arXiv:2601.07870v1 [cs.LG] 10 Jan 2026

HOSC：一种用于高保真隐式神经表示的具有饱和控制的周期激活函数

Michal Jan Wlodarczyk Danzel Serrano 华沙理工大学 新泽西理工学院michal.wlodarczyk@pw.edu.pl ds867@njit.eduPrzemyslaw Musialski 新泽西理工学院przem@njit.edu

摘要

正弦函数等周期激活函数凭借其振荡结构，能够在隐式神经表示（*INR*）中保留高频信息，但往往存在梯度不稳定性，且难以控制多尺度行为。我们提出带饱和控制的双曲振荡器（*HOSC*）激活函数，$HOSC(x) = \tanh(\beta \sin(\omega_0 x))$，它引入了一个显式参数 β，通过 $\beta\omega_0$控制激活函数的利普希茨界。这提供了一种直接调节梯度幅值的机制，同时保留了周期载波。我们进行了数学分析，并采用标准化训练方案，对图像、音频、视频、神经辐射场和有符号距离函数开展了全面的实证研究。与 *SIREN*、*FINER* 及相关方法的比较分析表明，*HOSC*在哪些方面具有显著优势，又在哪些方面表现相当。研究结果确立了 *HOSC* 作为一种适用于隐式神经表示的实用周期激活函数，并针对不同领域提供了超参数选择指南。代码请访问项目页面：https://hosc-nn.github.io/。

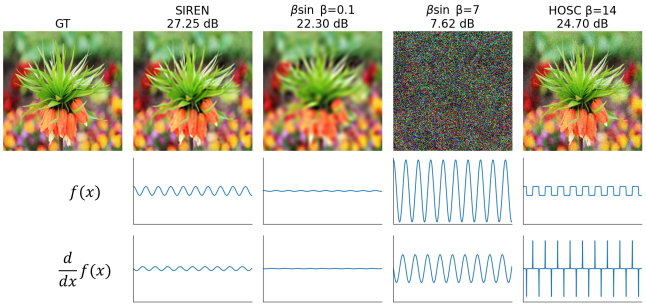

图 1。使用相同的坐标多层感知机和固定频率 $\omega_0 = 30$ 进行二维图像拟合，并改变激活函数。从左到右依次为：真实值、SIREN $\sin(\omega_0 x)$、缩放正弦函数 $\beta\sin(\omega_0 x)$，以及 HOSC $\tanh(\beta\sin(\omega_0 z))$。SIREN 和低-$\beta$ 正弦函数无法充分拟合高频细节，大-β 正弦函数会产生不稳定伪影，而 HOSC 通过结合高频支持与饱和的利普希茨受控梯度，恢复了锐利结构。

1. 引言

隐式神经表示（INR）已成为一种强大的信号编码范式，它以神经网络学习的连续基于坐标的函数取代离散网格 [23, 30, 32]。它们用基于坐标的神经网络 $\mathbb{R}^k \to \mathbb{R}^m$ 表示信号 s f_θ，并通过密集坐标样本进行训练，以最小化重建损失（加上正则化）。它们所具备的优势——无限分辨率、内存效率，以及对不规则几何形状的自然处理能力——推动了其在音频 [31]，视频 [6]，图像压缩 [8]，和三维重建 [20]领域的快速应用。

尽管取得了成功，隐式神经表示仍面临一项根本局限：标准多层感知机本质上难以学习高频成分，这一现象称为频谱偏差 [4, 14, 24]。目前已出现两种主要策略来缓解这一问题：位置编码方法 [20, 32]，通过变换输入坐标来增强高频学习；以及周期激活函数 [30]，直接提升对振荡结构的表征能力。

然而，在周期性激活设计中，仍存在一项根本的架构局限。SIREN [30]，中采用的固定周期正弦激活函数 $\sin(\omega_0 x)$，其激活函数级别的利普希茨常数为 $L_{\text{SIREN}} = \omega_0$，因此同一个参数 ω_0 同时控制频谱支撑和梯度尺度。FINER [17] 等变周期方法通过偏置初始化和依赖振幅的周期来提升频率灵活性，但并未提供一种简单的

全局 Lipschitz 参数：梯度行为仍由载波频率和内部调制决定。正如近期基准测试所指出的 [14], 周期激活函数缺少明确且易于理解的梯度尺度调节旋钮，限制了其对多模态的自适应能力。

为了解决这一问题，我们提出了带饱和控制的双曲振荡器（HOSC），这是一种带饱和控制的周期激活函数，通过组合式设计实现显式梯度调节，定义如下：

$$\mathrm{HOSC}_\beta(x) = \tanh\big(\beta \sin(\omega_0 x)\big),$$

其中，ω_0 设置载波频率，$\beta > 0$ 控制饱和强度。这种组合在保留周期结构的同时，产生紧致的激活层级利普希茨常数 $L_{\mathrm{HOSC}} = \beta\omega_0$。在实践中，$\omega_0$ 可以像 SIREN 中那样根据所需的频谱支撑进行选择，而 β 则作为一个简单易懂的参数，对梯度进行乘法缩放。因此，HOSC 提供了一种简单的激活层级机制，可按模态自适应梯度，无需改变编码或架构，并通过有界输出确保数值稳定性。

图1展示了这些行为在一幅二维图像上的表现：对于固定的网络和载波频率，SIREN 和低β 缩放正弦函数会产生过度平滑的重建，高β 缩放正弦函数则会变得不稳定；而 HOSC 通过结合相同的频谱支撑与受控的饱和梯度，实现了清晰的重建。总之，我们的贡献如下：

- 我们提出 HOSC，这是一种具有饱和控制的周期激活函数，它引入了显式的激活函数级 Lipschitz 参数 $L_{\mathrm{HOSC}} = \beta\omega_0$，将频率选择（$\omega_0$）与梯度缩放（$\beta$）分离。在受控消融实验中，我们表明，HOSC 保持稳定，并在 β范围内提升重建质量；而缩放正弦函数激活 $\beta\sin(\omega_0 x)$会在这些范围内崩溃。
- 我们提供了 SIREN 的即插即用替代方案，无需更改位置编码，与标准 SIREN 式初始化保持兼容，并维持有界输出，从而实现稳定优化。
- 我们在多种模态上验证了 HOSC，结果表明，与以往的周期激活函数（SIREN、FINER）相比，它在音频重建方面取得了显著的性能提升，在视频和图像方面也有可测量的改进，同时在神经辐射场和有符号距离函数上取得了有竞争力的结果。

2. 相关工作

2.1. 隐式神经表示、频谱偏差与周期激活函数

隐式神经表示（INR）通过坐标多层感知机对连续信号进行参数化，并已应用于图像、视频、音频和三维几何 [20, 30, 32]。其中一个核心挑战是频谱偏差：标准多层感知机首先拟合低频成分 [4, 5, 24], , 神经切线核理论对此的解释是核的特征值衰减 [10]。傅里叶特征 [32] 通过随机傅里叶嵌入映射坐标，从而将 NTK 重塑为具有可调带宽的平稳核，缓解了这一问题，并由此改变模型的频率支持集 。

SIREN [30] 从激活函数侧着手解决频谱偏差，采用固定周期正弦激活函数 $\sin(\omega_0 z)$，其激活层级利普希茨常数为 $L_{\mathrm{act}} = \omega_0$ 且输出有界。这改善了对导数和偏微分方程解的表示能力，但利普希茨常数完全由 ω_0 决定，因此频率支持集与梯度尺度相互耦合。增大 ω_0 会同时提高可表示的最高频率以及激活函数导数的上界，因而训练期间无法单独控制梯度幅值。

2.2. 变周期、高斯与学习型

一些研究沿着频率轴扩展了 SIREN。WIRE [27] 引入了复数伽博激活函数，具有更优的时频局部化，但代价是需要进行复数运算且参数数量更多。FINER [17] 和 FINER++ [33]提出了形式为 $\sin(\omega_0(|\mathbf{y}| + 1)\mathbf{y})$ 的变周期激活函数，并通过几何分析和神经正切核分析表明，扩大偏置初始化范围 $b \sim \mathcal{U}(-k,\ k)$可让网络选择具有不同有效频率的子函数。从概念上讲，FINER 通过结合可变周期与偏置范围 k来调节支持频率集，但没有引入显式的全局梯度参数或利普希茨参数：导数会随 $|z|$ 增大，且无法由单一标量界定。FINER++ 将这一变周期思路扩展至高斯和小波载波。

正交分支则完全摒弃了周期性。研究表明，高斯激活函数 [25]在若干隐式神经表示任务上优于正弦函数，而 GARF [7]则将高斯激活函数应用于 NeRF。这些工作强调局部化和梯度保留，而非周期结构。辛克函数及相关激活函数也曾从采样理论和泛函分析视角进行研究 [19, 28], , 但其附近的数值问题 $x = 0$限制了实际应用。近期，元学习激活函数字典（MIRE）[11]以及 KAN [18]等函数架构将激活函数的选择视为一个学习问题，证实激活几何（包括其利普希茨性质和平滑性）会显著影响隐式神经表示性能，但代价是需要额外的基础设施并增加复杂性。

INR-Bench [14]系统地评估了 9 项隐式神经表示任务中的 56 种坐标多层感知机变体和 22 个 KAN模型，并提供了常见激活函数的 L-利普希茨常数和 L-光滑常数（例如，$L_{\mathrm{Lip}} = \omega$ 正弦函数对应的 $(1/\sigma)e^{-1/2}$ 高斯函数对应的）。其主要结论是，周期性正弦函数在低维域中兼具良好的高频学习能力和固定 的激活函数利普希茨常数，这限制了——

其跨模态的多尺度适应性。这促使我们探索能够显式、可解释地控制 Lipschitz 常数和梯度尺度的激活函数。

2.3. 架构与编码方法

另一类研究不再关注激活函数的形状，而是通过架构或编码来实现多尺度表示。乘法滤波器网络和残差乘法滤波器网络 [9, 29]利用乘法滤波器和跳跃连接来表示多个尺度，而 ACORN 和 MINER [15, 26] 则在空间上或通过混合专家分解信号，以拟合十亿像素图像。InstantNGP [23] 和 Mip-NeRF 的变体 [2, 3] 表明，多分辨率哈希网格和抗锯齿编码能够显著加速 NeRF，并减轻高频伪影。BACON [16] 和抗锯齿有符号距离函数方法 [34]通过编码级约束实现带限或受细节层次控制的表示，而 NGSSF [21]则结合利普希茨约束多层感知机与调制编码，实现高斯尺度空间。神经点焦散等显式基元方法 [12]在某些现象（如焦散）的建模中绕过隐式神经表示，从而避开一些低频偏差。

这些方法主要作用于架构/编码轴：它们改变输入的映射与组合方式，而非简单坐标多层感知机的局部激活函数利普希茨性质。HOSC旨在与这类方法正交且兼容：它只修改激活函数，同时保持编码和架构不变。

3. 方法——HOSC

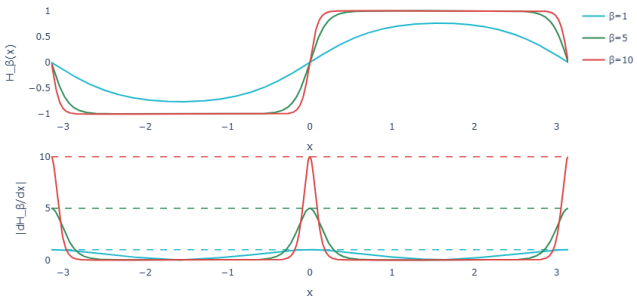

图2. HOSC梯度控制。对正弦函数应用双曲正切函数后，得到一种周期激活函数，可通过 β实现饱和控制。导数幅值及其利普希茨界为 $L = \beta\omega_0$ （虚线）。增大 β 可在保持输出有界的同时，增强零交叉处的梯度并使其更加集中。

3.1. 带饱和控制的双曲振荡器

我们的目标是引入一种机制，在不牺牲周期结构的前提下，解耦单个周期激活函数针对不同信号模态或训练机制的梯度幅度。将正弦载波包裹在双曲正切函数中，可以得到一种振荡激活函数，通过双曲正切实现饱和控制：

$$\mathrm{HOSC}_\beta(x) = \tanh(\beta \sin(\omega_0 x))$$

其中，$\beta > 0$ 控制饱和强度，ω_0 决定载波频率。这种组合带来两个关键性质：（1）输出有界 $|\mathrm{HOSC}_\beta(x)| \leq 1$；（2）可通过 β 参数调节梯度幅值。

图 2 展示了不同 β 取值下的 HOSC。当$\beta < 1$较小时，双曲正切饱和较弱，在零点附近呈现近似线性行为，而在极值处振荡幅度被压缩。当 $\beta > 5$较大时，饱和效应变得强烈，近似于具有陡峭跃迁的方波。这一特性使 HOSC 能够在平滑的周期函数（较低的 β）与阶跃式响应（较高的 β）之间进行插值，从而灵活适应对频率与细节有不同权衡需求的各类信号。

3.2. 梯度分析与利普希茨界

为理解 HOSC 的梯度流特性，我们使用链式法则计算其一阶导数。对于 $\mathrm{HOSC}_\beta(x) = \tanh(\beta\sin(\omega_0 x))$，可得：

$$\frac{d}{dx}\mathrm{HOSC}_\beta(x) = \beta\omega_0 \cos(\omega_0 x) \cdot \mathrm{sech}^2(\beta \sin(\omega_0 x))$$

其中，$\mathrm{sech}^2(z) = 1/\cosh^2(z) = 1 - \tanh^2(z)$。该导数具有三个关键特征：(1) 由 $\beta\omega_0$ 进行乘法缩放，(2) 通过 $\cos(\omega_0 x)$ 实现周期振荡，(3) 通过 $\mathrm{sech}^2(\cdot)$ 进行幅度调制，该调制取决于正弦载波的当前值。

激活层级利普希茨常数。 由于对所有 x，$|\cos(\omega_0 x)| \leq 1$；对所有 $z \in \mathbb{R}$，$0 < \mathrm{sech}^2(z) \leq 1$（在 $z = 0$ 处取等号），因此我们得到一个紧致的利普希茨界：

$$\left|\frac{d}{dx}\mathrm{HOSC}_\beta(x)\right| \leq \beta\omega_0$$

当 $\sin(\omega_0 x) = 0$（使 $\mathrm{sech}^2(0) = 1$）且 $\cos(\omega_0 x) = \pm 1$ 时，该界恰好取到；这在整数 k 对应的 $x = k\pi/\omega_0$ 处发生。激活层级利普希茨常数 $L = \beta\omega_0$ 界定了反向传播过程中每个非线性函数处的导数幅值。

3.3. 层与网络梯度界

在为激活函数确立标量激活层级利普希茨界 $L_{\mathrm{HOSC}} = \beta\omega_0$ 后，我们现在考察这一性质如何在多层网络中传播——这是理解 HOSC 对训练动态影响的关键步骤。

从激活函数到层。 对于一个标准的多层感知机层，其计算形式为 $\mathbf{y} = \mathrm{HOSC}_\beta(\mathbf{W}\mathbf{x} + \mathbf{b})$，链式法则将梯度分解为激活函数和线性分量。由于

激活层雅可比矩阵满足 $\|\nabla_{\mathbf{z}}\mathrm{HOSC}_\beta(\mathbf{z})\|_2 \le \beta|\omega_0|$，整个层的梯度有界于

$$\|\nabla_{\mathbf{x}}\mathbf{y}\|_2 \le \beta|\omega_0|\,\|\mathbf{W}\|_2$$

其中，$\|\mathbf{W}\|_2$ 是权重矩阵的谱范数。这表明，每一层的梯度尺度由激活参数 β 和权重幅值共同控制，因此从业者可以调整 β 来补偿较大的权重，而无需重新训练。

深度网络复合。 堆叠 L 个这样的层后，网络级 Lipschitz 常数满足以下乘法上界：

$$L_{\text{net}} \le \prod_{\ell=1}^{L} \beta|\omega_0|\,\|\mathbf{W}_\ell\|_2.$$

这个上界揭示了为何 β 可作为全局梯度缩放旋钮：将 β 减半，会使网络 Lipschitz 常数的这一上界缩小 2^{-L} 倍，从而能够对深层架构中的梯度幅值进行指数级控制。

3.4. 渐近行为与实践区间

为了理解 HOSC 在完整的 β频谱范围内的行为——从平滑、近线性的响应到陡峭的方波跃迁——我们分析其极限情形，并将其与实验中实际出现的区间联系起来。

平滑区间（$\beta \to 0$）。当 β较小时，对外层 tanh 进行泰勒展开可得 $\mathrm{HOSC}_\beta(x) \approx \beta\sin(\omega_0 x)$，此时恢复为幅度较小且导数幅值较低的平缓缩放正弦函数。在此区间内，梯度保持平滑且相对均匀；这与我们在视频和 NeRF 等高维坐标任务中观察到的行为在经验上相符，在这些任务中，较小的 β 值有助于实现稳定优化（第 4 节）。

尖锐区间（$\beta \to \infty$）。相反，随着 β 增大，激活函数逐渐趋近于方波极限 $\mathrm{HOSC}_\beta(x) \to \mathrm{sign}(\sin(\omega_0 x))$，其导数则集中在 $\sin(\omega_0 x)$ 零点附近的狭窄区域。这会产生高度局部化且幅值很大的梯度，非常适合拟合尖锐的跃变。在我们的一维音频实验中，中到较大的 β 值利用了这一效应，显著提升了对快速波形变化的重建效果（第 4 节）。

门控宽度。导数 $\mathrm{HOSC}'_\beta(x) = \beta\omega_0\cos(\omega_0 x)\mathrm{sech}^2(\beta\sin(\omega_0 x))$ 包含一个 $\mathrm{sech}^2(\beta\sin(\cdot))$ 因子，它会压缩远离正弦载波零点处的梯度。对于与峰值导数成比例的阈值，例如 $\tau = \kappa\beta\omega_0$（固定 $\kappa \in \in(0,1)$），可以证明（参见附录 A.4），超水平集 $\{x : |\mathrm{HOSC}'_\beta(x)| \ge \tau\}$ 的测度按 $1/\beta$ 的量级缩放。因此，增大 β 会缩小梯度较大的区域

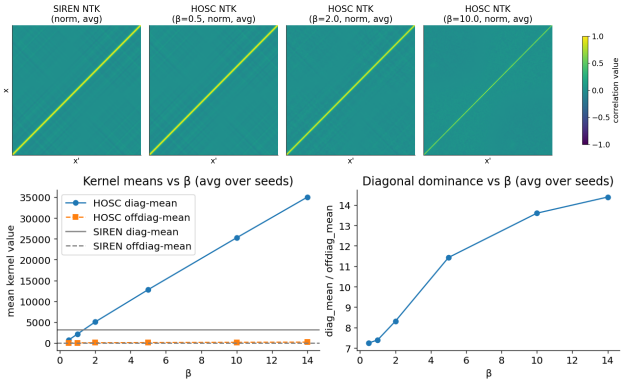

图 3. HOSC 的经验 NTK 行为。上：SIREN 和 HOSC（$\beta \in \{0.5, 2, 10\}$）的相关性归一化核，对多个随机种子的结果取均值。较大的 β 会使对角线更加尖锐，并减弱非对角结构。下：对角线均值和对角占优比（对角线 / 非对角线标准差）随 β的变化。两者均稳步上升，表明 β 是一个可控的核尺度和交互强度调节参数。

同时提高其最大幅值，从而在梯度强度与空间覆盖范围之间实现可控的权衡；我们在第 4 节针对不同模态采用了不同的权衡方式。

数学性质。完整起见，我们指出，HOSC 继承了正弦载波的 $2\pi/\omega_0$ 周期性和奇对称性，同时还具有有界输出 $\mathrm{HOSC}_\beta(x) \in [-1, 1]$ 以及 C^∞ 光滑性。激活函数级别的利普希茨界 $L_{\text{act}} = \beta\omega_0$ 是紧的：当 $x = k\pi/\omega_0$ 时，对于整数 k，有 $\sin(\omega_0 x) = 0$、$\mathrm{sech}^2(0) = 1$，以及 $|\mathrm{HOSC}'_\beta(x)| = \beta\omega_0$。

3.5. HOSC-β 的 NTK 视角

为补充利普希茨分析，我们考察一个具有共享随机参数的简单单隐藏层隐式神经表示中的经验 NTK。在一维网格 $\{x_i\}$ 上，我们计算 $K_\theta(x_i, x_j)$，比较 SIREN 与不同 β 下的 HOSC，并对多个随机种子取平均。

由于 β 会对 K_θ进行重新缩放，我们对角线归一化，并考察相关核 $\tilde{K}_{ij} = K_{ij}/\sqrt{K_{ii}K_{jj}}$.图 3 （上）显示，对于较小的 β，HOSC 的 NTK 与 SIREN 几乎无法区分；而较大的β 则会略微收窄对角带，并减弱非对角相关性。

我们用两个标量统计量概括这一行为（图3，下）：K_θ 的对角线均值 β几乎随之线性增加，而对角占优度量（对角线均值除以非对角线标准差）则单调增长。因此，在这一单层设置中，β 充当了可控的核尺度与交互强度调节旋钮，使核从小 β 时的类 SIREN 状态，平滑转变为较大 β时的自占优状态。

3.6. 与 SIREN 的关系

一个自然的问题是，HOSC 是否具备标准正弦激活函数之外的能力？换言之：为什么 *HOSC* 不能简化为经过缩放或重新参数化的 *SIREN*？根本区别在于参数耦合：SIREN 的单个参数 ω_0 同时决定频谱支撑（网络能够表示哪些频率）和反向传播过程中的梯度幅值。

这意味着，为了捕捉更高频率而增大 ω_0，必然会放大梯度，从而限制适应能力。HOSC 通过其组合式设计 $\tanh(\beta\sin(\omega_0 x))$ 打破了这种耦合：其中，ω_0 设定载波频率，而 β 则提供独立的乘法式梯度控制。这种解耦使模型能够在高梯度区域（$\beta > 5$）中稳定训练，而振幅缩放正弦网络在此类区域会出现不稳定。我们通过两项受控消融实验验证了这一差异。

消融实验 1：振幅缩放（β-正弦函数控制）。我们通过比较 HOSC 与直接进行振幅缩放的正弦函数 $\beta\sin(\omega_0 x)$ 在相同训练条件下的表现（图4），测试双曲正切包络是否不可或缺。尽管两者在较小的 β下表现相近，但未缩放的正弦函数会随着 β 增大而迅速变得不稳定，而 HOSC 则保持稳定。这证实了，在高梯度区域中实现可训练性的关键在于饱和包络，而非单纯的振幅缩放。

消融实验 2：频率调节（β–与–ω_0）。我们通过在匹配范围内大幅调节频率（图 β 用于 HOSC，ω_0 用于 SIREN），测试 SIREN 能否匹敌 HOSC。5）。尽管两者都呈现性能峰值，HOSC 的重建质量始终更高，这表明 β的梯度控制无法仅通过频率重新参数化来模拟。

这些消融实验证实：（1）HOSC 并非 β 缩放正弦函数；（2）HOSC 不等同于经过频率调节的 SIREN。双曲正切—正弦复合提供了一种独特的梯度控制机制，振幅缩放和频率重新参数化都无法复现。

3.7. 与 FINER 的关系

FINER [17]从不同角度处理频谱偏差。它并不改变固定频率正弦函数的利普希茨特性，而是引入了一种形式如下的可变周期激活函数：

$$\sigma_{\text{FINER}}(x) = \sin\big(\omega_0 \alpha(x)\, x\big), \quad \alpha(x) = |x| + 1, \qquad (1)$$

并通过扩大偏置初始化 $b \sim \mathcal{U}(-k,\ k)$来调整支持频率集。 [17] 中的几何分析和 NTK 分析表明，增大 k 会选择有效频率更高的子函数，扩大支持频率集$F_{\omega_0,\ k}$，并增强 NTK 对角占优性，从而提高高频收敛性能。

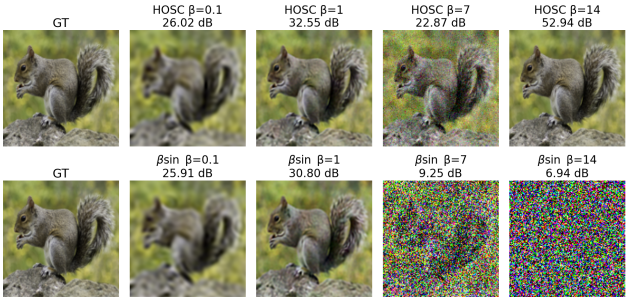

图4. β对 HOSC 和 β-sin 的影响。单张 DIV2K 图像，$\omega_0 = 30$。上排：$\beta \in \{0.1, 1, 7, 14\}$对应的真实值和 HOSC。下排：同一 β对应的真实值和 β-sin。所有实验均采用相同的隐式神经表示架构和训练设置。

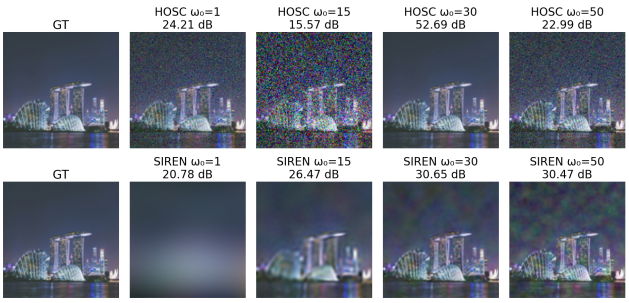

图 5。 ω_0 对 HOSC 和 SIREN 的影响。单张 DIV2K 图像。上排：使用固定的 $\beta = 14$ 和 $\omega_0 \in \{1, 15, 30, 50\}$ 的 HOSC。下排：使用 $\beta = 1$ 和相同 ω_0 值的 SIREN，并采用相同的隐式神经表示架构和训练设置。

相比之下，HOSC 保持载波频率不变，而是对正弦函数施加饱和非线性

$$\sigma_{\text{HOSC}}(x) = \tanh\big(\beta \sin(\omega_0 x)\big), \qquad (2)$$

从而得到紧致激活层级利普希茨常数 $\|\sigma'_{\text{HOSC}}\|_\infty = \beta\omega_0$ （第3节）。因此，HOSC 引入了一个显式的全局梯度尺度参数 β，同时使频谱支撑由 ω_0 和线性层决定。FINER 可被视为一种初始化层级频率调节器：它通过偏置范围 k 和可变周期 $\alpha(x)$ 扩大并重新分配频率集合。HOSC 是一种激活层级梯度调节器：它提供有界输出和可控的利普希茨常数，同时不改变底层频率集合。因此，这两种机制彼此正交，原则上可以组合使用。

4. 实际实验

我们考察 HOSC 在图像、音频、视频、NeRF 和有符号距离函数重建等任务上的重建质量。在每项实验中，我们首先通过消融研究探索 β 参数空间。随后，在方法比较中采用表现最佳的取值。我们将 HOSC 与四种基于激活函数的隐式神经表示方法进行比较，即高斯（Gauss） [25],



图 6. 图像拟合的定性比较。每行使用对应列所示的方法重建真实值数据。定量峰值信噪比每个示例展示的结果如下所示重建图像。HOSC 参数 $\beta = 14.0$。

小波（WIRE） [27], 正弦函数（SIREN） [30] 和 FINER [17] 激活函数，以及基于编码的方法，即傅里叶特征位置编码（PEMLP） [32]。我们在论文中报告了最具代表性的结果，并在附录中介绍详细信息。所有实验均采用统一的协议，在 A100/H100 GPU 上使用 SIREN 式初始化、Adam、相同的指标、固定的随机种子、确定性数据加载以及一致的训练计划。

4.1. 图像拟合

在二维图像重建任务中，我们旨在学习从二维输入像素位置到三维 RGB 输出颜色的映射 f: $\mathbb{R}^2 \to \mathbb{R}^3$，损失函数为网络输出与真实值之间的均方距离 L_2 （均方误差）。我们在 DIV2K 数据集 [1] 中的 512×512 分辨率下，对 16 张图像上的隐式神经表示重建性能进行评估。首先，我们对所有样本上的 β 进行消融实验。在与最先进方法比较时，我们采用平均表现最佳的 β。

结果。 如表 1所示，峰值信噪比随 β 稳步提升，并在 $\beta = 14.0$附近达到峰值，之后趋于饱和或略有下降，因此我们在后续所有二维图像实验中固定 $\beta = 14.0$ 。在此设置下，HOSC 的表现明显优于 PEMLP、Gauss、WIRE 和 SIREN，并且与 FINER 相当（表2）：在 16 张图像中，HOSC 有 9 张的表现与 FINER 持平或更优（见附录），但平均表现略逊于 FINER。这些结果证实，达到较高的 β 水平对于发挥 HOSC 的优势至关重要；在这一水平下，HOSC 缩小了与现有最强周期激活函数的大部分差距，同时相较于标准的基于正弦函数的隐式神经表示，取得了显著提升。

4.2. 音频拟合

我们评估 HOSC 在将多层感知机拟合到音频信号这一任务中的有效性。多层感知机学习一个隐式一维函数， f: $\mathbb{R}^1 \to \mathbb{R}^1$ 表示信号振幅随时间的变化。损失函数为均方误差。我们在两个示例上评估重建质量：巴赫《大提琴组曲第一号：前奏曲》（巴赫）的前 7 秒，以及一段男性演员从 0 数到 9 的 12 秒片段（计数）。论文中展示巴赫的结果，计数的结果则见附录。

结果。 表 β消融实验显示3， β在4–5 的范围内，在两种信号上都能获得最佳重建质量；我们将 $\beta = 5.0$ 固定为所有音频实验的统一设置。在此设置下，HOSC 显著优于所有基线方法（表 4）：与次优激活函数（SIREN）相比，HOSC 在巴赫上的均方误差降低了超过 4×，在计数任务上则几乎降低了一个数量级。这使音频设置成为最鲜明的例子，表明进入高β 区域能够实现仅靠正弦激活函数无法达到的保真度。

表 1. **图像 HOSC β-消融实验**。对 16 个样本取平均值。在实验表格中，我们用颜色标记各单元格，分别表[illegible]值和第[illegible]

指标 峰值信噪比 ↑	HOSC β 0.1	0.2	0.5	0.8	1.0	2.0	4.0	5.0
Mean	28.6934	31.8259	36.0805	37.2026	37.4394	38.7040	38.8871	38.6065
Std	3.3241	2.9944	2.8315	2.9084	2.9572	3.4891	3.2364	3.0213
	6.0	8.0	10.0	12.0	14.0	16.0	18.0	20.0
Mean	38.5514	39.2204	39.9796	40.1840	40.2247	39.9558	39.1992	38.9566
Std	2.9264	2.9735	2.8425	2.5163	2.3639	2.6917	2.6648	2.9618

表2。**图像拟合的定量比较**。均值和标准差为16个测试集示例的平均值。 $\beta = 14.0$。

示例 峰值信噪比 ↑	PEMLP	Gauss	WIRE	SIREN	FINER	HOSC
寺庙	26.3002	31.4108	28.0027	33.6837	35.8982	36.6901
市场	30.4807	35.8760	32.4138	40.2123	42.6105	40.3085
核桃	24.1596	32.2588	26.9073	34.45		37.0417
Mean	29.5525	35.6942	31.3934	38.5811	40.8816	40.2247
Std	3.7192	2.7251	3.0141	3.1956	3.1379	2.3639

表3. **音频 HOSC β-消融实验**。展示了巴赫和计数示例。完整的消融实验见附录

示例 均方误差 ↓	HOSC β 1.0	2.0	4.0	5.0	6.0	8.0
Bach	9.86e-06	2.029e-06	7.589e-07	1.319e-06	2.384e-06	2.138e-05
计数	4.328e-04	2.550e-04	6.851e-05	3.613e-05	9.355e-05	1.707e-04



图7. 视频拟合的定性比较。重建了一段猫的300帧视频。展示的是其中一帧，来自该视频。下方展示了各个示例重建帧的定量结果。HOSC $\beta = 0.5$。

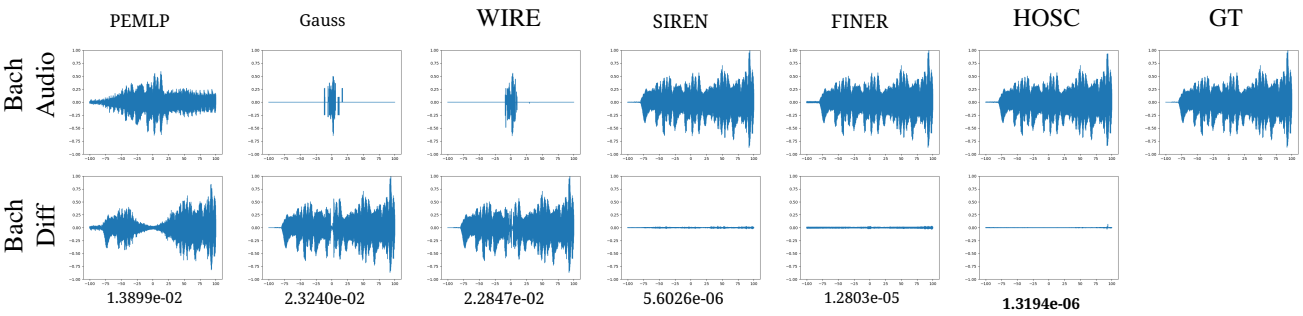

图 8. 音频拟合的定性比较。我们展示了重建信号（音频）以及预测信号与真实信号之间的差异（差异）。我们用均方误差量化差异。HOSC $\beta = 5.0$。

4.3. 视频拟合

我们评估 HOSC 在视频重建中的效果，其中多层感知机将函数 f 参数化为：$\mathbb{R}^3 \to \mathbb{R}^3$ 从坐标到 RGB 值的映射（x，y，t）。损失为均方误差（MSE）。我们使用一段分辨率为 512×512、包含 300 帧的猫视频，并报告逐帧及平均峰值信噪比/均方误差。所有方法采用相同的网络架构、优化器、训练计划和初始化；只有激活函数（以及 HOSC 的β）不同。我们首先对 β 进行消融实验，然后使用最佳的 β将 HOSC 与 PEMLP、Gauss、WIRE、SIREN 和 FINER 进行比较。

结果。 在 β的消融实验表5 中，最佳值清晰地出现在 $\beta = 0.5$；取值更小或更大都会降低重建质量。采用这一取值后，HOSC 的整体表现最佳（表6）：相较于 SIREN，峰值信噪比提升约 3 dB，均方误差减半，而非周期基线方法的表现则更差。这与我们的分析一致（参见第3.4节）：对于更高维的坐标输入，例如 $(x,\ y,\ t)$，相对较小的 β能够使 HOSC 保持在平滑饱和的类正弦状态，从而实现稳定且准确的重建。

表 4。**音频拟合的定量比较**。我们报告每个音频样本所达到的最低均方误差损失。$\beta = 5.0$

示例 均方误差 ↓	PEMLP	Gauss	WIRE	SIREN	FINER	HOSC
Bach	1.3899e-02	2.3240e-02	2.2847e-02	5.6026e-06	1.2803e-05	1.3194e-06
计数	4.1440e-03	7.4262e-03	7.4241e-03	3.9		

4.4. 神经辐射场（NeRF）

我们在 Blender基准（8 个场景）上评估 HOSC 和其他基于激活函数的方法 [20]。所有方法采用相同的架构、优化器、训练预算和 SIREN 式初始化；不使用位置编码。我们对 HOSC β 参数进行消融实验，并在相同预算下比较峰值信噪比、结构相似性和学习感知图像块相似度指标。

发现。 表 β消融实验（见表7）显示，NeRF 的性能在 $\beta \in [0.5, 1.0]$时达到峰值；取值更大时，峰值信噪比/结构相似性和学习感知图像块相似度会迅速下降，而 $\beta \geq 3$ 则会导致所有场景严重崩溃（表 7）。当 β 取较小值时，HOSC 在 Blender基准上的表现与 SIREN 和 FINER 相当（表8），各场景的峰值信噪比差异约为 0。1–0。2dB（FINER 为 29.84 dB，HOSC 为 29.74 dB，SIREN 为 29.71 dB）。因此，在 NeRF 中，HOSC 的主要优势在于提供了一个安全、显式的梯度尺度调节旋钮，能够保留基于正弦函数的隐式神经表示的频率支持，而非像我们的音频实验那样开辟一种新的精度水平。

表 5. **视频 HOSC β消融实验**。以猫视频为例。

	HOSC β				
指标	0.1	0.5	1.0	2.0	5.0
均方误差 ↓	7.7876e-04	5.5673e-04	6.7687e-04	1.5408e-02	1.6507e-01
峰值信噪比 ↑	37.3071	38.71367	37.821784	24.2534	13.9256

表6. 视频拟合的定量比较。$\beta = 0.5$

指标	PEMLP	Gauss	WIRE	SIREN	FINER	HOSC
均方误差 ↓	2.3615e-03	1.8921e-03	3.8377e-03			
峰值信噪比 ↑	32.3877	33.3167	30.2326	35.9106	32.1581	38.7137



图 9. 图像拟合的定性比较。每种方法下方展示了 NeRF 合成数据集中的乐高卡车示例的结构相似性指标[20]。HOSC $\beta = 1.0$。

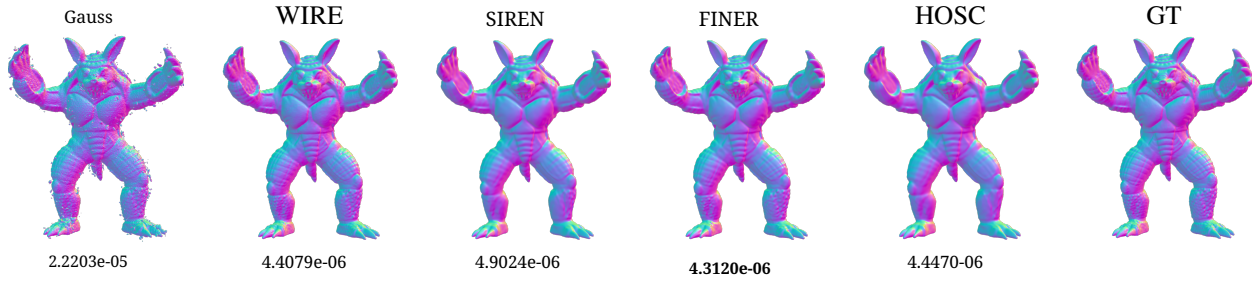

图 10. 有符号距离场（神经有符号距离场）拟合的定性比较。每个示例下方展示了倒角距离。 $\beta = 0.8$。完整结果、消融研究和实验细节见附录。

能够保留基于正弦函数的隐式神经表示的频率支持，而非像我们的音频实验那样开辟一种新的精度水平。

4.5. 有符号距离场

接下来，我们测试 HOSC 在三维有符号距离场（有符号距离函数）重建中的表现，其中多层感知机将一个函数参数化f: $\mathbb{R}^3 \to \mathbb{R}$ 其零水平集定义了一个曲面。遵循先前的工作[16, 34], , 我们训练所有方法以最小化点到曲面损失，并报告留出曲面样本上的倒角距离。我们在图 犰狳 形状上展示了具有代表性的定性比较，见图 10，并列出各方法的倒角距离。附录还报告了其他形状以及 HOSC 的小规模 β参数扫描结果。

研究发现。 在 犰狳上，HOSC 的倒角距离与周期性基线方法相当，且明显优于非周期性激活函数（图 10）：高斯函数的误差高出一个数量级，而 WIRE、SIREN、FINER 和 HOSC 的结果都集中在窄带范围内。

表 7. **NeRFHOSC β消融实验。**在 NeRF Synthetic 的八个样本上取平均 [20]

		HOSC β						
	统计量。	0.1	0.3	0.5	0.8	1.0	3.0	5.0
PSNR	均值。	26.4848	28.7892	29.3195	29.6461	29.7321	26.7590	21.6911
	Std.	4.5546	4.4083	4.4643	4.6295	4.6157	4.0048	3.1601
SSIM	均值。	0.9002	0.9300	0.9346	0.9372	0.9388	0.9033	0.7999
	Std.	0.0765	0.0652	0.0629	0.0640	0.0622	0.0579	0.0820
LPIPS	均值。	0.1157	0.0455	0.0393	0.0358	0.0353	0.0757	0.2355
	Std.	0.0835	0.0374	0.0318	0.0282	0.0277	0.0427	0.0768

5. 结论

我们提出了 HOSC，这是一种周期激活函数，它将梯度尺度与载波频率解耦，并加入了显式的梯度门控机制。通过解析推导，我们得到了紧致激活层级利普希茨界，证明高梯度区域会随着 $O(1/\beta)$ 而缩小，并将 β与经验神经切线核的结构联系起来；随着梯度局部化，该核会变得越来越对角占优。这些结果使 β 成为一种简单而有效的调节手段，可控制隐式神经表示的稳定性和局部性。

从经验来看，HOSC 可作为正弦激活函数的即插即用替代方案。在一维音频任务中，HOSC 的优势最为显著；与 SIRE N 相比，较高的载波频率在此类任务中更有利

表 8. **新视角合成的定量比较。**HOSC $\beta = 1.0$

	方法 椅子 鼓 榕树 热狗 乐高 材料							Mic	Ship
PSNR ↑	PEMLP 29.33 20.11 23.33 28.32 24.56					23.80		22.59	19.84
	Gauss	29.33	22.51	25.45	29.59	25.98	24.87	31.51	20.36
	WIRE	33.86	25.08	27.85	34.18	29.41	27.85	34.95	20.61
	SIREN	34.39	25.75	28.45	33.96	30.26	27.81	35.08	21.99
	FINER	34.77			33.58	30.76	27.74	35.05	22.04
	HOSC	34.28	25.83	28.36	33.77	30.49	27.86	35.28	21.98
SSIM ↑	PEMLP	0.950	0.823	0.907	0.925	0.893	0.907	0.931	0.712
	Gauss	0.940	0.851	0.932	0.919	0.892	0.902	0.970	0.696
	WIRE	0.976	0.915	0.958	0.971	0.949	0.945	0.982	0.747
	SIREN	0.978	0.923	0.960	0.969	0.958	0.946	0.982	0.783
	FINER	0.977	0.923	0.962	0.967	0.959	0.943	0.982	0.781
	HOSC	0.977	0.925	0.961	0.968	0.960	0.946	0.983	0.792
LPIPS ↓	PEMLP	0.039	0.180	0.078	0.054	0.085	0.069	0.121	0.236
	Gauss	0.042	0.109	0.057	0.082	0.061	0.086	0.024	0.224
	WIRE	0.015	0.053	0.033	0.018	0.034	0.032	0.015	0.187
	SIREN	0.014	0.046	0.029	0.023	0.024	0.030	0.0132	
	FINER	0.012	0.043	0.027	0.028	0.021	0.031	0.011	0.108
	HOSC	0.014	0.045	0.030	0.027	0.023	0.031	0.014	0.099

，在 FINER 中也是如此。在图像和视频任务中，HOSC 的表现类似一种温和的正则化方法，带来的提升虽小，却稳定可靠。在 NeRF 上，HOSC 的效果与经过频率调优的基线方法相当，同时在更激进的超参数设置下仍保持稳定。总体而言，我们的结果表明，针对可控的利普希茨常数和局部化梯度显式设计激活函数，与基于频率的隐式神经表示架构相辅相成。未来，我们计划开发激活函数和架构，进一步充分解耦频率内容与梯度控制，例如在高维神经场中采用可学习载波或空间自适应门控。

参考文献

[1] Eirikur Agustsson 和 Radu Timofte。NTIRE 2017 单幅图像超分辨率挑战：数据集与研究。载于 IEEE计算机视觉与模式识别会议（CVPR）研讨会论文集，2017年。6, 12[2] Jonathan T. Barron、Ben Mildenhall、Matthew Tancik、Peter Hedman、Ricardo Martin-Brualla 和 Pratul P. Srinivasan。Mip-NeRF：用于神经辐射场抗锯齿的多尺度表示。载于 IEEE/CVF国际计算机视觉会议（ICCV）论文集，第5855–5864页，2021年。3[3] Jonathan T. Barron、Ben Mildenhall、Dor Verbin、Pratul P. Srinivasan 和 Peter Hedman。Mip-NeRF 360：无界抗锯齿神经辐射场。载于 IEEE/CVF计算机视觉与模式识别会议（CVPR），第5460–5469页，2022年。3[4] Ronen Basri、Meirav Galun、Amnon Geifman、David Jacobs、Yoni Kasten 和 Shira Kritchman。非均匀密度输入下神经网络的频率偏差。载于国际机器学习会议（ICML）论文集，第685–694页。PMLR，2020年。1, 2[5] Yuan Cao、Zhiying Fang、Yue Wu、Ding-Xuan Zhou 和 Quanquan Gu。探究深度学习的频谱偏差。arXiv预印本 arXiv:1912.01198，2019年。2[6] Hao Chen、Bo He、Hanyu Wang、Yixuan Ren、Ser-Nam Lim 和 Abhinav Shrivastava。NerV：用于视频的神经表示。载于《神经信息处理系统进展》第34卷：2021年神经信息处理系统年会（NeurIPS 2021），第21557–21568页，2021年。1[7] Shin-Fang Chng、Sameera Ramasinghe、Jamie Sherrah 和 Simon Lucey。用于高保真重建和姿态估计的高斯激活神经辐射场。载于欧洲计算机视觉会议（ECCV），第264–280页，2022年。2[8] Emilien Dupont、Adam Goli´nski、Milad Alizadeh、Yee Whye Teh 和 Arnaud Doucet。COIN：利用隐式神经表示进行压缩。载于 ICLR 2021 神经压缩研讨会，2021年。arXiv:2103.03123。1[9] Rizal Fathony、Anit Kumar Sahu、Devin Willmott 和 J. Zico Kolter。乘法滤波器网络。载于国际表征学习会议（ICLR），2021年。3[10] Arthur Jacot、Franck Gabriel 和 Cl´ement Hongler。神经切线核：神经网络中的收敛与泛化。载于《神经信息处理系统进展》（NeurIPS），第 8580–8589 页。Curran Associates, Inc., 2018年。2[11] Dhananjaya J ayasundara、Heng Zhao、Demetrio Labate 和 Vishal M. Patel。MIRE：匹配的隐式神经表示。载于 IEEE/CVF计算机视觉与模式识别会议（CVPR），第 8279–8288 页，2025年。2[12] Georgios Kopanas、Thomas Leimk¨ uhler、Gilles Rainer、Cl´ement Jambon 和 George Drettakis。用于反射新视角合成的神经点焦散。ACM图形学汇刊，41(6)，2022年。2022年亚洲计算机图形学会议。3[13] 斯坦福大学计算机图形学实验室。斯坦福三维扫描存储库，2014年。http://graphics.stanford.edu/data/3Dscanrep/。14[14] Linfei Li、Fengyi Zhang、Zhong Wang、Lin Zhang 和 Ying Shen。INR-Bench：面向多领域回归与重建的隐式神经表示统一基准。arXiv 预印本 arXiv:2510.10188，2025年。1, 2[15] David B. Lindell、Dave Van Veen、Jeong Joon Park 和 Gordon Wetzstein。ACORN：用于神经场景表示的自适应坐标网络。ACM图形学汇刊（TOG），40(4)，2021年。3[16] David B. Lindell、Dave Van Veen、Jeong Joon Park 和 Gordon Wetzstein。BACON：用于多尺度场景表示的带限坐标网络。载于 IEEE/CVF计算机视觉与模式识别会议（CVPR），第 16241–16250 页，2022年。3, 8, 14[17] Zhen Liu、Hao Zhu、Qi Zhang、Jingde Fu、Weibing Deng、Zhan Ma、Yanwen Guo 和 Xun Cao。FINER：通过变周期激活函数调节隐式神经表示中的频谱偏置。载于 IEEE/CVF计算机视觉与模式识别会议（CVPR）论文集，第 2713–2722 页，2024年。1, 2, 5, 6, 12[18] Ziming Liu、Yixuan Wang、Sachin Vaidya、Fabian Ruehle、James Halverson、Marin Solja˘ci´c、Thomas Y. Hou 和 Max Tegmark。KAN：Kolmogorov-Arnold网络。载于国际学习表征会议（ICLR），2025年。口头报告。2[19] Lassi Meronen、Martin Trapp 和 Arno Solin。周期激活函数会诱导平稳性。《神经信息处理系统进展》（NeurIPS），34 2032，2021 2021 2 : – 。[20]Ben Mildenhall、Pratul P. Srinivasan、Matthew Tancik、Jonathan T. Barron、Ravi Ramamoorthi 和 Ren Ng。NeRF：将场景表示为神经辐射场以实现视图合成。载于欧洲计算机视觉会议（ECCV）论文集，第 405–421 页。Springer，2020年。1, 2, 7, 8, 14[21] Felix Mujkanovic、Ntumba Elie Nsampi、Christian Theobalt、Hans-Peter Seidel 和 Thomas Leimk¨uhler。神经高斯尺度空间场。ACM图形学汇刊，43(4)，2024年。2024年计算机图形学会议。3[22] Thomas M¨uller。tiny-cuda-nn，2021年。14

[23] Thomas M¨uller、Alex Evans、Christoph Schied 和 Alexan- der Keller。采用多分辨率哈希编码的即时神经图形基元。ACM图形学汇刊（TOG），41(4):102:1–102:15，2022年。1, 3, 14[24] Nasim Rahaman、Aristide Baratin、Devansh Arpit、Felix Draxler、Min Lin、Fred A. Hamprecht、Yoshua Bengio 和 Aaron C. Courville。论神经网络的频谱偏差。载于第 36 届国际机器学习会议论文集，第 5301–5310 页。PMLR，2019年。1, 2[25] Sameera Ramasinghe 和 Simon Lucey。超越周期性：迈向坐标多层感知机激活函数的统一框架。载于欧洲计算机视觉会议（ECCV），第 142–158 页，2022年。口头报告。2, 5, 12[26] Vishwanath Saragadam、Daniel LeJeune、Jasper Tan、Guha Balakrishnan、Ashok Veeraraghavan 和 Richard G. Baraniuk。MINER：多尺度隐式神经表示。载于欧洲计算机视觉会议（ECCV）论文集，第 318–333 页。Springer，2022年。3[27] Vishwanath Saragadam、Daniel LeJeune、Jasper Tan、Guha Balakrishnan、Ashok Veeraraghavan 和 Richard G. Baraniuk。WIRE：小波隐式神经表示。载于 IEEE/CVF计算机视觉与模式识别会议（CVPR）论文集，第 18507 18516 页，2023 2, 6, 12 – 。[28] Hemanth Saratchandran、Sameera Ramasinghe、Violetta Shevchenko、Alexander Long 和 Simon Lucey。隐式神经表示激活函数的采样理论视角，2024年。2[29] Shayan Shekarforoush、David B. Lindell、David J. Fleet 和 Marcus A. Brubaker。用于多尺度重建的残差乘性滤波网络。载于《神经信息处理系统进展》（NeurIPS），第 23848 23860 页，2022 3 – 。[30]Vincent Sitzmann、Julien N. P. Martel、Alexander W. Bergman、David B. Lindell 和 Gordon Wetzstein。采用周期激活函数的隐式神经表示。载于《神经信息处理系统进展》（NeurIPS），第 7462–7473 页。Curran Associates, Inc., 2020年。1, 2, 6, 12[31] Kun Su、Mingfei Chen 和 Eli Shlizerman。Inras：面向音频场景的隐式神经表示。载于《神经信息处理系统进展》，第 8144–8158 页。Curran Associates, Inc., 2022年。1[32] Matthew Tancik、Pratul P. Srinivasan、Ben Mildenhall、Sara Fridovich-Keil、Nithin Raghavan、Utkarsh Singhal、Ravi Ramamoorthi、Jonathan T. Barron 和 Ren Ng。傅里叶特征使网络能够在低维域中学习高频函数。载于《神经信息处理系统进展》（NeurIPS），第 7537–7547 页。Curran Associates, Inc., 2020年。1, 2, 6, 12[33] Hao Zhu、Zhen Liu、Qi Zhang、Jingde Fu、Weibing Deng、Zhan Ma、Yanwen Guo 和 Xun Cao。FINER++：构建一族用于激活隐式神经表示的变周期函数。arXiv 预印本 arXiv:2407 19434，2024 2 . . [34] Yiyu Zhuang、Qi Zhang、Ying Feng、Hao Zhu、Yao Yao、Xi- aoyu Li、Yan-Pei Cao、Ying Shan 和 Xun Cao。抗混叠采用编码细节层次的神经隐式曲面。在 *SIG-GRAPH* 亚洲 *2023* 会议论文集，第 1–10 页。ACM，2023 年。3，8

HOSC：一种用于高保真隐式神经表示的具有饱和控制的周期激活函数

补充材料

附录概述

本附录首先汇总 HOSC 激活函数的数学细节，包括第 3 节中各项结果的正式陈述与证明。后续章节介绍实验的实用信息：训练与实现细节、参数 β的消融实验，以及所有图像、音频、视频、NeRF 和有符号距离函数实验的扩展定量表格与定性可视化结果，另有一项额外的哈希网格实验。定性结果可通过入口页面（https://hosc-nn.github.io/）访问 HTML 实验套件。

A. HOSC 数学细节

本部分给出正式陈述及证明，以支持第 3 节中的数学论断。

符号说明、假设与术语

令 $\beta > 0$ 和 $\omega_0 > 0$。定义标量激活函数

$$\phi_{\beta,\omega_0}(x) = \tanh\big(\beta\, \sin(\omega_0 x)\big).$$

对于各层，使用 $z = Wx + b$ 并逐元素应用 ϕ。矩阵范数为谱范数（$\|\cdot\|_2$）；雅可比矩阵/算子范数为诱导的 ℓ_2-范数。“激活层级的利普希茨常数”是相对于预激活 z 而言的。**术语说明：**正文中的“解耦”是指引入一个独立旋钮 β来控制梯度尺度，而 ω_0 设定载波频率；激活层级的利普希茨常数等于 $\beta\omega_0$，并且取决于两者。

A.1. 激活函数微积分与利普希茨界

引理 A.1.1（导数）。对于 $f(x) = \phi_{\beta,\ \omega_0}(x)$，

$$f'(x) = \beta\,\omega_0\, \cos(\omega_0 x)\, \operatorname{sech}^2\big(\beta\, \sin(\omega_0 x)\big).$$

证明。根据链式法则，$u(x) = \beta\sin(\omega_0 x)$ 且 $\frac{d}{dx}\tanh u = $双曲正割$^2\, u \cdot u'$。 □

引理 A.1.2（紧致激活层级利普希茨界）。

$$\sup_{x\in\mathbb{R}} |f'(x)| = \beta\,\omega_0.$$

证明。由引理 A.1.1，$|f'| \le \beta\omega_0$ 由于 $|\cos| \le 1$ 且 $0 <$ 双曲正割$^2 \le 1$。当 $x = k\pi/\omega_0$ 时取等号，此时 $\sin = 0$ 且 $|\cos| = 1$。 □

推论 A.1.3（基本性质）。$\phi_{\beta,\ \omega_0}$ 是 $2\pi/\omega_0$-周期的、奇函数，在 $[-1, 1]$ 上有界，且 C^∞。

A.2. 层与网络界

命题 A.2.1（层雅可比界）。对于 $y = \phi_{\beta,\ \omega_0}(Wx + b)$，

$$\|\nabla_x y\|_2 \ \le\ \beta\,\omega_0\,\|W\|_2 \quad \text{for all } x.$$

证明。雅可比矩阵可分解为 $J = D(z)\,W$，其中 $D(z) = \operatorname{diag}(f'(z_i))$。由引理 A.1.2，$\|D(z)\|_2 \le \beta\omega_0$。利用次乘法性可得 $\|J\|_2 \le \beta\omega_0\|W\|_2$。 □

命题 A.2.2（网络利普希茨上界）。对于一个 L层多层感知机 y_L，其满足 $y_k = \phi_{\beta,\ \omega_0}(W_k y_{k-1} + b_k)$，

$$\|\nabla_x y_L(x)\|_2 \ \le\ \prod_{k=1}^{L} \big(\beta\,\omega_0\,\|W_k\|_2\big), \quad \forall x.$$

证明。对每一层应用命题 A.2.1，并对雅可比矩阵乘积应用次乘法性。 □

A.3. β 的渐近行为

命题 A.3.1（小-β 展开）。对于 $\beta \in (0,\ \beta_0]$，

$$\phi_{\beta,\omega_0}(x) = \beta\sin(\omega_0 x) - \frac{\beta^3}{3}\sin^3(\omega_0 x) + O(\beta^5)$$

关于 x 一致成立。

证明。对 $\tanh s = s - s^3/3 + O(s^5)$ 使用泰勒级数，其中 $s = \beta\sin(\omega_0 x)$，且 $|\sin| \le 1$。 □

命题 A.3.2（大-极限）。如果 $\sin(\omega_0 x) \neq 0$，则 $\phi_{\beta,\ \omega_0}(x) \to \operatorname{sign}(\sin(\omega_0 x))$，当 $\beta \to \infty$ 时；对几乎处处逐点成立，并且对于任意有限 p，在 $L^p{}_{\text{loc}}$ 中成立。

证明。当 $|y| \to \infty$ 时，$\tanh y \to \operatorname{sign}(y)$。取 $y = \beta\sin(\omega_0 x)$。支配收敛定理给出在远离零点的紧集上的局部$L^p$ 收敛；有界性by1提供了支配条件。 □

A.4. 门控区域测度

本节讨论论文中的第 3.4 节。设

$$g_\beta(x) = \beta\,\omega_0\,|\cos(\omega_0 x)|\ \operatorname{sech}^2\big(\beta\,|\sin(\omega_0 x)|\big),$$
$$u = \omega_0 x \in [0, 2\pi].$$

引理 A.4.1（比例阈值下的逐周期活跃集）。固定 $\kappa \in (0,1)$ 并定义比例阈值

$$\tau_\beta = \kappa\, \beta\, \omega_0.$$

在一个周期 $I = [0, 2\pi/\omega_0]$ 内，考虑超水平集

$$S_\beta = \{\, x \in I :\ g_\beta(x) \geq \tau_\beta \,\}.$$

Then

$$|S_\beta| \ \leq\ \frac{4}{\omega_0}\ \arcsin\Big(\min\Big\{1,\ \frac{\operatorname{arcosh}(1/\sqrt{\kappa})}{\beta}\Big\}\Big),$$

特别是 $|S_\beta| = O(1/\beta)$，其中 $\beta \to \infty$。
证明。令 $u = \omega_0 x$。不等式

$$g_\beta(x) = \beta\omega_0\, |\cos u|\ \operatorname{sech}^2\big(\beta |\sin u|\big)\ \geq\ \kappa\beta\omega_0$$

意味着 双曲正割^2(β|辛克函数 u|) $\geq \kappa$，因为 |余弦 u| ≤ 1。因此，属于 S_β 的必要条件是

$$\operatorname{sech}^2(\beta |\sin u|) \geq \kappa \iff \beta |\sin u| \leq c_\kappa = \operatorname{arcosh}\big(1/\sqrt{\kappa}\big).$$

因此

$$S_\beta\ \subset\ \Big\{\, x \in I :\ |\sin(\omega_0 x)| \leq c_\kappa/\beta \,\Big\}.$$

在 $u \in [0, 2\pi]$ 上，集合 $\{u : |$辛克函数 $u|\ \leq\ \varepsilon\}$ 由三个区间组成（分别靠近 0、 π 和 2π），总长度为 4 反正弦(ε)。通过 $x = u/\omega_0$ 映射回去可得

$$|S_\beta|\ \leq\ \frac{4}{\omega_0}\ \arcsin\Big(\min\{1,\ c_\kappa/\beta\}\Big).$$

由于 $\beta \to \infty$，我们有 $c_\kappa/\beta \to 0$，因此

$$\arcsin\Big(\min\{1,\ c_\kappa/\beta\}\Big) = \arcsin(c_\kappa/\beta) \sim c_\kappa/\beta,$$

这意味着 $|S_\beta| = O(1/\beta)$. □

A.5. 紧致性点

引理 A.5.1（最大斜率的取值位置）。$|f'(x)| = \beta\omega_0$ 当且仅当 $x = k\pi/\omega_0$， $k \in \mathbb{Z}$。 **证明**。同时要求 $|\cos(\omega_0 x)| = 1$ 和 $\sin(\omega_0 x) = 0$。 □

A.6. 多变量推广

对于逐元素激活函数， $z = Wx + b \in \mathbb{R}^m$ 和$y = \phi_{\beta,\omega_0}(z)$逐元素应用，其对角雅可比矩阵满足 $\|\nabla_z y\|_2 \leq \beta\omega_0$（引理A.1.2），与 W复合后即可得到命题A.2.1。

B. 实验细节

所有实验均按照论文中描述的相同实验方案进行。首先，对 HOSC β 参数进行消融研究。随后，与已知方法进行比较。对于每项实验，我们都会介绍实现和实验细节，并提供论文中省略、留待补充材料呈现的更多结果。

定性结果。为便于检查我们的结果，我们在补充材料中提供了一个 HTML 实验套件。入口页面（https://hosc-nn.github.io/）链接到图像、音频、视频、NeRF 和有符号距离函数拟合的各实验页面。每个页面展示与相应表格/图示相同的方法和测试案例，并提供交互式浏览功能（例如图像和视频滑块、音频播放–暂停控件），同时像定量结果中一样突出显示表现最佳的方法。

源代码。我们将公开发布本文所介绍的所有实验的完整源代码，包括训练脚本、评估流程以及 HTML 实验套件，以促进结果复现和进一步研究。

方法比较设置。我们将 HOSC 与 PEMLP (FF) [32], Gauss [25], WIRE [27],[30] 以及 FINER [17]等方法进行比较。对于每种基线方法，超参数均根据其原始论文中提供的规范，或其相应参考代码库中的实现进行配置。具体而言，对于 PEMLP (FF)，傅里叶特征映射共使用了 10 个频率。Gauss 的尺度参数设为 30。对于 WIRE，配置为将初始（ω_0）和隐藏（ω_h） ω 参数均设为 20，同时将缩放因子设为 10。对于 SIREN，采用统一设置：初始 ω_0 设为 30，后续所有逐层 ω 参数均统一设为 1。最后，FINER 的具体参数配置直接采用其官方实现中所描述的设置。

B.1. 图像拟合

数据。我们从 DIV2K 数据集 [1] 的测试集中选取了一个包含 16 个示例的子集。每幅图像的分辨率均为 512×512。值范围和输入像素坐标均缩放至 $[-1, 1]$。未进行额外的预处理。之后，我们在推理和结果收集阶段对这两个范围进行重新缩放。

网络与超参数。在所有图像拟合实验中，我们均使用一个包含 3 个隐藏层、每层宽度为 256 的多层感知机。我们训练 5000 轮，每次迭代都拟合真实图像中的每个像素，批量大小等于图像中的像素数量。我们使用 Adam 优化器，并为每种方法采用特定的学习率；对于每个训练样本，学习率始终保持一致。我们采用各激活函数推荐的学习率。对于 FINER、SIREN 和 HOSC,

表9。**音频 HOSC β-消融实验**。以巴赫和计数示例为展示对象。

示例 / 均方误差 ↓	HOSC β							
	0.1	0.2	0.5	0.8	1.0	2.0	4.0	5.0
Bach	6.6636e-04	2.2016e-04	1.7386e-05	1.0024e-05	9.8558e-06	2.0293e-06	7.5893e-07	1.3194e-06
计数	9.8662e-04	5.2919e-04	4.1339e-04	4.0398e-04	4.3283e-04	2.5501e-04	6.8511e-05	3.6126e-05
	6.0	8.0	10.0	12.0	14.0	16.0	18.0	20.0
Bach	2.3844e-06	2.1383e-05	9.4639e-05	4.3760e-04	2.3118e-02	2.4166e-02	2.4180e-02	2.4184e-02
计数	9.3554e-05	1.7069e-04	8.9352e-04	6.7790e-03	7.2318e-03	7.0961e-03	7.3637e-03	7.3835e-03

表10。**种子独立性**。HOSC 及相关方法的种子独立性评估。

示例 / 均方误差 ↓	SIREN		FINER		HOSC	
	Mean	Std	Mean	Std	Mean	Std
Bach	5.6905e-06	3.9998e-07	1.2704e-3	4.0961e-03	7.1032e-07	5.8887e-08
计数	3.853e-04	4.5750e-06	2.8612e-03	1.6933e-03	4.4506e-05	5.6223e-06

我们使用学习率 LR= $5e-04$；对于 Gauss 和 WIRE，使用学习率 LR= $5e-03$；对于 PEMLP，使用学习率 LR= $1e-3$。我们还对所有方法采用一致学习率进行了实验，发现上述学习率能够取得更好的重建质量。

硬件与运行时间。这些网络使用一块配备 40GB 显存的 NVIDIA A100 GPU 进行训练。我们训练 5,000 个训练轮次，拟合并评估单张图像约需 3 分钟。消融研究中，我们考察了 256 种配置，总训练时间约为 16 小时。对比研究耗时 4.8 小时。

B.2. 音频拟合

数据。为学习音频信号，我们使用两个样本：第一个样本是巴赫《第一号大提琴组曲》的前 7 秒，代表音乐数据；另一个是男性演员从 0 数到 9 的库存音频，代表人类语音数据。音频信号取自波形文件，采样率为每秒 44100 个样本。数值和坐标均缩放至 $[-1,1]$的范围。我们使用所有采样点作为训练数据。我们在 HTML 实验套件（https://hosc-nn.github.io/）中提供了计数样本的更多结果，并在表格中给出了扩展的 β 参数消融实验，见表3 和9。

网络与超参数。在所有音频拟合实验中，我们均使用一个包含 3 个隐藏层、每层宽度为 256 的多层感知机。我们训练 5000 轮次，每次迭代都使用完整音频波形进行拟合，批量大小等于采样点数。我们使用 Adam 优化器，学习率针对不同方法分别设置，与图像实验的设置完全相同。

硬件与运行时间。网络使用一块配备 40GB 内存的 NVIDIA A100 GPU 进行训练。我们训练 5,000 轮次，拟合并评估单个音频波形约需 2 分钟。消融研究包含 90 种配置，总训练时间约为 30 分钟。对比研究耗时 1 小时。

种子独立性。我们通过使用一组有限的指定随机种子训练音频重建模型，评估种子选择对结果质量的影响：$\{0,2,4,8,12,21,42,1337,3333,3407,2025\}$。我们发现，SIREN 和 HOSC 对随机种子的依赖都很小；而 FINER 由于其本身依赖初始化，因此受随机种子变化的影响显著更大，见表 10。

B.3. 视频拟合

数据。对于视频学习，我们使用一段包含 300 帧、分辨率为 512×512 的猫的视频。数值范围缩放至 $[-1,1]$。完整的定量结果见论文。

网络与超参数。对于所有视频拟合实验，我们均使用一个包含 5 个隐藏层、每层宽度为 1024 的多层感知机。我们训练 100 000 个训练轮次，每次迭代都对整个视频进行拟合，批量大小等于整个视频的大小。所有方法均使用 Adam 优化器，学习率为 $1e-04$。我们发现，针对图像和音频重建采用特定方法的学习率，所得结果并不明确。使用特定方法的学习率时，PEMLP 的峰值信噪比达到 37.12，优于 SIREN；但其他方法的峰值信噪比均难以超过 14.00。论文中我们展示了

表11。 **定量比较**。在八个 NeRF Synthetic [20] 样本上取平均值。

	统计量	PEMLP	Gauss	WIRE	SIREN	FINER	HOSC
PSNR	均值。	23.9840	26.1985	29.2248	29.7097	29.8378	29.7321
	Std.	3.4295	3.7701	4.9856	4.6224	4.5956	4.6157
SSIM	均值。	0.8811	0.8877	0.9303	0.9371	0.9371	0.9388
	Std.	0.0779	0.0852	0.0772	0.0653	0.0660	0.0622
LPIPS	均值。	0.1078	0.0856	0.0485	0.0355	0.0352	0.0353
	Std.	0.0681	0.0617	0.0575	0.0302	0.0313	0.0277

r一致学习率比较的结果 nly.

硬件与运行时间。 这些网络使用一块配备 80GB 内存的 NVIDIA H100 GPU 进行训练。我们训练 100 000 个训练轮次，拟合并评估单个视频约需 10 小时。消融研究中，我们考虑 5 种配置，总计需要 50 小时的训练时间。对比研究耗时 60 小时的训练时间。

B.4. 神经辐射场（NeRF）

数据。 在神经辐射场实验中，我们使用 NeRFSynthetic 数据集 [20]，该数据集包含 8 个示例。我们采用 WIRE 中描述的实验设置，仅使用 25 张图像进行训练，并将每张图像下采样至分辨率 200×200， $4\times$ 从原始分辨率 800×800进行下采样。

网络与超参数。 在所有 NeRF实验中，我们均使用一个具有 4 层、每层宽度为 182 的多层感知机。我们训练 37500 次迭代。我们使用 Adam 优化器，并采用经过调度的学习率，其起始值为 2e-4，最终值为 1e-15，所有方法均保持一致。由于学习率调度器的影响，我们不考虑特定于方法的学习率。

硬件与运行时间。 网络使用一块配备 40GB 显存的 NVIDIA A100 GPU 进行训练。我们训练 30 000 次迭代，拟合并评估单个 NeRF 大约需要 30 分钟。对于消融研究，我们考虑 56 种配置，总训练时间约为 30 小时。对比研究耗时 15 小时。

B.5. 有符号距离场

数据。 在有符号距离场实验中，我们使用来自斯坦福三维扫描存储库 [13], 的样本，包括犰狳、龙、泰国雕像和露西。模型顶点经过归一化，以适配单位立方体。

网络与超参数。 所有有符号距离函数实验均采用一个宽度为 256、包含 3 层的多层感知机。所有方法均使用 Adam 优化器，并采用预设的学习率 $1e-04$。采用由粗到细损失函数 [16] 进行训练。训练共进行 200 000 次迭代。每次迭代中，随机采样 10 000 个点并输入模型。为了进行可视化，提取 512^3 网格。

硬件与运行时间。网络使用一块配备 40GB 内存的 NVIDIA A100 GPU 进行训练。我们训练 200 000 次迭代，拟合并评估单个有符号距离函数约需 3 小时。消融研究中，我们考察了 11 种配置，总训练时间约为 30 小时。对比研究耗时 24 小时。

B.6. 与编码方法的正交性

为验证 HOSC 的可扩展性，并探讨激活函数设计与空间编码方法之间的关系，我们分别使用 PyTorch 和 *Tiny-Cuda-NN* [22] 实现了 HOSC，并在十亿像素级图像上进行训练。我们将 HOSC 与基于激活函数的方法（SIREN、PEMLP）以及 Instant-NGP 的 HashGrid编码进行比较 [23], ，后者代表了大规模空间数据领域的最先进水平。结果见图 12 和表 14、15。

数据。在十亿像素图像拟合实验中，我们使用了“戴珍珠耳环的少女”画作修复版，©Koorosh Orooj（CC BY-SA 4.0），分辨率为 8000×9302。

网络与超参数。 我们使用一个包含 3 个宽度为 256 的隐藏层的多层感知机。所有方法均使用 Adam 优化器，学习率为 1×10^{-4}。我们比较了四种方法：PEMLP、SIREN、HOSC，以及 Instant-NGP的哈希网格编码 [23]。训练共进行 1000 个训练轮次，批量大小为 2,000,000（占整个图像的 ≈3%）。

硬件与运行时间。 网络使用一块配备 40GB 显存的 NVIDIA A100 GPU 进行训练。训练 1000 个训练轮次，每种配置约需 3 小时。 β消融实验（表 14）包含 10 种配置，总计约需 30 小时。对比研究（表15）总共需要 21 小时。

结果：HOSC 与激活函数基线方法的比较。 HOSC 达到**23.45 dB 峰值信噪比**（PyTorch）和 22.71 dB（TCNN）（表 15），在这两种框架中均优于 SIREN（23.19 dB / 21.67 dB）和 PEMLP（23.03 dB / 22.51 dB）。

表 12. **有符号距离场 HOSC β 消融实验**。在 犰狳 形状上的实验。

	HOSC β										
指标	0.1	0.3	0.5	0.8	0.9	1.0	3.0	5.0	8.0	12.0	16.0
倒角	5.280e-06	4.814e-06	5.638e-06	4.447e-06	5.581e-06	4.324e-06	5.676e-06	6.057e-06	5.877e-06	5.293e-06	7.673e-06
IoU	0.9774	0.9778	0.9728	0.9817	0.9762	0.9821	0.9738	0.9729	0.9734	0.9808	0.9670

表 13. **有符号距离函数重建**。HOSC $\beta = \{0.8, 0.9, 1.0\}$

	样本	Gauss	WIRE	SIREN	FINER	HOSC $\beta = 0.8$	HOSC $\beta = 0.9$	HOSC $\beta = 1.0$
Chamfer ↓	犰狳	2.2065e-05	4.3975e-06	4.8958e-06	4.3178e-06	4.4470e-06	5.5813e-06	4.3237e-06
	龙	1.9011e-05	2.5875e-06	3.1763e-06	3.0937e-06	3.3799e-06	3.9059e-06	3.7843e-06
	Thai	3.7762e-05	2.8232e-05	3.1437e-05	2.9137e-05	3.0911e-05	3.2658e-05	3.1313e-05
	Lucy	2.5397e-05	2.1952e-05	2.2656e-05	2.2189e-05	2.2966e-05	2.3785e-05	2.2929e-05
IoU ↑	犰狳	0.9766	0.9799	0.9775	0.9809	0.9817	0.9762	0.9821
	龙	0.9506	0.9617	0.9584	0.9610	0.9546	0.9508	0.9498
	Thai	0.8310	0.8328	0.8338	0.8297	0.8366	0.8326	0.8229
	Lucy	0.8625	0.8719	0.8731	0.8692	0.8727	0.8675	0.8724

β-消融实验（表 14）表明其在 $\beta \in [0.5, 12.0]$方面具有稳健性，其中 PyTorch 在$\beta = 5.0$ 处达到峰值性能（23.45 dB），TCNN 则在 $\beta = 8.0$ 处达到峰值性能（22.71 dB）。视觉结果（图 12）显示，HOSC 生成的高保真重建结果在两种实现中均可与 SIREN 和 PEMLP 相媲美。

结果：HOSC 与空间编码（INGP）的比较。
Instant-NGP 的 HashGrid编码 达到 25.19 dB（表 15），比基于激活函数的方法高约 2 dB。这一差距并不意外：HashGrid 使用多分辨率哈希表显式建模空间相关性，而 HOSC （以及所有激活函数）则作用于单个坐标，不包含空间结构。值得注意的是，HOSC 和 HashGrid 是 **正交的设计选择**——HashGrid 是输入编码，HOSC 是激活函数——未来的工作有望将二者结合。

讨论。权衡取舍及 HOSC 的适用场景。HashGrid 在大规模空间数据（图像、NeRF）上表现出色，但需要进行特定领域的架构修改，并为哈希表额外占用内存。HOSC 则提供了一种更简单、基于激活函数的替代方案，无需专用编码即可适用于不同领域：同一个激活函数在图像（第 4.1 节）、音频（第 4.2 节）、视频（第 4.3 节）、NeRF（第 4.4 节）和有符号距离函数（第 4.5 节）上均取得了优异结果。对于优先考虑架构简洁性、跨领域通用性或通过 β实现显式梯度控制的应用，HOSC 是一个颇具吸引力的替代方案。在 PyTorch 和 TCNN 实现中表现稳定（表 15），进一步证明了 HOSC 的稳健性和实际适用性。

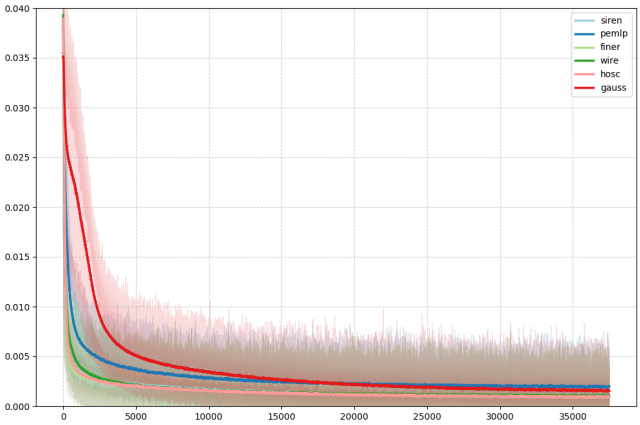

图11。**特定方法的损失曲线**。NeRF实验中所有示例的平均值。

表 14。**千兆像素 HOSC β-消融实验**。

	HOSC β				
框架	0.5	1.0	5.0	8.0	12.0
TCNN	21.05	21.73	22.57	22.71	22.69
PyTorch	23.20	23.09	23.45	22.70	22.50

表 15. **千兆像素重建的定量比较**。

	TCNN			PyTorch		
指标	INGP	PEMLP	SIREN	HOSC	PEMLP	SIREN	HOSC
PSNR	25.19	22.51	21.67	22.71	23.03	23.19	23.45



图 12。千兆像素重建的定性比较 。每行使用列中指定的方法重建真实值数据。定量 -
（峰值信噪比）结果展示 于下方重建图像的每个示例中。HOSC $\beta = 8.0$。适用于 TCNN， $\beta = 5.0$ 适用于 PyTorch。