由于我的轨迹预测是为了导航服务的，所以所有的2D图像上的轨迹预测统统不考虑，因为这些轨迹预测无法提供障碍物在三维空间中的位置。

要在三维空间上预测轨迹，只用视觉是不精确的，所以我的想法是激光雷达加视觉融合。阅读了《Human Motion Trajectory Prediction: A Surve》，选择了其中采取kitti数据集的几篇论文，因为这几篇论文采取了激光雷达加视觉的方式，和我的想法较为接近，所以深入学习这几篇论文。

1. DESIRE: Distant Future Prediction in Dynamic Scenes with Interacting Agents

1.1辅助论文《Learning Structured Output Representation using Deep Conditional Generative Models》

这篇论文中的网络结构被DESIRE用来作对比，因此也了解一下。这篇的结构称为CVAE，是一种用来推断在状态s下出现x的概率的网络结构。这种网络通常用来预测。