

# 基于改进型Transformer网络的图像去噪算法

陈勇<sup>1</sup>, 李松<sup>2</sup>, 李萌<sup>1</sup>, 毛秋吉<sup>2</sup>, 王云辉<sup>1</sup>, 张珂<sup>2</sup>

<sup>1</sup>云南电网有限责任公司, 云南 昆明

<sup>2</sup>云南电网有限责任公司红河供电局, 云南 红河

收稿日期: 2022年11月14日; 录用日期: 2022年12月14日; 发布日期: 2022年12月22日

## 摘要

对于一个图像处理系统, 包括图像的获取、处理、传输、接收、输出等环节, 都会存在不同程度的噪声, 使图像质量降低, 影响后续的处理。Transformers神经网络结构在自然语言和高级视觉任务上表现出显著的性能提升。虽然Transformer网络减轻了卷积神经网络在感受野和注意力机制上的不足, 但其计算复杂度随空间分辨率成二次方增长, 因此无法应用于大多数涉及高分辨率图像的图像恢复任务。本文基于改进型的Transformer网络架构, 通过改进的注意力机制以及像素重塑模块, 有效地降低了Transformer网络的计算复杂度, 使得模型能够支持更高分辨率的图像输入, 为模型提供了更好的图像细节, 使得模型在图像去噪上达到了优于以上传统方法和基于卷积神经方法的效果。

## 关键词

图像去噪, Transformer神经网络, 注意力机制, 高分辨率, 计算机视觉

# Image Denoising Algorithm Based on Improved Transformer Network

Yong Chen<sup>1</sup>, Song Li<sup>2</sup>, Meng Li<sup>1</sup>, Qiuji Mao<sup>2</sup>, Yunhui Wang<sup>1</sup>, Ke Zhang<sup>2</sup>

<sup>1</sup>Yunnan Power Grid Co., Ltd., Kunming Yunnan

<sup>2</sup>Yunnan Power Grid Co., Ltd. Honghe Power Supply Bureau, Kunming Yunnan

Received: Nov. 14<sup>th</sup>, 2022; accepted: Dec. 14<sup>th</sup>, 2022; published: Dec. 22<sup>nd</sup>, 2022

## Abstract

For an image processing system, including image acquisition, processing, transmission, reception, and output, there are varying degrees of noise that degrade the image quality and affect the subsequent processing. Transformer-based neural network structure shows significant performance

文章引用: 陈勇, 李松, 李萌, 毛秋吉, 王云辉, 张珂. 基于改进型 Transformer 网络的图像去噪算法[J]. 计算机科学与应用, 2022, 12(12): 2763-2771. DOI: 10.12677/csa.2022.1212280

improvement on natural language and advanced computer vision tasks. Although Transformer networks alleviate the deficiencies of convolutional neural networks in perceptual field and attention mechanisms, their computational complexity grows quadratically with spatial resolution and thus cannot be applied to most image recovery tasks involving high-resolution images. In this paper, based on the improved Transformer network architecture, the computational complexity of the Transformer network is effectively reduced by the improved attention mechanism and the pixel reshaping module, which enables the model to support higher-resolution image inputs and provides better image details for the model, making the model achieve better image denoising than the above traditional methods and convolutional-based neural methods.

## Keywords

Image Denoising, Transformer-Based Neural Network, Attentional Mechanisms, High Resolution, Computer Vision

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

图像恢复是从退化(如噪声、模糊、雨滴)的输入消除退化的任务,这是一个极具挑战性的问题[1]。图像去噪是许多其他图像工作的基础,如图像分割、边缘提取、图像识别等,图像中的噪声信息会很大程度影响后续工作的效果,图像去噪对于军工、医疗等领域也具有很高的研究价值,因此进行后续的工作之前通过图像去噪来去除干扰信息的必要的。

噪声可以理解为妨碍人的视觉器官或系统传感器对所接收图像源信息进行理解或分析的各种因素。对于一个图像处理系统,包括图像的获取、处理、传输、接收、输出等环节,都会存在不同程度的噪声,使图像质量降低,影响后续的处理。经典的图像去噪方法[2]有采用邻域平均法的均值滤波器非常适用于去除通过扫描得到的图像中的颗粒噪声的均值滤波器、根据图像的局部方差来调整滤波器的输出的自适应维纳滤波器、把数字图像或数字序列中一点的值用该点的一个领域中各点值的中值代换的中值滤波器、将开启和闭合结合起来用来滤除噪声的形态学滤波器、利用小波信号,保留图像细节信息的小波去噪滤波器等等。图像去噪算法可以在尽量保留图像细节特征的前提下对目标图像的噪声进行抑制,是图像预处理中不可缺少的操作,其处理效果的好坏将直接影响到后续图像处理和分析的有效性和可靠性。由于成像系统、传输介质和记录设备等的不完善,数字图像在其形成、传输记录过程中往往会受到多种噪声的污染。另外,在图像处理的某些环节当输入的图像对象并不如预想时也会在结果图像中引入噪声。这些噪声在图像上常表现为易引起较强视觉效果的孤立像素点或像素块。一般,噪声信号与要研究的对象不相关,它以无用的信息形式出现,扰乱图像的可观测信息。对于数字图像信号,噪声表示为或大或小的极值,这些极值通过加减作用于图像像素的真实灰度值上,对图像造成亮、暗点干扰,极大降低了图像质量,影响图像复原、分割、特征提取、图像识别等后继工作的进行。要构造一种有效抑制噪声的滤波器必须考虑两个基本问题[3]:能有效地去除目标和背景中的噪声;同时,能很好地保护图像目标的形状、大小及特定的几何和拓扑结构特征。

当前的图像去噪方法可以大致分为基于传统算法和基于深度学习方法[4]的两大类。传统方法包括双边滤波、中值滤波、小波变换等方法,主要思想是在去噪的基础上保持图像特征的平滑并且保证图像的边缘

信息被保留。由于卷积神经网络(CNN)在从大规模数据中学习可泛化的图像先验方面表现良好,因此这些模型已广泛应用于图像恢复和相关任务。最近,另一类 Transformers 网络结构在自然语言和高级视觉任务上表现出显著的性能提升。虽然 Transformer 网络减轻了卷积神经网络在感受野和注意力机制上的不足,但其计算复杂度随空间分辨率成二次方增长,因此无法应用于大多数涉及高分辨率图像的图像恢复任务。

本文基于改进型的 Transformer 网络架构,通过改进的注意力机制以及像素重塑模块,有效地降低了 Transformer 网络的计算复杂度,使得模型能够支持更高分辨率的图像输入,为模型提供了更好的图像细节,使得模型在图像去噪上的达到了优于以上传统方法和基于卷积神经网络的效果。

## 2. 相关工作

图像去噪算法的原理可以概括为通过图像本身的结构和特性将图像噪声与图像信息分开。传统的图像去噪算法包括空间像素特征去噪、变换域特征去噪、基于图像稀疏度的去噪方法、基于字典的学习方法和基于图像先验的方法等。在传统算法中,效果最好的算法一般被认为是 BM3D 算法[5], BM3D 算法结合了空间域算法的非局部滤波算法和变换域滤波算法,它在视觉效果和图像质量的客观评价方面都很出色。空间域的非局部平均算法和频域的小波变换去噪都是有效的,自然的想法是将这两种算法结合起来以达到更好的效果,而 BM3D 就是这样一种去噪方法。在传统的图像去噪领域, BM3D 既可用于视觉效果,也可用于客观的图像质量评估值。能达到最佳效果,被普遍认为是传统去噪算法中的最佳算法。非局部均值的滤波器与波浪类似, BM3D 首先将图像划分为较小的像素块,在选择参考切片后,它寻找其相似的像素块,将找到的像素块堆叠成一个三维块,并将这个三维块转换为一个变换域中的滤波,然后对三维块进行反变换,将其转换回空间域,最后将滤波后的三维块加入到空间域中。

卷积神经网络是由 LeCun 在 1989 年提出的,并首先在图像识别领域取得了巨大的成功。卷积神经网络首先在图像分类任务中取得了非常好的成绩。在这之后卷积神经网络被引入了许多研究领域,如图像去噪、图像分割等。该神经网络已经成为深度学习图像处理中最主流的网络。卷积神经网络可以从大规模的训练数据中自动学习输入数据的特征,并将学习的结果泛化到相同的未知数据中。基于深度学习的算法中比较成功且经典的算法有 2017 年提出的去噪卷积神经网络算法(Denoising convolutional neural networks, DnCNN) [6], 该算法以将非常深的架构、学习算法和正则化方法用于图像去噪的进展,使用残差学习策略, DnCNN 隐式去除隐藏层中潜在的干净图像。这个属性促使训练一个单一的 DnCNN 模型来处理几个通用的图像去噪任务,例如高斯去噪、单图像超分辨率和 JPEG 图像去块。一种更快更灵活的去噪卷积神经网络(Fast and Flexible Image Denoising Convolutional Neural Network, FFDNet) [7]是在 DnCNN 网络之后提出的算法, DnCNN 通过使用批量归一化和残差学习可以有效地去除均匀的高斯噪声,它可以在一定的噪声水平范围内抑制噪声。然而,真实的噪声并不是均匀的高斯噪声,它与信号有关,与每个颜色通道相关,并且是不均匀的,可能随空间位置的变化而变化[8]。在这种情况下, FFDNet 使用噪声估计图作为输入,在抑制均匀噪声和保留细节之间进行权衡,从而应对更复杂的真实场景。

## 3. 图像去噪算法

本节首先介绍图像去噪问题的模型,其次提出的去噪算法的流程以及改进后的 Transformer 网络的结构和内部模块,最后介绍网络训练相关的损失函数。

### 3.1. 图像去噪模型

图像被噪声污染的退化过程可以转化为一个退化函数和一个加性噪声叠加而成,数学上可以表示为:

$$g(x, y) = h(x, y) * f(x, y) + n(x, y) \quad (1)$$

上式中,  $h, f, n$  分别表示图像退化函数, 原始图像和加性噪声。尽管硬件设备受到各种噪声因素的影响, 但其综合分布非常接近于加性高斯白噪声, 因此基于高斯白噪声的去噪算法具有良好的鲁棒性, 且高斯噪声在空间域和频域上都易于处理。

### 3.2. 算法网络结构

本工作提出的算法的主要目标是针对图像去噪任务设计出一个基于 Transformer 模型的且能够处理高分辨率图像输入。算法整体的网络架构如图 1 所示:

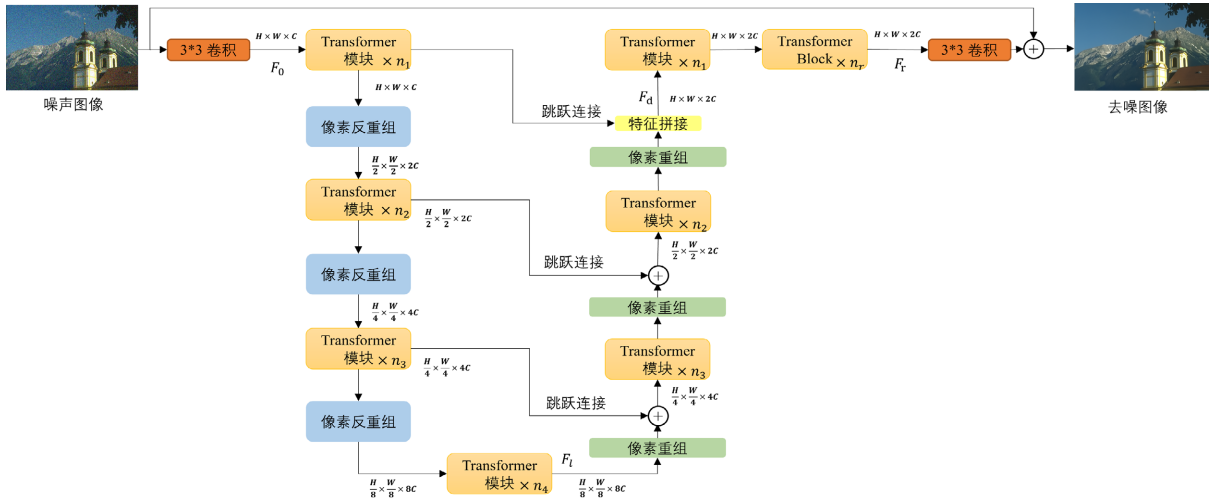


Figure 1. Algorithm network structure  
图 1. 算法网络结构

本文提出算法的整体流程如下:

首先针对给定的退化的图像  $I \in \mathbb{R}^{H \times W \times C_0}$ , 其中  $H \times W$  表示图像的空间维度,  $C_0$  表示输入图像的通道数。模型首先通过一个卷积层将输入图像嵌入到低维度的特征  $F_0 \in \mathbb{R}^{H \times W \times C}$ ,  $C$  表示特征层的维度。接下来, 这些浅层特征通过四级的对称编码器 - 解码器转换为深层特征  $F_d \in \mathbb{R}^{H \times W \times 2C}$ 。其中每一级的编码器 - 解码器包含多个基础的 Transformer 模块, 其数量自顶而下逐渐增加以保证模型的运算效率。从高分辨率的输入开始, 编码器逐级减少空间大小, 同时扩大特征通道的容量。而解码器以低分辨率的潜在特征  $F_l = \mathbb{R}^{\frac{H}{8} \times \frac{W}{8} \times 8C}$  作为输入, 逐步恢复高分辨率的显示。最后, 对于精炼后的特征  $F_r = \mathbb{R}^{H \times W \times 2C}$  采用卷积层来生成残差的图像  $I' \in \mathbb{R}^{H \times W \times C_0}$  并与退化后的图像进行相加, 以获得去噪恢复后的图像:  $\hat{I} = I + I'$ 。

#### 3.2.1. 位置编码

Transformer 模型缺乏一种解释输入序列中单词顺序的方法, 这与序列模型不同。为了处理这个问题, 转化器在编码器层和解码器层的输入中增加了一个额外的位置编码向量。其尺寸与嵌入特征的尺寸相同。这个向量采用了一种非常独特的方法让模型学习这个值。这个向量可以确定当前节点在图像中位置, 或者图像中不同节点的距离。这个位置向量的具体计算方法有很多, 本文中使用的位置编码的计算方法如下:

$$PE(pos, 2i) = \sin \frac{pos}{10000^{\frac{2i}{d_{model}}}}$$

$$PE(pos, 2i + 1) = \cos \frac{pos}{10000^{\frac{2i}{d_{model}}}}$$
(2)

其中,  $pos$  表示当前表示节点在图像中的位置,  $i$  是指向量中每个值的索引, 在偶数位置使用正弦编码, 在奇数位置使用余弦编码。

计算位置编码得到位置编码向量后, 将位置编码向量与嵌入特征向量进行相加, 作为输入送到模型的下一层。

### 3.2.2. 归一化模块

由于不同的特征具有不同数量级的数据, 它们对线性组合后的结果的影响所占比重就很不相同, 数量级大的特征显然影响更大。进行归一化可以协调在特征空间上的分布, 更好地进行梯度下降。

在 Transformer 网络中, 每一个子层在进行后续操作之前都有一个归一化模块, 归一化模块有很多种类, 但其目的都是统一的, 就是将输入转化为以均值为 0 方差为 1 分布的数据, 以便模型在训练过程中更好地收敛, 并具有更好的泛化性能, 且在将数据送入激活函数之前进行归一化, 会减少输入数据落在激活函数的饱和区的概率。

本文使用的模型采用的是层归一化(Layer Normalization)的方式, 通过在每一个样本上面计算均值和方差, 即在特征通道维度上做归一化, 其计算公式如下:

$$LN(x_i) = \alpha \times \frac{x_i - \mu_L}{\sqrt{\sigma_L^2 + \varepsilon}} + \beta \quad (3)$$

其中,  $LN(\cdot)$  表示层归一化函数;  $\alpha, \beta$  为可学习参数, 为还原参数, 在一定程度上保留原数据的分布;  $\mu_L, \sigma_L$  分别为输入  $x_i$  在特征通道维度上的均值和方差;  $\varepsilon$  是为了保证计算有效性的一个很小的常数。

### 3.2.3. 改进的注意力模块

注意力模块是 Transformer 模型最核心的内容, 每一个编码器中包含两层, 一个是自注意力层, 一个是前馈神经网络, 自注意力机制能够帮助当前的节点关注到全局的语义信息。原始 Transformer 模型的自注意力模块会计算出三个新的向量, 把这三个向量分别称为 Query、Key、Value, 这三个向量是用嵌入特征向量与一个矩阵相乘得到的结果, 这个矩阵是随机初始化的, 其值在误差反向传播的过程中会一直进行更新。然后需要计算的自注意力的分数值, 该分数值决定了当我们在某个位置节点对对输入部分的其他节点的关心程度。这个分数值的计算方法是 Query 与 Key 做点乘, 然后把得到的结果做一个 SoftMax 的计算。得到的结果即是每个节点对于当前位置的节点的相关性大小, 最后把 Value 的值与 SoftMax 的结果进行相乘并进行残差连接。

基于以上的注意力模块, 同时引入多头注意力模块, 同时初始化多组 Query、Key、Value 矩阵, 同时进行自注意力的计算, 并对计算结果进行拼接, 作为下一层的输入。通过多头注意力机制, 输入可以多次结合上下文的语义信息对当前的特征进行编码, 得到语义信息更加丰富的全局嵌入特征。

Transformer 的主要计算的主要开销来自注意力层, 常规的 Transformer 的计算和存储复杂度随输入图像的空间分辨率的增长呈二次增长。因此, 这对包含高分辨率图像的图像去噪任务来说是不可实行的。为了缓解这个问题, 本工作使用了改进的注意力机制, 它具有线性的复杂度。其工作原理是在通道维度上应用注意力机制而不是空间维度, 计算跨通道的协方差以生成一个包含全局上下文信息的注意力矩阵。

其计算过程如下: 改进的自注意力模块的输入是经过层归一化后的张量  $\mathbf{X} \in \mathbb{R}^{\hat{H} \times \hat{W} \times \hat{C}}$ , 首先通过一个  $1 \times 1$  的卷积聚合像素级跨通道上下文信息, 接着通过一个  $3 \times 3$  深度卷积来编码特征通道的空间上下文, 最终得到 Query (Q)、Key (K)、Value (V)。

$$\mathbf{Q} = W_d^Q W_p^Q \mathbf{X}, \mathbf{K} = W_d^K W_p^K \mathbf{X}, \mathbf{V} = W_d^V W_p^V \mathbf{X} \quad (4)$$

其中,  $W_p^{(\cdot)}$  表示  $1 \times 1$  的逐点卷积,  $W_d^{(\cdot)}$  表示  $3 \times 3$  的深度卷积。

其次，我们重塑 Query 和 Key 的投影，使它们的点积交互产生一个大小  $\mathbb{R}^{\hat{c} \times \hat{c}}$  转置的注意力矩阵，通过逐点卷积加上深度卷积的形式以及在特征维度上作用注意力模块可以有效地减少计算量，以减少高分辨率图像在模型前向计算过程中的计算量。改进的注意力模块具体计算流程如图 2 所示：

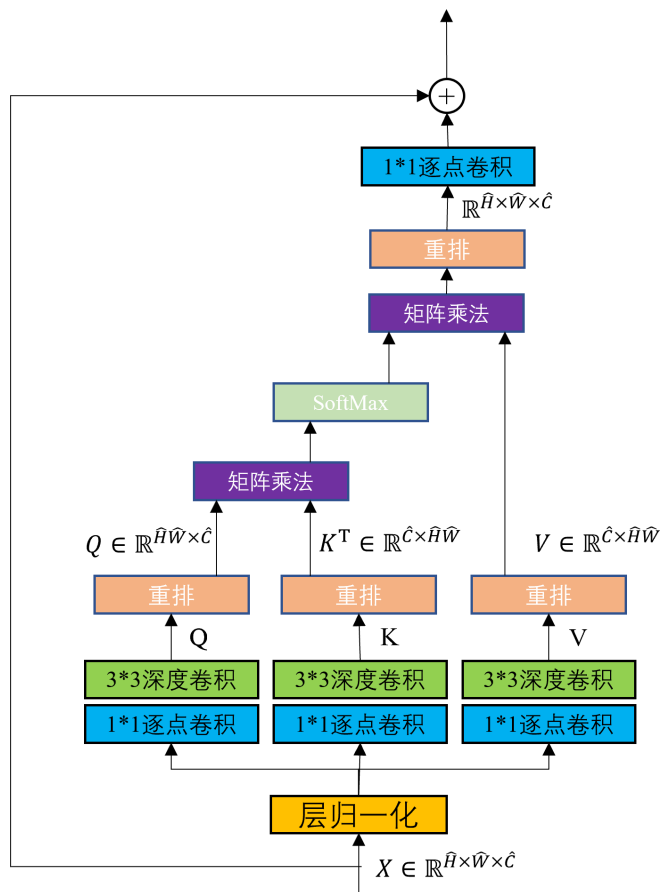


Figure 2. Improved attention module  
图 2. 改进的注意力模块

### 3.2.4. 像素重组模块

像素重组模块(PixelShuffle)是一种上采样的方法，可以替代传统的利用插值的方法和转置卷积的方法实现对缩小后的特征图进行有效的放大。像素重组是一种用于超分辨率模型的操作[3]，用于以  $\frac{1}{r}$  的步长实现高效的亚像素卷积。具体而言，它以形状张量重新排列元素  $(*, C \times r^2, H, W)$  到形状张量  $(*, C, H \times r, W \times r)$ 。与之对应的为像素反重组模块(PixelUnshuffle)，即一种替代池化的下采样方式，它以形状张量重新排列元素  $(*, C, H \times r, W \times r)$  到形状张量  $(*, C \times r^2, H, W)$ 。

### 3.3. 损失函数

本模型训练使用的损失函数为平均绝对误差(Mean absolute error, MAE)损失函数，通过最小化平均绝对损失函数[9]，对模型的参数进行迭代训练，像素级的平均绝对误差损失函数的计算公式为：

$$L_{MAE} = \frac{1}{WH} \sum_{i=1}^W \sum_{j=1}^H |x_{i,j} - \hat{x}_{i,j}| \tag{5}$$

其中,  $W, H$  分别为图像的宽度和高度,  $x_{i,j}$  表示原始图像在  $(i, j)$  的像素点值,  $\hat{x}_{i,j}$  表示模型输出的去噪图像在  $(i, j)$  位置的像素点值。

## 4. 实验结果与分析

### 4.1. 数据集

#### 4.1.1. 训练集

在训练过程中, 噪声图像的噪声为均值为零的独立同分布高斯噪声, 并使用 DIV2K、Flickr2K、Waterloo 数据集作为训练数据集[10]。DIV2K 数据集是用于图像恢复任务的高质量(2 K 分辨率)彩色图像数据集, DIV2K 数据集包含 800 张验证图像、100 张验证图像和 100 张测试图像; Flickr2K 数据集有 2650 张 PNG 图片, 包含人物、动物、风景; Waterloo 数据集包含 4744 张参考图像, 20 个失真类型。对于含噪图像去除的噪声主要是已知噪声等级的高斯白噪声, 本文将噪声分布划分为低、中、高三等级, 即标准差为  $\sigma = 15, 25, 50$  三个噪声等级进行实验。

#### 4.1.2. 测试集

本工作对使用加性高斯白噪声生成的合成基准数据集进行去噪实验, 用于测试本文算法效果的测试集分别选取的是灰度数据集(BSD68, Set12, Urban100)和彩色数据集(CBCD68, Kodak24, McMaster)数据集。测试集图像数据和训练集图像数据互不干扰, 相互独立。本文算法对此数据集进行去噪测试, 通过评价指标来衡量去噪算法的性能优劣。

### 4.2. 评价指标

目前在图像去噪任务中, 评价图像质量好坏可以通过测得的图像峰值信噪比(Peak Signal to Noise Ratio, PSNR)指标进行判定[11]。PSNR 指标的数值越大代表图像质量越好, 图形越清晰; 反之则表示图像越模糊, 视觉效果越差。

#### 峰值信噪比(PSNR)

峰值信噪比是图像处理领域中最常用的评价指标之一, 其利用像素点之间的误差来评价图像质量, 通过表示有效信号和背景噪声的强度之间的关系来表示图像质量:

$$\text{PSNR} = 10 \log_{10} \frac{\text{MAX}_1^2}{\text{MSE}} \quad (6)$$

其中,  $\text{MAX}_1$  是可以作为图像像素使用的最大像素值, 若待评价图形为宽度为  $n$  的单通道图像, 则  $\text{MAX}_1 = 2^n - 1$ , MSE 为均方误差:

$$\text{MSE} = \frac{1}{m * n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|I(i, j) - K(i, j)\|^2 \quad (7)$$

其中,  $I$  是原始图像,  $K$  是去噪后的图像,  $m$  和  $n$  分别是图像的长度和宽度,  $(i, j)$  为图像像素点的索引。

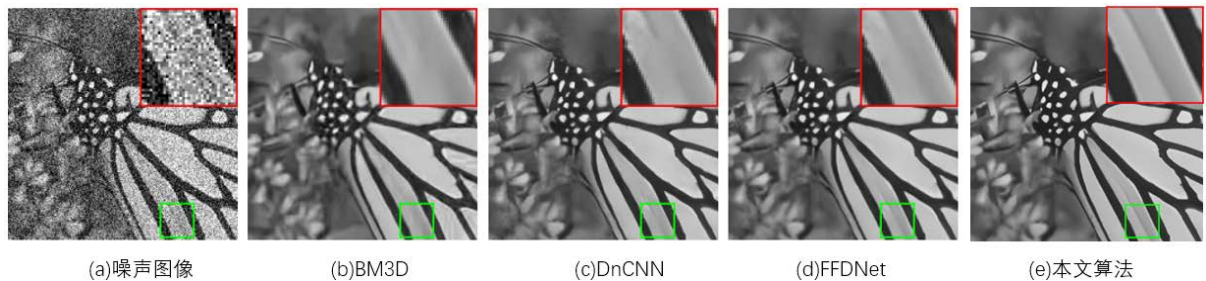
### 4.3. 训练细节

本文使用的模型采用四层编码器 - 解码器结构, 堆叠的 Transformer 模块个数从上到下依次为 4, 6, 6, 8, 对应的通道数分别为 48, 96, 192, 384, 实验模型是基于 Pytorch 框架进行搭建的。训练模型使用的优化器为 AdamW 优化器[12], 初始的学习率为  $3e^{-4}$ , 总的迭代次数为 150 K 次。

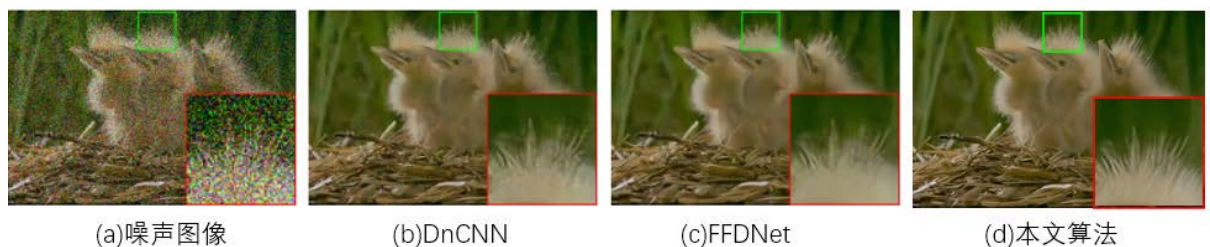
### 4.4. 实验结果

本工作首先分别针对灰度测试集以及彩色数据集上通过渲染的方式对不同算法的去噪效果进行了可

视化, 如图 3 以及图 4 所示。可以直观地看到本文算法的去噪算法在去噪效果和图像细节还原上的效果都要优于其他算法。



**Figure 3.** Comparison of the denoising effect of different methods on the grayscale test set with noise level  $\sigma = 50$   
**图 3.** 噪声水平  $\sigma = 50$  下不同方法在灰度测试集上的去噪效果对比



**Figure 4.** Comparison of the denoising effect of different methods on the color test set at noise level  $\sigma = 50$   
**图 4.** 噪声水平  $\sigma = 50$  下不同方法在彩色测试集上的去噪效果对比

除了定性的分析, 本工作还针对所有测试数据集进行了定量的实验, 表 1 比较了本文提出的去噪算法和主流的去噪算法针对不同程度噪声退化后的图片, 通过峰值信噪比(PSNR)进行衡量算法的效果。可以看出本文提出的算法相对比其他算法具有一定的优越性。

**Table 1.** Comparison of denoising results of each algorithm for degraded images with different noise levels (PSNR values)  
**表 1.** 各算法针对不同噪声程度退化图像的去噪结果比较(PSNR 值)

数据集	噪声程度	BM3D	DnCNN	FFDNet	本文算法
BCD68	$\sigma = 15$	31.07	31.73	31.63	31.81
	$\sigma = 25$	28.57	29.23	29.19	29.37
	$\sigma = 50$	25.60	26.23	26.29	26.45
Set12	$\sigma = 15$	32.37	32.86	32.75	33.06
	$\sigma = 25$	29.97	30.44	30.43	30.74
	$\sigma = 50$	26.72	27.18	27.32	27.66
Urban100	$\sigma = 15$	32.35	32.64	32.40	33.12
	$\sigma = 25$	29.70	29.95	29.90	30.70
	$\sigma = 50$	25.95	26.26	26.50	27.39
CBSD68	$\sigma = 15$	33.52	33.90	33.87	34.28
	$\sigma = 25$	30.71	31.24	31.21	31.66
	$\sigma = 50$	27.38	27.95	27.96	28.46



## Continued

Kodak24	$\sigma = 15$	34.28	34.60	34.63	35.26
	$\sigma = 25$	32.15	32.14	32.13	32.82
	$\sigma = 50$	28.46	28.95	28.98	29.75
McMaster	$\sigma = 15$	34.06	33.45	34.66	35.32
	$\sigma = 25$	31.66	31.52	32.35	33.07
	$\sigma = 50$	28.51	28.26	29.18	29.99

## 5. 总结

本文针对传统方法的不足以及基于卷积神经网络在注意力机制上的缺陷,为了更好地在去噪的同时还原出图像的纹理细节问题,采用了以 Transformer 为基础模型,并针对去噪问题对原模型进行了有效地改进。首先,通过改进的注意力机制以及像素重塑模块,有效地降低了 Transformer 网络的计算复杂度,使得模型能够支持更高分辨率的图像输入;其次,采用了像素重组模型以及残差连接的方式,有效地保留了更为丰富的特征信息,减少了编码器和解码器特征上的语义差距,使得模型在去噪细粒度上的处理效果更为准确。为了测试模型的性能,本文在多个测试集上对模型进行了定性和定量的实验,可以看出模型在图像去噪上达到了优于以上传统方法和基于卷积神经网络方法的效果。虽然本文提出的模型在去噪效果上优于其他算法,但模型的总体推理时间还有很大的提升空间,后续会在模型推理时间以及针对更真实的噪声数据进行模型的优化和探索。

## 基金项目

云南电网有限责任公司科技项目(YNKJXM20210149)。

## 参考文献

- [1] 赵高长, 张磊, 武风波. 改进的中值滤波算法在图像去噪中的应用[J]. 应用光学, 2011, 32(4): 678-682.
- [2] 方莉, 张萍. 经典图像去噪算法研究综述[J]. 工业控制计算机, 2010, 23(11): 73-74.
- [3] 仲点点, 吴杰, 余腾, 王翠, 汤佩卿. 改进小波阈值图像去噪方法[J]. 科学技术创新, 2022(26): 84-87.
- [4] 曾宪华, 李彦澄, 高歌, 赵雪婷. 基于残差编解码器的通道自适应超声图像去噪方法[J]. 电子与信息学报, 2022, 44(7): 2547-2558.
- [5] Zhang, K., Zuo, W., Chen, Y., et al. (2017) Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. *IEEE Transactions on Image Processing*, **26**, 3142-3155. <https://doi.org/10.1109/TIP.2017.2662206>
- [6] Dabov, K., Foi, A., Katkovnik, V., et al. (2007) Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. *IEEE Transactions on Image Processing*, **16**, 2080-2095. <https://doi.org/10.1109/TIP.2007.901238>
- [7] Zhang, K., Zuo, W. and Zhang, L. (2018) FFDNet: Toward a Fast and Flexible Solution for CNN-Based Image Denoising. *IEEE Transactions on Image Processing*, **27**, 4608-4622. <https://doi.org/10.1109/TIP.2018.2839891>
- [8] Shi, W., Caballero, J., Huszar, F., et al. (2016) Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 1874-1883. <https://doi.org/10.1109/CVPR.2016.207>
- [9] 冉光再, 徐雷, 李大双, 郭战岭. 基于图像差分 and 残差学习的 PCB 图像去噪算法[J]. 激光与光电子学进展, 2022, 59(12): 123-130.
- [10] 苏赋, 吕沁, 罗仁泽. 基于深度学习的图像分类研究综述[J]. 电信科学, 2019, 35(11): 58-74.
- [11] He, X.K., Wang, C., Zheng, R.Y., Sun, Z.B. and Li, X.W. (2022) GPR Image Denoising with NSST-UNET and an Improved BM3D. *Digital Signal Processing*, **123**, Article No. 103402. <https://doi.org/10.1016/j.dsp.2022.103402>
- [12] Li, P.L., Liang, J.L., Zhang, M.H., Fan, W. and Yu, G.Y. (2022) Joint Image Denoising with Gradient Direction and Edge-Preserving Regularization. *Pattern Recognition*, **125**, Article No. 108506. <https://doi.org/10.1016/j.patcog.2021.108506>