

城市自动驾驶红绿灯方案何去何从？



漠北的孤狼

电子科技大学 仪器科学与技术硕士

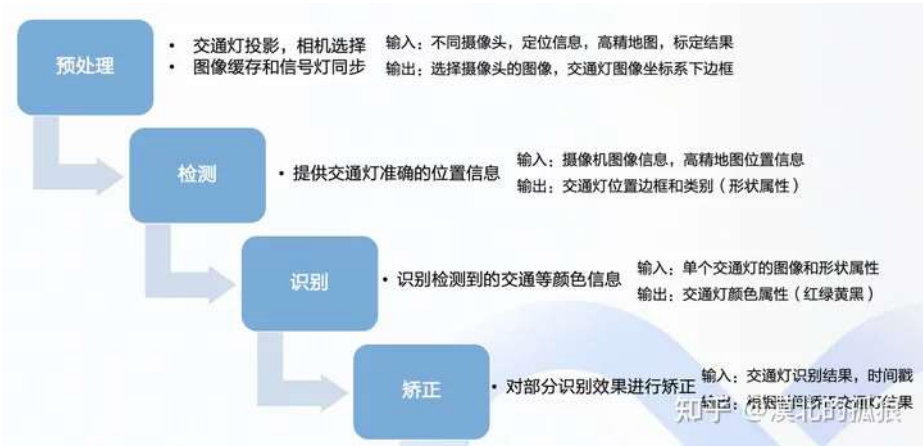
[关注他](#)

6 人赞同了该文章

目前自动驾驶的下半场，已经开始卷城市自动驾驶了，对于城市自动驾驶红绿灯检测是一个难点，特别是要覆盖的城市越来越多的情况下。

基于检测的方案

借助高精地图给出的位置信息，预先从相机图像中将包含红绿灯区域的大致位置取出来

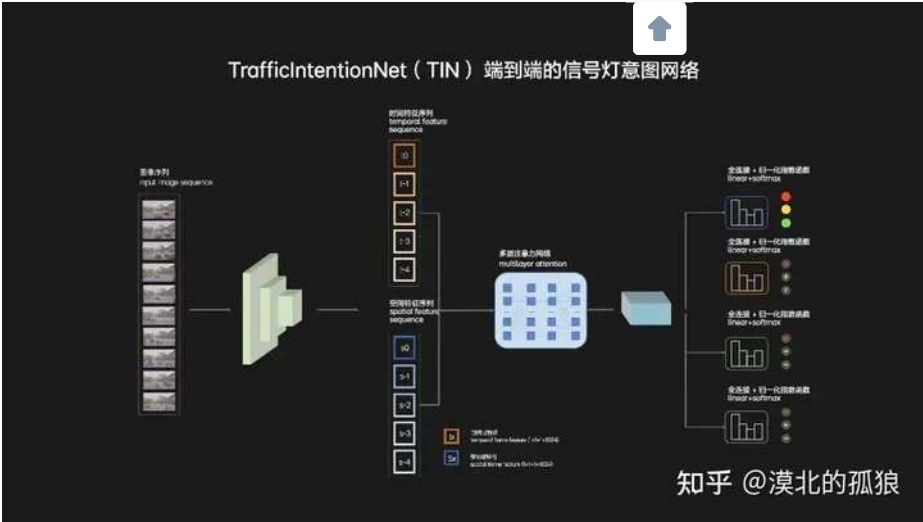


apollo 红绿灯 pipeline

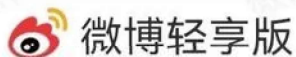
1: 该方案的优点, 方案比较成熟，有可以借鉴的东西，模型方向和数据方向都比较成熟（github.com/ApolloAuto/a...）。可以很快出一个不错的效果。2: 方案的缺点，长尾问题比较严重。

基于端到端的方案

1: 在理想科技日上 发布了关于红绿灯的新思考，直接是出通行概率。



理想科技日红绿灯pipeline



郎咸朋

理想自动驾驶负责人 努力探寻与创作美...

关注 70 粉丝 4716

2023-06-19 18:48

关于TIN网络的一些分享

在家庭科技日上我提到，相比于高速和快速路场景，城市场景与它们最大的差别就在于路口。不仅要准确识别路口的道路结构，还要准确的在路口按信号灯指示行驶。前者我们用NPN特征解决，在我上一个微博里介绍过它的特点。而后者我们用信号灯通行意图网络TIN（Traffic Intention Net）来解决，这是理想在城市自动驾驶场景的另一个重点工作。时间关系，我在家庭科技日上对这部分讲的比较粗略，但这个工作在我看来是跟NPN特征一样重要的，所以还是要跟大家简单分享一下。

不知道大家有没有发现一个小问题，TIN网络的英文名和中文名似乎对不上。按照字面直译，信号灯通行意图网络的英文应该叫Trafficlight Intention Net才对。我承认我不是英文母语，但也不至于犯这种错误。那这是为什么？大家看完这篇微博就会找到答案。

按信号灯通行，业界已经有一些常规的做法，我们来看一下。

知乎

个红绿灯的位置和颜色状态，这里用了“各个”是因为在复杂路口有很多个红绿灯会同时出现在一个画面上。

其次，匹配。将检测到的各个红绿灯与高精地图中的红绿灯位置做个匹配。匹配方式通常有两种。一种是将三维高精地图上的红绿灯位置投影到二维图像上做匹配，另一种是通过二维图像的连续帧恢复红绿灯的三维信息后再跟三维高精地图上的红绿灯做匹配。最后，决策。根据高精地图里红绿灯与其所控制车道的关联关系，以及感知到的灯的颜色状态，做出车辆行驶的决策。左右转、直行或者停止等待等。

这个做法理论上是OK的，但在实际使用中会遇到很多困难。

第一，红绿灯类型复杂。我国的红绿灯类型非常多非常复杂，其根本原因是标准不统一。根据我自己2013年在百度做街景时的经验，国家颁发的标准仅仅约束了红绿灯发光单元的具体要求（功率、大小、亮度等），并没有对样式做统一规范。当时百度大大小小做了400多个城市的街景，红绿灯是五花八门，最奇葩的要属天津的“进度条”红绿灯。如果要识别所有的红绿灯，确实需要收集大量的样本。而且很多红绿灯是“孤品”，用于训练的话样本量显得很少，需要做样本增强或者设计一些额外的规则。

第二，强依赖于高精地图。上面提到匹配和决策过程都需要高精地图的参与，如果红绿灯的位置、类型以及对应车道发生了变化，都需要及时更新高精地图，否则整个过程就会失败。

第三，强依赖于高精定位。如果定位不够准

知乎

是二维恢复三维，高精地图中的红绿灯位置与图像检测到的红绿灯位置就对不上了。更严重的是匹配错误，如果匹配到旁边的红绿灯上去了，看错了信号，就会造成交通事故。而定位漂移的情况在城市中发生的概率很大，特别是高楼林立的繁华街道，“城市峡谷”现象非常突出。高精定位也与高精地图有关系，因为有些做法的定位主要靠感知与高精地图定位要素的匹配完成。

以上的困难，实际上我们在早期的研发过程中都遇到过。为了解决红绿灯样本问题，团队使用自研的数据闭环系统设计了一套快速获取红绿灯样本的机制，当时已经覆盖了338个城市，标注了36.5万帧。为了解决定位漂移，我们也设计了很多后处理规则，来降低定位发生偏移的概率，甚至考虑换硬件。但是，就算这些解决了，高精地图的时效性还是一座大山立在面前。

团队最终讨论的解决方案，还是升维使用大模型解决。用一个端到端的网络（这里的端到端，是从传感器到行车决策）搞定所有问题。输入就是图像，而且只用图像，输出就是行车意图。并且使用Transformer架构，这样时序帧间的信息都可以一起纳入考虑。

在单帧特征提取时使用残差网络

（ResNet）获得高维的场景特征向量。对于历史帧特征向量，我们维护了两个特征序列，时间特征序列和空间特征序列，分别包含过去2秒和过去50米的帧特征。整合的特征序列会被送入多层注意力编码网络（Transformer）。在这里结合空间和时间全面的信息来保证意图识别的稳定性（对抗红

知乎

之后，经王廷按云辇画出牛马图。

不是红绿灯类型多吗，不检测了。不是高精地图更新难吗，不使用了。不是定位偏吗，不依赖了。

TIN网络的训练数据就是驾驶员通过这个路口的视频+刹车油门开合度和方向盘转角数据。刹车油门和方向盘转角，可以对应到当前的行车意图，简单说就是左右转、直行和等待等。在研发过程中，我们遇到的最大问题是训练数据。因为训练面临“冷启动”。还好有过去的积累，我们从6亿公里的训练样本存货中快速挖掘出来了大量驾驶员通行路口的数据，并实现了完全自动化标注。因为这个标注非常简单，一张图自动对应一个刹车油门方向盘数据，不需要人为再去判断任何内容。但是如果标注比较奇怪的红绿灯样本，人可能都得琢磨一阵，就谈不上自动化标注了。

这里再给一个小的技术细节。NPN特征覆盖的过程，也是TIN网络迭代的过程。收集路口NPN特征的同时，也在学习和训练TIN过路口的能力，这两个是伴生的。NPN+TIN这样的技术架构和方案，也蕴含着我们对城市道路的深刻理解。后续我们的规划是，将BEV大模型（包括NPN+TIN）跟规控模型打通，实现完整的端到端，即从传感器到执行器。

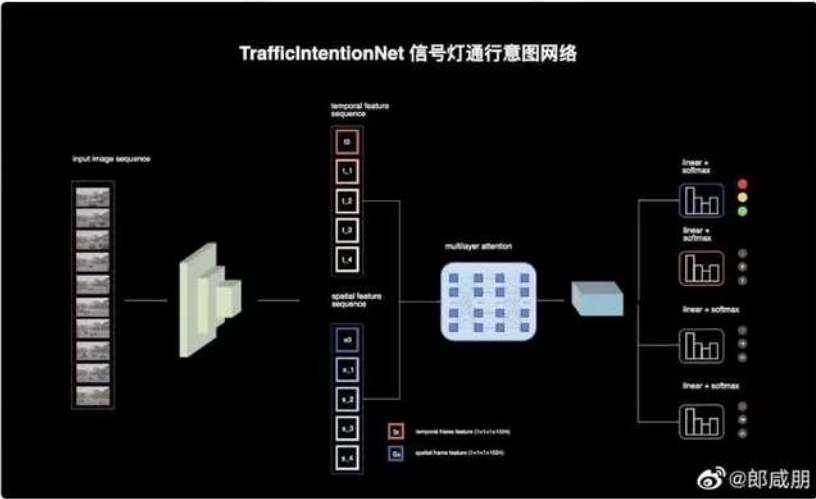
多亏我们做了正确的决定。就在刚刚过去的5-6月份，望京地区进行了红绿灯的更换，据不完全统计有80多个。如果不是使用

知乎

我心六就500多人，即下物么又新同情地图也弄不过来。而我们惊喜的发现，大部分换了红绿灯的路口，即使不补充数据，TIN网络依然有能力准确给出通行意图。这又是为什么？

回答这个问题，就回到了我们最初的问题，TIN为什么不叫Trafficlight Intention Net？是因为模型从图像中学习到的，并不仅仅是红绿灯信息和特征，而是整幅图像的信息与通行意图的对应关系。图像里有红绿灯，也有其他车辆和行人的移动，也有周围的环境，一张图包括的内容太多了。我们开车时实际上也并不需要每时每刻看红绿灯，跟着前车走也是一种方式对不对。我们其实并不知道大模型学到了什么，但它一定学到了什么，所以才会做出正确的通行意图，所以叫他Traffic Intention Net更合适。

大模型时代，产生了新的研发范式。我们真的不能再老眼光老思路去判断现在的技术和进展了，自动驾驶的发展一定比我们的想象还要快，我是坚信这一点的。



112

71

63

扫描二维码看更多精彩内容

分享来自微博轻享版客户端



知乎 @漠北的孤狼

发布于 2023-09-13 22:44 · IP 属地上海

自动驾驶 自动驾驶技术 目标检测

▲ 赞同 6 ▼ ● 2 条评论 ↗ 分享 ♥ 喜欢 ★ 收藏 📄 申请转载 ...



发布一条带图评论吧

2 条评论

默认 最新



爱吹牛的诚实小伙

类似小鹏目标识别率还是蛮不错，比如无遮挡情况下有80~100m能识别到红绿灯

09-14 · IP 属地陕西

● 回复 ♥ 喜欢



漠北的孤狼

作者 蓝 金

是的，检测的方案目前来看效果还是有的，主要关注的可能是如何泛化到更多的场景里

09-14 · IP 属地上海

● 回复 ♥ 喜欢

推荐阅读



自动驾驶汽车的车道检测

deeph... 发表于deeph...



Street Lanes Finder - 检测自动驾驶汽车的车道

相约机器人



竟是毫末率先量产，北京上路，量子位