

CAPRI-Net: Learning Compact CAD Shapes with Adaptive Primitive Assembly

摘要

CAPRI-Net [34], 是一个用于学习三维计算机辅助设计 (CAD) 模型的紧凑且可解释的隐式表示的自监督神经网络, 以自适应原始组件的形式表示。给定一个输入的三维形状, CAPRI-Net 通过构造实体几何 (Constructive Solid Geometry, CSG) 操作的曲面原始组件的组装来重建它。在没有任何真实的形状组件的情况下, 本文的自监督网络通过重建损失进行训练, 从而实现了具有清晰边缘和合理的 CSG 树的忠实三维重建。尽管 CAD 模型参数化性质使其在局部更加可预测, 但在形状级别上, 存在着许多结构和拓扑变化, 这对于目前用于三维形状的最先进神经网络模型而言构成了重大的可推广性挑战。本文的网络通过相对于每个测试形状的自适应训练来应对这一挑战。此外, 本文针对每一个单独的模型, 对在模型集合上预训练网络进行微调。本文在 ShapeNet 和 ABC 两个最大和最多样化的 CAD 数据集上评估了 CAPRI-Net 框架, 评估标准包括重建质量、清晰边缘、紧凑性和可解释性, 以展示在其神经网络在 CAD 模型重建领域的优越性。

关键词: 构造实体几何; 形状建模

1 引言

计算机辅助设计 (CAD) 模型在工程和制造领域广泛应用, 用于驱动与 3D 形状和几何相关的决策制定和产品演化。随着人工智能解决方案在所有相关领域的快速发展, 出现了几个 CAD 数据集 [14, 28, 30], 以支持几何深度学习领域的研究。

CAD 模型的一个常见特点是它们由沿着锐利边缘相交的明确定义参数化表面组成。尽管 CAD 形状的参数化性质使得它们在局部和基本级别上更加可预测, 但在形状级别上, 存在大量的结构和拓扑变化, 这对于当前的三维形状神经网络重建模型 [3, 5] 构成了显著的泛化性挑战。另一方面, 现有的用于基本形状拟合的网络通常关注于具有简单、因此有限的原始构件的抽象 [29], 从而限制了重建质量。

2 相关工作

2.1 基本形状检测和拟合

计算机图形学和计算机辅助几何设计领域已经对重建几何形状进行了广泛的工作。在给定原始的三维对象时, 不同的算法如 RANSAC [8] 和 Hough 变换 [10] 已被用于检测基本形状

[12]。基于 RANSAC 的方法已经在密集点云中检测多个基本形状的应用中得到了应用 [16,25]。然而，像 RANSAC 这样的基本基本形状拟合技术受限于预定义的基本形状类型，无法预测在复杂的 CSG 树中需要减去的形状。一些方法，如 [7,9]，随后从原始输入中提取了一个 CSG 树。Hough 变换已被用于检测点云中的平面 [1] 和圆柱体 [23]。然而，这些方法通常不具备泛化能力，因为它们需要针对每个形状不同的超参数。最近，基于神经网络的算法来在点云上检测和拟合基本形状 [15,27] 的方法得到了广泛应用。SPFN 使用监督学习首先检测基本形状类型，然后估计基本形状的参数。ParseNet 通过整合样条补丁和使用可微分的度量学习分割扩展了 SPFN。本文方法的不同之处在于，本文网络的训练过程不需要监督，就能够重建一个紧凑的基本形状组装表示。

2.2 基于隐式场表达的形状表示方法

神经隐式表面表示 [5,20] 因其在训练过程中使用较小的内存占用，同时能够生成复杂、高空间分辨率的三维形状而备受欢迎。这种表示已经在三维领域中用于零件理解 [4,32]、无监督单视图重建 [11,17,19] 以及场景或对象完成 [6,22] 等方面。最近，神经隐式表示还扩展到了 2D 图像和视频领域，用于单场景完成、图像生成模型、超分辨率和动态场景等应用。本文将神经隐式表面表示的使用扩展到 CAD 模型，并展示了本文的方法如何克服 CAD 设计所带来的一些挑战。

2.3 结构感知的三维形状表示

结构感知的三维表示通常由原子几何实体及其关系定义 [2]，其中原子实体可以表示语义部分或较低级别的几何基本构件。这些实体可以组合成完整的形状，如零件集合、树结构、场景图、分层图或结构程序。

在 CAPRI-Net 中，作者使用受限制的、隐式形式的二次方程来表示原子几何实体。像 Superquadrics [21] 这样的工作使用了无约束的显式形式的超二次方程，但会导致在 CAD 建模中难以解释的基本构件。此外，它们无法产生基于平面的交点来重建锐利特征。

设计用于学习 CSG 基本构件组装的神经网络包括 CSG-Net [26] 和 DeepCAD [31]，但是它们需要监督，还有其他一些 [3,13,24,33] 无监督的方法。在后者中，BSP-Net [3] 与本文的工作最为相关，但它们在几个重要方面存在差异：

- CAPRI-Net 可以从单个二次方程生成多个不同的可解释的基本构件，而 BSP-Net 只使用平面
- 与 BSP-Net 不同，本文的网络包括一个差分操作，非常适用于 CAD 建模
- 本文引入了一个新的损失项，鼓励使用差分操作，从而产生更紧凑、更自然的 CSG 组装

UCSG-Net [13] 也以无监督的方式学习 CSG 树，但只使用了盒子和球体这两种基本构件，这可能会显著限制形状重建质量。此外，与 BSP-Net 和 CAPRI-Net 不同，UCSG-Net 中的 CSG 操作的顺序是动态的，而不是固定的。然而，灵活的顺序或泛化性会增加组装学习任务的难度。正如本文将在接下来展示的，通过固定 CSG 操作的顺序（见图1），CAPRI-Net 往往比 UCSG-Net 产生更自然、更紧凑的 CSG 树。

在另一个同期工作中，CSG-Stump [24] 也遵循了固定的 CSG 操作顺序。此工作与本文工作的关键区别包括：

- 与 BSP-Net [3] 一样，CAPRI-Net 依赖于一个简单的二次方程来表示所有基本构件，而 CSG-Stump 需要为每种基本构件类型预设基本构件的数量
- CSG-Stump 使用一个不可微分的逆层来建模形状的差异，这限制了差分操作对简单基本构件的适用性
- 与 CSG-Stump 相比，本文的网络具有专门的可微分的差分层和一个新的损失项，以支持复杂凸形状的处理，从而提高了本文方法的 CSG 树的可解释性。

3 本文方法

3.1 本文方法概述

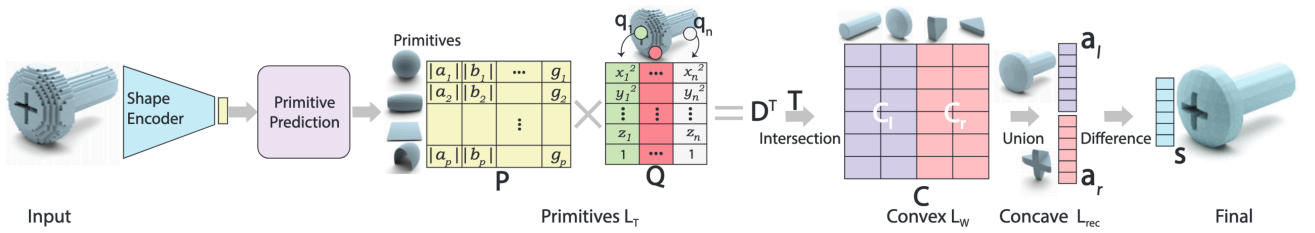


图 1. 方法示意图

给定一个输入的三维形状，可以是点云或体素，本文首先使用一个编码器将其映射到隐空间编码。这个隐空间编码用于预测包含在 \mathbf{P} 中的 p 个带有参数的基本组件。对于任何包含在矩阵 \mathbf{Q} 中的查询点 \mathbf{q}_j ，我们可以获得矩阵 \mathbf{D} ，指示了查询点到每个基本组件的近似有符号距离。选择矩阵 \mathbf{T} 用于从基本组件组件集中选择一小组基本组件，以便将凸形状分组到矩阵 \mathbf{C} 中，该矩阵指示了关于凸形状的查询点的内部/外部值。然后，我们在 \mathbf{C} 的每一半（即 \mathbf{C}_l 和 \mathbf{C}_r ）上执行 min 操作，将凸形状合并成两个（可能）凹形状，并获得左凹形状和右凹形状的内部/外部指示向量 \mathbf{a}_l 和 \mathbf{a}_r 。最后，执行差分操作，得到最终的逐点内部/外部指示 \mathbf{s} 。 $L_{\mathbf{T}}$, $L_{\mathbf{W}}$ 和 L_{rec} 是本文为网络定义的损失函数。

3.2 组件检测模块

从合适的编码器获取形状编码（大小 = 256）之后，该形状编码就会传递给本文的基本组件多层感知器（MLP）预测网络。对于每一个组件，本文使用二次方程进行表示

$$ax^2 + by^2 + cz^2 + dx + ey + fz + g = 0$$

MLP 输出一个大小为 $p \times 7$ 的矩阵。其中， p 为基本组件的数量，7 为二次方程的各个参数 ($a, b, c \geq 0$)（目的是简化学习任务）

3.3 构造实体几何操作

在本文中，构造实体几何操作有固定的顺序，即交集、并集、差分。下文将简要介绍前述的三种操作。

- 交集操作：获取凸形状

- $D_{\{n \times p\}}$: 有符号距离矩阵（带有 ReLU） \rightarrow 获得点到基本组件的距离
- $T_{\{p \times c\}}$: 选择矩阵（一开始是浮点数，当训练结束时变为二进制 01 矩阵）
- $C_{\{n \times c\}} = D_{\{n \times p\}} T_{\{p \times c\}}$: 用于在基本组件之间执行隐式交集操作 \rightarrow 获得凸形状
 - * 注意：只有当 $C(j, i) = 0$ 时，查询点 q_j 位于凸形 i 内部
 - * c : 凸形状的数量/最大交集数

- 并集操作：从凸形状获得复杂凹形状

- 定义 a_l 和 a_r （大小为 $n \times 1$ ）分别表示左侧凸形状和右侧凸形状，对应两个矩阵 C_l 和 C_r （大小为 $n \times \frac{c}{2}$ ）
- C_l （或 C_r ）包含了所有点相对于左侧（或右侧）凸形状的内部/外部指示，左侧的所有凸形状会被组装在一起形成一个形状，右侧同理
- 为了获得 a_l 、 a_r 、 C_l 和 C_r ，接下来将分别执行 min 操作，形状内部的点将被标记为 0，形状外部的点将被标记为大于 0

- 差分操作：当获取 a_l, a_r 之后，本文进行以下的差分操作来获得最终的形状

$$s^*(j) = \max(a_l^*(j), \alpha - a_r^*(j)) \begin{cases} 0 & \text{inside} \\ > 0 & \text{outside} \end{cases}$$

其中， s 用于重建形状，因为 $s(j)$ 表示查询点 q_i 是否在内部或外部。注意， α 需要是小而正的值，因为当点在形状内部时， s^* 接近于 0。

3.4 通过构造实体几何操作构建 CAD 模型网格

在获得基本构件 P 和选择矩阵 T 之后，我们可以将基本构件组装成凸形状，并执行 CSG 操作以输出 CAD 模型的网格。对于每个形状给定一个基本构件选择矩阵 T ，如果某个基本构件被某个凸形状选择，但它位于凸形状之外，那么它可能不会用于该凸形状的形成。我们采样一些接近重建形状表面的点，并获取它们的占用值 (occupancy values)。然后，我们从列表中移除每个基本组件，以测试是否会改变占用值。如果在移除一个基本组件后，所有点的占用值保持不变，我们将它从基本构件列表中丢弃。最后，我们通过在每个基本构件的有符号距离场上执行 Marching Cubes [18] 操作，构建未被移除的基本构件的表面。

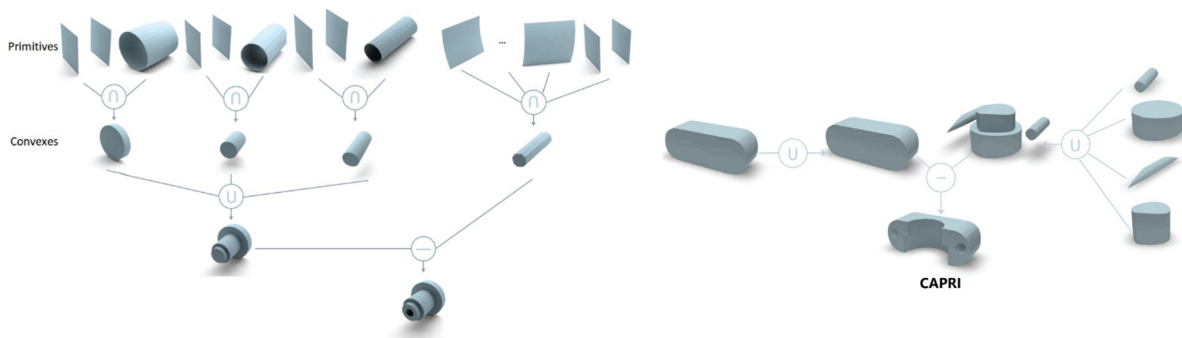


图 2. 通过 CSG 操作重建几何体

4 复现细节

4.1 与已有开源代码对比

本次复现参考了 CAPRI-Net 的官方代码，并在官方代码的基础上新增了以下内容：

- Transformer 解码器
- 更好的可视化实验结果生成代码

4.2 实验环境搭建

实验环境基于 CAPRI-Net 论文给出的原生环境，具体描述如下：

- Python 3.7
- PyMesh for Python 3.7 (使用 whl 文件安装，主要用于 CSG 操作) : [仓库地址](#)
- Pytorch
- trimesh: pip install trimesh ([仓库地址](#))
- PyMCubes: pip install PyMCubes ([仓库地址](#))

4.3 创新点

给 CAPRI-Net 引入了全新的 Transformer 解码模块

5 实验结果分析

实验在数据集 ABC 工件数据集 [14] 上进行，我们采用了 Marching Cube [18] 和 CSG 两种获得表面的方式进行了实验，Marching Cube 操作能够获得较为平滑的表面而损失了部分细节，而 CSG 操作可以获得较为锐利的特征但是对于原模型还原的忠实度有所欠缺。此外，我们还对 CAPRI-Net 在 ShapeNet 数据集上进行了测试，以检验结果是否和论文表述一致。我们发现，在台灯类上，CAPRI-Net 的还原效果逊于原文中展示的在其他类别上的结果。

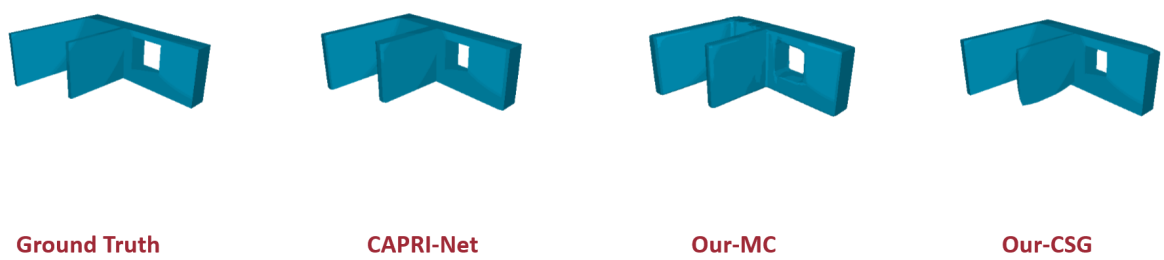


图 3. 实验结果 1

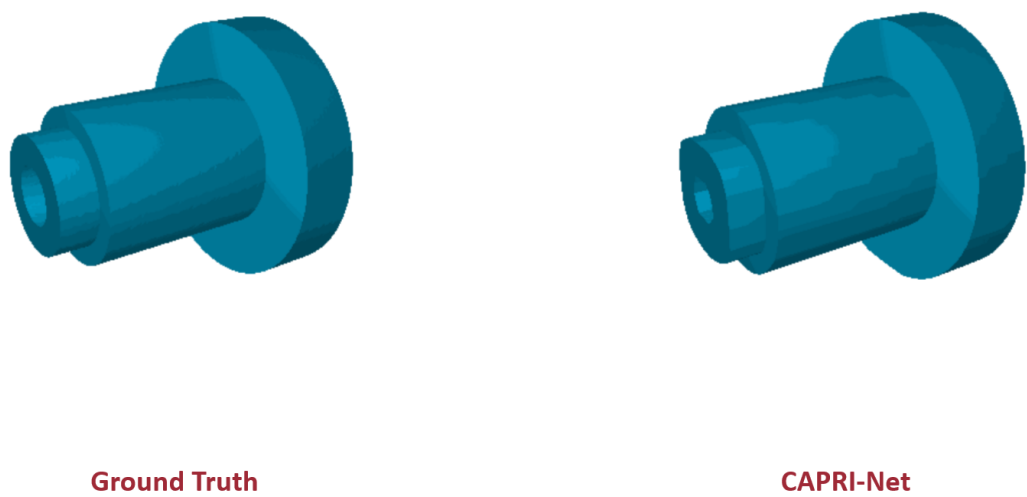


图 4. 实验结果 2

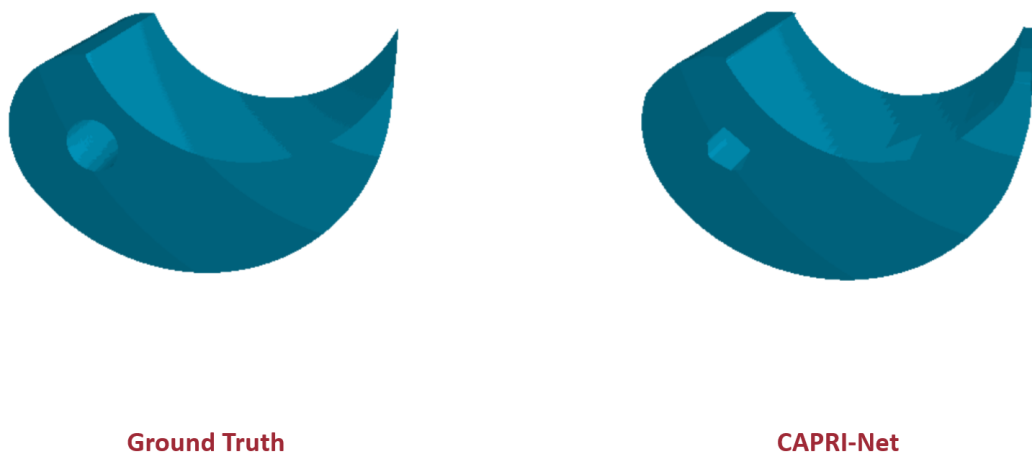


图 5. 实验结果 3

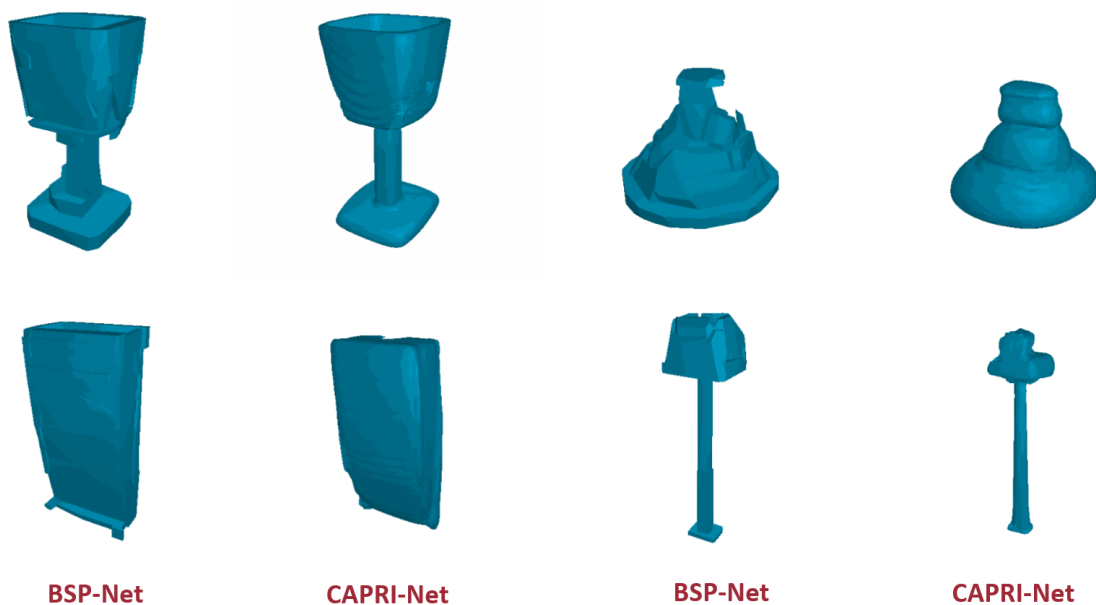


图 6. 实验结果 4

6 总结与展望

本次实验通过在原有 CAPRI-Net 的基础上添加的 Transformer 解码器来试图提升原文获取的几何表面和 Ground truth 的相似度。实验结果表明，获得的表面和原论文的结果相比提升微小。同时对于原论文的更多实验中，我们发现对于较为复杂的几何表面原论文的还原效果较差，如图 6 所示，还原的表面仍然存在非水密或非流形的情况。此外，本文的一大缺点是，需要的训练和推理的时间较长。本文需要对模型进行预训练，再针对每一个输入对模型进行单个优化，这导致此模型通用性和效率较差，这是一个可以改进的方向。

参考文献

- [1] Dorit Borrmann, Jan Elseberg, Kai Lingemann, and Andreas Nüchter. The 3d hough transform for plane detection in point clouds: A review and a new accumulator design. *3D Research*, 2:1–13, 2011.
- [2] Siddhartha Chaudhuri, Daniel Ritchie, Kai Xu, and Haotong Zhang. Learning generative models of 3d structures. *Computer Graphics Forum*, 39, 2020.
- [3] Zhiqin Chen, Andrea Tagliasacchi, and Hao Zhang. Bsp-net: Generating compact meshes via binary space partitioning. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 42–51, 2019.
- [4] Zhiqin Chen, K. Yin, Matthew Fisher, Siddhartha Chaudhuri, and Hao Zhang. Bae-net: Branched autoencoder for shape co-segmentation. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 8489–8498, 2019.
- [5] Zhiqin Chen and Hao Zhang. Learning implicit fields for generative shape modeling. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5932–5941, 2018.
- [6] Julian Chibane, Thiemo Alldieck, and Gerard Pons-Moll. Implicit functions in feature space for 3d shape reconstruction and completion. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6968–6979, 2020.
- [7] Tao Du, Jeevana Priya Inala, Yewen Pu, Andrew Everett Spielberg, Adriana Schulz, Daniela Rus, Armando Solar-Lezama, and Wojciech Matusik. Inversecsg: automatic conversion of 3d models to csg trees. *ACM Trans. Graph.*, 37:213, 2019.
- [8] Martin A. Fischler and Robert C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24:381–395, 1981.
- [9] Markus Friedrich, Pierre-Alain Fayolle, Thomas Gabor, and Claudia Linnhoff-Popien. Optimizing evolutionary csg tree extraction. *Proceedings of the Genetic and Evolutionary Computation Conference*, 2019.
- [10] Paul V. C. Hough. Machine analysis of bubble chamber pictures. 1959.
- [11] Yue Jiang, Dantong Ji, Zhizhong Han, and Matthias Zwicker. Sdfdif: Differentiable rendering of signed distance fields for 3d shape optimization. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1248–1258, 2019.
- [12] Adrien Kaiser, José Alonso Ybáñez Zepeda, and Tamy Boubekeur. A survey of simple geometric primitives detection methods for captured 3d data. *Computer Graphics Forum*, 38, 2018.

- [13] Kacper Kania, Maciej Zi ba, and Tomasz Kajdanowicz. UcsG-net - unsupervised discovering of constructive solid geometry tree. *ArXiv*, abs/2006.09102, 2020.
- [14] Sebastian Koch, Albert Matveev, Zhongshi Jiang, Francis Williams, Alexey Artemov, Evgeny Burnaev, Marc Alexa, Denis Zorin, and Daniele Panozzo. Abc: A big cad model dataset for geometric deep learning. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9593–9603, 2018.
- [15] Lingxiao Li, Minhyuk Sung, Anastasia Dubrovina, L. Yi, and Leonidas J. Guibas. Supervised fitting of geometric primitives to 3d point clouds. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2647–2655, 2018.
- [16] Yangyan Li, Xiaokun Wu, Yiorgos Chrysanthou, Andrei Sharf, Daniel Cohen-Or, and Niloy Jyoti Mitra. Globfit: consistently fitting primitives by discovering global relations. *ACM SIGGRAPH 2011 papers*, 2011.
- [17] Shichen Liu, Shunsuke Saito, Weikai Chen, and Hao Li. Learning to infer implicit surfaces without 3d supervision. In *Neural Information Processing Systems*, 2019.
- [18] William E. Lorensen and Harvey E. Cline. Marching cubes: A high resolution 3d surface construction algorithm. *Proceedings of the 14th annual conference on Computer graphics and interactive techniques*, 1987.
- [19] Michael Niemeyer, Lars M. Mescheder, Michael Oechsle, and Andreas Geiger. Differentiable volumetric rendering: Learning implicit 3d representations without 3d supervision. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3501–3512, 2019.
- [20] Jeong Joon Park, Peter R. Florence, Julian Straub, Richard A. Newcombe, and S. Lovegrove. Deepsdf: Learning continuous signed distance functions for shape representation. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 165–174, 2019.
- [21] Despoina Paschalidou, Ali O. Ulusoy, and Andreas Geiger. Superquadrics revisited: Learning 3d shape parsing beyond cuboids. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 10336–10345, 2019.
- [22] Songyou Peng, Michael Niemeyer, Lars M. Mescheder, Marc Pollefeys, and Andreas Geiger. Convolutional occupancy networks. *ArXiv*, abs/2003.04618, 2020.
- [23] Tahir Rabbani and Frank van den Heuvel. Efficient hough transform for automatic detection of cylinders in point clouds. 2005.
- [24] Daxuan Ren, Jianmin Zheng, Jianfei Cai, Jiatong Li, Haiyong Jiang, Zhongang Cai, Junzhe Zhang, Liang Pan, Mingyuan Zhang, Haiyu Zhao, and Shuai Yi. Csg-stump: A

- learning friendly csg-like representation for interpretable shape parsing. *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 12458–12467, 2021.
- [25] Ruwen Schnabel, Roland Wahl, and R. Klein. Efficient ransac for point-cloud shape detection. *Computer Graphics Forum*, 26, 2007.
 - [26] Gopal Sharma, Rishabh Goyal, Difan Liu, Evangelos Kalogerakis, and Subhransu Maji. Csgnet: Neural shape parser for constructive solid geometry. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5515–5523, 2017.
 - [27] Gopal Sharma, Difan Liu, Evangelos Kalogerakis, Subhransu Maji, Siddhartha Chaudhuri, and Radomir Mvech. Parsenet: A parametric surface fitting network for 3d point clouds. *ArXiv*, abs/2003.12181, 2020.
 - [28] Binil Starly, Akshay G. Bharadwaj, and Atin Angrish. Fabwave cad repository categorized part classes - cad 1 through 15 classes (part 1/3). 2019.
 - [29] Shubham Tulsiani, Hao Su, Leonidas J. Guibas, Alexei A. Efros, and Jitendra Malik. Learning shape abstractions by assembling volumetric primitives. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1466–1474, 2016.
 - [30] Karl D. D. Willis, Yewen Pu, Jieliang Luo, Hang Chu, Tao Du, J. Lambourne, Armando Solar-Lezama, and Wojciech Matusik. Fusion 360 gallery: A dataset and environment for programmatic cad reconstruction. *ArXiv*, abs/2010.02392, 2020.
 - [31] Rundi Wu, Chang Xiao, and Changxi Zheng. Deepcad: A deep generative network for computer-aided design models. *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 6752–6762, 2021.
 - [32] Rundi Wu, Yixin Zhuang, Kai Xu, Hao Zhang, and Baoquan Chen. Pq-net: A generative part seq2seq network for 3d shapes. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 826–835, 2019.
 - [33] Fenggen Yu, Qimin Chen, Maham Tanveer, Ali Mahdavi Amiri, and Hao Zhang. D²csg: Unsupervised learning of compact csg trees with dual complements and dropouts, 2023.
 - [34] Fenggen Yu, Zhiqin Chen, Manyi Li, Aditya Sanghi, Hooman Shayani, Ali Mahdavi-Amiri, and Hao Zhang. Capri-net: Learning compact cad shapes with adaptive primitive assembly. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 11758–11768, 2021.