1. 请说一下dbscan和kmeans的原理，为什么要这样做，这样做的目的是什么。
2. 深度学习雷达信号检测，一共有多少的数据量，数据大概长什么样，为什么用一维卷积不用二维卷积，识别的准确率是多少。这样做的目的是啥。

几十个G。上千万条。一条信号1024个点。做完时频后时频图存储量大。一维好的能到99，差的75。

训练集验证集好，测试集不好，最后发现是样式问题，测试集中有一些样式没取到，然后沟通后多取了一些样式。

有些在时域图上看起来一样，但是转换成时频图就不一样了。

置信度困难，之前一直用softmax效果不好，后来换成了双softmax，但是局限性不好，因为没有库外，没有猫狗之外的生物，后面改成了距离的置信度。

假设有10类，每个类有10000个样本，对这一万个样本用kmeans聚成1000个簇，然后如果要选5000个样本代表这一类，那么就从这1000个簇中每个簇选取5个样本，这5个样本是距离聚类中心点最近的点，这5000个样本点就能代表这一类，这样在算NN距离的时候需要计算的量就变少了，存储量少了一半。以上参数都是可以动态变化的。通过看论文。

1. dssd项目的目的是什么，遇到了什么难点，准确率分别是多少，进行了哪些改进，说一下dssd和ssd的原理，比之前的比如说faster rcnn要好在哪，还有比这更好的小目标检测系统吗。

DSOD。