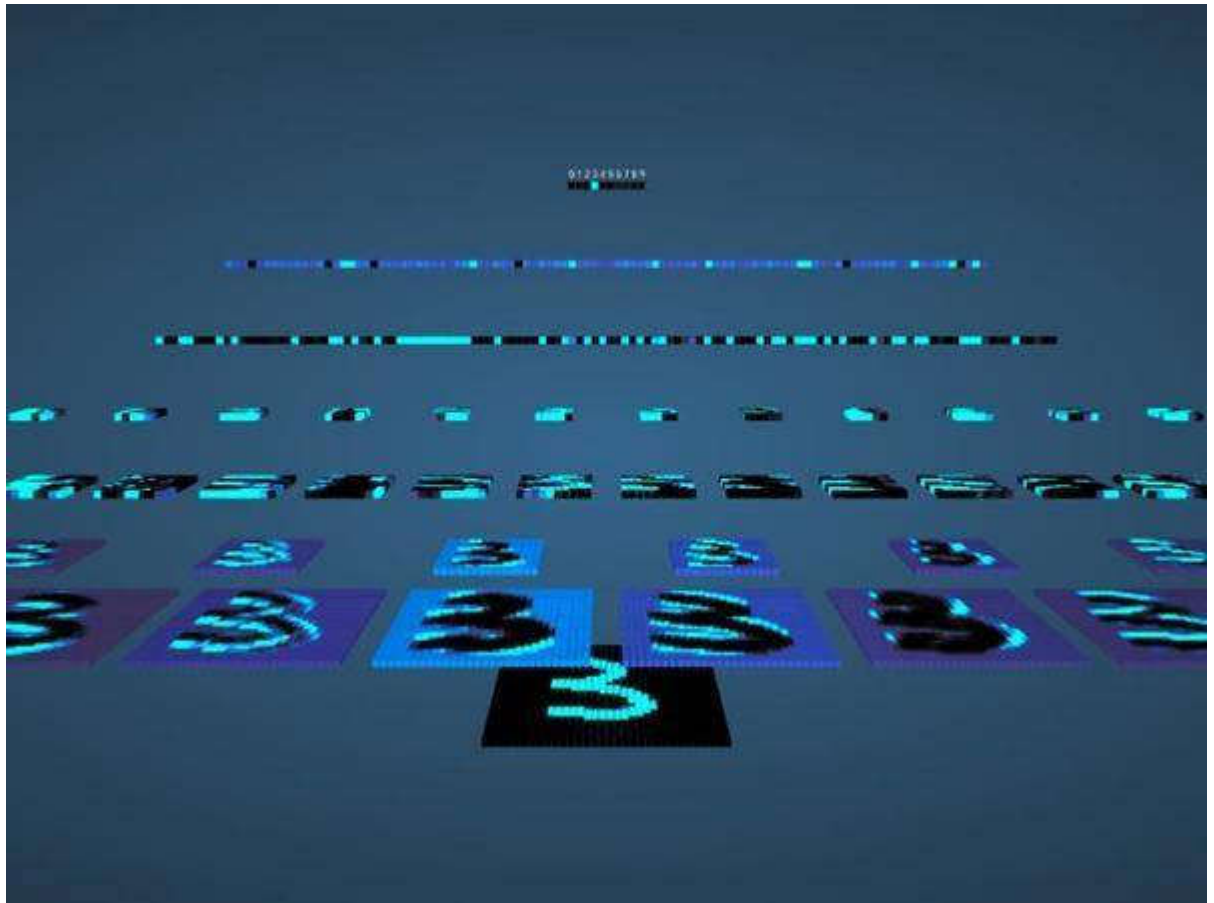


人工智能玄乎其玄，内部究竟如何？带你一窥卷积神经网络内部



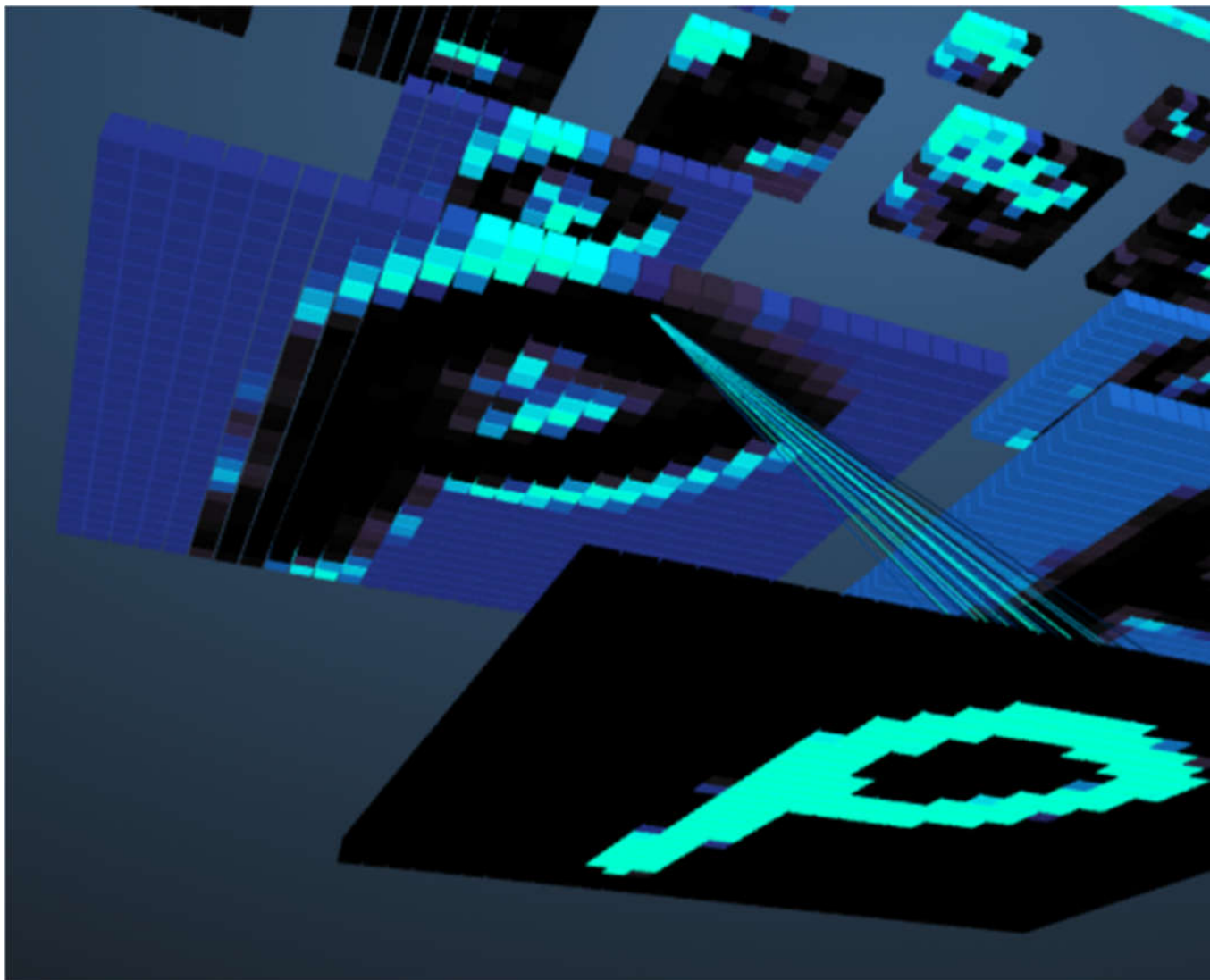
本文选自 POPSCI，作者 Dave Gershgorin，机器之心编译出品，编译：柒柒。

Adam Harley 的卷积神经网络 3D 视觉化模型能够让你一窥人工智能的内部工作。

除了邪恶 AI 将会统治地球这样的惊悚事件之外，人工智能领域让局外人感到害怕。Facebook 的人工智能负责人 Yann LeCun 这样比喻人工智能：人工智能就是个有百万个结点的黑盒子，其内部工作原理对大部分人而言仍然是一个谜题，但现在，我们有机会一窥究竟。

瑞尔森大学的硕士候选人 Adam Harley 创建了一个交互式视觉化模型，能够帮助解释卷积神经网络（一种用来分析图片的人工智能程序）内部是如何工作的。

正如在该交互式视觉化模型中可看到的，神经网络按顺序分层进行工作。底部是输入，即计算机试图进行解释的原始内容——在该情况中，就是你写下的数字——而顶层是输出，即计算机的最终结论。**两者中间是数学功能层，每一层都将最重要的识别信息进行压缩，并传导至下一层。**



卷积神经网络读取图片识别数字的视觉化效果

输入端（最底部的一排）的绿色像素对应于你写下的数字，这些黑色像素代表着背景色，而数字就需要从这些背景中识别出来。如果它试图检测一张脸，那么 3 就代表这张脸，黑色就代表着相片的背景。在每一阶段，我们都将看到这些图片在经历每一步后会变成什么样，而不是看到步骤本身。

在神经网络中，前几层主要关注边缘和形状之类的情况，会抽取出整体的视觉效果，找出不同的识别特征，将之提取出来以便找出能够将图案与周边环境区分出来的特征。

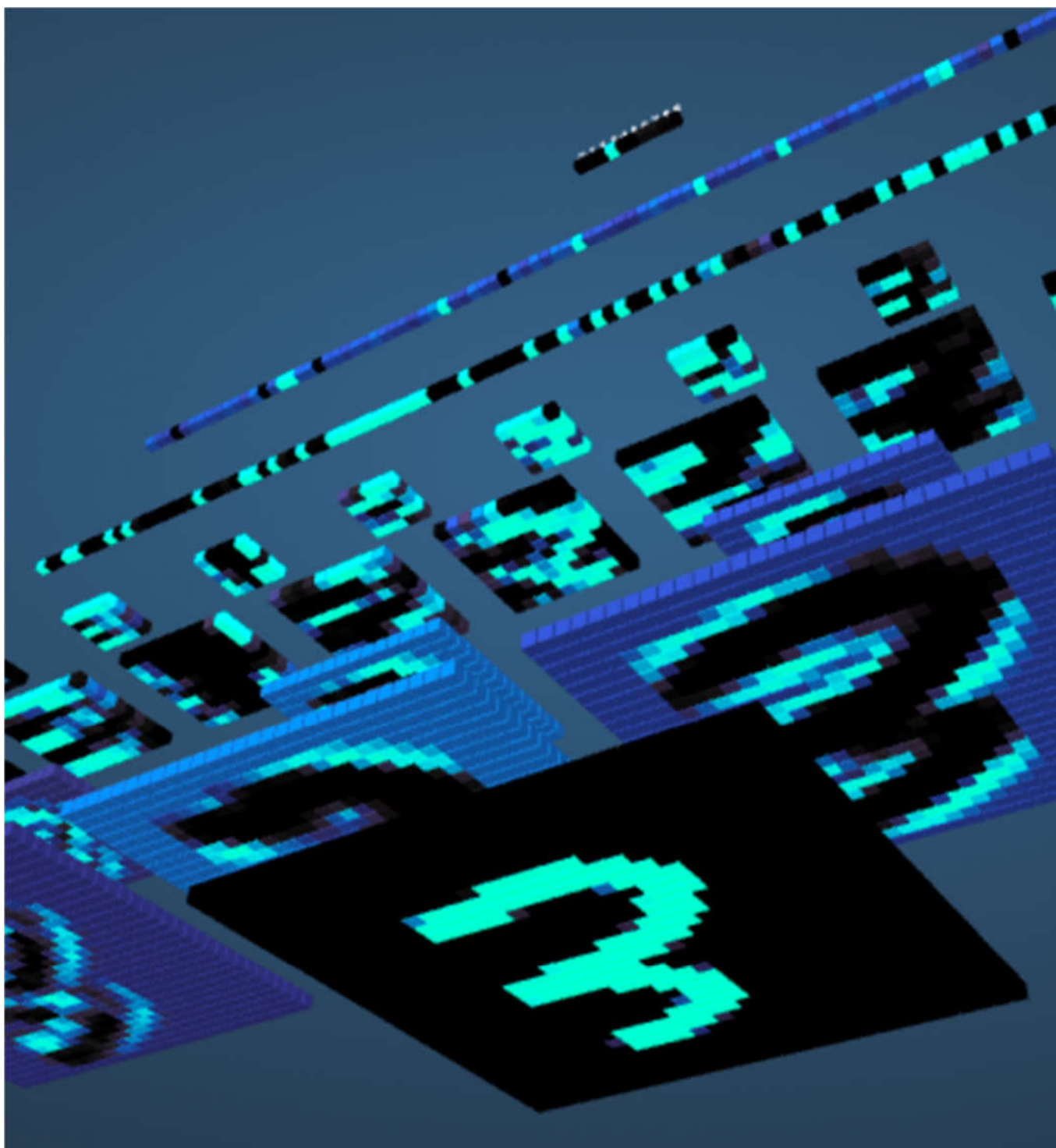
这些层的每一层都会预先识别这些数据，也就是称之为训练的过程。训练通常意味着在机器上运行不是百万也有成百上千的样例图片来告诉机器不同类型的 3 是什么样子。这一过程对所有类型的机器学习和人工智能都一样，只是源材料不同。谷歌会采取使用其服务的人们的随机声音样本来训练其声音识别软件，而脸书会用人不同角度的图片训练其面部识别算法。

训练需要在机器上运行几百万个样本。

经过第一层传输的数据会被第二层进行简化（称之为缩减采样层，因为这一层会减少数据的复杂度）。之后第三层会再次分析其形状，这层就类似第一层的卷积层。该神经网络有两层卷积层，而有些更复杂的网络会有 10 层以上的卷积层。

这组形状和边缘之后会经过处理，并与一组预定的输出进行匹配，最终得出结论有很大的可能使用者画的是个 3（或者可能是个 8）。你能看

到数据通过层层处理时的颜色区别。你画的绿色数字最终是绿色的比特信息，并（有很大希望）指出了正确的输出。



卷积神经网络内部

在 **Harley** 的模型中，计算机能够简单地识别出数字，非常像是原始的卷积神经网络被用来在 **ATM** 机上识别支票存款。最前沿的人工智能要复杂得多，能够以 **97%** 的正确率识别人脸。