

自动驾驶学习笔记（6） | 人工智能：自动驾驶强力助推器

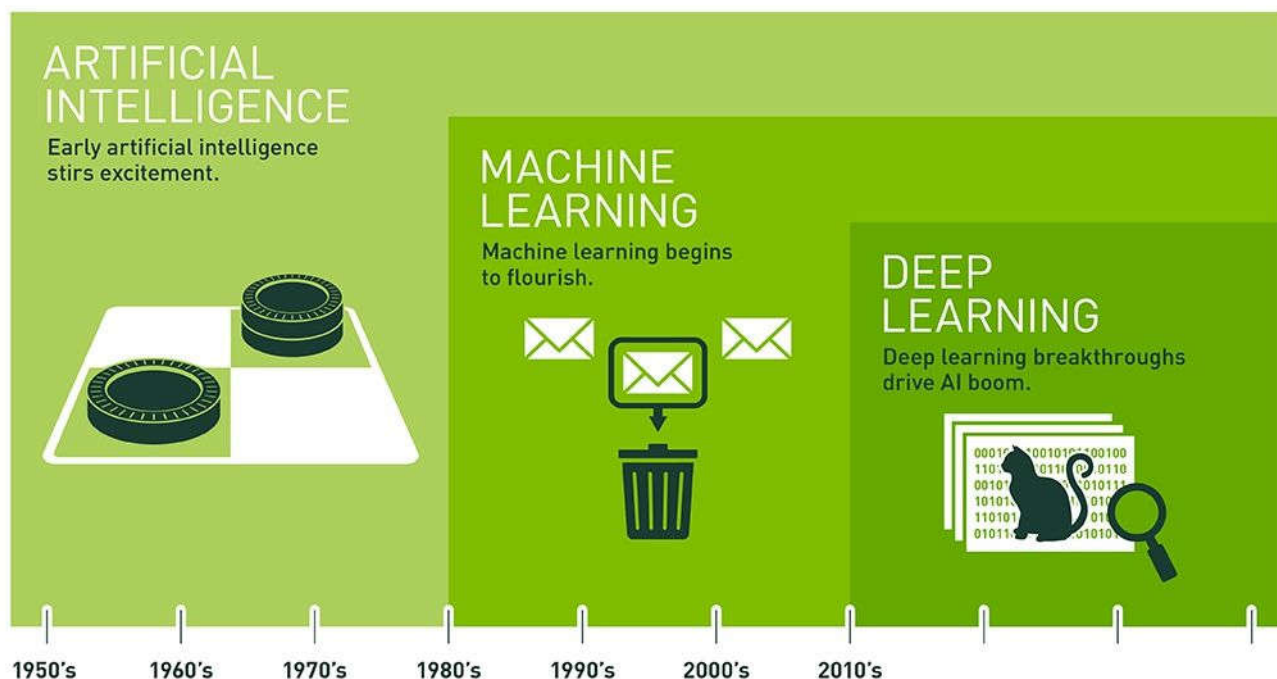
引言

如果没有人工智能加持，自动驾驶难以取得新的进展和突破。由于人工智能是一个很宽泛的主题，在自动驾驶领域，目前应用最广泛的人工智能技术是基于视觉的深度学习技术，特别是卷积神经网络（Convolutional Neural Network）用于实时目标检测和跟踪。此外，人工智能在自动驾驶决策、规划、控制等方面的研究也方兴未艾。

可以说时至今日，在汽车行业从事自动驾驶相关岗位，没听过深度学习都不好意思和同行打招呼。尽管自动驾驶是一个系统工程，不是所有环节都需要研究深度学习算法，但是了解一些基本概念还是大有裨益的。

深度学习

Learning) 和深度学习 (Deep Learning) 是依次包含的关系：人工智能范围最大，核心问题包括感知、推理、规划、学习、交流等。机器学习是实现人工智能的途径之一，与**概率统计学**有着密切的联系。深度学习是机器学习的一种技术方式，它使用独特的仿生神经网络进行训练和决策，是目前机器学习最热门的研究领域。



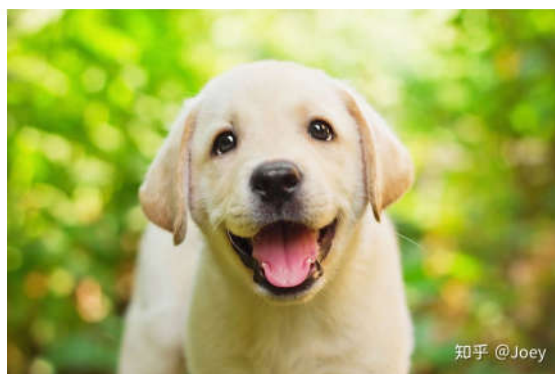
Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.

知乎 @Joey

图片来源：<https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/>

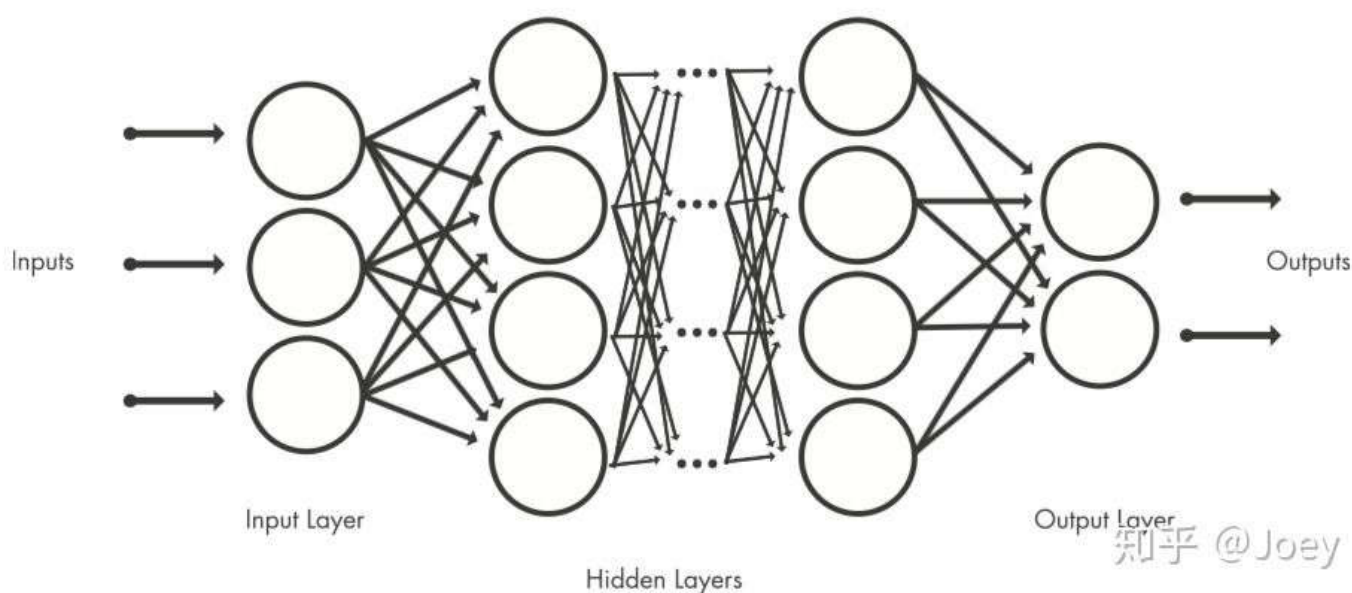
机器学习和深度学习的本质是通过大量数据的训练找到一个“函数”，然后使用这个函数去根据特定输入（文本、语音、图像等）得到输出。对于深度学习来说，这个函数的形式是神经网络。

设想一下，如何编程让计算机能识别出图片中的小狗？



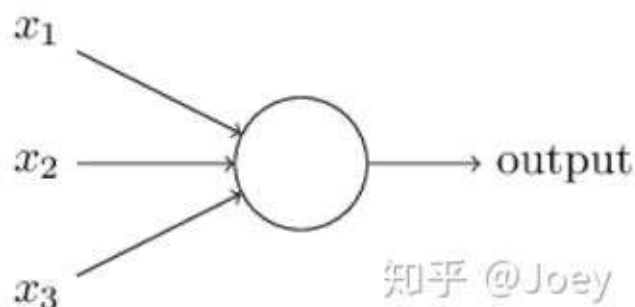
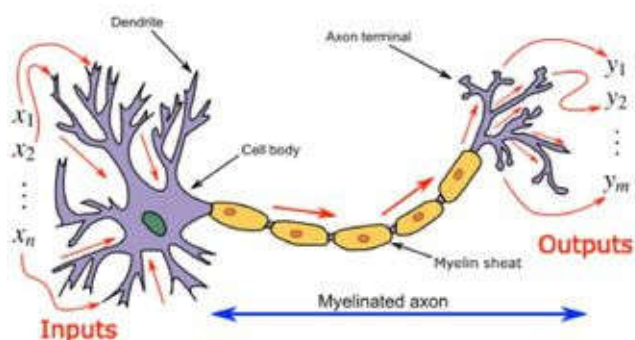
通常的基于规则（Rule Based）的编程无法完成这个任务，小狗可以有各种各样的品种、姿态、特征、所处环境等，没有一套普遍适用的规则可以识别所有的小狗。而基于神经网络的深度学习出发点是试图让计算机通过大量数据的学习能够拥有人类特有的“直觉”。

深度神经网络结合多个非线性处理层，并行使用简单元素操作。它由一个输入层、多个隐藏层和一个输出层组成。各层通过节点或神经元相互连接，每个隐藏层使用前一层的输出作为其输入。隐藏层可达上百层，因此得名“深度”学习。



图片来源：https://www.mathworks.com/content/dam/mathworks/tag-team/Objects/d/80879v00_Deep_Learning_ebook.pdf

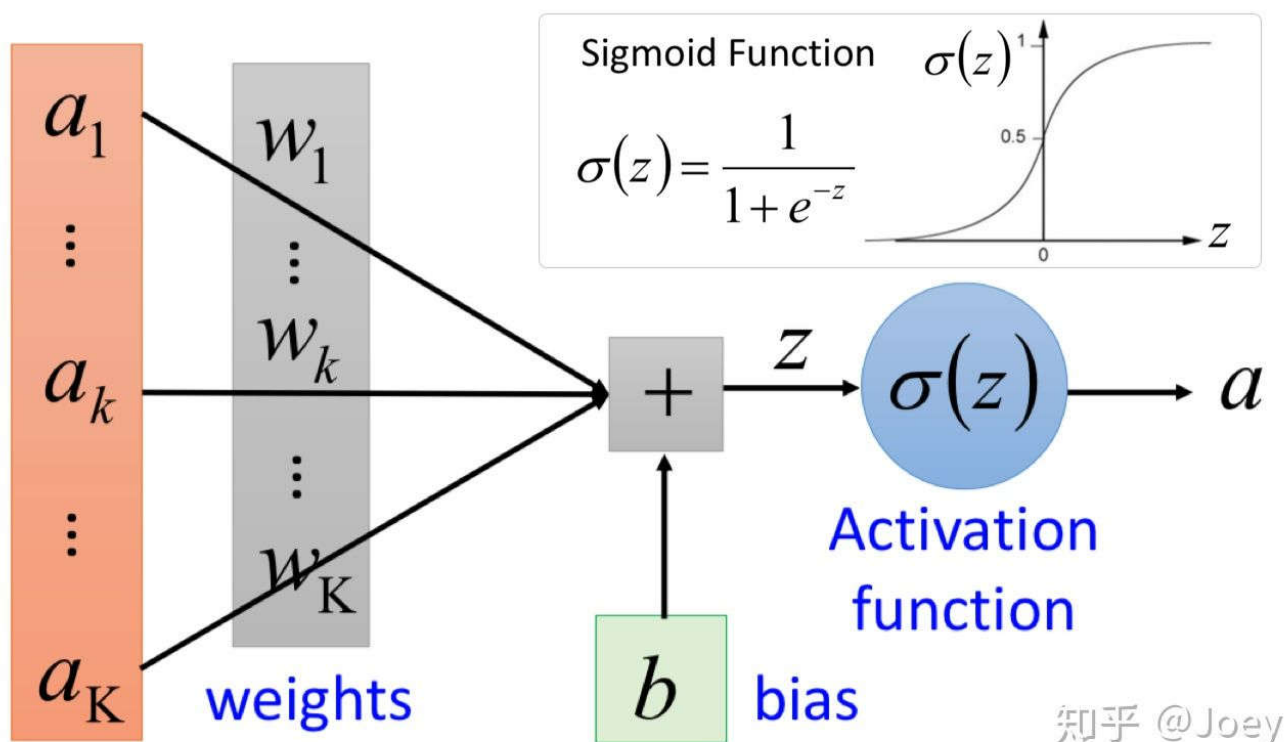
网络中节点神经元 (Neuron) 的一种数学表示是感知机 (Perceptron) , 它有若干个输入和一个二进制输出 (0或1) , 是神经网络的基本单元。



神经元和感知机

另外一种常用的神经元模型是Sigmoid Neuron, 它的数学模型包含权重 (weight) 和偏置 (bias) 两类参数和一个激活函数 (Activation function) 。与感知机模型不同的是, Sigmoid Neuron的输出可以是0到1之间连续变化的任意实数, 这样的好处是避免了信号传递过程中难以控制的跳变, 神经网络参数的微小调整使得输出的变化也会是较小的, 神经网络的学习会更加稳定。

$$z = a_1 w_1 + \cdots + a_k w_k + \cdots + a_K w_K + b$$

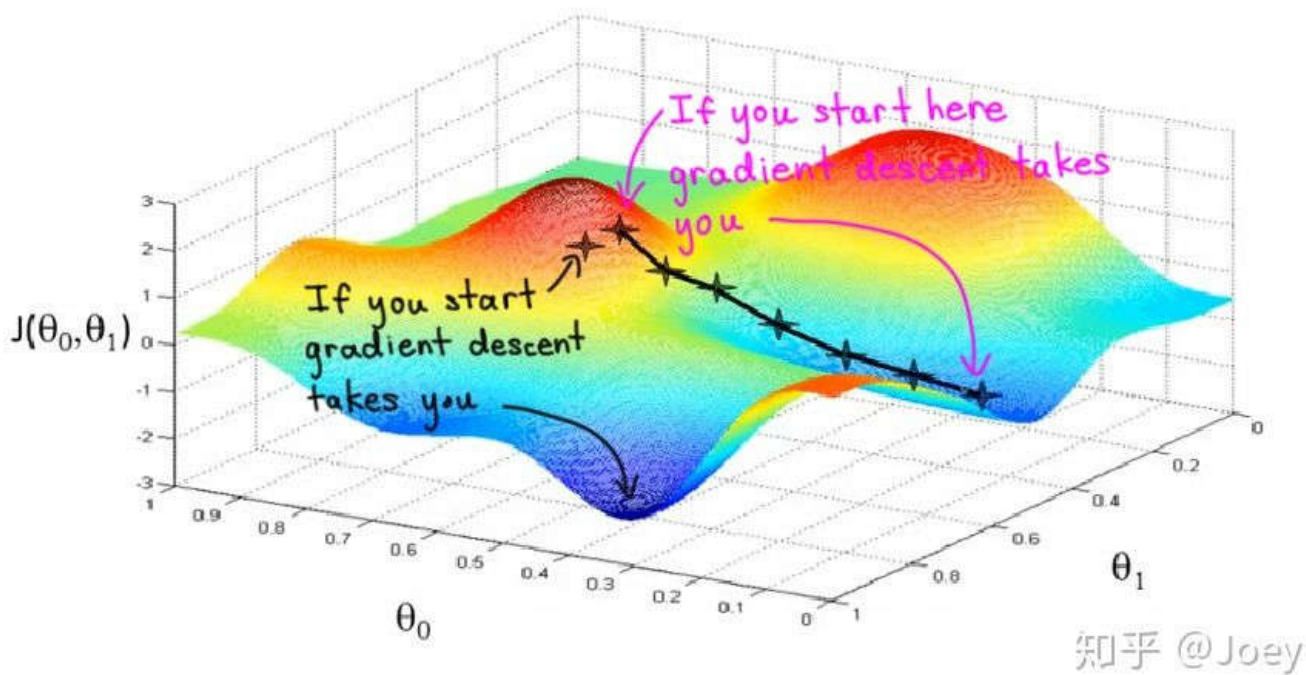


图片来源：李宏毅Deep Learning Tutorial

这些简单的神经元模型就像0和1对于计算机，与非门（NAND Gate）对于电路一样重要，是构筑深度神经网络的基石。

可以从理论上证明单隐藏层的神经网络就可以表达所有的函数，因此神经网络具有强大的普适性。层数越多，神经网络可以拟合的函数就更加复杂和精确。

设计一个深度学习系统首先需要设计一个神经网络架构（多少输入输出变量？多少层？每层多少神经元？），然后根据应用场景使用训练数据集（Training Dataset）确定网络参数（每个神经元连接权重、偏置），目前优化参数最常用的数学方法是梯度下降法（Gradient Descent）。**梯度下降的迭代过程其实是在寻找损失函数（Loss Function）最小化最优解的过程，优化神经网络参数的过程，也就是学习的过程。**



梯度下降

深度学习神经网络的形式多种，常见的有循环神经网络（Recurrent Neural Network:

RNN）和卷积神经网络（Convolutional Neural Network: CNN）。循环神经网络带记忆功能，常用于自然语言处理、语音识别、机器翻译等。卷积神经网络常用于数字图像处理，在自动驾驶中有着重要的作用。

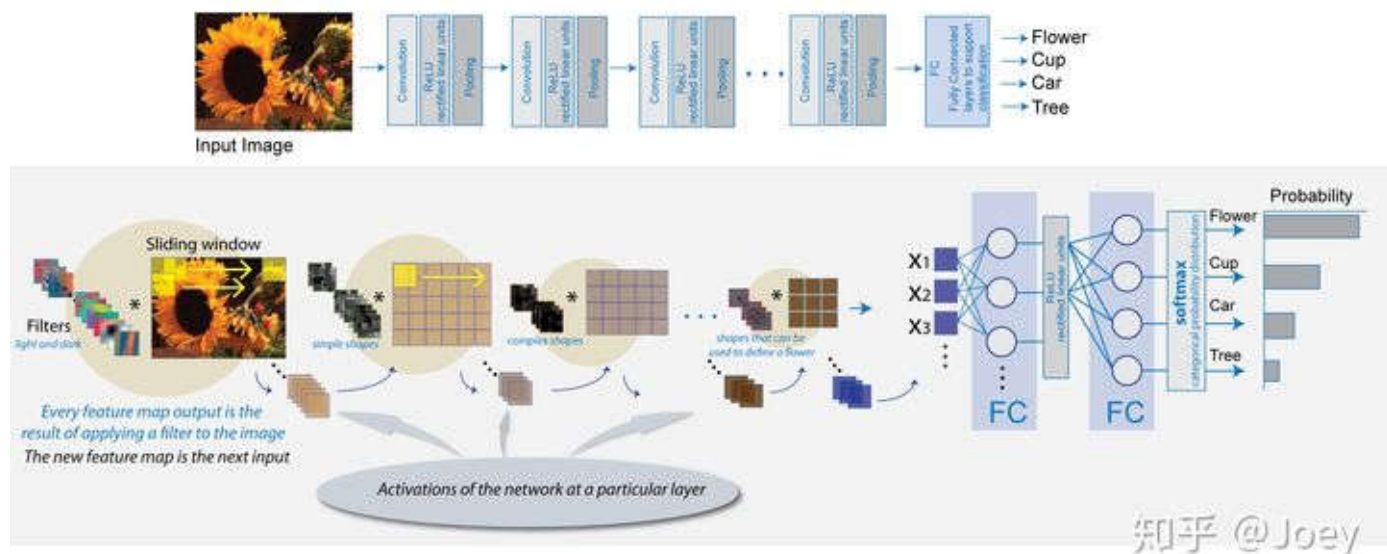
下面简单介绍卷积神经网络和应用举例。

卷积神经网络

卷积神经网络是图像和视频深度学习的最流行算法之一。自动驾驶摄像头使用卷积神经网络的训练后进行道路、行人、车辆等目标的识别。

像其他神经网络一样 卷积神经网络中一个输入层 一个输出层和中间的多个

线性单元 (ReLU)、池化 (pooling) 和完全连接层 (Fully Connected Layer)，最终得到输出，也就是得到图像的分类以及对应的概率。



图片来源:

<https://www.mathworks.com/help/deeplearning/ug/introduction-to-convolutional-neural-networks.html?>

searchHighlight=convolutional%20neural%20networks&&am
p&s tid=doc srchtitle

卷积将输入图像放进一组卷积过滤器，每个过滤器激活图像中的某些特征。修正线性单元通过将负值映射到零和保持正数值，实现更快、更高效的训练。池化通过执行非线性下采样（downsampling），减少网络需要学习的参数个数，从而简化输出。完全连接层观察上一层的输出（其表示了更高级特征的激活映射）并确定这些特征与哪一分类最为吻合。例如，如果该程序预测某一图像的内容为狗，那么激活映射中的高数值便会代表一些爪子或四条腿之类的高级特征。同样地，如果程序测定某一图片的内容为鸟，激活映射中的高数值便会代表诸如翅膀或鸟喙之类的高级特征。大体上来说，完全连接层观察高级特征和哪一分类最为吻合和拥有怎样的特定权重，因此当计算出权重与先前层之间的点积后，你将得到不同分类的正确概率。

这几种操作在几十层或几百层上反复进行，每一层都学习检测不同的特征。随着网络的层数增加，每一层对于前一层次的抽象表示更深入。

通过可视化工具，可以窥见卷积神经网络内部节点通过学习后都检测哪些特征，这对于理解神经网络的工作方式以及如何提高设计的网络的性能有有很大的帮助。

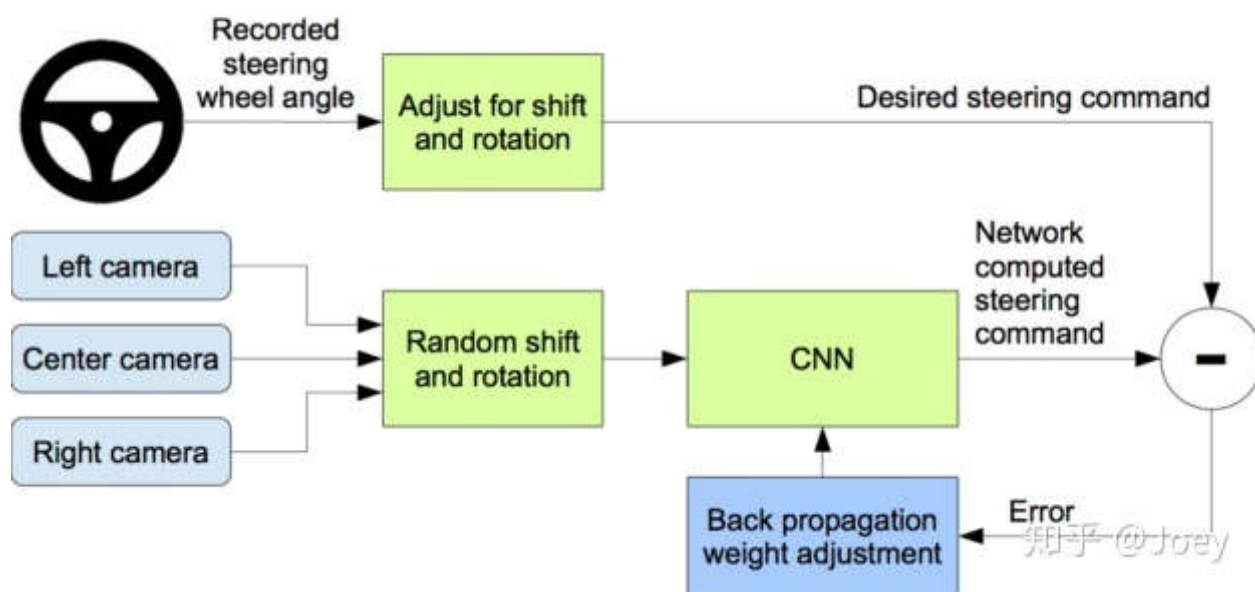
arxiv.org/pdf/1311.2901... (Visualizing and Understanding Convolutional Networks)

应用举例：端到端控制End-to-End

一般来说，自动驾驶系统需要感知、决策、控制严密的系统逻辑设计，而端到端控制直接从感知通过神经网络映射至执行机构控制。

根据NVIDIA公开的论文，该公司设计了一套基于深度学习的端到端自动驾驶控制系统PilotNet，只需要通过驾驶数据（摄像头数据+驾驶员转向控制数据）的训练，卷积神经网络即可学会根据前视摄像头实时图像得出合适的转向控制。

PilotNet训练方法和一般深度学习训练方法无异：驾驶员转向指令的数据通过反向传播调整卷积神经网络的权重参数。



PilotNet训练方法

完成训练后，在测试中神经网络就能够根据摄像头的图像生成转向指令。测试路径不一定需要清晰的车道标线，也没有设计规则让系统识别车道标线，经过数据训练的系统会选择出神经网络认为重要的“道路特征”（如道路轮廓、道路边缘、车辆等）并根据“道路特征”确定转向控制量。也就是说尽管没有明确告诉系统道路边界对于自动驾驶必不可少，神经网络通过学习后会自动在测试中探测道路特征，这是深度学习神奇的地方。



PilotNet根据摄像头输入的图像输出转向指令

由于端到端内在逻辑较抽象，目前仅限特定道路的试验验证，还需要从理论上论证该方法的可行性和鲁棒性，探索debug的方法，并进行更大范围的公开道路测试。

本篇初步介绍了深度学习的概念和深度学习在自动驾驶中的应用，从中可以看出深度学习的强大能力。本篇的不足之处是略去了很多数学推导，如果想要深入理解深度学习，基本的数学推导论证是需要耐心看下来的。

除了深度学习，业界还有强化学习（Reinforce Learning）、迁移学习（Transfer Learning）等新理念，具体内容请读者们查阅参考资料和论文。

愿大家都能从学习中获得乐趣！

[1] <https://www.mathworks.com>

[2] neuralnetworksanddeeplearning.com

[3] cnblogs.com/subconsciou...

[4] mp.weixin.qq.com/s?...

[5] devblogs.nvidia.com/cat...

[6] <http://cs231n.stanford.edu>