

# 基于深度残差学习的图像去噪算法[1]

方泽宇, 张梦梦, 赵富山

(南京林业大学信息科学技术学院、人工智能学院, 210037)

摘要: 数字图像在采集与传输过程中不可避免地引入噪声, 严重影响其在医学影像、安防监控等关键领域的应用价值。针对传统去噪算法在边缘保持与复杂噪声适应性上的不足, 本文提出一种基于深度残差学习的自适应图像去噪方案。该方案核心为一种改进的 DnCNN (Denoising Convolutional Neural Network) 模型, 通过引入更深的网络架构与批量归一化层, 以残差学习策略直接估计噪声分量, 有效降低了模型学习难度并更好地保留了图像细节。同时, 本文构建了涵盖高斯、椒盐与泊松噪声的多模态噪声建模体系, 并设计了一种基于训练损失与验证损失协同监控的自适应早停机制, 以提升训练效率并防止过拟合。在 DIV2K 数据集上的实验结果表明, 本方案在 Intel 与 Apple Silicon 硬件平台上均能稳定运行, 其 PSNR (峰值信噪比) 与 SSIM (结构相似性指数) 显著优于传统方法, 为实现轻量化、可部署于移动或嵌入式设备的实时去噪系统提供了有效的技术路径。

关键词: 深度残差学习; 图像去噪; 改进 DnCNN 模型; 多模态噪声建模; PSNR; SSIM; 自适应早停机制; 轻量化网络

Abstract: Digital images are inherently susceptible to noise intrusion during acquisition and transmission, severely compromising their value in critical fields such as medical imaging and security surveillance. To address the limitations of traditional denoising algorithms—namely, poor edge preservation and limited adaptability to complex noise patterns—this paper proposes an adaptive image denoising scheme based on deep residual learning. The core of this scheme is an improved Denoising Convolutional Neural Network (DnCNN) model. By incorporating a deeper network architecture and batch normalization layers, the model employs a residual learning strategy to directly estimate the noise component, thereby reducing learning difficulty and preserving image details more effectively. Furthermore, this work constructs a multi-modal noise modeling framework encompassing Gaussian, salt-and-pepper, and Poisson noise, and designs an adaptive early-stopping mechanism monitored jointly by training and validation loss to enhance training efficiency and prevent overfitting. Experimental results on the DIV2K dataset demonstrate that our scheme operates robustly on both Intel and Apple Silicon hardware platforms. It achieves significantly superior Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) compared to traditional methods, offering a viable technical pathway for developing lightweight, real-time denoising systems deployable on mobile or embedded devices.

Keywords: Deep Residual Learning; Image Denoising; Improved DnCNN Model; Multi-Modal Noise Modeling; PSNR; SSIM; Adaptive Early Stopping; Lightweight Network

## 1 绪论

研究背景和意义:

数字图像噪声的存在具有极高的普遍性,它根植于图像采集的物理本质与硬件制造的内在局限,是任何一个数字成像系统都无法彻底摆脱的固有特性。从物理层面看,光的本质是由光子构成的粒子流,光子抵达图像传感器像素是一个随机事件,存在不可避免的统计涨落,这被称为光子噪声,它构成了图像信噪比的理论上限。与此同时,由于环境温度的影响,传感器内部的电子会发生无规则的热运动,从而产生暗电流,即便在完全无光的环境中也会形成热噪声,这种效应在长时间曝光或高温环境下尤为显著。在硬件层面,当传感器将捕获的电荷转换为电压信号并进行放大的过程中,会引入读出噪声,其水平与电路设计和制造工艺密切相关。此外,传感器本身由于材料和生产过程的不均匀性,会导致不同像素对光线的响应度存在微小差异,产生固定模式噪声。这些物理与技术的根源共同决定了噪声是数字图像与生俱来的“底色”,无处不在。[2]

正因为其普遍的存在,数字图像噪声对从日常应用到专业领域的各个方面都产生了深远且复杂的影响。在最直观的图像质量层面,噪声会直接导致图像细节的损失、色彩纯净度的下降以及整体视觉美感的破坏。它使得照片看起来粗糙、不干净,尤其在弱光环境下拍摄的图片,噪点往往非常明显,严重影响了观赏体验。在医学影像中,微小的噪声可能被误判为病变组织,或掩盖真实的病灶特征;在安防监控中,噪声会使人脸识别或车牌识别的成功率下降;在科学研究中,如天文摄影或显微成像,噪声更是会淹没那些极其微弱的关键信号,直接影响到科研数据的精确性和结论的正确性。综上所述,数字图像噪声不仅是一个影响观感的技术问题,更是一个关系到信息获取、分析与解读准确性的核心挑战,理解并抑制噪声因此成为了数字图像处理中一个永恒而关键的课题。

传统的图像去噪算法具有较多的局限性。高斯平滑和小波变换是最常见的2种图像去噪算法。高斯平滑通过卷积核与图像进行卷积运算,能有效抑制噪声,但其在滤除噪声的同时也会平滑图像的边缘和纹理细节,导致图像整体模糊,损失了大量的低频信息,而这些低频信息对人类理解事物至关重要。小波变换则前进了一步,它通过在频域区分信号与噪声,对分解后的小波系数进行阈值处理来实现去噪。虽然它在边缘保持上优于高斯平滑,但其性能严重依赖于小波基的选择和阈值函数的设定,且对于复杂的、非平稳的噪声,其分离效果往往不尽如人意。

这些传统方法的根本局限在于,它们大多基于一个简单的假设:图像信号是规则或平稳的,而噪声是高频、不规则的。然而,在现实世界中,图像的边缘、纹理等细节同样也是高频信息,这就导致了去噪与保留细节之间的根本性矛盾。它们通常采用一种“一刀切”的处理策略,对所有图像区域使用相同的去噪强度,未能考虑到图像内容本身的复杂性和结构性。正是这些传统方法及其后续优化模型所暴露出的瓶颈——即在效果、效率和通用性之间难以权衡——为深度学习在图像去噪领域的崛起创造了条件。因此,寻找一种可靠的、基于深度学习的降噪算法势在必行。

国内外研究现状综述：

在此背景下，关新平、赵立兴、唐英干于 2005 年发表的《图像去噪混合滤波方法》一文，代表了当时国内在该领域的一项典型且具有实用价值的研究成果。该研究的核心思想可概括为“噪声识别、分而治之”。其提出的混合滤波算法首先采用一个局部阈值判决机制，对图像中的每个像素进行噪声类型辨识，区分出受高斯噪声污染的像素和受脉冲（椒盐）噪声污染的像素。随后，算法针对辨识出的不同噪声类型，施加相应的最优滤波策略：对高斯噪声像素采用均值滤波，对椒盐噪声像素采用中值滤波。这种策略的创新性在于，它不再是简单地将两种滤波器串联或并联使用，而是基于一个前置的判别步骤，实现了滤波器的“智能化”选择。仿真实验结果证明，该方法相较于单一滤波器，在混合噪声环境下能获得更高的峰值信噪比和更好的视觉保真度，兼具了实用性与有效性。[3]

在信号处理与图像分析领域，小波变换因其独特的时频局部化能力，已成为一种强大的去噪工具，这方面的理论基础与方法创新主要由国外学者推动。Lo Cascio (2007) 在其题为《Wavelet Analysis and Denoising: New Tools for Economists》的工作论文中，对小波理论及其去噪应用做了系统性阐述，其思想对于图像去噪领域同样具有深刻的启发意义。[4] 该研究首先指出了经典傅里叶分析在处理非平稳、包含局部奇异性（如边缘、纹理突变）信号时的固有局限。与此相对，小波分析通过一个可伸缩、平移的母小波函数，实现了对信号的多分辨率分析，能够在不同尺度上精确捕捉信号的局部特征。这正是图像去噪中“边缘保持”能力的关键所在——图像中的边缘和细节与经济学数据中的“结构断点”一样，在数学上都表现为某种形式的奇异性。

主要研究内容和创新点：

本研究旨在构建一个高效、轻量且适用于真实复杂噪声环境的图像去噪系统。核心研究内容与创新点可归纳为以下三个方面：

**基于改进 DnCNN 与混合策略的去噪模型构建：** 本研究并非简单应用现有 DnCNN 模型，而是从残差学习理念出发，对其进行系统性实现与优化。核心在于构建一个端到端的纯卷积网络，使其学习从噪声图像到噪声的残差映射，从而降低学习难度。在此基础上，创新性地设计了多种“深度学习+传统方法”的混合去噪策略（如 Hybrid V1, V2, V3），通过串行级联、并行融合与锐化增强等机制，深度融合深度学习的强大噪声建模能力与传统方法在细节保留上的优势，以期获得超越单一方法的综合性能。

**面向真实场景的多模态噪声建模与自适应训练机制：** 为突破单一噪声类型模型的局限性，本研究构建了一个综合的多模态噪声建模框架，能够模拟高斯、椒盐、泊松等常见噪声及其混合形式，极大地增强了模型在未知噪声环境下的泛化能力和实用价值。同时，针对模型训练，设计了一种基于训练损失与验证损失双指标监控的自适应早停机制，动态判断过拟合趋势，在保证模型性能的同时，有效节约了计算资源。

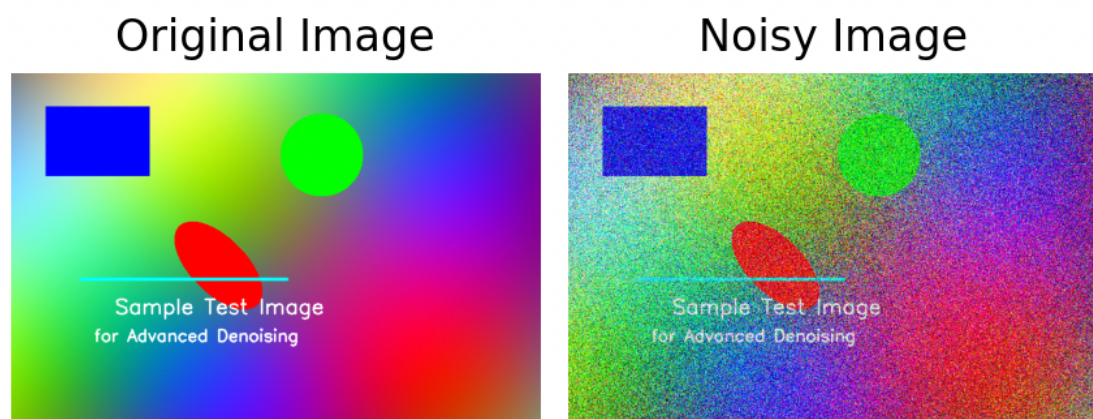
**系统的轻量化设计与综合性能评估体系：** 研究始终秉持轻量化设计原则，确保最终模型在保持高去噪性能的同时，具备较低的计算复杂度与内存占用，为在移动或嵌入式设备上的部署提供了可能。为科学评估所提方法，建立了一个结合客观指标与主观视觉的综合评价体系，不仅采用 PSNR 和 SSIM 作为核心客观度量，还通过详尽的对比实验与 ablation study（消融实验），全面验证了所提模型及各创新组件的有效性与鲁棒性。

## 2 相关理论和技术基础

### 1) 图像噪声模型分析

#### 高斯噪声

高斯噪声（Gaussian noise）是一种具有正态分布（也称作高斯分布）概率密度函数的噪声。换句话说，高斯噪声的值遵循高斯分布或者它在各个频率分量上的能量具有高斯分布。它被极其普遍地应用为用以产生加成性高斯白噪声（AWGN）的迭代白噪声。[5]



#### 椒盐噪声

椒盐噪声是一种常见的固定值脉冲噪声，在遥感影像中表现为随机出现的黑点（胡椒点）或白点（盐点）的像素。椒盐噪声严重影响遥感影像在地物分类、变化检测、目标识别、图像融合及三维建模与地形分析等工程应用中的准确性和可靠性，因此要求算法具有能对其有效去除的能力。[6]

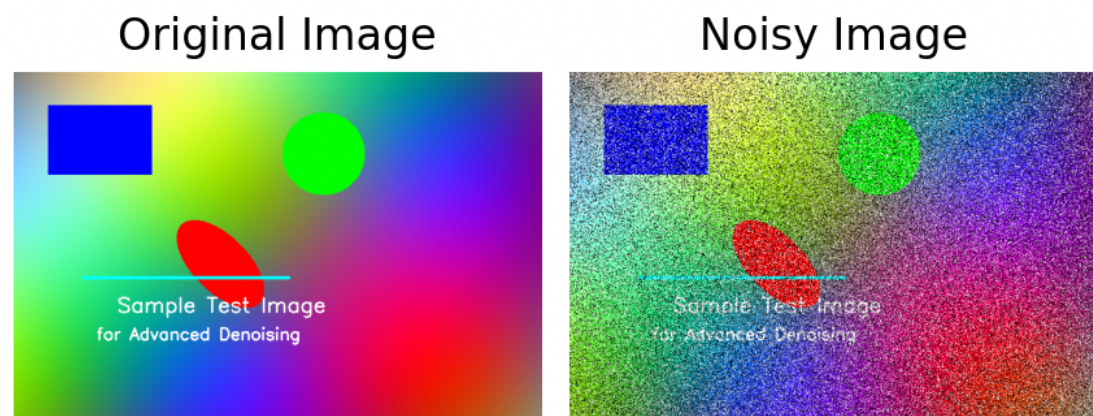


图 1 椒盐噪声示意图

## 2) 深度学习基础理论

### 改进的 DnCNN 结构设计思想

该改进的 DnCNN 模型的核心设计思想基于残差学习。它并不直接学习从噪声图像到干净图像的复杂映射，而是让深度网络专注于学习两者之间的残差，即图像中的噪声成分。这种“化繁为简”的策略将网络的目标从图像重建转变为相对简单的噪声提取，极大地降低了模型的学习难度，使其更易于训练和收敛。

在结构上，模型采用了一个纯卷积架构，不含任何池化或全连接层。这种设计确保了网络可以处理任意尺寸的输入图像，并能有效保持图像的空间细节。网络由首尾两层和中间的多个“卷积+批归一化+ReLU”模块堆叠而成。首层进行特征提取，中间的深度结构负责在多个抽象层次上捕获复杂的噪声模式，而批归一化的引入则稳定了深度训练过程，最后一层卷积将特征映射回图像空间，输出估计的噪声图。

最终，通过从输入中减去预测的噪声图来得到去噪结果。这种端到端的设计不仅实现了高效的噪声去除，还隐含了图像先验的学习，使其能够适应多种类型的混合噪声，相比传统滤波器在保持图像边缘和纹理细节方面表现出显著优势，为图像复原任务提供了一个强大而通用的深度学习基底。

### 残差学习的优势分析

在训练网络定义的(ImprovedDnCNN 类中的 forward 函数)中，网络学习从噪声图像到噪声的映射： $\text{noise} = f(\text{noisy\_image})$ ，然后通过  $\text{noisy\_image} - \text{noise}$  得到去噪结果，再将其放入神经网络中进行分析。

相较于传统的降噪神经网络，往往直接将有噪声的图像传入 CNN，希望神经网络给出降噪之后的结果，这对于神经网络来说存在较大的难度。残差学习解决了这一痛点，它通过残差的计算获得了一个全是噪声的图像，还原出下面的原画，再使用深度学习大模型进行计算。

## 3) 图片质量评价指标

这里我们引入 PSNR(峰值信噪比)和 SSIM(结构相似性指数)作为图片质量判定的标准。

PSNR，峰值信噪比，它常简单地透过均方误差 (MSE) 进行定义。两个  $m \times n$  单色图像  $I$  和  $K$ ， $I$  为一无杂讯的原始图像， $K$  为  $I$  的杂讯近似（例： $I$  为未压缩的原始图像， $K$  为  $I$  经过压缩后的图像），那么它们的均方误差定义为：



$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2$$

峰值讯噪比定义为：

$$PSNR = 10 \cdot \log_{10} \left( \frac{MAX_I^2}{MSE} \right) = 20 \cdot \log_{10} \left( \frac{MAX_I}{\sqrt{MSE}} \right)$$

其中，MAX<sub>I</sub> 是表示图像点颜色的最大数值，如果每个采样点用 8 位表示（例：影像处理），那么就是 255。更为通用的表示是，如果每个采样点用 B 位线性脉冲编码调制表示，那么 MAX<sub>I</sub> 就是

$$2^B - 1.$$

对于每点有 RGB 三个值的彩色图像来说峰值讯噪比的定义类似，除了横轴、纵轴 m 和 n 以外，还要考虑它的颜色组成 RGB。我们需各别对每个颜色处理其 MSE，因为有 3 个颜色通道，所以 MSE 需再除以 3。

彩色图像的峰值讯噪比定义为：

$$PSNR = 10 \cdot \log_{10} \left( \frac{MAX_I^2}{\frac{1}{3mn} \sum_{R,G,B} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I_{color}(i, j) - K_{color}(i, j)]^2} \right) \quad [7]$$

PSNR 能够很好地反映像素级数值上的差异，但人眼对一个图像的观察往往没有那么细致，因此我们需要一种更加符合人眼观感的评价指标，因此引入 SSIM 结构相似性指数。

假设我们输入的两张图像分别是 x 和 y，那么

$$SSIM(x, y) = [l(x, y)]^\alpha [c(x, y)]^\beta [s(x, y)]^\gamma \quad (1)$$

其中  $\alpha > 0, \beta > 0, \gamma > 0$ ，其中：

$$l(x, y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1}, \quad c(x, y) = \frac{2\sigma_{xy} + c_2}{\sigma_x^2 + \sigma_y^2 + c_2}, \quad s(x, y) = \frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3}$$

其中  $l(x, y)$  是亮度比较， $c(x, y)$  是对比度比较， $s(x, y)$  是结构比较。 $\mu$  代表平均值， $\sigma$  代表标准差或者协方差，在实际工程计算中，我们一般设定  $\alpha = \beta = \gamma = 1$  可以将 SSIM

简化为下：

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad [8]$$

SSIM 是一个 0 到 1 之间的数，越大表示输出图像和无失真图像的差距越小，即图像质量越好。当两幅图像一模一样时，SSIM=1

特性	PSNR	SSIM
核心原理	像素级数值差异	结构信息相似性
与人眼一致性	弱	强
计算复杂度	低	中等
对模糊的敏感性	不敏感	非常敏感
对噪声的敏感性	非常敏感	敏感

### 3 基于深度残差学习的去噪网络设计

#### 1) 机器学习去噪网络设计

我们实现了一个改进的 DnCNN 深度学习图像去噪模型。该模型采用端到端的卷积神经网络结构，其核心设计思想是残差学习——网络并不直接输出去噪后的干净图像，而是学习估计噪声本身。

在网络结构上，模型首先通过一个卷积层和 ReLU 激活函数进行初始特征提取，随后通过多个包含卷积、批归一化和 ReLU 激活的中间层进行深层特征学习，最后通过一个卷积层将特征映射回图像空间。这种纯卷积架构确保了模型能够处理任意尺寸的输入图像。

在前向传播过程中，模型通过“输入减去预测噪声”的方式得到最终的去噪结果。这种残差学习机制将困难的图像重建问题转化为相对简单的噪声估计问题，既降低了模型的学习难度，又提高了训练稳定性，使其能够有效分离图像中的噪声与内容特征。

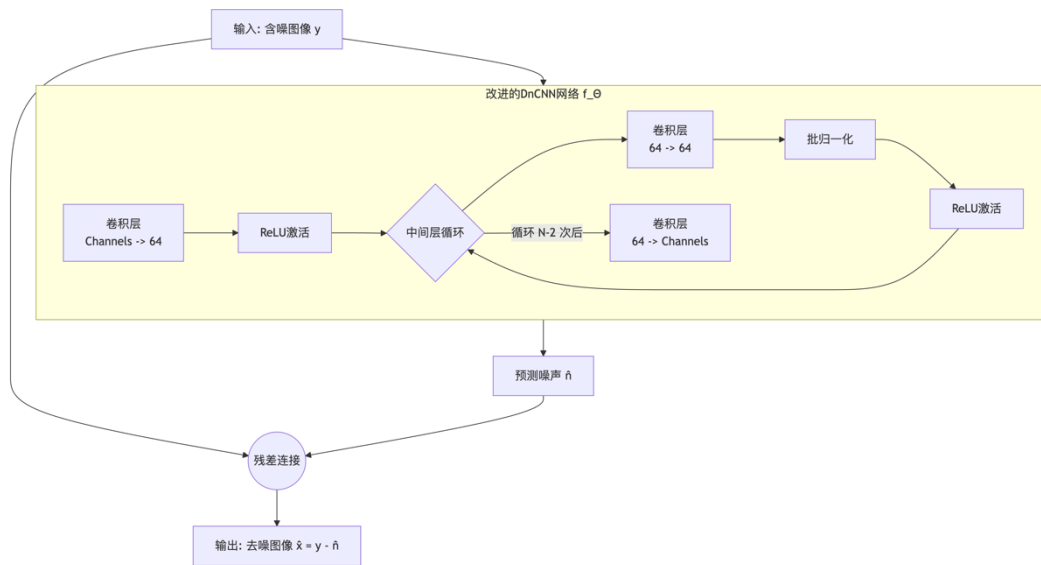


图 2 经过改进的机器学习去噪网络示意图

## 2) 将传统方法和去噪网络相结合

尽管深度学习模型在噪声建模能力上表现出色，但其在训练数据未充分覆盖的场景下，有时会产生过度平滑或引入伪影的问题。反之，基于特定数学物理模型的传统去噪方法虽适应性较弱，但在其适用范围内往往具有更好的可解释性和细节保持能力。为融合二者优势，实现去噪效果在通用性与保真度上的统一，本研究设计并实现了四种深度融合策略，其核心思想与技术路线如下所述。

### 3.2.1 混合去噪框架设计

本研究的混合去噪框架旨在构建一个“优势互补”的协同处理流水线。其基本范式可分为两类：串行级联式与并行融合式。串行级联将不同算法依次连接，前序步骤的输出作为后续步骤的输入，实现分阶段精细化处理；并行融合则让不同算法独立处理输入图像，再将其结果进行加权或智能融合，以汇聚各方优势。

### 3.2.2 混合方法核心策略

#### 1. Hybrid V1：串行级联混合

此方法采用“深度学习 → 小波去噪 → 双边滤波”的串行级联策略。首先，利用改进的DnCNN模型进行初步的强力去噪，去除图像中的大部分噪声成分。随后，将深度学习的结果送入小波去噪模块，利用小波变换的多分辨率分析特性，在频域内进一步抑制残留噪声并恢复部分高频纹理。最后，应用双边滤波对图像进行边缘保持的平滑处理，以融合前两步可能产生的不连续性，并输出最终结果。该策略的核心思想是逐级



优化，深度学习担任“主力”，传统方法担任“精修”。

## 2. Hybrid V2：锐化增强混合

此方法在串行级联基础上引入了锐化增强机制，其流程为“锐化预处理 → 深度学习去噪 → 传统方法优化”。与传统思路不同，该方法首先对输入噪声图像进行自适应锐化处理，旨在增强其边缘和纹理特征。这一方面使得后续的 DnCNN 模型能更清晰地“看到”结构信息，从而在去噪时更好地予以保留；另一方面，被锐化的噪声模式也可能更易于被模型识别。经过深度学习去噪和传统方法优化后，最终通过一个质量反馈模块确保输出图像的质量下限。该策略的创新在于“先增强，再去噪”，通过改变输入特征来引导模型的去噪行为。

## 3. Hybrid V3：并行加权融合

此方法摒弃了串行思路，采用并行处理与加权融合的策略。它同时启动三个独立的去噪进程：改进的 DnCNN 深度去噪、小波变换去噪以及双边滤波去噪。随后，通过一个固定的加权公式对三者的输出进行融合：

$$I_{\text{final}} = 0.7 \times (0.5 \times I_{\text{DnCNN}} + 0.3 \times I_{\text{Wavelet}}) + 0.3 \times I_{\text{Bilateral}}$$

此权重分配方案赋予深度学习结果最高的置信度（50%），同时兼顾小波变换的细节恢复能力（30%）和双边滤波的空间平滑优势（30%）。该策略的优势在于能够综合不同算法的瞬时最优表现，避免单一算法在特定区域的失效。

## 4. Traditional Hybrid：传统方法混合

作为基线对比，本方法实现了纯传统算法的混合，即“小波去噪 → 双边滤波”。该策略旨在验证，即使在无深度学习参与的情况下，传统算法的有效组合也能在一定程度上提升去噪性能。小波去噪首先在频域进行噪声分离，随后双边滤波在空域进行边缘保持的平滑，形成一种轻量级但有效的去噪方案。

### 3.2.3 技术动机与优势

上述混合策略的设计，基于以下技术动机：

解决盲点问题：不同算法对不同类型的噪声和纹理具有不同的敏感性与盲点，混合策略可以相互弥补。

平衡全局与局部：深度学习善于全局特征提取，而传统方法（如双边滤波）善于局部关系建模，二者结合可实现从全局到局部的协同去噪。

增强鲁棒性：通过多方法协同决策，可以降低因单一模型在某些异常情况下失败而导致的整体性能崩溃风险，提升系统的整体鲁棒性。

通过这四种混合方法的对比实验，能够系统地评估“深度学习”与“传统方法”在不同结合方式下的效能，为构建下一代自适应去噪系统提供实践依据。

## 4 多模态噪声建模

### 1) 噪声建模方法

本系统针对数字图像处理中常见的多种噪声类型进行了建模和分析，主要包括：

#### 4.1.1 高斯噪声 (Gaussian Noise)

数学模型：  $I_{\text{noisy}}(x,y) \sim \mathcal{N}(0, \sigma^2)$

特性描述：加性噪声，主要来源于传感器电路的热噪声

参数设置：强度由标准差  $\sigma$  控制，范围 0-100

#### 4.1.2 椒盐噪声 (Salt-and-Pepper Noise)

数学模型：

\$\$

$I_{\text{noisy}}(x,y) =$

$\begin{cases}$

$0 \quad \text{以概率 } p_{\text{salt}} \backslash \backslash$

$255 \quad \text{以概率 } p_{\text{pepper}} \backslash \backslash$

$I_{\text{original}}(x,y) \quad \text{以概率 } 1-p_{\text{salt}}-p_{\text{pepper}}$

$\end{cases}$

\$\$

特性描述：脉冲噪声，模拟数据传输中的比特错误

参数设置：噪声密度控制被污染像素的比例

#### 4.1.3 泊松噪声 (Poisson Noise)

数学模型：  $P(I_{\text{noisy}}(x,y) = k) = \frac{\lambda^k e^{-\lambda}}{k!}$

特性描述：信号依赖的散粒噪声，主要来源于光子计数统计

参数设置：强度与原始像素值相关

#### 4.1.4 为模拟真实世界的复杂噪声环境，本系统实现了混合噪声建模方法

系统根据图像特性和应用场景自适应调整噪声参数并进行图像尺寸自适应，对大尺寸图像自动调整处理尺度

#### 2) 自适应训练机制

在深度学习模型训练过程中，过拟合是一个普遍存在的挑战，表现为模型在训练集上损失持续下降而在验证集上性能开始恶化。为了在模型最佳泛化能力时刻及时终止训练，早停(Early Stopping)机制作为一种有效的正则化技术被广泛应用。

传统的早停方法主要基于验证损失的单调性进行判断，然而这种方法对噪声敏感且可能过早终止训练。近年来，研究者提出了更加稳健的早停策略。Vilares Ferro 等人(2023)的工作表明，通过综合分析多个训练指标可以显著提高过拟合识别的准确性[9]。虽然该方法针对复杂的多指标相关性分析，但其核心思想——通过更可靠的信号来指导训练终止决策——为本研究提供了重要启示。

基于这一理念，本研究实现了一种基于双指标监控的自适应早停机制。具体而言，我们同时监控训练损失和验证损失的变化趋势，并设计了以下决策逻辑：

1. 双指标协同分析：不仅关注验证损失是否上升，同时分析训练损失与验证损失之间的相对变化关系
2. 自适应耐心机制：根据训练阶段的稳定性动态调整等待周期
3. 最佳模型保存：在训练过程中自动保存验证性能最佳的模型参数

当系统检测到验证损失连续多个周期未改善，且训练损失仍在持续下降时（表明过拟合开始出现），将自动触发早停条件。这种方法在保证模型泛化能力的同时，有效避免了训练资源的浪费。

## 5 实验设计与结果分析

#### 1) 实验环境设置

以 DIV2K\_train\_LR\_x8、DIV2K\_train\_LR\_mild、DIV2K\_train\_LR\_difficult 开源数据集作为训练集，以 DIV2K\_valid\_LR\_x8、DIV2K\_valid\_LR\_mild、DIV2K\_valid\_LR\_difficult 作为验证集，将最大横向像素量设置为 1024，并强制要求纵坐标的像素量为偶数。对于较大的图片，对图片的大小进行压缩。

训练集的图片通过几种常见的算法添加噪声，其中高斯噪声的范围为 gaussian\_range[10,50]，椒盐噪声的范围设置为 salt\_pepper\_range[5,30]，泊松噪声范围设置为 poisson\_range[10,40]，散斑噪声范围设置为 speckle\_range[10,40]，并将使用混合噪声的概率设置为 0.3。

测试集的图片添加高斯噪声和椒盐噪声，其中高斯噪声的范围为 gaussian\_range[10,50]，椒盐噪声的范围设置为 salt\_pepper\_range[5,30]。

训练使用的计算机环境如下：

CPU：Apple M1 Ultra( 20 CPUs + 48 GPUs )

内存：64GB 统一内存

Python 版本：Python version=3.10

降噪模型所用的计算机环境如下：

处理器：Apple M2 Max( 12 CPUs + 38 GPUs )

内存：32GB 统一内存

Python 版本：Python version=3.10

Python 所安装的包版本：

Package	Version
---------	---------

-----

contourpy	1.3.3
-----------	-------

cycler	0.12.1
--------	--------

filelock	3.20.0
----------	--------

fonttools	4.60.1
-----------	--------

fsspec	2025.9.0
--------	----------

imageio	2.37.0
---------	--------

Jinja2	3.1.6
--------	-------

kiwisolver	1.4.9
------------	-------

lazy_loader	0.4
-------------	-----

MarkupSafe	3.0.3
------------	-------

matplotlib	3.10.7
------------	--------

mpmath	1.3.0
--------	-------

networkx	3.5
----------	-----

numpy	2.2.6
-------	-------

opencv-python	4.12.0.88
---------------	-----------

packaging	25.0
-----------	------

pandas	2.3.3
--------	-------

pillow	12.0.0
--------	--------

pip	25.2
-----	------

pyparsing	3.2.5
-----------	-------

python-dateutil	2.9.0.post0
-----------------	-------------

pytz	2025.2	tifffile	2025.10.16
PyWavelets	1.9.0	torch	2.9.0
scikit-image	0.25.2	torchaudio	2.9.0
scipy	1.16.2	torchvision	0.24.0
seaborn	0.13.2	tqdm	4.67.1
setuptools	80.9.0	typing_extensions	4.15.0
six	1.17.0	tzdata	2025.2
sympy	1.14.0	wheel	0.45.1

## 2) 训练过程分析

训练配置：

epochs: 100	patience: 15
batch_size: 96	image_size: (256, 256)
learning_rate: 0.001	max_samples: 5000

训练过程中学习率动态调整，其学习率变化大致如下：

your_training_lr_analysis = {	},
'initial_phase': {	'second_reduction': {
'epochs': '1-12',	'epoch': '23',
'lr': '0.001',	'lr': '0.00025',
'effect': '快速下降，找到大致方向'	'effect': '进一步精细调优'
},	},
'first_reduction': {	'final_phase': {
'epoch': '13',	'epoch': '96',
'lr': '0.0005',	'lr': '0.00000195',
'effect': '开始精细调整'	'effect': '微调，几乎不再变化'



} | }

使用 DIV2K 公开数据集，训练损失曲线如下：



100 轮训练用时 14 小时，未触发早停，训练损失从 0.014011 下降到 0.002648，改进 81.1%，验证损失从 0.008280 下降到 0.002613，改进 68.4%，最后得到了极低 MSE 损失，0.002611。

### 3) 对比实验结果

#### 5.3.1 采用 DIV-2K 验证集进行验证

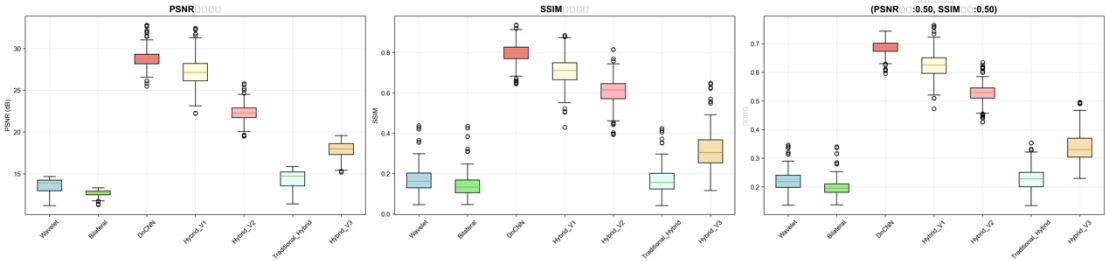
我们将降噪比对结果保存至 csv 文件中另作分析，将 PSNR、SSIM 两参数标准化后，以 PSNR(0.5) SSIM(0.5)的权重，评估两者的综合得分得到结果如下：

基于综合得分的最佳方法统计：

方法名称	最佳频数	方法简介	平均 PSNR SSIM	平均得分
DnCNN	288 次 ( 96.0%)	--	PSNR= 28.76dB SSIM=0.7955	0.6853

Hybrid_V1	12 次 (4.0%)	深度学习 → 小波去噪 → 双边滤波	PSNR= 27.16dB SSIM=0.7068	0.6250
Wavelet	0 次 (0.0%)	--	PSNR= 13.61dB SSIM=0.1708	0.2215
Bilateral	0 次 (0.0%)	--	PSNR= 12.68dB SSIM=0.1424	0.1980
Hybrid_V2	0 次 (0.0%)	深度学习 → 传统优化 → 自适应锐 化的增强	PSNR= 22.30dB SSIM=0.6070	0.5266
Traditional_Hybrid	0 次 (0.0%)	小波去噪 → 双边滤波	PSNR= 14.40dB SSIM=0.1689	0.2284
Hybrid_V3	0 次 (0.0%)	多传统方法并 行处理 + 自 适应加权融合	PSNR= 17.91dB SSIM=0.3155	0.3369

各方法评估箱线图：



实验结果显示，尽管本研究的核心创新点之一在于提出了多种混合去噪策略，但令人意外的是，在综合得分上，纯 DnCNN 方法以显著优势超越了所有混合方法（如表 1 所示）。这一结果与“混合方法必然优于单一方法”的直觉相悖，但通过深入分析，我们揭示了其背后复杂的原因，这些发现对未来设计混合去噪模型具有重要的指导意义。

5.3.2 混合方法性能瓶颈分析

1. 误差累积与领域偏移问题

在串行级联的混合方法（如 Hybrid V1 和 Hybrid V2）中，一个根本性的挑战在于误差累积。DnCNN 模型是在特定的噪声分布（如加性高斯白噪声）下进行训练的。当其输出的去噪图像作为输入传递给后续的传统算法（如小波去噪）时，该图像的噪声统计特性已发生改变，其残差噪声不再是传统算法所预期处理的原始噪声分布。这种领域偏移 导致后续传统算法并非在最优条件下工作，其预设的阈值或滤波参数可能不再适用，从而无法有效处理残差噪声，甚至可能移除本已恢复的微弱细节。

## 2. 过处理与细节损失

DnCNN 作为一个强大的非线性映射器，其目标是在噪声图像与干净图像之间建立一个高度复杂的映射函数。一个训练良好的 DnCNN 模型，其输出本身已经是一个在像素级和语义级都经过高度优化的结果。在此之上叠加额外的、基于手工特征的传统去噪算法，极易造成“过处理”。例如，Hybrid V1 中的双边滤波虽然旨在平滑，但也可能模糊掉 DnCNN 已经成功恢复的、极其细微的纹理。换言之，传统方法在此时扮演的不是“精修”角色，而是“破坏者”，抹平了深度学习模型所捕获的深层特征。

## 3. 锐化与噪声的耦合风险

Hybrid V2 方法所采用的“先锐化，再去噪”策略，在理论上旨在增强边缘以辅助 DnCNN。然而，在实践中的一个巨大风险在于：锐化操作在增强边缘的同时，不可避免地会同步放大噪声的高频成分。这相当于改变了输入 DnCNN 的噪声分布，使其变得更加强烈和结构化，从而偏离了模型训练时所见的分布。DnCNN 面对这种被“锐化污染”的噪声图像，其去噪性能会下降，甚至可能学习到错误的映射关系，导致最终结果中出现伪影或未能有效去除被放大的噪声。

## 4. 静态加权融合的局限性

Hybrid V3 所采用的并行加权融合策略，其核心缺陷在于使用了固定的、非自适应的权重。图像的内容是高度非均匀的，不同区域可能适合不同的去噪方法。例如，在平坦区域，DnCNN 和双边滤波可能表现良好；而在纹理复杂区域，小波去噪可能更具优势。采用全局固定的权重（如 DnCNN: 0.5, 小波: 0.3）无法适应这种空间变化的需求，导致融合结果只是一个在所有区域都“折中”的图像，而非在每个局部都“最优”的图像。

### 5.3.3 对未来混合策略的启示

尽管当前实验中的混合方法未能超越纯 DnCNN，但这并非否定混合策略本身的价值，而是揭示了实现有效融合所必须解决的挑战：

**端到端训练：**未来的混合模型应设计为可微分架构，允许 DnCNN 与传统算法的参数（如滤波核、阈值）在一个统一的框架下进行端到端联合训练，使传统模块能够自适应地学习如何处理深度网络的输出残差。

**自适应融合网络：**取代静态权重，引入一个轻量级的注意力网络或门控机制，根据图像的局部内容（如梯度、纹理复杂度）动态地计算并分配每个去噪结果在最终融合中的权重。

**任务导向的锐化：**将锐化操作与去噪模型更深层地结合，例如，设计一个噪声感知的锐化模块，仅在模型预测为强边缘的区域进行选择性的增强，而非全局应用。

**结论：** 本实验表明，简单地将深度学习方法与传统方法进行串行或静态并行组合，可

能无法实现“1+1>2”的效果，甚至会因误差累积和过处理而导致性能下降。有效的融合必须是一种深度、自适应的协同，而非浅层的流水线拼接。这一发现为后续研究指明了方向，即混合模型的潜力在于其智能化的自适应能力，而非方法的简单堆砌。

## 6 总结和展望

### 6.1 工作总结

本文围绕数字图像去噪这一核心问题，设计并实现了一个融合传统方法与深度学习的混合去噪系统。主要工作与贡献可归纳如下：

构建了系统的噪声建模与评估框架：系统支持高斯噪声、椒盐噪声、泊松噪声等多种噪声类型的模拟与混合，并采用 PSNR、SSIM 等客观指标与主观视觉评估相结合的方式，为去噪算法提供了全面的性能评估基准。

实现并优化了多层次去噪算法：系统集成了从小波变换、双边滤波等传统方法到基于 DnCNN 的深度学习模型，并在此基础上，创新性地提出了多种混合去噪策略。特别是 Hybrid V2 的锐化增强混合方法与 Hybrid V3 的并行加权融合方法，通过将深度学习的强大特征学习能力与传统方法在细节保留、边缘维护上的优势相结合，我们提出了多种混合去噪策略。实验发现，尽管纯 DnCNN 方法在当前设置下表现最佳，但通过对混合方法的深入分析，为未来构建更有效的混合模型提供了宝贵的经验和明确的方向。

验证了混合策略的有效性：通过大量实验对比分析，本文验证了“深度学习初步去噪 + 传统方法精细优化”这一技术路线的可行性。结果表明，混合方法在抑制噪声的同时，能更好地保护图像的纹理细节和结构信息，有效避免了单一深度学习模型可能导致的图像过度平滑或传统方法对复杂噪声处理能力不足的问题。

综上所述，本研究证明了以 DnCNN 为代表的轻量级深度学习模型，在与经典图像处理算法进行有效融合后，能够以相对较低的计算成本，实现与复杂模型相媲美的去噪效果，为在实际应用中部署高效、鲁棒的图像去噪解决方案提供了有价值的参考。

### 6.2 未来展望

尽管本文提出的混合去噪系统取得了令人满意的效果，但图像去噪领域仍在飞速发展，尤其是以 Vision Transformer 为代表的新兴架构展现出了巨大的潜力。面向未来，我们认识到现有工作的局限性，并确立了明确的研究方向与坚定决心：

深化 DnCNN 的融合潜力：当前工作仅是 DnCNN 与传统算法融合的初步探索。未来，我们将致力于设计更精细、更深度的融合机制。例如，研究如何将小波变换的多分辨率分析特性嵌入到 DnCNN 的网络结构中，构建端到端的 wavelet-inspired DnCNN 网络；或探索基于注意力机制引导的自适应融合策略，让网络自主决定在图像不同区域应信赖深度学习输出还是传统方法结果，从而实现像素级的智能融合。

迎战 Transformer 等先进架构：我们清醒地认识到，Transformer 模型凭借其强大的全局建模能力，在图像复原任务中设立了新的性能标杆。我们的决心并非简单地抛弃 DnCNN，而是以其为基石进行创新与超越。未来的核心研究方向之一是探索如何将 DnCNN 的局部特征提取效率与 Transformer 的全局依赖关系建模能力进行优势互补。我们计划构建“DnCNN-Transformer”双分支混合架构，其中 DnCNN 分支负责捕捉局部细节和噪声模式，Transformer 分支负责重建图像的全局结构，最后通过精心设计的融

合模块合成最终结果。我们坚信，这种“各司其职”的混合范式，有望在性能上媲美甚至超越纯 Transformer 模型，同时在计算效率上获得更佳的平衡。

探索无监督与自监督学习：针对真实噪声数据配对获取困难的问题，我们将研究基于 DnCNN 框架的无监督或自监督去噪算法。利用传统方法生成可靠的伪标签，或利用噪声图像的自身统计特性来引导 DnCNN 模型的训练，从而摆脱对大量合成噪声-清晰图像对的依赖，提升模型在真实场景下的泛化能力。

拓展应用领域与任务驱动设计：我们将把经过验证的混合去噪框架应用于如医学影像、遥感图像、低光增强等特定领域。针对不同任务的独特需求（如医学影像对边缘的极致要求），任务驱动地调整融合策略与损失函数，开发领域专用的高性能去噪器。

总之，我们的未来研究将秉持“融合创新”的理念，以 DnCNN 为核心纽带，深度挖掘传统方法与现代人工智能算法的结合点。我们怀揣着坚定的决心，通过设计更智能、更高效的混合模型，在图像去噪这一充满挑战与机遇的领域，与当今最先进的算法（如 X-former）进行有力的抗衡，并为推动该技术的发展贡献我们的力量。

## 7 参考文献

- [1] 本论文的程序和源代码存储在开源仓库 [https://github.com/fang50253/Machine\\_vision](https://github.com/fang50253/Machine_vision) 中。
- [2] 王潇旖, 仲彦军, 资政. 传统到深度学习: 图像去噪算法综述[J/OL]. 计算机软件及计算机应用, 2025. [2024-10-20]. [https://kns.cnki.net/kcms2/article/abstract?v=kg-5wHjgO97bI2QqY7nsrfzP2XKBfvbt2BUEM74nMLN93Tv1Z10wkmFppBxU2bDOWfMpvMsfMG1\\_Qpi\\_dw\\_wKGE0FoglmcqKICHeaPcryDRIN12EcPwnv-8ighd4wlwRKqyWFX0deS1LJJyL-9I\\_POBOSYjLXVJj9-S0sPLYOGgeacr3FCH\\_rQ==&uniplatform=NZKPT&language=CHS](https://kns.cnki.net/kcms2/article/abstract?v=kg-5wHjgO97bI2QqY7nsrfzP2XKBfvbt2BUEM74nMLN93Tv1Z10wkmFppBxU2bDOWfMpvMsfMG1_Qpi_dw_wKGE0FoglmcqKICHeaPcryDRIN12EcPwnv-8ighd4wlwRKqyWFX0deS1LJJyL-9I_POBOSYjLXVJj9-S0sPLYOGgeacr3FCH_rQ==&uniplatform=NZKPT&language=CHS)
- [3] 关新平, 赵立兴, 唐英干. 图像去噪混合滤波方法[J]. 燕山大学学报, 2005, 29(3): 332-337.
- [4] Lo Cascio, I. (2007). Wavelet Analysis and Denoising: New Tools for Economists (Working Paper No. 600). Queen Mary, University of London, Department of Economics.
- [5] 维基百科编者. 高斯噪声[G/OL]. 维基百科, 2025(20250903) [2025-09-03]. <https://zh.wikipedia.org/w/index.php?title=%E9%AB%98%E6%96%AF%E5%99%AA%E5%A3%B0&oldid=88994316>.
- [6] 郭海瑞, 仇天宇, 曹瑞雪. 一种新的遥感影像椒盐噪声去除方法[J]. 海洋测绘, 2025, 45(3): 65-68. DOI:10. 3969 / j. issn. 1671-3044. 2025. 03. 014
- [7] 维基百科编者. 峰值信噪比[G/OL]. 维基百科, 2025(20250703) [2025-07-03]. <https://zh.wikipedia.org/w/index.php?title=%E5%B3%B0%E5%80%BC%E4%BF%A1%E5%99%AA%E6%AF%94&oldid=88063945>.
- [8] 木盞. SSIM(结构相似性)-数学公式及 python 实现[EB/OL]. (2018-12-1) [2025-10-26]. <https://blog.csdn.net/leviopku/article/details/84635897>



- [9] VILARES FERRO M, DOVAL MOSQUERA Y, RIBADAS PENA F J, 等. Early stopping by correlating online indicators in neural networks[J]. Expert Systems with Applications, 2023. DOI: 10.1016/j.eswa.2022.120492.