

PyTorch

深度学习入门

曾凡壹◎著



$$\text{Sigm}(x) = \frac{1}{1 + e^{-x}} \quad \nabla L = \left(\frac{\partial L}{\partial w_1}, \frac{\partial L}{\partial w_2} \right)$$
$$y = \sqrt{z_1^2 + z_2^2} \quad \hat{y} = \vec{x} \cdot \vec{w} \quad (x^{(1)}, y^{(1)})$$
$$f(x) = \begin{cases} 0, & x \leq 0 \\ +\infty, & x > 0 \end{cases} \quad P(y|0) = \vec{0} \cdot (1 - \vec{0})^T$$
$$\begin{bmatrix} 2 & 4 & 5 \\ 7 & 6 & 3 \end{bmatrix} \quad \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$
$$x \quad x \quad \phi \quad e$$
$$+ \quad + \quad + \quad +$$

TURING

图灵原创

PyTorch

深度学习入门

曾凡壹◎著



$$\text{Sign}(x) = \frac{1}{1 + e^{-x}} \quad \nabla L = \left(\frac{\partial L}{\partial w_1}, \frac{\partial L}{\partial w_0} \right)$$

$$y = \sqrt{z_1^2 + z_2^2} \quad \hat{y} = \vec{x} \cdot \vec{w} \quad (x^{(1)} y^{(1)})$$

$$f(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases}$$

$$\begin{bmatrix} 2 & 4 & 5 \\ 7 & 6 & 3 \end{bmatrix}$$

$$x \times \phi e$$

常州大学图书馆
藏书章

RFID

人民邮电出版社
北京

图书在版编目 (CIP) 数据

PyTorch深度学习入门 / 曾芄壹著. — 北京: 人民邮电出版社, 2019.9

(图灵原创)
ISBN 978-7-115-51919-1

I. ①P… II. ①曾… III. ①机器学习 IV. ①TP181

中国版本图书馆CIP数据核字(2019)第188297号

内 容 提 要

本书用浅显易懂的语言,图文并茂地讲解了深度学习的基础知识,从如何挑选硬件到神经网络的初步搭建,再到实现图片识别、文本翻译、强化学习、生成对抗网络等多个目前最流行的深度学习应用。书中基于目前流行的PyTorch框架,运用Python语言实现了各种深度学习的应用程序,让理论和实践紧密结合。本书适合打算入门深度学习的人群以及对Python语言有初步了解的读者阅读。

-
- ◆ 著 曾芄壹
责任编辑 王军花
责任印制 周昇亮
 - ◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路11号
邮编 100164 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
北京鑫正大印刷有限公司印刷
 - ◆ 开本: 800×1000 1/16
印张: 15 彩插: 2
字数: 346千字 2019年9月第1版
印数: 1-3 500册 2019年9月北京第1次印刷
-

定价: 59.00元

读者服务热线: (010)51095183转600 印装质量热线: (010)81055316

反盗版热线: (010)81055315

广告经营许可证: 京东工商广登字 20170147号



曾芃壹

现为中山大学数据科学与计算机学院在读硕士，主要研究兴趣有深度学习、语音识别、推荐系统、自动犯罪侦查等。熟悉C、C++、Java、Python等多种程序设计语言，Flask建站技术以及PyTorch、TensorFlow深度学习框架，在简书上写了多篇PyTorch的文章，深受读者欢迎。



微信连接



回复“深度学习”查看相关书单



微博连接

关注@图灵教育 每日分享IT好书



QQ连接

图灵读者官方群I: 218139230

图灵读者官方群II: 164939616

图灵社区

iTuring.cn

在线出版, 电子书, 《码农》杂志, 图灵访谈

站在巨人的肩上
Standing on Shoulders of Giants



iTuring.cn

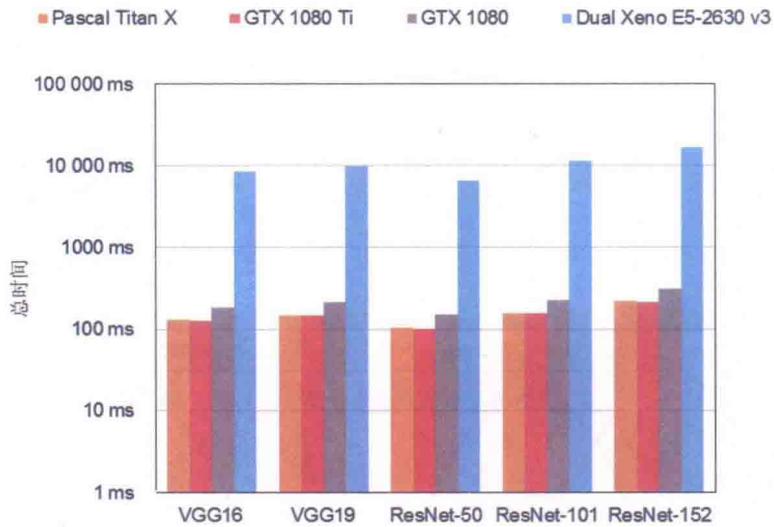


图 1-1 不同处理器在深度卷积神经网络任务中的表现

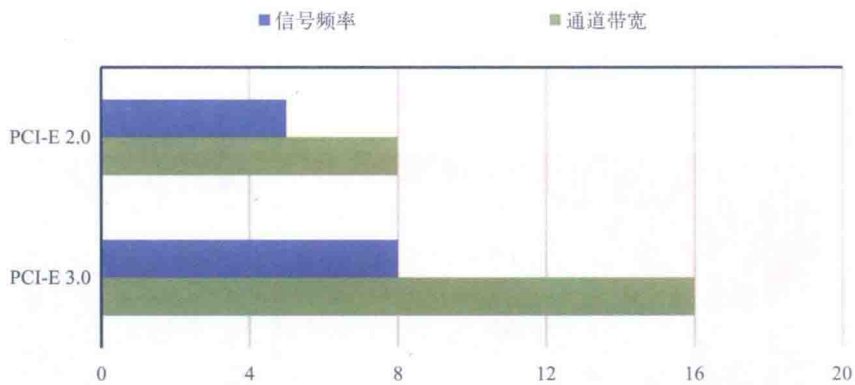


图 1-3 PCI-E 2.0 与 PCI-E 3.0 的性能比较

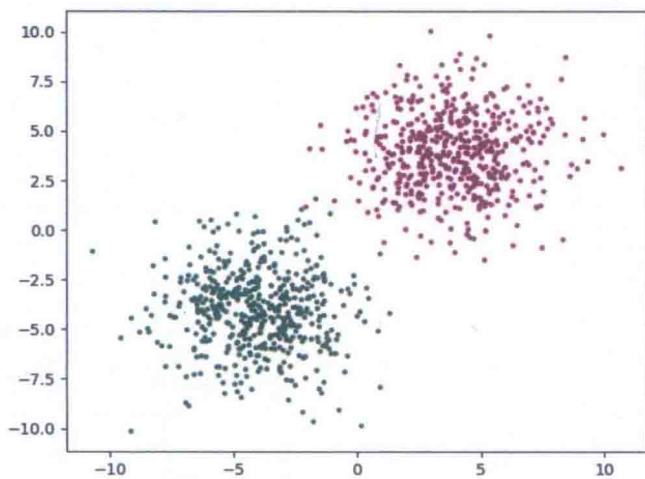


图 3-23 自动生成的数据分布

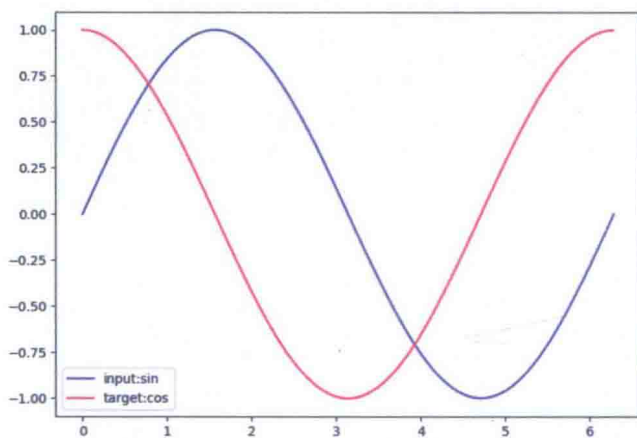


图 5-6 sin 和 cos 函数分布

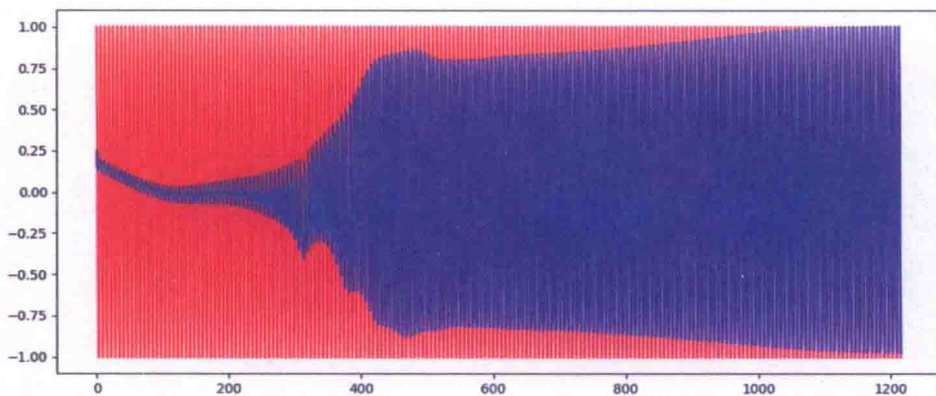


图 5-7 RNN 实例训练结果

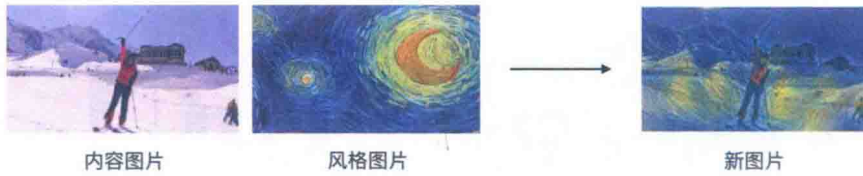


图 8-1 风格迁移示意图



图 8-9 损失权重对比图

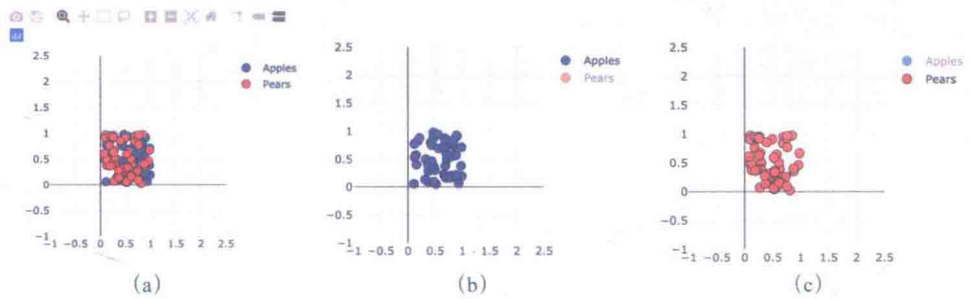


图 11-9 visdom 绘制二维散点图

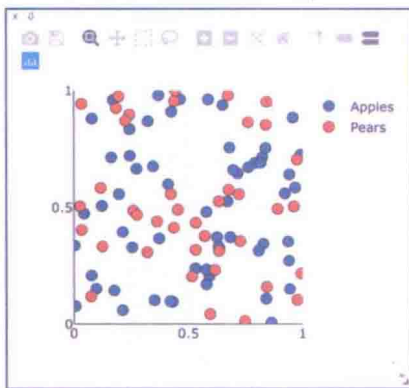


图 11-10 更新散点图

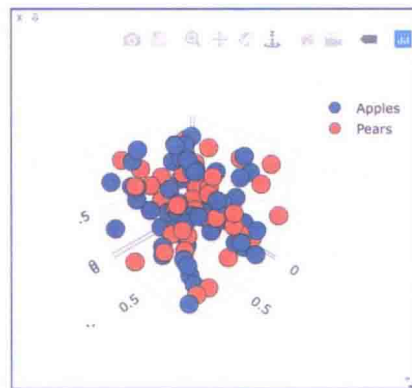


图 11-11 visdom 的三维散点图

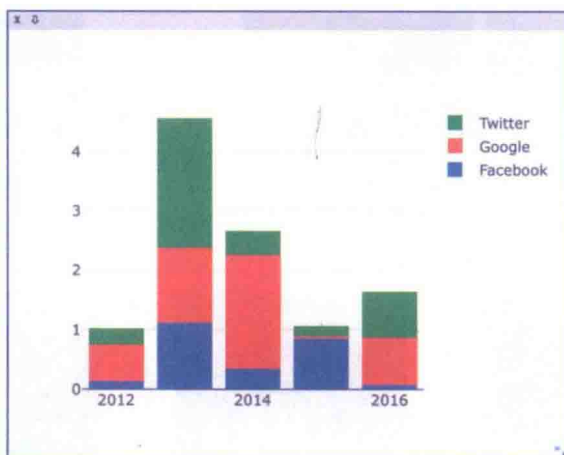


图 11-15 visdom 多标签直方图



图 11-19 TensorBoard 的多曲线图

前言

概述

近年来，学术界和商业界掀起了人工智能的热潮，这股热潮背后的推动力量正是深度学习技术。如今，在图像、语音和自然语言等数据的处理方面，基于深度学习的算法均以绝对的优势碾压传统算法。为了让研究者和开发人员能够更快地构建各式各样的深度学习模型，各大公司纷纷为 Python 语言开发自己的深度学习依赖库，其中较为出名的有谷歌公司的 TensorFlow、Facebook 公司的 PyTorch、Apache 公司的 MXNet、微软公司的 CNTK 以及百度公司的 PaddlePaddle。从 2017 年 1 月 Facebook 公司开源 PyTorch 以来，短短两年时间，PyTorch 便与老牌框架 TensorFlow 势均力敌。PyTorch 的迅猛发展得益于其丰富且全面的开发文档、活跃的社区和易学易用的特点。它不但对深度学习的初学者非常友好，还能高效地完成各式各样的深度学习任务并结合 Caffe2 快速部署到生产环境中。

想要读懂本书，只需具备微积分、线性代数以及 Python 编程的基本知识即可。读者可将本书作为入门深度学习技术的第一本书，通过 PyTorch 框架实现当今最潮的深度学习算法。

本书结构

全书共 12 章，分为基础篇（第 1 章~第 3 章）、实战篇（第 4 章~第 8 章）和高级篇（第 9 章~第 12 章）3 个部分，各章的知识要点如下。

第 1 章讲解如何打造深度学习的硬件环境，在不同系统下配置 PyTorch 的运行环境。

第 2 章介绍张量的基本知识以及 PyTorch 中处理张量的常用方法，为后面构建深度神经网络打下基础。

第 3 章介绍深度学习的基本概念，用 PyTorch 实现线性回归、逻辑回归、分类以及基于深度卷积神经网络的 MNIST 手写字体识别。

第 4 章介绍 4 个目前比较流行的图片识别模型，并利用 PyTorch 实现 AlexNet 的迁移学习模型。

第 5 章介绍序列转序列模型的基础知识，用 PyTorch 构建一个将法语翻译成英语的神经翻译机。

第 6 章介绍生成对抗网络模型的基本概念，通过 PyTorch 实现一个二次元人物头像的生成器。

第 7 章浅谈强化学习中基于策略的算法和基于值的算法，用 PyTorch 在 Gym 平台上实现一个简单的强化学习智能体。

第 8 章讲解风格迁移的基本概念，用 PyTorch 借助已训练好的 VGG19 模型来实现风格迁移。

第 9 章介绍如何自定义神经网络层以及如何使用 C++ 加载 PyTorch 的模型。

第 10 章介绍 ONNX，并讲解如何使用 ONNX 将 PyTorch 模型迁移至其他框架中。

第 11 章讲解 PyTorch 可视化工具 TensorBoard 和 visdom 的基本用法，以及模型可视化工具 Netron 的使用方法。

第 12 章介绍如何用 PyTorch 创建多进程任务和多 GPU 并行计算。

基础篇可以帮助读者充分了解深度学习的基础知识和 PyTorch 最基本的使用方法。实战篇可以让读者对不同的深度学习应用场景有粗略的认知，通过代码来快速上手，读者甚至可以通过简单修改本书提供的代码来实现自己的个性化应用。高级篇主要讲解在生产环境下部署时需要用到的一些技巧以及 PyTorch 的可视化方法。本书的配套代码可在图灵社区（iTuring.cn）的本书主页中免费注册下载。

致谢

感谢王军花和武芮欣两位女士，是她们给了我这次写作的机会以及专业的修改意见。

感谢我的妻子，是她给了我支持与陪伴，让我拥有充足的写作空间和时间。以此书纪念我们从相识到成婚的这段时光。

感谢我的父母，他们在我的成长道路上默默给予爱和鼓励，才使我最终完成此书。

曾芄壹

2019 年 4 月于广东梅州

目 录

第一部分 基础篇

第 1 章 准备工作	2
1.1 硬件配置	2
1.2 在 Mac OS X 系统下配置 PyTorch 运行环境	6
1.3 在 Ubuntu 系统下配置 PyTorch 运行环境	8
1.4 在 Windows 系统下配置 PyTorch 运行环境	14
第 2 章 Tensor 基础	17
2.1 Tensor	17
2.2 Autograd	30
第 3 章 深度学习基础	35
3.1 机器学习	35
3.2 线性回归	38
3.3 非线性回归	53
3.4 逻辑回归	58
3.5 多元分类	66
3.6 反向传播	70
3.7 卷积神经网络	72
3.8 手写字体识别	78
3.9 fastai 手写字体识别	86

第二部分 实战篇

第 4 章 迁移学习	90
4.1 经典图像模型	90
4.2 迁移学习实战	100
4.3 使用 fastai 实现迁移学习	109
第 5 章 序列转序列模型	111
5.1 循环神经网络模型	111
5.2 神经翻译机简介	119
5.3 利用 PyTorch 构造神经翻译机	122
第 6 章 生成对抗网络	138
6.1 生成对抗网络概览	138
6.2 使用生成对抗网络生成二次元头像	142
6.3 使用 TorchGAN 生成二次元头像	149
第 7 章 深度强化学习	153
7.1 深度强化学习	153
7.2 基于策略的算法	155
7.3 基于值的算法	157
7.4 Gym 简介	161
7.5 Q-Learning 实战	163
第 8 章 风格迁移	168
8.1 风格迁移原理	168
8.2 风格迁移实践	174

第三部分 高级篇

第 9 章 PyTorch 扩展.....	184	第 11 章 PyTorch 可视化.....	201
9.1 自定义神经网络层.....	184	11.1 使用 visdom 实现 PyTorch 可视化.....	201
9.2 C++加载 PyTorch 模型.....	189	11.2 使用 TensorBoard 实现 PyTorch 可视化.....	213
第 10 章 PyTorch 模型迁移.....	193	11.3 使用 Netron 显示模型.....	221
10.1 ONNX 简介.....	193	第 12 章 PyTorch 的并行计算.....	223
10.2 使用 ONNX 将 PyTorch 模型 迁移至 Caffe2.....	196	12.1 多进程.....	223
10.3 使用 ONNX 将 PyTorch 模型 迁移至 Core ML.....	199	12.2 多 GPU 并行计算.....	231

第一部分

基础篇

$$\pi_i = \frac{e^{x_i}}{\sum_{j=1}^k e^{x_j}}$$

$$\text{sigm}(x) = \frac{1}{1 + e^{-x}}$$

$$y = \sqrt{z_1^2 + z_2^2}$$

$$f(x) = \begin{cases} 0, & x \leq 0 \\ x, & x > 0 \end{cases}$$

$$P(y|0) = \delta(y - 0)$$

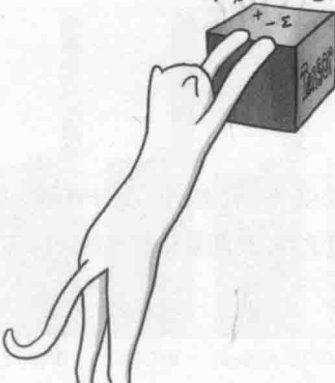
$$w^{t+1} = w^t - \frac{dL}{dw^t} \times \delta$$

$$\nabla L = \left(\frac{\partial L}{\partial w_1}, \frac{\partial L}{\partial w_2} \right)$$

$$\begin{bmatrix} 2 & 4 & 5 \\ 7 & 6 & 3 \end{bmatrix} \hat{y} = \hat{x} \cdot \hat{w}$$

$$\hat{x} \times \phi \ e$$

$$+ \epsilon$$



$$\frac{1}{3} \times \frac{1}{2}$$

2017.7.15

第 1 章

准备工作

2017 年 1 月，Facebook 人工智能研究团队在 GitHub 上开源了 PyTorch，随后迅速占领了 GitHub 热度榜榜首。如今，PyTorch 已经更新至 1.0 版本，成为众多科研机构研究深度学习的首选框架。PyTorch 具有简洁优雅的接口和先进的动态图设计，可以快速且灵活地构造复杂的神经网络模型。同时，PyTorch 具有与其他流行框架（如 TensorFlow 和 Keras 等）比肩的运行速度。通过 ONNX 接口，可以将 PyTorch 模型导入 Caffe2，兼容各种终端环境。另外，PyTorch 拥有完整易读的文档和非常活跃的社区，这也使它成为深度学习入门的首选框架。

本章主要为大家介绍以下内容：

- 如何挑选适合做深度学习的硬件
- 如何在 macOS 系统下配置 PyTorch 运行环境
- 如何在 Ubuntu 系统下配置 PyTorch 运行环境
- 如何在 Windows 系统下配置 PyTorch 运行环境

1.1 硬件配置

工欲善其事，必先利其器。为了满足深度学习任务中“高计算量”这一特殊需求，我们需要配置一台适合的计算机。作为一个深度学习的初学者，在配置计算机的时候不仅要考虑硬件的性能和价格，而且要考虑其扩展性。

1. 显卡

为了加速计算机的图形处理速度，纳斯达克（NVIDIA）公司在 1999 年发布 GeForce 256 图形处

理芯片时提出了 GPU (Graphic Processing Unit) 的概念, 此后又提出了 GPGPU (General Purpose Graphic Processing Unit) 的概念, 即利用 GPU 进行一些通用计算, 从而解决 CPU 浮点运算能力不足的问题。也正是因为 GPGPU 的兴起, 促使了深度神经网络的蓬勃发展。

NVIDIA 公司推出的 CUDA (Compute Unified Device Architecture) 是一种使用 GPU 进行通用计算的架构。在 CUDA 架构的基础上, NVIDIA 发布了 cuDNN 库, 用于支持其显卡对深度神经网络 (Deep Neural Network, DNN) 的加速计算。现在众多的深度学习框架 (如 PyTorch、Caffe、TensorFlow 和 Theano) 都支持 cuDNN 加速。

为什么 GPU 对于深度神经网络如此重要? 我们可以让 GPU 和 CPU 处理相同的深度神经网络任务, 然后进行一次对比。深度卷积神经网络 (Convolutional Neural Network, CNN) 是一种经常被用于处理图片数据的深度神经网络, 这里采用 16 张 $3 \times 224 \times 224$ 的彩色图片作为一组数据, 分别传入不同的深度卷积神经网络中, 不同处理器之间的表现存在差异。

图 1-1 演示了 GPU 和 CPU 在处理相同的深度卷积神经网络任务时性能上的巨大差距。在 Pascal Titan X、GTX 1080 Ti 和 GTX 1080 这三个 GPU 上分别处理 VGG16、VGG19、ResNet-50、ResNet-101 和 ResNet-152 时, 所用的时间在 101 ms 至 314 ms 之间, 而 CPU 处理器 Dual Xeno E5-2630 v3 在处理相同任务时所用的时间在 6627 ms 至 16 872 ms 之间。在处理上述任务时, Pascal Titan X 比 Dual Xeno E5-2630 v3 快了 50 倍到 76 倍!

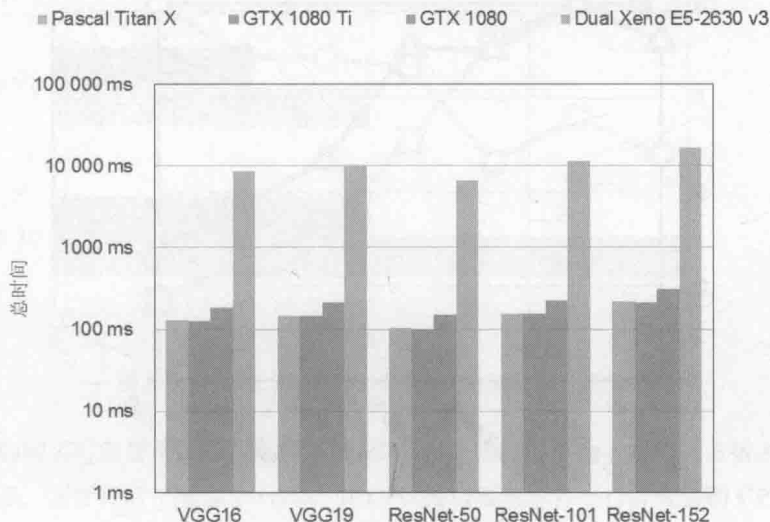


图 1-1 不同处理器在深度卷积神经网络任务中的表现 (另见彩插)