

人工智能 | AI 掘金，先从这座金矿开始.....

原创：IDZ英特尔开发人员专区2017-09-15



学习和研究 AI 的专业人士，都知道到哪里去看论文掘金，今天，小编分享几篇英特尔计算智能协同研究院 (ICRI-CI) 专家发布在金矿 arxiv.org 里面的一些专业文章，抛砖引玉。祝大家继续烧脑，学习愉快。

近年来，深度学习脱颖而出，成为完成各项人工智能任务的领先技术。深度学习是一种十分先进的方法，适用于多个领域，包括对象识别与验证、文本理解与翻译、问题回答等。此外，深度学习还有望在多种（以前认为不可能的）全新用途中发挥重要作用，比如自动驾驶。

尽管深度学习解决复杂问题的能力一次次得到了充分的验证，但在深度学习为何奏效、具备哪些真正的能力、何时奏效（以及何时不奏效）等方面依然存在许多谜团。无论是理论研究还是用于实践，都必须首先解开这些谜团，才能了解如何安全有效地使用这些方法。尽管人们对此展开了深入研究，并获得了重要洞察，但大部分谜团依然没有解开。我们普遍认为深度学习就像是一种“神奇的魔法”，但我们其实并不了解个中原理。缺乏真正的了解将导致科学家无法深入探索，开发人员也因此

而担忧大家会不会让运行机制和弱点都不得而知的系统来驱动自动驾驶汽车？

自 2012 年成立以来，英特尔计算智能协同研究院 (ICRI-CI) 一直大力支持机器学习和深度学习方面的研究。我们集合了六位最优秀的 ICRI-CI 深度学习研究人员解决“深度学习为何奏效以及如何奏效”的挑战，旨在深入了解深度学习、提供有关深度网络运行原理的洞察，并展示他们在深度学习的表现方式、局限性和潜能等方面的重要研究成果。

这次挑战取得了可喜的成果，他们撰写了五篇报告，介绍如何解决深度学习的各种难题。报告总结了研究人员在重大会议和期刊上发布的著作内容，以及专为撰写此文所收集的最新研究成果。相关难题包括深入了解深度网络为何奏效，何时奏效（以及何时不奏效），深度网络的几何表达方式所产生的影响，以及实现的深度网络的可理解性。

了解深度网络为何奏效，何时奏效？

（以及何时不奏效）

1

Naftali Tishby 和 Ravid Schwartz-Ziv 在[通过信息打开深度神经网络黑匣子](#)一文中对深度网络进行了深入研究，采用的方法包括分析信息理论属性、查看各层关于输入输出的信息内容，以及建议网络在每一个层级都明确尝试优化压缩与预测之间的信息瓶颈 (IB) 折衷。

另外文中指出，用于培训网络的随机梯度下降（SGD）时代可划分成两个明显的阶段：缓慢表现压缩和快速经验错误最小化。之后他们在关于隐藏层的计算优势方面提出了一个新的理论论据。

2

Shai Shalev-Shwartz、Ohad Shamir 和 Shaked Shamma 在**基于梯度的深度学习的失败**中尝试深入了解与常用方法和算法相关的困难和局限性。

文中列举出四类导致现有算法失败或遭遇重大困难的问题、通过实际体验证明这些问题的失败之处、提供相关的理论洞察解释问题来源，并提出解决这些问题以提升性能的补救措施。

3

Amnon Shashua、Nadav Cohen、Or Sharir、Ronen Tamari、David Yakira 和 Yoav Levine 在**通过层级张量分解分析和设计卷积网络**中分析了深度卷积网络的表现属性。

他们通过层级张量分解的等价性，研究卷积网络中不同架构特性（深度、宽度、池化几何体、互联、重叠操作等）的表现效率和归纳偏置。他们的研究结果为验证卷积网络的有效性指明了方向，并提供了新的网络设计工具。

4

Nathan Srebro、Behnam Neyshabur、Ryota Tomioka 和 Ruslan Salakhutdinov 在**深度学习中的几何优化和隐式正则化**中指出，用于神

神经网络培训的优化方法通过隐式正则化对深度学习模型的泛化能力起着关键的作用。

他们经过论证表明，泛化能力不取决于网络规模，而是取决于部分其他的隐形控制。他们之后对深度网络的参数空间几何体进行了研究，建议根据该几何体调整优化算法，并因此证明了不断变化的实证优化程序可显著提高泛化性能。

5

Shie Mannor、Tom Zahavy 和 Nir Baram 在灰色化**黑匣子：了解 DQN** 中展示了以非盲法分析深度 Q 网络 (DQN) 的方法和工具。他们建议采用一种新模型——半聚合马尔可夫决策过程 (SAMDP)，和一种可自动学习的算法。

他们使用这些工具，表明 DQN 学习的特性能够以层级的方式聚合状态空间，进而解释其成功之处。另外，他们还能够深入观察网络，了解和描述 DQN 针对三款不同的 Atari2600 游戏所学习的策略，并提供在增强学习过程中解释、调试和优化深度神经网络的方法。