

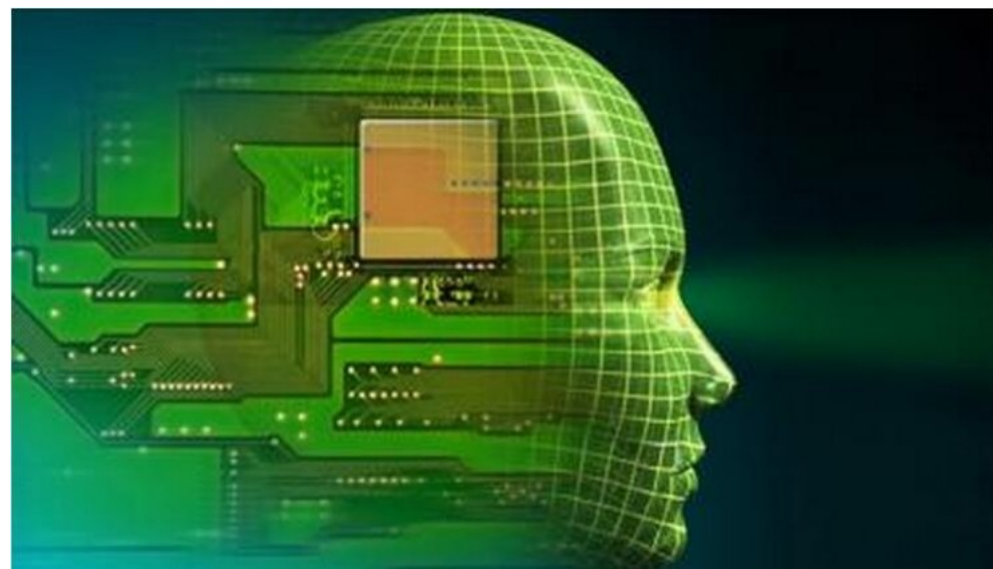
深度：如何将深度学习和自动驾驶结合在一起？

本文作者：李星宇

2016-06-25 08:56

导语：为什么要自己做深度学习芯片？本文会告诉你答案。


雷锋网按：本文整理自地李星宇在《人工智能和自动驾驶技术论坛》上的演讲。李星宇，现地平线机器人汽车业务总监，前飞思卡尔应用处理器汽车业务市场经理，原士兰微电子安全技术专家。



via : [wrox.cn](#)

以下是演讲内容的整理：

我是来自于底线线机器人汽车业务的负责人李星宇，地平线机器人成立有一年的时间，创始人余凯是百度深度学习研究院的院长，也算是国内第一个深度学习的研究院是他建立的。他成立了地平线这家公司，专注于深度学习技术的开发，所以在今天我想跟大家分享的主题就是：《如何将深度学习和自动驾驶结合在一起？》。



李星宇

专栏作者

地平线汽车业务总监，前飞思卡尔应用处理器汽车业务市场经理，原士兰微电子安全技术专家。

发私信

当月热门文章

最新文章

NewGen 张璐：硅谷看好的热点技术

iPhone 7 Plus的这颗长焦镜头将改变人工智能的消费生态

控制权的争夺战即将开始

如果没有“3”主定是场剧”

为何大数据公司很多，AI公司很少？

刷月饼事件是规则问题，与你观没什么关系

热门搜索

AR

平板电脑

中国移动

OPPO

宅客

NASA

GoF



- Galaxy S6
- 新能源
- Fitbit
- 张小龙

背景：复杂的驾驶环境正是深度学习的机会

应该说在今年的AlphaGo和李世石的对决当中，大家对于机器人的关注已经到了一个非常高的水准，其实自动学习已经有不少的时间。今年丰田有做过一个Demo的系统，他做了是8辆车并没有被教任何的驾驶规则，就是随意往前开，这8辆车会置于一个训练系统之下，经过八个小时的训练之后，八个小时下来没有碰撞，这真是了不起的事情。

通过资本界我们来看，这些10亿美元的投入都能真金白银地证明整个业界对于深度学习和增强学习技术的看好。

为什么我们需要把这样一个技术运用在自动驾驶领域？千言万语可以归结为一句话就是**复杂性**。我们看到很多人在很多时候会质疑自动驾驶这件事情，其实他的质疑觉得就是复杂，你可以看到在十字路口，非常复杂的路况。包括中国进入汽车社会不久，驾驶技术有待提高的特殊的国情，变道非常频繁，拐弯很多时候也很暴力。**这样一个场景其实对于传统的ADAS有很大的挑战性，而这种复杂性恰恰是深度学习的一个优势。**

当然很多人也会问，在自动驾驶领域是不是人就没有作用了？我们很长一段时间在自动驾驶领域，人和车的关系其实并不是一个简单的服务与被服务的关系，而是人和马之间的关系。意味着是什么？其实马并不知道在一个大的方向该怎么走，是快一点还是慢一点，这个是需要人来控制和调节。而马能做的就是，无论你告诉还是不告诉，前面是悬崖就会停下来，意味着自动驾驶可以在微观的一个层面，能够把风险降到极低的水平。而在高层次的体验层面，更多的是以跟人配合的关系，它会去理解你。**所以在自动驾驶领域很重要的一个话题就是对于驾驶习惯的学习。**自动驾驶习惯的学习不仅仅是一个标准的驾驶，还包括特殊驾驶风格的学习。

这是一个非常标准的美国高速公路管理局的对于自动驾驶的分析。应该说，从我们跟车厂的沟通里面可以看到，现在研发的重点还是基于Level3，就是高度的无人驾驶。它的主要不同跟Level4的不同在于，**一个就是保证在任何情况下都可以做紧急碰撞的防止，第二个就是在条件良好的道路上部分实现自动驾驶。**

从现在实践的情况来看，高速公路上应该不是有太大的问题。现在主要解决的是高速公路上的一些特殊的案例和条件比较好的城区主干上的案例，比如上海的中环、内环、外环的情况下，尤其红绿灯情况下需要一些特殊案例的深度学习。

深度学习为什么受到重视？

我们介绍了一个背景之后，稍微回来谈一下深度学习为什么会受到重视。大家现在都觉得“深度学习”这个词特别火热。

• 第一个为什么它会受到重视？

应该说深度学习出现的历史其实非常久，最久的历史可以追溯到50年代末，在漫长的发展时期里面是几经起伏，有它的高潮的时候也有低谷的时代。从现在的情况来看深度学习最具吸引力的一点**就是端到端的学习**。

举例，以自动学习的系统去识别一个图片，比如这个图片上是一个人，它会把图片作为输入一个名字，你可以看到人脸的信息量通常是几兆，几百万个字节，最终输出的就是几个字节，就是端到端，可以把非常复杂的非结构化的数字转化为精简数字化的表达，包括语音、语言、图像、视频都是非结构化的，包括金融的交易产生的数据都是非结构化的。

从目前的情况来看，**第一个它非常适合大数据时代**。在过去的年代，其实大家也是在研究深度学习的算法，那个时候还不叫深度学习，叫深度神经网络。但是发现很难把它调校得好，因为那个时候学的都是一个小的，几千个样本，想把它调到一个非常好的水准，其实是不可能的。后来大数据出来以后，这个性能就飙升，最终的一个例子就是2012年的时候爱莱斯和他的老师参加了图片分类识别的比赛，一下子从过去的74%提高到了85%，非常大的一个进步。也就是说现在机器对于图片的识别其实是优于人眼，所以实践效果很好。比较通俗的说法，对于这种行为的一种模拟，这种说法可能不是特别的准确，事实上跟它相关联的网络，按它的类似性来说是非常复杂的，而且现在没有搞得特别复杂。

机器学习技术前沿，注意力模型。图片上是对于人脸识别过程的一个描述，如果你抛开技术不看，你直接看，你会看到图像识别的时候，每过一个识别阶段，由很多个识别阶段组成。它的图像的信息会减少一些，最开始会把图像的彩色性去掉，只留下一个归录信息。

第二步会把**视录中间的去掉，只留下边缘线条的信息**，然后线条会变成点，这是从低位空间向高位空间不断的扭曲的过程，到最后它的图片信息就会精简到一个字母，所以这就叫多层的卷积神经网络。大家都知道这样一个复杂的关系势必会带来数据极大的产生，因为肯定需要非常大的参数。

自动学习技术也在不断地往前推进，我们现在看到各个领域都全面推进。

比如注意力模型，什么叫注意力模型呢？举例，在鸡尾酒会上看到很多的人，但是你只跟其中一个或者几个人谈，这个时候你的听觉系统只会注意到你想关注到那个人的话，而把其他人的给屏蔽掉。解决了在复杂的语义环境下，会对特别的语音进行聚焦。当我们对这个妇女说，我要求你把飞盘作为聚焦目标的时候，机器系统可以把飞盘反射出来，也就证明它准确能够聚焦到你想要它聚焦的部分，这个是非常关键的，因为这个系统是非常复杂的，而你只需要自动驾驶的系统关注车辆、车道线、人这些关键的信息，这样一个技术可以使得它很有效的处理。

除了注意力模型以外还有其他的进展，包括长时短时的记忆力模型，这个是做什么用？它会使深度学习，不仅仅是在空间，二维平面上有足够的学习，它能记住过去发生的事情。刚才丰田的小车的Demo里面可以看到，它对瞬间场景的处理做得很好，**但是有没有想过如果场景的处理手法依赖于过去的历史怎么办？**比如我们玩《魔兽》和《星际争霸》，下一步的动作取决于你的历史状态，这个时候要求你的深度神经网络需要足够的记忆力，能够记住过去的状态，这个记忆力模型就是解决这个问题包括序列化等等。

• 增强学习

什么是增强学习？

打一个比较形象的比方，如果过去的系统我们当做是计划经济的话，那么增强学习就是市场经济，可以理解它是一种结果导向的技术。准确描述这是一个智能主体跟环境不断博弈而来优化的一个过程。比如刚才的丰田的小车是一个增强学习的典型的案例，如果它碰撞了给它惩罚，如果没有碰撞给它奖励，多轮的循环会加强这些行为，达到一个期待值。深度学习可以用来做环境的感知，而增强学习可以用来做控制的东西，这样就可以构成一个完整的自动驾驶系统。

这是一个非常典型的汽车的自动驾驶系统的框架图，在图像的左侧有各种各样的传感输入，这个数据来源于三个，汽车外围的环境数据，汽车本身的状态数据，比如车速、转向，还有车内人的数据，是不是疲劳驾驶，是不是要转弯。一个是感知的融合，一个是决策。

深度学习让自动驾驶从感知到控制

深度学习在中间的感知和决策都是有非常大的优势，可以进一步把这个事情分解一下。

我们跟过内的主机厂接触，由于现在国内有一些主机厂已经开始比较深的自动驾驶的研发，去讨论大量的细节。我们可以简单把它划分成这样的一个板块，就是**态势感知和决策控制**。

• 态势感知和决策控制

态势感知一部分就是对外的态势，就是你要对整个的环境要有感知，对内就是对驾驶员的意图要有感知，包括意图的判断。包括对外部物理的辨识，哪些是固定障碍物，哪些是栏杆，哪些是可行驶区域。对于驾驶，包括状态的识别其实非常重要。这也是**整个自动驾驶业界流派之争的一个焦点**。

以谷歌为代表的，他们想做的是全自动驾驶，但是我们还是要一步一步来，这面临有一个半自动驾驶的过程，谷歌觉得很难做到很好的人工和自动驾驶的切换，没有办法把这件事情做得特别的可靠。

其实这个我认为要这样来看，首先**主机厂其实是一定会按照循序渐进的过程来走**，因为他们是造车的。第二个在这件事情上**深度学习可以很好地帮助缓解这个问题**。我们不能说一定能全部解决，至少能够极大地缓解这件事情。因为深度学习对于驾驶员状态的综合理解比其他的手段要来得更好，因为它是基于大量对于驾驶员技术分析的基础上建立对它的一个综合理解，包括他的驾驶习惯、意图等。

决策控制包括几个部分。就是对于车辆来讲有一个局部的运动路径的规划，另外规划里面还有一个驾驶风格的情况，比如需要运动感强一点的，都不一样。最后一个是执行器控制，状态之间形成一个闭环去调校。

应该说在现在有很多初创的公司，包括像英伟达这样的公司做大量的尝试，如何把整个的系统用深度学习的方法进行优化。英伟达做得比较暴力，直接忽略了中间的决策这一部分，是直接把传感器的输入和车的状态的输入直接关联，做这样一个预测。也就是说Y就是执行器刹车、油门、转向，输入就是车的信息、环境的信息，非常暴力，不做任何中间人工智能的分析。

• 自动驾驶还有一个非常重要的点，就是如何进行测试。

整个自动驾驶来讲主要的投入还是在测试这块，并不是做一套软件系统就可以了。特斯拉做得很讨巧，直接在量产上进行测试，所以短时间内积累了超过1.76亿英里的速度。但是还是需要一个模拟系统去做这件事情，有一个模拟系统，如何在模拟系统各个主体能够真实的模拟，这其实也是一个增强学习可以帮助你

的。当你创造了一个虚拟环境的时候，你希望用多台的计算设备去模拟一个目标，模拟几十个人，模拟路牌等等，每一个都是动态的过程，它们不地变。比如红绿灯会不断地变，车辆动态的变化，这样一个复杂的结构，其实很适合于深度学习和增强学习来做，包括结合模拟。

在这个方面德国的车厂宝马和奥迪走得比较靠前，当然包括丰田其实也是一样，做大量的工作，这个是奥迪的汽车大脑的构架图，AUDI的zFAS。zFAS系统模块配备了Mobileye的EyeQ3移动处理器以及英伟达TegraK1芯片。**奥迪的观念非常清晰，自动驾驶的关键还在学习。**

地平线在这块领域也是进行了大量的工作，我们有一个品牌叫做雨果，这是汽车的一个开放平台。

（播放视频）这个是雨果做的整个道路的语义理解，它可以把道路各个不同种类的物体运用不同色彩标识出来，比如栅栏是红色，道路是绿色，树木是深绿，汽车是紫色，这样一个道路综合语义理解对于智能驾

驶系统是很关键的。很多道路也不是结构化的道路，比如乡村的道路其实没有车道线，栅栏也没有，你没有办法通过车道线去自动驾驶，这个时候你要自己去决定该走哪一块。

这是地平线做的识别，左上角是英伟达的，很遗憾不能秀一下动态的视频。地平线是在北京做的测试，在这个路口的人其实非常的密集，地平线做了识别，并且对于出租车里面的、仅仅露一个脑袋的人也可以识别。

对于车辆检测，地平线从去年8月开始，长期保持在KITTI排名第一的识别率。Densebox就是测试的名字。

地平线是一家专注于算法和芯片开发的公司，我们在最下端的解决方案，我们更愿意跟业界合作伙伴，各类的公司和主机厂一起合作，一起做最后的产品出来。地平线的算法团队其实蛮国际化的，助力的算法团队来自于像Facebook、百度、中科院。我认为在深度学习领域，我们中国第一次有机会基本上跟国际的主流的开发的机构处于差不多的一个水平线上，不能说是绝对的齐平，但是不会有太大的差别。

我们为什么要自己做深度学习的芯片？

刚才赵总也分享了，为什么我们需要深度学习的芯片，提到很多国内公司在宣称自己在做这个，这是一件好事，因为大家都知道现在的计算结构不合理。

比如说举一个例子，深度学习的输入是高度密集化的，这种密集化输入不太适合DSP这种，是流水线的，GPU的一个问题在哪里？

在多输入之后的下一步处理，在数据缓冲方面做得不好。很不幸的是深度学习网络它的输入是前后关联的，在中间隐藏层的时候，这个卷积盒是要跟前置进行卷积计算的，这就意味着首先要取参数、计算，这就会导致在正常的一般结构下会大量的有外出效应。

第二个因为带宽永远是有限的。

第三个是功耗问题，这些都是非常实际的问题。

还有一点就是，**深度神经网络的参数非常多。**

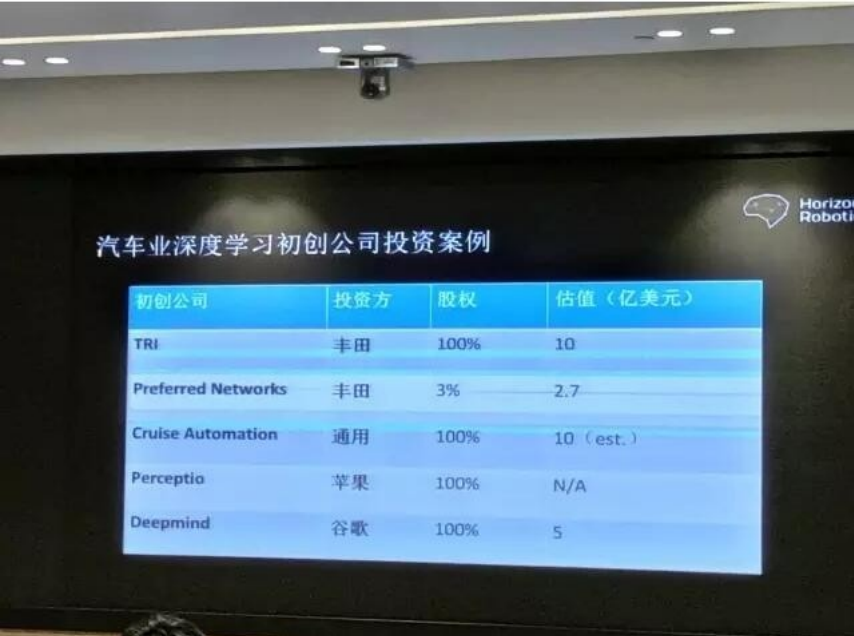
整个业界都意识到，如果要在深度学习领域能够做的话，像地平线得自己做芯片有点类似于手机业界，在大约十年前的时候，手机业界还没有一个想法说一定要做自己的芯片，但是今天大家已经看得非常清楚，**如果要想做起来必须做自己的芯片**，苹果、三星、华为都是自己的，如果你不这么做只是使用高通的平台，你还是可以做，但是很不幸你是第二梯队。所以在深度学习也是一样，你还可以用英伟达、高通的芯片，可是这不是第二名和第一名的差距，这是第二梯队和第一梯队的差距。因为算法在演进，你如果用别人的话肯定是慢一拍的，这是一个问题。

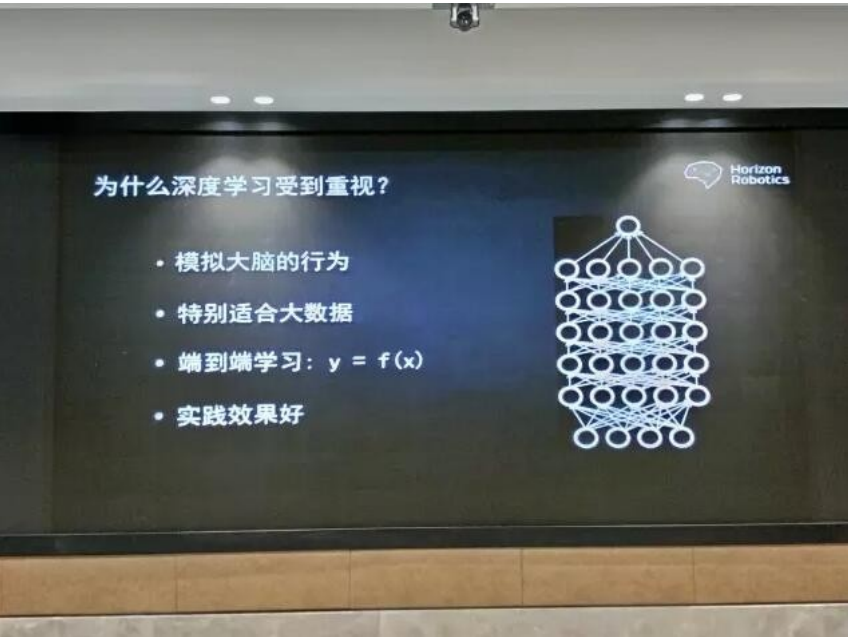
国内有不少公司在做，包括有一家公司叫做（深鉴科技），最近他们发表了成绩，就是DPU，深度学习的处理器，他们在单位功耗下的性能表现是英特尔之前CPU的24000倍，是英伟达GPU的3000倍。而他们所做的优化的工作，恰恰是刚才我提到的问题，**对于压缩，对于访问的优化，降低了功耗。**

地平线也是一样，我们非常清晰认识到这个问题，所以我们的做法也是这样，我们地平线也是在做自己的芯片。我们希望在短时间内能够把整个的能耗比提升千倍。

地平线现在专注的深度学习的技术的应用点有哪些？四个大块，语音、语言、视觉、控制。应该说，这四个纬度可以构成大量的创新的应用，不仅仅是自动驾驶，还包括家庭的服务器的监控等等。我们的目标是，希望在未来可以用我们自己的算法芯片，跟广大下游的伙伴进行合作，比如空调公司，比如扫地机器人公司等等，能够覆盖超过1000个品类的智能产品，我们的目标是为现代的智能硬件真正的赋予智能，所以我们的目标是定义物联网时代的大脑。

我的分享就到这里，我们也非常欢迎大家来跟地平线合作，我们会提供平台解决方案，而不是某个具体品类和最终的产品，谢谢大家。







雷锋网(公众号：雷锋网)注：本文为雷锋网整理文章，转载请联系授权并标注出处和作者，不得删减内容。

雷锋网原创文章，未经授权禁止转载。详情见[转载须知](#)。

5人收藏

分享：

相关文章

深度学习

自动驾驶

芯片

汽车

中科院

深鉴科技

地平线机器人



如何利用机器学习预测房价？



一文看懂深度学习在语音合成&增强上的应用



Innoviz CEO专访：拿下6500万美元后，如何用固



六张图带你全方位看懂EMNLP 2017，神经网络和



知乎的「野心与终局」



龙芯推出新一代处理器，离Intel 还有多大差距？



钱晨：为什么“智能音箱”可能会被做臭？



我知道的周航

文章点评：

我有话要说.....

☐ 同步到新浪微博

提交

最新评论

- 

小芹菜 回复: 范伦丁•鲍勃 06月26日 11:21

多谢指出，文章说根据速记整理有点错误。小学毕业的人得好好说话不是。

(0) 回复
- 

卢卡 06月25日 10:55

很不错

(0) 回复
- 

范伦丁•鲍勃 06月25日 09:53

这篇文章水平太低，错别字太多。level 3竟然打成leve 13，你是不是英语没小学毕业？深鉴科技写成深建科技，你不知道这个公司名称你不会输入拼音百度一下？

(0) 回复

热门关键字

热门标签 微信小程序平台 微信小程序在哪 CES 2017 CES 2016年最值得购买的智能硬件 2016 互联网 小程序 微信朋友圈 抢票软件 智能手机 智能家居 智能手环 智能机器人 智能电视 360智能硬件 智能摄像机 智能硬件产品 智能硬件发展 智能硬件创业 黑客 白帽子 大数据 云计算 新能源汽车 无人驾驶 无人机 大疆 小米无人机 特斯拉 VR游戏 VR电影 VR视频 VR眼镜 VR购物 AR 直播 扫地机器人 医疗机器人 工业机器人 类人机器人 聊天机器人 微信机器人 微信小程序 移动支付 支付宝 P2P 区块链 比特币 风 高盛 人脸识别 指纹识别 黑科技 谷歌地图 谷歌 IBM 微软 乐视 百度 三星s8 腾讯 三星Note8 小米MIX 小米Note 华为 小米 阿里巴巴 苹果 MacBook Pro iPhone Facebook GAIR IROS 双创周 云栖大会 智能硬件公司 智能硬件 QQ红包 支付宝红包 敬业福 支付宝敬业福 支付宝集五福 Waymo 虚拟现实 深度学习 人工智能 中国银联 蚂蚁金服 WRC CNCC material design nubia 系统 机械硬盘 1200美元 facebook开发者平台 酷我k1 eero 雷锋网ppt 川普 中国 锤子发布会2016 量子点电视优缺点 mt8173 iphone7plus 摄像头测评 树莓派用win10能干嘛 mindbody 更多

联系我们 关于我们 加入我们 意见反馈 投稿