

团队介绍：

我队由三位成员组成，分别是：

- 来自北京大学的夏雄尉，研究方向是深度学习在视觉领域的应用；
- 来自北京信息科技大学的赵明珠，方向是无人驾驶；
- 来自中搜网络技术有限公司的软件工程师雷蕾，方向是无人驾驶。

在本次的比赛中，我们发挥的并不好，很多赛前准备的想法因为时间比较仓促都没有实现，非常可惜，但这却是一次很好的锻炼的机会，也积攒了一些经验。

作品介绍：

我们采用的深度学习框架是 caffe，由于我之前在项目中使用 caffe 比较多，所以在框架的使用上并没有太大的问题，但是系统环境不能直接支持网上开源版本的 caffe，而且 caffe 的依赖库比较多，所以在比赛前半段的时间都浪费在了环境和框架的配置上。

我们本来的想法是使用一个轻量级的目标检测框架 pvanet，detection 和 classification 一起训练，这种训练方式也是目前做 detection 的主流方式；pvanet 它的设计思想是 less channels with more layers，相比 rfcn 和 faster rcnn，可以达到更快的检测速度且保持精度不降。理想很美好现实却很骨感，本以为可以很好的跑出结果，结果切换到本次比赛的数据集，怎么跑都跑不通，需要调整的参数都已经调整了，可在训练过程中始终有一个错误无法解决。眼看时间一点点流逝，为了避免最后没有结果，我只能转战其他框架，这里我使用了山世光老师团队的 seetaface 去单独做 detection，微调了参数后跑出了 detection 的结果，但是结果并不是很理想，部分图片没有检测到人脸，检查发现是部分图片由于有遮挡所以没有检测出人脸。得到了 detection 的结果后，接下来就是做 classification 了。

Classification 的思路是先根据 groundtruth 把人脸部分 crop 下来，resize 到一定大小后送入 cnn 中进行训练，测试阶段就根据 seetaface 的检测结果将人脸区域做同样的处理进行测试，但是这里存在一个问题，首先 seetaface 的检测结果是有误差的，会产生误差累积，其次部分图片没有检测出人脸，或者图片中人脸漏检，这里只能很 naïve 的认为没有检测出人脸的图片的分类结果为 mask；没有检测出人脸的图片 bbox 的结果我本来的想法是根据训练数据算出一个平均的 bbox 位置作为返回结果，这是没办法的办法，后来由于时间不够也只能作罢。这里我用的

Cnn 的结构与 VGG16 类似，因为考虑到这里是二分类任务，且数据量很小，如果使用原始的 VGG16 可能会产生过拟合，所以我用了我之前使用过的一个浅层 VGG，即层数比 VGG16 要浅，且加入了 batch normalization。由于做 classification 任务的时候时间已经非常紧张了，虽然我们拿了 50% 的数据，但是由于时间问题也没用上，只用上了 20% 的数据，所以最后的结果也很不理想，还有一些其他辅助分类的 loss 也因为时间问题没有用上，非常遗憾。

总体来说本次比赛我队的结果并不理想，团队的任务分配上做的不是很好，相关经验也不够丰富，导致时间非常仓促，但这却是一次非常棒的体验，第一次和队友连续 12 小时编程解决问题，累并快乐着。比赛中印象最深的莫过于队友的支持和鼓励，还有和参赛选手们一起埋头奋战的氛围，这是平时做项目很难体会到的感觉，所以今后有类似的比赛一定会再次参加。在这次比赛中也见识到了很多大牛，认识到我们还有许多不足，还需多多积累。