

被《硅谷》选中的“压缩算法”，为什么说下一个机器学习杀手应用将由此诞生



转自：大数据文摘 | BigDataDigest

编译：jiaxu、李馨瑜、云舟

移动互联网时代，人类生产的新数据正以指数级别增长，数据中心越来越大，并消耗着地球上难以想象的巨大能耗，但人类依然可能面临着“数据无处存放”的境地。

因此，让文件变小一点，这绝对是个值得投资的点子。

神剧《硅谷》剧组就选中了这个点子。在剧中，几个主角的得以创业成功的核心算法——“魔笛”即是选中了“压缩”这一多数人都能理解但对技术要求颇高的领域：允许用户在线将图片压缩至一半的大小，并依然保持其清晰度。

“魔笛算法”也一度引发了不少跟风打造属于自己的“魔笛手”，此外，包括谷歌等人工智能巨头公司近年来也在这个领域频频发力。

2016年，谷歌相关团队推出了一款叫 RAISR (Rapid and Accurate Image Super-Resolution) 的图像压缩技术。希望以人工智能技术作为压缩路径，将图片大小压缩到源文件的 $1/4$ ，但基本不会改变图片的清晰度。



尽管引来如此高的关注，但很多人还是不认可它会在未来的发展中取得压倒性的优势。谷歌的 RAISR 算法也被不好业内认识吐槽说“压缩一张图片耗时久”、“压缩后图片清晰度没有那么多高”等等。

谷歌机器学习工程师、Jetpac 创始人 Pete Warden 近期发布了一篇文章，认为在短期内压缩算法可能会成为机器学习应用最可能出现的杀手级应用的领域，他从“压缩”这一技术的起源开始，希望告诉大家问什么压缩技术的前景是如此的广阔。

大数据文摘对这篇文章进行了编译。

OSDI 上的一篇文章“神经自适应型内容感知互联网视频传输”激发了我对这一领域的兴趣。

Neural Adaptive Content-aware Intern

Authors:

Hyunho Yeo, Youngmok Jung, Jaehong Kim, Jinwoo Shin, and Dongsu Han, *KAIST*

Abstract:

Internet video streaming has experienced tremendous growth over the last few decades. However, the bandwidth resource. Consequently, user quality of experience (QoE) suffers inevitably when video delivery framework that utilizes client computation and recent advances in deep neural networks (DNNs) for high-quality video. The use of DNNs enables us to enhance the video quality independent to the network conditions. This paper addresses several challenges, such as client heterogeneity, interaction with bitrate adaptation, and network congestion. 3G and broadband network traces shows the proposed system outperforms the current state of the art by providing the same bandwidth budget or saving 17.13% of bandwidth while providing the same user QoE.

Open Access Media

USENIX is committed to Open Access to the research presented at our events. Papers and proceedings are available for free and open to all. Any video, audio, and/or slides that are posted after the event are also free and open to all.

Presentation Audio



[Download Audio](#)

[Yeo PDF](#) [View the slides](#)

文章链接：

<https://www.usenix.org/conference/osdi18/presentation/yeo>

这篇文章告诉我们：通过神经网络，在带宽相同的条件下，用户体验将能提高 43%，或者在带宽缩减 17%的条件下可以保持体验的质量基本不变。还有其他同类的论文将类似方法运用在了生成压缩或自适应图像压缩等方面。他们都展示出了惊人的成果。那么，为何我们不把机器学习更广泛地应用在文件压缩方面呢？

我们尚未拥有性能配套的计算设备



上文所提到的所有方法都需要相对较大的神经网络，并且所需的计算量与像素数量成正比。这意味着大型图片文件或每秒帧数较高的视频文件可能需要比当前广泛使用的移动设备更强的计算能力。

一致，但也能够从原图像继承一些信息。如果你希望结果与原图像尽可能接近，你也可以从风格化方面考虑，例如为每个场景生成简图。这些技术的共同特征在于他们提取出了输入文件中对于我们人类重要的部分，并省略了余下的部分。

不仅仅是图像



语言世界也有类似的趋势。语音识别正在迅速改进，合成语音的能力也在迅速提高。识别可以看作是将音频压缩为自然语言文本的过程，而合成则恰好相反。你可以想象将对话高度压缩为自然语言文字而非音频。诚然，我们不需要达到那种要求，但似乎通过扩展我们对于表达的认知类型，我们可以实现在低带宽情况下的更好的结果。

我甚至还发现一些有趣的将机器学习应用在文本本身上的可能性。Andrej Karpathy 的 CHAR-RNN（递归神经网络）展示了神经网络对某些模板进行模仿的优越性，而对于压缩技术来说，预测也是相同的道理。考虑到普通的 HTML 页面有不少冗余，这一情况下 GZIP 压缩方法似乎有很大机率能被机器学习改进，尽管我认为这没有机器学习对文本的处理那么容易。

对于压缩的研究已立项拨款



在我创业期间，营销失败的经历让我学到了一件事：如果已经有一笔预算要购买你所推销的产品，该产品的销售会容易得多。预算的存在意味着公司在是否应该花钱购买解决方案上已经赢得了艰难的斗争，现在唯一的问题是购买哪种解决方案。

这也是我认为机器学习可以在这个领域取得巨大突破的原因之一，因为制造商已经拥有专门用于视频和音频压缩的工程师、资金和芯片。如果我们能够证明将机器学习添加到现有解决方案中能够以一些可度量的方式（例如质量、速度或功耗）改进它们，那么它们将被快速采用。

带宽将会花费用户和运营商的大量的成本，质量和电池寿命将会成为产品的卖点，因此采用机器学习改进压缩的优势将比其他运用更为显而易见。

许多激动人心的研究表明这种方式将会非常高效，而且我乐观地认为还有许多应用尚未被发掘。我希望机器学习在压缩中的应用将发展成为一门核心技术。