**中图法分类号**:TP18;TP391.41　　**文献标识码**:A　　**文章编号**:1006-8961(2026)10-3779-23

**论文引用格式**:Zhang S N, Sun Z, Yu H, Gao Z H and Ding G A. 2026. HRDRL-Net: a deep reinforcement learning framework for high-resolution reconstruction of low-dose computed tomography images. Journal of Image and Graphics, 31(10):3779-3801(张胜楠, 孙正, 于寒, 高章硕, 丁港澳. 2026. 面向低剂量CT图像高分辨率重建的深度强化学习框架HRDRL-Net. 中国图象图形学报, 31(10):3779-3801)[DOI:10.11834/jig.250502]

# 面向低剂量CT图像高分辨率重建的深度强化学习框架HRDRL-Net

张胜楠[1],孙正[1,2*],于寒[1],高章硕[1],丁港澳[1]

1. 华北电力大学电子与通信工程系,保定　071003; 2. 华北电力大学河北省电力物联网技术重点实验室,保定　071003

**摘　要**:**目的**　低剂量计算机断层扫描(low-dose computed tomography,LDCT)在降低辐射剂量方面具有重要临床价值,但其图像质量常受噪声与伪影影响。现有改善LDCT图像质量的深度学习方法在模型泛化能力、计算效率及对噪声的自适应性方面仍存在不足。为此,提出一种基于深度强化学习的高分辨率重建框架(high resolution deep reinforcement learning network,HRDRL-Net),旨在实现噪声抑制与结构保留的有效平衡。**方法**　将LDCT去噪任务建模为序列决策过程,采用异步优势动作评价算法作为智能体训练基础。通过构建双路径多分支协同架构与低剂量噪声抑制模块,并设计融合像素级误差、梯度相似性及局部方差约束的复合奖励函数,动态引导智能体在丰富动作空间中学习自适应去噪策略。**结果**　在Mayo与Piglet公开数据集上的实验表明,HRDRL-Net重建图像的定量指标优于主流基线方法。在Mayo测试集上,与基线模型相比,HRDRL-Net的峰值信噪比平均提高约1.3%,结构相似性指数平均提高约0.7%,梯度幅度相似性偏差降低约6%。在Piglet数据集上,峰值信噪比与结构相似性指数较基线模型平均提高约1.5%和0.8%。消融实验证实了复合奖励、完整动作集、双路径架构及多分支模块对方法性能的有效性。**结论**　HRDRL-Net能够有效抑制LDCT图像噪声并保留关键解剖结构与纹理细节,在重建质量、泛化能力与计算效率之间取得了良好平衡。

**关键词**:低剂量计算机断层扫描(LDCT);图像重建;深度学习;深度强化学习(DRL);多尺度双重注意力

# HRDRL-Net: a deep reinforcement learning framework for high-resolution reconstruction of low-dose computed tomography images

Zhang Shengnan[1], Sun Zheng[1,2*], Yu Han[1], Gao Zhangshuo[1], Ding Gang'ao[1]

1. *Department of Electronic and Communication Engineering, North China Electric Power University, Baoding　071003, China*;
2. *Hebei Key Laboratory of Power Internet of Things Technology, North China Electric Power University, Baoding　071003, China*

**Abstract**: **Objective** Low-dose computed tomography (LDCT) has become an important imaging modality for reducing patient radiation exposure in diagnostic and screening protocols. However, the inherent trade-off between radiation dose and image fidelity introduces severe quantum noise and streak artifacts, which degrade anatomical visibility and reduce diagnostic confidence. Although deep-learning-based reconstruction methods, especially convolutional neural networks (CNNs) and generative adversarial networks (GANs), have made remarkable progress in learning direct mappings from LDCT to normal-dose CT (NDCT), they still suffer from several intrinsic limitations. These include a tendency toward

**收稿日期**:2025-10-21;**修回日期**:2026-03-02;**预印本日期**:2026-03-09

*通信作者:孙正　sunzheng@ncepu.edu.cn

**基金项目**:国家自然科学基金项目(62571188)

**Supported by**: National Natural Science Foundation of China (62571188)

excessive smoothing that erodes fine pathological detail, sensitivity to distribution shifts between training and clinical data, high computational complexity, and a static processing paradigm that cannot adapt to the spatially heterogeneous noise characteristics typical of LDCT scans. To address these challenges, this paper introduces HRDRL-Net (high-resolution deep reinforcement learning network), a novel framework that reformulates LDCT reconstruction as an adaptive, pixel-level decision-making process. The goal is to achieve an optimal, context-aware balance between aggressive noise suppression and faithful preservation of diagnostically important structures. **Method** HRDRL-Net is grounded in the theory of Markov decision processes (MDPs) and implemented using the asynchronous advantage actor-critic (A3C) algorithm. Its core innovation is a dual-path, multi-branch cooperative architecture. The upper path contains a dedicated low-dose noise suppression module (LDNM) with four branches—dilated convolution, identity mapping, detail enhancement, and noise suppression—combined with a dual-attention mechanism that includes both channel and spatial attention for multi-scale semantic feature extraction and refinement. The lower path uses a predefined Sobel operator followed by convolutional layers to explicitly extract and enhance structural edge information. Features from the two paths are fused adaptively through a gating mechanism, producing a rich state representation for the agent. A comprehensive action space of 14 predefined image-processing operators is further designed, including Gaussian filters of different strengths, edge-preserving bilateral filters, median filtering, non-local means filtering, guided filtering, and adaptive sharpening. This allows the agent to select the most appropriate local processing strategy for each region. Training is guided by a composite reward function with three components: 1) a pixel-level fidelity term that minimizes mean squared error, 2) a gradient-structure similarity term that enforces edge and texture consistency, and 3) a local variance penalty term that selectively suppresses noise in homogeneous regions without over-smoothing high-gradient anatomical details. Both the policy network (actor) and the value network (critic) are implemented as fully convolutional networks, enabling efficient per-pixel action-probability prediction and state-value estimation. The entire system is trained end to end: the agent iteratively interacts with the LDCT image environment, receives rewards based on its reconstruction actions, and updates its policy to maximize cumulative future quality gains. **Result** HRDRL-Net was rigorously evaluated on two publicly available benchmark datasets, the Mayo Clinic LDCT Challenge Dataset and the Piglet Dataset. It was compared with five state-of-the-art deep learning methods: residual encoder-decoder convolutional neural network (RED-CNN), dual domain generative adversarial network, DU-GAN), CT-former, self-adaptive weight embedded lightweight network (SWELNet), and multi-stage multi-order context aggregation (MMCA). Quantitative results on the Mayo testset show that HRDRL-Net achieved a peak signal-to-noise ratio (PSNR) of 32.319 ± 1.016 dB, a structural similarity index measure (SSIM) of 0.897 ± 0.020, and a gradient magnitude similarity deviation (GMSD) of 0.063 ± 0.008, outperforming all competing methods. Compared with the baselines, this corresponds to an average improvement of ~1.3% in PSNR, ~0.7% in SSIM, and a ~6% reduction in GMSD. On the Piglet dataset, the model achieved a PSNR of 32.263 ± 1.158 dB and an SSIM of 0.900 ± 0.009, confirming strong cross-domain generalization. Statistical significance was verified by Friedman tests ($p < 0.001$ for all metrics on both datasets) followed by post hoc Nemenyi tests. Ablation studies further validated each core design component: the composite reward improved PSNR by ~0.7 dB compared with a pixel-only reward; the full action set yielded an approximately 1.2% SSIM gain over a basic action set; the dual-path architecture reduced GMSD by ~2.3% relative to single-path variants; and the complete LDNM improved PSNR by 0.8 dB over a simple dilated-convolution baseline. Visual analysis, supported by residual maps and edge-intensity profiles, shows that HRDRL-Net achieves thorough noise removal in homogeneous regions while preserving fine vasculature, organ boundaries, and tissue textures, whereas competing methods tend to leave residual noise, blur structures, or distort anatomy. In addition, HRDRL-Net remains computationally practical, requiring only 0.3 GB of GPU memory and offering competitive inference speed, thereby balancing reconstruction quality and resource efficiency. **Conclusion** This work presents HRDRL-Net as a deep reinforcement learning framework that addresses the fundamental tension between noise suppression and detail preservation in LDCT reconstruction. By formulating reconstruction as a sequential, adaptive decision-making problem, HRDRL-Net moves beyond static processing pipelines. Its combination of a dual-path feature extractor, a rich action space, and a composite reward function enables intelligent pixel-level adaptation to local image content. Extensive experiments demonstrate its superiority over current methods in quantitative metrics, visual quality, statistical significance, and generalization, while maintaining practical efficiency. The frame-

work therefore provides a robust, high-fidelity, and clinically promising solution, and it points to a new direction for intelligent, adaptive processing in medical imaging.

**Key words**: low-dose computed tomography (LDCT); image reconstruction; deep learning; deep reinforcement learning (DRL); multi-scale dual attention

## 0 引　言

计算机断层扫描(computed tomography, CT)在医学诊断中应用广泛,但其临床应用存在辐射剂量与图像质量的平衡问题。低剂量CT(low-dose CT, LDCT)虽然能降低辐射风险,但会引入显著的量子噪声与条状伪影,影响对细微病变的识别准确性(张永高 等,2025),这一局限性推动了LDCT重建方法的持续发展。

早期方法主要围绕投影域滤波与迭代重建展开,其中投影域方法(Ehman等,2012)在原始信号端进行噪声抑制,易损失高频诊断信息;迭代重建算法(Park等,2024;Yu等,2021;Gui等,2023)虽然提升了图像质量,却因计算复杂、参数调优困难且依赖人工先验,难以适应复杂的噪声环境。近年来,基于深度学习的重建方法显著提升了性能。例如,残差编码—解码卷积神经网络(residual encoder-decoder convolutional neural network, RED-CNN)(Chen 等, 2017)通过编码器—解码器结构学习噪声映射,但在强噪声环境下泛化能力不足;CT-former(Wang等,2023)利用Transformer架构增强远程依赖建模,但其重叠推理机制导致计算成本高昂;基于Swin Transformer V2与卷积特征融合的方法(利铭康 等,2025)通过并行特征提取,在保证去噪性能的同时提升了计算效率;多尺度大核注意力特征融合网络(宋霄罡 等,2025)则利用大核可分离卷积块在维持轻量化设计的同时,兼顾类Transformer的全局建模能力与卷积的局部感知优势;自适应权重嵌入式轻量级半监督网络(self-adaptive weight embedded lightweight network, SWELNet)(Wang等,2025)通过多尺度特征融合与半监督策略提升泛化性能,但依赖仿真与真实数据分布的一致性;多阶段多阶上下文聚合(multi-stage multi-order context aggregation, MMCA)框架(Li等,2025)通过上下文聚合增强特征表达,但其边缘增强模块难以适应多样化解剖结构。

生成对抗网络(generative adversarial network, GAN)的引入为LDCT重建提供了新思路(Zhao等,2025;Li等,2024)。例如,双域GAN(dual domain GAN, DU-GAN)通过图像域和梯度域判别器提升结构保持能力,并利用CutMix技术可视化不确定性(Huang等,2022);尺度敏感GAN通过多尺度判别机制识别异质噪声(Wang等,2024b);轻量化GAN则探索了实时处理的可行性(Hojjat等,2024)。此外,无监督方法如CycleGAN(Li 等, 2021; Hossain 等, 2025)和GAN-CIRCLE(You等,2020)尝试摆脱对配对数据的依赖,推动去噪与超分辨率重建任务的联合优化。然而,现有GAN方法多采用静态前馈处理,难以根据图像局部特征动态调整策略,对噪声与细节的权衡仍然依赖人工参数设定(Sadia 等, 2024)。

深度强化学习(deep reinforcement learning, DRL)将图像重建建模为序列决策过程,通过智能体与环境的交互实现自适应优化(Furuta等,2020;乌兰 等,2025)。该范式已在CT技术中用于重建算法的参数调优(Patwari 等, 2022)和扫描策略优化(Wang等,2024a),展现出动态决策潜力。然而,现有DRL方法尚未充分解决LDCT重建中局部特征感知与多目标权衡的问题。

针对上述局限,本文提出一种高分辨率深度强化学习重建框架(high resolution deep reinforcement learning network, HRDRL-Net),其创新性体现在3个方面:首先,将像素级恢复策略引入图像重建过程,通过多尺度双重注意力机制,实现了对局部特征与全局语义的协同感知;其次,设计融合像素保真度、梯度结构相似性与局部统计特性的复合奖励函数,解决了噪声抑制与细节保留的权衡问题;最后,构建包含多种图像处理操作的强化学习动作空间,使模型能够根据图像内容自适应选择最优处理策略。这一框架突破了传统静态处理的限制,实现了基于图像特征的动态优化,为LDCT重建提供了新的技术路径。

# 1 方 法

如图1所示，所提方法的技术流程包含4个核心环节：噪声分布标准化、特征协同提取、强化学习决策和迭代终止控制。首先，通过Anscombe变换将LDCT图像中的泊松噪声转换为高斯噪声，以提高后续处理的数值稳定性；然后，构建基于异步优势动作评价（asynchronous advantage actor-critic，A3C）框架（Mnih等，2016）的智能体，通过多尺度特征提取与复合奖励机制实现自适应去噪；最终，通过状态评估模块动态控制重建过程，经Anscombe逆变换输出高质量重建结果。

## 1.1 强化学习框架设计

本文构建的HRDRL-Net强化学习重建框架严格遵循马尔可夫决策过程，其核心要素包括智能体、环境与两者交互的闭环决策流程。智能体完成从观察状态到生成动作的决策过程。环境为输入的LDCT图像；状态由智能体深度提取的多尺度特征表示；动作选自涵盖多类型滤波与增强算子的预设空间；奖励由像素级误差、梯度相似性及局部方差约束的复合函数计算，用于引导策略优化。通过智能体与环境的持续交互，迭代优化策略与价值评估网络。

采用A3C框架作为智能体训练基础的理论依据如下：首先，LDCT图像重建需在像素/局部块级别进行细粒度决策，其状态空间维度高且关联性强。A3C的异步多线程架构能通过并行探索有效打破高维图像状态样本间的相关性，提高训练效率与策略鲁棒性。其次，本文任务的动作空间包含多种连续、多维度滤波与增强操作。传统深度Q网络（deep Q-network，DQN）（Mnih等，2015）方法需对连续动作进行离散化，易导致维度灾难，且难以刻画动作价值在图像不同区域（如边缘与均匀组织）的显著差异。而A3C的Actor-Critic架构可直接输出连续动作空间上的策略分布，更适合此类细粒度自适应决策。最后，在优化稳定性方面，近端策略优化（proximal policy optimization，PPO）（Schulman等，2017）方法依赖批量样本统计与复杂约束调参，可能增加推理延迟与调优难度。A3C通过优势函数实现更稳定的策略更新，在保证性能的同时实现更简洁、更适合CT图像噪声环境下的高效稳定训练。

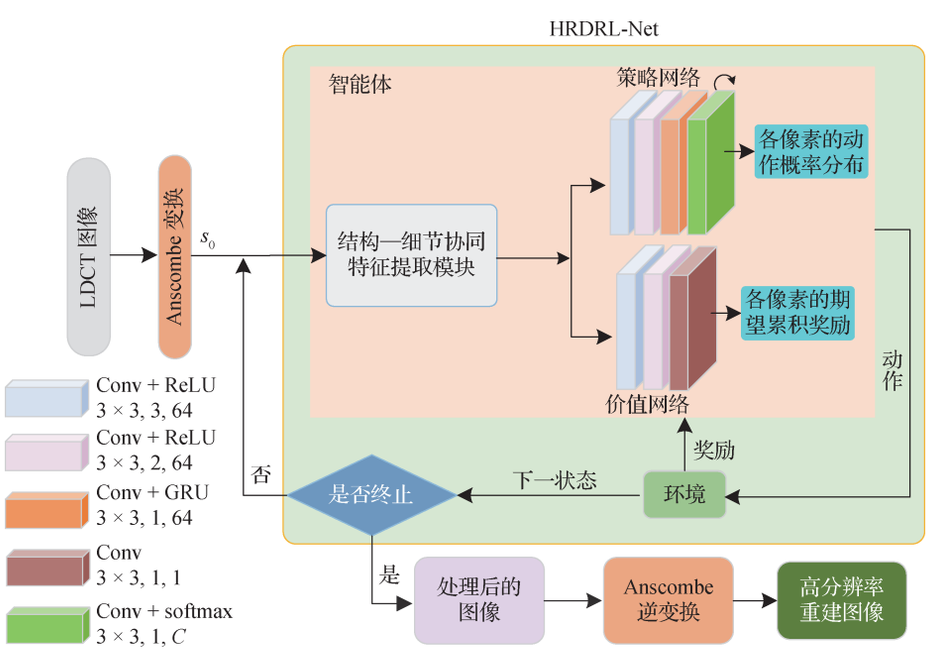



图1 基于DRL的LDCT图像重建方法原理示意图

Fig. 1 Schematic diagram of the principle of LDCT image reconstruction method based on DRL

## 1.2 智能体架构设计

智能体由特征提取模块、策略网络和值网络3个核心组件构成。特征提取模块采用结构—细节协同设计，通过双路径架构分别处理语义特征和边缘信息；策略网络基于全卷积网络构建，输出每个像素的动作概率分布；值网络采用相同架构，评估状态

价值函数。两个网络通过参数共享机制协同优化，既保证决策效率，又确保价值评估准确性。

1.2.1　结构—细节协同特征提取模块

结构—细节协同特征提取（structure-detail collaborative feature extractor，SDCFE）模块基于多尺度特征相融合设计，通过将多尺度卷积、边缘增强和双重注意力相结合，在有效抑制噪声的同时保持和增强血管及微小病灶等关键结构和细节。

如图2所示，输入图像通过双路径并行处理进行特征提取，其中上路径用于捕获图像的多尺度语义信息，下路径提取和增强图像的结构边界信息。这种分离式设计可以确保边缘特征在深层网络中不会丢失，为后续细节增强提供结构指导。在上路径中，输入图像首先经过3×3卷积层进行基础特征提取，之后特征图输入低剂量噪声抑制（low-dose noise mitigation，LDNM）模块中，通过四分支并行架构进行多尺度特征处理：1）膨胀卷积分支：采用并行化非对称膨胀率设计（dilation = 1，2，5），通过特征拼接整合多尺度感受野信息，有效扩大其范围。2）身份映射分支：通过1×1卷积保留原始特征信息，防止梯度消失并维持信息完整性。3）细节增强分支：采用多尺度设计，其中小尺度分支通过1×3和3×1的级联卷积捕获血管、小病变等精细结构；中尺度分支通过两个3×3的级联卷积处理中等大小的解剖结构。两个分支的特征在通道维度拼接后，经过融合卷积层整合，精确增强细节区域。4）噪声抑制分支：通过两个连续的3×3卷积层构建深层特征，针对LDCT图像中的噪声模式进行抑制。

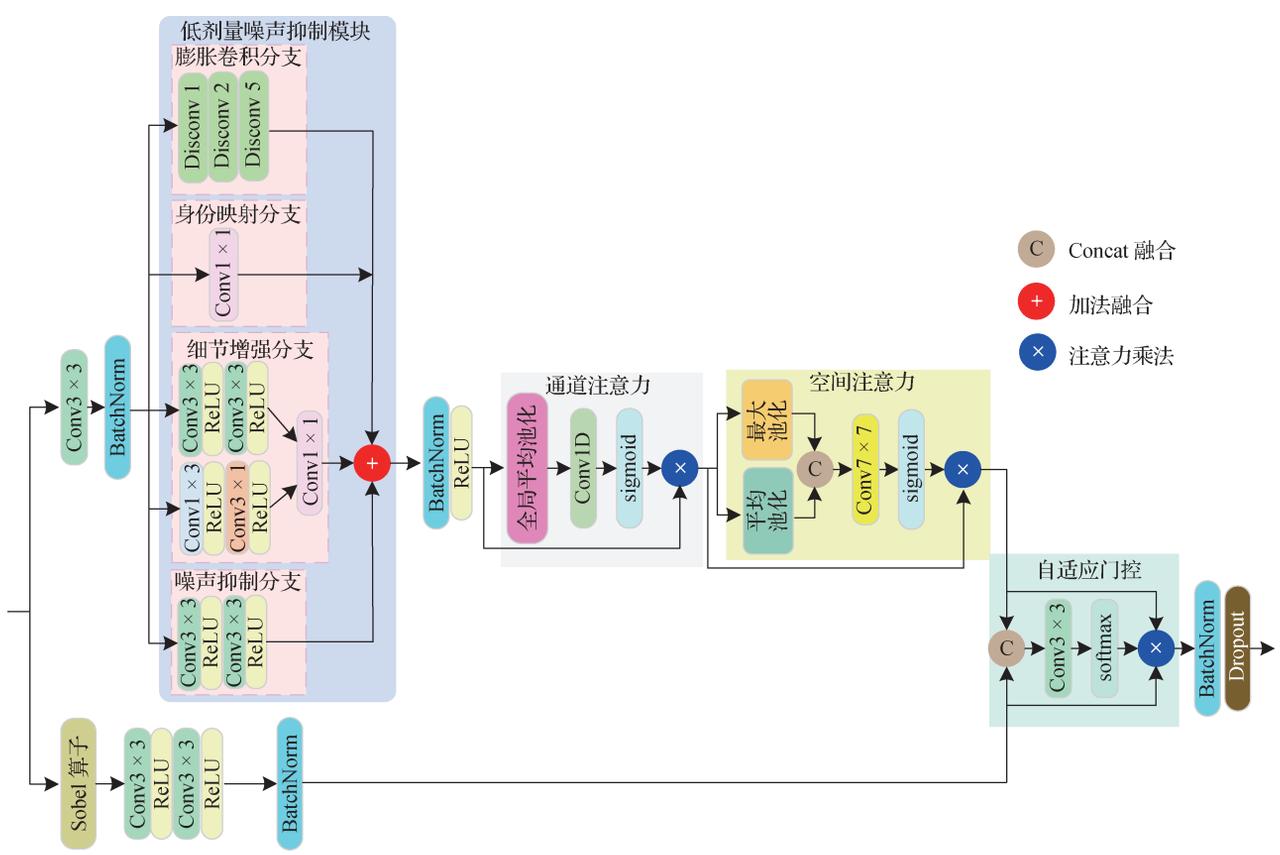



图2　SDCFE模块的结构示意图

Fig. 2　Structural schematic diagram of the SDCFE module

上述4个分支的输出通过元素级加法融合，形成富含多尺度信息的特征图。进而，通过双重注意力策略，从通道与空间两个维度协同优化特征表征能力：通道注意力（efficient channel attention，ECA）分支利用一维卷积自适应调整通道权重，增强重要特征通道的表达；空间注意力（spatial attention，SA）分支结合平均池化和最大池化的空间描述子，经过7×7卷积生成空间注意力权重图，精确反映图像中各位置的重要性分布。在下路径中，通过预定义的Sobel算子对输入图像进行边缘检测，其后接两级3

×3卷积层深化边缘特征表达。

最后，上路径输出的特征图与下路径提取的边缘特征通过自适应门控机制融合，动态平衡语义特征与边缘特征的贡献比例，避免特征冲突导致的细节损失，为后续强化学习决策提供高精度状态表征。

1.2.2 策略网络和值网络

策略网络和值网络均以当前图像状态作为输入，但分别优化不同的目标函数：策略网络专注于动作选择的最优化，而值网络则准确评估状态价值，二者协同工作，共同推动智能体学习最优去噪策略。

策略网络同时处理所有像素的状态，以SDCFE模块输出的多尺度特征图为输入，前向计算首先经过两层卷积操作：第1层为步长3的3×3卷积（输出64通道），完成特征的初步压缩与编码；第2层为步长2的3×3卷积（输出64通道），进一步细化特征表达并调整特征图尺寸。随后，连接步长1的门控循环单元（gated recurrent unit，GRU）层，建模像素之间的空间依赖关系，同时消除相邻智能体间的干扰。网络的最后一层通过卷积生成对应动作的特征映射，经过softmax函数归一化后，得到各像素的动作概率分布，具体为

$$\pi(a_{i,t}\left|s_{i,t};\theta_{\pi}\right.)=\frac{\exp(L_{k,i})}{\sum_{l=1}^{14}\exp(L_{l,i})} \tag{1}$$

式中，$i=1,2,\cdots,N$表示像素索引，$N$是总像素数，$t$为时间步，$a_{i,t}$和$s_{i,t}$分别是第$i$个像素的当前动作和当前状态，$\theta_{\pi}$是策略参数，$k=1,2,\cdots,14$表示动作索引，$L_{k,i}$是第$i$个像素对应动作$a_k$的原始输入分数。训练时对$\pi(a_{i,t}\left|s_{i,t};\theta_{\pi}\right.)$随机采样，为每个像素匹配适配的去噪操作，在保证探索多样性的同时，推动策略逐步收敛至最优。

值网络与策略网络的输入相同，其第1层为步长2的3×3卷积（输出32通道），对输入特征进行初步降维与编码；第2层为步长1的3×3卷积（输出1通道），完成特征的最终整合。值网络的输出是每个像素状态的期望累积奖励$V(s_{i,t};\theta_v)$，其中$\theta_v$是值网络的参数，其输出的1通道特征图对应每个像素的状态价值。在初始状态下，每个像素的状态表示为$s_{i,0}=\boldsymbol{I}(i)$，其中$\boldsymbol{I}(i)$是输入图像$\boldsymbol{I}$中第$i$个像素的强度值。训练过程中，时间步$t$的折扣总奖励定义为

$$R^{(t)}=r^{(t)}+\gamma r^{(t+1)}+\gamma^2 r^{(t+2)}+\cdots+\gamma^{n-1}r^{(t+n-1)}+\gamma^n r^{(t+n)} \tag{2}$$

式中，$r^{(t)}$是即时奖励，$\gamma$是归一化折扣因子。通过最小化预测状态价值与目标价值（当前奖励+下一状态价值的折扣值）的均方误差（mean squared error，MSE），提高状态价值评估的准确性，为策略网络的动作选择提供可靠的价值参考。

### 1.3 奖励函数的设计

LDCT图像重建的本质是在强噪声背景下恢复出与标准剂量CT（normal dose CT，NDCT）在结构、纹理与诊断信息上高度一致的图像。传统的像素级奖励通过最小化当前图像与目标图像之间的差异来抑制噪声，然而过度依赖均方误差往往导致图像过度平滑，模糊边缘结构，损失重要的解剖结构和病理细节。为此，本文从图像质量评价理论出发，设计多维度复合奖励函数，以引导智能体在训练过程中学习针对不同噪声模式与局部图像特征的差异化处理策略，动态平衡噪声抑制与结构保留。

首先，基于信息论中的保真度准则，建立像素级奖励项$r_{\text{pixel}}$，直接衡量当前图像$\boldsymbol{I}_t$与参考图像$\boldsymbol{I}_{\text{ref}}$之间的差异，定义为

$$r_{\text{pixel}}=s\cdot[V_{\text{pixel}}(\boldsymbol{I}_t)-V_{\text{pixel}}(\boldsymbol{I}_{t-1})] \tag{3}$$

式中，$s$是调整奖励幅度的缩放因子，$V_{\text{pixel}}$是像素级价值函数，具体为

$$V_{\text{pixel}}(\boldsymbol{I}_t)=-\left\|\boldsymbol{I}_t-\boldsymbol{I}_{\text{ref}}\right\|_2^2 \tag{4}$$

$r_{\text{pixel}}$作为奖励函数的基础组成部分，通过误差的时序差分形式提供基础优化方向，促使智能体持续优化去噪效果。

其次，基于人类视觉系统特性的结构相似性理论，构建梯度相似性奖励项$r_{\text{gradient}}$，具体为

$$r_{\text{gradient}}=SSIM(\nabla\boldsymbol{I}_t,\nabla\boldsymbol{I}_{\text{ref}}) \tag{5}$$

式中，$SSIM(\cdot,\cdot)$表示结构相似度计算，$\nabla\boldsymbol{I}$表示图像$\boldsymbol{I}$的梯度，计算式为

$$\nabla\boldsymbol{I}_t=[S_x(\boldsymbol{I}_t),S_y(\boldsymbol{I}_t)]^{\mathrm{T}} \tag{6}$$

式中，$S_x$和$S_y$为$x$和$y$方向的Sobel算子。$r_{\text{gradient}}$通过计算当前图像与参考图像在梯度域的相似性，保护高频结构信息，尤其是在噪声较强的区域，保护血管分支、组织边界等高梯度区域的结构完整性。

最后，基于图像统计特性，设计局部方差惩罚项$r_{\text{var}}$，定义为

$$r_{\text{var}}=\sum_p M_{\text{smooth}}(p)\left[Var_{\Omega_p}(\boldsymbol{I}_{t-1}(p))-Var_{\Omega_p}(\boldsymbol{I}_t(p))\right] \tag{7}$$

式中，$p$是图像中的像素，$M_{\text{smooth}}$是梯度感知掩码，用于降低高梯度区域的惩罚强度，它基于原始LDCT图像$I_{\text{LDCT}}$的梯度信息生成，以区分边缘区域与平滑区域，具体为

$$M_{\text{smooth}}(p) = \mathbf{1}(\left\| \nabla \boldsymbol{I}_{\text{LDCT}}(p) \right\|_2 < \tau) \tag{8}$$

式中，$\mathbf{1}(\cdot)$是指示函数，$\tau$为梯度阈值。该掩码在平滑区域（低梯度区域）值为1，在边缘区域（高梯度区域）值为0。$Var_{\Omega_p}(\boldsymbol{I}_t(p))$是图像$\boldsymbol{I}_t$中以像素$p$为中心的局部邻域$\Omega_p$内的方差，具体为

$$Var_{\Omega_p}(\boldsymbol{I}_t(p)) = \frac{1}{N_{\Omega_p}} \sum_{q \in \Omega_p} [\boldsymbol{I}_t(q) - \mu_{\Omega_p}]^2 \tag{9}$$

式中，$N_{\Omega_p}$为邻域内的像素总数，$\mu_{\Omega_p}$为邻域内的灰度均值。$r_{\text{var}}$通过计算局部区域的方差，自适应地区分真实解剖纹理（高方差高梯度）与噪声（高方差低梯度），并通过平滑掩码减少其对整体图像质量的影响，实现更精准的噪声抑制，在抑制残留噪声的同时保持合理的纹理特征。

复合奖励函数定义为

$$r_{\text{total}} = \lambda_1 r_{\text{pixel}} + \lambda_2 r_{\text{gradient}} + \lambda_3 r_{\text{var}} \tag{10}$$

式中，$\lambda_1$、$\lambda_2$和$\lambda_3$是权重参数，通过网格搜索优化确定，确保各目标间的平衡。该复合奖励函数通过协同优化像素级保真度、结构相似性与局部统计特性，旨在为智能体提供一个能够区分并自适应平衡图像中均匀区域噪声抑制与高梯度细节保留的多维度优化目标。

### 1.4 动作空间设计

本文针对LDCT图像的特性，基于图像退化模型设计动作空间。如表1所示，动作集包含14种图像处理操作，分为基础去噪和结构保持两类，覆盖了图像恢复的主要技术路径。其中，高斯滤波类动作针对加性噪声模型，双边滤波保持边缘结构，非局部均值滤波利用图像自相似性先验。锐化类动作基于图像增强理论，补偿滤波过程中的细节损失。深度网络增强动作通过数据驱动方式学习残差映射。这种涵盖多种滤波类型与强度的动作空间设计，旨在赋予智能体自适应处理不同区域、不同强度噪声的灵活决策能力。

**表1 动作空间设置**

**Table 1 Action space settings**

| 类别 | 动作类型 | 参数 | 作用 |
|---|---|---|---|
| 基础去噪 | 无动作 | | 状态保持 |
| | 高斯滤波 | $5 \times 5, \sigma = 0.5$ | 轻度噪声去除 |
| | 高斯滤波 | $5 \times 5, \sigma = 1.5$ | 重度噪声去除 |
| | 双边滤波 | $5 \times 5, \sigma_c = 0.1, \sigma_s = 5$ | 基于像素值相似性和空间邻近性的保边去噪，平滑均匀区域的同时保护弱边缘 |
| | 双边滤波 | $5 \times 5, \sigma_c = 1, \sigma_s = 5$ | 增强颜色域平滑能力，在保留明显边缘的同时，处理强度变化较大的噪声区域 |
| | 中值滤波 | $5 \times 5$ | 去除脉冲噪声 |
| 结构保持 | 非局部均值滤波 | $h = 5, 10, 15$ | 结构保护去噪 |
| | 引导滤波 | $eps = 0.01$ | 边缘保持平滑 |
| | 自适应锐化 | $\alpha = 1.5, 2.5$ | 粗细结构增强 |
| | 卷积增强 | | 改善图像清晰度 |
| | 残差增强 | | 改善图像细节 |

基础去噪部分包含6个核心动作：以无操作动作为基准选项，保持当前状态不变，为智能体提供“不作为”的选项；两个不同方差（$\sigma = 0.5$或$\sigma = 1.5$）的$5 \times 5$高斯滤波器分别用于处理不同强度的噪声区域；两个$5 \times 5$双边滤波器通过调节颜色空间参数（颜色空间标准差$\sigma_c$）和坐标空间参数（坐标空间标准差$\sigma_s$）实现差异化平滑效果，两种不同参数配置的双边滤波组合使用，能够适应CT图像中不同组织的密度差异；$5 \times 5$中值滤波器则专门针对脉冲噪声。

此外，针对CT图像特有的解剖结构特征，设计了8个结构保持动作：非局部均值滤波（滤波强度参数$h$ = 5, 10, 15）利用图像自相似性保护组织结构；引导滤波（正则化参数$eps$ = 0.01）在平滑均匀区域的同时保持边缘清晰度；自适应锐化滤波（锐化因子$\alpha$ = 1.5或2.5）通过高斯差分方法增强边缘和细节；卷积增强采用3层卷积网络进行清晰度提高，其中第1层为特征提取层，使用9×9卷积核将3通道输入映射到64通道特征；第2层为非线性映射层，使用5×5卷积核将64通道特征映射到32通道；第3层为重建层，使用5×5卷积核将32通道特征重建为3通道输出，参数量约为57 K。残差增强采用20层残差网络进行细节恢复，第1层为初始特征提取层，使用3×3卷积核将3通道输入映射到64通道特征；第2层到第19层为中间卷积层，每层使用3×3卷积核保持64通道特征不变；第20层为最终重建层，使用3×3卷积核将64通道特征重建为3通道输出；最后通过全局残差连接，将原始输入图像与学习到的残差相加得到增强结果，参数量约为665 K。

在图像复原过程中，算法采用自适应去噪策略，即首先进行全局噪声抑制，继而转向局部细节修复，最终完成边缘结构优化，形成层次分明的处理流程。对一幅图像在各时间步的处理效果如图3所示，图中第1行呈现了复原效果的演进过程：前两个时间步（$t$ = 1和$t$ = 2）图像质量显著提升，后续步骤（$t$ = 3，$t$ = 4和$t$ = 5）则进入优化稳定阶段。第2行中，以目标图像作为视觉参考，其余图像则展示了每个时间步执行的具体操作。具体而言，在初始阶段（$t$ = 1），面对高强度噪声污染，智能体优先选择强滤波（如高斯滤波、双边滤波以及引导滤波）以获得最大奖励。随着处理过程的推进（$t$ = 2和$t$ = 3），由于连续强滤波会导致图像模糊，因此智能体自适应地调整策略，在平滑区域采用像素级微调，同时在边缘区域切换至引导滤波和中值滤波以保持结构清晰度。在最终阶段（$t$ = 4和$t$ = 5），图像整体呈现稳定的蓝紫色调，表明噪声已基本消除，此时智能体的操作与奖励反馈达到动态平衡，重建过程逐步收敛。

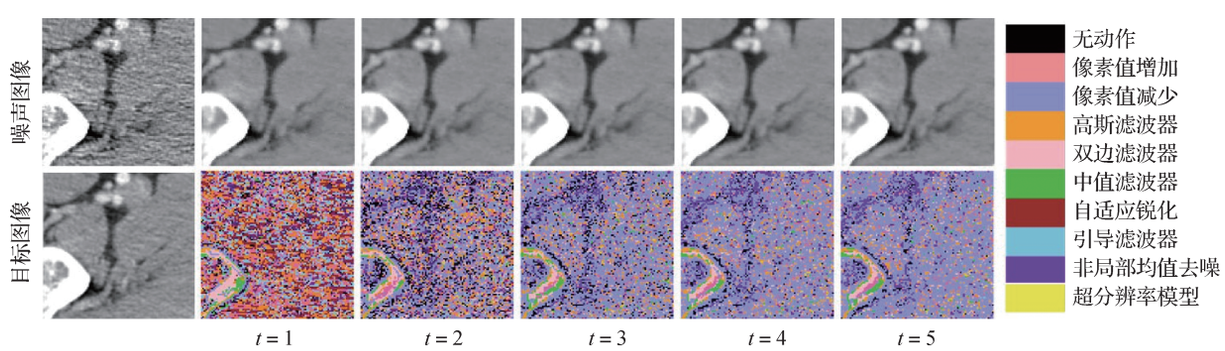



图3　CT图像的逐步恢复过程

Fig. 3　Gradual recovery process of CT images

### 1.5　算法流程及收敛性分析

HRDRL-Net算法将初始状态$s_0$设置为经Anscombe变换后的待去噪图像$\boldsymbol{I}_{\mathrm{LDCT}}$。$\boldsymbol{I}_{\mathrm{LDCT}}$送入像素层智能体后，首先通过SDCFE模块进行特征化，其中第一个通道保留原始图像，其余通道存放多尺度特征以及上一时间步的隐藏状态，用于捕获结构与细节线索。在时间步$t$，对于第$i$个像素位置，智能体根据局部状态$s_{i,t}$（包括当前像素、邻域特征、隐藏状态等）从动作集中采样动作$a_{i,t}$。环境状态根据动作类型执行相应的算子，包括像素平移、各类滤波、自适应锐化、卷积增强、残差增强等。只在选择该动作的像素处写回处理结果，得到下一状态$s_{i,t+1}$。同时，利用参考图像计算逐像素奖励$r_{i,t}$。若尚未达到终止条件，则继续执行状态→动作→奖励→更新的循环；若已达到最大步数$T$，则输出最终状态$s_T$，再经Anscombe逆变换得到高分辨率图像。

HRDRL-Net算法流程如下：

输入：待去噪图像$\boldsymbol{I}_{\mathrm{LDCT}}$，总训练回合数$P$，每回合最大时间步$T$，学习率$\alpha$，折扣因子$\gamma$，奖励权重$\lambda_1$、$\lambda_2$和$\lambda_3$。

输出：策略网络参数$\theta_\pi$，价值网络参数$\theta_v$。

1）for $p$ = 1, 2, …, $P$ do

//迭代循环$P$个训练回合。

2）$s_0 \leftarrow \boldsymbol{I}_{\mathrm{LDCT}}$//使用Anscombe变换和窗宽窗位

初始化图像。

3）for $p = 1, 2, \cdots, T$ do

//每个回合循环$T$个时间步。

4）根据状态$s_{i,t}$，进入智能体后进行SDCFE特征化，使用策略网络生成状态转移概率$\pi(a_{i,t}|s_{i,t};\ \theta_\pi)$。

5）采样动作$a_{i,t}$并执行，得到下一个状态$s_{i,t+1}$和$r_{i,t+1}$。

6）将$(s_{i,t}, a_{i,t}, s_{i,t+1}, r_{i,t+1})$追加到缓冲中。

7）end for //该回合结束

8）for $t = T - 1, \cdots, 0$ do

//反向计算累积回报与优势函数。

9）计算折扣累积汇报$R_t \leftarrow r_{t+1} + \gamma \cdot V(s_{t+1};\ \theta_v)$。

10）计算优势函数后令$A_t \leftarrow R_t - V(s_t;\ \theta_v)$。

11）累积策略梯度$\mathrm{d}\theta_\pi \leftarrow \mathrm{d}\theta_\pi + \nabla_{\theta_\pi} \log \pi(a_t|s_t;\ \theta_\pi) \cdot A_t$和累积价值梯度$\mathrm{d}\theta_v \leftarrow \mathrm{d}\theta_v + \nabla_{\theta_v}(R_t - V(s_t;\ \theta_v))^2$。

12）end for

//梯度累计结束。

13）使用$d\theta_\pi$和$d\theta_v$更新策略网络和价值网络参数。

14）$\theta_\pi \leftarrow \theta_\pi + \alpha \cdot \mathrm{d}\theta_\pi, \theta_v \leftarrow \theta_v + \alpha \cdot \mathrm{d}\theta_v$。

15）end for

//所有训练回合结束。

本文算法满足马尔可夫决策过程的基本假设，状态转移满足马尔可夫性，奖励函数有界。通过广义优势估计降低策略梯度方差，使用异步并行采样提升数据多样性。理论分析表明，在折扣因子$\gamma \in (0, 1)$条件下，算法能收敛至局部最优策略。

## 2　实验设置与数据准备

### 2.1　数据集

本文分别采用Mayo Clinic LDCT挑战赛数据集（https://www. aapm. org/GrandChallenge/LowDoseCT/）和Piglet数据集（Li等，2023）进行方法验证。两个数据集均提供精确配对的LDCT与NDCT图像，确保了监督学习的可靠性。

Mayo数据集是评估CT算法性能的国际金标准，其遵循临床标准扫描协议采集，包含多例患者的腹部CT扫描数据，每例均提供精确时空配准的两种剂量图像：NDCT（100%剂量水平）作为参考真值，以及通过仿真技术获得的25%剂量水平的LDCT图像作为样本的输入。数据集纳入了不同体型特征（BMI（body mass index）范围18. 5～32. 4）、年龄分布（22～76岁）和性别比例的受试者，完整覆盖了肝脏、肾脏、脾脏等主要腹部器官。原始图像尺寸为512 × 512像素，层厚3 mm。

为了验证HRDRL-Net在数据稀缺情况下的有效性，本文采用Mayo数据集进行了不同训练集规模的对比实验，结果如表2所示。

**表2　训练集规模不同时HRDRL-Net的性能对比（均值±标准差）**

**Table 2　Performance comparison of HRDRL-Net with varying training set sizes（mean±standard deviation）**

| 患者数量 | 图像对数 | 测试集PSNR/dB | 测试集SSIM |
|---|---|---|---|
| 9 | 2 135 | 32.319 9±1.016 8 | 0.897 7±0.020 6 |
| 7 | 1 691 | 32.318 7±1.015 5 | 0.897 2±0.020 4 |
| 5 | 1 134 | 32.319 4±1.014 9 | 0.897 4±0.019 8 |
| 3 | 766 | 32.319 0±1.016 6 | 0.896 9±0.020 1 |
| 2 | 553 | 32.319 4±1.015 8 | 0.897 2±0.020 2 |
| 1 | 224 | 32.092 1±1.017 0 | 0.890 3±0.021 0 |

实验表明，训练集仅包括2名患者（553对图像）时，HRDRL-Net在测试集上的性能已趋于稳定，与使用全部9名患者（2 135对图像）训练的性能无显著差异。基于此，为了充分证明HRDRL-Net的高数据利用效率和泛化能力，从数据集中随机选取2名患者共553对图像作为HRDRL-Net的专用训练集。同时，为确保所有对比的基线模型在最优条件下进行训练，且确保对比的严格公平性，使用完整的9名患者数据（共2 135对图像）作为基线模型的训练集。所有方法均在完全相同的独立测试集（由另外的病例组成，共243对图像）上进行评估，保证了训练集与测试集在患者级别上的完全隔离，杜绝了数据泄露。图像预处理包括：将原始CT值转换为Hounsfield单位（HU），并采用临床诊断中为优化腹部软组织对比度而常用的窗宽窗位设置（窗宽160 HU，窗位240 HU）进行数值截断，随后归一化至[0，1]区间。

Piglet数据集是采用Discovery CT750 HD（GE

Healthcare, Chicago, IL, USA)高清螺旋CT系统对仔猪样本胸部进行扫描构建的专业数据集。其中管电流曝光量为75 mAs的LDCT图像和300 mAs的NDCT图像形成严格配对的实验数据。所有图像均经过专业放射科医师的质量控制,确保解剖结构的清晰可辨和剂量信息的准确标注。随机选取了758对图像对作为训练集,并单独设置153对图像作为独立测试集。预处理流程与Mayo数据集保持一致,采用肺部窗(窗宽1 500 HU,窗位-600 HU)进行处理。该数据集用于评估方法在跨中心、不同扫描协议下的泛化能力。

### 2.2 网络训练策略与动态分析

网络训练采用A3C框架,其总损失函数由策略损失$\mathcal{L}_{\text{policy}}$与值函数损失$\mathcal{L}_{\text{value}}$加权求和构成,具体为

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{policy}} + \eta\mathcal{L}_{\text{value}} \tag{11}$$

式中,$\eta$为平衡两项损失的权重因子,其具体数值通过交叉验证确定。

策略损失基于策略梯度理论构建,旨在最大化期望累积奖励,其定义为

$$\mathcal{L}_{\text{policy}} = -\frac{1}{N}\sum_{i=1}^{N}\{[\log\pi(a_{i,t}|s_{i,t};\theta_{\pi})]A(a_{i,t},s_{i,t})\} \tag{12}$$

式中,$N$为图像总像素数,$A(a_{i,t},s_{i,t})$是优势函数,用于衡量当前动作相对于平均水平的优势程度。该损失函数利用优势函数对每个像素位置的动作选择概率进行加权,确保网络能够精细地学习并优化图像的局部特征处理策略。

值网络的优化目标是最小化状态价值的估计误差,为策略更新提供稳定的基线,其损失函数定义为

$$\mathcal{L}_{\text{value}} = \frac{1}{2N}\sum_{i=1}^{N}[A(a_{i,t},s_{i,t})]^2 \tag{13}$$

优势函数采用广义优势估计(generalized advantage estimation, GAE)(Schulman等,2017)进行计算,其中折扣因子$\gamma = 0.99$,平滑参数$\lambda = 0.95$,以平衡策略梯度的估计偏差与方差。

策略损失与值函数损失通过反向传播进行协同优化:策略损失驱动智能体优先选择高期望回报的动作,而值函数损失则提高状态价值估计的精度。这种双支路优化机制确保网络在充分探索动作空间的同时,能依赖准确的价值评估进行决策,从而达成噪声抑制与结构细节保留的最佳平衡。

训练超参数设置如下:采用Adam优化器,其自适应学习率特性适合处理策略梯度中的非平稳问题。初始学习率设为0.001,该值基于常见深度学习任务的经验设定,并经初步实验验证,能在训练稳定性和收敛速度间取得平衡。批次大小设置为64,这是考虑GPU内存容量与训练效率后的典型选择。训练共进行30 000轮,每轮包含5个时间步。此轮数设定是为了确保模型充分收敛,并通过验证集监控确认了其必要性。学习率采用指数衰减(衰减因子0.9),以在训练后期实现精细调优。数据增强采用随机剪裁,剪裁块尺寸为64×64像素,滑动间隔为64像素,以增加样本多样性和防止过拟合。所有超参数均基于训练过程中的奖励收敛趋势与模型在训练集上的性能表现,通过网格搜索与手动调优相结合的方式确定。

训练过程的动态可通过奖励函数曲线进行观察。图4展示了在所述超参数下,模型在5次独立训练中的奖励函数变化曲线,图中实线为5次训练的平均值,灰色填充区域为对应的标准差。可以看出,模型性能在前5 000轮内迅速提升,奖励值达到约320;随后在约12 500轮内,奖励值在318~325的区间内波动并趋于稳定,表明模型已收敛至一个高性能的稳定策略。

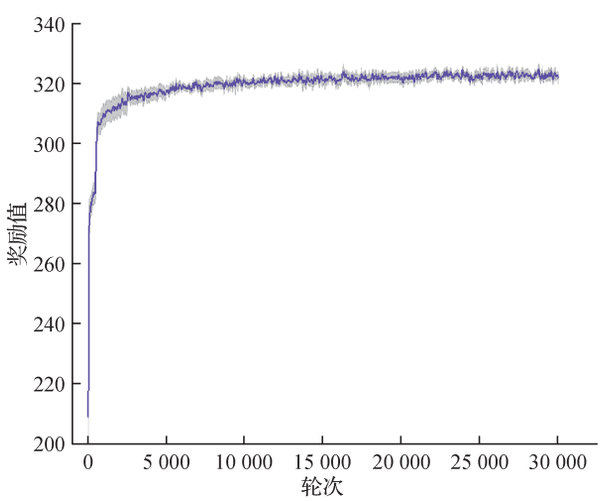



图4 HRDRL-Net在5次独立训练中的奖励平均值和标准差

Fig. 4 Average rewards and standard deviations of HRDRL-Net after five indepenclent training sessions

### 2.3 性能评价方法

训练完成后,为全面评估HRDRL-Net,设计了多维度对比实验。

#### 2.3.1 基线方法与比较实验设计

本文选取了5种基于LDCT图像到高质量图像

直接映射的代表性深度学习模型作为基线进行对比，其任务目标与HRDRL-Net一致，均为从LDCT图像重建出高质量图像，具有可比性。包括：RED-CNN（Chen 等，2017）、CT-former（Wang 等，2023）、SWELNet（Wang 等，2025）、MMCA（Li 等，2025）和DU-GAN（Huang 等，2022）。其中，RED-CNN是代表性残差网络，采用残差连接和跳跃连接构建编码器—解码器结构；DU-GAN是代表性GAN模型，通过双判别器网络同时学习全局结构与局部细节的差异；CT-former基于Transformer架构，利用Token2Token机制和膨胀移位操作以捕获长程依赖关系；SWELNet是轻量级半监督学习网络，结合了多尺度卷积特征提取模块、递归模块以及自适应权重学习的softmax特征融合模块；MMCA是代表性多尺度上下文聚合网络。所有对比方法均使用预处理后的Mayo数据集进行训练与评估（如2.1节所述），其实现均基于原文公开代码，训练策略与超参数设置严格遵循原始论文设置或论文推荐值，如表3所示，以确保对比的公平性。

### 2.3.2 消融实验设计

为验证HRDRL-Net架构设计的合理性并分析核心组件的有效性，分别设计了系统级和组件级消融实验。系统级实验评估核心模块组合的有效性，组件级实验则分析共享卷积模块设计的合理性，具体设计如表4所示。每个消融设置均保持其他部分与完整模型一致。

### 2.3.3 评价指标

采用峰值信噪比（peak signal-to-noise ratio，PSNR）、结构相似性指数（structural similarity index measure，SSIM）和梯度幅度相似性偏差（gradient magnitude similarity deviation，GMSD）评估整体重建质量。PSNR通过计算图像间的均方误差衡量重建保真度，其值越高，表明重建图像的失真越小。SSIM从亮度、对比度和结构3个维度评估图像间的相似性，其取值范围为[0，1]，值越大，表示重建图像与标准图像在视觉结构上越相似。GMSD通过衡

**表3 基线方法原理及训练参数设置**

**Table 3 Principle of baseline methods and the setting of training parameters**

| 方法 | 网络结构 | 训练策略及超参数 | | | | | |
|---|---|---|---|---|---|---|---|
| | | 损失 | 优化器 | 学习率 | 批次 | 轮数 | Mayo受试数 |
| RED-CNN | 由自编码器、反卷积网络和跳跃连接组成的残差编码器—解码器CNN | MSE | Adam | $10^{-5}$ | 16 | 100 | 9 |
| CT-former | 采用无卷积的编码器—解码器结构，核心包含Token2Token膨胀模块和Transformer块 | MSE | Adam | $10^{-5}$ | 16 | 4 000 | 9 |
| SWELNet | 由包含监督分支和无监督分支的学生模型以及通过指数移动平均更新的教师模型构成，包括多尺度卷积特征提取模块、递归模块和特征融合模块 | 复合损失（CBLoss + SSIM + Edge + Perceptual + TV + Teacher） | AdamW | $10^{-5}$ | 2 | 200 | 24（8标注 + 16无标注） |
| MMCA | 由5个阶段组成的纯卷积网络，每个阶段包含特征嵌入模块和多个堆叠的多阶上下文聚合模块，并通过融合边缘增强特征保持高分辨率表征 | MSE + 多尺度感知损失 | AdamW | $2 \times 10^{-4}$ | 8 | 200 | 9 |
| DU-GAN | 由一个去噪生成器和两个U-Net判别器构成，其中一个判别器在图像域工作，另一个在梯度域工作，共同约束生成器输出兼具全局结构保真度和局部细节清晰度的去噪图像 | 像素损失 + 梯度损失 + 梯度生成损失 + 图像生成损失 + 一致性正则化损失 | Adam | 生成器：$1.81 \times 10^{-4}$<br>判别器：$7.12 \times 10^{-5}$ | 64 | 105 | 20 |

表4　消融实验设计细节

Table 4　Details of the ablation experiment design

| 实验类别 | 消融项目 | | 说明 | 实验目的 |
|---|---|---|---|---|
| 系统级 | 奖励函数 | 像素奖励 | 以当前图像与目标图像之间的均方误差作为奖励函数 | 评估多目标奖励设计对提高网络收敛性和整体性能的作用 |
| | | 移除结构相似度奖励 | 仅保留像素误差和局部方差惩罚，分析梯度奖励对模型性能的影响 | |
| | | 移除方差惩罚 | 仅保留像素误差和梯度相似性奖励，分析方差惩罚对模型性能的影响 | |
| | | 复合奖励 | 由像素级奖励、梯度相似性奖励和局部方差的惩罚项 | |
| | 动作集 | 基础动作 | 高斯滤波、中值滤波、像素增减 | 验证各动作子集的有效性 |
| | | 移除网络增强 | 增加双边滤波、引导滤波、自适应锐化 | |
| | | 完整动作集 | 增加非局部均值滤波、卷积增强、残差增强 | |
| 组件级 | 双路径并行处理 | 仅保留上路径 | 移除下路径，保留上路径 | 验证结构—细节协同特征提取模块中双路径特征融合架构对图像特征提取的有效性 |
| | | 仅保留下路径 | 移除上路径，保留下路径 | |
| | | 双路径 | 完整双路径处理 | |
| | LDNM模块的四分支结构 | 仅保留膨胀卷积 | 仅以膨胀卷积作为基础特征提取 | 验证结构—细节协同特征提取模块中LDNM模块四分支结构设计的合理性 |
| | | 仅移除身份映射 | 提供原样基线 | |
| | | 仅移除细节增强 | 恢复细小结构与轮廓清晰度 | |
| | | 仅移除噪声抑制 | 平滑噪声的随机扰动 | |
| | | 完整的四分支 | 完整的四分支特征提取 | |
| | 双重注意力机制 | 移除注意力 | 移除双重注意力模块 | 评估结构—细节协同特征提取模块上分支中双重注意力机制对特征选择和信息增强的影响 |
| | | 仅保留ECA | 移除SA，仅保留ECA | |
| | | 仅保留SA | 移除ECA，仅保留SA | |
| | | 保留ECA和SA | 包含ECA和SA的双重注意力 | |
| | 下路径中的边缘提取算子 | Sobel | 一阶梯度核 | 验证重建图像质量对下路径中边缘提取算子的敏感性 |

量图像梯度幅值的相似性来反映感知质量，其值越低表示重建质量越高（檀结庆 等，2021）。

采用边缘斜率（edge slope，ES）和边缘保持指数（edge preservation index，EPI）量化细节保持能力。其中ES通过计算图像中选定边缘区域的灰度变化率，定量评价边缘的锐利程度。其计算方式为：在跨越边缘的剖面上，选取灰度值从10%到90%的上升区间，通过线性拟合得到斜率值。ES值越大，表明边缘越锐利，结构保持能力越强。EPI用于量化去噪前后边缘结构的保持程度，定义为去噪图像与参考图像在边缘区域梯度幅值的相关系数，其取值范围为[0, 1]，越接近1表示边缘保持能力越优。

此外，为了验证HRDRL-Net与基线方法性能差异的统计显著性，采用非参数的Friedman检验分别对Mayo和Piglet数据集中50个样本的PSNR、SSIM和GMSD指标进行全局显著性分析，显著性水平$\alpha$ = 0.05。当Friedman检验拒绝原假设（即$p$值小于0.05）时，进一步使用Nemenyi事后检验进行两两比

较，分析具体的组间差异，并报告其临界差（critical difference，CD）值。

2.3.4　实验环境配置

所有实验在配备Intel Core i5-8250U处理器、8 GB内存及NVIDIA Tesla P100 GPU（16 GB显存）的计算机平台上完成，操作系统为Windows 11 64位。软件环境以Python 3.8.10作为开发语言，依托PyTorch 1.8.1深度学习框架，并基于CUDA 11.3与cuDNN 8.2.0实现GPU加速计算。

## 3　实验结果与分析

### 3.1　Mayo数据集实验结果

图5—图6展示了对Mayo测试集中随机选取的5个样本的重建结果对比，包括重建图像（图5）、残差图（图6(a)）及边缘强度曲线（图6(b)）。其中，重建图像中每个样本的第2行为框选区域的放大显示，边缘强度曲线是沿该区域水平中心线提取的。残差图反映了重建图像与NDCT参考图像之间的像素级差异，颜色越浅表示差异越小。边缘强度曲线则用于量化评估边缘锐利度与结构连续性，与NDCT曲线越接近表明保真度越高。

观察可知，原始LDCT图像因辐射剂量降低导致严重的量子噪声和条形伪影，其残差图整体色深，边缘强度曲线与NDCT曲线差异较大。RED-CNN和DU-GAN在低对比度区域（圆圈标注）残留明显噪声与伪影，残差图中对应区域颜色较深，边缘强度曲线波动显著。CT-former和MMCA虽然能抑制大部分噪声，但存在过度平滑效应，导致边缘强度曲线出现整体偏移与峰值衰减。SWELNet重建结果整体残差较小，但局部仍存在零星噪声斑点，对应边缘强度曲线存在局部断点。相比之下，HRDRL-Net在噪声抑制与细节保留方面取得了更好的平衡。其重建图像在血管—肝脏交界等精细结构处边缘清晰、纹理完整，残差图整体色浅且均匀，边缘强度曲线与NDCT参考曲线重合度最高，尤其在血管—肝脏交界等关键区域的曲线梯度保持最为完整。综合视觉结果与量化曲线可知，HRDRL-Net凭借其复合奖励函数对结构相似性的约束以及自适应动作选择机制，在抑制噪声的同时最大限度地保持了边缘锐度与纹理连续性，表现出最优的结构保真度。

图7和表5的定量评价结果表明，不同网络架构在LDCT图像重建任务中展现出明显的性能差异。

作为早期深度学习方法代表，RED-CNN采用编码器—解码器结构并引入短路连接，通过残差学习实现初步去噪（PSNR较LDCT图像提高约3 dB），但其浅层网络结构限制了复杂噪声模式的建模能力，尤其在处理非均匀分布的量子噪声时表现欠佳，这体现在其相对较低的边缘保持能力上。DU-GAN通过U-Net结构的判别器在图像域和梯度域进行联合优化，双域训练机制使其在保持解剖结构连续性方面表现突出，其ES优于RED-CNN，表明对抗训练有助于生成更锐利的轮廓，然而GAN固有的训练不稳定性导致模型容易出现模式崩溃，影响重建结果的可靠性。CT-former无卷积的Token2Token设计通过长程依赖建模显著提升了特征表示能力，在边缘锐化方面表现较好，说明Transformer架构在捕获全局结构关系上的优势，但自注意力机制带来的平方级计算复杂度导致推理效率降低。SWELNet采用半监督学习范式，通过自适应权重机制平衡有标签和无标签数据的贡献，在数据充足时展现出优异的泛化能力，但其性能对无标签数据质量具有较强依赖性。MMCA通过阶段多阶上下文聚合模块，在高分辨率特征图上实现局部细节与全局上下文的协同优化，其ES达到较高水平，体现了多尺度特征融合对边缘增强的积极作用，但其级联式网络结构导致参数量大幅增加。

相比之下，HRDRL-Net通过融合多尺度双重注意力机制和深度强化学习框架，在多个维度实现突破：多尺度特征提取模块有效捕获不同频段的噪声特征；双重注意力机制协同优化空间和通道维度的重要特征；强化学习的动态决策机制实现像素级自适应处理。这种综合优势使其不仅在传统指标（SSIM、PSNR和GMSD）上领先，更在边缘量化指标上表现突出：ES显著高于其他方法，表明重建图像具有最锐利的边缘过渡；EPI达到最高值，证明该方法在去除噪声的同时最大限度地保留了原始解剖结构。这种优势在血管分支等精细结构的保持方面尤为明显，验证了强化学习框架通过自适应动作选择实现噪声抑制与细节保护动态平衡的有效性。此外，HRDRL-Net仅使用2例受试者的数据进行训练，在训练数据量显著少于基线方法的条件下，不仅在整体性能上超越现有方法，同时保证了相对稳定的指标分布，证明了该方法在性能和稳定性之间取得

了良好平衡。

图 8 的箱线图分析结果显示，HRDRL-Net 的 SSIM 和 PSNR 不仅整体处于最高区间（SSIM 中位数 0.895，PSNR 中位数 32.4 dB），其四分位距（SSIM IQR = 0.027 9，PSNR IQR = 0.979 3 dB）也显著小于其他方法，表明该方法在不同病例上的表现最为稳定。这种稳定的性能分布特性与 RED-CNN（PSNR IQR = 1.02 dB）和 SWELNet（PSNR IQR = 1.05 dB）等方法的较大波动形成鲜明对比。

为了全面评估各方法的计算效率，分别记录了 6 种方法的每秒浮点运算次数（floating point operations per second，FLOPs）、GPU 运行时间和 GPU 内

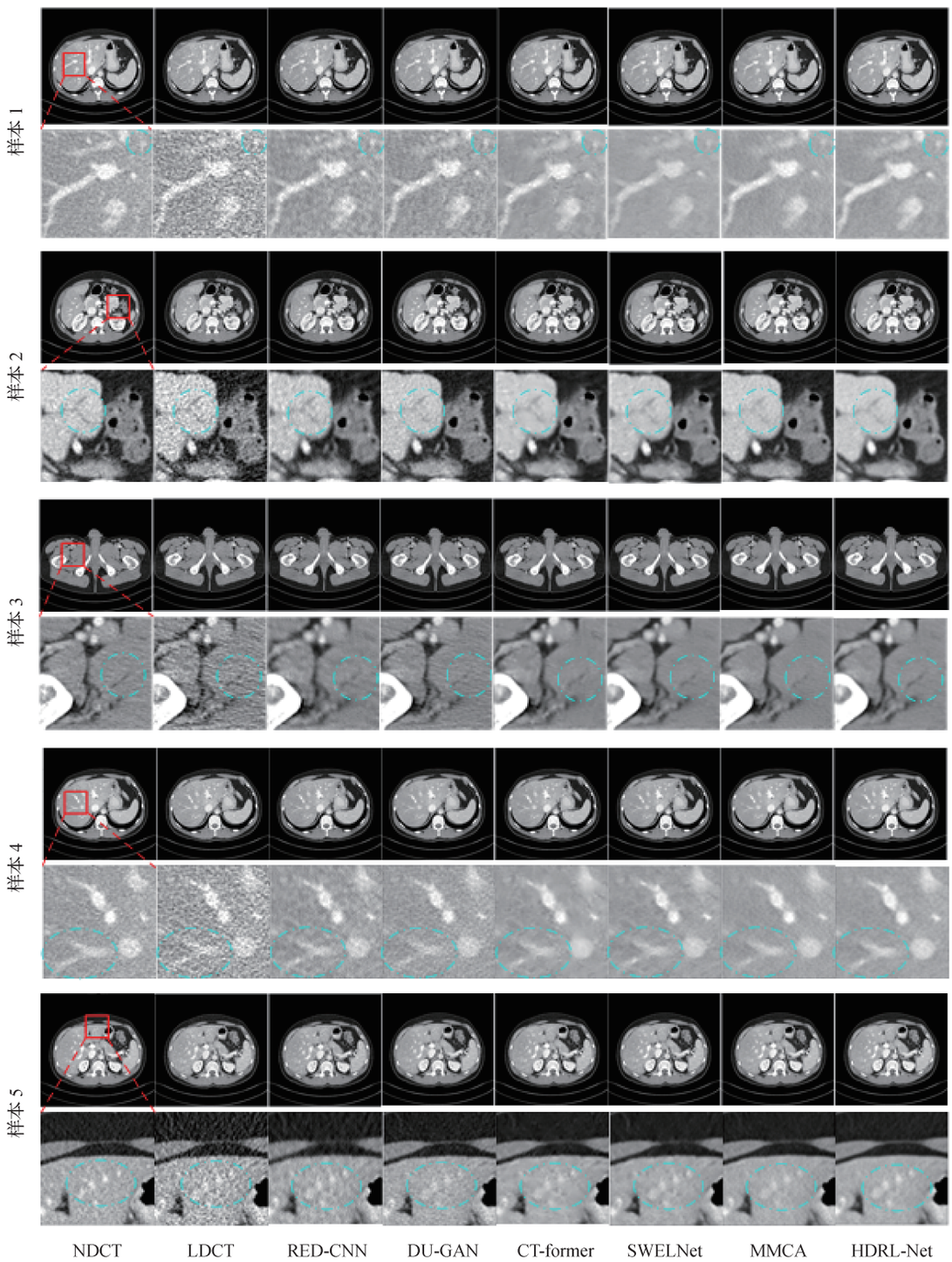



图 5　Mayo 测试集中随机选取样本的重建图像

Fig. 5　Reconstructed images of randomly selected samples from the Mayo testset

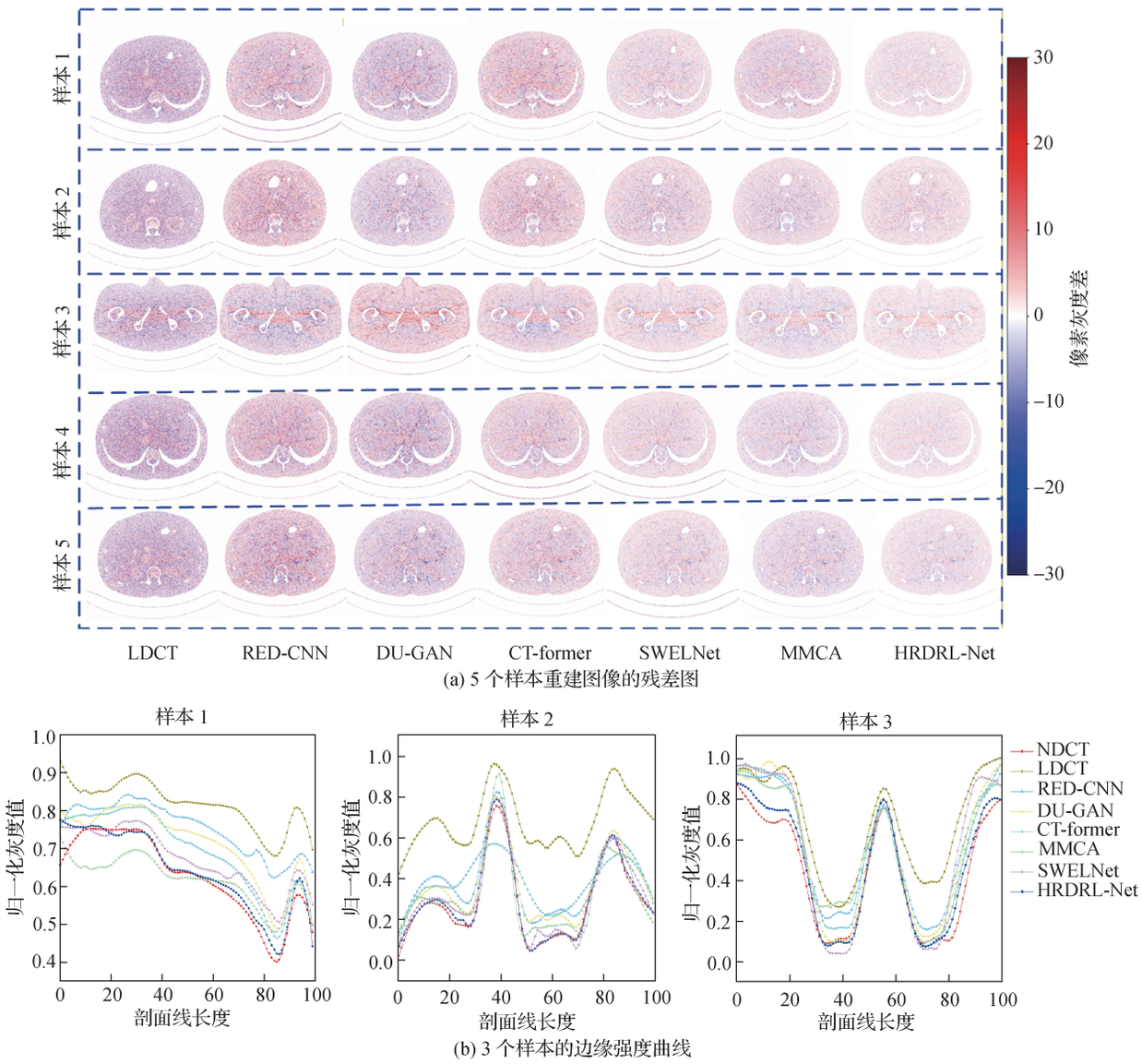



图 6　对 Mayo 测试集中随机选取样本的图像重建结果

Fig. 6　Image reconstruction results for randomly selected samples from the Mayo testset ((a) residual maps of the reconstructed images for five samples; (b) edge intensity profiles of three samples)

存消耗。如图 8 所示，HRDRL-Net 在各个方面均实现了显著优化，相比于 RED-CNN、DU-GAN、CT-former 等基线方法，在计算效率与图像重建质量方面取得了较好的平衡，能够在保持适度计算复杂度的同时，提供较快的处理速度和合理的内存使用。

### 3.2　Piglet 数据集实验结果

图 9 展示了随机选取的 Piglet 测试集样本的重建结果。在方框和箭头标注的感兴趣区域，HRDRL-Net 展现出优异的细节保持能力。原始图像中，感兴趣区域内的解剖结构因严重噪声干扰而几乎无法辨识。RED-CNN、DU-GAN、SWELNet 和 MMCA 虽然能够有效抑制噪声，但在处理过程中损失了大量细节信息，导致输出图像出现明显的模糊。这种现象主要源于这些方法缺乏有效的细节保护机制，无法在去噪过程中准确区分噪声与真实解剖结构。CT-former 通过 Transformer 架构的长程依赖建模能力，在细节保留和边缘增强方面表现相对较好，但仍存在细微结构恢复不完整的问题。

表 5 和图 10 的指标统计结果表明，HRDRL-Net 在 PSNR、SSIM 和 GMSD 3 项指标上均显著领先于其他方法，同时，其 SSIM 和 PSNR 整体处于最高区间（SSIM 中位数 0.902 5，PSNR 中位数 32.228 3 dB）。在衡量结构细节保持能力的关键指标上，HRDRL-Net 同样展现出卓越性能：其 ES 与 EPI 均显著优于

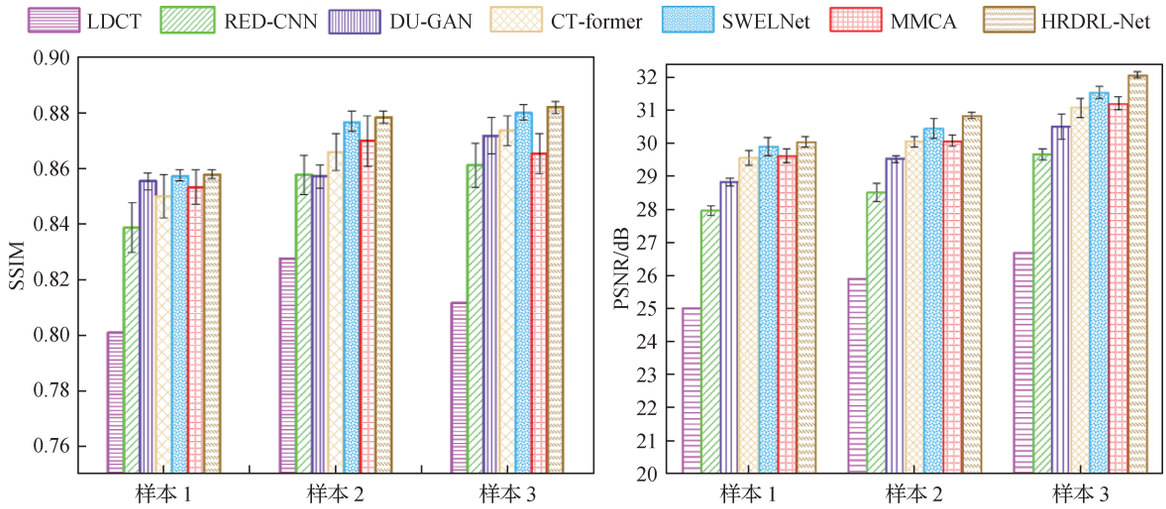



图7　对Mayo测试集中随机选取的3个样本图像重建结果的评价指标

Fig. 7　Evaluation metrics for image reconstruction results of three randomly selected samples from the Mayo testset

**表5　Mayo和Piglet测试集中全部样本图像重建结果的平均指标值(均值±标准差)**

**Table 5　Average metrics (mean ± standard deviation) of all sample image reconstruction results in the Mayo and Piglet testsets**

| 方法 | SSIM | | PSNR/dB | | GMSD | |
|---|---|---|---|---|---|---|
| | Mayo | Piglet | Mayo | Piglet | Mayo | Piglet |
| LDCT | 0.839 1±0.030 4 | 0.857 9±0.014 4 | 27.244 2±1.179 4 | 29.573 8±1.183 9 | 0.107 0±0.014 5 | 0.110 0±0.023 8 |
| RED-CNN | 0.877 1±0.022 7 | 0.889 1±0.009 6 | 30.237 4±1.016 6 | 31.229 4±1.233 9 | 0.071 2±0.008 3 | 0.094 9±0.020 7 |
| DU-GAN | 0.890 3±0.022 0 | 0.891 4±0.010 4 | 31.300 6±1.099 3 | 31.601 5±1.128 6 | 0.068 4±0.008 5 | 0.095 8±0.021 1 |
| CT-former | 0.889 1±0.022 2 | 0.896 1±0.009 7 | 31.511 7±0.960 2 | 31.780 5±1.159 2 | 0.067 3±0.008 1 | 0.094 1±0.021 6 |
| SWELNet | 0.894 0±0.021 2 | 0.884 9±0.010 0 | 32.174 5±1.066 9 | 31.108 1±1.126 4 | 0.068 3±0.008 0 | 0.094 7±0.021 8 |
| MMCA | 0.892 6±0.023 8 | 0.893 0±0.009 5 | 31.736 9±1.096 2 | 31.678 2±1.180 6 | 0.066 5±0.009 0 | 0.094 3±0.020 8 |
| HRDRL-Net | **0.897 2±0.020 2** | **0.899 6±0.009 1** | **32.319 4±1.015 8** | **32.262 5±1.158 2** | **0.062 7±0.007 8** | **0.093 8±0.020 6** |

| 方法 | ES (灰度/像素) | | EPI | |
|---|---|---|---|---|
| | Mayo | Piglet | Mayo | Piglet |
| LDCT | 1.12±0.25 | 1.28±0.21 | 0.712±0.045 | 0.748±0.038 |
| RED-CNN | 1.87±0.18 | 1.95±0.16 | 0.845±0.032 | 0.857±0.028 |
| DU-GAN | 2.05±0.19 | 2.10±0.17 | 0.868±0.030 | 0.875±0.026 |
| CT-former | 2.11±0.17 | 2.18±0.15 | 0.873±0.029 | 0.882±0.025 |
| SWELNet | 2.15±0.18 | 2.01±0.16 | 0.880±0.031 | 0.866±0.027 |
| MMCA | 2.18±0.20 | 2.12±0.18 | 0.882±0.033 | 0.878±0.029 |
| HRDRL-Net | **2.34±0.16** | **2.30±0.14** | **0.902±0.028** | **0.896±0.024** |

注:加粗字体表示各列最优结果。

对比方法。这一优势在感兴趣区域尤为明显。具体而言,相较于性能次优的CT-former,HRDRL-Net在保持血管壁厚度、肺泡间隔等亚毫米级结构方面展现出更强的能力,其更高的ES与EPI值分别从边缘锐利度与结构保真度两个维度提供了量化证明。

这种跨数据集的性能优势,源于HRDRL-Net方法设计的固有特性:其多尺度特征融合机制能够同时建模噪声的局部统计特性与组织的全局结构上下文,为精准区分噪声与真实边缘提供了基础;而其基于强化学习的自适应细节增强策略,则能根据局部

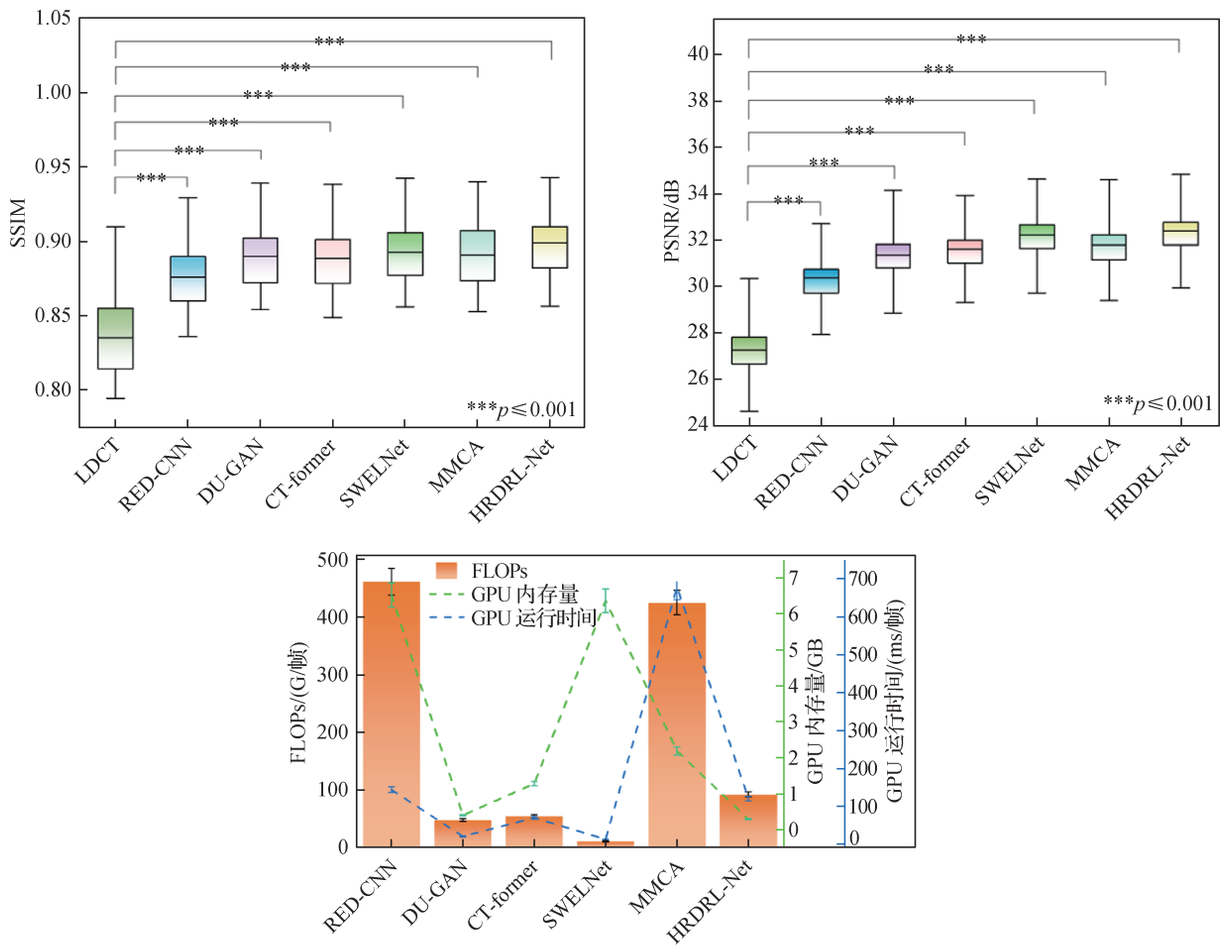



图 8　对 Mayo 测试集中全部样本进行图像重建的评价指标统计结果

Fig. 8　The statistical results of the evaluation indicators for image reconstruction of all samples in the Mayo testset

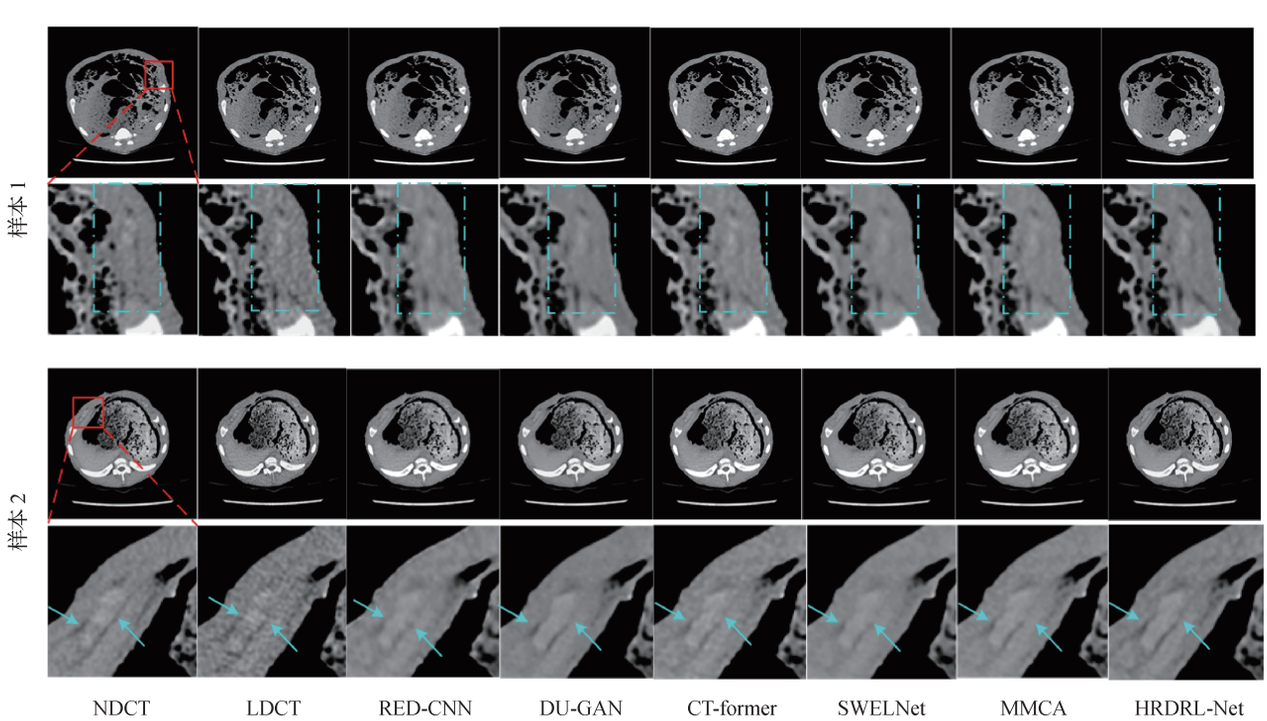



图 9　对 Piglet 测试集中随机选取的两个样本进行图像重建的结果

Fig. 9　Image reconstruction results for two randomly selected samples from the Piglet testset

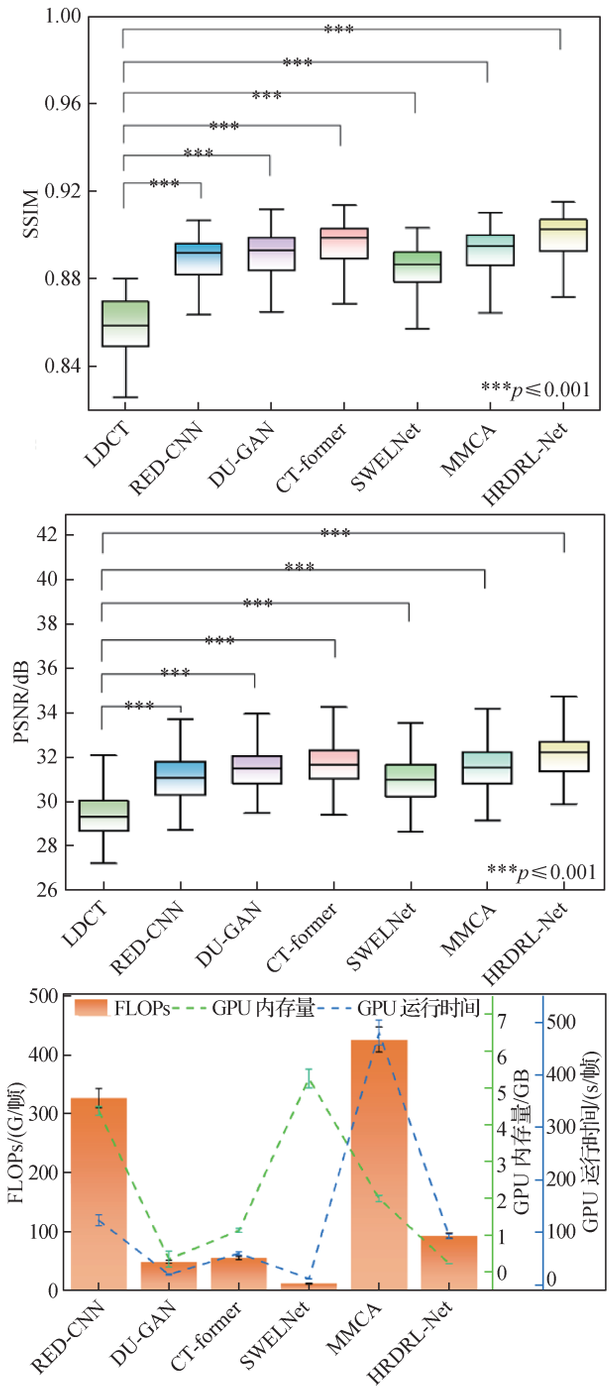



图 10　对 Piglet 测试集中全部样本进行图像重建的评价指标统计结果

Fig. 10　Statistical results of the evaluation indicators for image reconstruction of all samples in the Piglet testset

图像特征动态调整处理强度，避免了传统方法在均匀区域欠平滑或在复杂边缘处过平滑的固有问题。因此，该方法能够在抑制均匀区域噪声（高 PSNR 与低 GMSD）的同时，有效保护并增强对诊断重要的细微结构（高 ES 与 EPI），从而实现了噪声抑制与细节保留的良好平衡。

### 3.3　Friedman 检验结果

表 6 和表 7 所示 Friedman 检验结果表明，所有指标在两个数据集上的 $p$ 值均小于 0.001，说明不同方法之间至少存在一对具有统计学意义上的显著差异。在算法数量 $k = 6$、样本数 $N = 50$ 的条件下，查得 $q_\alpha = 2.850$，计算得到 CD = 1.066 4。图 11(a)的 CD 图显示，HRDRL-Net 在 GMSD 指标上的平均排名显著优于基线方法，表明其性能优势具有统计稳定性。图 11(b)的热力图进一步证实 HRDRL-Net 在所有数据集和指标中均保持最优排名，体现了方法的强泛化能力。

统计显著性结果表明，不同方法在 LDCT 图像重建任务上存在本质的性能差异。HRDRL-Net 的显著优势可归因于其核心设计：多尺度特征融合机制使模型能够同时建模局部噪声模式和全局结构特征，从而在 PSNR 和 SSIM 指标上获得稳定提升；而基于强化学习的动态决策机制则通过像素级自适应处

**表 6　不同数据集性能差异的 Friedman 检验结果**

**Table 6　Friedman test results for performance differences across different datasets**

| 指标 | Mayo | | Piglet | |
|---|---|---|---|---|
| | $\chi^2$ | $p$ 值 | $\chi^2$ | $p$ 值 |
| PSNR/dB | 200.28 | < 0.001 | 208.28 | < 0.001 |
| SSIM | 210.14 | < 0.001 | 92.28 | < 0.001 |
| GMSD | 145.28 | < 0.001 | 77.63 | < 0.001 |

**表 7　不同方法在 Mayo 和 Piglet 数据集上的平均排名（50 个样本，显著性水平 $\alpha = 0.05$）**

**Table 7　Average ranking of different methods on the Mayo and Piglet datasets (50 samples, significance level $\alpha = 0.05$)**

| 方法 | Mayo | | | Piglet | | |
|---|---|---|---|---|---|---|
| | PSNR/dB | SSIM | GMSD | PSNR/dB | SSIM | GMSD |
| HRDRL-Net | **1.1** | **1.3** | **1.2** | **1.0** | **1.0** | **1.0** |
| SWELNet | 2.1 | 2.1 | 3.9 | 5.9 | 5.7 | 3.5 |
| MMCA | 3.1 | 2.9 | 2.1 | 3.1 | 3.3 | 2.8 |
| CT-former | 3.8 | 4.9 | 3.1 | 2.1 | 2.5 | 2.1 |
| DU-GAN | 4.8 | 3.7 | 4.4 | 3.9 | 4.2 | 5.8 |
| RED-CNN | 5.9 | 6.0 | 5.8 | 4.8 | 4.9 | 4.3 |

注：加粗数据表示各列最优结果。

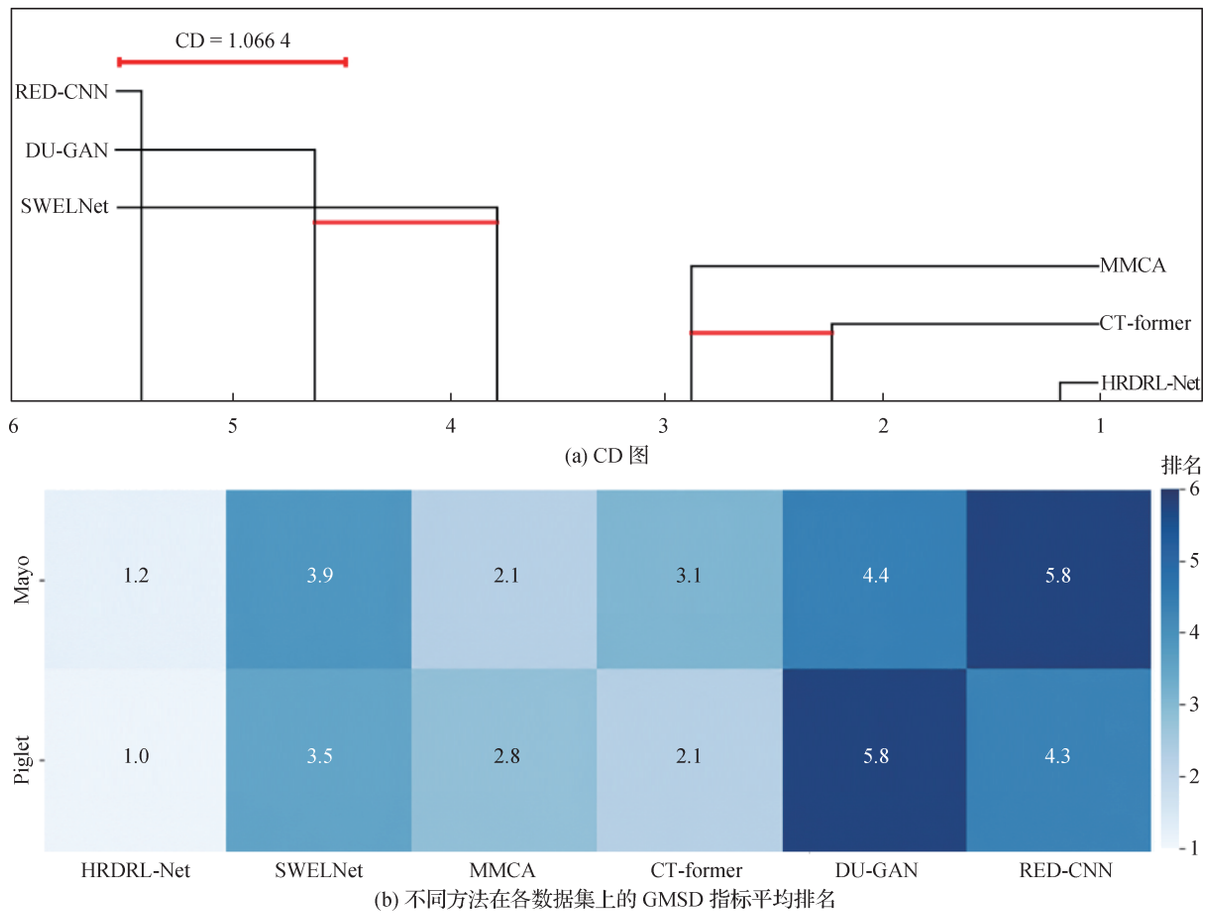



图 11　Friedman 检验结果

Fig. 11　Friedman test results ((a) CD plot; (b) average ranking of GMSD metrics for different methods on each dataset)

理，有效平衡了噪声抑制与细节保留的矛盾，这体现在 GMSD 指标的显著改善上。相比之下，传统方法采用静态前馈与单目标优化范式，缺乏协同优化机制，在统计比较中呈现出明显的性能分层。

### 3.4　消融实验结果

本节基于 Mayo 测试集的指标统计结果（表 8）和图像重建结果（图 12），从系统和组件两个层面分析 HRDRL-Net 架构的各关键设计对图像重建性能的影响。

#### 3.4.1　系统级消融实验结果

1）奖励函数设计。奖励函数实验中，完整复合奖励函数在所有评估指标上均达到最优，且测试时间适中。与仅使用像素奖励相比，复合奖励在结构保真度（SSIM 提升 0.012 9）、噪声抑制（PSNR 提升 0.686 6 dB）和感知质量（GMSD 降低 0.007 6）方面均有显著改善，表明其能有效缓解单纯像素损失导致的图像过度平滑与细节丢失问题，验证了其设计的有效性。移除结构相似度奖励导致 SSIM 下降至 0.888 3，PSNR 降低 0.372 9 dB，且 GMSD 增至 0.067 7，表明该组件对保持结构完整性与纹理细节至关重要。尽管测试时间缩短至 0.42 s，但性能下降明显。移除方差惩罚项后，性能介于像素奖励与完整奖励之间，GMSD 增至 0.063 1，表明其有助于提升图像清晰度、抑制局部噪声与伪影。视觉分析进一步证明，复合奖励函数在噪声抑制与细节保留方面达到最优平衡，恢复的图像更接近 NDCT 参考图像。

2）动作集配置。动作集实验中，完整动作集在所有指标上均优于其他配置。基础动作集虽然测试时间较短（0.46 s），但其 SSIM 和 PSNR 较低，GMSD 较高，视觉上表现为噪声抑制不足且图像出现明显过度平滑，重要解剖细节和纹理信息丢失严重。移除网络增强动作后，性能有所提升，但细节保真度和结构清晰度仍不足，细小血管和组织边界恢复不完整。这体现了网络增强操作在处理复杂噪声模式和提升细节恢复能力中的关键作用。

表8　对Mayo测试集进行消融实验所得图像重建结果的性能指标(均值±标准差)

Table 8　Performance metrics (mean ± standard deviation) of image reconstruction results obtained from ablation experiments on the Mayo testset

| 实验类别 | 消融项目 | | SSIM | PSNR/dB | GMSD |
|---|---|---|---|---|---|
| 系统级 | 奖励函数 | 像素奖励 | 0.884 3±0.023 5 | 31.632 8±1.049 6 | 0.070 3±0.008 6 |
| | | 移除结构相似度奖励 | 0.888 3±0.021 6 | 31.946 5±1.010 1 | 0.067 7±0.007 6 |
| | | 移除方差惩罚 | 0.892 5±0.021 7 | 32.153 2±0.952 4 | 0.063 1±0.008 1 |
| | | 复合奖励 | **0.897 2±0.020 2** | **32.319 4±1.015 8** | **0.062 7±0.007 8** |
| | 动作集 | 基础动作 | 0.885 6±0.023 1 | 31.641 6±1.059 8 | 0.069 8±0.008 4 |
| | | 移除网络增强 | 0.888 9±0.020 8 | 32.022 3±1.043 2 | 0.063 6±0.007 9 |
| | | 全部动作集 | **0.897 2±0.020 2** | **32.319 4±1.015 8** | **0.062 7±0.007 8** |
| 组件级 | 双路径并行处理 | 仅保留上路径 | 0.893 6±0.020 9 | 32.173 8±1.028 7 | 0.063 2±0.008 1 |
| | | 仅保留下路径 | 0.889 0±0.022 9 | 31.973 1±1.050 2 | 0.064 9±0.008 8 |
| | | 双路径 | **0.897 2±0.020 2** | **32.319 4±1.015 8** | **0.062 7±0.007 8** |
| | LDNM模块的四分支结构 | 膨胀卷积 | 0.884 2±0.023 1 | 31.492 1±1.032 8 | 0.072 8±0.008 8 |
| | | 移除细节增强 | 0.890 1±0.022 7 | 31.674 2±1.072 3 | 0.066 1±0.007 9 |
| | | 移除噪声抑制 | 0.891 6±0.021 9 | 31.785 3±0.993 4 | 0.065 7±0.008 1 |
| | | 移除身份映射 | 0.893 0±0.020 8 | 31.821 9±1.062 8 | 0.064 9±0.007 3 |
| | | 完整的四分支 | **0.897 2±0.020 2** | **32.319 4±1.015 8** | **0.062 7±0.007 8** |
| | 双重注意力机制 | 移除注意力 | 0.894 9±0.021 7 | 32.044 3±1.027 3 | 0.063 8±0.007 8 |
| | | 移除SA | 0.895 3±0.022 9 | 32.122 4±1.026 1 | 0.063 7±0.007 6 |
| | | 移除ECA | 0.896 1±0.020 7 | 32.099 6±1.012 3 | 0.063 2±0.008 0 |
| | | ECA + SA | **0.897 2±0.020 2** | **32.319 4±1.015 8** | **0.062 7±0.007 8** |
| | 边缘提取算子 | Sobel | 0.897 2±0.020 2 | **32.319 4±1.015 8** | **0.062 7±0.007 8** |
| | | Prewitt | **0.897 8±0.022 8** | 32.317 9±1.026 0 | 0.062 9±0.007 6 |
| | | Canny | 0.897 3±0.020 6 | 32.319 3±1.016 2 | 0.062 7±0.007 9 |
| | | Laplacian | 0.897 2±0.020 8 | 32.318 7±1.019 4 | 0.062 8±0.008 0 |

注:加粗字体表示各列各消融项目的最优结果。

3.4.2　组件级消融实验结果

1)双路径并行处理。双路径架构在各项指标上均最优,可以高效抑制噪声的同时显著提升细小血管和器官边界的恢复质量。仅使用上路径时,其性能与完整双路径接近,但推理速度下降,表明上路径承担了主要特征提取功能;仅使用下路径时,各项指标显著降低,细小血管等精细结构恢复明显不足,说明下路径在细节增强和结构补充方面发挥着重要作用。

2)LDNM模块结构。LDNM模块的完整四分支结构达到最佳性能,能够有效平衡噪声抑制与细节保留。仅使用膨胀卷积时噪声抑制不足且结构模糊;移除细节增强分支后SSIM显著下降,图像中出现细微纹理丢失;移除噪声抑制分支则导致GMSD明显上升,视觉可见局部噪声残留和伪影;移除身份映射后性能全面下降,重建图像出现局部失真。

3)双重注意力机制。针对注意力机制的消融实验结果表明,通道注意力与空间注意力结合取得最优性能,在保持较高计算效率的同时,显著提升了结构保持能力和图像视觉自然度。移除空间注意力后PSNR降至32.122 4 dB,视觉上边缘连续性略有下降;移除通道注意力后图像局部对比度下降,微小结

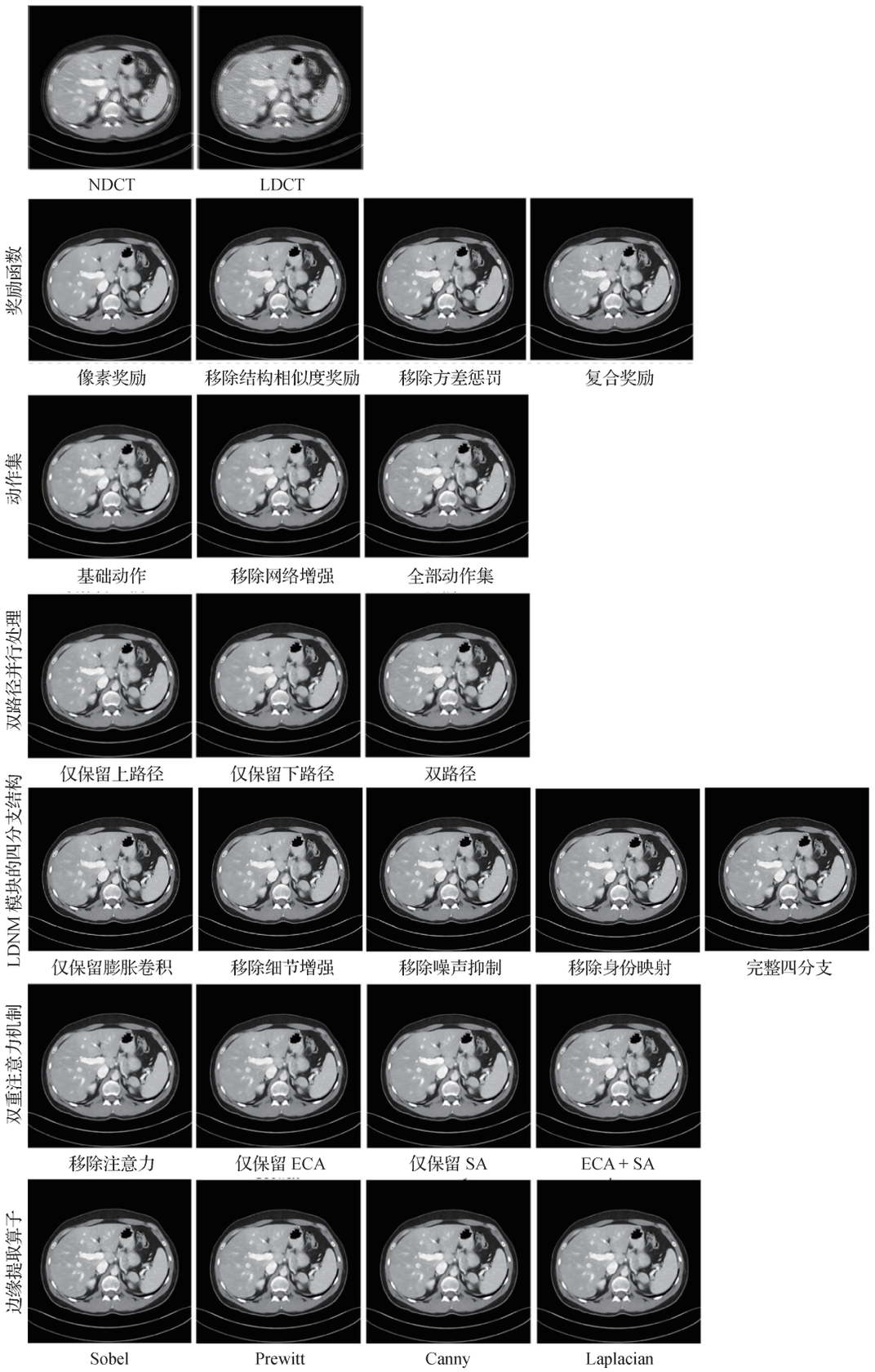



图 12　对 Mayo 测试集进行消融实验所得图像重建结果

Fig. 12　Image reconstruction results obtained from ablation experiments on the Mayo testset

构可见度降低。

4)边缘提取算子。边缘提取算子对比显示，Sobel、Prewitt、Canny及Laplacian 4种算子的性能差异小于0.15%。具体而言，4种算子取得的SSIM均值均在0.897 2～0.897 8之间，PSNR均值在32.317 9～32.319 4 dB之间，GMSD均值在0.062 7～0.062 9之间，且其标准差范围相互重叠，重建图像的视觉质量亦无显著区别。这表明网络对边缘提取算子的选择具有一定的鲁棒性。从实际应用和医学重建需求来看，Sobel算子采用梯度加权方式，对解剖边缘的刻画更清晰，抗噪能力优于Prewitt；与Canny相比，结构简单、实现方便，结果更稳定；而相比于Laplacian，其能保留边缘方向信息，不易放大噪声伪影。因此，在各项指标相当的情况下，综合算子特性与医学图像重建的实际需求，本文选用Sobel算子。

## 4 结 论

针对LDCT图像重建中噪声抑制与结构保留难以平衡的问题，本文提出了一种基于深度强化学习的高分辨率重建框架HRDRL-Net。该框架通过双路径特征协同提取、复合多目标奖励函数与自适应动作选择机制，实现了对图像局部特征的动态优化处理。实验结果表明，HRDRL-Net在Mayo与Piglet公开数据集上均取得显著优于现有方法的性能，在PSNR、SSIM、GMSD与EPI等指标上表现突出，且具有统计显著性。同时，模型在训练数据有限条件下仍保持稳定性能，展现出良好的数据效率与跨解剖部位泛化能力。

本研究主要聚焦于验证HRDRL-Net在标准LDCT重建任务中的基本框架有效性，实验在公开基准数据集上进行。尽管在跨数据集实验中展现了初步的泛化潜力，但要推动其向临床实用转化，仍需在以下方面开展深入研究：1)在连续噪声水平谱系及更广泛解剖结构(如头颅、关节)上测定模型的性能边界，特别需要评估其在复杂病理区域(如肿瘤、钙化灶)的可靠性；2)结合临床先验知识，分析模型在关键解剖与病理区域的决策依据，增强其临床可信度；3)进一步平衡动态决策的精度与速度，探索轻量化架构与推理加速策略，以满足临床实时性需求。

## 参考文献(References)

Chen H, Zhang Y, Kalra M K, Lin F, Chen Y, Liao P X, et al. 2017. Low-dose CT with a residual encoder-decoder convolutional neural network. IEEE Transactions on Medical Imaging, 36(12): 2524-2535 [DOI: 10.1109/TMI.2017.2715284]

Ehman E C, Guimarães L S, Fidler J L, Takahashi N, Ramirez-Giraldo J C, Yu L F, et al. 2012. Noise reduction to decrease radiation dose and improve conspicuity of hepatic lesions at contrast-enhanced 80-kV hepatic CT using projection space denoising. American Journal of Roentgenology, 198(2): 405-411 [DOI: 10.2214/AJR.11.6987]

Furuta R, Inoue N and Yamasaki T. 2020. PixelRL: fully convolutional network with reinforcement learning for image processing. IEEE Transactions on Multimedia, 22(7): 1704-1719 [DOI: 10.1109/TMM.2019.2960636]

Gui Y J, Zhao X, Bai Y J, Zhao R G, Li W Q and Liu Y. 2023. Low-dose CT iterative reconstruction based on image block classification and dictionary learning. Signal, Image and Video Processing, 17(2): 407-415 [DOI: 10.1007/s11760-022-02247-7]

Hojjat M, Shayegan M J and Ghadami O. 2024. Low-dose CT image denoising based on EfficientNetV2 and Wasserstein GAN//Proceedings of the 10th International Conference on Web Research (ICWR). Tehran, Iran, Islamic Republic of: IEEE: 195-201 [DOI: 10.1109/ICWR61162.2024.10533377]

Hossain T, Arafat M Y, Jawad T, Siddiqua A and ul Masood M R. 2025. Performance comparison of CycleGAN for image-to-image translation in multiple domains//Proceedings of the 2nd International Conference on Next-Generation Computing, IoT and Machine Learning. Gazipur, Bangladesh: IEEE: 1-6 [DOI: 10.1109/NCIM65934.2025.11160223]

Huang Z Z, Zhang J P, Zhang Y and Shan H M. 2022. DU-GAN: generative adversarial networks with dual-domain U-Net-based discriminators for low-dose CT denoising. IEEE Transactions on Instrumentation and Measurement, 71: #4500512 [DOI: 10.1109/TIM.2021.3128703]

Li J F, Wang L, Wang S X, Li Y K and Qi F Z. 2025. MMCA: multi-stage multi-order context aggregation framework for LDCT denoising. Applied Intelligence, 55(10): #694 [DOI: 10.1007/s10489-025-06553-8]

Li M K, Liu W and Chen W D. 2025. Image denoising via Swin Transformer V2 and feature fusion U-Net. Journal of Image and Graphics, 30(11): 3524-3534 (利铭康, 柳薇, 陈卫东. 2025. Swin Transformer V2和特征融合的U-Net图像去噪方法. 中国图象图形学报, 30(11): 3524-3534) [DOI: 10.11834/jig.240659]

Li Q, Li R R, Li S Z, Wang T, Cheng Y B, Zhang S M, et al. 2024. Unpaired low-dose computed tomography image denoising using a progressive cyclical convolutional neural network. Medical Phys-

ics，51(2)：1289-1312［DOI：10.1002/mp.16331］

Li Z H，Zhou S W，Huang J Z，Yu L F and Jin M W. 2021. Investigation of low-dose CT image denoising using unpaired deep learning methods. IEEE Transactions on Radiation and Plasma Medical Sciences，5(2)：224-234［DOI：10.1109/TRPMS.2020.3007583］

Li Z Y，Liu Y，Chen Y，Shu H Z，Lu J and Gui Z G. 2023. Dual-domain fusion deep convolutional neural network for low-dose CT denoising. Journal of X-Ray Science and Technology，31(4)：757-775［DOI：10.3233/XST-230020］

Mnih V，Badia A P，Mirza M，Graves A，Lillicrap T，Harley T，et al. 2016. Asynchronous methods for deep reinforcement learning//Proceedings of the 33rd International Conference on Machine Learning. New York，USA：PMLR：1928-1937

Mnih V，Kavukcuoglu K，Silver D，Rusu A A，Veness J，Bellemare M G，et al. 2015. Human-level control through deep reinforcement learning. Nature，518(7540)：529-533［DOI：10.1038/nature14236］

Park M，Lim S，Kim H，Kim J Y and Lee Y. 2024. Optimization of smoothing parameter for block matching and 3D filtering algorithm in low-dose chest and abdominal computed tomography images. Applied Radiation and Isotopes，210：#111374［DOI：10.1016/j.apradiso.2024.111374］

Patwari M，Gutjahr R，Raupach R and Maier A. 2022. Limited parameter denoising for low-dose X-ray computed tomography using deep reinforcement learning. Medical Physics，49(7)：4540-4553［DOI：10.1002/mp.15643］

Sadia R T，Chen J and Zhang J. 2024. CT image denoising methods for image quality improvement and radiation dose reduction. Journal of Applied Clinical Medical Physics，25(2)：#e14270［DOI：10.1002/acm2.14270］

Schulman J，Wolski F，Dhariwal P，Radford A and Klimov O. 2017. Proximal policy optimization algorithms.

Song X G，Zhang P F，Liu W B，Lu X F and Hei X H. 2025. Image superresolution reconstruction based on multiscale large-kernel attention feature fusion network. Journal of Image and Graphics，30(4)：1084-1099（宋霄罡，张鹏飞，刘万波，鲁晓锋，黑新宏. 2025. 多尺度大核注意力特征融合网络的图像超分辨率重建. 中国图象图形学报，30(4)：1084-1099）［DOI：10.11834/jig.240042］

Tan J Q，Zhu X C and Cai M Q. 2021. New mixed loss function based on GMSD. Journal of Hefei University of Technology（Natural Science），44(4)：472-477，502（檀结庆，朱星辰，蔡蒙琪. 2021. 基于GMSD的新混合损失函数. 合肥工业大学学报(自然科学版)，44(4)：472-477），502）［DOI：10.3969/j.issn.1003-5060.2021.04.007］

Wang D Y，Fan F L，Wu Z，Liu R，Wang F and Yu H Y. 2023. CTformer：convolution-free Token2Token dilated vision transformer for low-dose CT denoising. Physics in Medicine and Biology，68(6)：#065012［DOI：10.1088/1361-6560/acc000］

Wang J P，Fan H，Wu Z Y，Du Q，Li M，Zheng J，et al. 2025. Self-adaptive weight embedded lightweight network using semi-supervised learning for low-dose CT image denoising. IEEE Transactions on Radiation and Plasma Medical Sciences，9(7)：890-904［DOI：10.1109/TRPMS.2025.3541169］

Wang T Y，Lucka F and van Leeuwen T. 2024a. Sequential experimental design for X-ray CT using deep reinforcement learning. IEEE Transactions on Computational Imaging，10：953-968［DOI：10.1109/TCI.2024.3414273］

Wang Y L，Han Z F，Zhang X，Shangguan H，Zhang P C，Li J，et al. 2024b. Scale-sensitive generative adversarial network for low-dose CT image denoising. IEEE Access，12：98693-98706［DOI：10.1109/ACCESS.2024.3425606］

Wu L，Liu Q，Huang Z G and Zhang L H. 2025. A review of research on offline reinforcement learning. Chinese Journal of Computers，48(1)：156-187（乌兰，刘全，黄志刚，张立华. 2025. 离线强化学习研究综述. 计算机学报，48(1)：156-187）［DOI：10.11897/SP.J.1016.2025.00156］

You C Y，Li G，Zhang Y，Zhang X L，Shan H M，Li M Z，et al. 2020. CT super-resolution GAN constrained by the identical，residual，and cycle learning ensemble（GAN-CIRCLE）. IEEE Transactions on Medical Imaging，39(1)：188-203［DOI：10.1109/TMI.2019.2922960］

Yu W，Peng W，Yin H，Wang C X and Yu K H. 2021. Low-dose computed tomography reconstruction regularized by structural group sparsity joined with gradient prior. Signal Processing，182：#107945［DOI：10.1016/j.sigpro.2020.107945］

Zhang Y G，Liu J and Gao J B. 2025. Emphasizing the clinical application of deep learning image reconstruction for CT. Chinese Journal of Radiology，59(1)：5-8（张永高，刘杰，高剑波. 2025. 重视深度学习CT图像重建算法的临床应用. 中华放射学杂志，59(1)：5-8）［DOI：10.3760/cma.j.cn112149-20240306-00103］

Zhao F X，Liu M Z，Xiang M R，Li D F，Jiang X，Jin X C，et al. 2025. Unsupervised and self-supervised learning in low-dose computed tomography denoising：insights from training strategies. Journal of Imaging Informatics in Medicine，38(2)：902-930［DOI：10.1007/s10278-024-01213-8］

## 作者简介

张胜楠，女，硕士研究生，主要研究方向为深度强化学习和图像高分辨率重建技术。E-mail：zzssnn2000@163.com

孙正，通信作者，女，教授，主要研究方向为光电仪器和图像处理。E-mail：sunzheng@ncepu.edu.cn

于寒，女，硕士研究生，主要研究方向为深度学习和医学图像配准。E-mail：17360794016@163.com

高章硕，女，硕士研究生，主要研究方向为深度学习和超声成像。E-mail：gao_zhangshuo@163.com

丁港澳，男，硕士研究生，主要研究方向为深度学习的定量光声成像技术。E-mail：15136418713@163.com