arXiv:2301.05187v1 [cs.CV] 5 Jan 2023

WIRE：小波隐式神经表示

Vishwanath Saragadam、Daniel LeJeune、Jasper Tan、
Guha Balakrishnan、Ashok Veeraraghavan、Richard G.
Baraniuk 莱斯大学https://vishwa91.github.io/wire

摘要

隐式神经表示（*INR*）近年来推动了众多视觉相关领域的发展。*INR*的性能很大程度上取决于其多层感知机（*MLP*）网络所采用的非线性激活函数。研究人员探索了多种非线性函数，但遗憾的是，当前旨在实现高精度的*INR*也存在鲁棒性较差的问题（例如对信号噪声、参数变化等较为敏感）。受调和分析启发，我们开发了一种兼具高精度和鲁棒性、且不存在这种权衡的新型*INR*。**小波隐式神经表示（WIRE）**采用连续的**复Gabor小波**激活函数。众所周知，该函数在时频上具有最佳集中性，并且在表示图像方面具有出色的偏置特性。通过大量实验（图像去噪、图像修复、超分辨率、计算机断层扫描重建、图像过拟合，以及基于神经辐射场的新视角合成）证明，*WIRE*在*INR*精度、训练时间和鲁棒性方面达到了新的最先进水平。

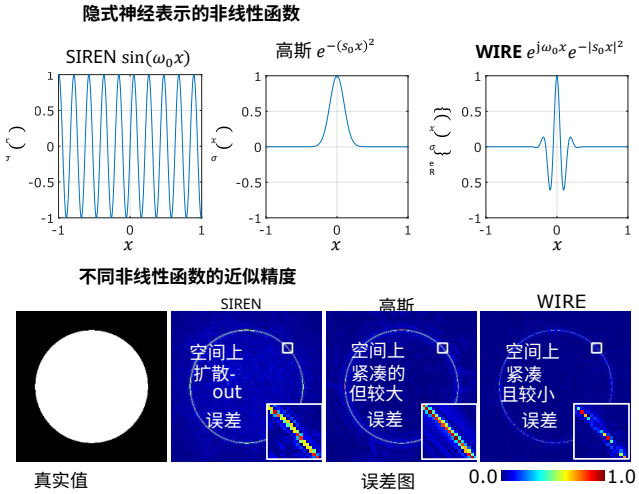


图1. **小波隐式神经表示（WIRE）**。我们提出一种基于连续复Gabor小波的新型隐式神经表示（INR）非线性函数，它对视觉信号具有很强的表示能力。上排展示了两种常用的非线性函数：采用正弦非线性函数的 SIREN、采用高斯非线性函数的方法，以及使用连续复Gabor小波的 WIRE。WIRE兼具正弦函数的频率紧致性和高斯非线性函数的空间紧致性。下排展示了近似含有明显边缘的图像时的误差图。SIREN会产生全局振铃伪影，而高斯非线性函数则会在边缘处产生紧致但较大的误差。WIRE产生的误差最小且空间上最为紧致。因此，WIRE能够快速、准确地学习表示，同时对噪声和数据欠采样具有鲁棒性。

1. 引言

隐式神经表示（INR）通过在一组数据点上学习连续函数，已成为一种前景广阔的通用信号处理框架。隐式神经表示由多层感知机（MLP）构成，其中线性层与逐元素非线性激活函数交替排列。得益于多层感知机，隐式神经表示不受局部性偏置的影响，而这种偏置可能会限制卷积神经网络（CNN）的性能。因此，隐式神经表示在众多与视觉相关的领域推动了最先进水平的发展，包括计算机图形学 [24, 29, 30]，图像处理 [11]，逆问题 [43]，和信号表示 [42]。

目前，隐式神经表示仍面临诸多障碍，限制了其应用。首先，对于某些应用，尤其是涉及三维体数据等高维数据的应用，要将隐式神经表示拟合到较高精度仍可能耗时过长（数十秒），这实时应用。其次，隐式神经表示对信号噪声或测量值不足并不鲁棒。事实上，文献中大多数关于隐式神经表示的研究都假设信号噪声几乎为零且数据量充足。我们在自己的实验中发现，当前的隐式神经表示方法在去噪或超分辨率等任务上效果不佳。最后，隐式神经表示的表示精度仍有提升空间，尤其是在精细细节方面。

在本文中，我们开发了一种新颖、更快速、更精确且更鲁棒的隐式神经表示，以解决这些问题并将其性能提升至新的水平。为此，我们从调和分析中汲取灵感，并重新审视多层感知机所使用的非线性激活函数。近期研究表明，隐式神经表示可以被解释为一种结构化信号表示字典 [53], ，其中激活函数的非线性决定了字典中的原子。例如，正弦激活函数会生成一种信号的伪傅里叶变换表示，使其在频域中达到最大集中 [53]。

从过去四十年的调和分析研究中可以得出一个重要结论：对于典型视觉任务 [26]中出现的信号，傅里叶方法并非最优。这类信号，例如照片中的自然图像，使用在时空-频率上达到最优集中的小波 原子表示时，能够更加简洁、稳健。众所周知，由小波原子构成的稀疏组合非常适合表示图像；相关开创性工作可见于计算机视觉（例如拉普拉斯金字塔）、计算神经科学 [32], 以及 JPEG2000 压缩标准。

本文介绍了 **小波隐式神经表示（WIRE）**，这是一种基于 复*Gabor*小波 激活函数的新型隐式神经表示（参见图 1）。通过广泛的实验，我们证明 WIRE 在隐式神经表示的精度、训练时间和鲁棒性方面达到了新的最先进水平。我们展示了 WIRE 增强的鲁棒性对于解决困难的视觉逆问题尤为有用，包括图像去噪（鲁棒性）、图像修复和超分辨率（更优的插值），以及二维计算机断层扫描（CT）重建（解决更高维的逆问题）。在图像过拟合和学习点云占据体等信号表示任务中，WIRE 的表现也优于其他隐式神经表示。最后，我们展示了 WIRE 如何利用神经辐射场（NeRF），从极少的训练视图中实现更快速、更稳健的新视角合成 [29] 。

2. 相关工作

逆问题的正则化。逆问题涉及根据一组线性或非线性测量值估计信号。测量值不可避免地会受到噪声（例如相机读出噪声或光子噪声）的影响，或者问题本身是病态的，因此需要进行正则化。正则化有多种形式，包括岭回归、套索回归 [46], 全变分（TV） [10], 和基于稀疏性的 [7] 技术，这些方法旨在惩罚信号或其某种变换的 ℓ_1范数。过去十年间，数据驱动正则化方法得到发展，包括基于超完备字典的 [4] 方法和基于生成网络的 [31, 36,37] 方法。对于严重病态的问题，经典的基于模型的方法并不适用，而数据驱动方法则严重依赖数据。

卷积神经网络（CNN）。在过去十年中，CNN 一直是计算机视觉领域最受欢迎的神经网络架构。研究表明，这类网络具有强烈的隐式偏置，倾向于生成类似图像的信号。深度图像先验（DIP） [48]及其变体 [15, 21]便展示了这一点：即使没有任何先验训练数据，它们也能在图像相关线性逆问题上取得出色成果。然而，这类基于 CNN 的先验依赖于离散网格状的信号表示，因此不适用于新视角合成等问题，也不适用于求解常微分方程和偏微分方程；此外，它们也难以扩展到维度极高的信号，例如三维断层成像体、十亿像素图像或大型点云。

深度图像先验。神经网络，尤其是 CNN，会因其特定架构（例如 UNet [38]）而表现出隐式偏置，这意味着即使未经训练的神经网络也可用于正则化。基于这一点，研究者构建了深度图像先验（DIP） [48] ，其生成的输出往往具有图像的外观。其核心思想是将正则化重新表述为针对每个问题实例优化网络权重。DIP 的性能显著优于经典正则化方法。然而，DIP 只有在过参数化时才能取得良好性能，并且依赖于信号的网格状离散表示，这意味着它无法扩展到高维信号，例如包含大量点的点云。深度解码器 [21] 和 DeepTensor [39], 在一定程度上解决了计算成本问题，但它们仍要求将信号定义为规则数据网格，例如二维矩阵或三维张量。

隐式表示。 隐式神经表示是基于多层感知机（多层感知机）的连续可学习函数近似器。隐式神经表示的连续特性在处理不规则采样的信号（例如点云）时尤为吸引人。自从隐式神经表示首次广泛应用于计算机图形学中的新视角合成以来 [29], ，它们已遍及视觉与信号处理的几乎所有领域，包括渲染 [24], 、计算成像 [6, 12], 、医学成像 [51],和虚拟现实 [16]。

标准神经网络中广受欢迎的 ReLU 非线性函数，已有实证表明会导致隐式神经表示的近似精度较差。为此，人们对多层感知机进行了多种改进，包括所谓的位置编码 [30, 44], ， 以及各种非线性函数的选择，例如正弦函数 [42] 和高斯函数 [34]。与之密切相关的一项工作是基于 Gabor小波的乘法滤波网络（MFN），其每一层的输出会进行多重

相乘，并经过 Gabor 滤波器。由此得到的输出是指数数量级的 Gabor小波组合，因此具有很大的容量。此外，人们还提出了许多架构改动，利用视觉信号的多尺度特性来加速隐式神经表示的训练过程，包括自适应块分解 [27], kilo-NeRF [35], 以及预测信号的拉普拉斯金字塔 [40]。

得益于这些众多进展，隐式神经表示如今几乎可以瞬间在信号上完成训练 [30]。然而，这类隐式神经表示的高容量使其缺乏鲁棒性——这意味着其信号表示十分脆弱，导致模型对噪声和信号同样过拟合。本文提出将复Gabor小波用作非线性函数，它特别适合提升隐式神经表示的鲁棒性。

小波变换与Gabor小波。 傅里叶变换将信号分解为一系列正弦波之和，这些正弦波具有无限的空间支撑，因此没有空间紧致性的概念。小波变换通过将信号分解为短振荡脉冲（称为小波）的平移和缩放版本的线性组合，弥补了这一不足。与傅里叶变换相比，小波通常能更快地逼近信号和图像 [17], ，因此常用于图像压缩 [18, 41] ，并作为图像 [20] 和视频 [49]逆问题的鲁棒先验。本文表明，由于小波在空间和频率上具有紧支撑，因此能带来更快的逼近速度，是隐式神经表示中非线性函数的普遍优选。

3. 小波隐式表示

3.1. 隐式神经表示的详细介绍

考虑一个隐式神经表示函数 $F_\theta: \mathbb{R}^{D_i} \mapsto \mathbb{R}^{D_o}$，它将 D_i 个输入维度映射到 D_o 个输出维度，其中 θ 表示多层感知机的可调参数。目标是构造 F_θ，使其能够逼近感兴趣的函数 $g(\mathbf{x})$，即 $g(\mathbf{x}) \approx F_\theta(\mathbf{x})$。例如，$g(\mathbf{x})$ 可以简单地表示真实图像，即将坐标映射到像素值的函数。将 $F_\theta(\cdot)$ 建模为一个 M 层多层感知机时，每一层的输出如下所示：

$$\mathbf{y}_m = \sigma(W_m \mathbf{y}_{m-1} + \mathbf{b}_m), \tag{1}$$

其中，σ 是非线性函数（或非线性激活函数）；W_m, $\mathbf{b}_m$ 是第 m^{th} 层的权重和偏置；$\mathbf{y}_0 = \mathbf{x} \in \mathbb{R}^{D_i}$ 是输入坐标，$\mathbf{y}_{M+1} = W_{M+1}\mathbf{y}_M + \mathbf{b}_{M+1}$ 是最终输出。

非线性激活函数 σ 在隐式神经表示的表示能力中起着关键作用（见图1）。两种主要选择包括 SIREN 中使用的周期函数 $\sigma(x) = \sin(\omega_0 x)$ [42], ，以及 Ramasinghe等人使用的高斯非线性 $\sigma(x) = e^{-(s_0 x)^2}$ ——二者都能实现显著高于 ReLU 的表示精度。然而，它们强大的表示能力也有弊端，因为它们表示噪声的精度几乎与表示图像一样高。我们的目标是提出一种非线性函数 σ，它适用于图像、视频和三维体数据等视觉信号，但对类噪声信号的拟合效果较差。

3.2. WIRE

基于Gabor小波能够实现最优时频紧致性的这一认识，我们提出了小波隐式表示（WIRE），它采用**连续复Gabor小波 ψ 作为其激活非线性**：

$$\sigma(x) = \psi(x; \omega_0, s_0) = e^{j\omega_0 x} e^{-|s_0 x|^2}, \tag{2}$$

其中，ω_0 控制小波的频率，s_0 控制其扩展范围（或宽度）。第一层的激活形式为

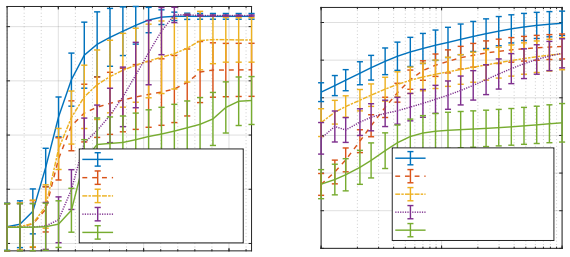
$$\mathbf{y}_1 = \psi(W_1 \mathbf{x} + \mathbf{b}_1; \omega_0, s_0), \tag{3}$$

它们是母Gabor小波 ψ 在由 W_1 和 $\mathbf{b}_1$决定的尺度和平移下的副本。因此，WIRE 的构建单元取自小波原子字典。我们令隐式神经表示（INR）的权重及其输出均为复数值，以便在整个过程中保留相位关系；对于实信号，只需取输出的实部并舍弃虚部即可。正如小波兼具空间紧致性和频率紧致性，WIRE 凭借复指数项而具有 SIREN 等周期性非线性函数的优势，并通过高斯窗项实现空间紧致性；请回顾图1。此外，与 SIREN 不同，WIRE 无需精心挑选一组初始权重（见图 3），因为高斯窗会使每一层的输出在空间上紧致；使用与参数 ω_0无关的默认神经网络初始化（权重均匀随机初始化）也能产生高质量结果 s_0。

3.3. WIRE 的隐式偏置

神经切线核视角。 如前所述，我们希望找到一种隐式神经表示，使其能够很好地拟合视觉信号，但相较之下不易拟合噪声。受 [53], 提出比较经验神经切线核（NTK）的特征函数，以 [22]理解其逼近性质的启发，我们通过神经切线核梯度流来比较对含噪自然图像的拟合。隐式神经表示的神经切线核梯度流能够准确捕捉神经网络训练早期的行为，因此，在去噪等通过提前停止进行正则化的任务中，训练早期的行为决定了隐式偏置。在宽神经网络的惰性训练机制下 [25], 时刻的拟合图像 $t \geq 0$ 的值为

$$F_{\theta_t}(\mathbf{x}) = [(I - e^{-tK})g](\mathbf{x}), \tag{4}$$



(a) 神经切线核下的隐式偏置梯度流

(b) 标准隐式神经表示中的隐式偏置。

图2。**去噪中的隐式偏置**(a) 有限宽度隐式神经表示的经验神经切线核揭示了隐式神经表示的隐式偏置。神经切线核梯度流的早期轨迹表明，WIRE 比噪声更快地收敛到图像，相较于所有其他非线性函数表现更优。柱状条表示数据集上的一个标准差。(b) 标准训练的早期迭代表现与(a)部分中神经切线核梯度流的相对性能高度一致。此外，在训练剩余阶段，WIRE 始终优于其他非线性函数。

其中， I 是恒等算子， K 是作用于图像空间域的神经切线核算子，而 g 是正在拟合的图像。

在图中。2a，我们使用经验有限宽度神经切线核，对去噪任务应用神经切线核梯度流，以拟合加入了 $\mathcal{N}(0,\ 0。052)$ 独立同分布逐像素加性噪声的原始图像。由于计算神经切线核的开销很大，我们在 $64 \times 64 \times 3$ 张 Tiny I mageNet 图像上进行评估 [2]。将 WIRE 与其他隐式神经表示进行比较后，我们发现，正如预期，WIRE 在训练早期更倾向于学习图像中的信号，而非噪声，并且能够快几个数量级地收敛到几乎任意给定的峰值信噪比（PSNR）。

实证评估。我们在 Kodak无损真彩色图像集的 24 张 $768 \times 512 \times 3$ 图像 [1], 上再次进行 $\mathcal{N}(0,\ 0。052)$ 加性噪声下的去噪任务，其设置类似于针对真实隐式神经表示的神经切线核分析，结果见图2b。我们采用与神经切线核示例相同的隐式神经表示，但使用常规神经网络梯度优化进行训练，而非神经切线核梯度流。与前述结果相同，WIRE 的表现远胜其他隐式神经表示，以快一个数量级的速度收敛至相同的峰值信噪比。

3.4. 选择参数 ω_0， s_0

WIRE 的性能主要由常数 ω_0、 s_0 决定，它们分别控制正弦函数的频率和高斯函数的宽度。在一系列广泛的参数取值下，WIRE 的表现均优于 SIREN 和高斯非线性函数。图3 展示了 WIRE 在不同参数下达到的近似精度。我们将隐藏层数量设为 3，隐藏

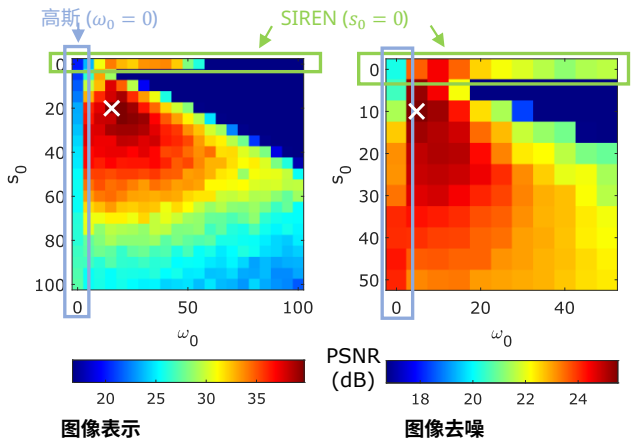


图3。**WIRE 对参数选择具有鲁棒性**。上图展示了在不同的 ω_0 和 s_0 设置下，图像表示和去噪的准确率。方框标出了特殊情况：$\omega = 0$ 对应高斯非线性函数， $s_0 = 0$ 对应 SIREN。在图像表示和图像去噪任务中（以白色十字标记），WIRE 的准确率均高于 SIREN 和高斯。此外，在很大范围的参数 ω_0， s_0 选择下，WIRE 都表现出优异性能，这表明 WIRE 对超参数并不过分敏感。

特征数量设为 256。当 $\omega_0 = 0$时，我们使用高斯非线性函数。当 $s_0 = 0$时，我们使用正弦非线性。当两个参数均为零时，我们使用 ReLU 非线性函数。对于去噪任务，我们添加了光子噪声，其强度相当于每像素最多 50 个光子。从图3 中可以看出，WIRE 的表现优于 SIREN、高斯和 ReLU。此外，在 ω_0、 s_0 的较大取值范围内，WIRE 在图像表示和去噪方面的性能都更出色。对ω_0、s_0 的具体取值不太敏感，这意味着即使缺乏关于图像或噪声统计特性的精确信息，也可以使用 WIRE。

WIRE的其他形式。对于复数权重不可行的问题，可以将WIRE实例化为复Gabor小波的虚部（或实部），$\psi(x; \omega_0, s_0) = \sin(\omega_0 x)e^{-(s_0 x)^2}$。注意，设置 $s_0 = 0$会得到SIREN中使用的正弦非线性函数 [42] ，而$\omega_0 = 0$则会得到高斯非线性函数 [34] ，这表明WIRE继承了先前提出的非线性函数的优良特性。WIRE的另一种形式是常数-Q Gabor小波，其中 $\omega_0 s_0 = Q$，由此得到恒定分数带宽（$\omega/\delta\omega$）。常数-Q Gabor小波常用于音乐分析 [8, 47]、小波变换 [26], 和拉普拉斯变换。由于只有一个参数，在固定 Q的情况下，超参数调优会更加简单。

3.5. 多维局部化

如上所述，WIRE将Gabor母小波 ψ 逐元素应用于线性变换 $W_m\mathbf{y}_{m-1}$的输出。因此，每个单元的输出仅沿一个方向进行空间定位，该方向由 W_m 对应的行决定（见图11 上图）。虽然这种高度各向异性的空间定位适用于某些类型的数据，但许多自然数据（例如摄影图像）最适合用各向同性和各向异性空间定位原子（例如小波和曲波 [9]，或小波和楔波[50]）的组合来表示。为了实现沿多个方向的空间定位，我们在Gabor母小波中加入 $D_m - 1$额外的高斯窗：

$$\mathbf{y}_m = \psi(W_m^{(1)}\mathbf{y}_{m-1} + \mathbf{b}_m^{(1)}; w_0, s_0) \cdot e^{-\sum_{k=2}^{D_m} |s_0(W_m^{(k)}\mathbf{y}_{m-1}+\mathbf{b}_m^{(k)})|^2}. \quad (5)$$

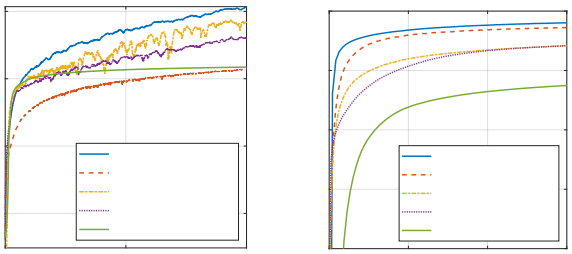
在二维场景中，例如自然图像数据，得到的第一层激活将呈现 Gabor小波与曲波的混合形态（见图11 底部）。正如我们将在下文第4.4节中看到的，二维 WIRE 表示更丰富的空间定位显著提升了其性能（见图12 和图13）。

4. 实验

与最先进的技术相比，WIRE 能够更快地为所有信号类别学习表示。此外，WIRE 非常适合解决一大类逆问题：测量值的数量远少于信号的维度，或者测量值受到噪声干扰。对于以下所有实验，我们都在 PyTorch 中实现了优化过程 [33]，并使用 Adam优化器 [23]。代码运行于一台配备 64GB RAM 和一块显存为 8GB 的 Nvidia RTX 2080 Ti 图形处理器（GPU）的系统单元上。除非另有说明，我们使用了一个 ℓ_2 损失函数来衡量测量值与隐式神经表示（INR）输出之间的差异。没有使用其他正则化。我们使用学习率调度器，在训练轮次结束时将初始学习率衰减为0。1 。

4.1. 信号表示

隐式神经表示的一项常见应用是表示信号。本实验评估了两项任务：表示图像和表示占据体积 [28]。在这两种情况下，我们均使用具有三个隐藏层、每层宽度为 300 个特征的多层感知机，并测试所有非线性函数。对于 WIRE，为了抵消实部和虚部导致的参数数量翻倍，我们将参数数量减半。具体做法是将隐藏特征数量减少 $\sqrt{2}$。针对每种非线性函数所选用的参数以及学习率



(a) 图像表示　　(b) 体积表示

图 4. **WIRE 学习速度更快。** 上方两个图展示了图像（图 5上排）和占据体（图 5下排）随时间变化的表示精度。得益于 Gabor 小波对视觉信号具有很高的近似能力，WIRE 能够更快地实现高精度，因此适合用于表示视觉信号。

经过选择，以获得最快的逼近速度。具体而言，我们为 $\omega_0 = 20$，$s_0 = 10$ WIRE 选择 $\omega_0 = 40$，为 SIREN 选择 $s_0 = 30$，为高斯选择。此外，我们还与乘性频率网络（MFN）进行比较 [19]。对于占据体积，我们在一个 $512 \times 512 \times 512$网格上进行采样，将体积内的每个体素赋值为 1，将体积外的体素赋值为 0。我们对图像评估峰值信噪比（PSNR）和结构相似性指数（SSIM） [52]，并对占据体积评估交并比（IOU）。

图 4 展示了图像（柯达 数据集）和占据体（泰国雕像）的近似精度随时间的变化。WIRE 不仅达到最高精度，而且速度远快于其他方法。图 5 展示了示例图像经过 1.6 分钟后的最终表示，以及泰国雕像经过 30 分钟后通过移动立方体算法构建的三维网格。WIRE 对图像（43.2 dB）和占据体（0.99）均实现了最高精度，这印证了我们的假设：采用 Gabor 非线性的隐式神经表示具有更高的近似精度。

4.2. 求解二维图像的逆问题　　ges

WIRE 的归纳偏置有利于表示图像，因此可用于求解线性逆问题。为展示 WIRE 作为强图像先验的优势，我们展示了它在图像去噪、单图像超分辨率和多图像超分辨率方面的表现。

图像去噪。为了评估隐式神经表示表示噪声信号的鲁棒性，我们在 DIV2K 数据集中的一张高分辨率彩色图像上学习了一个表示 [3]。我们模拟了光子噪声：在每个像素处独立采样一个泊松随机变量，最大平均光子计数为30，读出计数为2，得到的输入峰值信噪比为 17.6 dB。随后，我们学习了一个

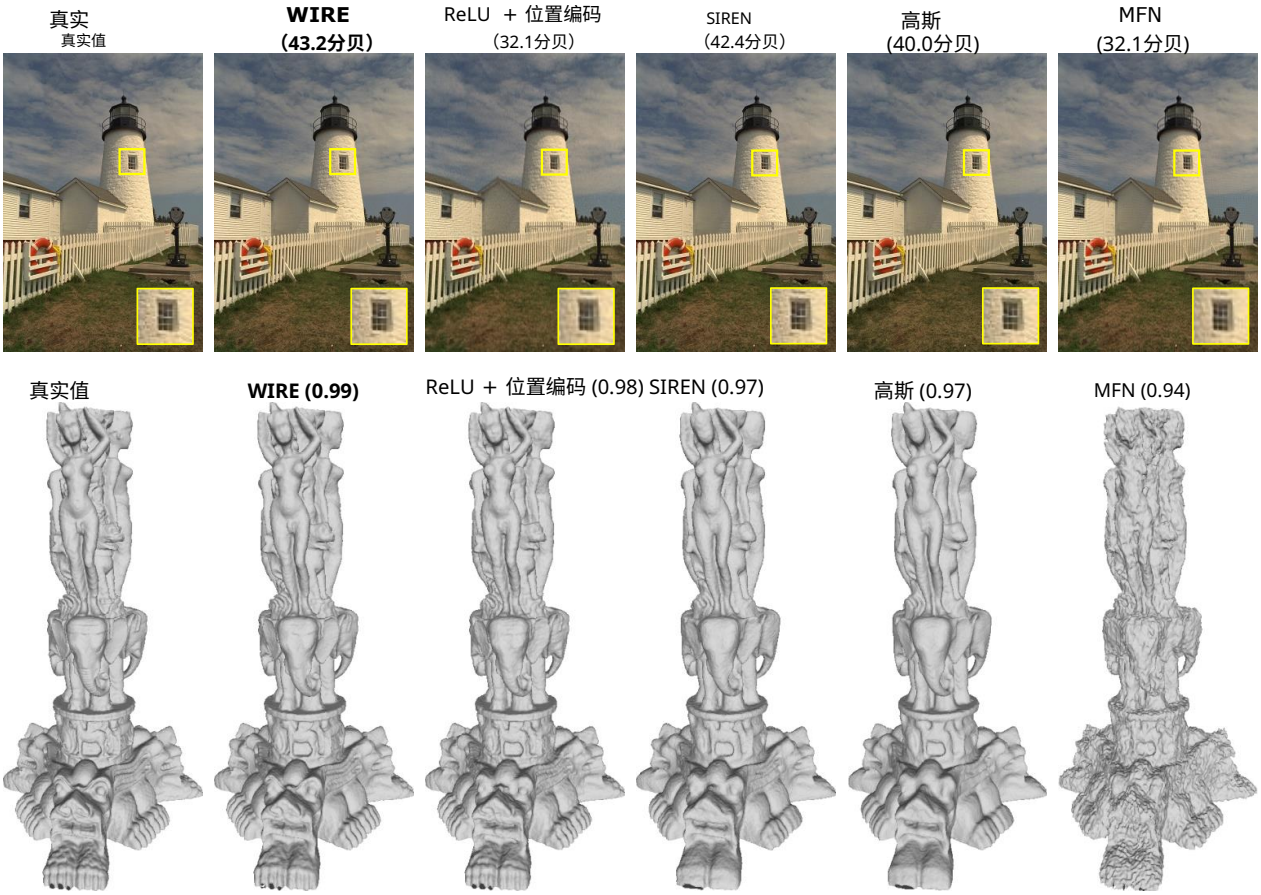


图 5. **WIRE 具有很强的表示能力**。 上方结果展示了第一行中的图像表示，以及第二行中使用不同非线性函数和占据体积生成的网格。对于这两类数据，WIRE 都实现了最高的表示精度，凸显了其作为信号模型的优势。

针对这张含噪图像，使用各种非线性函数学习其表示。所有情况下，我们都选用一个包含两个隐藏层、每层有 256 个特征的多层感知机。我们还将去噪结果与深度图像先验（DIP）[48]进行了比较。图6 展示了每种非线性函数对应的最终结果及其指标。WIRE 生成的图像最清晰，残留噪声也最少。从定性角度看，WIRE 的结果与 DIP 相似，这表明 WIRE 具有使其适用于逆问题的归纳偏置。

图像超分辨率。 隐式神经表示充当插值函数，因此，具有良好隐式偏置的隐式神经表示有望提升超分辨率效果。我们通过在一幅 DIV2K 图像上实现 $4\times$ 超分辨率来验证这一假设。前向算子可表示为 $\mathbf{y} = A_4\mathbf{x}$ 其中 A_4 实现了一个 $4\times$ 下采样算子（无混叠）。随后，我们将 $\mathbf{x}$ 建模为隐式神经表示的输出，以求解清晰图像。图7 展示了采用不同方法对蝴蝶图像进行超分辨率处理的结果。WIRE 生成的结果最为清晰，蝴蝶的触角和翅膀细节锐利分明。与去噪任务一样，WIRE 的结果与 DIP 相似，证明了 WIRE 的广泛适用性。

当需要在不规则网格上进行数据插值时，隐式神经表示尤其具有优势。这类情况的一个例子是多图像超分辨率，其中图像彼此之间存在平移和旋转。图 21 展示了一个使用四幅图像进行 $4\times$ 超分辨率的示例（因此压缩率为 25% ），图像来自柯达数据集 [1]，并模拟了图像之间微小的亚像素运动。此时，前向算子为 $\mathbf{y}^k = A_4^k\mathbf{x}$，其中 A_4^k表示对 k^{th} 图像进行下采样、平移和旋转。图中的可视化结果表明，WIRE 达到了最高精度，并且在重建高频分量方面表现更佳。相比之下，高斯非线性会导致重建结果模糊，而 SIREN 则会产生振铃伪影。

计算机断层扫描（CT）重建。强信号先验对于求解欠约束问题至关重要，CT 重建就是其中一例。我们模拟了—

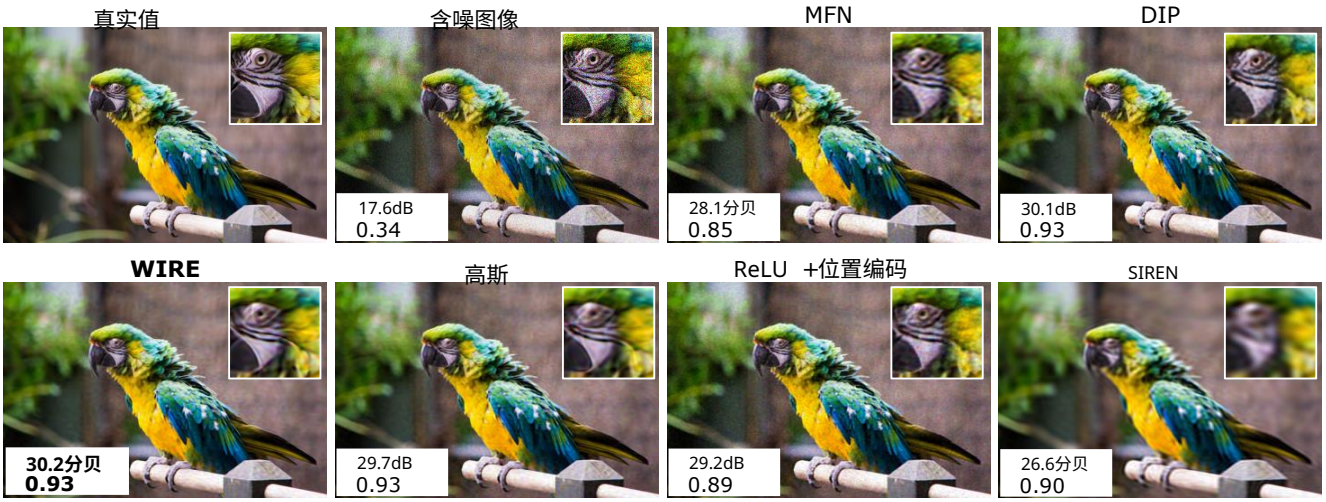


图6。**WIRE 对噪声具有鲁棒性**。WIRE 独有的一项强大特性是对含噪数据的鲁棒性。在此，我们展示了添加散粒噪声后的图像表示，其输入峰值信噪比为 17.6 dB。在各种方法中，WIRE 的峰值信噪比和结构相似性指数均高于其他任何表示，因此能够自然地实现去噪。

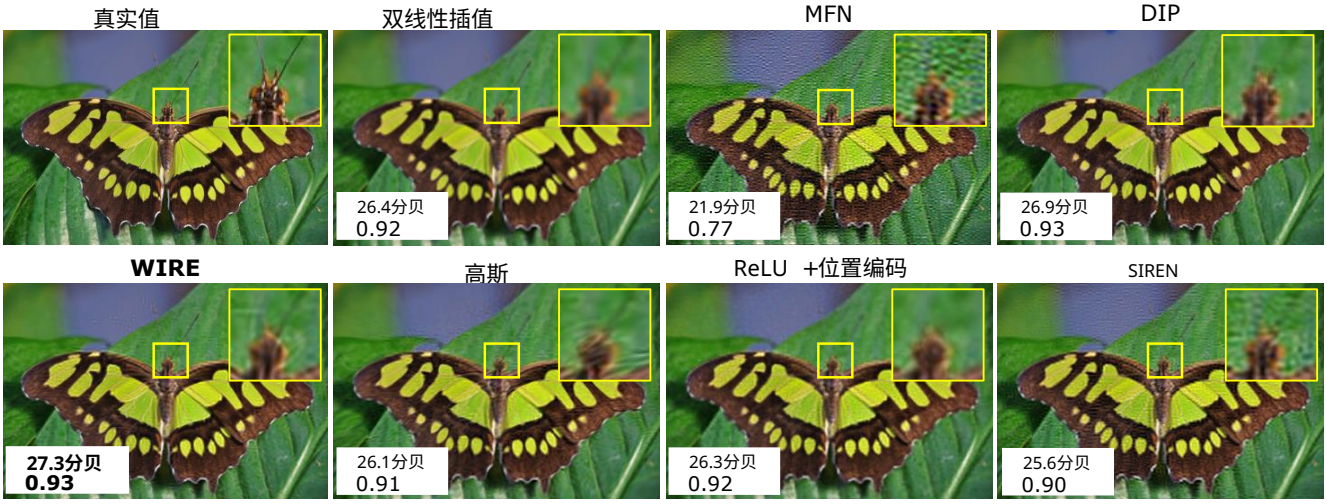


图7。**WIRE 用于单图像超分辨率**。上图展示了采用多种方法进行 4× 单图像超分辨率的结果。得益于其强大的隐式偏置，WIRE 实现了最清晰的重建，且重建指标在定量上更高

100 次 CT 测量值，针对一张 256 × 256 X 射线结直肠图像 [13]。图9 展示了采用不同方法得到的最终重建结果。WIRE 的重建结果最为清晰，特征也十分鲜明。SIREN 的表现次之，但存在条纹伪影，这是未正则化重建中意料之中的现象。高斯非线性函数得到的结果过于平滑。因此，WIRE 可作为一种稳健的先验，用于处理含噪且欠采样测量值的逆问题。

4.3. 学习神经辐射场

隐式神经表示已成功应用于基于神经辐射场（NeRF）的新视角合成 [29]。给定来自稀疏视点集的图像，目标是渲染一张来自训练集之外的不同视点的图像。NeRF通过训练一个通用隐式神经表示来实现这一目标，该表示以三维位置和观察方向作为输入（因此输入为5维），并输出该位置处的透射率和颜色。随后，沿着穿过每个视点透镜（针孔）的射线进行积分，生成图像。最简单的NeRF架构由位置编码以及两个配备ReLU的多层感知机构成，分别用于预测透射率和颜色值。我们展示了无需任何位置编码的WIRE能够在更少的训练轮次内取得更高质量的结果。我们使用合成鼓数据集 [29] 训练NeRF进行重建。每张图像都被下采样至以下分辨率

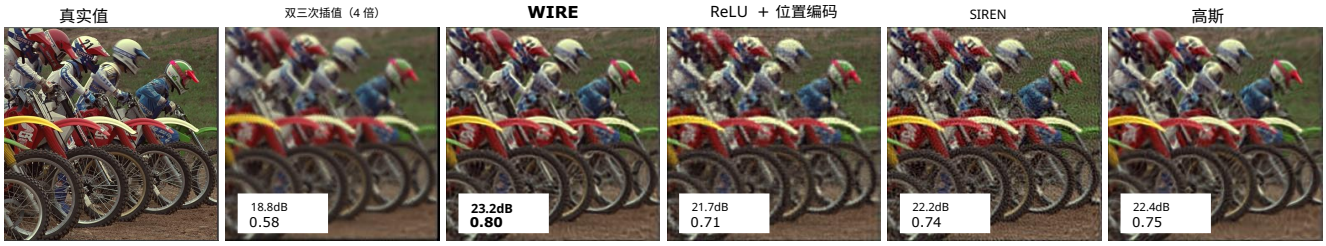


图 8. **多图像超分辨率**。隐式神经表示特别适合处理不规则网格上的数据，例如通过多个亚像素位移拍摄的图像。上图展示了使用 4 张以不同亚像素位移和旋转拍摄的图像进行 4× 超分辨率的结果。随后，我们求解了一个联合逆问题，其中高分辨率图像被建模为隐式神经表示的输出。无论从定量还是定性角度来看，WIRE 都能生成最佳重建结果，这表明 WIRE 对视觉信号具有良好的插值特性。

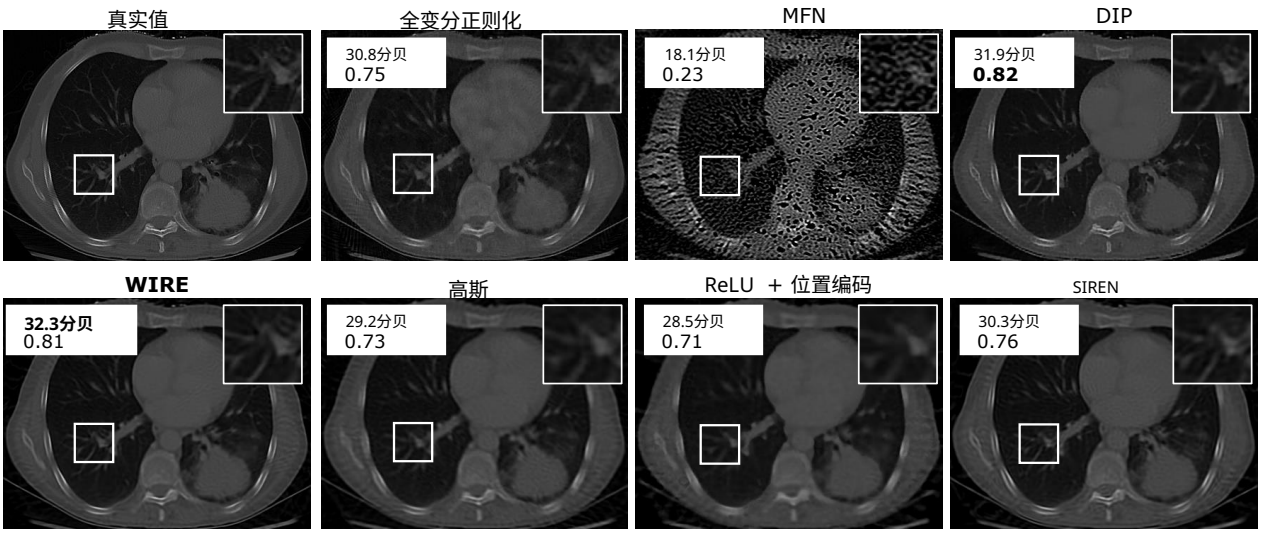


图 9. **计算机断层扫描重建**。对于含噪声的欠采样数据，逆问题需要强大的信号先验，才能实现稳健重建。这里，我们展示了采用多种方法对一幅 256 × 256 图像（压缩率为2.5× ）进行 100 个角度的计算机断层扫描重建。WIRE 得到的重建结果清晰锐利，呈现出了其他方法的重建结果中模糊不清或伴有振铃伪影的特征。因此，WIRE 是一种强大的图像信号先验，能够解决大量逆问题。



图 10. **使用神经辐射场进行新视角合成**。隐式神经表示在新视角合成方面展现出最大的潜力，其中每个三维体素的透射率和颜色都被建模为隐式神经表示的输出。在此，我们展示了 WIRE 非常适合进行新视角合成，且无需额外的位置编码。WIRE 不仅以更少的训练轮次实现了更高的精度（+0.2dB），还捕捉到了其他非线性函数遗漏的细节，例如连接叠音钹与支架的杆，以及钹上的各向异性反射。

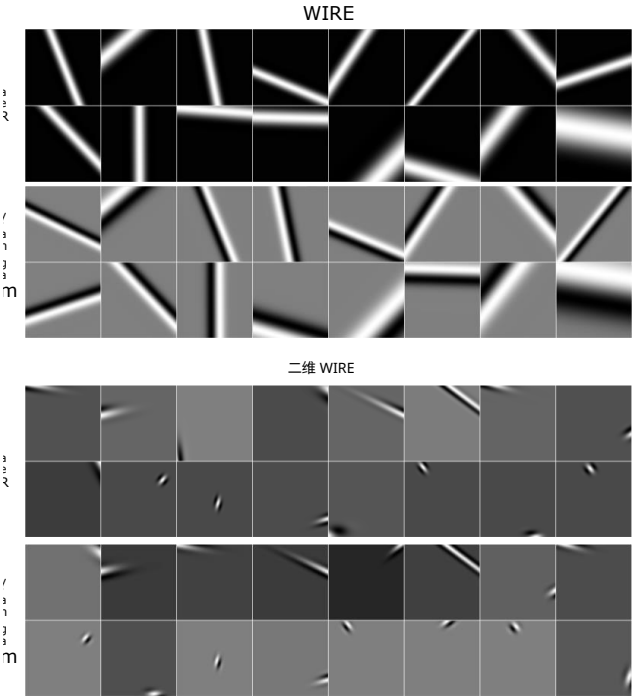


图 11. **多维 WIRE 的第一层输出**。上图展示了 WIRE 和二维 WIRE 经过第一个隐藏层后的输出。我们观察到，由于第二个高斯窗，二维 WIRE 的输出在空间上更为紧凑；而 WIRE 则呈现出与高斯窗正交的埃隆门控结构。

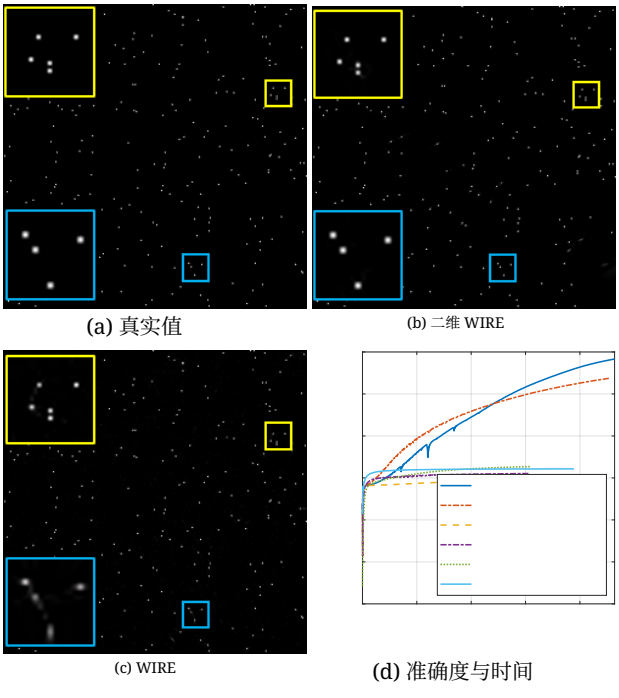


图 12。**多维定位**。二维 WIRE 的空间紧凑特性使其能够表示稀疏图像。在此示例中，我们展示了包含 256 个非零值的 256×256 图像的表示。二维 WIRE 能够清晰地表示每个点，而 WIRE 往往会使特征沿两个轴中的一个发生模糊。与其他方法相比，WIRE 和二维 WIRE 都能快速收敛，如右下角的图表所示。

的 200×200。为展示 WIRE 的优势，我们仅使用 25 张图像训练辐射场，而非默认的 100 张图像。我们使用“torch-NGP”代码库 [45] 训练 NeRF 模型。对于所有实验，我们都选择了一个 4 层多层感知机，每层宽度为 128 个特征。我们选择参数和学习率，以便在验证数据集上最快提升近似精度。图 10 展示了采用各种非线性函数的结果。WIRE 达到了最高精度（+0.2dB），且精度提升最快。WIRE 学到了其他非线性函数输出中缺失的特征，例如连接叠音钹与支架的杆，以及钹上的各向异性反射。

4.4. 多维 WIRE 对比

正如我们在第 3.5 节所讨论的，WIRE 可以实例化为多维非线性函数。二维 WIRE 的一个优点是，其激活在两个空间轴上的分布往往都较为紧凑（见图 11）。这使其能够更准确地拟合由空间上紧凑的结构组成的信号，例如许多自然图像。图 12展示了二维 WIRE 在表示由点奇异性组成的图像时，具有更广泛空间定位能力的优势。与会模糊掉部分点的 一维 WIRE 相比，二维 WIRE 得到的表示更加清晰。

多维 WIRE 的空间定位能力对于更稳健地求解逆问题也极具优势。图13 比较了 WIRE 和二维 WIRE 在去噪、CT 重建及 $4\times$ 超分辨率任务中的表现。在所有任务中，ω_0 和 s_0 母小波均经过选择以取得最佳性能，且 WIRE 和二维 WIRE 的参数总数相同。总体而言，二维 WIRE 的准确度高于 WIRE。在肺部 CT 重建中，二维 WIRE 学到的特征也更加清晰，而 WIRE 学到的特征往往较为模糊。同样，在超分辨率任务中，二维 WIRE 学到了更出色的高频特征，用于表示花朵的中心。

5. 结论

我们提出并验证了 WIRE 的优势：它为隐式神经表示配备了复Gabor小波激活非线性函数。通过大量实验，我们证明 WIRE：(a) 具有更强的表示能力，(b) 能以更快的速度达到更高的精度，(c) 具有强大的归纳偏置，因此非常适合解决具有挑战性的逆问题。其他激活非线性函数在很大程度上各有优势：SIREN

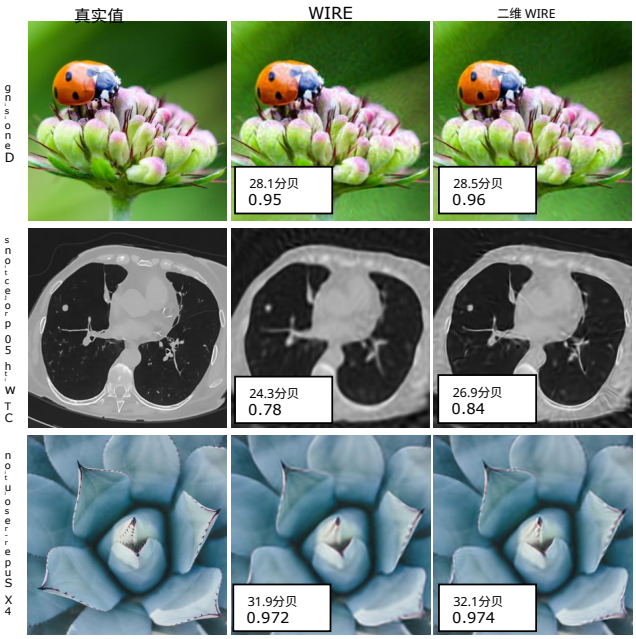


图 13。**多维 WIRE 的性能**。上图展示了采用配备 WIRE 和二维 WIRE 的隐式神经表示解决的各种线性逆问题。总体而言，二维 WIRE 在峰值信噪比和结构相似性指数方面表现更佳。从视觉上看，得益于激活的紧凑性，二维 WIRE 的 CT 重建和超分辨率结果明显更加清晰。

具有很强的表示能力且训练速度快，但在逆问题正则化方面表现不佳。位置编码的表示能力较弱，但适用于新视角合成。高斯非线性函数更适合去噪任务。然而，WIRE 兼具上述所有激活非线性函数的优势，因此目前是用于信号表示和求解逆问题的首选隐式神经表示方案。

6. 致谢

本研究获得了美国国家科学基金会资助，项目编号为 CCF-1911094、IIS-1838177 和 IIS-1730574；美国海军研究办公室资助，项目编号为 N00014-18- 12571、N00014-20-1-2534，以及多学科大学研究计划项目 N00014-20-1-2787；美国空军科学研究办公室资助，项目编号为 FA9550-22-1-0060；以及范内瓦·布什教职奖学金（美国海军研究办公室项目，编号为 N00014-18-1-2047）。

A. 附录 1：实验细节

A.1. WIRE 初始化

像 SIREN 这样的隐式神经表示 [42]在很大程度上依赖初始化来获得准确的表示。WIRE 除默认的均匀权重外，无需任何特殊初始化。不过，由于 WIRE 包含复正弦项，采用类似 SIREN 的初始化能带来些许好处。

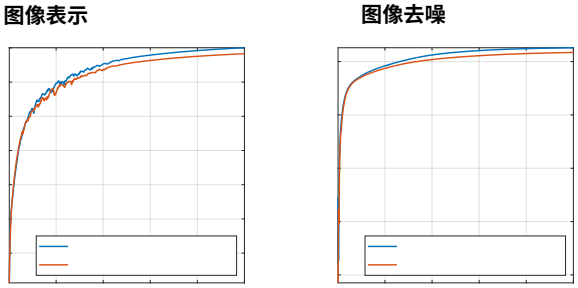


图 14. **初始化的影响**。这些曲线展示了在 SIREN 式权重初始化和标准初始化下，图像表示与图像去噪（输入峰值信噪比为 20 dB）的近似精度随训练轮次的变化。WIRE 对初始权重具有鲁棒性，但采用 SIREN 式初始化时性能略有提升。

为了了解这种依赖关系，我们评估了图像表示（无噪声）和图像去噪（图像噪声为 20dB）任务中的近似精度。这里，类似 SIREN 的权重初始化是指第一层权重从$\mathcal{U}(-1/N, 1/N)$中抽取，其余各层的权重则从 $\mathcal{U}(-\sqrt{6/(\omega_0 N)}\ \sqrt{6/(\omega_0 N)})$中抽取，其中$N$是输入特征的数量，$\mathcal{U}(a,b)$是定义在 $[a,b]$上的均匀分布。正态权重初始化则是指所有层的权重均从 $\mathcal{U}(-1/\sqrt{N}, 1/\sqrt{N})$中抽取。图 14 比较了采用类似 SIREN 的初始化和标准初始化时的表示精度。我们发现，两种情况下的趋势都几乎相同；类似 SIREN 的初始化最多可将精度提高 1dB。因此，WIRE 对初始参数总体上具有较强的鲁棒性，因而能够轻松地在很大的超参数范围内进行调优 ω_0, s_0。

A.2. WIRE 层可视化

Gabor小波能够实现独特的时空-频率局部化，我们观察到 WIRE 也继承了这一特性。为检验这一假设，我们对由一个包含两个隐藏层、每层有 181 个隐藏特征的多层感知机构成的 WIRE 的输出进行了可视化。随后，我们针对一幅包含所有空间频率和方向的西门子星测试图像学习了一种表示。图15 展示了方差最小的 64 个隐藏特征对输入图像产生的实部和虚部输出。每个第一层特征的输出由不同方向的一维 Gabor小波组成，而第二层特征的输出则由稀疏分布的图像组成。

图 16 展示了西门子星测试图像在不同非线性函数下各层的输出，以及最终的近似图像。第二层的稀疏输出显然是 WIRE 独有的。高斯的输出看起来没那么稀疏，而 SIREN 和采用位置编码的 ReLU 则会产生密集的输出。这直接影响对信号高频部分的近似能力——

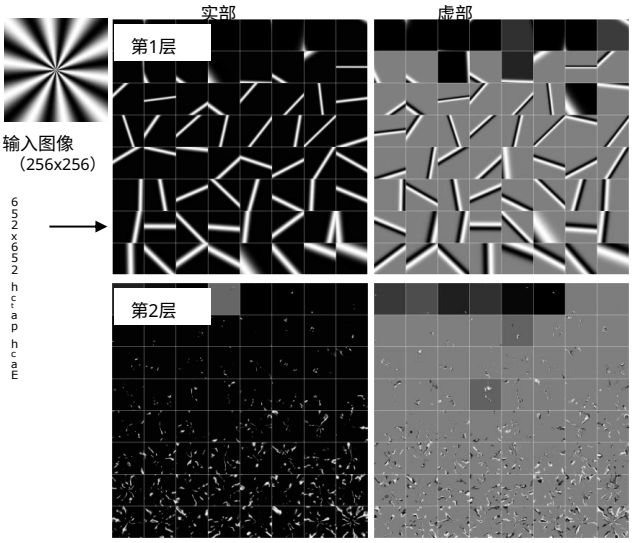


图15. **WIRE的层输出**。 输入图像为西门子星测试图像，其中包含所有空间频率和各个角度的信息。图中图块展示了采用WIRE非线性函数的双隐藏层多层感知机中，第一层和第二层各个隐藏特征的输出（尺寸与图像相同）。如图 16所示，WIRE能生成稀疏图像，从而实现较高的图像表征能力。

底行的最终结果表明，WIRE 输出的稀疏特性使其能够实现较高的近似精度，并在图像中心（空间频率最高的区域）呈现出定性上更好的特征。其次是高斯，因为它在各层产生的输出稀疏度仅次于 WIRE。SIREN 和采用位置编码的 ReLU 都会在中心产生模糊的输出，这主要是由于其输出不具紧致性。WIRE 能够将图像分解为稀疏图像的线性组合，因此在参数数量相同的情况下具有较高的表征能力；我们将在下一节通过实验证实这一点。

A.3. 对训练参数的敏感性

WIRE是一种很有前景的隐式神经表示模型，具有较高的表征精度，并且对各种训练参数都很稳健。本节通过若干敏感性分析展示WIRE的有效性。

学习率的影响。在较宽的学习率范围内，WIRE 都能表现出良好的性能。为了了解性能变化趋势，我们使用不同的非线性函数，对添加了噪声（输入峰值信噪比为 20 dB）的图像进行表示学习。我们使用了一个包含 2 层、每层有 256 个隐藏特征的多层感知机。图17展示了不同学习率下表示的最高峰值信噪比。与其他方法相比，WIRE 的准确率稳定且显著更高。有趣的是，最高准确率是在较高学习率2×10^{-2}下取得的。在深度图像先验 [48]中也观察到了这一现象：较大的学习率带来了更强的正则化。这种相似的表现意味着 WIRE 具有很强的归纳偏置，因此适用于求解逆问题。

层数的影响。 图18a 展示了不同非线性函数下，图像表征精度随隐藏层数变化的曲线。所有情况下，隐藏特征数均设为256。随着层数增加，我们降低了学习率，以避免发散。除隐藏特征数为0的情况外，WIRE始终优于其他方法；这正如预期，因为Gabor小波能够实现较高的图像逼近精度。有趣的是，当隐藏层数较多（≥ 3）时，WIRE的表现与SIREN和高斯非线性函数相近。这是意料之中的，因为网络在层数如此之多时具有很强的表征能力。然而，层数过多会带来较高的计算开销，并且常常导致学习过程不稳定。因此，在大多数情况下，对于少量到中等数量的隐藏层，WIRE都是可靠的选择。

特征数量的影响。图18b 展示了图像表示在不同隐藏特征数量下的近似精度。所有情况下，隐藏层数量均固定为两层。当隐藏特征数量非常少时，WIRE 的表现与其他非线性函数相似，此时所有模型都缺乏足够的表达能力。当特征数量超过 128 时，WIRE 的表现优于其他方法，MFN [19] 紧随其后。

A.4. 逆问题

计算机断层扫描重建。 我们在第 4 节展示了计算机断层成像（CT）如何受益于隐式神经表示的归纳偏置。这里，我们研究测量值数量的影响。图19a 展示了我们在实验中使用的真实图像。我们对原始512×512 图像 [5] 进行 BM3D [14] ($\sigma = 0.1$) 去噪，以去除条纹伪影。随后，我们模拟了投影数量不同的 CT 重建。在每种情况下，我们都使用一个包含三个隐藏层、每层有 256 个隐藏特征的多层感知机。我们在规则网格上对隐式神经表示进行采样，先生成图像，再使用Radon变换获取正弦图。从图19b的准确度曲线可以看出，WIRE 的峰值信噪比高于其他所有非线性函数。图20 展示了每种非线性函数在不同投影数量下的重建结果。即使投影数量较少，重建在视觉效果上仍然更出色，这对于减少拍摄时的 X 射线照射尤为有益。

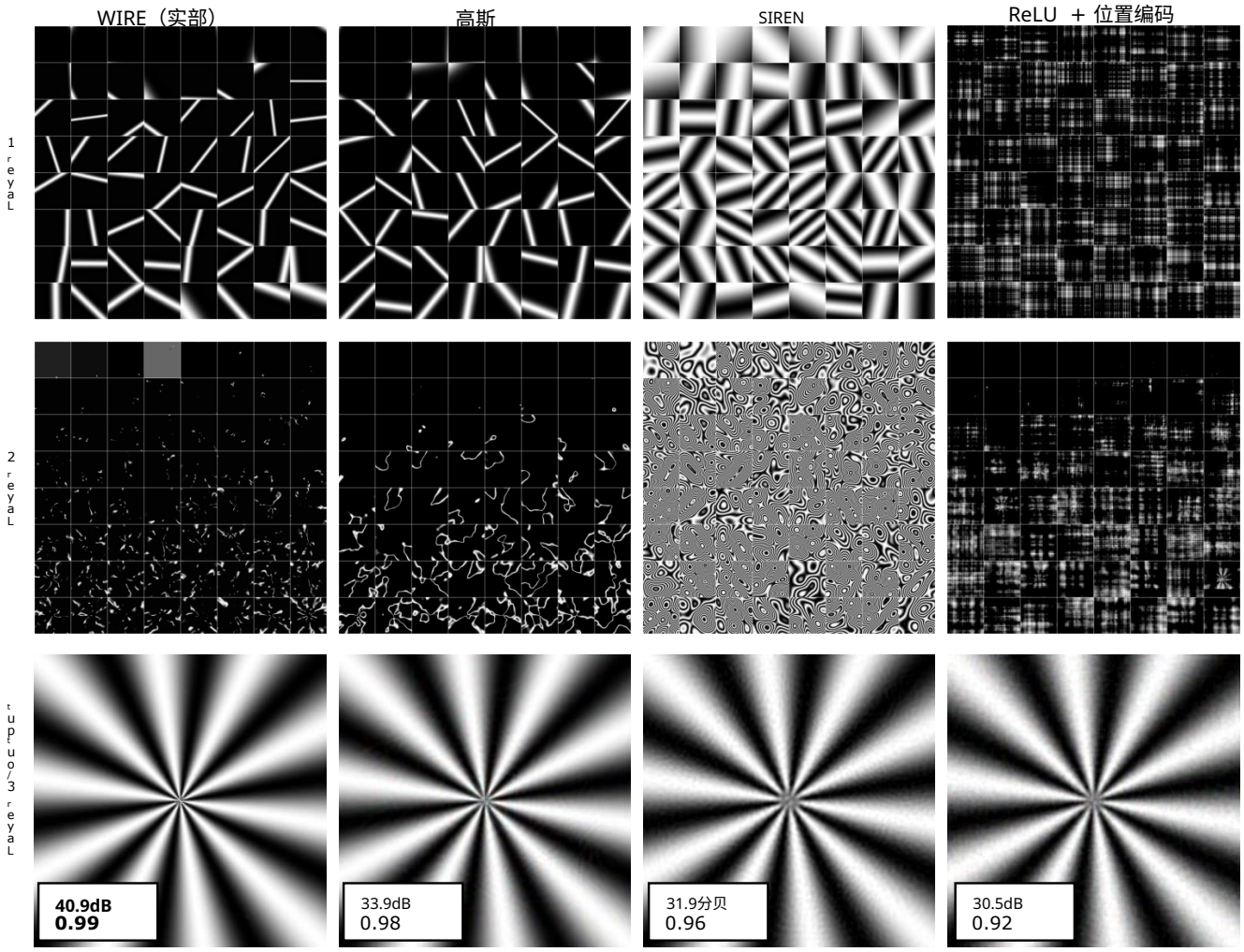


图16。 **隐藏层输出的可视化**。 上图展示了图15中西门子扇区测试图像的两个层中隐藏特征的输出。 15。WIRE 生成的图像具有独特的稀疏性，因此能够高精度地表示图像中的高频部分（扇区中心）。

多图像超分辨率。 我们在第4节展示了多图像超分辨率的一个结果。这里，我们提供有关该实验的更多细节。图 21 展示了来自 Kodak 数据集的 512×768 维真实图像 [1]。我们通过对每幅 $4\times$ 下采样图像施加小幅平移和旋转，模拟了总共四幅低分辨率图像，从而使各帧之间产生亚像素运动。我们假设高分辨率帧 A_k $\mathbf{x}$ **与每幅低分辨率帧** $\mathbf{y}_k$ 之间的变换已知。我们将高分辨率图像 $\mathbf{x}$ 表示为隐式神经表示的输出。在每种情况下，该隐式神经表示都包含三个隐藏层，每层有 256 个隐藏特征。随后，我们求解一个线性逆问题，以估计高分辨率图像。图 21展示了每种非线性函数的重建输出及其指标。插图展示了摩托车辐条的重建效果。从视觉效果来看，WIRE 生成最清晰的特征，且没有任何振铃伪影。此外，WIRE 的重建精度至少提高 1dB，结构相似性指数高出 0.04。

A.5. 神经辐射场

实现细节。在所有实验中，我们使用了`torch-ngp` 软件包 [45]，它实现了多种训练神经辐射场的方法。该架构由两个网络组成，分别用于预测每个体素处的透射率（sigma）和颜色。这两个网络均由一个包含四个隐藏层、每层有 182 个隐藏特征的 MLP 构成。颜色 MLP 将位置 (x,y,z) 和方向（θ，ϕ）作为输入，而透射率 MLP 仅将位置作为输入。与其他所有实验一样，为了考虑复权重导致的参数数量翻倍，我们为 WIRE 使用了 $182/\sqrt{2} = 128$ 隐藏特征

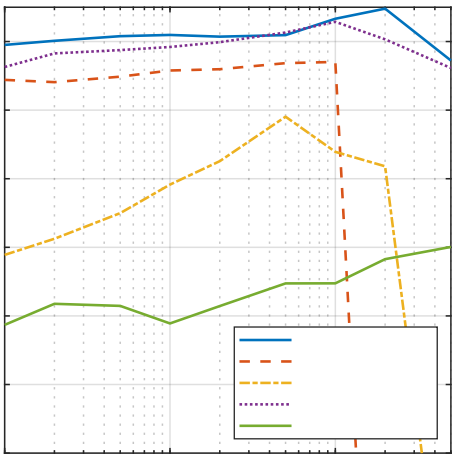
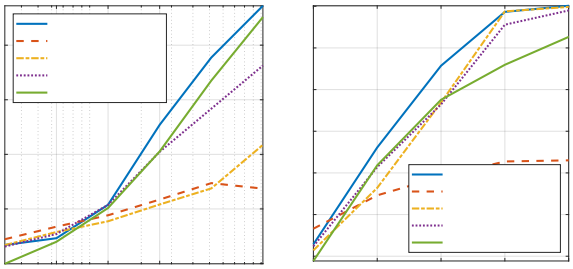


图 17. **学习率的影响**。上图展示了使用各种非线性函数表示含噪图像（输入峰值信噪比为 20 dB）时的近似精度。WIRE 对学习率具有鲁棒性，并在较高学习率 2×10^{-2} 下取得最佳效果。



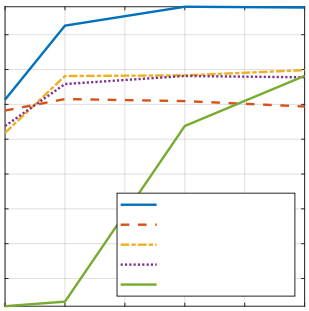
(a) 峰值信噪比与特征数量的关系　(b) 峰值信噪比与层数的关系

图 18. **参数数量的影响**。上图展示了使用不同数量的 (a) 隐藏特征和 (b) 隐藏层表示图像时的近似精度。当隐藏特征数量不少于 128 且层数不少于 1 时，WIRE 的表现优于其他非线性函数；当层数超过 3 时，其表现与 SIREN 和高斯非线性函数几乎相同。

。我们将图像下采样 $4\times$，以确保模型和训练数据能够装入图形处理单元（GPU）的内存中。在图10所示的结果中，我们随机选取了 25 张图像来训练 NeRF，并用 100 张图像对其进行验证。WIRE 的学习率设为 4×10^{-4}，所有其他非线性函数的学习率设为 2.5×10^{-4}，并在总计 2500 个训练轮次内将其降低至 $0.1\times$ 初始值。除 ReLU 外，我们没有对其他非线性函数使用任何形式的位置编码，因为我们希望展示各非线性函数本身的能力。



(a) 真实图像



(b) 峰值信噪比与投影数量的关系

图 19. **投影数量不同时的计算机断层扫描**。 (a) 显示了 512×512 我们在计算机断层扫描实验中使用的肺部真实值 X 射线图像 [5]。我们对原始图像进行了去噪处理，以去除条纹伪影。(b) 显示了不同非线性函数下准确率随测量次数的变化情况。总体而言，WIRE 的表现远胜于所有其他方法。

图像数量的影响。 图 22a 显示了鼓数据集使用全部 100 张图像进行训练时，准确率随训练轮次的变化情况。WIRE 在 2500 个训练轮次内达到最高准确率，且比其他方法收敛更快。图22b 显示了准确率随训练图像数量的变化情况。使用 25、50 和 75 张图像时，WIRE 比次优竞争方法 SIREN 高出 0.1 dB；使用 100 张图像训练时，则高出 0.4 dB。

图23 展示了鼓的一个重建视图。我们将图像数量从 25 张增加到 100 张，然后从一个新视角渲染图像。从视觉效果来看，WIRE 生成的结果最令人满意，包括钹及其支架的清晰特征，以及鼓上平滑的鼓面。相比之下，高斯非线性函数会产生云状伪影，而 SIREN 则会产生高频伪影，图像数量较少时尤为明显。ReLU+位置编码需要使用全部 100 张图像，并且训练轮次远超 2500 轮，才能重建这些组件。总的来说，即使训练样本数量较少，WIRE 也是训练辐射场的一种稳健方案。

参考文献

[1] 柯达无损真彩色图像集。 http://r0k.us/graphics/kodak/，1999年。访问日期：2022-11-09。 4，6，12[2] Tiny ImageNet视觉识别挑战赛。 https://www.kaggle.com/c/tiny-imagenet，2015年。访问日期：2022-11-03。 4[3] Eirikur Agustsson和Radu Timofte。NTIRE 2017单图像超分辨率挑战赛：数据集与研究。载于 *IEEE*计算机视觉与模式识别（*CVPR*），2017年7月。 5[4] Michal Aharon、Michael Elad和Alfred Bruckstein。K-SVD：一种用于设计超完备字典的算法

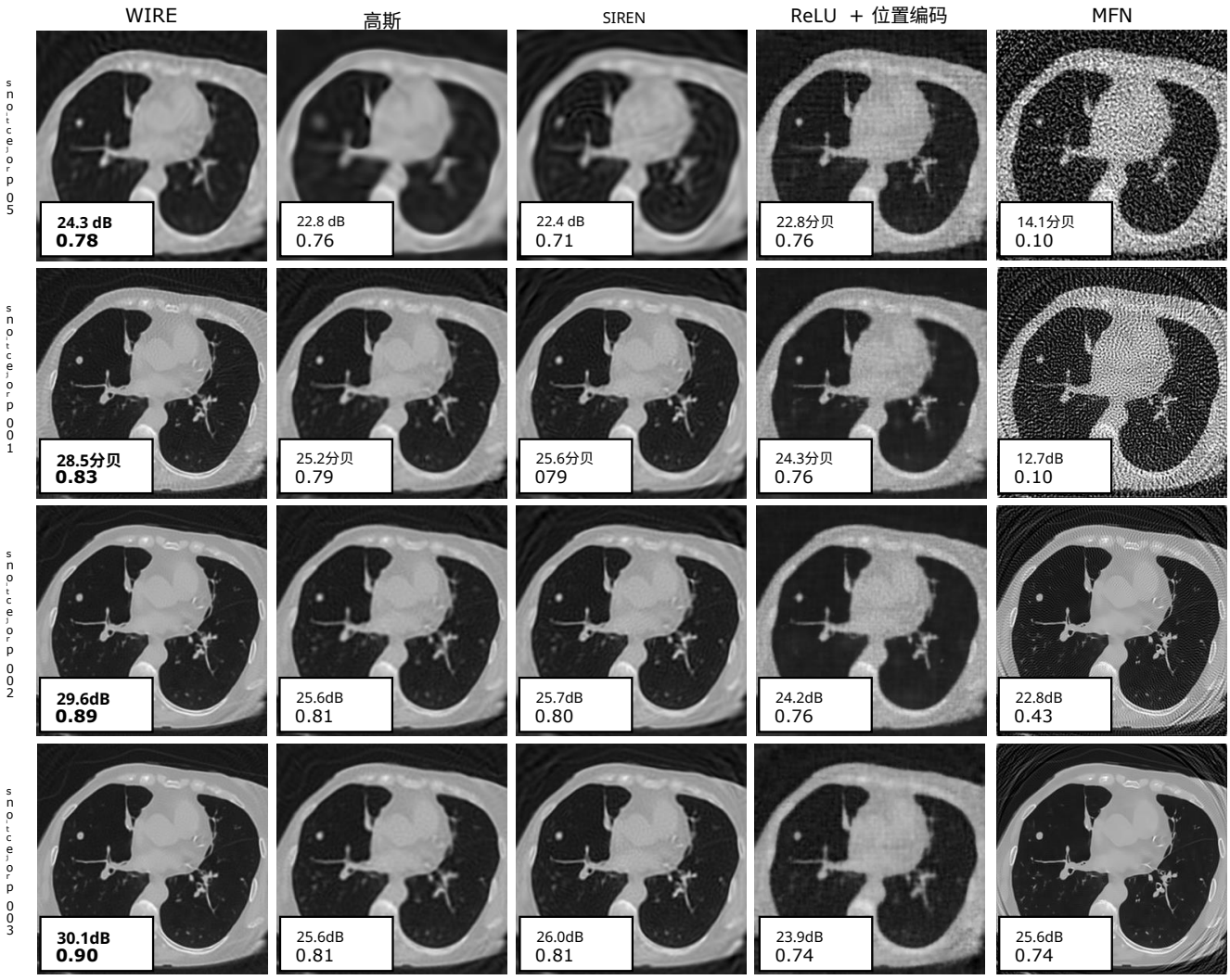


图20。 **投影数量对CT准确度的影响**。 上方图像展示了对图19a所示肺部图像进行重建的结果。 19a。即使只有50个投影（10% 测量值），WIRE的表现也优于所有其他方法，且视觉效果良好。

或稀疏表示 4311 4322, 2006 2 ( ): *rans. Signal Process.* [5] III、Geoffrey McLennan、Luc Bidaut、Michael F. McNitt-Gray、Charles R. Meyer、Anthony P. Reeves、Binsheng Zhao、Denise R. Aberle、Claudia I. Hen- schke、Eric A. Hoffman、Ella A. Kazerooni、Heber MacMa- hon、Edwin J. R. van Beek、David Yankelevitz、Alberto M. Biancardi、Peyton H. Bland、Matthew S. Brown、Roger M. Engelmann、Gary E. Laderach、Daniel Max、Richard C. Pais、David P.-Y. Qing、Rachael Y. Roberts、Amanda R. Smith、Adam Starkey、Poonam Batra、Philip Caligiuri、Ali Farooqi、Gregory W. Gladish、C. Matilda Jude、Reginald F. Munden、Iva Petkovska、Leslie E. Quint、Lawrence H. Schwartz、Baskaran Sundaram、Lori E. Dodd、Charles Fenimore、David Gur、Nicholas Petrick、John Freymann、Justin Kirby、Brian Hughes、Alessi Vande Casteele、Sangeeta Gupte、Maha Sal- lam、Michael D. Heath、Michael H. Kuhn、Ekta Dharaiya、Richard Burns、David S. Fryd、Marcos Salganicoff、Vikram Anand、Uri Shreter、Stephen Vastagh、Barbara Y. Croft 和 Laurence P. Clarke。肺部图像数据库联盟（lidc）与图像数据库资源计划（idri）：一套已完成的肺结节 CT 扫描参考数据库。医学物理学，38(2):915–931，2011年。11，13[6] Benjamin Attal、Eliot Laidlaw、Aaron Gokaslan、Changil Kim、Christian Richardt、James Tompkin 和 Matthew O'Toole。T¨orf：用于动态场景视图合成的飞行时间辐射场。神经信息处理系统进展，34:26289–26301，2021年。2[7] Richard G Baraniuk、Volkan Cevher、Marco F Duarte 和 Chinmay Hegde。基于模型的压缩感知。*IEEE*信息论汇刊，56 4 1982 2001, 2010 2 ( ): –. Judith C Brown。常数 q 谱变换的计算。《美国声学学会杂志》，89 434，1991 1 4254():–. [8] [9] Emmanuel JCand`es 和 David L Donoho。具有分段 c² 奇点的曲波新紧框架及物体的最优表示

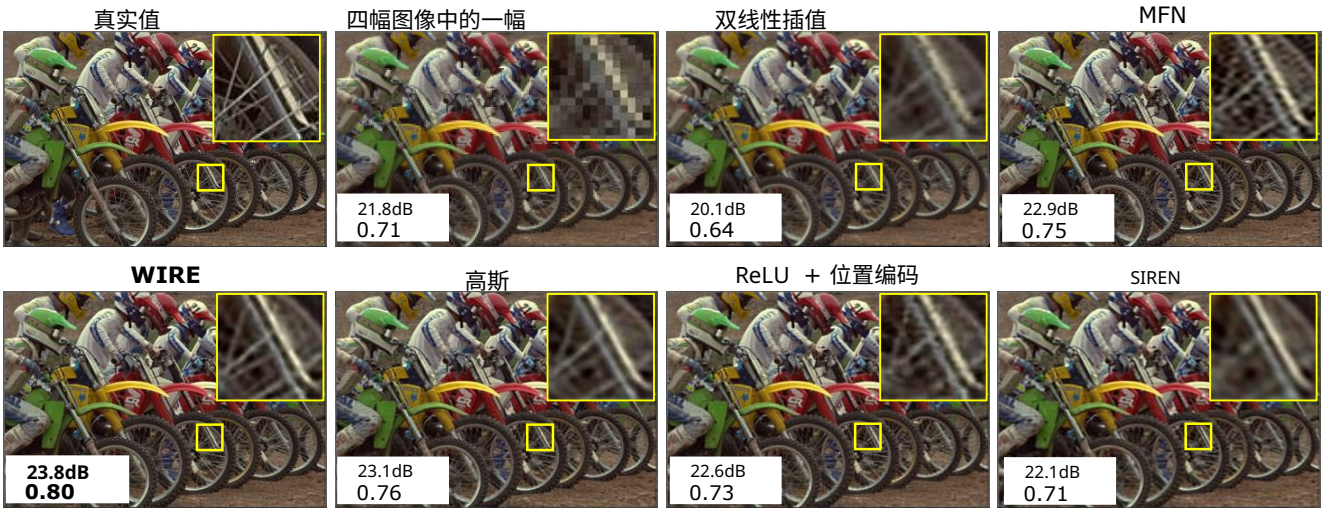
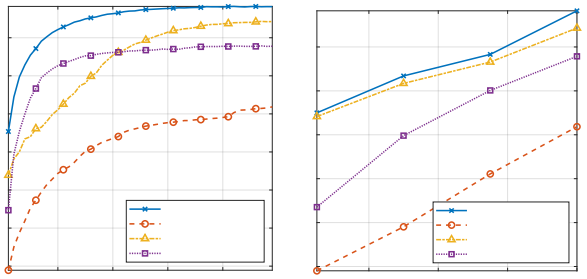


图21. **多图像超分辨率**。上图展示了多帧超分辨率，其中每个子帧都通过已知的微小亚像素偏移模拟而成。WIRE 实现了最高的重建准确度，重建效果在视觉上也更好。



（a）准确度与训练轮次的关系（100张图像）。

（b）准确度与图像数量的关系。

图22. **鼓数据集上的 NeRF 准确度**。（a）展示了使用100张图像训练神经辐射场时，不同非线性函数在每个训练轮次的训练准确度。WIRE 的准确度比次高的 SIREN 高出0。4dB，而且在训练100张图像时收敛迅速。（b）展示了准确度随图像数量的变化情况，WIRE 在所有图像数量下均优于其他方法。

。《纯数学与应用数学通讯》，57 2 219 266, 2004 5 ():–. Antonin Cha

[10] mbolle。全变分最小化算法及其应用。《数学成像与视觉》，20 1 89 97, 2004 2 ( ): – . Yinbo Chen、Sifei Liu 和

[11] Xiaolong Wang。利用局部隐式图像函数学习连续图像表示。载于*IEEE*计算机视觉与模式识别（*CVPR*），2021 1 。

[12] Chiun-Hong Chien 和 Jake K Aggarwal。用于表示三维物体的体/表面八叉树。《计算机视觉、图形学与图像处理》，36(1):100–113，1986年。2[13] Kenneth Clark、Bruce Vendt、Kirk Smith、John Freymann、Justin Kirby、Paul Koppel、Stephen Moore、Stanley Phillips，David Maffitt、Michael Pringle 等。癌症影像档案库（TCIA）：维护和运营公共信息存储库。《数字成像杂志》，26(6):1045–1057，2013年。7[14] Kostadin Dabov、Alessandro Foi、Vladimir Katkovnik 和 Karen Egiazarian。利用亮度-色度空间中带分组约束的稀疏三维协同滤波进行彩色图像去噪。发表于电气电子工程师学会国际图像处理会议（ICIP），2007 11 。Mohammad Zalbagi Darestani 和

[15] Reinhard Heckel。利用未经训练的神经网络加速 MRI。IEEE信号处理汇刊，7:724–733，2021年。2[16]Nianchen Deng、Zhenyi He、Jiannan Ye、Budmonde Duinkharjav、Praneeth Chakravarthula、Xubo Yang 和 Qi Sun。FOV-NeRF：面向虚拟现实的凹视神经辐射场。IEEE可视化与计算机图形学汇刊，28 11 3854 3864, 2022 2 ( ): – 。Ronald A. DeVore。非线性

[17] 逼近。《数值数学学报》，7 51 150, 1998 3 : – 。

[18] Ronald A DeVore、Bjorn Jawerth 和 Bradley J Lucier。通过小波变换编码进行图像压缩。IEEE信息论汇刊，38(2):719–746，1992年。3[19] Rizal Fathony、Anit Kumar Sahu、Devin Willmott 和 J Zico Kolter。乘性滤波网络。发表于国际学习表征会议，2020 5, 11 。Lihan He 和 Lawrence

[20] Carin。利用基于小波的贝叶斯压缩感知中的结构。IEEE信号处理汇刊，57 9 3488 3497, 2009 3 ( ): – 。

[21] Reinhard Heckel 和 Paul Hand。深度解码器：通过未经训练的非卷积网络实现紧凑图像表示。发表于国际学习表征会议，2018 2 。

[22] Arthur Jacot、Franck Gabriel 和 Clement Hongler。神经切线核：神经网络中的收敛与泛化。载于 S. Bengio、H. Wallach、H. Larochelle、K. Grauman、N. Cesa-Bianchi 和 R. Garnett 编，《神经信息处理系统进展》，2018 3 。

[23] Diederik P. Kingma 和 Jimmy Ba。Adam：一种用于

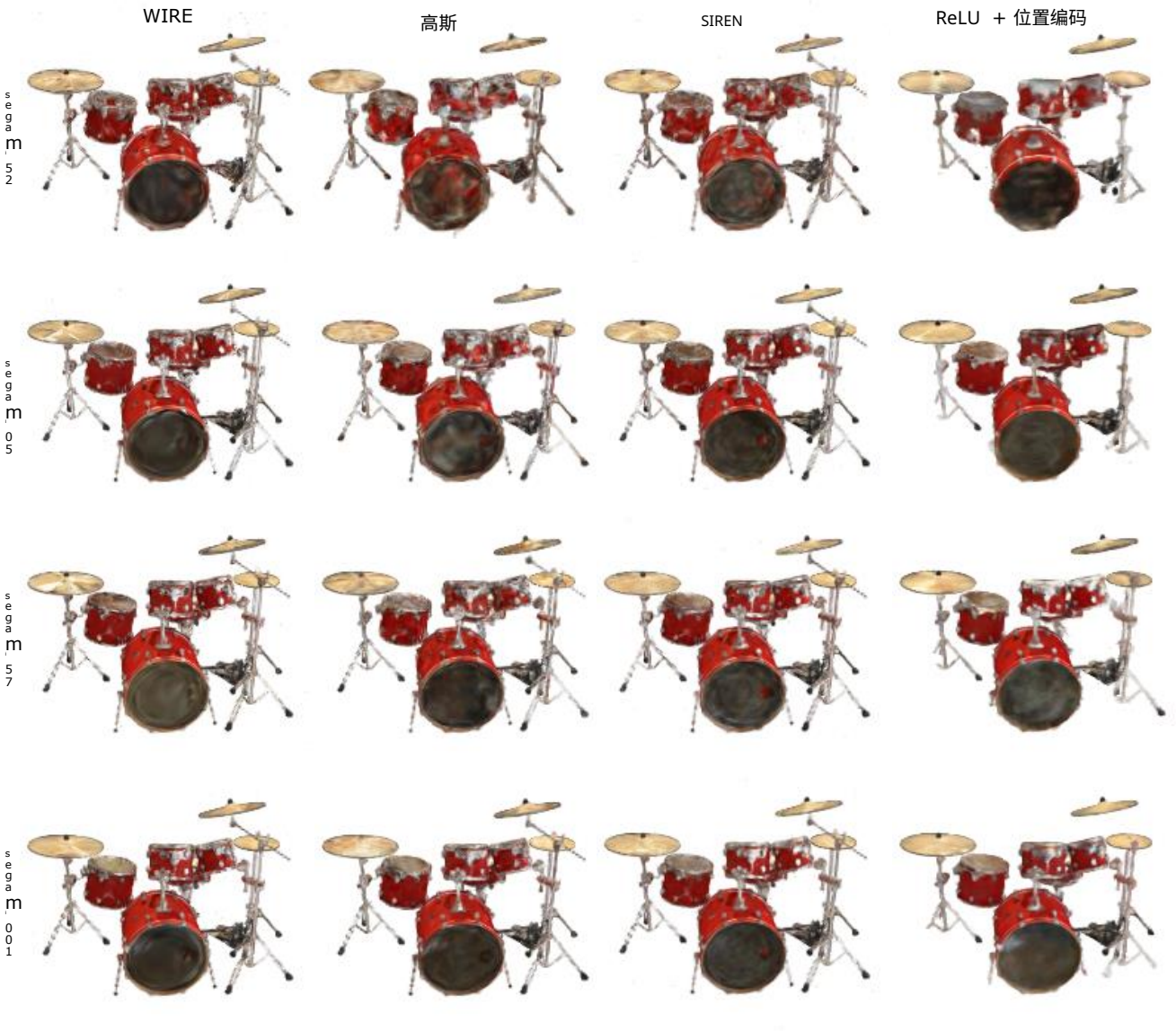


图 23. **图像数量对 NeRF 精度的影响**。上图展示了使用神经辐射场渲染的图像，涵盖不同的非线性函数（按列排列）和不同数量的图像（按行排列）。对于所有图像数量，WIRE 的重建效果在视觉上都优于其他所有方法，这得益于其强大的归纳偏置，能够学习场景辐射场的高频特征。

随机优化。发表于国际学习表征会议，2015年。5[24] Alexandr Kuznetsov、Krishna Mullia、Zexiang Xu、Miloˇs Haˇsan 和 Ravi Ramamoorthi。NeuMIP：多分辨率神经材料。*ACM*计算机图形学汇刊，40(4):1–13，2021年。1，2[25] Jaehoon Lee、Lechao Xiao、Samuel Schoenholz、Yasaman Bahri、Roman Novak、Jascha Sohl-Dickstein 和 Jeffrey Pennington。任意深度的宽神经网络在梯度下降下演化为线性模型。发表于神经信息处理系统进展，2019年。3

[26] St´ephane Mallat. 小波信号处理之旅。Elsevier，1999年。2，4[27] Julien NP Martel、David B Lindell、Connor Z Lin、Eric R Chan、Marco Monteiro 和 Gordon Wetzstein。Acorn：用于神经场景表示的自适应坐标网络。 *arXiv* 预印本 *arXiv:2105.02788*，2021年。3[28] Lars Mescheder、Michael Oechsle、Michael Niemeyer、Sebastian Nowozin 和 Andreas Geiger。占用网络：在函数空间中学习三维重建。载于 *IEEE*计算机视觉与模式识别（*CVPR*），2019年。5[29]Ben Mildenhall、Pratul P Srinivasan、Matthew Tancik,

Jonathan T Barron、Ravi Ramamoorthi 和 Ren Ng。NeRF：将场景表示为用于视图合成的神经辐射场。载于 IEEE 欧洲计算机视觉会议（ECCV），2020年。1，2，7[30] Thomas M¨uller、Alex Evans、Christoph Schied 和 Alexander 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》，41(4):102:1–102:15，2022年。1，2，3[31] Anh Nguyen、Jeff Clune、Yoshua Bengio、Alexey Dosovitskiy 和 Jason Yosinski。即插即用生成网络：在潜空间中进行图像的条件迭代生成。载于 IEEE计算机视觉与模式识别（CVPR），2017年。2[32] Bruno A Olshausen 和 David J Field。通过为自然图像学习稀疏编码，涌现简单细胞的感受野特性。《自然》，381(6583):607–609，1996年。2[33] Adam Paszke、Sam Gross、Francisco Massa、Adam Lerer、James Bradbury、Gregory Chanan、Trevor Killeen、Zeming Lin、Natalia Gimelshein、Luca Antiga、Alban Desmaison、Andreas Kopf、Edward Yang、Zachary DeVito、Martin Raison、Alykhan Tejani、Sasank Chilamkurthy、Benoit Steiner、Lu Fang、Junjie Bai 和 Soumith Chintala。PyTorch：一种命令式风格、高性能的深度学习库。载于神经信息处理系统进展。2019 5。Sameera [34] 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 IEEE 欧洲计算机视觉会议（ECCV），2021年。2，3，4[35] Christian Reiser、Songyou Peng、Yiyi Liao 和 Andreas Geiger。KiloNeRF：利用数千个微型多层感知机加速神经辐射场。arXiv 预印本 arXiv:2103 13744，2021年 3 。 。[36] JH Rick Chang、Chun-Liang Li、Barnabas Poczos、BVK Vijaya Kumar 和 Aswin C Sankaranarayanan。一个网络解决所有问题——使用深度投影模型求解线性逆问题。载于 IEEE国际计算机视觉会议（ICCV），2017 2。Yaniv Romano、[37] Michael Elad 和 Peyman Milanfar。小引擎也能成大事：去噪正则化（RED）。《SIAM成像科学期刊》，10 4 1804 1844, 2017 2 ( ): – 。[38] Olaf Ronneberger、Philipp Fischer 和 Thomas Brox。U-Net：用于生物医学图像分割的卷积网络。载于国际医学图像计算与计算机辅助干预会议，2015年。2[39] Vishwanath Saragadam、Randall Balestriero、Ashok Veeraraghavan 和 Richard G Baraniuk。DeepTensor：结合深度网络先验的低秩张量分解。arXiv 预印本 arXiv:2204.03145，2022年。2[40] Vishwanath Saragadam、Jasper Tan、Guha Balakrishnan、Richard Baraniuk 和 Ashok Veeraraghavan。MINER：多尺度隐式神经表示。载于欧洲会议。计算机视觉，2022 3。Jerome M [41] Shapiro。利用小波系数零树进行嵌入式图像编码。IEEE信号处理汇刊，41(12):3445–3462，1993年。3[42] Vincent Sitzmann、Julien Martel、Alexander Bergman、David Lindell 和 Gordon Wetzstein。采用周期激活函数的隐式神经表示。神经信息处理系统进展，2020 1, 2, 3, 4, 10。Yu [43] Sun、Jiaming Liu、Mingyang Xie、Brendt Wohlberg 和 Ulugbek S Kamilov。Coil：用于成像逆问题的基于坐标的内部学习。arXiv 预印本 arXiv: *2102 05181*, 2021 1。 .[44] Matthew Tancik、Pratul P. Srinivasan、Ben Mildenhall、Sara Fridovich-Keil、Nithin Raghavan、Utkarsh Singhal、Ravi Ra-mamoorthi、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。神经信息处理系统进展，2020年。2[45] Jiaxiang Tang。Torch-ngp：instant-ngp 的 PyTorch 实现，2022年。https://github.com/ashawkey/torch-ngp.[46] 9, 12 Robert Tibshirani。通过套索回归进行回归收缩与选择。英国皇家统计学会B辑（方法论），58 1 267 288, 1996 2 ( ): – 。[47] Massimiliano Todisco、Hector Delgado 和 Nicholas Evans。恒定Q倒谱系数：自动说话人验证中的欺骗攻击对策。《计算机语音与语言》，45 516 535, 2017 4 : – 。Dmitry Ulyanov、[48] Andrea Vedaldi 和 Victor Lempitsky。深度图像先验。载于电气电子工程师学会计算机视觉与模式识别会议（CVPR），2018 2, 6, 11。Michael B. Wakin、Jason N. [49] Laska、Marco F. Duarte、Dror Baron、Shriram Sarvotham、Dharmpal Takhar、Kevin F. Kelly 和 Richard G. Baraniuk。用于视频表示与编码的压缩成像。载于PCS论文集，2006年。3[50] Michael B Wakin、Justin K Romberg、Hyeokho Choi 和 Richard G Baraniuk。分片平滑图像的小波域近似与压缩。IEEE图像处理汇刊，15 5 1071 1087, 2006 5 ( ): – 。[51] Yuehao Wang、Yonghao Long、Siu Hin Fan 和 Qi Dou。用于机器人手术中可变形组织立体三维重建的神经渲染。载于国际医学图像计算与计算机辅助干预会议，2022年。2[52] Zhou Wang、Alan C Bovik、Hamid R Sheikh 和 Eero P Simoncelli。图像质量评价：从误差可见性到结构相似性。IEEE图像处理汇刊，13(4):600–612，2004年。5[53] Gizem Y¨uce、Guillermo Ortiz-Jim´enez、Beril Besbinar 和 Pascal Frossard。隐式神经表示的结构化字典视角。载于电气电子工程师学会计算机视觉与模式识别会议（CVPR），2022年。2，3