CoreDiff：用于低剂量 CT 去噪与泛化的上下文误差调制广义扩散模型

高琪，*IEEE* 学生会员，李子龙，*IEEE* 学生会员，张俊平，*IEEE* 高级会员，张毅，*IEEE* 高级会员和单宏明，*IEEE* 高级会员

arXiv:2304.01814v2 [eess.IV] 6 Oct 2023

摘要—**低剂量计算机断层扫描（CT）图像会因光子饥饿和电子噪声而出现噪声和伪影。近年来，一些研究尝试利用扩散模型，解决以往深度学习去噪模型面临的过度平滑和训练不稳定问题。然而，由于涉及大量采样步数，扩散模型的推理时间较长。最近，冷扩散模型对经典扩散模型进行了推广，并具有更高的灵活性。受冷扩散启发，本文提出一种用于低剂量 CT（LDCT）去噪的新型上下文误差调制广义扩散模型，称为 CoreDiff。首先，CoreDiff 使用低剂量CT图像替代随机高斯噪声，并采用一种新颖的均值保持退化算子来模拟 CT 退化的物理过程。由于以信息丰富的低剂量CT图像作为采样过程的起点，该方法显著减少了采样步数。其次，为缓解采样过程中不完美的恢复算子所造成的误差累积问题，我们提出一种新颖的上下文误差调制恢复网络（CLEAR-Net）。该网络能够利用上下文信息约束采样过程，避免结构失真，并调制时间步嵌入特征，使其与下一时间步的输入更好地对齐。第三，为了尽可能少地使用资源，快速将训练好的模型泛化到新的、未见过的剂量水平，我们设计了一种单次学习框架，仅使用一幅与常规剂量 CT（NDCT）配对或非配对的低剂量CT图像，使 CoreDiff 更快、更好地实现泛化。在四个数据集上进行的大量实验结果表明，与其他方法相比，CoreDiff 的去噪和泛化性能更优，且推理时间在临床上可接受。源代码可在此获取：**https://github.com/qgao21/CoreDiff。

索引词—**低剂量 CT、去噪、扩散模型、单次学习。**

Q. Gao 和 H. Shan 任职于复旦大学脑启发智能科学与技术研究所和教育部脑科学前沿科学中心（中国上海 200433），同时任职于上海脑科学与脑启发技术中心（中国上海 201602）（电子邮箱：qgao21@m.fudan.edu.cn；hmshan@fudan.edu.cn）

Z. Li 和 J. Zhang 任职于复旦大学计算机科学学院上海智能信息处理重点实验室（中国上海 200433）（电子邮箱：zilongli21@m.fudan.edu.cn；jpzhang@fudan.edu.cn）

Y. Zhang 就职于中国四川省成都市四川大学网络空间安全学院，邮编 610065（电子邮箱：yzhang@scu.edu.cn）

I. 引言

COMPUTED 算机断层扫描（CT）是临床诊断中广泛使用的一种成像方式。然而，CT扫描中的X射线电离辐射可能导致脱发和癌症等健康风险 [1], [2]。在临床实践中，可通过降低管电流来减少辐射剂量。不幸的是，由此产生的低剂量 CT（LDCT）图像含有严重噪声和伪影，影响放射科医生的诊断。若能获取原始数据，厂商特定的正弦图预处理或迭代重建算法可以有效去除低剂量CT图像中的噪声。然而，正弦图预处理可能导致边缘模糊和分辨率损失，而迭代重建方法则面临高昂的计算成本 [3]–[5]。此外，由于商业隐私，研究人员通常无法获取原始数据。与这些方法不同，图像后处理算法 [6]–[8] 可直接处理重建后的图像，并因其无需访问原始数据、即插即用的特性而日益受到关注。例如，Ma 等人利用先前常规剂量扫描中的冗余信息，计算非局部均值（NLM）低剂量CT图像去噪所需的非局部权重 [9]。Li等人利用对模体数据进行重复扫描所获得的解析噪声图，改进了NLM算法，从而能够根据CT图像的局部噪声水平自适应地进行去噪 [10]。Sheng等人提出了一种基于三维块匹配（BM3D）的低剂量兆伏级CT图像去噪算法。该算法利用BM3D去噪后残余纹理信息生成显著性图，以增强软组织的视觉显著性 [11]。

近年来，许多研究致力于开发用于低剂量CT图像后处理的深度学习（DL）技术，并取得了令人瞩目的效果 [12], [13]。最初，一些研究人员通过最小化去噪图像与常规剂量CT（NDCT）图像之间的逐像素损失来优化编码器-解码器网络；其中一个代表性模型是残差编码器-解码器卷积神经网络（RED-CNN） [14]。此外，Xia 等人将 RED-CNN 集成到参数依赖框架（PDF-RED-CNN）中，以适用于多种几何结构和不同剂量水平 [15]。尽管这些方法具有出色的去噪性能，但往往会导致图像过度平滑 [16], [17]。为缓解这一问题，

©2023 IEEE。允许个人使用本材料。任何其他用途均须获得 IEEE 的许可，无论是在当前还是未来的任何媒体中，包括为广告或促销目的重印/再版本材料、创作新的合集、转售或重新分发至服务器或列表，或在其他作品中重复使用本作品的任何受版权保护部分。数字对象标识符：10.1109/TMI.2023.3320812。

一些研究采用生成对抗网络（GAN）来尽可能保留与常规剂量CT图像相近的更多纹理和细节 [18]。例如，Yang等人结合了 Wasserstein生成对抗网络和感知损失（WGAN-VGG），以生成更逼真的去噪图像 [19]。Huang等人提出了一种双域生成对抗网络（DU-GAN），用于学习去噪图像与常规剂量CT图像之间的全局和局部差异 [20]。然而，由于生成对抗网络具有对抗性，通常难以训练，并且需要仔细设计优化方法和网络架构，以确保收敛 [21]。

近年来，扩散模型凭借出色的图像生成性能受到了广泛关注 [22]–[27]，，兼具多种生成模型的优势：具有类似变分自编码器的良好分布覆盖能力，且生成质量优于生成对抗网络 [28]–[30]。然而，由于扩散模型从高斯噪声逐步生成图像，需要进行多次迭代采样，因此推理的计算成本高昂；例如，去噪扩散概率模型（DDPM） [23]需要 1,000 个采样步。这限制了扩散模型在各种实时场景中的应用，尤其是在医学成像领域 [31], [32]。一些研究致力于加速扩散模型，使其更具实用性。例如，Nichol 和 Dhariwal 提升了 DDPM 的对数似然表现，并将采样步数减少至 100 [25]。Xia等采用快速常微分方程求解器加速 DDPM，用于低剂量 CT 图像去噪，仅需 50 个采样步 [32]。尽管推理时间在一定程度上有所缩短，这些改进后的扩散模型仍是在经典扩散模型的理论框架内，着重权衡性能与采样速度。近年来，一种称为冷扩散的广义扩散模型扩展了经典扩散模型：它通过预定义的退化算子逐步使图像退化，例如添加各种类型的噪声、模糊、下采样等 [33], [34]。冷扩散使用可学习的恢复算子来逆转扩散过程，并通过“恢复—再退化”采样过程生成图像。尽管冷扩散允许自定义扩散过程，其性能仍受所学习的恢复算子制约。在实际应用中，学习得到的恢复算子可能并不完美，经过多次采样迭代后，恢复图像与真实图像之间可能产生累积误差，并导致不可忽视的逐像素偏差。

本文受冷扩散启发，提出了一种用于低剂量CT去噪的上下文误差调制广义扩散模型（CoreDiff）。为加速采样，我们开发了一种均值保持退化算子，适用于低剂量CT去噪任务，并将低剂量CT图像分别作为扩散过程（正向）的终点和采样过程（反向）的起点。这样，由于低剂量CT图像（温态）比随机高斯噪声（热态）包含更多信息，采样步数得以大幅减少。为缓解采样过程中恢复算子不完善所导致的累积误差，我们进一步提出了一种新颖的上下文误差调制恢复网络（CLEAR-Net），该网络能够利用相邻切片中的丰富上下文信息，减轻z轴上的结构失真，并通过误差调制模块校正输入图像与时间步嵌入特征之间的错位。最后，得益于所提出的均值保持退化算子，我们设计了一种单次学习框架，可利用一张与常规剂量CT（NDCT）配对或非配对的低剂量CT图像，快速将CoreDiff泛化至新的、未见过的剂量水平。

总之，本研究的贡献如下。第一，我们提出了一种用于低剂量CT去噪的新型广义扩散模型 CoreDiff，其中生成的扩散过程模拟了 CT图像退化的物理过程。据我们所知，这是首项将冷扩散模型拓展至低剂量CT去噪的研究。第二，我们提出了一种新型复原网络 CLEAR-Net，通过利用相邻切片之间的上下文信息约束采样过程，并根据最新预测校准时间步嵌入特征，从而减轻累积误差。第三，我们进一步设计了一种单次学习框架，可快速、轻松地将训练好的 CoreDiff 适配到新的、未见过的剂量水平。只需一张与常规剂量CT（非）配对的低剂量CT图像即可完成适配。第四，在四个测试数据集上的大量实验结果表明，所提出的 CoreDiff 性能优越，且推理时间达到临床可接受水平，每张切片仅需 0.12 秒。

II. 方法

本节首先介绍冷扩散模型的基本原理以及误差累积问题。随后，我们在广义扩散模型框架下介绍用于低剂量CT去噪的 CoreDiff，其中包括一种新的均值保持扩散过程和一种新的上下文误差调制复原网络（CLEAR-Net），最后介绍用于快速泛化的单次学习框架。

A. 预备知识：冷扩散

冷扩散模型是一种广义扩散模型 [33]，，它将高斯噪声的扩散与采样扩展到任意类型的退化，例如添加各种类型的噪声、模糊、下采样等。具体而言，给定一幅来自训练数据分布 Q的图像 $\boldsymbol{x}_0$，使用定制的退化算子 $\mathcal{D}(\cdot)$，逐步将图像 $\boldsymbol{x}_0$（冷态）退化为从随机初始分布 P（热态）采样得到的图像 $\boldsymbol{x}_T$；例如例如高斯分布，其中 T是扩散的总时间步数。然后，将扩散过程中任意时间步 t 的图像 $\boldsymbol{x}_t$ 定义为 $\boldsymbol{x}_t = \mathcal{D}(\boldsymbol{x}_0, \boldsymbol{x}_T, t)$，其中 t 对应退化程度，且算子 $\mathcal{D}(\cdot)$对任意 t都应连续。在低剂量CT去噪的背景下，添加噪声是最相关的退化算子。对于添加噪声的扩散过程，冷扩散模型中的退化算子与经典扩散模型中使用的相同，定义如下：

$$\boldsymbol{x}_t = \mathcal{D}(\boldsymbol{x}_0, \boldsymbol{x}_T, t) = \sqrt{\alpha_t}\boldsymbol{x}_0 + \sqrt{(1-\alpha_t)}\boldsymbol{x}_T, \quad (1)$$

其中，$\boldsymbol{x}_T$ 是具有已知分布的随机噪声，且 $\alpha_t < \alpha_{t-1}$、$\forall\, 1 \le t \le T$。

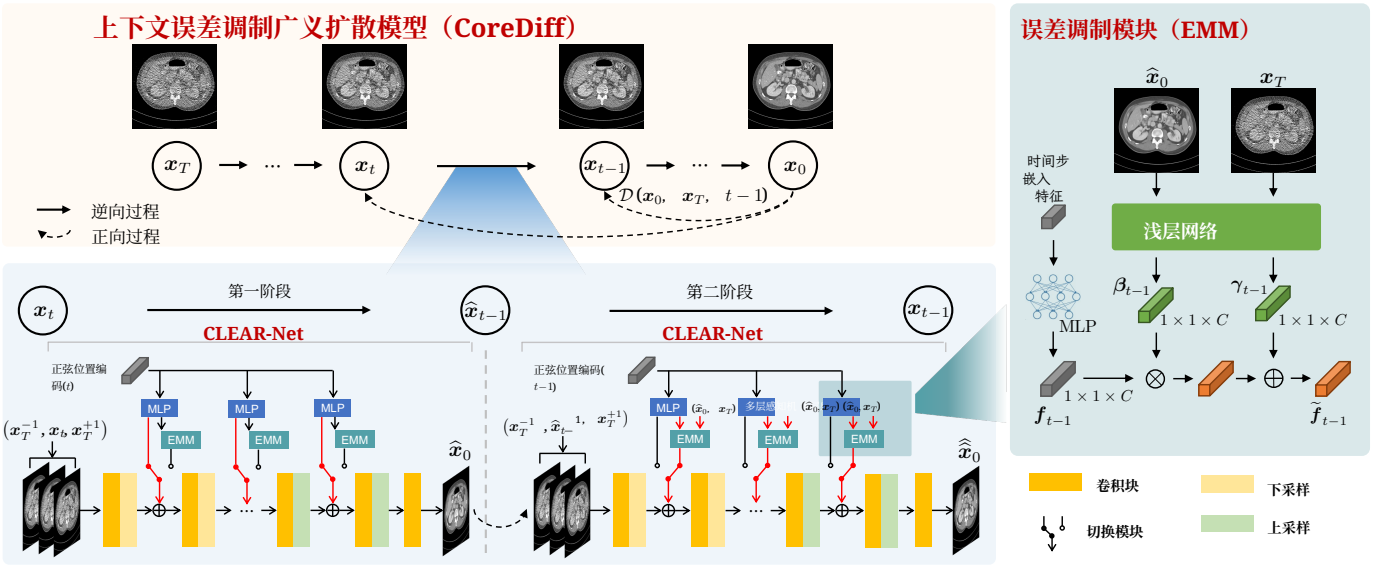

图1。 所提出的用于低剂量 CT 去噪的 CoreDiff 概览。引入的广义扩散模型利用一种新颖的退化算子，在扩散过程中模拟 CT图像退化的物理过程。所提出的 CLEAR-Net 能够减轻累积误差，并针对每个时间步采用两阶段方式进行训练；CLEAR-Net 的一个关键特征是误差调制模块（EMM），它能够利用最新的预测和给定的低剂量CT图像校准时间步嵌入特征。

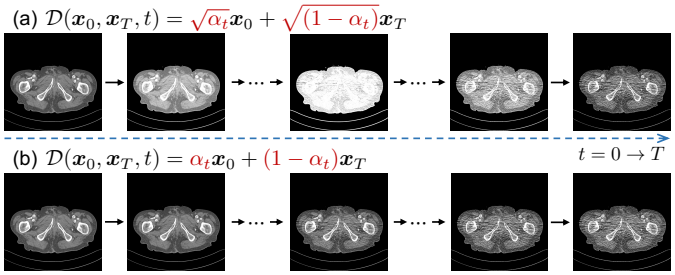

图2。 对比（a）公式 (1) 中的退化算子与（b）公式 (6)中提出的算子。所提出的算子实现了均值保持过程，模拟了 CT 退化的物理过程。

在逆向过程中，我们首先从 P中采样 $\boldsymbol{x}_T$ ，然后使用恢复 算子 $\mathcal{R}(\cdot)$来逆转扩散过程，可表示如下：

$$\widehat{\boldsymbol{x}}_0 = \mathcal{R}(\boldsymbol{x}_t, t) \approx \boldsymbol{x}_0. \tag{2}$$

实际上， $\mathcal{R}(\cdot)$ 是一个由 $\boldsymbol{\theta}$ 参数化的神经网络，可通过以下目标函数进行优化：

$$\min_{\boldsymbol{\theta}} \mathbb{E}_{\boldsymbol{x}_0\sim Q, \boldsymbol{x}_T\sim P} \|\mathcal{R}_{\boldsymbol{\theta}}(\mathcal{D}(\boldsymbol{x}_0, \boldsymbol{x}_T, t), t) - \boldsymbol{x}_0\|. \tag{3}$$

请注意，对于任意 t， $\mathcal{R}_{\boldsymbol{\theta}}(\cdot)$ 都可以直接从 $\boldsymbol{x}_T$ 生成恢复图像 $\widehat{\boldsymbol{x}}_0$ 。然而，我们强调，这种一步预测可能会生成模糊的图像 $\widehat{\boldsymbol{x}}_0$ ，并造成严重的细节损失 [33]。

为解决这一问题，遵循经典扩散模型中的退火采样算法 [23], [25], [28]，冷扩散模型采用“复原-再退化”采样算法，通过总计 T 个采样步逐步生成图像。时间步 $t-1$ 的图像 $\widehat{\boldsymbol{x}}_{t-1}$ 可根据预测结果 $\widehat{\boldsymbol{x}}_0$ 计算如下：

$$\widehat{\boldsymbol{x}}_{t-1} = \mathcal{D}(\widehat{\boldsymbol{x}}_0, \boldsymbol{x}_T, t-1). \tag{4}$$

尽管这种迭代采样算法能够生成比单步预测更清晰的图像，但 $\boldsymbol{x}_0$ 与 $\widehat{\boldsymbol{x}}_0$ 之间的预测误差可能导致 $\widehat{\boldsymbol{x}}_{t-1}$ 与时间步 $t-1$之间发生错位。因此，随着采样过程中误差不断累积， $\mathcal{R}_{\boldsymbol{\theta}}(\cdot)$的预测偏差可能会因错位而进一步加剧。Bansal 等人提出了一种改进的采样算法，以减少这种累积误差 [33]:

$$\boldsymbol{x}_{t-1} = \boldsymbol{x}_t - \mathcal{D}(\widehat{\boldsymbol{x}}_0, \widehat{\boldsymbol{x}}_T, t) + \mathcal{D}(\widehat{\boldsymbol{x}}_0, \widehat{\boldsymbol{x}}_T, t-1), \tag{5}$$

其中 $\widehat{\boldsymbol{x}}_T = \left(\boldsymbol{x}_t - \sqrt{\alpha_t}\widehat{\boldsymbol{x}}_0\right)/\sqrt{(1-\alpha_t)}$。尽管式（5）中改进后的采样算法缓解了误差累积问题，并已证明能够改善图像质量 [33], [34], 但它无法纠正输入与其对应时间步之间的错位，这可能导致像素值出现不可忽略的偏移。

B. 所提出的 CoreDiff 模型

图 1 展示了 CoreDiff 的整体架构，其中包括一个以低剂量CT图像作为扩散过程终点的广义扩散模型、一个用于模拟 CT 退化物理过程的新型均值保持退化算子，以及一个用于解决冷扩散中累积误差和错位问题的新型 CLEAR-Net。

1）用于低剂量 CT 的广义扩散模型：以往基于扩散的低剂量 CT 去噪方法 [31], [32] 通常将扩散过程描述为添加高斯噪声，并使用低剂量CT图像作为条件来预测相应的常规剂量CT图像。然而，需要注意的是，CT图像中噪声的统计特性十分复杂，不能简单地用高斯分布来建模。此外，对于均值为零的噪声，通常认为干净图像代表多组噪声测量值的期望 [35]–[37]。在.....中

算法 1 CoreDiff 的训练

输入： 配对的正常剂量/低剂量CT图像集 $I=\{(\boldsymbol{x}_0,\ \boldsymbol{x}_T)_i\}_{i=1}^N$，总时间步数 T

输出： 已训练的 CLEAR-Net $\mathcal{R}_{\boldsymbol{\theta}}$

初始化： 随机初始化 CLEAR-Net $\mathcal{R}_{\boldsymbol{\theta}}$

重复

  采样 $(\boldsymbol{x}_0,\boldsymbol{x}_T)\sim I$

  采样 $t\sim$均匀$(\{1,\ldots,T\})$

  计算 $\boldsymbol{x}_t^{\mathrm{c}}$ 根据公式 (6) 并 拼接(·) ▷第一阶段

  $\widehat{\boldsymbol{x}}_0\leftarrow\mathcal{R}_{\boldsymbol{\theta}}(\boldsymbol{x}_t^{\mathrm{c}},\ t)$

  通过 $\widehat{\boldsymbol{x}}_{t-1}^{\mathrm{c}}$ 式（4）计算，并拼接(·) ▷第二阶段

  $\widehat{\widehat{\boldsymbol{x}}}_0\leftarrow\mathcal{R}_{\boldsymbol{\theta}}\left(\widehat{\boldsymbol{x}}_{t-1}^{\mathrm{c}},t-1,\mathcal{F}_{\boldsymbol{\phi}}(\widehat{\boldsymbol{x}}_0,\boldsymbol{x}_T)\right)$

  更新 $\boldsymbol{\theta}$，$\boldsymbol{\phi}$ 根据式（9）

直到收敛

算法2 CoreDiff的采样

输入： 测试用低剂量CT图像 $\boldsymbol{x}_T$

输出： 去噪图像 $\boldsymbol{x}_0$

加载训练好的 CLEAR-Net $\mathcal{R}_{\boldsymbol{\theta}}$

$\boldsymbol{x}_T^{\mathrm{c}}\leftarrow$拼接$(\boldsymbol{x}_T^{-1},\boldsymbol{x}_T,\boldsymbol{x}_T^{+1})$

对于 $t=T,T-1,\ldots,1$ **执行**

  $\boldsymbol{x}_0\leftarrow\mathcal{R}_{\boldsymbol{\theta}}(\boldsymbol{x}_t^{\mathrm{c}},t,\mathcal{F}_{\boldsymbol{\phi}}(\boldsymbol{x}_0,\boldsymbol{x}_T))$

  根据公式（$\boldsymbol{x}_{t-1}^{\mathrm{c}}$ 计算，式（5），并进行拼接（·）

结束循环

低剂量CT去噪任务中，我们认为常规剂量CT图像$\boldsymbol{x}_0$ 代表其一组低剂量CT对应图像的期望 $\{\boldsymbol{x}_T^i\}_{i=1}^N$。然而，如图2(a)所示，我们发现，由于 $\sqrt{\alpha_t}$ 与 $\sqrt{(1-\alpha_t)}$ 之和并不总是等于 1，依据公式（1）计算得到的中间图像期望会偏离 $\boldsymbol{x}_0$，并在扩散过程中出现明显的 CT值漂移。因此，公式（1）中广泛采用的退化算子偏离了因剂量降低而产生 CT 退化的实际物理过程。

与以往基于扩散的方法不同，这些方法将低剂量CT去噪任务转化为条件图像生成任务，以随机高斯噪声作为扩散过程的终点，并需要大量步骤才能生成精确估计的图像；我们提出了一种用于低剂量CT去噪的通用扩散模型，将低剂量CT图像作为扩散过程的终点，即 $\boldsymbol{x}_T$。为使扩散过程模拟CT图像退化的物理过程，我们引入了一个新的退化算子 $\mathcal{D}(\cdot)$，定义如下：

$$\boldsymbol{x}_t=\mathcal{D}(\boldsymbol{x}_0,\boldsymbol{x}_T,t)=\alpha_t\boldsymbol{x}_0+(1-\alpha_t)\boldsymbol{x}_T,\tag{6}$$

其中，图像 $\boldsymbol{x}_t$ 在每个时间步都保留了低剂量CT图像 $\boldsymbol{x}_T$ 特有的噪声统计特性。如图2(b)所示，采用这一算子的另一优点是，它能够确保扩散过程中的中间图像 $\boldsymbol{x}_t$ 保持相同的期望值 $\boldsymbol{x}_0$，且不会引入额外的CT值偏移。因此，我们将这一退化算子称为均值保持退化算子。需要指出的是，在实际场景中，由于噪声复杂且存在伪影，该算子可能无法严格保持均值。这一特性不仅使CoreDiff的扩散过程与低剂量CT图像退化过程保持一致，对我们的单次学习框架也至关重要。低剂量CT图像 $\boldsymbol{x}_T$ 可视为冷态（干净图像）与热态（随机噪声）之间的中间状态，我们称之为温态。如第II-A节所述，冷扩散和经典高斯扩散模型的采样过程都从随机高斯噪声开始，逐步降低图像中的噪声，直至生成$\widehat{\boldsymbol{x}}_0$ 。因此，要生成噪声水平与低剂量CT图像相近的图像，它们需要大量采样步数；而低剂量CT图像包含常规剂量CT图像的基本语义信息。因此，所提出的CoreDiff可以从温态开始采样，只需较少的 T，而无需从随机高斯噪声开始。

2）上下文误差调制恢复网络（CLEAR-Net）：为缓解冷扩散中由复原网络不完美导致的累积误差和错位，我们提出了一种新型复原网络，称为 CLEAR-Net。基于“复原-再退化”采样算法，如图 1 所示，我们将训练过程中的每个时间步分为两个阶段：1）在第一阶段，我们首先使用公式（6）获得退化图像 $\boldsymbol{x}_t$，然后使用 CLEAR-Net $\mathcal{R}_{\boldsymbol{\theta}}(\cdot)$ 估计 $\widehat{\boldsymbol{x}}_0$；2）在第二阶段，我们执行公式（4）中的再退化操作，基于最新预测 $\widehat{\boldsymbol{x}}_0$计算$\widehat{\boldsymbol{x}}_{t-1}$，然后使用同一网络 $\mathcal{R}_{\boldsymbol{\theta}}(\cdot)$预测常规剂量CT图像。CLEAR-Net 的创新之处主要有两点。一方面，受 [5], [31],中所用上下文信息的启发，我们引入来自 $\boldsymbol{x}_T$ 相邻切片的上下文信息，以减轻采样过程中的结构失真。更具体地说，我们假设 $\boldsymbol{x}_T$ 的相邻切片为$\boldsymbol{x}_T^{-1}$ 和 $\boldsymbol{x}_T^{+1}$，分别对应其上层和下层切片。我们在每个时间步将 $\boldsymbol{x}_t\in\mathbb{R}^{1\times H\times W}$ 与起始时刻的相邻切片 $\boldsymbol{x}_T^{-1}$ 和 $\boldsymbol{x}_T^{+1}$ 沿通道维度拼接，得到$\boldsymbol{x}_t$的上下文版本，即 $\boldsymbol{x}_t^{\mathrm{c}}=\mathrm{Concat}(\boldsymbol{x}_T^{-1},\boldsymbol{x}_t,\boldsymbol{x}_T^{+1})\in\mathbb{R}^{3\times H\times W}$。由于相邻切片在采样过程中保持不变，因此可约束网络 $\mathcal{R}_{\boldsymbol{\theta}}(\cdot)$生成连续的 z轴结构。

另一方面，CLEAR-Net 利用误差调制模块（EMM）校准网络输入 $\widehat{\boldsymbol{x}}_{t-1}$ 与 $t-1$的时间步嵌入特征之间的错位。具体而言，我们的 EMM 是一个逐特征线性调制模块 [38]–[40]，用于调制时间步嵌入特征，其中时间步 $t-1$ 处的调制因子估计如下：

$$\boldsymbol{\beta}_{t-1},\boldsymbol{\gamma}_{t-1}=\mathcal{F}_{\boldsymbol{\phi}}(\widehat{\boldsymbol{x}}_0,\boldsymbol{x}_T),\quad\widehat{\boldsymbol{x}}_0=\mathcal{R}_{\boldsymbol{\theta}}(\boldsymbol{x}_t,t),\tag{7}$$

其中，$\mathcal{F}_{\boldsymbol{\phi}}(\cdot)$ 是一个由 $\boldsymbol{\phi}$ 参数化的浅层网络，用于根据最新预测$\widehat{\boldsymbol{x}}_0$ 和初始输入低剂量CT图像 $\boldsymbol{x}_T$估计调制因子。随后，$t-1$ 的时间步嵌入特征按如下方式进行调制：

$$\widetilde{\boldsymbol{f}}_{t-1}=\boldsymbol{\beta}_{t-1}\boldsymbol{f}_{t-1}+\boldsymbol{\gamma}_{t-1},\ \boldsymbol{f}_{t-1}=\mathrm{MLP}(\mathrm{SinPE}(t-1)),\tag{8}$$

其中，SinPE(·)表示时间步 $t-1$ 的正弦位置编码，$\boldsymbol{f}_{t-1}$是来自

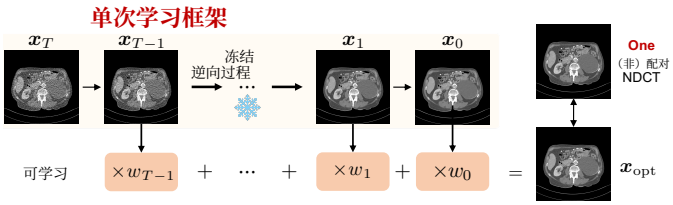

图 3. 用于快速泛化的单次学习框架。

~多层感知机（MLP）的时间步嵌入特征，而 $\boldsymbol{f}_{t-1}$ 是调制后的特征。请注意，所提出的 EMM 仅用于第二阶段，调制特征用于 $\mathcal{R}_{\boldsymbol{\theta}}(\cdot)$ 中每次上采样/下采样操作之后。

通过所提出的 CLEAR-Net，我们的 CoreDiff 的最终训练目标定义如下：

$$\min_{\boldsymbol{\theta},\boldsymbol{\phi}} \mathbb{E}\Big[\| \underbrace{\mathcal{R}_{\boldsymbol{\theta}}(\boldsymbol{x}_t^{\mathrm{c}}, t) - \boldsymbol{x}_0}_{\text{Stage I}} \|^2 + \| \underbrace{\hat{\hat{\boldsymbol{x}}}_0 - \boldsymbol{x}_0}_{\text{Stage II}} \|^2\Big], \tag{9}$$

其中，$\hat{\hat{\boldsymbol{x}}}_0$ 表示第二阶段中时间步 $t-1$ 时 $\mathcal{R}_{\boldsymbol{\theta}}(\cdot)$ 的输出；即，$\hat{\hat{\boldsymbol{x}}}_0 = \mathcal{R}_{\boldsymbol{\theta}}(\hat{\boldsymbol{x}}_{t-1}^{\mathrm{c}},\ t-1,\ \mathcal{F}_{\boldsymbol{\phi}}(\hat{\boldsymbol{x}}_0,\ \boldsymbol{x}_T))$，且 $\hat{\boldsymbol{x}}_{t-1}^{\mathrm{c}} = \text{拼接}(\boldsymbol{x}_T^{-1},\ \hat{\boldsymbol{x}}_{t-1},\ \boldsymbol{x}_T^{+1})$。

最后，在 CoreDiff 的采样过程中，每个时间步仅执行一次退化算子和恢复算子。利用训练好的 CLEAR-Net，我们采用式（5）中的改进采样算法，并根据式（6）中的退化算子替换系数，以推断最终的去噪图像。训练和采样（推理）过程分别如算法 1 和 2所示。

C. 用于快速泛化的单次学习

临床实践中采集的低剂量 CT图像因设备和协议不同而各不相同。如何利用尽可能少的资源，使一个已训练模型快速适应此前未见过的新剂量水平是一个重要的临床问题 [15], [41], [42]。

在此，我们设计了一种单次学习（OSL）框架，专门用于已训练的 CoreDiff，只需 T 个可学习参数，并且仅用一张低剂量 CT图像即可完成训练，如图 3 所示。借助式（6）中引入的均值保持退化算子，CoreDiff 的“复原-再退化”过程能够逐步去除图像中的噪声和伪影，生成一系列均值相同、噪声程度各异的去噪图像。我们的想法是整合这些图像，为新的、此前未见过的剂量水平生成视觉效果最佳的去噪图像，具体实现如下：

$$\boldsymbol{x}_{\text{opt}} = \sum_{t=0}^{T-1} w_t \boldsymbol{x}_t \ \text{s.t.} \ \sum_{t=0}^{T-1} w_t = 1,\ \forall t, w_t \geq 0, \tag{10}$$

其中，$w_t,\ t=0,\ \ldots\ldots,\ T-1$ 是用于在每一步协同融合图像的可学习权重，$\boldsymbol{x}_{\text{opt}}$ 表示新剂量水平下的最优去噪图像。在训练这一单次学习框架时，我们冻结 CLEAR-Net 的参数，只学习 $w_t,\ t=0,\ \ldots\ldots,\ T-1$。因此，我们可以将单幅图像划分为多个图像块来训练该框架。尽管新的低剂量 CT 和常规剂量CT图像并未配对，我们的 OSL框架不会引入结构失真，因为所有 $\boldsymbol{x}_t$ 都对应于同一幅常规剂量CT图像。为确保 $\boldsymbol{x}_{\text{opt}}$ 具有更好的视觉感知且不过度平滑，我们使用感知损失 [19] 来指导这些 T 参数的学习。

III. 实验与结果

A. 数据集

我们在实验中使用了四个数据集，涵盖不同剂量、中心和对象。

1) Mayo 2016数据集：我们使用“2016 NIH-AAPM-Mayo Clinic 低剂量 CT 大挑战赛”数据集进行训练和测试 [43], , 其中包含来自 10 位患者的 5,936 张厚度为 1 mm 的常规剂量 CT 切片。我们随机选取 9 位患者作为训练集，其余 1 位作为测试集。为获得不同剂量水平的图像，我们使用经过验证的低剂量 CT 模拟算法，采用广受认可的“Pois- son+Gaussian”噪声模型生成低剂量投影 [44]:

$$\boldsymbol{p}_{\text{ld}} = \ln \frac{I_0}{\text{Poisson}(I_0 \exp(-\boldsymbol{p}_{\text{hd}})) + \text{Gaussian}(0, \sigma_e^2)}, \tag{11}$$

其中，$\boldsymbol{p}_{\text{ld}}$ 和 $\boldsymbol{p}_{\text{hd}}$ 分别表示低剂量和常规剂量投影。I_0 表示入射光子数，设为 1.5×10^5。σ_e^2 表示电子噪声的方差，根据 [15]固定为 10。随后，使用滤波反投影（FBP）算法重建图像。在本实验中，我们模拟了 50%、25%、10% 和 5%剂量数据，其中 5%对应超低剂量情况 [45]。为进行公平比较，所有深度学习方法均使用 25%或 5%剂量进行训练和测试。我们使用测试数据中的 50%和 25%剂量来验证单次学习框架的泛化性能。

2）Mayo 2020 数据集：为考察不同方法在同一中心数据集上的新剂量水平泛化性能，我们使用了 Mayo Clinic 于 2020 年最新发布的“低剂量 CT 图像与投影数据”进行外部测试，该数据集称为 Mayo 2020数据集 [46]。该数据集包含来自两家厂商的 299 次扫描，其中头部和腹部扫描提供 25%剂量数据，胸部扫描提供 10%剂量数据。我们随机选取了 5 次胸部扫描和 5 次腹部扫描，共包含 800 张图像，用于混合剂量水平测试。

3）仔猪数据集：为进一步考察不同方法在不同中心数据集上的泛化性能，我们还使用了一个真实的仔猪数据集，该数据集由 GE Discovery CT750 HD 扫描仪采集，共包含 850 张 CT图像 [47]。该数据集提供了与每次常规剂量CT扫描对应的 50%、25%、10% 和 5%剂量扫描。本实验选用了两种剂量数据：25% 和 10%。

4) 模体数据集：我们还使用了一个公开的真实模体数据集（Gammex 467 CT 模体）来评估所提方法的临床实用性。该数据集包含使用胸部协议 [48] 扫描的 9 种不同剂量数据（从 33 到 499 毫安秒）。本实验选取了两种剂量数据：271 毫安秒（54.31%）和 108 毫安秒（21.64%）。对于每次扫描，均选取第 10 至第 21 层，以确保所有圆柱形植入物都清晰可见。

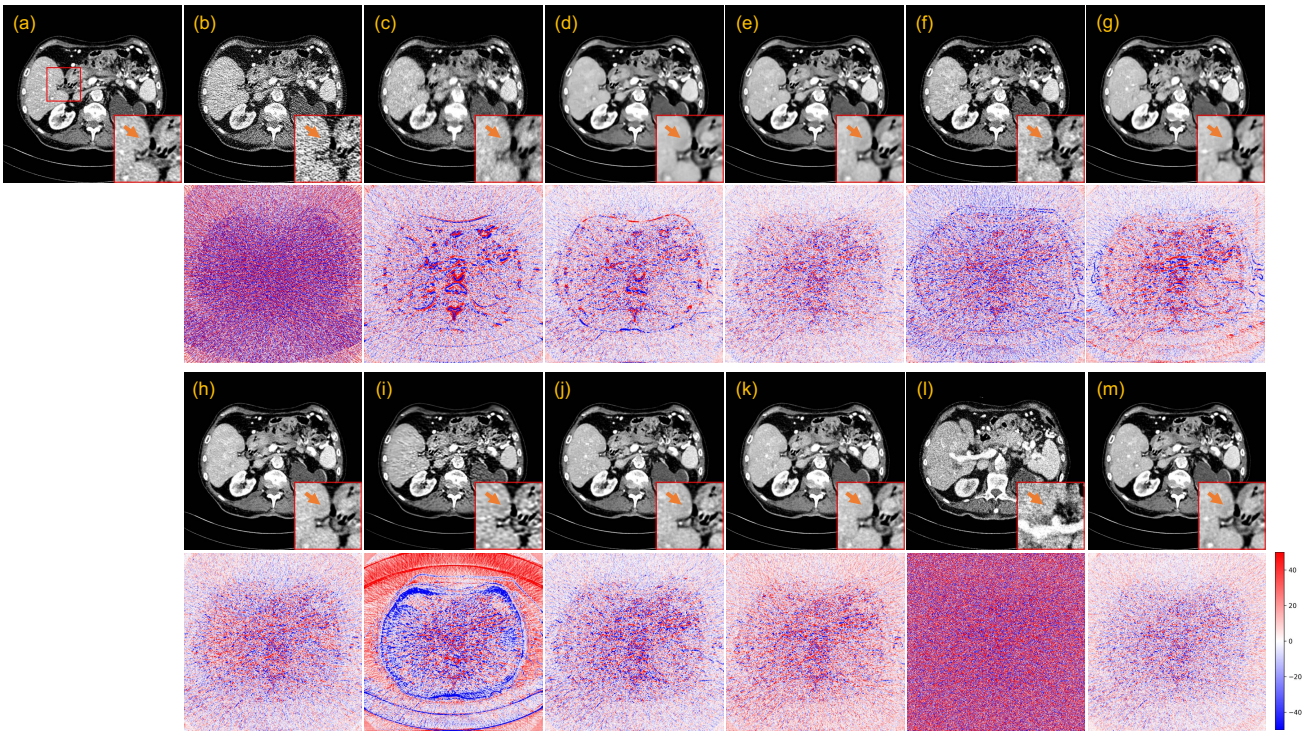

图 4. Mayo 2016数据集中25%剂量腹部 CT 图像的定性结果。（a）常规剂量CT图像（真实值），（b）滤波反投影，（c）惩罚加权最小二乘，（d）RED-CNN，（e）PDF-RED-CNN，（f）WGAN-VGG，（g）CNCL-U-Net，（h）DU-GAN，（i）DDM²，，（j）IDDPM-1000，（k）IDDPM-50，（l）IDDPM-10，以及（m）CoreDiff-10（**我们的方法**）。显示窗为 [-100, 200] HU。放大红色感兴趣区以进行视觉比较，橙色箭头指向一个病灶。

表 I来自 MAYO 2016数据集的25%剂量测试数据上的定量结果（均值±标准差）

	峰值信噪比(dB)	SSIM	均方根误差(亨氏单位)	FSIM	VIF	噪声质量测量(dB)
FBP	34.16±1.76	0.7763±0.0605	40.0±8.0	0.9603±0.0108	0.618±0.062	32.40±2.18
PWLS	38.26±1.29	0.9449±0.0091	24.7±3.5	0.9794±0.0053	0.633±0.052	29.94±2.20
RED-CNN	39.29±1.53	0.9599±0.0106	22.1±4.4	0.9799±0.0044	0.574±0.033	29.64±1.98
PDF-RED-CNN	42.94±1.32	0.9685±0.0103	14.4±2.3	0.9884±0.0038	0.667±0.049	35.96±1.85
WGAN-VGG	40.12±0.98	0.9419±0.0118	19.9±2.3	0.9836±0.0031	0.593±0.038	32.97±1.39
CNCL-U-Net	40.91±1.05	0.9598±0.0118	18.2±2.2	0.9836±0.0040	0.606±0.045	32.23±1.57
DU-GAN	41.50±1.22	0.9591±0.0121	17.0±2.5	0.9875±0.0032	0.660±0.050	34.55±1.89
DDM^2	37.83±1.11	0.8773±0.0402	25.9±3.6	0.9766±0.0037	0.591±0.048	28.93±1.71
IDDPM-1000	41.30±1.20	0.9593±0.0116	17.4±2.5	0.9860±0.0036	0.626±0.048	34.20±1.85
IDDPM-50	41.49±1.18	0.9582±0.0122	17.0±2.4	0.9871±0.0034	0.634±0.047	34.77±1.79
IDDPM-10	33.02±1.29	0.6934±0.0912	45.2±7.0	0.9664±0.0075	0.479±0.040	29.79±1.29
CoreDiff-10（**我们的方法**）	**43.92±1.33**	**0.9744±0.0087**	**12.9±2.1**	**0.9919±0.0026**	**0.724±0.047**	**37.84±1.76**

B. 实现细节

按照 [49], ，我们使用 U-Net 作为所提出的 CLEAR-Net 的骨干网络，它由两个下采样块、一个中间块、两个上采样块和一个输出卷积层组成。CLEAR-Net 的输入大小为 $3 \times 512 \times 512$，包含相邻切片的上下文信息。我们使用 Adam 优化器对 CoreDiff 进行优化，学习率为 2×10^{-4}，训练总迭代次数为 150k。α_1，……，α_T设置为从 0。999 线性变化至 0。我们基于 MIRT 工具箱 [50]进行了数据模拟。我们使用 PyTorch 实现 CoreDiff，并在一块 NVIDIA RTX 3090 图形处理器（24GB）上以 4 为小批量大小进行训练。对于单次学习框架的训练，我们将每幅图像划分为 81 个大小为 256×256 、步幅为 32 的图像块。训练时使用的小批量大小为 8，学习率设置为 2×10^{-3}，总训练迭代次数为 3k。在测试阶段，我们根据公式（ 512×512 ），将 CoreDiff 输出的尺寸为10的图像与学习得到的权重直接相乘，从而获得最优去噪图像。

我们将 CoreDiff 与四类型的低剂量 CT 去噪方法进行了比较，包括 1）迭代重建算法：惩罚加权最小二乘模型（PWLS）[4]；2）基于RED-CNN的方法：RED-CNN [14]和 PDF-RED-CNN [15]；3）基于GAN的方法：WGAN-VGG [19]、DU-GAN [20], 和基于 U-Net 的内容-噪声互补学习（CNCL-U-Net） [51]；以及 4）基于扩散的方法：用于去噪扩散磁共振成像的去噪扩散模型（DDM^2）和改进的 DDPM（IDDPM） [25]。我们将所比较的基于深度学习的方法的超参数设置为遵循-

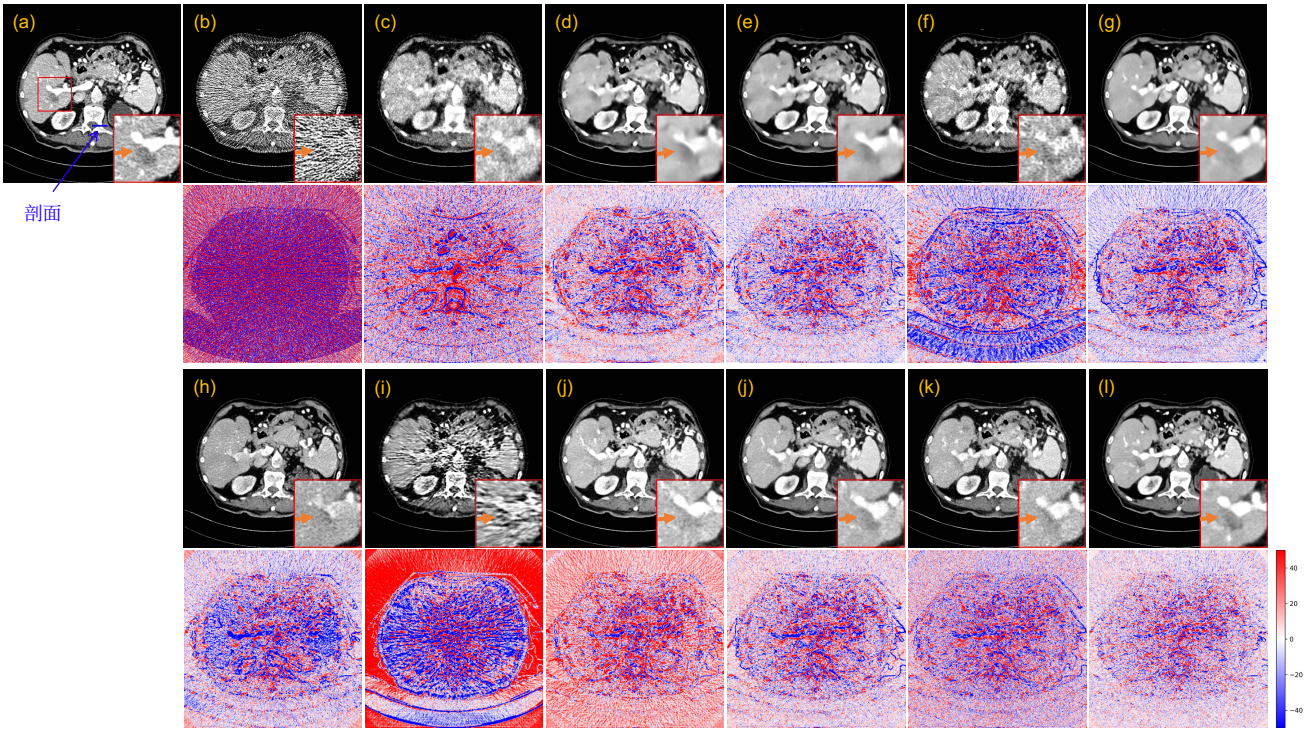

图5. Mayo 2016数据集中5%剂量腹部计算机断层扫描图像的定性结果。(a) 常规剂量CT图像（真实值），(b) 滤波反投影，(c) 惩罚加权最小二乘，(d) RED-CNN，(e) PDF-RED-CNN，(f) WGAN-VGG，(g) CNCL-U-Net，(h) DU-GAN，(i) DDM^2，(j) IDDPM-1000，(k) IDDPM-50，(l) IDDPM-10，以及 (m) CoreDiff-10（**我们的方法**）。显示窗为 [-100, 200] 亨氏单位。红色感兴趣区已放大以供视觉比较，橙色箭头指向一个病灶。

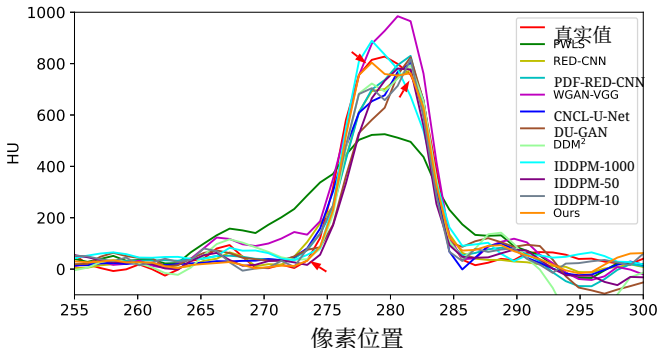

图6. 不同方法得到的图5中蓝线的剖面图。

依据原论文或官方开源代码，而 PWLS 的所有超参数均遵循 [3] 的开源代码中提供的设置。具体而言，PWLS 的总迭代次数设为 20。我们还使用 Mayo 2016数据集中的所有剂量水平的训练数据额外训练了一个 PDF-RED-CNN，称为 PDF-RED-CNN*；我们将原论文中使用的 7 个几何和剂量条件参数调整为 1 个参数，即剂量水平。对于 DDM^2 训练，我们将逐切片预训练的 MRI 去噪模型替换为 Noise2Sim [52]，，这是一种设计完善的低剂量 CT 去噪模型。我们参考了一些专注于为低剂量CT图像去噪任务开发扩散模型的工作，对 IDDPM 进行了修改 [31], [32]。在每个时间步，将低剂量CT图像与采样图像沿通道维度拼接。然后我们输入拼接-将去噪图像输入网络，以引导 IDDPM 生成相应的去噪图像。按照 IDDPM 原论文的建议，我们将 $T = 1000$ 用于 IDDPM 训练，随后在推理时分别采用 1000、50 和 10 个采样步数进行比较；所得模型分别命名为 IDDPM-1000、IDDPM-50 和 IDDPM-10。CoreDiff 在训练和推理时使用相同的步数；除非另有说明，$T = 10$ 用于 CoreDiff，所得模型命名为 CoreDiff-10。

我们使用三种常用的客观图像质量评估指标对去噪性能进行定量评价：峰值信噪比（PSNR）、结构相似性指数（SSIM）和均方根误差（RMSE）。此外，我们还使用了三种新型客观图像质量评估指标，*i.e.* 特征相似性指数（FSIM） [53],视觉信息保真度（VIF） [54],和噪声质量指标（NQM） [55],，这些指标已证明与放射科医生对医学图像所作的主观评价具有更好的一致性 [56]。PSNR、SSIM、FSIM、VIF 和 NQM 越高，RMSE 越低，表示性能越好。除非另有说明，所有指标均基于 [-1000,1000] 亨氏单位的 CT 窗宽计算。

C. Mayo 2016数据集上的性能比较

在本小节中，我们评估了不同模型在 Mayo 2016数据集的 25% 和 5%剂量测试数据上的去噪性能；需要注意的是，所有模型也都使用相同剂量的数据进行了训练。

表 II

Mayo 2016数据集的 5%剂量测试数据上的定量结果（均值±标准差）及单幅图像的去噪计算时间。所报告的深度学习方法的耗时不包括滤波反投影重建所需的时间

	峰值信噪比（dB）	SSIM	均方根误差（亨氏单位）	FSIM	VIF	噪声质量测量（dB）	时间（秒）
FBP	25.49±2.15	0.4310±0.0908	109.5±25.4	0.8427±0.0378	0.379±0.062	23.53±2.70	-
PWLS	34.87±0.95	0.8736±0.0084	36.3±3.9	0.9560±0.0073	0.479±0.045	25.58±2.03	3.50
RED-CNN	37.43±0.95	0.9363±0.0136	27.0±3.0	0.9665±0.0059	0.466±0.036	27.43±1.02	0.01
PDF-RED-CNN	39.25±1.22	0.9445±0.0147	22.0±3.1	0.9737±0.0071	0.527±0.049	29.91±1.98	0.01
WGAN-VGG	34.68±0.77	0.8821±0.0199	37.0±3.3	0.9528±0.0065	0.384±0.036	24.24±1.29	0.01
CNCL-U-Net	38.34±1.07	0.9341±0.0155	24.4±3.0	0.9684±0.0066	0.493±0.045	28.91±1.71	0.02
DU-GAN	37.39±1.13	0.9287±0.0161	27.2±3.5	0.9708±0.0063	0.495±0.046	27.80±1.86	0.01
DDM^2	29.21±1.64	0.5908±0.0804	70.5±13.6	0.9206±0.0180	0.399±0.054	21.10±2.23	28.3
IDDPM-1000	37.31±1.14	0.9177±0.0263	27.5±3.6	0.9695±0.0068	0.482±0.046	28.01±2.03	94.2
IDDPM-50	37.95±1.28	0.9170±0.0362	25.6±3.8	0.9725±0.0064	0.500±0.047	29.04±2.01	4.67
IDDPM-10	34.80±2.52	0.8063±0.1125	38.1±12.8	0.9702±0.0056	0.469±0.041	28.20±1.86	0.96
CoreDiff-10	**40.71±1.26**	**0.9576±0.0123**	**18.6±2.7**	**0.9830±0.0048**	**0.597±0.050**	**32.36±1.95**	0.12

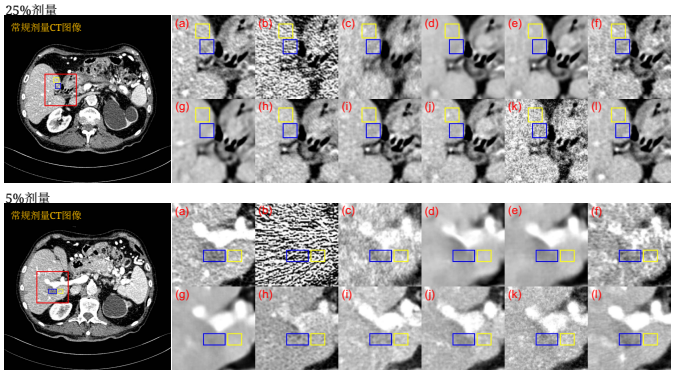

图7. 用于计算对比噪声比的选定感兴趣区的定性结果。（a）常规剂量CT图像（真实值），（b）滤波反投影，（c）惩罚加权最小二乘，（d）RED-CNN，（e）PDF-RED-CNN，（f）WGAN-VGG，（g）CNCL-U-Net，（h）DU-GAN，（i）IDDPM-1000，（j）IDDPM-50，（k）IDDPM-10，以及（l）CoreDiff-10（**我们的方法**）。蓝色感兴趣区从病灶区域中提取，黄色感兴趣区从背景中提取。包含上述两类感兴趣区的红色感兴趣区被放大，以便进行视觉比较。

表III不同方法下病灶感兴趣区的对比噪声比和平均像素值比较

	25%剂量		5%剂量	
	CNR	Mean	CNR	Mean
真实值	1.0456	106.6960	2.2971	64.6133
FBP	0.3482	106.0759	0.2179	74.5219
PWLS	1.0207	109.3893	1.3154	91.1295
RED-CNN	**1.9308**	110.8473	**6.1873**	80.9604
PDF-RED-CNN	1.5268	110.6010	5.4688	78.2443
WGAN-VGG	1.0120	101.5679	1.0363	71.3906
CNCL-U-Net	1.0309	119.2151	4.2600	91.2229
DU-GAN	0.9932	115.3734	1.2580	83.6046
IDDPM-1000	0.7855	112.3634	1.2247	115.3525
IDDPM-50	1.0991	116.8319	2.9835	99.0598
IDDPM-10	0.3222	115.6313	1.7250	91.9786
CoreDiff-10（**我们的方法**）	1.4406	**109.2925**	4.5688	**70.2043**

1）25%剂量评估：图 4 展示了25%剂量测试数据中由不同方法去噪后的一个代表性切片，用于视觉比较。橙色箭头指示了红色感兴趣区域（ROI）中的病灶位置。尽管基于RED-CNN的方法能够有效去除低剂量 CT 图像中的噪声，但往往会模糊精细细节。在基于生成对抗网络的方法中，WGAN-VGG引入了天鹅绒状伪影，而DU-GAN提供了更接近常规剂量CT图像的纹理。CNCL-U-Net保留了最多的细节，但其残差图显示，在骨边缘预测方面存在明显差异。在基于扩散的模型中，DDM^2 表现出明显的伪影和CT值漂移。我们推测，这种现象可能是由于DDM^2 假设图像噪声服从高斯分布，而这与CT图像的实际噪声分布存在偏差。无论是在纹理保留还是细节保持方面，IDDPM和我们的CoreDiff都优于其他对比方法。对于IDDPM，将采样步数 T 降至50对去噪性能影响不大。然而，当 T 降至10时，由于采样严重不足，该模型产生了最差的结果。此外，IDDPM-1000/-50会抹除关键的病灶信息，而我们的CoreDiff能够很好地保留这些信息。残差图证实，我们的方法具有最小的预测偏差。

表 I 展示了所有方法的定量结果。我们的 CoreDiff 优于所有基于深度学习的方法和迭代重建算法。值得注意的是，我们的方法在所有指标上均大幅领先于排名第二的方法（PDF-RED-CNN）。

2) 5%剂量评估：图 5 展示了5%剂量测试数据的定性结果。在这种超低剂量场景下，由于光子饥饿效应，FBP图像受到极其严重的噪声和条纹伪影影响，因此无法用于临床诊断。部分去噪方法的去噪性能大幅下降。图 5 显示，基于RED-CNN的方法和CNCL-U-Net产生了过度平滑的结果。此外，PWLS和WGAN-VGG均在去噪图像中引入了明显的伪影。除基于扩散的模型外，DU-GAN取得了最佳性能。然而，DU-GAN的去噪结果缩小了病灶大小。除IDDPM-10和DDM^2，外，其他基于扩散的模型在超低剂量去噪任务中始终表现出色，展现出在低剂量CT去噪方面的巨大潜力。其中，我们的CoreDiff在残差图和放大的感兴趣区域方面均表现出最佳去噪性能。此外，图 6 展示了不同方法的剖面结果，如图 5中常规剂量CT图像的蓝线所示。红色箭头表明，我们的CoreDiff能够更好地保持CT值

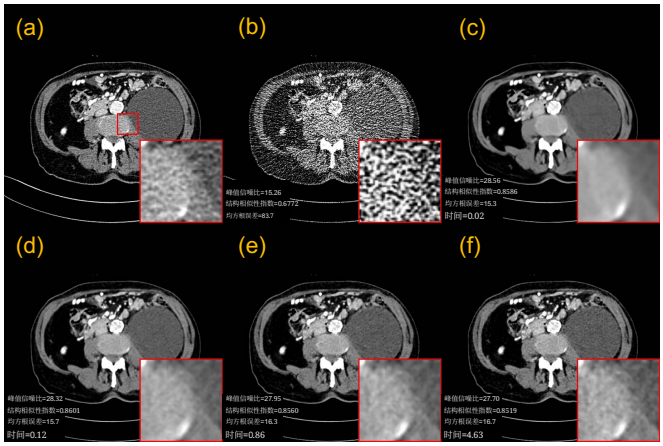

图 8. 针对 CoreDiff 不同 $\boldsymbol{T}$ 设置的消融研究。(a) 常规剂量CT图像（真实值），(b) 滤波反投影，(c) $\boldsymbol{T}=\mathbf{1}$，(d) $\boldsymbol{T}=\mathbf{10}$，(e) $\boldsymbol{T}=\mathbf{50}$，以及 (f) $\boldsymbol{T}=\mathbf{250}$。显示窗为 [-100, 200] 亨氏单位。放大红色感兴趣区中的组织边界以供比较。定量结果列于左下角，并根据显示窗计算。时间（秒）为 CoreDiff 在整个测试集上针对不同 $\boldsymbol{T}$ 设置的平均推理时间。

优于其他方法。

表 II 给出了 5%剂量测试数据的定量结果。我们的 CoreDiff 同样优于所有竞争方法。平均而言，与排名第二的 PDF-RED-CNN 相比，我们的 CoreDiff 的峰值信噪比提高约 +1.46 dB，结构相似性指数提高+1.39%，均方根误差降低 15.45%。此外，表 II 还报告了不同方法对单幅图像进行去噪的计算时间。CoreDiff 的推理速度远快于基于扩散的模型，已达到临床可接受的水平。

此外，我们采用对比噪声比（CNR） [57], [58] 来评估图 4 和 5中低对比度病灶的可检测性。病灶与背景感兴趣区之间的对比噪声比越高，检测到低对比度病灶的概率就越大。如图 7所示，我们从两个切片中仔细选取了蓝色病灶感兴趣区和黄色背景感兴趣区，并在表 III中列出了采用不同方法去噪后各感兴趣区的对比噪声比值。可以看出，RED-CNN 和 PDF-RED-CNN 的对比噪声比值分别位居前两位，而我们的方法排名第三。尽管如此，如图 7所示，RED-CNN 和 PDF-RED-CNN 都会模糊病灶边缘，而病灶边缘对于医生进行疾病分期及判断病灶良恶性至关重要。考虑到在许多临床实践中，CT值常用于区分健康组织与病变组织，我们还在表 III中计算了病灶感兴趣区的平均像素值。值得注意的是，我们的 CoreDiff 所得病灶感兴趣区 CT值最接近真实值。

D. 消融实验

我们进行了消融研究，以考察不同的 T 设置以及 CoreDiff中所有组件的影响。所有模型均使用Mayo 2016数据集中的5% 剂量数据进行训练和测试。

表 IV5%剂量下 Mayo 2016数据集测试数据的不同模块消融研究

Warm	CLEAR-Net		PSNR (dB)	SSIM	RMSE (HU)
	CTX	EMM			
-	-	-	39.96±1.46	0.9501±0.0143	26.8±3.5
✓	-	-	41.59±1.57	0.9628±0.0123	22.2±3.1
✓	✓	-	42.56±1.52	0.9680±0.0109	19.9±2.7
✓	✓	✓	**43.04±1.55**	**0.9713±0.0101**	**18.8±2.7**

表 V使用配对切片或非配对切片（PAIR. VS UNPA.）的一次性学习框架中的权重分布

Dose	Type	$w_0 \longrightarrow w_9$
50%	配对切片。	0.33 0.21 0.14 0.08 0.03 0.03 0.04 0.04 0.05 0.05
	非配对。	0.33 0.22 0.14 0.09 0.03 0.03 0.04 0.04 0.04 0.04
25%	配对。	0.44 0.29 0.12 0.04 0.02 0.01 0.02 0.02 0.02 0.02
	非配对。	0.50 0.28 0.11 0.03 0.01 0.01 0.01 0.01 0.02 0.02

1）不同T 设置：我们评估了CoreDiff在 $T=1$、10、50和250 下进行训练和推理时的性能。图8 展示了不同 T 设置下的去噪图像。当 $T=1$时，CoreDiff退化为单步复原，导致去噪图像出现边缘模糊。随着 T 增大，组织边界逐渐变得清晰，但推理时间也相应增加。此外，随着 T 增大，复原网络的预测误差也会累积。例如，当 $T=1$时，峰值信噪比和均方根误差值最高，对应的像素级误差较少，且图像过度平滑。当$T=10$时，CoreDiff的结构相似性指数最高，去噪图像在视觉上最接近真实值。然而，当设置 $T \geq 50$ 时，CoreDiff的定量性能会逐渐下降。尽管采用了CLEAR-Net，T 较大时，累积误差仍不可忽视。因此，综合考虑去噪图像的锐利度、定量性能和方法的推理时间，$T=10$ 是适用于CoreDiff的设置。

2）不同组件的消融实验：我们进一步考察了采用均值保持退化算子的广义扩散过程、引入的上下文信息以及 CLEAR-Net 中 EMM 的作用。为公平比较，我们将原始冷扩散的总采样步数设为 T，以此作为基线。为简洁起见，将所提出的以低剂量 CT图像（温态）为初始状态的广义扩散过程简称为*Warm*，并将上下文信息的引入简称为*CTX*。表IV 给出了不同组件的定量比较结果，表明所有提出的组件都显著提升了 CoreDiff 的整体去噪性能。平均而言，包含全部组件的 CoreDiff 相较于基线，峰值信噪比提高约 +3.08 dB，结构相似性指数提高 +2.23%，均方根误差降低 29.85%。

E. 面向新剂量和新数据集的单次泛化

我们进一步在四个不同的测试数据集上进行了实验，以评估我们的一次性学习框架的有效性。该框架可用于考察模型对以下数据的泛化能力：1）来自同一数据集的不同剂量（Mayo 2016数据集），

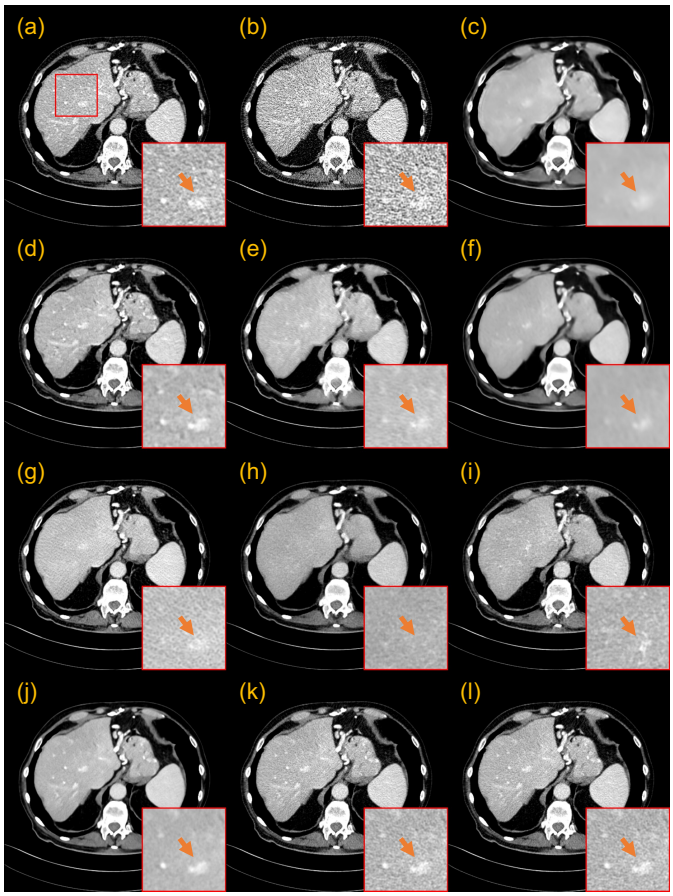

图 9. 使用 5%剂量训练数据训练的不同方法对 50%剂量CT图像进行去噪的定性结果，训练数据和测试数据均来自 Mayo 2016数据集。(a) 常规剂量CT图像（真实值），(b) 滤波反投影，(c) RED-CNN，(d) PDF-RED-CNN*，(e) WGAN-VGG，(f) CNCL-U- Net，(g) DU-GAN，(h) IDDPM-1000，(i) IDDPM-50，(j) CoreDiff-10（**本文方法**），(k) CoreDiff+OSL^{P}-10（**本文方法**），以及 (l) CoreDiff+OSL^{u}-10（**本文方法**）。显示窗为 [-100, 200] 亨氏单位。放大显示红色感兴趣区以进行视觉比较，橙色箭头指向一个关键细节。

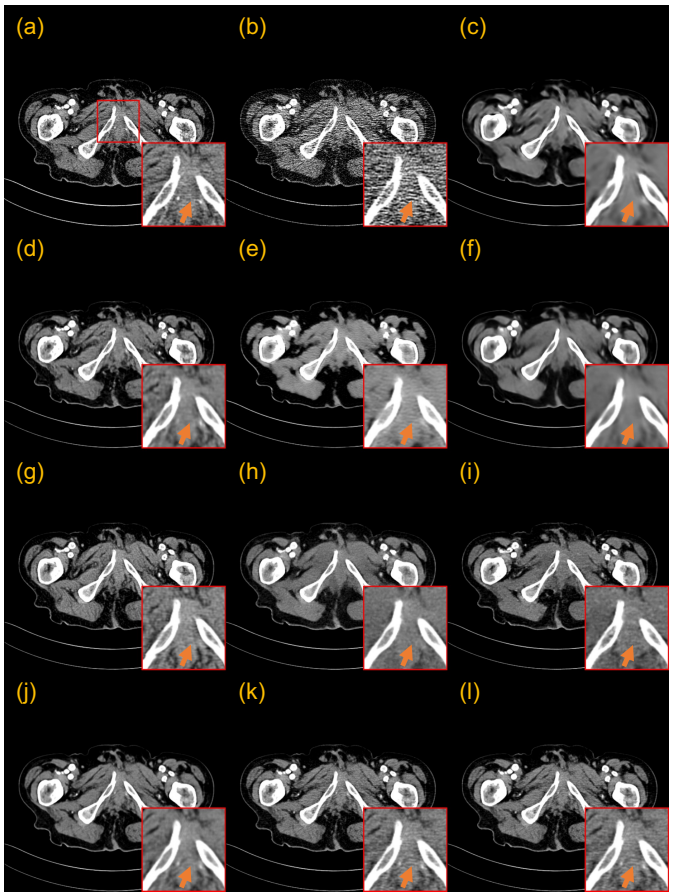

图 10. 使用不同方法对 5%剂量训练数据训练后，对 25%剂量CT图像进行去噪的定性结果；训练数据和测试数据均来自 Mayo 2016数据集。(a) 常规剂量CT图像（真实值），(b) 滤波反投影，(c) RED-CNN，(d) PDF-RED-CNN*，(e) WGAN-VGG，(f) CNCL-U- Net，(g) DU-GAN，(h) IDDPM-1000，(i) IDDPM-50，(j)CoreDiff-10（**本文方法**），(k) CoreDiff+OSL^{P}-10（**本文方法**），以及 (l) CoreDiff+OSL^{u}-10（**本文方法**）。显示窗为 [-100, 200] 亨氏单位。红色感兴趣区已放大，以便进行视觉比较；橙色箭头指向一个关键细节。

2）来自同一中心的不同采集组（Mayo 2020数据集）；3）来自不同中心的不同物种（仔猪数据集）；以及 4）来自不同中心的模体数据（模体数据集）。除 PDF-RED-CNN*外，本小节中的所有实验所用模型（包括我们的方法）均使用 Mayo 2016数据集中的 5% 剂量数据进行训练。具体而言，PDF-RED-CNN* 使用 Mayo 2016数据集中的全部剂量数据进行训练。此外，泛化实验所用测试数据的剂量因数据集而异。对于任一数据集中新的剂量，我们仅选取一张新的低剂量CT图像和一张（非）配对常规剂量CT图像，用于训练式（10）中的 OSL 框架权重。若新的低剂量CT图像与常规剂量CT图像配对，则所得 OSL 模型称为 CoreDiff+OSL^{P}。为降低对新训练数据配对的要求，我们还考虑了非配对情形，即低剂量CT图像和常规剂量CT图像在不同时间采集。为此，我们选取对应低剂量CT图像下方两层的一张常规剂量CT图像作为训练标签，以模拟非配对训练中的轻微位移；所得 OSL 模型称为 CoreDiff+OSL^{u}。

1）在 Mayo 2016数据集上对新剂量水平进行泛化：我们考察了 CoreDiff 对新剂量水平的泛化能力：模型使用 5% 剂量数据训练，并在来自 Mayo 2016数据集的 50% 和 25%剂量测试数据上进行评估。请注意，将 PDF-RED-CNN*、CoreDiff+OSL^{P} 和 CoreDiff+OSL^{u} 与其他方法进行比较并不公平，因为它们使用了额外的 50% 和 25%剂量训练数据。

表V展示了CoreDiff学习得到的权重+OSL^{P}和CoreDiff+OSL^{u}。配对训练和非配对训练得出的权重分布高度相关，表明OSL框架适用于临床应用。

图9 展示了50%剂量低剂量CT图像的定性结果。由于测试图像中的噪声水平高于训练图像，RED-CNN和CNCL-U-Net都会使图像变得平滑。DU-GAN和IDDPM虽然成功保留了图像纹理信息，但血管等某些细节有所丢失。尽管我们的CoreDiff倾向于去除过多的噪声，但它能够保留关键的解剖细节。我们突出显示了我们的CoreDiff+OSL^{P} 和CoreDiff+OSL^{u} 呈现出

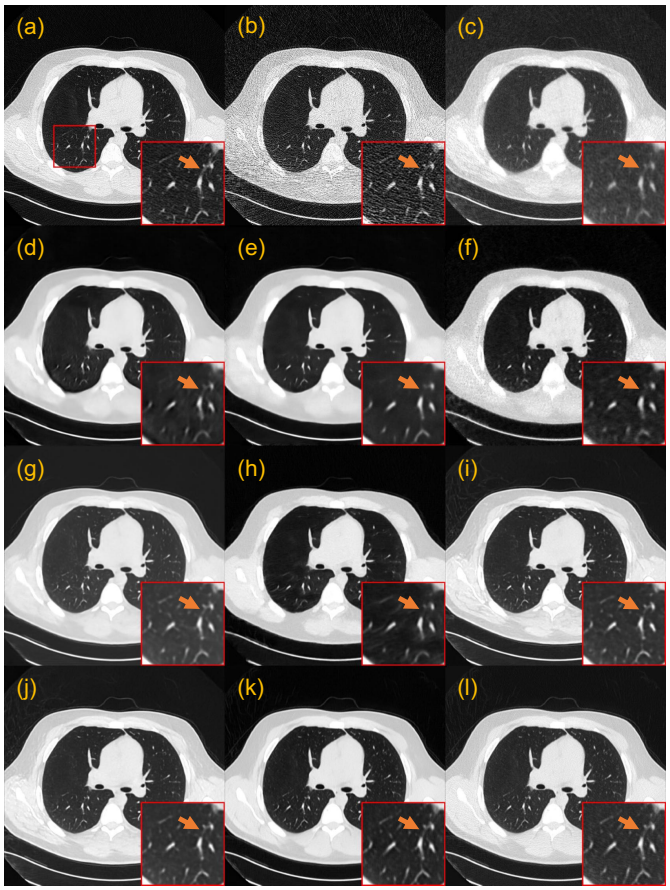

图11。Mayo 2020数据集中一幅10%剂量胸部CT图像的定性结果。(a) 常规剂量CT图像（真实值），(b) 滤波反投影，(c) 惩罚加权最小二乘，(d) RED-CNN，(e) PDF-RED-CNN*，(f) WGAN-VGG，(g) CNCL-U-Net，(h) DU-GAN，(i) IDDPM-1000，(j) IDDPM-50，(k) CoreDiff-10（**我们的方法**），以及 (l) CoreDiff+OSL^{u}-10（**我们的方法**）。显示窗为 [-1350, 150] 亨氏单位。红色感兴趣区已放大以便进行视觉比较，橙色箭头指向一处关键细节。

视觉感知效果更接近真实值。相比之下，PDF-RED-CNN* 可能会引入一些失真。图 10展示了25%剂量低剂量CT图像的定性结果，可以得出类似的结论。

2）对Mayo 2020数据集的泛化：我们进一步评估了所提出的 CoreDiff 和 CoreDiff+OSL^{u} 在 Mayo 2020 数据集的 25%剂量和 10%剂量数据上的有效性。为尽可能贴近临床应用场景，我们分别选取一幅5%剂量胸部切片和一幅25%剂量腹部切片，以及与之不成对的常规剂量切片，训练两个单样本模型。结合这两个独立的单样本模型，我们能够快速将 CoreDiff 泛化到混合剂量水平的测试数据。

图11 展示了 Mayo 2020数据集中一个具有代表性的胸部切片的去噪结果。所有方法都在一定程度上去除了噪声。基于 RED-CNN 的方法会平滑图像，导致感兴趣区域内肺部的诸多细节丢失。PDF-RED-CNN* 会造成一定的细节丢失，部分原因是我们为10%剂量设置的入射光子数与 Mayo Clinic 使用的数值存在差异。与基于 RED-CNN 的方法相比，惩罚加权最小二乘保留了更多信息，但可能无法充分抑制噪声。在生成对抗网络

表 VIMayo 2020数据集上的定量结果（均值±标准差）

	峰值信噪比(dB)	SSIM	均方根误差(亨氏单位)
FBP	29.06±8.72	0.6246±0.3002	108.7±84.2
PWLS	31.96±5.29	0.7653±0.1947	60.3±32.3
RED-CNN	32.57±4.83	0.7650±0.1918	54.7±27.3
PDF-RED-CNN*	33.69±6.20	0.7735±0.1986	52.2±31.0
WGAN-VGG	31.09±4.44	0.7475±0.1922	63.5±29.9
CNCL-U-Net	33.15±4.87	0.7612±0.1954	51.3±25.9
DU-GAN	32.73±5.42	0.7546±0.2030	55.7±30.6
IDDPM-1000	32.41±4.29	0.7849±0.1540	53.8±23.8
IDDPM-50	32.67±4.48	0.7887±0.1543	52.7±24.2
CoreDiff-10	34.10±5.28	0.7964±0.1746	**46.8±24.5**
CoreDiff+OSL^{u}-10	**34.13±5.37**	**0.8107±0.1626**	**46.8±24.9**

方法中，CNCL-U-Net 的表现最佳，但也会导致背景值偏移。基于扩散的方法在噪声抑制和图像保真度之间实现了良好的平衡。最后，结果证明，所提出的 OSL框架有助于 CoreDiff 生成与真实值纹理高度相似的图像。

表 VI给出了定量结果。在比较的方法中，PDF-RED-CNN* 的峰值信噪比最高，这得益于它使用多剂量数据进行训练；而 IDDPM 的结构相似性指数最高，表明其对结构信息的保留效果更好。令人惊讶的是，即使仅使用 5% 剂量数据进行训练，且未采用额外的 OSL 框架，我们的 CoreDiff 仍取得了更好的性能。我们推测，这是因为更合理的基于扩散的方法能够让模型逐步去除图像中的噪声，并避免结构失真。此外，上下文误差调制模块使 CoreDiff 对输入变化更加稳健。OSL 框架进一步提升了峰值信噪比和结构相似性指数，这表明 CoreDiff+OSL^{u} 能够通过学习最优去噪图像，生成视觉上更加真实的纹理和结构信息。

3）在仔猪数据集上的泛化：我们还研究了所提出的单次框架在提升仔猪数据集上泛化性能方面的效果。我们在仔猪数据集的 25% 和 10% 剂量数据上测试了所有训练好的模型。我们还从该仔猪数据集中选取了一张切片及其（非）配对的正常剂量切片，用于训练我们的 CoreDiff+OSL^{p} 和 CoreDiff+OSL^{u}。

图 12 和 13 展示了25%剂量和10%剂量下两个具有代表性切片的去噪结果。在10%剂量的情况下，RED-CNN、WGAN-VGG和CNCL-U-Net会使去噪图像严重模糊。DU-GAN和IDDPM虽然避免了整体平滑，但仍导致重要局部细节丢失。PDF-RED-CNN* 和CoreDiff成功保留了精细细节。结合OSL框架后，CoreDiff的去噪图像+OSL^{p} 和CoreDiff+OSL^{u} 呈现出最接近常规剂量CT图像的纹理。CoreDiff+OSL^{p} 和CoreDiff+OSL^{u} 的定量性能也优于CoreDiff，并胜过其他方法。在25%剂量的情况下，令人惊讶的是，除CoreDiff+OSL^{p} 和CoreDiff+OSL^{u} 之外，其余方法的定量指标甚至都比滤波反投影更差。这一观察结果凸显了基于深度学习的模型在跨多中心、多物种计算机断层扫描数据泛化方面的局限性

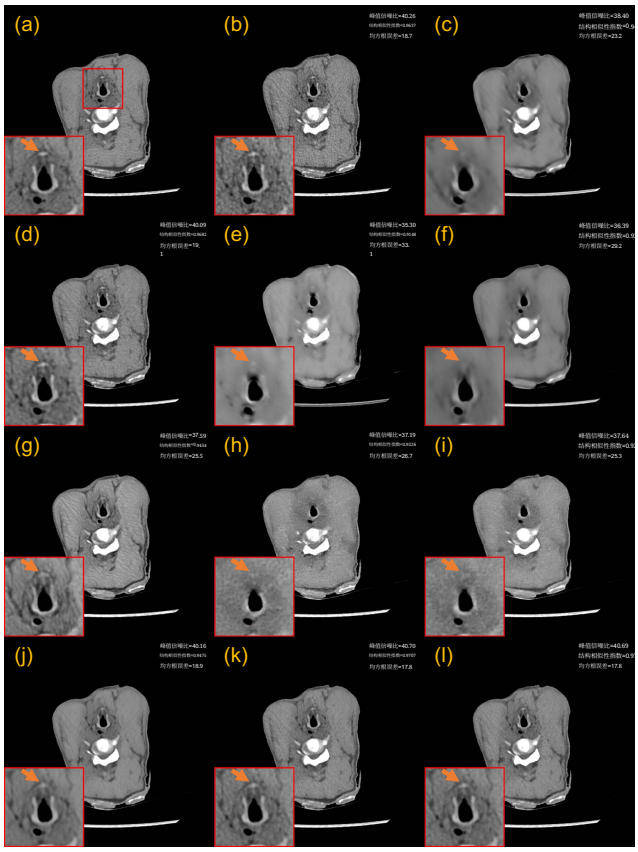

图12。 不同方法对25%剂量仔猪CT图像进行去噪的定性结果。（a）常规剂量CT图像（真实值），（b）滤波反投影，（c）RED-CNN，（d）PDF-RED-CNN*，（e）WGAN-VGG，（f）CNCL-U-Net，（g）DU-GAN，（h）IDDPM-1000，（i）IDDPM-50，（j）CoreDiff-10（**我们的方法**），（k）CoreDiff+OSL^{P}-10（**我们的方法**），以及（l）CoreDiff+OSL^{U}-10（**我们的方法**）。显示窗为 [-100, 200] 亨氏单位。红色感兴趣区经过放大以进行视觉比较，橙色箭头指向关键细节。

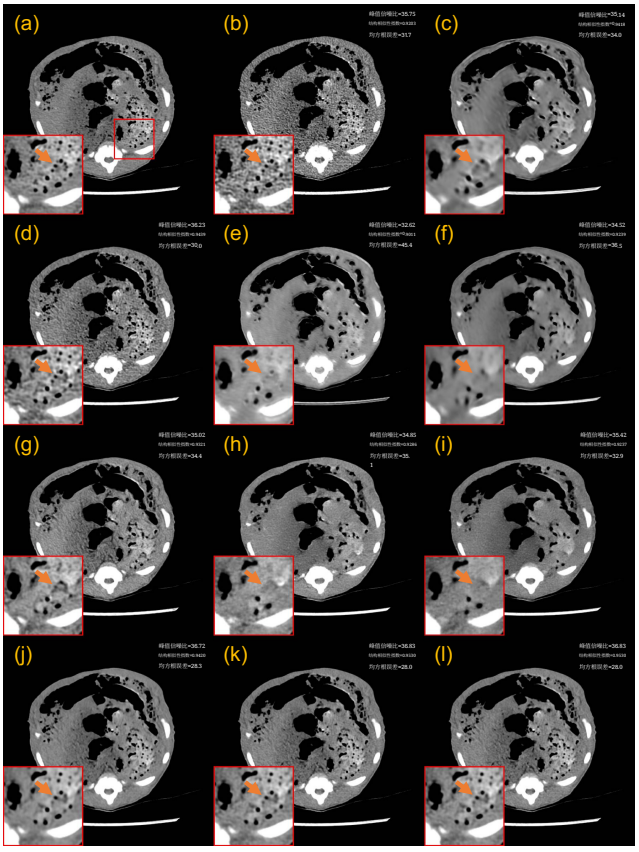

图 13. 不同方法对10%剂量仔猪CT图像进行去噪的定性结果。(a) 常规剂量CT图像（真实值），(b) 滤波反投影，(c) RED-CNN，(d) PDF-RED-CNN*，(e) WGAN-VGG，(f) CNCL-U-Net，(g) DU-GAN，(h) IDDPM-1000，(i) IDDPM-50，(j) CoreDiff-10（**本文方法**），(k) CoreDiff+OSL^{P}-10（**本文方法**），以及 (l) CoreDiff+OSL^{U}-10（**本文方法**）。显示窗为 [-100, 200] 亨氏单位。放大红色感兴趣区域以进行视觉比较，橙色箭头指向关键细节。

化。然而，在我们的OSL框架中，CoreDiff+OSL^{P} 和 CoreDiff+OSL^{U}生成的去噪图像始终表现出优异的视觉质量和定量性能。

4）在模体数据集上的泛化：为进一步考察所提出CoreDiff、CoreDiff+OSL^{P}和CoreDiff+OSL^{U}的临床实用性，我们在模体数据集上开展了泛化实验，并选择通道化Hotelling观察者（CHO），对不同方法在低对比度信号检测任务上的性能进行定量评估。已有研究探索了CHO在评估不同计算机断层扫描图像重建算法的低对比度检测任务性能方面的效用，结果表明其与人类观察者具有高度相关性 [59]–[61]。采用信噪比（SNR）衡量CHO使用不同方法去噪图像作为输入时的检测性能；SNR越大，性能越好。如图14 和15所示，我们将蓝色有信号感兴趣区域和黄色无信号背景感兴趣区域输入CHO进行信号检测。所有训练模型均在该模体数据集中的54.31%和21.64%剂量数据上进行测试。对于PDF-RED-CNN*，我们使用最接近的对应剂量（25%和50%）作为参数依赖框架中的条件输入。

表 VII以不同方法去噪后的图像作为输入的通道化Hotelling观察者SNR

	271 毫安秒（54.31%）	108 毫安秒 (21.64%)
NDCT	10.92	10.92
FBP	9.09	7.71
RED-CNN	5.72	4.49
PDF-RED-CNN*	7.95	6.70
WGAN-VGG	7.51	4.95
CNCL-U-Net	8.09	5.68
DU-GAN	7.45	7.71
IDDPM-1000	4.37	4.10
IDDPM-50	2.58	2.49
CoreDiff-10	8.24	8.03
CoreDiff+OSL^{P}-10	**11.48**	**9.40**
CoreDiff+OSL^{U}-10	11.39	9.36

图 14 和 15 展示了不同方法对剂量分别为 54.31% 和 21.64% 的图像进行去噪后的代表性切片。RED-CNN 和 CNCL-U-Net 模糊了圆柱形植入物的边缘，而 WGAN-VGG 和 IDDPM 则引入了明显的伪影。DU-GAN 对去噪图像的噪声抑制不足。尽管 PDF-RED-CNN* 优于其他对比方法，但与我们的 CoreDiff 相比，在 21.64% 剂量场景下，其保留边缘细节的能力仍显不足。所提出的 OSL 框架有助于模型生成接近常规剂量 CT图像的背景纹理

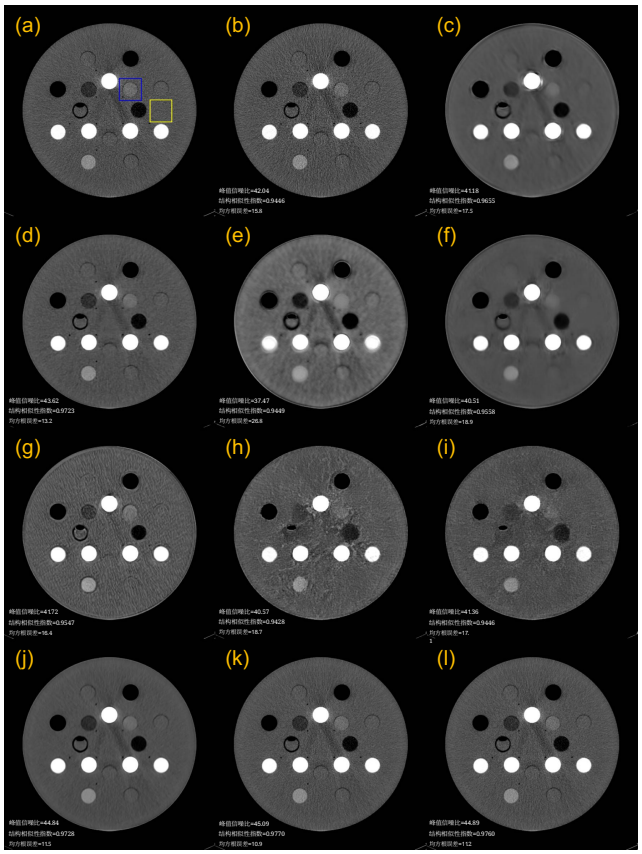

图14。不同方法对54.31%剂量下模体CT图像进行去噪的定性结果。(a) 常规剂量CT图像（真实值），(b) 滤波反投影，(c) RED-CNN，(d) PDF-RED-CNN*，(e) WGAN-VGG，(f) CNCL-U-Net，(g) DU-GAN，(h) IDDPM-1000，(i) IDDPM-50，(j) CoreDiff-10（**我们的方法**），(k) CoreDiff+OSL^{P}-10（**我们的方法**），以及(l) CoreDiff+OSL^{U}-10（**我们的方法**）。显示窗为 [-100, 200]亨氏单位。选取蓝色的有信号感兴趣区域和黄色的无信号背景感兴趣区域作为CHO信号检测的输入。

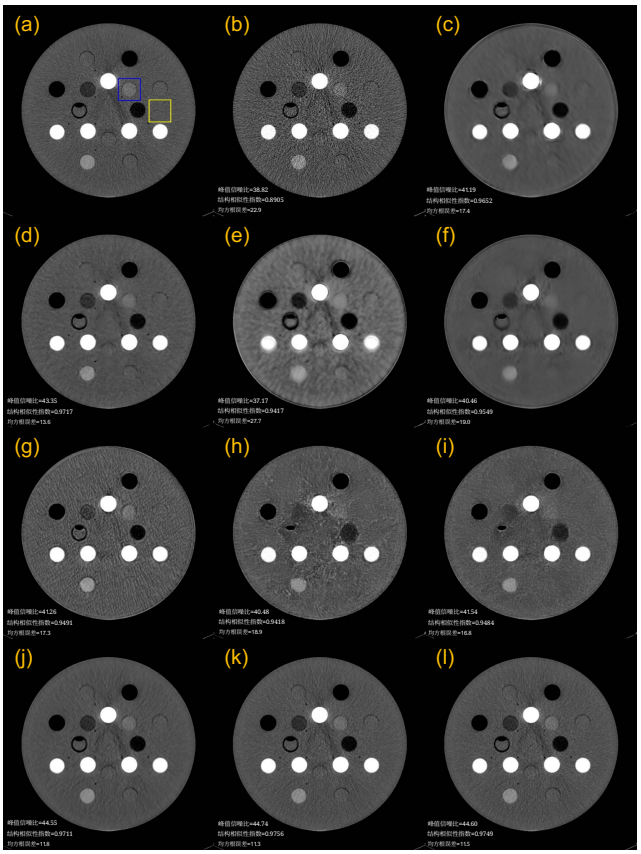

图 15. 不同方法对剂量为 21.64% 的模体CT图像进行去噪的定性结果。(a) 常规剂量CT图像（真实值），(b) 滤波反投影，(c) RED-CNN，(d) PDF-RED-CNN*，(e) WGAN-VGG，(f) CNCL-U-Net，(g) DU-GAN，(h) IDDPM-1000，(i) IDDPM-50，(j) CoreDiff-10（**本文方法**），(k) CoreDiff+OSL^{P}-10（**本文方法**），以及 (l) CoreDiff+OSL^{U}-10（**本文方法**）。显示窗为 [-100, 200] 亨氏单位。选择蓝色的信号存在感兴趣区域和黄色的信号缺失背景感兴趣区域作为 CHO 的输入，以进行信号检测。

，从而获得更理想的视觉效果。

表VII 展示了不同方法针对特定任务的定量性能。除CoreDiff外，其他方法的性能甚至不及滤波反投影，这可能是由于测试数据（模体）与训练数据（患者）之间存在分布偏移。与定性分析一致，我们的CoreDiff性能仍然最佳。与OSL框架结合后，模型针对特定任务的性能进一步提升，在54.31%剂量场景下甚至超过了常规剂量CT。

IV. 讨论

A. 关于 CoreDiff 优势的讨论

实验结果表明，在定量、定性和特定任务性能方面，我们的 CoreDiff 优于其他竞争模型，展示了广义扩散模型在低剂量CT去噪中的潜力。需要指出的是，即使采用 1,000 步采样，为低剂量CT去噪任务改进的 IDDPM 仍逊于我们的 CoreDiff。下面讨论所提出的 CoreDiff 相较于其他基于扩散模型的低剂量CT

从以下三个方面介绍去噪方法。1）确定性采样过程带来的益处：近年来，已有多种基于扩散模型的低剂量CT 去噪方法被提出 [31], [32]。然而，其中大多数遵循经典的高斯扩散模型框架，将低剂量CT图像去噪任务转化为条件生成任务。高斯扩散模型的采样过程可统一表示为随机微分方程（SDE），以确保对目标分布的良好覆盖 [28]。但需要注意的是，图像去噪任务与图像生成任务不同，它只需在去噪图像与干净图像之间建立一一映射。在本研究中，我们基于广义扩散模型构建了确定性采样过程，不仅加快了采样速度，也有利于图像去噪。2）所提出的均值保持退化算子的优势：图2 显示，随着剂量降低，高斯扩散模型中常用的退化算子会偏离 CT图像退化的物理过程。所提出的均值保持退化算子不仅在扩散过程的每个时间步引入低剂量CT图像特有的噪声和伪影，还

确保中间图像的每个像素的 CT值都与常规剂量CT图像中的对应像素具有相同的均值。3）所提出的 *CLEAR-Net* 带来的益处：由于扩散模型追求像素级精度，因此在图像去噪任务中，采样过程中的误差累积问题尤为突出。本研究提出了一种新型复原网络 CLEAR-Net，它能够利用上下文信息约束采样过程，防止结构失真，并调节时间步嵌入特征，使其在下一时间步与输入更好地对齐。

B. 关于单次学习框架的讨论

我们还在四个数据集上评估了所提出的单次学习框架，大量实验结果证实，即使资源极少，该框架相较于现有方法仍具有很强的泛化能力。在此，我们将 5% 视为超低剂量情形，而大多数临床扫描实际使用的剂量都超过 5%。然而，当测试数据的剂量低于某个训练数据的剂量时，CoreDiff 在经过 10 个采样步数后，即使未使用 OSL，其输出也已是我们能够获得的最佳去噪结果。因此，仍需进一步研究如何使 CoreDiff 能够利用高剂量可用训练数据，泛化到低剂量测试数据。近期工作，如 GMM-U-Net [41] 和 PDF [15]，也致力于解决泛化问题。我们的工作与它们的主要区别如下。1）它们基于以往的深度学习框架，而这些框架在低剂量CT成像质量方面存在局限。我们首次探索了如何增强扩散模型在复杂低剂量CT去噪场景中的鲁棒性。2）PDF 在训练阶段将剂量、成像几何参数及其他信息作为条件输入，以增强其重建各种测试数据的鲁棒性。然而，对于几何参数和剂量水平在训练期间未曾出现过的测试数据，其性能会受到影响。GMM-U-Net 则采用高斯混合模型（GMM）来表征低剂量CT图像的噪声分布，因此需要针对不同数据集凭经验选择不同数量的高斯模型及损失函数权重。相比之下，我们的单次学习框架训练仅使用一幅新的低剂量CT图像和一幅（非）配对常规剂量CT图像，无需针对测试数据凭经验调整超参数，从而能够快速适应未见过的新剂量水平。3）我们验证了 CoreDiff 在要求极高的超低剂量成像任务中的潜力。相比之下，它们所使用的测试数据剂量均高于我们的超低剂量水平。

C. 局限性讨论

我们承认本研究存在一些局限性。首先，在实际场景中，由于低剂量CT图像中的噪声均值可能并不严格为 0，因此，所提出的退化算子并不是严格意义上的“均值保持退化算子”。低剂量CT图像中的噪声成因十分复杂，可能受到多种因素的影响，包括散射、束硬化、患者运动、金属植入物等 [62]。因此，还需要进一步改进退化算子，以更好地近似实际的低剂量 CT 物理退化过程。其次，尽管我们的 CoreDiff 采样速度显著快于 DDM^2 和 IDDPM，但其推理时间仍是基于 RED-CNN 和生成对抗网络的模型的 10 倍。近期有一些旨在加速扩散模型的技术，或许可以整合到我们的 CoreDiff 中，以进一步减少采样步数。例如，Stable Diffusion [63] 所采用的潜在空间采样技术，以及一致性模型所使用的知识蒸馏技术 [64]。然而，采样步数的减少意味着单次学习框架将使用更少的中间图像，因此需要在泛化性能与推理时间之间进行权衡。第三，由于测试数据仅包含两个病灶，开展任务驱动读者研究无法获得具有统计显著性的结果。我们计划在未来的研究中纳入更大的测试患者数据集，以开展多读者、多病例研究（MRMC）研究。

V. 结论

在本研究中，我们提出了一种用于低剂量CT图像去噪的新型上下文误差调制广义扩散模型，称为 CoreDiff。所提出的 CoreDiff 采用：(i) 将低剂量CT图像作为扩散过程的信息性端点；(ii) 引入均值保持退化算子，以模拟 CT 退化的物理过程；(iii) 提出复原网络 CLEAR-Net，以减轻累积误差和错位；以及 (iv) 设计单次学习框架，以提升泛化能力。实验结果证明了 CoreDiff 的有效性，尤其是在超低剂量情形下。值得注意的是，我们的 CoreDiff 模型仅需 10 个采样步数，因而在临床应用中比经典扩散模型快得多。

参考文献

[1] R. Smith-Bindman 等人，“常见计算机断层扫描检查相关的辐射剂量及其终生归因癌症风险”，《内科学档案》，第169卷，第2078–2086页，2009年。[2] A. Sodickson 等人，“成人重复 CT、累积辐射暴露及 CT 相关的辐射诱发癌症风险”，《放射学》，第251卷，第175–184页，2009年。[3] Q. Xie 等人，“通过利用噪声生成机制实现稳健的低剂量 CT 正弦图预处理”，《*IEEE* 医学影像汇刊》，第36卷，第12期，第2487–2498页，2017年。[4] J. Wang、T. Li、H. Lu 和 Z. Liang，“用于正弦图噪声抑制和低剂量 X 射线计算机断层扫描图像重建的惩罚加权最小二乘方法”，《*IEEE* 医学影像汇刊》，第25卷，第10期，第1272–1283页，2006年。[5] H. Shan 等人，“通过从二维训练网络进行迁移学习实现用于低剂量 CT 的三维卷积编码器-解码器网络”，《*IEEE* 医学影像汇刊》，第37卷，第6期，第1522–1534页，2018年。[6] M. Aharon、M. Elad 和 A. Bruckstein，“K-SVD：一种用于设计稀疏表示超完备字典的算法”，《*IEEE* 信号处理汇刊》，第54卷，第11期，第4311–4322页，2006年。[7] P. F. Feruglio、C. Vinegoni、J. Gros、A. Sbarbati 和 R. Weissleder，“用于吸收光学投影断层成像的三维块匹配随机噪声滤波”，《医学与生物学物理》，第55卷，第18期，第5401页，2010年。[8] Y. Chen 等人，“基于快速字典学习处理改善腹部肿瘤低剂量 CT 图像”，《医学与生物学物理》，第58卷，第16期，第5803页，2013年。[9] J. Ma 等人，“利用既往常规剂量扫描进行低剂量计算机断层扫描图像复原”，《医学物理学》，第38卷，第5713–5731页，2011年。

[10] Z. Li 等人，“基于局部噪声水平的自适应非局部均值滤波用于 CT 去噪”，《医学物理学》，第41卷，第1期，第011908页，2014年。[11] K. Sheng、S. Gou、J. Wu 和 S. X. Qi，“经去噪和纹理增强的兆伏级计算机断层扫描图像用于改善软组织显影度”，《医学物理学》，第41卷，第10期，第101916页，2014年。[12] H. Shan 等人，“模块化深度神经网络与商业算法在低剂量 CT 图像重建中的性能比较”，《自然·机器智能》，第1卷，第269–276页，2019年。[13] G. Wang、J. C. Ye 和 B. De Man，“用于断层图像重建的深度学习”，《自然·机器智能》，第2卷，第748页，2020年，第12期，第737页。[14] H. Chen 等人，“采用残差编码器-解码器卷积神经网络的低剂量 CT”，《IEEE 医学影像汇刊》，第36卷，第2524–2535页，2017年。[15] W. Xia 等人，“采用 PDF 的 CT 重建：适用于多种几何结构和剂量水平数据的参数依赖框架”，《IEEE 医学影像汇刊》，第40卷，第11期，第3065–3076页，2021年。[16] B. Kim、M. Han、H. Shim 和 J. Baek，“基于卷积神经网络的图像去噪方法的性能比较：损失函数对低剂量 CT 图像的影响”，《医学物理学》，第46卷，第3906–3923页，2019年。[17] M. Nagare、R. Melnyk、O. Rahman、K. D. Sauer 和 C. A. Bouman，“用于 CT 图像去噪的偏差抑制损失函数”，载于《IEEE 声学、语音与信号处理国际会议论文集》，2021年，第1175–1179页。[18] E. Kang、H. J. Koo、D. H. Yang、J. B. Seo 和 J. C. Ye，“用于多期冠状动脉CT血管成像的循环一致对抗去噪网络”，《医学物理学》，第46卷，第550–562页，2019年。[19] Q. Yang 等人，“利用 Wasserstein距离和感知损失的生成对抗网络进行低剂量 CT 图像去噪”，《IEEE 医学影像汇刊》，第37卷，第1348–1357页，2018年。[20] Z. Huang、J. Zhang、Y. Zhang 和 H. Shan，“DU-GAN：采用双域 U-Net 判别器的低剂量 CT 去噪生成对抗网络”，《IEEE仪器与测量汇刊》，第71卷，第1–12页，2021年。[21] S. Zhao、H. Ren、A. Yuan、J. Song、N. Goodman 和 S. Ermon，“深度生成模型中的偏差与泛化：一项实证研究”，载于《神经信息处理系统进展》，第31卷，2018年。[22] J. Sohl-Dickstein、E. Weiss、N. Maheswaranathan 和 S. Ganguli，“利用非平衡热力学进行深度无监督学习”，载于《国际机器学习会议论文集》，2015年，第2256–2265页。[23] J. Ho、A. Jain 和 P. Abbeel，“去噪扩散概率模型”，载于《神经信息处理系统进展》，第33卷，2020年，第6840–6851页。[24] J. Song、C. Meng 和 S. Ermon，“去噪扩散隐式模型”，载于《国际学习表征会议论文集》，2021年。[25] A. Nichol 和 P. Dhariwal，“改进的去噪扩散概率模型”，载于《国际机器学习会议论文集》，2021年，第8162–8171页。[26] H. Chung、D. Ryu、M. T. McCann、M. L. Klasky 和 J. C. Ye，“利用预训练二维扩散模型解决三维逆问题”，载于《IEEE 计算机视觉与模式识别会议论文集》，2023年。[27] H. Chung、B. Sim、D. Ryu 和 J. C. Ye，“利用流形约束改进逆问题的扩散模型”，载于《神经信息处理系统进展》，2021年。[28] Y. Song、J. Sohl-Dickstein、D. P. Kingma、A. Kumar、S. Ermon 和 B. Poole，“通过随机微分方程进行基于分数的生成建模”，载于《国际学习表征会议论文集》，2021年。[29] P. Dhariwal 和 A. Nichol，“ 扩散模型在图像合成方面胜过生成对抗网络 ”，载于《神经信息处理系统进展》，第34卷，2021年，第8780–8794页。[30] L. Yang 等人，“扩散模型：方法与应用综述”，arXiv:2209.00796，2022年。[31] Q. Gao 和 H. Shan，“CoCoDiff：用于低剂量 CT 图像去噪的上下文条件扩散模型”，载于《SPIE论文集》，第12242卷，2022年。[32] W. Xia、Q. Lyu 和 G. Wang，“利用去噪扩散概率模型进行低剂量 CT 20× 加速”，arXiv:2209.15136，2022年。[33] A. Bansal 等人，“冷扩散：无噪声地反演任意图像变换”，载于《神经信息处理系统进展》，2023年。[34] H. Yen、F. G. Germain、G. Wichern 和 J. L. Roux，“用于语音增强的冷扩散”，arXiv:2211.02527，2022年。[35] N. Joshi 和 M. F. Cohen，“观赏雷尼尔山：用于多图像去噪、锐化和去雾的幸运成像”，载于《计算摄影国际会议论文集》，2010年，第1–8页。[36] J. Lehtinen 等人，“Noise2Noise：无需干净数据的图像复原学习”，载于《国际机器学习会议论文集》，2018年，第2965–2974页。[37] D. Wu、K. Gong、K. Kim、X. Li 和 Q. Li，“仅使用含噪训练样本进行医学影像去噪的共识神经网络”

《医学图像计算与计算机辅助干预国际会议论文集》，2019年，第741–749页。[38] E. Perez、F. Strub、H. De Vries、V. Dumoulin 和 A. Courville，“FiLM：通过通用条件层进行视觉推理”，载于《AAAI人工智能会议论文集》，第32卷，第1期，2018年。[39] X. Huang 和 S. Belongie，“利用自适应实例归一化实时进行任意风格迁移”，载于《IEEE国际计算机视觉会议论文集》，2017年，第1501–1510页。[40] V. Dumoulin、J. Shlens 和 M. Kudlur，“艺术风格的学习表征”，载于《国际学习表征会议论文集》，2017年。[41] D. Li、Z. Bian、S. Li、J. He、D. Zeng 和 J. Ma，“用于稳健CT图像重建的噪声特性建模无监督网络”，《IEEE 医学影像汇刊》，第41卷，第12期，第3849–3861页，2022年。[42] H. Shan、U. Kruger 和 G. Wang，“一种用于低剂量 CT 的新型迁移学习框架”，载于《SPIE论文集》，第11072卷，2019年，第513–517页。[43] B. Chen、S. Leng、L. Yu、D. Holmes III、J. Fletcher 和 C. McCollough，“开放的CT患者投影数据库”，载于《SPIE论文集》，第9783卷，2016年，第330–335页。[44] D. Zeng 等人，“基于高剂量扫描的简单低剂量X射线CT模拟”，《IEEE核科学汇刊》，第62卷，第5期，第2226–2233页，2015年。[45] T. Zhao、M. McNitt-Gray 和 D. Ruan，“用于超低剂量CT去噪和肺气肿筛查的卷积神经网络”，《医学物理学》，第46卷，第9期，第3941–3950页，2019年。[46] T. R. Moen 等人，“低剂量CT图像与投影数据集”，《医学物理学》，第48卷，第2期，第902–911页，2021年。[47] X. Yi 和 P. Babyn，“利用条件生成对抗网络进行锐度感知低剂量CT去噪”，《数字成像杂志》，第31卷，第655–669页，2018年。[48] I. Zhovannik 等人，“从扫描仪中学习：影像组学中的偏差降低和特征校正”，《临床与转化放射学》，第19卷，第33–38页，2019年。[49] S. Guo、Z. Yan、K. Zhang、W. Zuo 和 L. Zhang，“面向真实照片卷积盲去噪”，载于《IEEE计算机视觉与模式识别会议论文集》，2019年，第1712–1722页。[50] L. Shi 等人，“CT图像重建开源工具包综述”，《X射线科学与技术杂志》，第28卷，第4期，第619–639页，2020年。[51] M. Geng 等人，“用于医学图像去噪的内容-噪声互补学习”，《IEEE 医学影像汇刊》，第41卷，第2期，第407–419页，2021年。[52] C. Niu 等人，“基于相似性的自监督深度学习噪声抑制”，《IEEE 医学影像汇刊》，2022年。[53] L. Zhang、L. Zhang、X. Mou 和 D. Zhang，“FSIM：用于图像质量评估的特征相似性指数”，《IEEE图像处理汇刊》，第20卷，第8期，第2378–2386页，2011年。[54] H. R. Sheikh 和 A. C. Bovik，“图像信息与视觉质量”，《IEEE图像处理汇刊》，第15卷，第2期，第430–444页，2006年。[55] N. Damera-Venkata、T. D. Kite、W. S. Geisler、B. L. Evans 和 A. C. Bovik，“基于退化模型的图像质量评估”，《IEEE图像处理汇刊》，第9卷，第4期，第636–650页，2000年。[56] A. Mason 等人，“客观图像质量指标与放射科医生对MR图像诊断质量评分的比较”，《IEEE 医学影像汇刊》，第39卷，第4期，第1064–1072页，2019年。[57] R. Gutjahr 等人，“临床剂量水平下基于光子计数的CT人体成像：对比噪声比与尸体研究”，《调查放射学》，第51卷，第7期，第421页，2016年。[58] P. C. Park 等人，“用于改善多期计算机断层扫描中肝胆肿瘤对比噪声比和可检测性的强化模式映射技术”，《医学物理学》，第47卷，第1期，第64–74页，2020年。[59] D. Van de Sompel、M. Brady 和 J. Boone，“基于信号对比噪声比和通道化Hotelling观察者，对数字乳腺断层合成中FBP、SART和ML进行基于任务的性能分析”，《医学图像分析》，第15卷，第1期，第53–70页，2011年。[60] A. Wunderlich、F. Noo、B. D. Gallas 和 M. E. Heilbrun，“图像质量研究中通道化Hotelling观察者性能的精确置信区间”，《IEEE 医学影像汇刊》，第34卷，第2期，第453–464页，2014年。[61] L. Yu 等人，“在CT多层阅读的低对比度检测任务中，二维通道化Hotelling观察者与人类观察者之间的相关性”，《医学物理学》，第44卷，第8期，第3990–3999页，2017年。[62] K. Kalisz、J. Buethe、S. S. Saboo、S. Abbara、S. Halliburton 和 P. Rajiah，“心脏CT伪影：物理原理与解决方案”，《放射学图谱》，第36卷，第7期，第2064–2083页，2016年。[63] R. Rombach、A. Blattmann、D. Lorenz、P. Esser 和 B. Ommer，“利用潜在扩散模型进行高分辨率图像合成”，载于《IEEE计算机视觉与模式识别会议论文集》，2022年，第10684–10695页。[64] Y. Song、P. Dhariwal、M. Chen 和 I. Sutskever，“一致性模型”，载于《国际机器学习会议论文集》。

,” in , 2023年。