

# Uber AI实验室总结「神经进化」研究：如何利用遗传算法优化网络？

转载

[AI科技大本营](#)



于 2017-12-20 00:00:00 发布



1499



收藏



本文授权转自雷克世界（微信：ROBO\_AI）

编译 | 嗯~阿童木呀

现如今，在深度学习领域中，我们通过随机梯度下降（SGD）对具有多个层和数百万个连接的深度神经网络（DNN）进行常规训练。许多人认为，SGD所具有的对梯度进行有效训练的能力是它有能力对深度神经网络进行训练的关键所在。然而，在我们新近发布的五篇论文中，我们支持这样一种新提出的观点，神经进化（neuroevolution）也是一个用于解决强化学习（RL）问题的训练深度神经网络的有效方法，其中，通过进化算法优化神经网络。Uber涉及的领域很多，其中机器学习可以改进其操作，而开发一系列涵盖神经进化在内的强有力学习方法将帮助我们实现我们的使命，即开发出更安全、更可靠的交通解决方案。

## 遗传算法——用于训练深度神经网络的有效替代方案

使用我们发明的一种新技术来使深度神经网络进化，在这个过程中，我们发现一个极其简单的遗传算法（GA）可以对超过400万个参数的深度卷积网络进行训练，从而使其能够从像素点进行Atari游戏，而且在许多游戏中的性能表现要远超过现代深度强化学习（RL）算法（例如DQN和A3C）或进化策略（ES），与此同时，还因为具有更好的并行化而运行地更快。这个结果是令人惊讶的，我们之所以这样认为是因为，一方面是没有期待这个不是基于梯度的遗传算法能够很好地扩展到如此大的参数空间中，另一方面是认为在强化学习中使用遗传算法的性能表现是不可能匹配或超越现有技术的。

我们进一步表明，现代遗传算法改进提高了系列遗传算法的能力，例如全新性搜索，也可以在深度神经网络规模下运作，并且可以促进探索以解决欺骗性问题（那些具有挑战局部最优的问题），这些问题阻碍了诸如Q-learning（DQN）、政策梯度（A3C）、进化策略和遗传算法等奖励最大化算法。

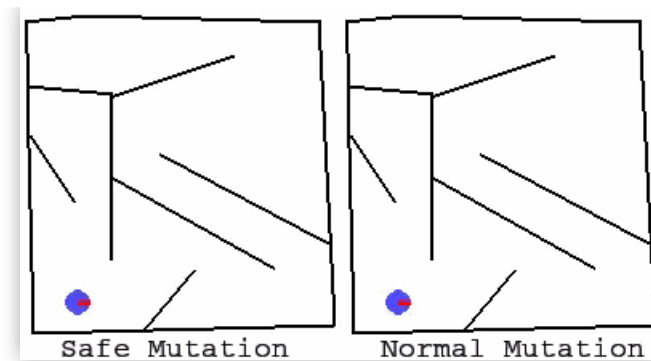


左：GA策略在Frostbite游戏中得到了10500分，而DQN、A3C、ES等得分不足1000分。

右：GA策略在Asteroids游戏中表现非常好，平均分超过了DQN和ES，但不及A3C。

### 通过梯度计算实现安全突变

在一篇论文中，我们将展示梯度是如何与神经进化相结合，以提高进化循环和极其深的深度神经网络的能力，实现了超过一百层的深度神经网络的进化，这要远远超过以前通过神经进化实现的水平。我们通过计算网络输出相对于权重的梯度（即，不同于传统深度学习中的误差梯度）来实现这一点，使得随机突变的校准能够更精细地对最敏感的参数进行处理，从而解决大型网络中随机突变的主要问题。



两个动画分别显示了一个解决迷宫问题的单一网络中的一组变异（从左下角到左上角），正常突变大多不能到达终点，而安全突变很大程度上保留了这种能力，同时还产生了多样性，这说明安全变异具有显著的优势。

### 进化策略（ES）如何与随机梯度下降（SGD）相关联？

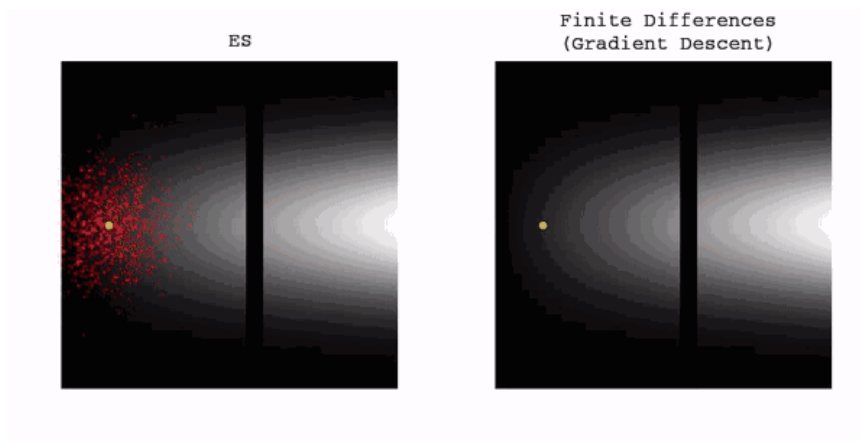
在一篇论文中，我们对OpenAI的一个团队首次提出的一个发现做出了补充，神经元进化的进化策略多样性可以在深度强化学习任务上有效优化深度神经网络。然而，迄今为止，这个结果的更为广泛的影响仍然受到质疑。基于进化策略，为了进行进一步探索，我们通过一个全面的研究，深入了解了ES与SGD的关系，研究了ESIST梯度近似与MNIST数据集上由SGD计算的每个小批量的最佳梯度的近似程度，以及这个近似值所必须要表现良好的程度。结果证明，如果提供足够的计算来改善其梯度近似，ES可以在MNIST上达到99%的精确度，这暗示了为什么ES将越来越成为深度强化学习中的一个强有力的竞争者，其中，在并行计算增加的情况下，没有任何方法可以获得完美的梯度信息。

### ES不仅仅是传统的有限差分法（finite differences）

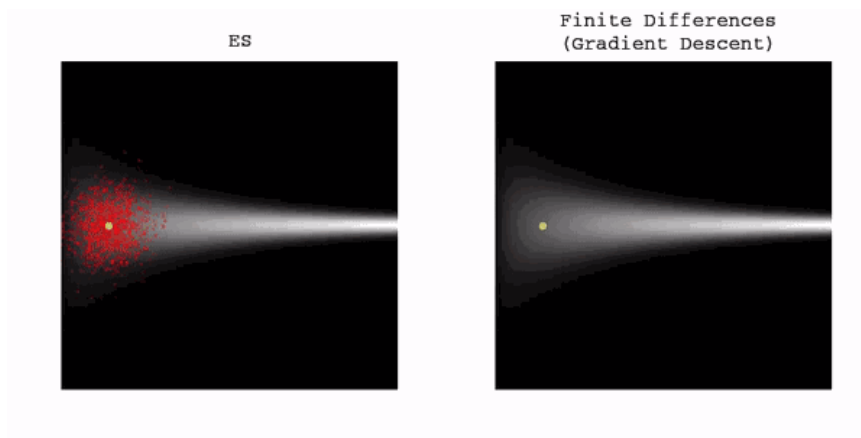
为了加强进一步的理解，经同行业研究证实，ES（具有足够大的干扰大小参数）的行为与SGD不同，因为它优化的是由概率分布所描述的策略的预期奖励（搜索空间中的一片云），而SGD优化的是单一策略的奖励（搜索空间中的一个点）。这种变化使得ES访问搜索空间的不同区域，无论好坏（这两种情况都有相关说明）。对参数干扰总体进行优化的另一个结果是，ES获得了不是通过SGD获得的鲁棒性属性。强调ES在参数总体上所进行的优化，也就强调了ES和贝叶斯方法之间的有趣联系。

```
class="video_iframe" data-vidtype="2" allowfullscreen="" frameborder="0" data-ratio="1.875" data-w="480"
scrolling="no" data-src="http://v.qq.com/iframe/player.html?
vid=s13306eg55g&width=654&height=367.875&auto=0" style="display: block; width: 654px !important;
height: 367.875px !important;" width="654" height="367.875" data-vh="367.875" data-vw="654"
src="http://v.qq.com/iframe/player.html?vid=s13306eg55g&width=654&height=367.875&auto=0"/>
```

由TRPO学习的步行者权重引发的随机干扰，导致它与由ES进化出来的相同质量步行者发生的随机干扰相比，步态明显更不稳定。原始的步行者位于每个九宫格中心。



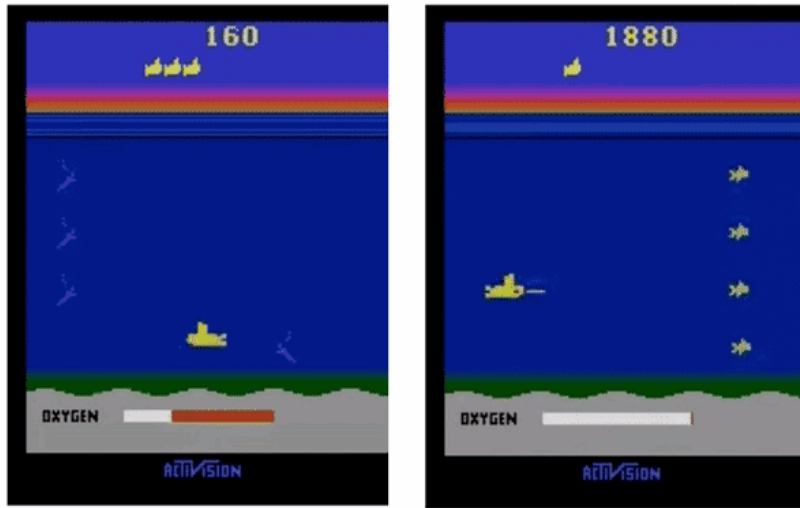
传统有限差分（梯度下降）不能跨越低适应性的窄通道，而ES能够轻松跨越它，在右侧找到一个更高的适应性。



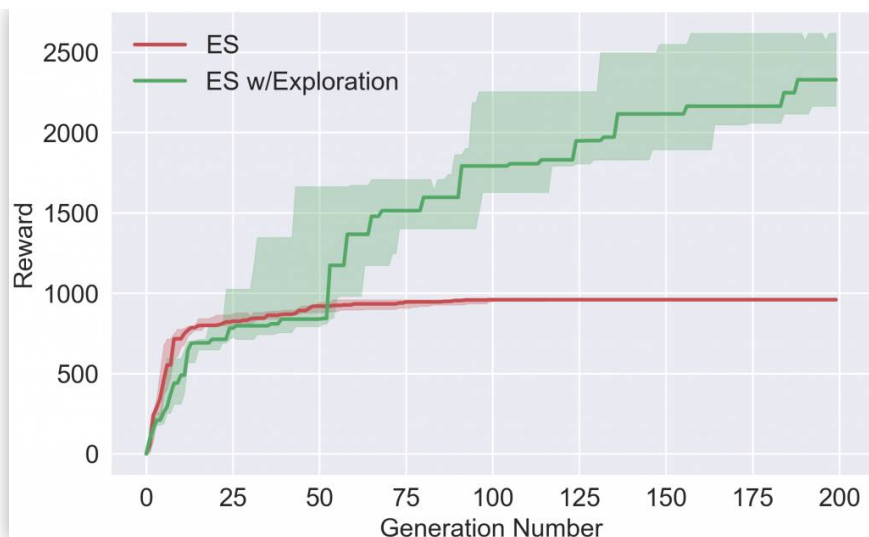
当高适应性路径通道时，ES停滞不前；而传统的有限差分（梯度下降）毫无迟疑地穿过了相同的路径。这与上面的视频一起显示出了两种方法的差异和折中。

### 改善ES中的探索（exploration）

有关深度神经进化的一个令人兴奋的结果是，之前为神经进化开发的工具集合现在已经成为加强神经网络训练的备选方法。我们通过引入一种新的方法来探索这种可能性，这种方法将ES的优化能力和可扩展性与神经进化的独特方法——通过激励不同的智能体的不同表现来促进在强化学习领域的探索。这种基于群体的探索与强化学习中的单一智能体传统不同，包括最近在深度强化学习中的探索。我们的实验显示，增加这种新的探索方式，可以在许多需要探索以避免欺骗性的局部最优域中提高ES的性能表现，包括一些Atari游戏和Mujoco模拟器中的类人动作任务。

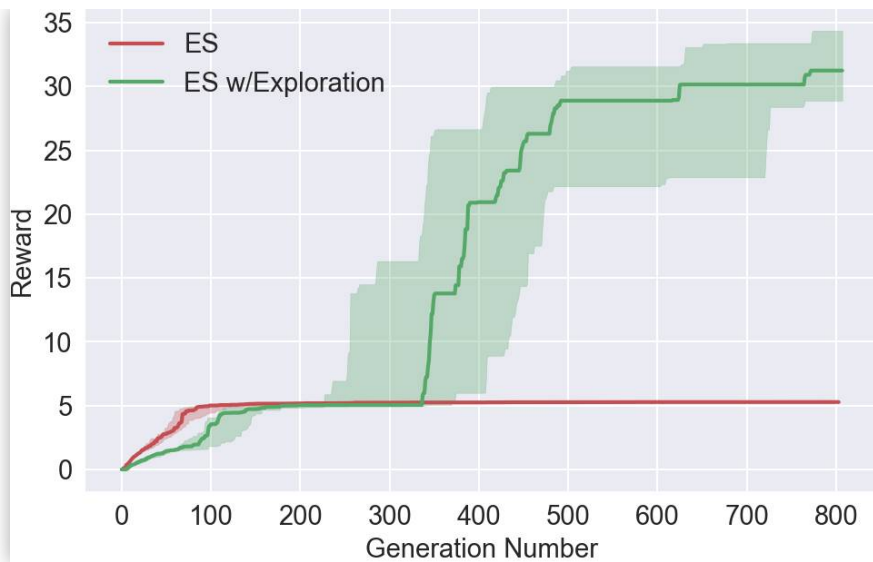


使用我们的超参数，ES很快收敛于局部最优，而没有考虑氧气（oxygen），因为这样做暂时放弃了获得奖励。然而，通过探索，智能体学会了寻找氧气，从而在未来获得更高的奖励。值得注意的是，Salimans等人在2017年没有报告，使用超参数，ES遇到了这个特别的局部优化，但是没有探索的那个点可能会被无限期地困在局部最优（而探索则能够帮助它摆脱困境），正如我们的论文所显示的那样。



智能体的任务是尽量向前跑，而ES则永远都学不会避开这个陷阱，然而，有了探索的压力，其中的一个智能体学会了在陷阱中寻找出路。

```
class="video_iframe" data-vidtype="2" allowfullscreen="" frameborder="0" data-ratio="1.875" data-w="480"
scrolling="no" data-src="http://v.qq.com/iframe/player.html?
vid=j052163o0qc&width=654&height=367.875&auto=0" style="display: block; width: 654px !important;
height: 367.875px !important;" width="654" height="367.875" data-vh="367.875" data-vw="654"
src="http://v.qq.com/iframe/player.html?vid=j052163o0qc&width=654&height=367.875&auto=0"/>
```



## 结论

总而言之，对于那些有兴趣转向深度网络的神经进化的研究人员来说，有几个重要的考虑因素：首先，这些类型的实验较之前需要更多的计算；对于这些新论文中的实验，在每次运行中，我们经常使用数百甚至数千个同步CPU。但是，对更多CPU或GPU的渴求不应被视为一种负担；从长远来看，将进化扩展到大规模并行计算中心的简单性意味着，神经进化可能最有利于利用即将到来的世界。

新的结果与以前在低维度神经进化中所观察到的结果大有不同，它们有效地推翻了多年的直觉，特别是高维度的搜索的影响上。正如在深度学习中发现的那样，考虑到某种复杂性，搜索似乎实际上在高维度方面变得更加容易，因为它不易受局部最优的影响。虽然深度学习领域熟悉这种思维方式，但它的含义才刚刚开始神经进化中被理解。

神经进化的再度出现是一个很好的例子，说明旧算法与现代计算量的结合可以很好地发挥作用。神经进化的可行性是有趣的，因为在神经进化社区开发出的许多技术能够即刻在神经网络规模上得到应用，每个技术提供不同的工具以解决那些具有挑战性的问题。此外，正如我们的论文所显示的，神经进化搜索与SGD不同，因此机器学习工具箱中提供了有趣的可替代方法。我们想知道，深度神经进化是否会像深度学习一样经历复兴。如果是这样的话，那么2017年可能就标志着这个时代的开端，而今后会发生什么，且拭目以待！

以下是我们发布的五篇论文链接，有兴趣的读者可以点击链接查看详情：

《深度神经进化：遗传算法是强化学习中用于训练深度神经网络的一种有效可替代方法》，链接：<http://eng.uber.com/wp-content/uploads/2017/12/deep-ga-arxiv.pdf>

《通过输出梯度对深度和循环神经网络进行安全突变》，链接：<http://eng.uber.com/wp-content/uploads/2017/12/arxiv-sm.pdf>

《关于OpenAI进化策略和随机梯度下降之间的关系》，链接：[http://eng.uber.com/wp-content/uploads/2017/12/ES\\_SGD.pdf](http://eng.uber.com/wp-content/uploads/2017/12/ES_SGD.pdf)

《进化策略不仅仅是一个传统的有限差分近似》，链接：<http://eng.uber.com/wp-content/uploads/2017/12/arxiv-fd.pdf>

《通过一群寻找新奇事物的智能体来改进对深度强化学习中进化策略的探索》，链接：<http://eng.uber.com/wp-content/uploads/2017/12/improving-es-arxiv.pdf>

原文来源 | Uber Engineering

作者 | Kenneth O. Stanley、Jeff Clune

## 热文精选

深度学习高手该怎样炼成？这位拿下阿里天池大赛冠军的中科院博士为你规划了一份专业成长路径

深度学习框架哪家强？MXNet称霸CNN、RNN和情感分析，TensorFlow仅擅长推断特征提取

如何成为一名全栈语音识别工程师？

Twitter大牛写给你的机器学习进阶手册



**CSDN**  
不止于代码

# 2017中国开发者 有奖大调查

全方位洞察中国开发者生态圈  
线上征集数据截止时间：2017年12月31日



扫码立即参与