

结合隐式几何编码与 Lipschitz 线性约束的 保守包围盒构造方法

张冰钰^{1,2,3}, 况立群^{1,2,3}, 熊风光^{1,2,3}, 孙凡淑^{1,2,3}, 焦世超^{1,2,3}

(1. 中北大学计算机科学与技术学院, 山西 太原 030051;
2. 机器视觉与虚拟现实山西省重点实验室, 山西 太原 030051;
3. 山西省视觉信息处理及智能机器人工程研究中心, 山西 太原 030051)

摘要: 目前主流的包围盒方法在三维场景渲染、光线追踪和碰撞检测等任务中广泛应用,但在拟合复杂几何形状时存在空间利用率低、拟合精度不足等问题,难以确保严格的保守性,并在降低误检率方面仍有改进空间。为解决上述问题,提出一种结合隐式几何编码与 Lipschitz 约束的保守包围盒构造方法,隐式几何编码通过位置编码将输入坐标映射至高维空间,从而捕捉局部及全局的几何信息,提升包围盒的适应性;随后,引入可训练的 Lipschitz 线性约束层,动态调整 Lipschitz 常数以控制梯度变化,并结合 Lipschitz 正则化损失与动态加权交叉熵损失,在优化边界拟合的同时降低假阳性率。实验结果表明,该方法在多个三维模型上均能实现假阳性率为 0,且相比基准方法,误检率最高降低 3.1%,单条光线查询方法提高 1.7 ms,为高精度保守包围盒拟合提供了一种高效、稳健的解决方案。

关键词: 保守包围盒; Lipschitz 约束; 隐式几何编码; 光线追踪; 碰撞检测

中图分类号: TP 391.41

DOI: 10.11996/JGj.2095-302X.2026010152

文献标识码: A

文章编号: 2095-302X(2026)01-0152-10

Conservative enclosing box construction algorithm based on implicit geometric coding with Lipschitz linear constraints

ZHANG Bingyu^{1,2,3}, KUANG Liqun^{1,2,3}, XIONG Fengguang^{1,2,3}, SUN Fanshu^{1,2,3}, JIAO Shichao^{1,2,3}

(1. School of Computer Science and Technology, North University of China, Taiyuan Shanxi 030051, China;

2. Shanxi Provincial Key Laboratory of Machine Vision and Virtual Reality, Taiyuan Shanxi 030051, China;

3. Shanxi Province's Vision Information Processing and Intelligent Robot Engineering Research Center, Taiyuan Shanxi 030051, China)

Abstract: Currently the mainstream enveloping box methods are widely used in 3D scene rendering, ray tracing, and collision detection tasks; however, these methods suffer from the problems of low space utilization and insufficient fitting accuracy in fitting complex geometries, which are difficult to ensure strict conservatism and still have room for improvement in reducing false detection rates. To address these issues, a conservative bounding-box construction method combining implicit geometric coding and Lipschitz constraints was proposed. Implicit geometric coding mapped the input coordinates to a high-dimensional space via position coding, thus capturing local and global geometric information and improving bounding-box adaptability. A trainable Lipschitz-constrained linear layer was introduced to dynamically adjust Lipschitz constants control gradient changes, and Lipschitz regularization loss was combined with dynamically weighted cross-entropy loss to reduce the FP rate while optimizing the boundary fitting. The experimental results demonstrated that the method can achieve a false-negative rate of 0 on multiple 3D models

收稿日期: 2025-05-08; 定稿日期: 2025-09-17; 通信作者: 况立群, E-mail: kuang@nuc.edu.cn

Received: 8 May, 2025; Finalized: 17 September, 2025; Corresponding author: KUANG Liqun, E-mail: kuang@nuc.edu.cn

基金项目: 国家自然科学基金(62272426); 山西省科技重大专项计划“揭榜挂帅”项目(202201150401021); 山西省基础研究项目(202303021212189)

Foundation items: National Natural Science Foundation of China (62272426); Shanxi Provincial Science and Technology Major Special Programs “Listed and Commanded” Project (202201150401021); Basic Research Program of Shanxi Province (202303021212189)

and reduce the false-detection rate by up to 3.1% compared to the benchmark method, and improve the single-ray query method by 1.7 ms, providing a highly efficient and robust solution for high-precision conservative bounding box fitting.

Keywords: conservative bounding box; Lipschitz constraint; implicit geometric encoding; ray tracing; collision detection

随着三维计算机图形学的快速发展, 三维场景的高效渲染越来越重要。在三维场景渲染中, 包围盒是一种常用的加速结构, 广泛应用于光线追踪和碰撞检测等任务, 通过构建简单的几何形状来包围物体, 简化计算并提高效率^[1-2]。目前主流的包围盒结构包括轴对齐包围盒(Axis Aligned Bounding Box, AABB)、有向包围盒(Oriented Bounding Box, OBB)、球形包围盒(Bounding Sphere, Sphere)以及 k-离散有向包围盒(k-Discrete Oriented Polytopes, k-DOP)等, 如图 1 所示。最基础的包围盒结构为 AABB, 由于其边界与坐标轴对齐, 结构相对简单, 适合于静态场景和规则几何体^[3]。OBB 结构^[4]允许包围盒根据物体的形状和方向自由旋转, 能够更紧密地包围复杂物体, 显著提高拟合精度。Sphere 结构^[5]采用球体来包围物体, 具有极高的计算效率和旋转不变性, 在动态场景下无需频繁更新, 但对于细长或不规则物体的包围效果较差, 空间利用率较低。k-DOP 结构^[6]选取 k 个固定方向的平面来构建包围体, 相较于 AABB 具有更高的拟合精度, 同时避免了 OBB 复杂的旋转计算, 随着方向数量 k 的增加, 拟合效果进一步提升。

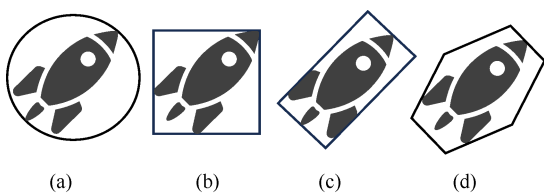


图 1 经典包围盒类型((a) 包围球; (b) AABB 包围盒; (c) OBB 包围盒; (d) k-DOP 包围盒)

Fig. 1 Classic bounding box types ((a) Bounding Sphere; (b) AABB bounding box; (c) OBB bounding box; (d) k-DOP bounding box)

针对单一包围盒在复杂场景下的局限性, 研究者提出了混合包围盒结构, 以结合简单包围盒的快速构建能力与复杂包围盒的高效剔除特性, 从而提升碰撞检测与光线追踪渲染的效率^[7-8]。其中, Sphere-AABB-OBB 混合层次包围盒结构通过外层 Sphere 和 AABB 进行快速排除, 并在此基础上利用内层 OBB 进行精确相交检测, 以优化计算开销

并提高检测准确性^[9]。但在处理高度不均匀的几何结构时, 可能导致层次划分不均, 进而影响剔除效率, 此外, AABB 在面对斜向或细长物体时, 受到表面积启发成本的影响较大, SABINO 等^[10]提出 ODOPs(Orthogonal Discrete Oriented Polytopes)结构通过动态优化 OBB 方向选择策略, 有效减少冗余空间, 提高包围盒的空间利用率。

尽管上述包围盒结构在光线追踪渲染与碰撞检测中显著提升了计算效率, 但在复杂几何场景下, 其空间利用率较低, 导致资源浪费并限制整体性能优化。为克服这一问题, 研究者引入神经网络方法, 以提高包围盒拟合的精度和适应性。BEHERA 和 MISHRA^[11]提出一种基于神经方向距离场的物体表示方法, 并结合包围盒优化策略, 提升光线追踪任务中的求交精度, 以致计算复杂度较高, 面临较大的存储和计算负担。为了提高计算效率, N-BVH(Neural Bounding Volume Hierarchy)多分辨率哈希网格神经模型对三维模型表面进行高效编码, 并快速判断光线与三维模型表面的相交情况, 同时获取交点处的各种属性^[12]。但难以有效满足包围盒的保守性约束, 即所构建的包围盒必须完全包含目标物体的几何结构, 从而确保任何可能与物体相交的点和光线等都不会被错误剔除^[13]。在快速预览渲染或早期剔除阶段, 研究者更关注计算速度和光线求交效率, 而不是每条光线的精确交点。因此, 允许包围盒不完全包围物体的几何边界。

对于要求图像质量的终渲染阶段, 严格遵守保守性约束仍然不可或缺。在高精度的阴影计算、全局光照计算和可见性分析等模块中, 严格的保守性约束尤为重要, 如果包围盒未能完整覆盖物体, 易导致渲染错误, 影响最终效果的真实性和完整性。因此, 如何在保证保守性的前提下, 提高拟合精度和查询效率, 仍是包围盒构建中亟待解决的核心问题。在几何处理领域, 包围笼的生成方法提供了保守的边界控制机制, 使外部控制点可驱动目标模型的整体状态变化^[14]。XU 和 HARADA^[15]提出的 DRFC(Deforming Radiance Fields with Cages)包围笼方法将该结构引入隐式神经辐射场, 通过控制笼点变形

联动内部场景几何与颜色,拓展了笼形结构在神经建模中的适用性。神经包围笼方法^[16]进一步通过隐式表示增强变形表达力,在保持整体结构的同时保留高频细节,克服了传统笼子大变形下细节丢失和表面失真的问题,与本文的神经包围盒在思想上具有一致性。此外,在光线追踪与渲染领域,LIU等^[17]利用不对称损失函数引导网络学习保守性约束,通过对假阴性赋予更高的惩罚权重,有效降低假阴率,从而提高边界拟合的完整性。为了实现严格保守性,LUDWIG和CAMPEN^[18]进一步提出在神经网络拟合过程中引入自适应仿射算法,动态调整网络参数,提高模型对复杂边界的适应能力。现有的方法模型结构大多仅依赖多层感知机(MultiLayer Perceptron, MLP)进行保守包围盒的拟合,但在处理高维复杂几何体时,模型的拟合精度与计算效率仍有较大提升空间,难以在严格控制假阴率的同时有效优化假阳率。因此,如何在保证假阴率为零的前提下,进一步优化假阳率并提升计算效率,仍是神经包围盒构建中的关键挑战。

为了更精准地拟合三维模型的边界,本文提出一种隐式几何编码器与 Lipschitz 约束相结合的神神经包围盒方法。利用隐式特征提取网络对输入数据进行全局特征提取,位置编码器将输入的三维坐标映射至高维空间,增强对复杂几何结构的表征能力。随后,引入 Lipschitz 线性约束层,通过控制梯度变化提高模型的稳定。最后,经过解码过程,将隐式特征转换为空间占据概率,从而实现三维模型边界的高效拟合。本文的主要贡献如下:

1) 为了解决传统包围盒方法在复杂几何形状拟合中的低精度和局部细节缺失问题,提出了结合隐式几何编码的神经包围盒结构,通过隐式特征提取全局几何特征,Fourier 位置编码增强网络对复杂几何结构的表达能力,并利用 Tinycudann (Tcnn) 加速网络提取局部细节,提高模型对细微边界的拟合精度,从而实现更紧密和高效的包围盒构建。

2) 针对神经隐式表示方法在边界拟合过程中梯度不稳定、泛化能力不足和过拟合的问题,引入可训练的 Lipschitz 约束线性层,通过施加连续性约束提高数值稳定性。同时,结合 Lipschitz 正则化损失控制模型的 Lipschitz 常数,避免梯度爆炸或梯度消失,并采用动态加权的交叉熵损失函数适应不同类别样本的不均衡性,从而在降低假阳率的同时,提升模型的泛化能力并抑制过拟合现象。

1 方法

本文提出基于隐式几何编码(ImplicitNet)和 Lipschitz 约束的包围盒结构,使用神经网络来学习紧密和保守的三维模型边界,网络框架如图 2 所示。架构由隐式几何编码、Lipschitz 线性约束和解码器构成。首先,根据对应的查询方式随机生成多个采样点,并对其坐标进行 Fourier 位置编码,实现高维映射,通过 3 层 MLP 实现隐式特征提取;然后,将提取的特征经过多层 Lipschitz 约束线性变换,并通过可训练的 Lipschitz 常数动态控制每层的 Lipschitz 连续性,模型的 Lipschitz 常数通过额外的正则化项进行优化^[19],抑制梯度爆炸并提高

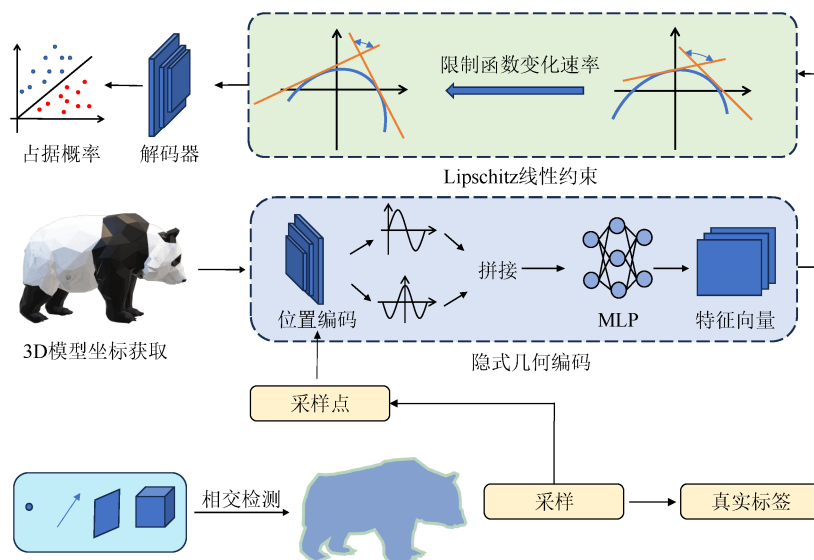


图 2 模型总体网络架构图

Fig. 2 Model general network architecture diagram

数值稳定性; 最终, 经过解码器将特征映射至标量场表示得到空间点的占据概率。占据概率与采样点的真实标签作比较, 依据高于设定阈值的点集构建包围盒。

1.1 隐式几何编码器与解码器

MLP 的输入通常为低维特征, 仅包含原始三维坐标, 其表达能力有限, 需要大量参数和长时间训练才能学习复杂形状, 并且在训练过程中优先学习低频信息, 而对高频细节的学习能力较弱。这导致模型难以有效表示复杂几何结构, 尤其是在边界区域, 最终表现为边界模糊化, 无法精确拟合物体的真实轮廓。因此, 本文提出隐式几何编码器, 并引入 Fourier 位置编码, 来克服 MLP 在低维坐标输入下难以捕捉高频信息的局限。该方法能够使神经网络更有效地学习高频函数的细节^[20], 提升模型的表达能力。随后再经过 MLP 进一步学习隐式几何信息^[21]。隐式几何编码结构如图 3 所示。

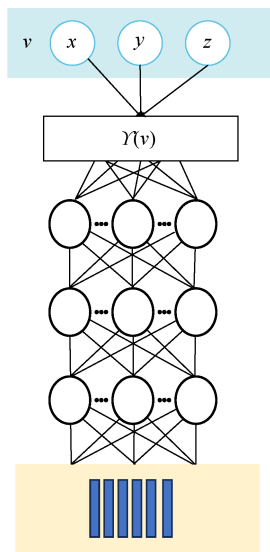


图 3 隐式几何编码模块
Fig. 3 ImplicitNet module

首先将输入的三维坐标信息映射到一个高维的潜在特征空间, 提取和表达三维几何结构, 给定三维空间中的输入点 $v=(x, y, z)$, 对其施加 Fourier 编码 $\gamma(v)$, 并将其转换为高维特征表示, 即

$$\gamma(v) = [v, \{\sin(2^i \pi x)\}_{i=0}^N, \{\cos(2^i \pi x)\}_{i=0}^N] \quad (1)$$

位置编码作为特征映射, 将输入的三维坐标映射到高维空间, 但其本质上仅是固定的变换, 并不具备特征学习能力。为了进一步提取和组织有效的几何信息, 高维特征经过一个全连接 MLP, 通过非线性变换对输入特征进行学习和优化^[22]。具体而

言, 该 MLP 由 3 层全连接神经网络组成, 每层包含 64 个神经元, 并采用 ReLU 作为激活函数, 从而增强模型的非线性表达能力。最终, 编码器输出一个 64 维的潜在几何特征向量, 该向量编码了输入点的空间占据信息、局部几何属性及隐式结构信息, 即

$$F = f_{\theta}(\gamma(v)) \in \mathbb{R}^{64} \quad (2)$$

式中: f_{θ} 为可由可训练参数 θ 表示的神经网络映射函数。

在解码阶段, 采用轻量级 MLP 从编码特征中提取空间占据信息。具体而言, 解码器将编码特征作为输入, 并映射到一个标量输出, 其为该点的占据概率。出于时间成本的考量, 本文采用一个 2 层全连接 MLP, 每层包含 32 个神经元, 并在最后施加 Sigmoid 激活函数, 得到空间占据信息。编码器与解码器的 MLP 均基于 Tcnn 高效训练。

1.2 Lipschitz 约束线性层

由于普通线性层的参数不受约束, 模型可能出现剧烈的梯度波动, 从而导致过拟合或边界震荡, 尤其在三维模型的神经表示学习中, 边界区域的决策往往缺乏稳定性, 导致假阳率和假阴率较高, 影响包围盒的精度。此外, 传统神经网络在处理不同尺度的三维数据时, 易受到输入数据尺度变化的影响, 从而降低泛化能力。Lipschitz 约束结构如图 4 所示。

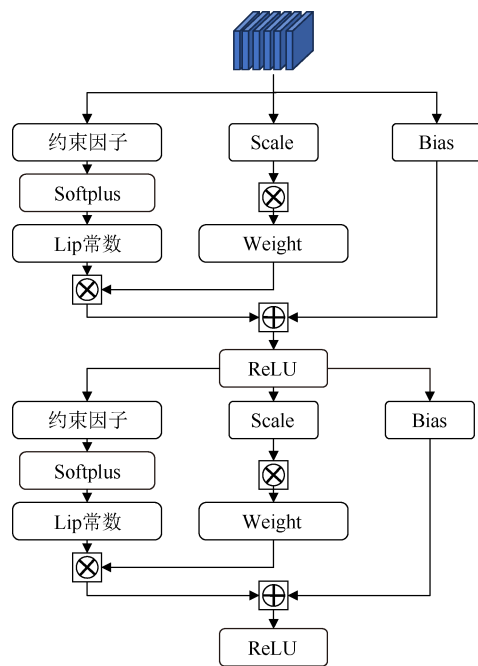


图 4 Lipschitz 约束模块
Fig. 4 Lipschitz constraints module

Lipschitz 线性约束方法能够较好地解决上述问题, 通过对权重范数施加约束, 使得网络 Lipschitz 常数保持在合理范围内, 从而限制参数的变化幅度, 提高梯度更新的稳定性^[23-24]。然而, 在三维模型的神经表示学习中, 计算时间和效率是一个至关重要的因素, 传统的 Lipschitz 约束网络往往通过硬约束方式来控制网络的 Lipschitz 常数, 这种方法虽然有效, 但往往伴随着较高的计算成本。考虑到时间成本, 本文采用高效的 Lipschitz 线性约束模块, 不同于传统的 Lipschitz 约束网络, 本文模型能够在训练过程中动态学习 Lipschitz 常数, 从而在保证计算效率的同时提升几何表示能力。每层的 Lipschitz 常数设为可训练参数, 并在训练过程中自动调整, 以适应不同的数据和任务。

具体来说, Lipschitz 模块旨在限制函数的输出变化率不超过某个常数, 从而保证模型对输入扰动的鲁棒性和稳定性。通过结合动态权重缩放与可学习的 Lipschitz 常数, 在保证特征表达能力的同时约束函数变化的全局上界。对于函数 $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$, 存在任意的 x_1 和 x_2 , 其 Lipschitz 常数 L 为以下函数的最小值

$$\|f(x_1) - f(x_2)\| \leq L \|x_1 - x_2\| \quad (3)$$

在范数 L_∞ 下, 全连接层 $f(x)$ 的常数的上界为矩阵 W 的最大行绝对和, 计算权重矩阵的每行绝对值和, 并取最大值作为初始 Lipschitz 约束 c , 得到 Lipschitz 常数。再通过 Softplus 函数将 Lipschitz 常数转换为正值, 即

$$c = \max_i \|W_i\|_1 = \max_i \sum_{j=1}^n |W_{ij}| \quad (4)$$

$$lipc = \text{Softplus}(c) = \ln(1 + e^c) \quad (5)$$

通过反向传播, 网络会根据损失函数对 Lipschitz 常数 $lipc$ 的影响进行优化, 在每次梯度更新时, $lipc$ 的值通过更新参数 c 进行调整, 限制权重矩阵的变化范围。Softplus 变换的引入避免硬约束带来的数值不稳定性, 同时允许网络动态调整 Lipschitz 约束的强度。然后对权重矩阵的每一行计算其绝对和, 设第 i 行的绝对和为 s_i , 为确保该行的权重不超过 Lipschitz 上界, 计算其缩放因子得到

$$s_i = \sum_1^i |W_{ij}| \quad (6)$$

$$scale_i = \min(lipc/s_i, 1) \quad (7)$$

在 Lipschitz 线性约束中, 通过对每一行的权重按照特定因子进行缩放, 得到用于线性映射的有效权重 \tilde{W} , 从而确保该层在 L_∞ 范数下的 Lipschitz 常数为 $lipc$ 。将 Lipschitz 线性约束层与 ReLU 激活函数顺序连接后, 整体模块的 Lipschitz 常数保持不变。

1.3 损失函数

在多标签分类任务中, 面对正负样本不平衡的情况, 传统的二元交叉熵损失函数常常无法有效处理正负样本的不平衡, 导致训练过程中模型对少数类别的预测能力较差。为了解决这一问题, 本文提出一种结合动态加权的二元交叉熵损失函数和 Lipschitz 正则化项的方法, 动态调节正负样本的权重, 并引入 Lipschitz 连续性约束, 进一步提高模型在复杂样本分布下的表现。

动态加权的二元交叉熵损失函数通过引入样本类别的不同权重, 使得正负样本的损失对模型训练的影响不对称。模型在训练时, 根据样本的类别调整损失的大小, 在正负样本数量极为不平衡的情况下, 增加少数类别的权重, 以确保其对最终损失函数的贡献更大, 从而缓解正负样本不平衡带来的影响。在加权的基础上, 损失函数仍采用传统的二元交叉熵形式进行计算。对于每个样本, 计算正类和负类的交叉熵损失, 并根据类别权重进行加权, 最终的动态加权的二元交叉熵损失为

$$L_{\text{BCEwithWeight}} = -\frac{1}{N} \sum_{i=1}^N (\omega_+ y_i \log \hat{y}_i + \omega_- (1 - y_i) \log(1 - \hat{y}_i)) \quad (8)$$

式中: 2 个类别权重参数 ω_+ 和 ω_- 分别表示正样本和负样本的权重, 正类权重 ω_+ 固定为 1, 负类权重 ω_- 初始为 1, 初始控制参数 d 为 0, 每经历 250 000 次迭代进行一次调整, 当第 k 次调整时, 负类权重为

$$\omega_-^{(k)} = \frac{1}{20k}, k \geq 1 \quad (9)$$

随着训练进行, 负类权重呈单调递减趋势, 导致模型逐渐降低对负类样本的惩罚力度, 该策略迫使模型在训练后期更加关注正类样本的学习, 从而缓解类别分布不平衡带来的偏差。为了进一步提高模型的鲁棒性, 并抑制模型参数的过度变化, 引入了 Lipschitz 正则化项, 其损失由所有 Lipschitz 约束层的 Lipschitz 常数乘积计算得到, 如式(10)所示。最终, Lipschitz 正则化项与动态加权的二元交叉熵损失函数结合, 形成总损失函数为

$$L_{\text{Lipschitz}} = \prod_{l=1}^L \|W_l\|_{\text{lipc}} \quad (10)$$

$$L_{\text{total}} = L_{\text{BCEwithWeight}} + \lambda * L_{\text{Lipschitz}} \quad (11)$$

式中: L 表示网络层数; $\|W_l\|_{\text{lipc}}$ 为对应网络层的 Lipschitz 常数; λ 表示 Lipschitz 约束的权重超参数。

2 实验分析

2.1 数据集及评价指标

为验证基于 Lipschitz 约束的神经包围盒拟合结果的有效性, 本文在 ShapeNet 数据集^[25]及斯坦福大学三维扫描存储库数据集进行实验, 部分模型如图 5 所示。ShapeNet 是一个大规模的三维形状数据集, 包括 ShapeNetCore 和 ShapeNetSem 2 个子集。其中, ShapeNetCore 包含约 55 000 个三维模型, 涵盖 55 个常见类别。而 ShapeNetSem 包含超过 270 000 个三维模型, 并提供了更丰富的语义标签, 包括物体属性和材料信息等。斯坦福大学的三维扫描存储库的三维模型主要由激光扫描仪获取, 涵盖了从简单物体到复杂雕像的多样化几何结构, 为渲染算法的开发和验证提供了丰富的测试样本。存储库中共有 9 个模型, 涵盖宗教雕像、动物、手稿和文化符号等类别。



图 5 数据集模型图片((a) ShapeNet 数据集部分模型;
(b) 斯坦福大学三维扫描存储库)

Fig. 5 Dataset model images ((a) ShapeNet partial models;
(b) The Stanford 3D scanning repository)

三维模型均经过体素化处理作为输入, 选取假阴率(False Negative Rate, FNR)、假阳率(False Positive Rate, FPR)和单条光线完成一次查询所用的时间(T_{avg})作为主要评估指标。其中, 假阴率指真实属于目标物体内部的空间区域中, 被模型错误判断为外部区域的比例, 即

$$FNR = \frac{FN}{TP + FN} \quad (12)$$

式中: TP 表示模型正确预测为占据的点数; FN 表示真实占据但模型预测为非占据的点数。

假阳率表示预测为物体内部但实际为空区域所占的比例, 反映模型拟合的紧致程度, 假阳率越低, 表明冗余空间越少, 效率越高, 即

$$FPR = \frac{FP}{TN + FP} \quad (13)$$

式中: TN 表示模型正确预测为非占据的点数; FP 表示真实非占据但模型预测为占据的点数。

单条光线查询耗时用于评估空间查询任务的计算效率, 指模拟一次完整光线与包围结构交互所需的平均时间, 通过多次计算取均值, 保证其统计稳定性, 即

$$T_{\text{avg}} = \frac{1}{1000} \sum_{i=1}^{1000} t_i \quad (14)$$

式中: t_i 表示光线查询的总耗时。

上述指标均通过比较预测结果与真实三维体素数据之间的差异获得。其中, 预测输出采用固定阈值进行二值化判断, 真实标签来源于标准体素模型, 通过逐点查询得出每个空间位置的真实占据状态, 两者的逐点差异被统计为假阴、假阳或正确预测, 进而计算上述比例指标。

为了模拟实际渲染任务中光线与目标物体的交互行为, 本文在评估阶段引入基于空间点的光线查询机制。光线随机生成, 每条光线均在其路径上进行等间距采样 500 个点, 形成一组具有方向性的空间点集, 并批量输入模型, 以计算单条光线的查询时间。为避免光线方向对评估结果产生偏差, 本文采用统一采样范围与步长, 保证每轮评估中光线数量与采样密度一致, 并采用 1 000 条光线的平均结果作为单条光线查询时间, 避免因少量光线的特殊性导致误差放大。

2.2 实验设置

硬件环境为: Intel(R) Xeon(R) Gold 6152 CPU+NVIDIA GeForce RTX 3090+30 GB RAM。

软件环境为: Ubuntu 20.04+CUDA 12.1+PyTorch 2.3.0+Python 3.9.20。

模型训练最大迭代次数为 50 000 000 轮, 训练数据由 50 000 个区域组成, 每个区域的采样点数设定为 500。优化器采用 Adam, 学习率设为 0.000 1, 动态加权的二元交叉熵损失函数中 Class Weight 每 250 000 次迭代进行一次更新, 损失函数中 Lipschitz 正则化的权重为 0.01。

2.3 对比实验

为了评估所提出的包围盒结构在拟合三维模型任务中的表现, 本文使用 ShapeNet 数据集和斯坦福大学三维扫描存储库进行实验。对于 ShapeNet 数据集, 使用全部 55 个常用类别, 并在每类中随机选取 3 个几何复杂度差异显著的代表性模型进行测试。由于该数据集规模庞大, 若在全部模型上运行将带来极高的计算成本, 故采用此抽样策略, 在保证几何多样性的同时控制计算成本。对于斯坦福大学三维扫描存储库, 使用了其中全部 9 个三维模型。实验中与多个经典方法进行了点、盒、平面和光线查询, FNR 严格为 0, FPR 取所有模型的平均值, 结果见表 1 和图 6。实验选取 AABB, OBB, Sphere 和 k-DOP 共 4 种经典包围盒结构, 以及结合神经网络的优化包围盒的 3 种方法 NN^[17], NBE^[17]与 SCNI^[18], 还有 DRFC^[12]包围笼生成方法作为对比方法, 并在 Point, Box, Plane 和 Ray 共 4 种不同的查询方法下评估模型性能。其中, NN

方法基于 MLP, 优化网络参数来提高对复杂形状的适应能力; NBE 在 NN 方法基础上加入早停机机制; SCNI 方法在网络训练拟合过程中运用自适应仿射技术。实验结果表明, 在经典包围盒结构中, k-DOP 结构相较于 AABB, OBB 和 Sphere 结构能够提供更紧致的包围盒, 但存在较大的误差。在经过神经网络优化后, 各包围盒方法 FPR 明显降低, 且可整体保持在 3%至 10%的较低水平, 表现出更强的精度一致性与拟合稳定性。与神经网络方法相比, 本方法在 4 种查询方法上误差均有下降。其中, 在点查询中, 相较于经典包围盒算法误差最高降低 74.2%, 相较于神经网络方法最高降低 2.7%; 在盒查询中, 相较于经典包围盒算法误差最高降低 52.6%, 相较于神经网络方法最高降低 3.4%; 在平面查询中, 相较于经典包围盒算法误差最高降低 60.6%, 相较于神经网络方法最高降低 3.7%; 在光线查询中, 相较于经典包围盒算法误差最高降低 65%, 相较于神经网络方法最高降低 2.4%。

表 1 不同包围盒方法的 FPR 对比

Table 1 Comparison of FPR for different methods of bounding boxes

方法	ShapeNet 数据集				斯坦福大学三维扫描存储库			
	Point	Box	Plane	Ray	Point	Box	Plane	Ray
AABB	49.3	45.4	55.2	69.2	66.6	51.2	49.6	70.2
OBB	42.1	49.2	58.4	63.3	51.2	41.9	45.3	65.6
Sphere	54.2	56.7	64.9	70.3	77.4	56.3	59.3	71.9
k-DOP	34.5	40.9	41.6	62.3	45.2	51.8	42.3	68.3
NN	4.2	6.1	7.6	6.9	5.9	7.5	8.2	8.7
NBE	3.9	5.5	6.2	7.5	5.5	7.5	8.4	8.3
SCNI	4.2	4.6	6.4	7.9	5.6	7.8	9.6	9.3
DRFC	11.6	12.4	23.1	29.9	10.4	16.6	26.7	32.8
Ours	3.5	4.1	4.3	5.5	3.2	4.4	5.9	6.9

注: 加粗数据表示最优值。

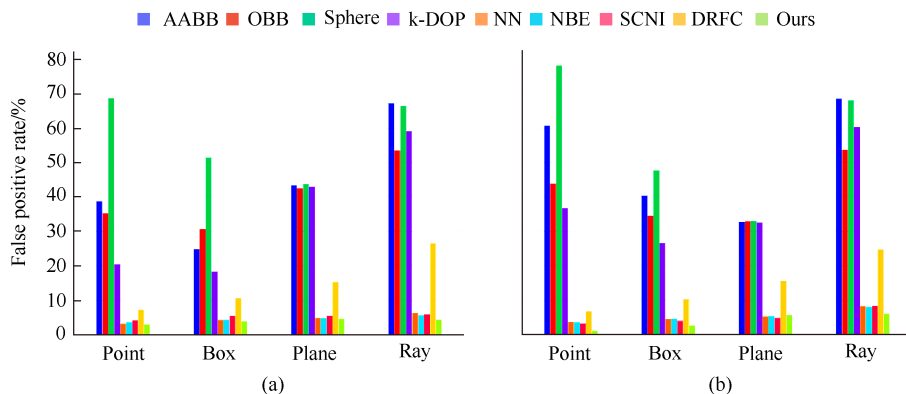


图 6 各种包围盒方法的 FPR 对比图((a) ShapeNet 数据集; (b) 斯坦福大学三维扫描存储库)

Fig. 6 Comparison of FPR for each bracketed box methodology ((a) ShapeNet dataset; (b) Stanford 3D scanning repository)

本文针对 Bunny 模型的 4 种查询方式(Point, Box, Plane, Ray)与基准方法 NBE 进行对比实验, 记录不同迭代次数下的 FNR 变化情况。由于迭代次数过多, 因此每 50 000 次迭代选取一次数据进行记录与绘制, 如图 7 所示。在训练过程中, 本文方

法相较于基准模型展现出更优的收敛性, FN 率始终保持在较低水平。当 FNR 逐步收敛至 0 时, 基准模型仍存在显著的漏检现象, 而本文方法在不同查询策略下均能稳定维持 FNR 为 0。说明本文方法能够有效地降低 FNR, 提高包围盒的适用性和精度。

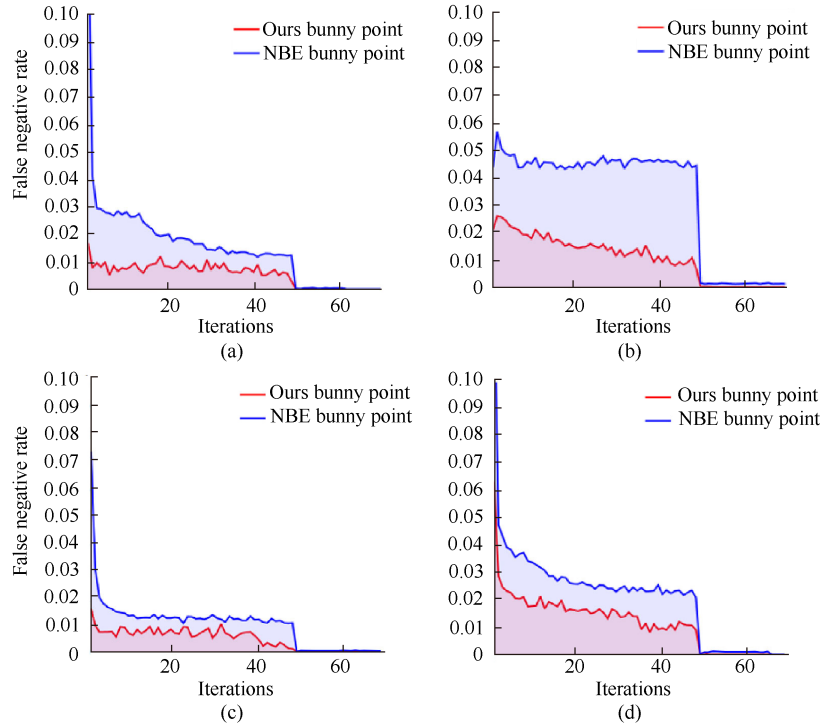


图 7 基准模型与本文方法 FNR 对比

Fig. 7 Comparison of FNR between the benchmark model and the method in this paper ((a) Point; (b) Plane; (c) Box; (d) Ray)

单条光线查询的时间对比见表 2。神经网络方法在查询时间上无法与经典包围盒方法匹敌, 但是与目前主流的 k-DOP 包围盒的查询速度相当, 并且神经网络方法在紧密性方面有着更大的优势。本文方法相较于其他基于神经网络拟合包围盒的优化方法, 其计算开销得到了控制, 相比于 NN 方法, 本文方法的查询时间下降了 25.56 ms, 相比于基准 NBE 方法, 时间下降了 1.70 ms。实验表明, 本文方法在时间与精度之间取得了良好的平衡, 在高精度包围盒拟合任务中具备一定的应用价值。

表 2 单条光线查询时间对比/ms

Table 2 Comparison of single ray query time/ms

方法	时间
AABB	2.83
OBox	2.85
Sphere	2.49
k-DOP	43.26
NN	71.38
NBE	47.52
SCNI	55.85
DRFC	65.44
Ours	45.82

注: 加粗数据表示最优值。

2.4 实验可视化

基于神经网络的包围盒方法 SCNI, NNE, NN 以及本文提出的包围盒方法可视化分析如图 8 所示。

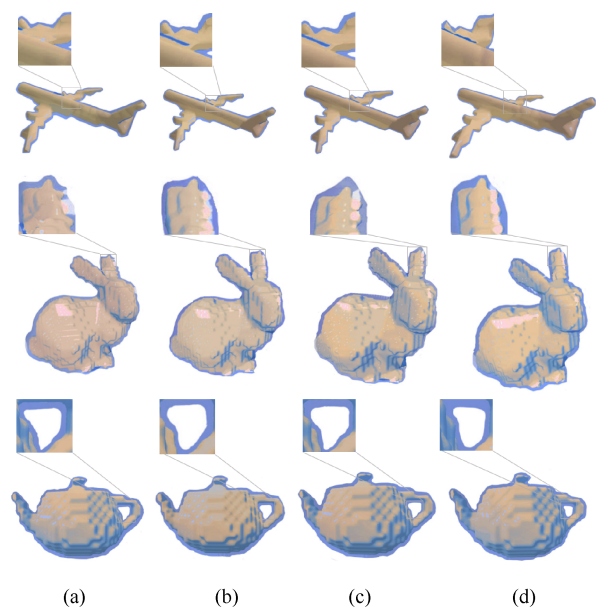


图 8 模型包围盒可视化对比

Fig. 8 Visual comparison of model bounding boxes ((a) Ours; (b) SCNI; (c) NN; (d) NBE)

其中, Airplane 模型结构规则、Bunny 模型细节丰富、Teapot 模型曲面复杂且具有典型的中空结构, 涵盖了不同类型的三维结构形态, 具有代表性。SCNI, NNE 和 NN 方法在对复杂结构建模时, 均存在一定程度的冗余包围区域, 尤其在细节边界或曲率较大的区域, 拟合边界与真实模型之间仍存在可见偏差。本方法在多个三维模型上均表现出稳定的拟合效果, 所生成的神经包围盒能够有效减少冗余空间, 紧密包围原模型, 同时使得 FNR 为 0, 并有效降低 FPR。实验结果表明, 本方法不仅适用于简单规则的物体, 对于复杂形态物体同样具备良好的适应能力, 在高精度包围盒拟合任务中展现了较高的应用价值。

2.5 消融实验

为了评估 Lipschitz 线性约束与隐式几何编码在神经包围盒中的作用, 验证本文方法 2 个模块的有效性, 针对 Bunny 模型的光线查询任务, 研究隐式几何编码器和 Lipschitz 约束在降低 FPR 和光线查询时间上的作用。消融实验结果见表 3。以 Baseline 方法作为对比, 引入隐式几何编码器后, FPR 降低了 1.2%, 时间减少了 3.1 ms, 说明隐式几何编码器在提高包围盒拟合精度的同时, 能够优化计算效率。仅引入 Lipschitz 约束后, FPR 降低了 1.7%, 计算时间下降了 2.7 ms, 说明 Lipschitz 约束有助于提高模型的收敛稳定性和泛化能力, 使得包围盒拟合更加精确。

表 3 消融实验结果对比

Table 3 Comparison of ablation experiment results

方法	FPR/%	时间/ms
Baseline	7.6	46.3
ImplicitNet	6.4	43.2
Lipschitz	5.9	43.6
ImplicitNet+Lipschitz	4.9	45.1

注: 加粗数据表示最优值。

当隐式几何编码器和 Lipschitz 约束联合使用时, FPR 相比 baseline 方法下降了 3.1%, 相比单独使用隐式几何编码器和 Lipschitz 方法也分别下降了 1.5% 和 1.0%, 表明两者的结合能够最大程度地减少错误包围, 提高拟合精度。此外, 本文方法相比 Baseline 方法计算效率提升了 1.7 ms。

消融实验表明隐式几何编码结构和 Lipschitz 约束在优化包围盒拟合精度方面均有积极作用, 两者的联合使用能够在降低 FPR 的同时, 保持较优的计算效率, 从而实现更高效且精确的神经包围盒

拟合。

2.6 参数敏感性实验

本文采用结合动态加权的二元交叉熵损失函数与 Lipschitz 正则化项的方法, 引导模型学习更加区分性强、边界平滑的隐式几何特征。为了平衡 2 种损失项对训练目标的影响, 引入超参数 λ 控制 Lipschitz 正则项在总损失中的权重。为系统评估不同 λ 设置对模型性能的影响, 在 ShapeNet 数据集的 Airplane 模型上进行实验, 将 λ 从 0 到 0.09 以 0.01 为间隔逐步变化, 并观察对应的假阳率 FPR 表现。如图 9 所示, 随着 λ 增大, 模型边界逐渐收敛、预测更平滑, 但正则过强时会限制模型的表达能力, 导致假阳率上升。实验结果表明, 当 $\lambda = 0.01$ 时, 模型在保持假阴率为 0 的同时取得最低的假阳率, 整体性能最优。因此, 本文在后续所有实验中统一采用该权重设置。

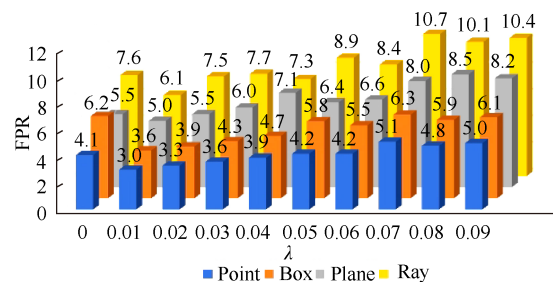


图 9 参数 λ 在 Airplane 模型上的性能比较

Fig. 9 Performance comparison of parameter λ on airplane model

同时, 为评估动态加权策略中参数 k 对模型性能的影响, 设计了若干种负类权重更新方案, 设置不同的权重衰减公式中的参数 c 来间接反映 k 对模型的影响, 从而调节负类权重下降速率, 实质上等价于对 k 变化趋势的控制。设置参数 c 分别为 20, 40, 60 和 80, 在训练过程中模拟不同的动态调整策略, 影响每次更新时负类权重的下降速度, 实验结果如图 10 所示, 随着参数 c 的增大, 4 类查询方

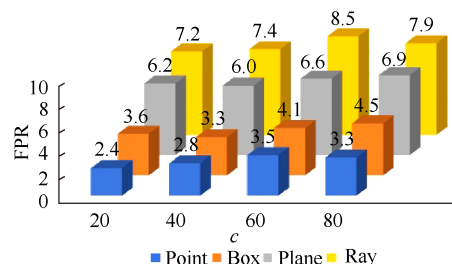


图 10 参数 c 在 Airplane 模型上的性能比较

Fig. 10 Performance comparison of parameter c on airplane model

式的 FPR 均呈现一定程度的上升趋势, 说明当负类权重下降过慢时, 模型更倾向于扩大预测区域, 导致更多负类样本被错误判为正类。实验结果表明, 参数 $c=20$ 时, 在维持零假阴率的同时, 实现了最小的假阳率, 显著提升模型的性能表现。

3 结论与展望

本文针对神经包围盒在光线查询任务中的计算开销和误检问题, 分析和总结了现有方法在不同几何结构下的适应性及其局限性。针对这些问题, 本文提出了一种结合 Lipschitz 线性约束与隐式几何编码结构的神经包围盒优化方法, 该方法能够在保证紧密包围目标模型的同时, 降低误检率, 并优化查询效率。实验结果表明, 所提出的方法在不同查询任务下均能保证 FNR 降低到 0, 形成保守的包围盒结构, 并在 FPR 与查询时间之间取得良好的平衡。由于本文研究聚焦于保守神经包围盒在静态三维模型中的高精度与高效率拟合, 当前工作主要基于静态场景展开。未来将进一步拓展方法在动态环境与复杂几何结构中的适应能力, 以增强其在实时渲染、光线追踪与物理仿真等实际应用中的广泛适用性与工程价值。

参考文献 (References)

- [1] MEISTER D, OGAKI S, BENTHIN C, et al. A survey on bounding volume hierarchies for ray tracing[J]. *Computer Graphics Forum*, 2021, 40(2): 683-712.
- [2] MEISTER D, KULKARNI P, VASISHTA A, et al. HIPRT: a ray tracing framework in HIP[J]. *Proceedings of the ACM on Computer Graphics and Interactive Techniques*, 2024, 7(3): 44.
- [3] 杨帆. 基于 B+树存储的 AABB 包围盒碰撞检测算法[J]. *计算机科学*, 2021, 48(S1): 331-333, 348.
YANG F. Collision detection algorithm of AABB bounding box based on B+ tree[J]. *Computer Science*, 2021, 48(S1): 331-333, 348 (in Chinese).
- [4] CHANG J W, WANG W P, KIM M S. Efficient collision detection using a dual OBB-sphere bounding volume hierarchy[J]. *Computer-Aided Design*, 2010, 42(1): 50-57.
- [5] 成居宝, 杜娟, 刘丽琴, 等. 基于数控机床特性的碰撞检测算法研究[J]. *组合机床与自动化加工技术*, 2020(8): 101-105, 110.
CHENG J B, DU J, LIU L Q, et al. Research on collision detection algorithm based on characteristics of CNC machine tools[J]. *Modular Machine Tool & Automatic Manufacturing Technique*, 2020(8): 101-105, 110 (in Chinese).
- [6] 刘超, 蒋夏军, 施慧彬. 一种快速的双重层次包围盒碰撞检测算法[J]. *计算机与现代化*, 2018(5): 6-10.
LIU C, JIANG X J, SHI H B. A fast collision detection algorithm based on dual bounding volume hierarchy[J]. *Computer and Modernization*, 2018(5): 6-10 (in Chinese).
- [7] CHAO W, ZHANG Z L, YONG L, et al. Improved hybrid bounding box collision detection algorithm[J]. *Journal of System Simulation*, 2019, 30(11): 4236-4243.
- [8] ZHAO W, QU H Y. Application of improved algorithm based on sphere OBB hybrid hierarchical bounding box in Teaching[J]. *Journal of Physics: Conference Series*, 2021, 1732(1): 012080.
- [9] 林菲, 邹玲, 张聪. 基于混合层次包围盒的快速碰撞检测算法[J]. *计算机仿真*, 2023, 40(9): 454-457.
LIN F, ZOU L, ZHANG C. A fast collision detection algorithm based on hybrid hierarchical bounding boxes[J]. *Computer Simulation*, 2023, 40(9): 454-457 (in Chinese).
- [10] SABINO R, VIDAL C A, CAVALCANTE-NETO J B, et al. Building oriented bounding boxes by the intermediate use of ODOPs[J]. *Computers & Graphics*, 2023, 116: 251-261.
- [11] BEHERA A P, MISHRA S. Neural directional distance field object representation for uni-directional path-traced rendering[C]//The 14th International Conference on Computing Communication and Networking Technologies. New York: IEEE Press, 2023: 1-6.
- [12] WEIER P, RATH A, MICHEL É, et al. N-BVH: neural ray queries with bounding volume hierarchies[C]//ACM SIGGRAPH 2024 Conference Papers. New York: ACM, 2024: 99.
- [13] ZESCH R S, MODI V, SUEDA S, et al. Neural collision fields for triangle primitives[C]//SIGGRAPH Asia 2023 Conference Papers. New York: ACM, 2023: 76.
- [14] STRÖTER D, THIERY J M, HORMANN K, et al. A survey on cage-based deformation of 3D models[J]. *Computer Graphics Forum*, 2024, 43(2): e15060.
- [15] XU T H, HARADA T. Deforming radiance fields with cages[C]//The 17th European Conference on Computer Vision. Cham: Springer, 2022: 159-175.
- [16] PENG Y C, YAN Y C, LIU S Q, et al. CageNeRF: cage-based neural radiance field for generalized 3D deformation and animation[C]//The 36th International Conference on Neural Information Processing Systems. Red Hook: Curran Associates Inc., 2022: 2277.
- [17] LIU S W, FISCHER M, YOO P D, et al. Neural bounding[C]//ACM SIGGRAPH 2024 Conference Papers. New York: ACM, 2024: 98.
- [18] LUDWIG I, CAMPEN M. Strictly conservative neural implicits[J]. *Computer Graphics Forum*, 2024, 43(7): e15241.
- [19] LIU H T D, WILLIAMS F, JACOBSON A, et al. Learning smooth neural functions via Lipschitz regularization[C]//ACM SIGGRAPH 2022 Conference Proceedings. New York: ACM, 2022: 31.
- [20] TANCİK M, SRINIVASAN P P, MILDENHALL B, et al. Fourier features let networks learn high frequency functions in low dimensional domains[C]//The 34th International Conference on Neural Information Processing Systems. Red Hook: Curran Associates Inc., 2020: 632.
- [21] YANG G D, BELONGIE S, HARIHARAN B, et al. Geometry processing with neural fields[C]//The 35th International Conference on Neural Information Processing Systems. Red Hook: Curran Associates Inc., 2021: 1722.
- [22] GROPP A, YARIV L, HAIM N, et al. Implicit geometric regularization for learning shapes[EB/OL]. [2025-01-08]. <https://dl.acm.org/doi/abs/10.5555/3524938.3525293>.
- [23] PRACH B, BRAU F, BUTTAZZO G, et al. 1-Lipschitz layers compared: memory, speed, and certifiable robustness[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2024: 24574-24583.
- [24] NEACȘU A, PESQUET J C, BURILEANU C. EMG-based automatic gesture recognition using lipschitz-regularized neural networks[J]. *ACM Transactions on Intelligent Systems and Technology*, 2024, 15(2): 26.
- [25] CHANG A X. ShapeNet: an information-rich 3D model repository[EB/OL]. [2025-01-11]. <http://arxiv.org/abs/1512.03012>.