中国研究生智慧城市技术赛，这个比赛是基于深度学习

任务：通过全景照片定位拍摄相机所在的坐标位置。和普通的图像定位不太一样，这个是通过图像定位相机所在的位置。

数据集：简单说一下数据集，一共有八个不同的场景，有室外，还有室内，每个场景都拍摄了一系列照片。照片呢，就是通过一款全景相机360度拍摄6张照片，由于相机是鱼眼相机，得到的是有畸变的照片。在拍摄的同时，还利用一款定位仪对相机进行了位置标定，得到了x,y,z三个空间坐标点。

任务：给了训练集和测试集。训练集是图像加位置标签。测试集只有图像。我们的任务就是通过训练及训练模型，使得模型在测试集上得到的位置误差尽可能小。

工作：

1.使用opencv将每个位置点的6张全景图片矫正畸变，然后拼接成一张照片。

2.搭建神经网络。使用Tensorflow搭建了很多网络，各种尝试，最终使用inception\_V3。

3.首先在其他数据集上对模型进行了预训练，然后迁移到训练集上进行训练，最后得到了不错的结果。

·去畸变：首先用棋盘格照片进行相机标定：得到相机的参数矩阵，畸变系数，旋转矩阵和平移向量。然后使用这些系数进行去畸变。所用到的函数都是opencv模块里面的鱼眼相机去畸变函数。

·图像拼接：因为图像拼接的原理已经很成熟了，也不是比赛的重点，我上Github上找了一个程序，修改了下参数就运行了，效果还不错。基于SUFT(SIFT尺度不变特征变换的改进)，尺度不变性：不仅在任何尺度下拍摄的物体都能检测到一致的关键点，而且每个被检测的特征点都对应一个尺度因子。 速度上快了3倍左右，但是精确度小了一些。

·对每幅图进行特征点提取

·对对特征点进行匹配

·进行图像配准：两张图像转换为同一坐标系下，求得变换矩阵，实现图像配准。

·把图像拷贝到另一幅图像的特定位置

·对重叠边界进行特殊处理，去裂缝。

·搭建神经网络：期间尝试了很多神经网络，有自己搭建的，有比较经典的。尝试了GoogLeNet的inception\_v1到v\_4，还尝试了vgg和残差网路，最后我得硬件条件下inception\_v3的效果最好。我把最后的分类器改为了回归，然后添加了一层全连接层，效果好了一些。我没有GPU，想办法使用Google云盘上的免费GPU来着，只有一块，云盘空间有限，所以跑不起来大型的网络。选择的inception\_v3整个训练模型只有200M左右，VGG的话800多M。残差网络和v4版本训练的有一些慢。所以综合考虑选择了v3。

期间困难很多，我觉得有个队友帮我的话应该会好很多。没有硬件GPU，只是也都是现学的，看书看博客刷视频，从头到尾一直自己在弄。这个有很浪费时间，每次尝试一个新的想打的周期有点长，所以最后的成绩不能说很好，但是还算挺满意了。