

# 深度学习如何助力自动驾驶？

自动驾驶是人工智能的典型应用场景。人在驾驶时容易疲惫、分心、斗气，出现事故时反应速度也不够，让机器驾驶不但让人减轻压力，同时也会更安全。

如何实现自动驾驶呢？说起来并不复杂，给汽车装上各种传感器（如图 1 所示），让他可以感知道路和周围的环境，将这些信息送入一个中央控制器，由控制器分析应该实施的驾驶操作（如转向、加油、刹车等），即可实现自动驾驶。

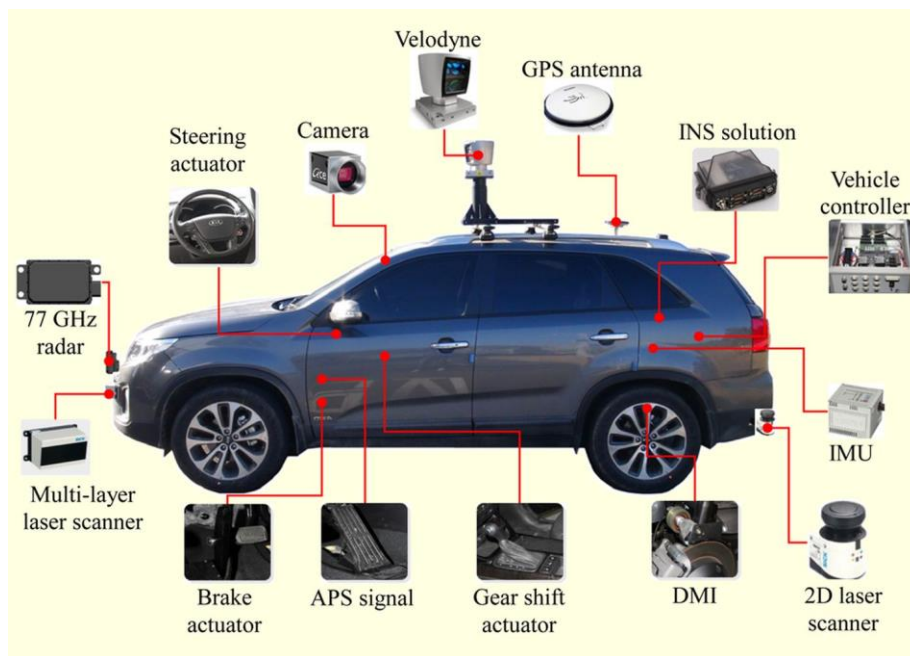


图 1：自动驾驶汽车身上的传感器

虽然概念很简单，但在实施起来并不容易。首先，对各个传感器得到的信号进行分析和理解本身就是个复杂的问题；第二，这些信号要集成起来，形成对车身所处位置、环境和运动状态的整体理解也并不容易；第三，需要基于当前驾驶状态做出相应决策，特别是当出现突发状况时的紧急动作，这又是个挺困难的的任务。更糟糕的是，上述感知-理解-决策过程中的任何一个环节都可能出错，如果错误累积起来，有可能产生严重后果。

另一种方案是把传感器的信号送入一个神经网络，让网络自动生成驾驶动作。这一方案里，神经网络相当于一个“老司机”，看到路况反射式地做出应对，用不着仔细分析，行为成自然。

这一方案早在上个世纪 80 年代就有人提出来[2]，然而直到深度学习兴起以后，特别是大规模数据库的出现，这一方法才取得较好的结果[3, 4]。图 2 是 Nvidia 研究者发布的一个“端到端”自动驾驶系统，该系统接入三个摄像头的信号输入，通过一个卷积神经网络（CNN）直接生成驾驶操作。网络训练采用人类驾驶员的操作数据，并利用了数据增强技术以提高系统的可扩展性。



图 2: Nvidia 的自动驾驶网络[4]

研究者发现，虽然神经网络直接输出的是驾驶动作，但在其内部确实对图像进行了某种理解。图 3 给出了在乡间道路上驾驶时，CNN 内部对图片的分析。由第一层和第二层的卷积核输出结果可以看到，CNN 网络试图在理解图片中的道路，并基于此在更高层做出驾驶决策。

这种端到端方法的好处是系统简洁，但模型训练较为困难，有可能出现不可理解的驾驶决策。值得一提的是，不论是哪种自动驾驶技术，在复杂场景下都可能引发风险，因此，在技术没有完全成熟之前，开辟专用车道也许比技术本身更重要。

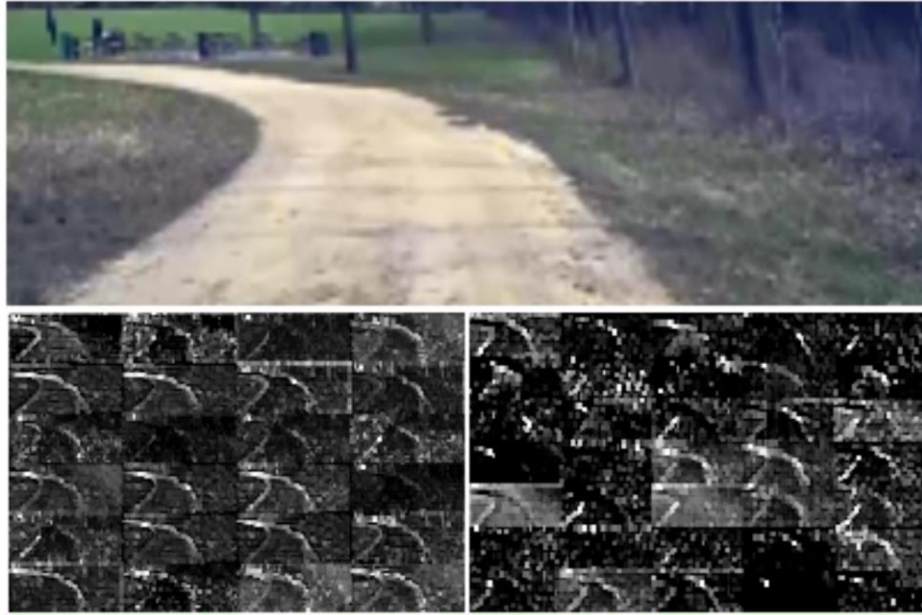


图 3: 卷积神经网络可以自动检测道路。上图是输入图像, 左下角是 CNN 第一层的输出, 右下角是 CNN 第二层的输出。

- [1] 自动驾驶汽车涉及哪些技术? <https://www.zhihu.com/question/24506695>
- [2] D. A. Pomerleau. Alvin: An autonomous land vehicle in a neural network. Technical report, DTIC Document, 1989.
- [3] Chen, C.; Seff, A.; Kornhauser, A.; Xiao, J. (2015). DeepDriving: Learning Affordance for Direct Perception in Autonomous Driving. ICCV. pp. 2722-2730.
- [4] Bojarski, M. et al. (2016). End to End Learning for Self-Driving Cars. arXiv:1604.07316.