

# 小学期实践报告：三维点云模型补全算法设计与实践

软 73 沈冠霖 2017013569

## 1. 背景介绍

点云是 3D 计算机视觉处理的常见数据，但是由于遮挡，吸光材料等多种原因，获取的点云数据一般不完整，因此点云补全技术十分有意义。

而目前许多点云补全技术，比如 PCN[1],都是有监督的，需要完整的点云作为 ground truth。但是，获取完整的点云 ground truth 十分难，需要使用 CAD 来绘制。

但是,现在已经有 ShapeNet 等有完整 ground truth 的点云数据集了，因此，使用有完整 ground truth 的数据集，来帮助没有 ground truth 的无标数据进行补全，很有意义。

## 2. 思路介绍

首先，为了实现和验证方便，我们采用 PCN[1]作为我们网络的 backbone，也就是采用 encoder-decoder 结构。

之后，我们参考了多模态深度学习论文 ACMR[2]的思路，尝试利用 triplet loss 学习到一个变换，来把有标，无标数据映射到同一个特征空间，进而改进效果。我们的详细网络结构如下：

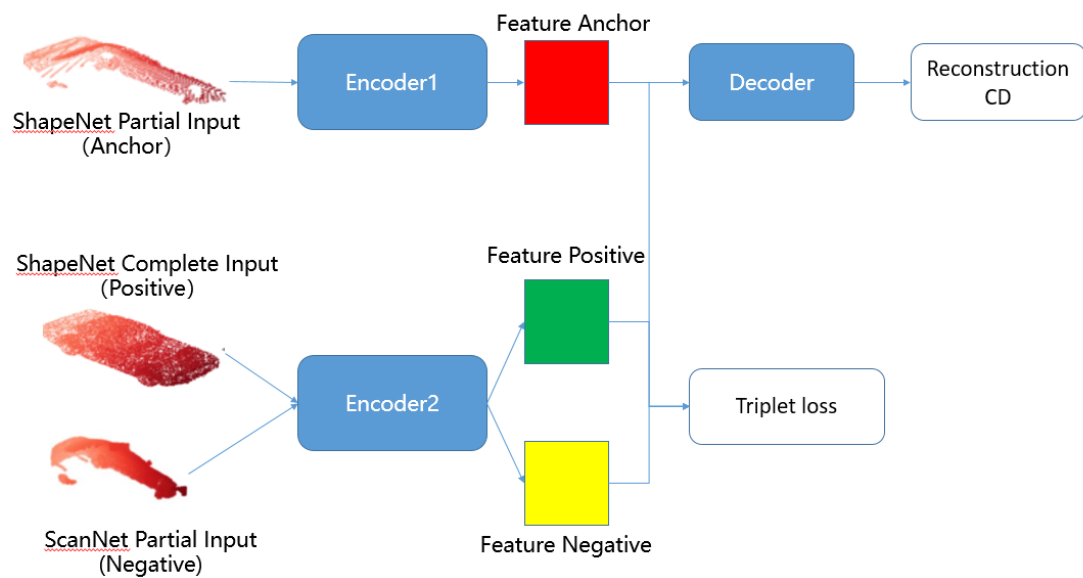


图 1: 整体网络架构

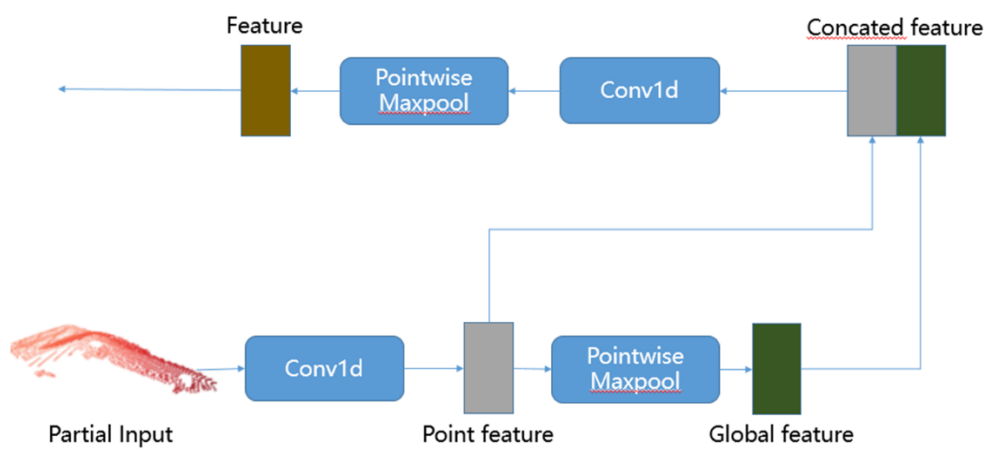


图 2: 我们的 encoder 架构, 和 PCN 一致

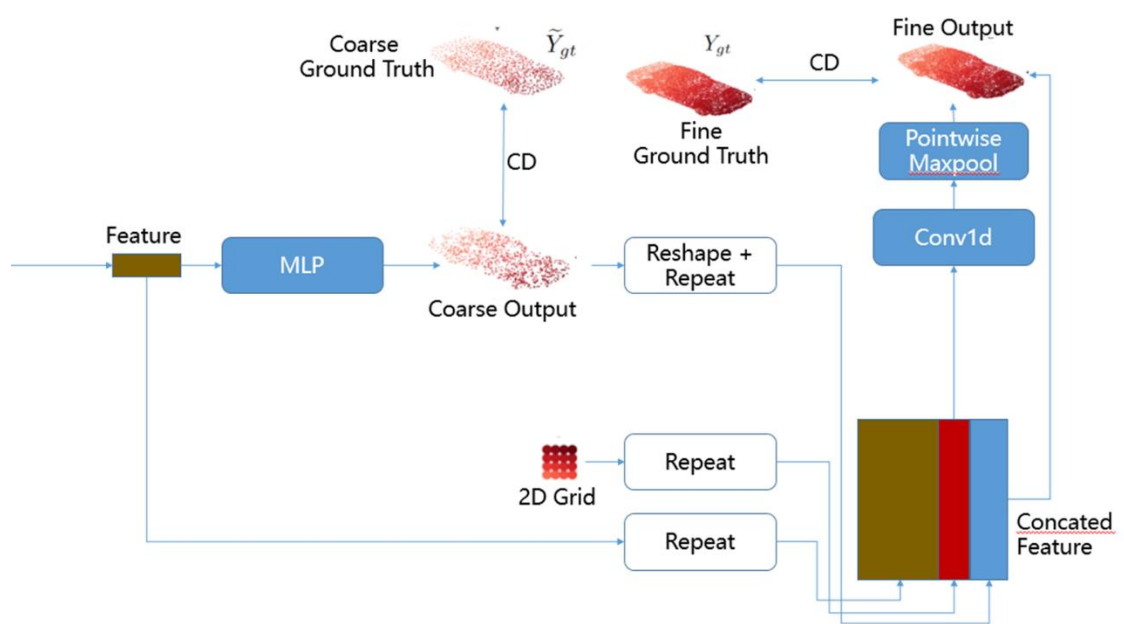


图 3: 我们的 decoder 结构, 和 PCN 一致

对于数据集, 为了便于验证我们思路的合理性, 我们暂时只用了有标的 shapenet 数据集, 如果这个思路能改进 (至少不负面影响) 有标数据的点云补全, 才有可能改进无标数据的点云补全。我们采用了标准 shapenet 数据集和经过处理的 shapenet completion 数据集 (将 shapenet 的每个完整点云在八个不同视角进行裁剪, 得到八个对应的不完整点云数据)。

我们的评价标准采用 Chamfer Distance (CD), 度量两个点云最近点的距离之和。

### 3. 实验结果

首先, 我参考 github 的 PCN 代码, 实现了 pytorch 版本的 PCN, 用于 backbone 和 baseline。

我们测试了三种 triplet loss: triplet loss, cosine triplet loss (triplet loss 的欧氏距离改成余弦距离), 带归一化的 cosine triplet loss。在调试过 margin, 权重等超参数之后, 我们得到, 带归一化的 cosine triplet loss 较

好。各种 triplet loss 的结果如下：

	Airplane	Carbinet	Car	Chair	Lamp	Sofa	Table	Watercraft	整体
PCN	9.79	22.70	15.78	22.54	18.78	28.39	19.96	11.16	18.22
PCN (我实现的)	4.82	24.54	9.94	24.60	28.28	16.10	27.38	15.21	19.22
GRNet (SOTA)	6.12	16.90	8.27	12.23	10.22	14.93	10.08	5.86	10.64
Triplet loss	5.80	22.82	9.81	25.91	30.77	17.65	27.41	16.15	19.66
Cosine loss	5.13	21.16	9.23	23.74	27.66	16.16	25.85	14.37	18.82
正则化 Cosine loss	<b>5.33</b>	<b>22.02</b>	<b>9.30</b>	<b>22.65</b>	<b>26.41</b>	<b>16.48</b>	<b>24.38</b>	<b>14.32</b>	<b>18.36</b>

图 4：shapenet 数据集下的结果

	Chamfer Distance
PCN	35.66
Triplet loss	36.74
Cosine loss	35.96
正则化 Cosine loss	<b>35.94</b>

图 5：shapenet completion 数据集下的结果



Partial



Ground truth



PCN



我的



Partial



Ground truth



PCN



我的

图 6-1、6-2: 不同网络的可视化结果对比  
其中我的可视化用带 norm 的 cosine triplet loss, 数据集是 shapenet

综上，对比 CD，使用我的方法，最优也并不明显优于 PCN。而且对比可视化结果，我的补全结果并不好，猜测是 triplet loss 方法降低了网络的鲁棒性，带来了一定程度的“过拟合”，这种思路并不是很可行，需要一定改进。

而且，triplet loss 仅仅改进 encoder，在 decoder 阶段有标-无标数据没有进一步融合，我们需要进一步探索如何在 decoder 阶段改进网络。

## 4. 下一步工作

- 1.在 Encoder 部分，深入分析 triplet loss 失败的原因，更改 triplet loss 的采样方法，或者尝试分类标签等其他任务。
- 2.参考其他重建，分割等类型的任务，找到一种方法让无标数据在 decoder 阶段参与训练。
- 3.如果我们的设计能改进 ShapeNet 的补全效果，那么就引入 ScanNet，尝试适配网络以真的改进 ScanNet 的补全效果。

## 5. 心得体会

这个项目是我第一次进行基于深度学习的 3D 计算机视觉研究，虽然并未完成，而且目前结果不是特别乐观，但是我还是收获很大。

一方面，我学习了许多科研所需的技能，充分了解了基于深度学习的 3D 计算机视觉研究。我阅读了 PCN 等点云补全方面的许多经典，前沿论文，以及各种可以参考的其他论文，提升了文献检索和阅读能力，了解了 3D 计算机视觉，尤其是点云领域。我用 pytorch 基于 github 的 PCN 项目实现了自己的 pytorch，以及自己实现了各种 triplet loss 和各种其他的实验，锻炼了科研所需的代码能力。除此之外，我也了解了网络绘制，结果可视化，linux

服务器操作等多种科研所需的技能。

另一方面，我体验到了，科研会经常遇到失败和挫折，也会有一些阶段性成绩，这让我对未来的科研生活有了更深的了解和认识。

感谢刘老师和组内温欣，刘鑫海等学长的帮助，希望之后能继续深入科研，做出属于自己的成果。

## 6. 参考文献

- [1] Yuan, W., Khot, T., Held, D., Mertz, C., Hebert, M.: PCN: point completion network. In: 3DV 2018. (2018)
- [2] *Bokun Wang, Yang Yang, Xing Xu, Alan Hanjalic, and Heng Tao Shen. 2017. Adversarial Cross-Modal Retrieval. In Proceedings of the 25th ACM international conference on Multimedia (MM '17). Association for Computing Machinery, New York, NY, USA, 154–162. DOI:<https://doi.org/10.1145/3123266.3123326>*