信本书的出版能够激发更多研究者的兴趣，推动 AI 技术在中国的发展和应用。

邹谋炎

中国科学院大学教授，研究员，洪堡学者

推荐序三

让机器具有人类的智能是科学家们从计算机诞生开始就一直在努力的方向，但是由于选择了基于规则的算法，效果一直得不到大的提升，论文中经常以效果比乱猜好作为结论。卷积神经网络的发明者 Geoffrey Hinton 在 20 世纪 70 年代就已经提出了今天的深度学习理论，但是限于当时的计算能力，一直不被人重视。21 世纪以来，随着 NVIDIA GPU 的广泛应用，神经网络发挥了它应有的价值，成为今天人工智能的代表性成就，Hinton 也被尊称为鼻祖。

当 GPU 在深度学习领域大范围使用时，研究人员遇到了一个重大的问题——要写大量的复杂的神经网络代码，这带来了巨大的困难。在这个历史性的关键时刻，贾扬清同学开发的 Caffe 适时地出现了，Caffe 让只要会 C++ 编程的人员就可以编写深度学习代码，一下子就降低了深度学习的门槛。随后 Caffe 得到了广泛使用，并且获得了社区的广泛支持，也得到了 NVIDIA 的大力支持，获得了充足的发展，几乎可以说不知道 Caffe 就不能说会深度学习。

Caffe 把深度学习的门槛降低了很多，但是实际上依旧需要了解大量的代码细节才能对其进行修改，而深度学习又是一个计算密集的应用，如何写出高效的代码也非常重要。卜居做过许多有关 Caffe 的工作，包括优化卷积算法，非常了解 Caffe 框架的各个细节，他编写的《深度学习：21 天实战 Caffe》一书非常详细、专业。

卜居用人类的恋爱过程来比喻深度学习的学习过程，从初识、热恋到升华，很让人称道。在初识阶段，从深度学习的概念、历史开始，介绍深度学习和基本理论与传统机器学习算法的不同，也包含了业界对深度学习的反思。在热恋阶段，在具体实操方面，从 Caffe 的安装开始介绍，到具体运行 mnist 数据集；从 Caffe 的目录结构、不同层（功能）和数据抽象的实现细节，到如何求解一个深度学习模型，卜居都一一精确地解读。在升华阶段，卜居详细地解说了 Caffe 支持的 NVIDIA GPU 加速工具 CUDA 和 cuDNN，然后介绍了 Caffe 可视化方法，以及如何在生产环境中部署训练好的 Caffe 模型。

本书虽然以 21 天起名，但是其真实内容是需要读者每天 24 小时学习才能够完全掌握的，相信读者会一天 24 小时手不释卷。

我郑重地向大家推荐此书。

风辰

并行计算领域专家，深度学习平台架构师

前言

缘起

2014年，我开始在阿里云进行深度学习平台优化，使用了开源框架 Caffe，其间在多种 GPU 服务器上部署运行，同时也根据内部业务需求修改源码，研究基于 FFT 的卷积层加速方法（有幸与 Facebook AI Research Yann LeCun 老爷子同时做同一件事）。在学习和修改 Caffe 源码过程中根据个人理解写了数篇博客，本是无心插柳，孰料有一天博文视点编辑@永恒的侠少找到我，建议将博客内容进一步扩展为一本深度学习入门书，贴近实际，让更多读者走近这个如火如荼的领域。理想远大，现实残酷，互联网公司的工作强度大，常常加班，还要抽时间梳理写作思路，迟迟不能交出一份满意的书稿。熬夜艰辛，码字劳苦，时而思如泉涌，时而困顿踌躇，磨蹭半年有余，总算付梓。

在写作过程中，从头到尾重新审视了一遍近年大火的深度学习技术，提炼了经典论文的精华，深深为其近 30 年的坎坷历程唏嘘不已。我们目前看到的深度学习模型和基本理论可能比我们的年龄都要大，受限于当时计算能力不足，数据集也相对匮乏，深度学习一度陷入研究低谷，直到云计算、大数据时代的到来，日益增长的数据和计算能力为深度学习提供了适宜的温度和土壤，也为其突飞猛进的成长提供了充足的养料。

我们赶上了好时代。在物质和文化如此繁荣之时，有更多的人愿意思考未来——一个充满人工智能技术的时代。自动驾驶汽车、智能机器人、无人机……很多科幻电影中的技术正在成为现实，这一切都得益于深度学习技术和相应软硬件系统的发展进步。通过本书内容，读者将逐步走进深度学习，了解其过去、现在和未来。

本书读者

本书非常适合以下读者：

- 对人工智能、机器学习感兴趣的读者；
- 希望用深度学习完成设计的计算机或电子信息专业学生；

- 准备开设机器学习、深度学习实践课的授课老师；
- 学习过 C++，希望进一步提升编程水平的开发者；
- 刚入坑的机器学习、语音、机器视觉、智能机器人研发或算法工程师。

致谢

首先要感谢@永恒的侠少从茫茫人海中找到了我，出版深度学习入门书籍的想法我们一拍即合。然而，真正执笔才发现困难重重，业余时间常常被加班挤占，进度一拖再拖，在侠少不厌其烦一而再再而三地敦促下才将毫无条理的博客风格文章整理成书，自己也从中重新认识了深度学习，克服了拖延症，战胜了懒惰。

特别感谢我的爱人@晓曼在我情绪低落无心写作时给予的悉心照顾和支持，一次次谈心让我转换新的思路，跳出局部极小值点，坚定迈向下一个目标。

感谢远在家乡的父母，他们年过花甲，不懂我写的“深度学习”是什么，但他们知道我忙，知道我在为了理想付出，每次打电话都记挂着我，一次次叮嘱少熬夜。掩卷深思，自己欠父母太多，只愿能抽出更多的时间陪陪他们。

在阿里云工作已有近两年时间，非常 Nice 的同事们也对我的成长起到至关重要的作用，感谢老大@长仁提供良好的学习平台和环境，让我有机会接触最新架构 CPU、GPU、Phi、FPGA，通过一次次折腾机器，不断打怪升级攒经验。愿 2016 年我们的阿里云高性能计算服务 (<https://www.aliyun.com/product/hpc>) 随深度学习技术一起牛 X 一起飞！

感谢 CaffèCN 社区 (<http://www.caffecn.cn/>) 的@辛淼博士创建国内最前沿的深度学习交流平台，在这里聚集了一批始终代表先进生产力的 CS 硕博研究生、企业一线工程师，通过学术和技术讲座，让不同的思维火花碰撞，触发无限可能，欢迎更多的有志之士加入！

感恩互联网时代，感恩深度学习领域的先驱者 Geoffrey Hinton、Yann LeCun、YoushuaBengio，感恩青年才俊 Alex Krizhesky、YangqingJia 将生冷的深度学习理论转化为开源代码吸引了大量的热血青年前赴后继投身该领域，共建社区。

赵永科（笔名：卜居）

2016 年 4 月于杭州蓝湾 Coffee

CSDN 博客地址：<http://blog.csdn.net/kkk584520>

电子邮箱：zhaoyongke@yeah.net

目 录

上篇 初见

第 1 天 什么是深度学习	2
1.1 星星之火，可以燎原	3
1.2 师夷长技	4
1.2.1 谷歌与微软	4
1.2.2 Facebook、亚马逊与 NVIDIA	5
1.3 中国崛起	6
1.3.1 BAT 在路上	6
1.3.2 星光闪耀	7
1.3.3 企业热是风向标	8
1.4 练习题	9
第 2 天 深度学习的过往	10
2.1 传统机器学习的局限性	10
2.2 从表示学习到深度学习	11
2.3 监督学习	12
2.4 反向传播算法	13
2.5 卷积神经网络	15
2.6 深度学习反思	17
2.7 练习题	18
2.8 参考资料	18
第 3 天 深度学习工具汇总	19
3.1 Caffe	19
3.2 Torch & OverFeat	20
3.3 MxNet	22
3.4 TensorFlow	22
3.5 Theano	24

3.6	CNTK	24
3.7	练习题	25
3.8	参考资料	26
第 4 天	准备 Caffe 环境	27
4.1	Mac OS 环境准备	27
4.2	Ubuntu 环境准备	28
4.3	RHEL/Fedora/CentOS 环境准备	29
4.4	Windows 环境准备	29
4.5	常见问题	32
4.6	练习题	32
4.7	参考资料	33
第 5 天	Caffe 依赖包解析	34
5.1	ProtoBuffer	34
5.2	Boost	38
5.3	GFLAGS	38
5.4	GLOG	39
5.5	BLAS	40
5.6	HDF5	41
5.7	OpenCV	42
5.8	LMDB 和 LEVELDB	42
5.9	Snappy	43
5.10	小结	43
5.11	练习题	49
5.12	参考资料	49
第 6 天	运行手写体数字识别例程	50
6.1	MNIST 数据集	50
6.1.1	下载 MNIST 数据集	50
6.1.2	MNIST 数据格式描述	51
6.1.3	转换格式	53
6.2	LeNet-5 模型	60
6.2.1	LeNet-5 模型描述	60

6.2.2	训练超参数	65
6.2.3	训练日志	66
6.2.4	用训练好的模型对数据进行预测	76
6.2.5	Windows 下训练模型	76
6.3	回顾	78
6.4	练习题	79
6.5	参考资料	79
	篇尾语	80

中篇 热恋

第 7 天	Caffe 代码梳理	82
7.1	Caffe 目录结构	82
7.2	如何有效阅读 Caffe 源码	84
7.3	Caffe 支持哪些深度学习特性	86
7.3.1	卷积层	86
7.3.2	全连接层	89
7.3.3	激活函数	91
7.4	小结	99
7.5	练习题	99
7.6	参考资料	100
第 8 天	Caffe 数据结构	101
8.1	Blob	101
8.1.1	Blob 基本用法	102
8.1.2	数据结构描述	108
8.1.3	Blob 是怎样炼成的	109
8.2	Layer	125
8.2.1	数据结构描述	126
8.2.2	Layer 是怎样建成的	127
8.3	Net	136
8.3.1	Net 基本用法	136
8.3.2	数据结构描述	139
8.3.3	Net 是怎样绘成的	139

8.4	机制和策略	146
8.5	练习题	147
8.6	参考资料	148
第 9 天	Caffe I/O 模块	149
9.1	数据读取层	149
9.1.1	数据结构描述	149
9.1.2	数据读取层实现	150
9.2	数据变换器	155
9.2.1	数据结构描述	155
9.2.2	数据变换器的实现	156
9.3	练习题	171
第 10 天	Caffe 模型	172
10.1	prototxt 表示	173
10.2	内存中的表示	176
10.3	磁盘上的表示	176
10.4	Caffe Model Zoo	178
10.5	练习题	180
10.6	参考资料	180
第 11 天	Caffe 前向传播计算	181
11.1	前向传播的特点	181
11.2	前向传播的实现	182
11.2.1	DAG 构造过程	182
11.2.2	Net Forward 实现	190
11.3	练习题	192
第 12 天	Caffe 反向传播计算	193
12.1	反向传播的特点	193
12.2	损失函数	193
12.2.1	算法描述	194
12.2.2	参数描述	195
12.2.3	源码分析	195

12.3 反向传播的实现	203
12.4 练习题	205
第 13 天 Caffe 最优化求解过程	207
13.1 求解器是什么	207
13.2 求解器是如何实现的	208
13.2.1 算法描述	208
13.2.2 数据结构描述	210
13.2.3 CNN 训练过程	218
13.2.4 CNN 预测过程	225
13.2.5 Solver 的快照和恢复功能	227
13.3 练习题	230
第 14 天 Caffe 实用工具	231
14.1 训练和预测	231
14.2 特征提取	241
14.3 转换图像格式	247
14.4 计算图像均值	254
14.5 自己编写工具	257
14.6 练习题	257
篇尾语	258

下篇 升华

第 15 天 Caffe 计算加速	260
15.1 Caffe 计时功能	260
15.2 Caffe GPU 加速模式	262
15.2.1 GPU 是什么	262
15.2.2 CUDA 是什么	263
15.2.3 GPU、CUDA 和深度学习	263
15.2.4 Caffe GPU 环境准备	264
15.2.5 切换到 Caffe GPU 加速模式	268
15.3 Caffe cuDNN 加速模式	269
15.3.1 获取 cuDNN	270

15.3.2	切换到 Caffe cuDNN 加速模式	270
15.3.3	Caffe 不同硬件配置性能	272
15.4	练习题	273
15.5	参考资料	273
第 16 天	Caffe 可视化方法	275
16.1	数据可视化	275
16.1.1	MNIST 数据可视化	275
16.1.2	CIFAR10 数据可视化	277
16.1.3	ImageNet 数据可视化	278
16.2	模型可视化	279
16.2.1	网络结构可视化	279
16.2.2	网络权值可视化	281
16.3	特征图可视化	288
16.4	学习曲线	295
16.5	小结	298
16.6	练习题	298
16.7	参考资料	299
第 17 天	Caffe 迁移和部署	300
17.1	从开发测试到生产部署	300
17.2	使用 Docker	302
17.2.1	Docker 基本概念	302
17.2.2	Docker 安装	303
17.2.3	Docker 入门	305
17.2.4	Docker 使用进阶	312
17.3	练习题	317
17.4	参考资料	317
第 18 天	关于 ILSVRC 不得不说的一些事儿	318
18.1	ImageNet 数据集	318
18.2	ILSVRC 比赛项目	319
18.2.1	图像分类 (CLS)	320
18.2.2	目标定位 (LOC)	320

18.2.3	目标检测 (DET)	321
18.2.4	视频目标检测 (VID)	322
18.2.5	场景分类	322
18.3	Caffe ILSVRC 实践	323
18.4	练习题	326
18.5	参考资料	326
第 19 天	放之四海而皆准	327
19.1	图像分类	327
19.1.1	问题描述	327
19.1.2	应用案例——商品分类	330
19.2	图像中的字符识别	332
19.2.1	问题描述	332
19.2.2	应用案例——身份证实名认证	333
19.3	目标检测	337
19.3.1	问题描述	337
19.3.2	最佳实践——运行 R-CNN 例程	337
19.4	人脸识别	340
19.4.1	问题描述	340
19.4.2	最佳实践——使用 Face++ SDK 实现人脸检测	342
19.5	自然语言处理	343
19.5.1	问题描述	343
19.5.2	最佳实践——NLP-Caffe	344
19.6	艺术风格	350
19.6.1	问题描述	350
19.6.2	最佳实践——style-transfer	352
19.7	小结	354
19.8	练习题	354
19.9	参考资料	355
第 20 天	继往开来的领路人	356
20.1	Caffe Traps and Pitfalls	356
20.1.1	不支持任意数据类型	356
20.1.2	不够灵活的高级接口	357