

融合多注意力机制的生成对抗网络图像去模糊

李秋良

新疆师范大学数学科学学院, 新疆 乌鲁木齐

收稿日期: 2026年2月7日; 录用日期: 2026年3月7日; 发布日期: 2026年3月16日

摘要

生活中的图像模糊通常由相机抖动、物体运动等多因素造成, 且模糊核未知, 这使得图像去模糊成为病态逆问题。现有的深度学习方法虽能取得较好效果, 但在处理真实、复杂的模糊时, 往往存在细节恢复不足和结果过于平滑的问题。为此, 本文提出了一种新颖的融合多注意力机制的生成对抗网络。该网络设计了融合自适应特征激励策略与增强通道注意力方法的多尺度金字塔架构, 同时提出了新的多尺度空间特征增强模块增强模型特征提取能力。我们在不同基准数据集(如GoPro和HIDE)上进行了大量实验, 定性与定量结果均表明, 本方法在恢复锐利边缘和真实纹理方面优于传统神经网络方法。

关键词

深度学习, 注意力机制, 生成对抗网络, 图像去模糊

Multi-Attention Fusion Generative Adversarial Network for Image Deblurring

Qiuliang Li

School of Mathematical Sciences, Xinjiang Normal University, Urumqi Xinjiang

Received: February 7, 2026; accepted: March 7, 2026; published: March 16, 2026

Abstract

Image deblurring represents a challenging ill-posed inverse problem, typically caused by multi-factorial origins such as camera shake and object motion with unknown blur kernels. While current deep learning approaches achieve considerable performance, they often exhibit limitations in reconstructing fine details and tend to produce over-smoothed outputs when handling real-world, complex blur patterns. To mitigate these issues, this paper introduces a novel generative adversarial network

incorporating a multi-attention mechanism. The proposed framework employs a multi-scale pyramid architecture that integrates an adaptive feature excitation strategy with an enhanced channel attention methodology, alongside a newly designed multi-scale spatial feature enhancement module to augment the model's representational capacity.

Keywords

Deep Learning, Attention Mechanism, Generative Adversarial Network, Image Deblurring

Copyright © 2026 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

图像去模糊属于计算机视觉中的底层视觉任务，其目的是在减少目标图片的信息丢失的基础上，从低质量图像中恢复清晰度更高的图像；模糊原因来自图像捕捉过程中受相机抖动或场景内运动等，其主要包含运动模糊、失焦模糊、高斯模糊和混合模糊。借助扩散模型方程可将模糊过程转化为如下公式：

$$\frac{\partial u}{\partial t} = k(t) \Delta u = k(t) \left(\frac{\partial^2 u(x, y, t)}{\partial x^2} + \frac{\partial^2 u(x, y, t)}{\partial y^2} \right)$$

其中， $k(t)$ 是与时间相关的模糊扩散系数，其用来控制模糊扩散的速度， Δu 是进行图像各向同性模糊扩散的拉普拉斯算子；方程的解 u 表示不同模糊情景下清晰图像随时间 t 模糊的结果。

基于上述背景，本文提出了融合注意力机制的生成对抗网络图像去模糊模型，提升图像去模糊的效果。本文的主要贡献如下：

1. 提出了一种基于多注意力机制融合生成对抗网络的图像去模糊模型，提升了图像去模糊的效果。
2. 通过在多个公开数据集上进行实验，验证了所提方法在图像去模糊任务中的优越性能。

2. 相关工作

2.1. 传统图像去模糊

传统图像去模糊方法通过建模图像退化过程的先验信息，并利用优化算法来求解逆问题。其思想是将模糊图像 I_b 看作是清晰图像 I_s 与模糊核 k 的卷积，并加上噪声 n ，即 $I_b = I_s \otimes k + n$ 。

模糊核估计和非盲去模糊是常见的传统图像去模糊方法。在模糊核估计方面，主要包含基于边缘预测与基于显著性的方法。估计出模糊核后非盲去模糊转化为解卷积问题，常用维纳滤波[1]、Richardson-Lucy 算法[2]以及基于稀疏先验的正则化优化方法[3]。

尽管传统方法具有深厚数理基础以及较强可解释性。然而，由于手工设计先验在应对复杂场景模糊去除时，其先验假设会失效，从而导致估计出的模糊核不准确问题，使得产生振铃效应和不自然的平滑效果。

2.2. 基于生成对抗网络的图像去模糊

生成对抗网络在图像去模糊领域取得了颠覆性的进展。

Kupyn [4]成功验证了 GAN 通过学习模糊 - 清晰图像对之间的映射关系，能够恢复出视觉感知上更

清晰的图像，超越了传统优化方法的局限。Neji [5]将 CycleGAN 的对抗学习能力精准适配于文档图像盲去模糊任务。该方法首次实现了如运动模糊、失焦模糊等条件下对文档图像的恢复。随后，Zhao [6]对经典的 CycleGAN 框架进行了针对盲图像去模糊任务的改造，提出了轻量级域转换单元(LDCU)作为生成器核心，通过引入基于批量归一化(BN)的元结构和轻量级编码器-解码器(LED)设计，取代了原 CycleGAN 中不适合去模糊的实例归一化(IN)和复杂网络结构，有效解决了直接使用 CycleGAN 导致的色彩失真和伪影问题。Fanous [7]针对连续运动扫描显微成像的特殊性结合 GAN 的生成对抗学习机制设计了图像去模糊网络。其生成器采用 U-Net 架构，通过跳跃连接将编码器的低级特征与解码器的高级特征融合，从而在去模糊过程中有效保留细胞边界、纹理等细节信息；判别器则使用 PatchGAN 结构，专注于评估图像局部区域的真实性，迫使生成器恢复出更逼真的高频成分。

因而，GAN 已将图像去模糊带入了一个数据驱动、端到端优化的新阶段，但仍面临训练稳定性等挑战。

2.3. 基于注意力机制的图像去模糊

近年来，注意力机制在图像去模糊研究中受到了广泛关注，并逐渐成为提升去模糊模型性能的关键技术之一。

该机制受人类视觉系统中注意力分配机制的启发，能够引导神经网络自适应地聚焦于图像中信息丰富的图像块。Chen [8]设计了结合通道注意力和像素注意力的多尺度去模糊网络模型，使网络更关注模糊区域和重要通道信息。通道注意力通过全局平均池化和卷积层计算通道权重；像素注意力进一步强化对局部模糊区域的关注。

随着自注意力机制在底层视觉任务中的突破。Cui [9]通过设计高效的空間注意力模块(SAM)和频率注意力模块(FAM)，分别增强模型在空间域和频率域的模糊去除能力。Lee [10]在注意力模块中同时引入通道注意力和空间注意力。其根据输入特征自适应调整注意力分支和非注意力分支的权重，同时使用深度过参数化卷积替代普通卷积，增强特征表达能力。

Zhang [11]设计多尺度特征交叉融合模块融合编码器中不同层级的特征，增强多尺度特征表达能力。其设计的区域自注意力模块(RSAM)通过使用条带池化(Strip Pooling)捕捉水平和垂直方向的非均匀模糊区域，赋予不同模糊区域不同权重。并融合深度残差小波变换模块将空间特征转换到频域(小波域)，从而利用高频子带恢复边缘和纹理细节来提升图像去模糊效果。

注意力机制的引入使得图像去模糊模型能够更好地捕捉长距离依赖关系和局部细节，从而显著提升了去模糊效果。但其计算复杂度与更高效的注意力机制去模糊模型仍需要探索。

2.4. 基于扩散模型的图像去模糊

近年来，扩散模型的先验学习和细节生成能力使其迅速被引入图像去模糊领域。其将模糊过程建模为从随机噪声向清晰图像逐步去噪的过程。

Chen [12]提出了一种基于扩散模型的图像去模糊方法，其将扩散模型作为潜在空间图像先验生成器。其在潜在空间中生成富含纹理细节的特征，并且融合交叉注意力机制引导 Transformer 主干网络进行去模糊。Wu [13]通过隐式扩散模型模拟连续运动轨迹，生成多样化的模糊图像，通过增强去模糊模型的训练数据来提升模型去模糊的泛化能力。

Feng [14]将扩散模型作为精细化残差生成器来进行图像去模糊。模型首先用一个轻量级网络进行预去模糊，得到一个基础结果；然后扩散模型通过学习“清晰图像”与“预去模糊结果”之间的细节差异提升去模糊图片的质量。Li [15]利用分数扩散模型学习清晰图像的概率分布，通过反向随机微分方程逐

步重建清晰图像。除此以外, 其还设计了适用于真实世界场景且无需成对训练数据的去模糊方法。Xie [16] 使用扩散模型从单张模糊图像运动先验特征, 模拟事件相机提供的运动信息, 普通相机也能获得事件相机的去模糊优势, 而无需真实事件输入。

总体而言, 基于扩散模型的图像去模糊方法通过其强大的生成能力, 显著提升了去模糊图像的质量和细节还原水平, 但是其高计算成本和复杂的训练过程需要优化。

3. 本文模型

本文模型框架如图 1 所示, 生成器将模糊图像去模糊, 然后通过判别器反馈的信息优化生成器的图像去模糊性能。

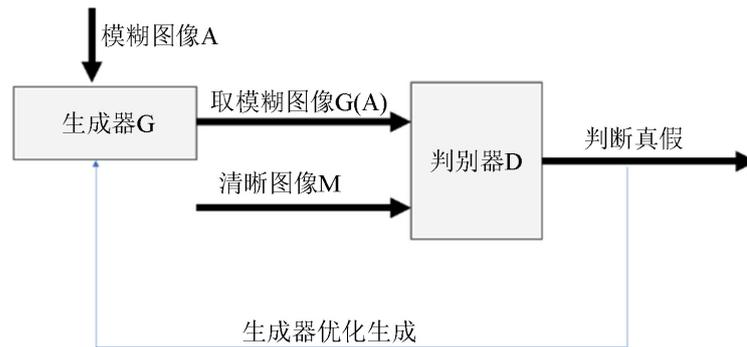


Figure 1. Model flowchart
图 1. 模型流程

3.1. 生成对抗网络的生成器

生成器流程图

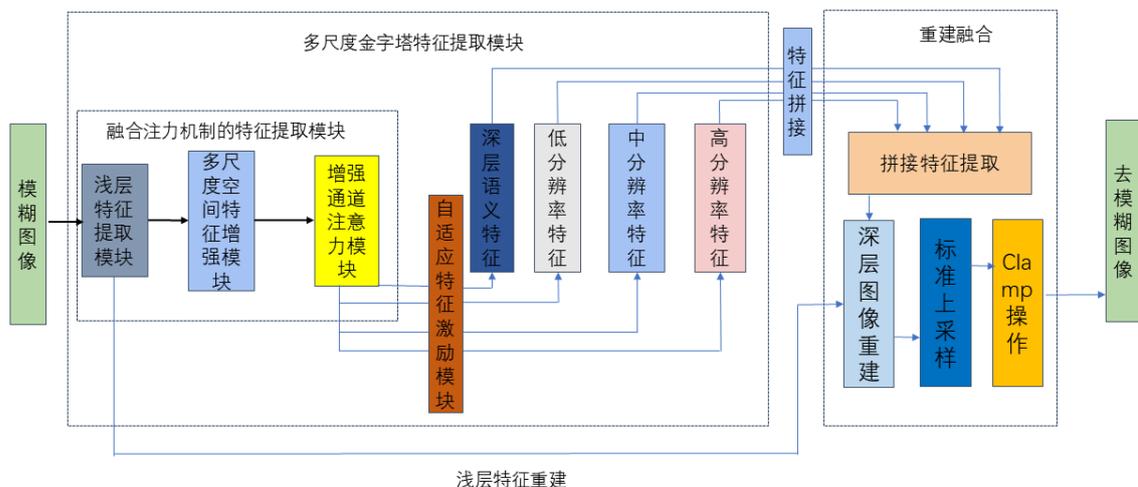


Figure 2. Generator flowchart
图 2. 生成器流程图

本任务的生成器主要由新设计的融合多注意力机制的金字塔结构特征提取模块以及包含多头上采样部件的图像重建模块构成。其流程如图2所示。

具体来说,模糊图像首先通过多尺度金字塔特征提取模块提取不同尺度特征,该模块首先通过浅层特征提取模块提取为后续特征重建准备的浅层特征,随后通过多尺度空间特征增强模块(Multi-scale Spatial Feature Enhancement Module)和增强通道注意力模块(Enhanced Channel Attention)构成的融合注意力机制的特征提取模块提取浅层输入特征 F_{input}^c 的5个尺度的特征信息 $EAFPN_{feature}^c$ 。其数学表达如下所示:

$$EAFPN_{feature}^c = EAFPN_{model}(F_{input}^c, \text{block0}, \text{block1})$$

其中, $EAFPN_{model}$ 表示基于金字塔结构创建的多尺度特征提取模块,其浅层特征提取模块由二维卷积块 block0 和 block1 实现。

随后通过包含深层卷积块和混合上采样的自适应特征模块将不同尺度的特征恢复到同一空间维度,以此作为图像去模糊的基础。

$$Upsample_{feature}^c = Upsample_{new}(FPN_{scale_factor}, \text{head_net}_{feature}^c)$$

随后,使用张量拼接函数,融合不同层统一空间维度特征信息,进行初步的信息重建。其数学表达如下所示:

$$feature = \text{cat}(FPN_{head}_{feature}^c, \text{dim} = 1)$$

最后,通过两次 smooth 操作对图像的信息整合。该操作通过卷积网络处理后并且通过基础的张量拼接操作实现:

$$smooth_{feature}^c = smooth_{net}(F_{input}^c, \text{norm_layer}, \text{upsample}_{smoothed})$$

其中 $upsample_{smoothed}$ 表示基本的上采样操作。在经过上述处理后,通过对 $smooth_{feature}^c$ 进行 clamp 操作获得去模糊的图像。

1. 多尺度空间特征增强模块

多尺度空间特征增强模块通过对输入特征进行多尺度的空间卷积操作,获得不同层次的细致特征信息。其流程如图3所示。

多尺度空间特征增强模块将输入特征转化为三个分支,在各自分支提取细粒度不同的特征,来提升模型特征提取的能力。

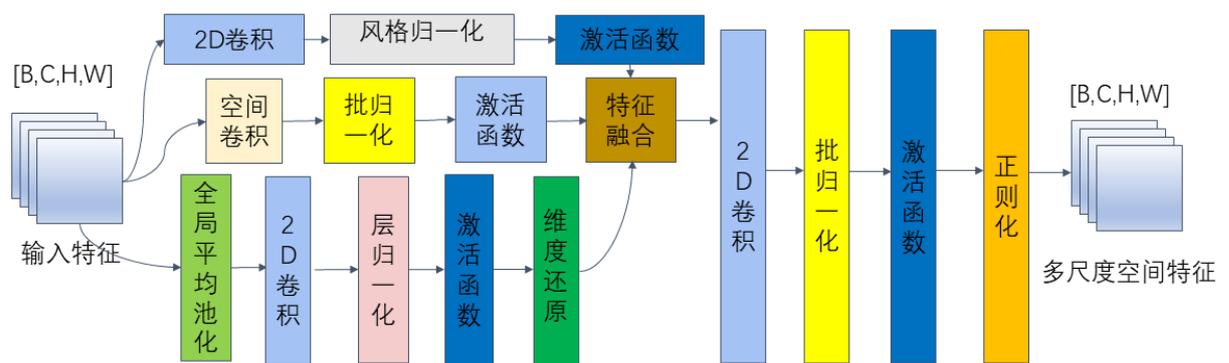


Figure 3. Multi-scale spatial feature enhancement module

图3. 多尺度空间特征增强模块

首先, 输入特征 X_{input} 通过 2D 卷积操作获得不同尺度的特征信息 X_{conv} , 随后通过风格归一化和激活函数获得高频细节信息, 为后续图像重建提供纹理特征, 其数学表达如下所示:

$$\begin{cases} \text{Conv1x1:} & Y^{(1)} = W * X, \quad W \in \mathbb{R}^{C_{out} \times C_{in} \times 1 \times 1} \\ \text{InstanceNorm:} & Y^{(2)} = \gamma \frac{Y^{(1)} - \mu(Y^{(1)})}{\sigma(Y^{(1)})} + \beta \\ \text{ReLU:} & Y^{(3)} = \max(0, Y^{(2)}) \end{cases}$$

其中, Conv1x1 表示 1×1 卷积操作, W 为卷积核权重, InstanceNorm 表示实例归一化操作, μ 和 σ 分别为每个样本每个通道的均值和标准差, γ 和 β 为可学习的缩放和偏移参数; ReLU 表示整流线性单元激活函数, 用于引入非线性变换; C_{in} 和 C_{out} 分别表示输入和输出通道数。

其次, 使用空洞卷积获得不同尺度下的局部特征信息。其数学表达如下所示:

$$\begin{cases} x_2 = \mathcal{F}_{\text{atrous}}^{(1)}(x) = \text{ReLU}(\text{BN}(W_2 *_{d_1} x)) \\ x_3 = \mathcal{F}_{\text{atrous}}^{(2)}(x) = \text{ReLU}(\text{BN}(W_3 *_{d_2} x)) \\ x_4 = \mathcal{F}_{\text{atrous}}^{(3)}(x) = \text{ReLU}(\text{BN}(W_4 *_{d_3} x)) \end{cases}$$

其中, $\mathcal{F}_{\text{atrous}}^{(i)}$ 表示第 i 个空洞卷积分支, W_i 为对应的卷积核权重, $*_{d_i}$ 表示空洞率为 d_i 的空洞卷积操作, BN 表示批量归一化操作, ReLU 表示整流线性单元激活函数。

接着, 通过全局平均池化提取图像块的局部级语义特征, 然后通过卷积、归一化和激活函数增强特征表达能力, 使全局特征与其他分支的特征图尺寸保持一致, 其数学表达如下所示:

$$\begin{cases} x_5 = \text{Interpolate}(\mathcal{F}_{\text{global}}(x), \text{size} = (H, W), \text{mode} = \text{bilinear}) \\ \mathcal{F}_{\text{global}}(x) = \text{ReLU}(\text{LayerNorm}(\text{Conv}_{1 \times 1}(\text{AdaptiveAvgPool2d}_1(x_5)))) \end{cases}$$

最终将不同尺度分支的特征图在通道维度上进行拼接, 然后通过 1×1 卷积进行通道降维和特征整合, 其数学表达如下所示:

$$\begin{cases} Y = \mathcal{F}_{\text{conv_out}}(\text{Concat}(x_1, x_2, x_3, x_4, x_5; \text{dim} = 1)) \\ \mathcal{F}_{\text{conv_out}}(Z) = \text{Dropout}_{0.5}(\text{ReLU}(\text{BN}(W_{\text{out}} * Z))) \end{cases}$$

在图像去模糊任务中, 该模块通过显式建模不同空间尺度上的特征表达, 使网络能够有效应对由相机抖动、物体运动等造成的、在图像中分布不均且尺度各异的模糊模式, 为后续的特征融合与清晰图像重建提供鲁棒且信息丰富的多尺度特征基础。

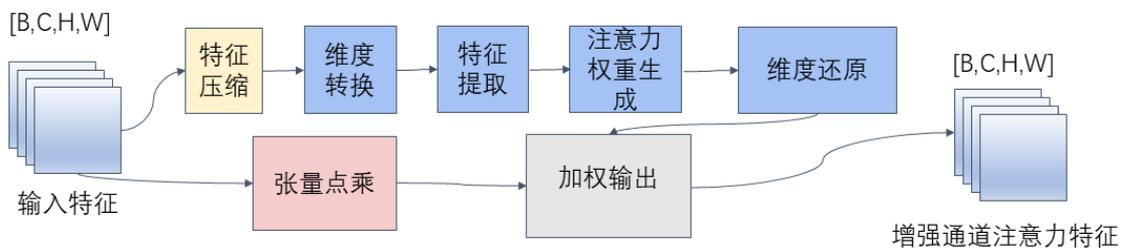


Figure 4. Enhanced channel attention
图 4. 增强通道注意力

2. 增强通道注意力模块

增强通道注意力模块目的是在网络浅层自适应获得通道特征；其流程如图 4 所示。

其功能是在浅层提取重要的通道信息为图像信息恢复做准备。

首先，使用全局平均池化计算输入特征逐通道的像素位置平均值，从而把单通道的二维特征压缩为一维标量特征。该操作进行求和平均消除输入特征空间位置维度的影响，为后续增强通道注意力权重的提取提供全局信息。其数学表达如下所示：

$$X_{avepool} = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x(i, j), X_{avepool} \in \mathbb{R}^{b \times C \times 1 \times 1}$$

随后，维度转换 *squeeze* 和 *permute* 方法将 4 维张量转换为 3 维张量，为后续特征的卷积提取做准备，其数学表达如下所示：

$$\begin{cases} X_{squeeze} = squeeze(X_{avepool}, dim = -1) \in \mathbb{R}^{b \times C \times 1} \\ X_{transform} = permute(X_{squeeze}, 0, 2, 1) \in \mathbb{R}^{b \times 1 \times C} \end{cases}$$

其中 *squeeze* 是将 $X_{avepool}$ 的最后一个维度的数据移除得到形状为 $[b, c, 1]$ 的 $X_{squeeze}$ 特征，*permute* 之后，将维度转换后的特征张量进行一维卷积 Conv1D 提取局部通道信息，并借助 Sigmoid 激活函数获得注意力权重 X_W 。其数学表达如下所示：

$$X_W = \sigma(\text{Conv1D}(X_{transform}))$$

其中 sigmoid 函数用来归一化处理特征，使得重要通道接近一不重要通道接近零。

最后将原始输入与张量广播后的增强通道注意力权重逐个元素相乘后获得重要通道的注意力信息。其数学表达如公式所示：

$$X_{enhanced} = X \odot \text{Broadcast}(X_{AW}) \in \mathbb{R}^{B \times C \times H \times W}$$

增强通道注意力在图像去模糊中的作用是通过输入特征增强通道注意力信息，从而提升模型对关键特征的关注度。借助这种方式，模型能够更好地捕捉到图像中的细节和纹理信息，提升去模糊效果。

3. 自适应特征激励模块

该模块包含局部特征提取与自适应特征增强部分，局部特征提取由步长为 3 的两个卷积层实现。其流程如图 5 所示：

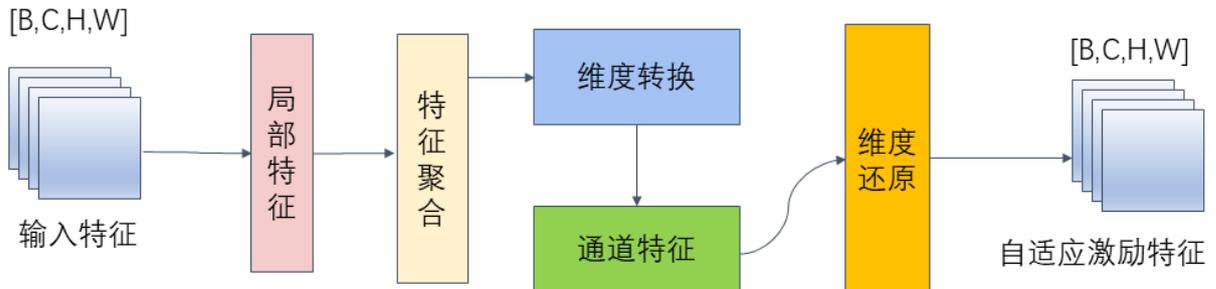


Figure 5. Adaptive feature excitation module
图 5. 自适应特征激励模块

首先通过特征变换操作获得输入特征 X_{input} 的通道权重向量。通过全局平均池化压缩空间维度进行特

征聚合与维度转换，其数学表达如下所示：

$$z_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W X_{\text{input}}^{(c)}(i, j)$$

随后通过两个全连接层深层捕获通道特征，首先通过降维减少参数量，再通过升维恢复原始通道，并使用 ReLU 和 Sigmoid 激活函数引入非线性，其数学表达如下所示：

$$s = \sigma(W_2 \cdot \delta(W_1 \cdot z))$$

其中， W_1 和 W_2 分别为两个全连接层的权重矩阵， δ 表示 ReLU 激活函数。

最后将学习到的通道权重与原始输入特征进行逐通道相乘，实现特征重校准获得自适应激励特征，其数学表达如下所示：

$$\hat{X}^{(c)} = s_c \cdot X_{\text{input}}^{(c)}$$

该模块通过对特征融合过程中的特征进行自适应调节，能够有效地增强关键特征的表达能力，同时抑制冗余信息，从而提升图像去模糊的效果。通过这种方式，模型能够更好地捕捉到图像中的细节和纹理信息，提升去模糊效果。

3.2. 生成对抗网络的判别器

判别器主要将去模糊图像与真实清晰图像进行对比，获得生成器优化信息，帮助生成器优化模糊图像去模糊。本文判别器结构如图 6 所示：

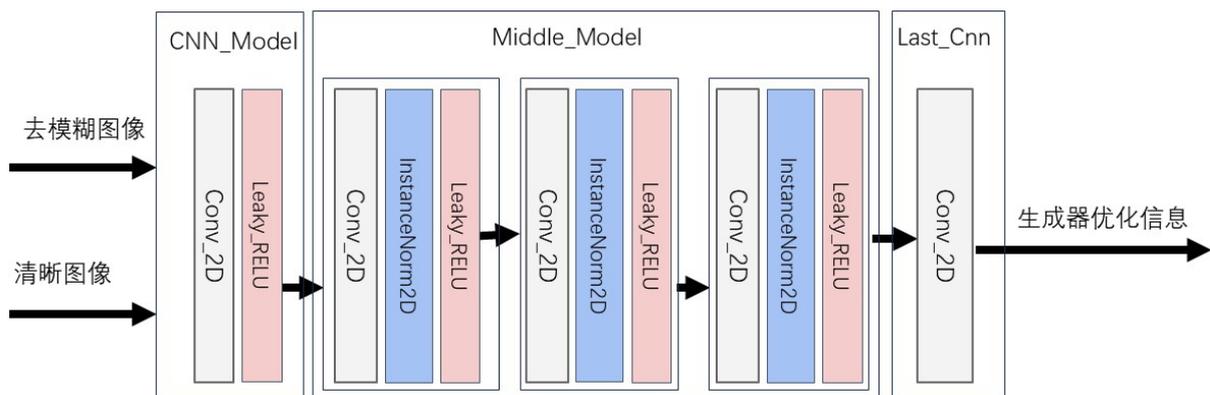


Figure 6. Discriminator network flowchart

图 6. 判别器网络流程图

首先，清晰图像和去模糊图像分别作为输入特征， F_{input}^c 会分别通过卷积神经网络操作获取浅层特征。

$$Shallow_{\text{feature}}^c = Cnn_{\text{model}}(F_{\text{input}}^c, \text{norm_layer}, \text{num_filters})$$

其中， Cnn_{model} 表示由卷积层和激活函数构成的浅层特征提取块，该模块通过简单的卷积操作来获得输入图像的全局特征信息。

其次判别器通过中间特征提取块 $Middle_{\text{Cnn}}$ 提取图像的深层次的局部特征信息。该部分有多个卷积层、批量归一化层以及激活函数复合而成

$$Middle_{\text{feature}}^c = Middle_{\text{Cnn}}(F_{\text{input}}^c, \text{norm_layer}, \text{num_filters})$$

最后通过仅带有二维卷积的卷积神经网络来输出为真实图像的概率反馈给生成器进行图像去模糊的优化。

$$last_{\text{feature}}^c = last_{\text{Cnn}}(F_{\text{input}}^c, \text{num_filters})$$

其中, $last_{\text{model}}$ 表示由卷积层构成的特征提取块。

3.3. 生成对抗网络的损失函数

网络复合损失函数主要包含生成器损失、判别器损失。生成器损失函数主要包含内容损失、对抗性损失和感知损失。

$$L_G = L_{\text{content}} + L_{\text{adversarial}} + L_{\text{perceptual}}$$

其中, L_{content} 表示内容损失, $L_{\text{adversarial}}$ 表示对抗性损失, $L_{\text{perceptual}}$ 表示感知损失。其数学表达如下所示:

$$\begin{aligned} L_{\text{content}} &= \frac{1}{CHW} \|X_{\text{generated}} - X_{\text{target}}\|_1 \\ L_{\text{adversarial}} &= -\mathbb{E}_{z \sim p_z} [\log D(G(z))] \\ L_{\text{perceptual}} &= \sum_{l=1}^L \frac{1}{C_l H_l W_l} \|\phi_l(X_{\text{generated}}) - \phi_l(X_{\text{target}})\|_2^2 \end{aligned}$$

在生成器损失函数中, 内容损失(L_{content})用于衡量生成图像($X_{\text{generated}}$)与目标图像(X_{target})在像素级别的差异, 本文采用 L1 范数计算并除以 CHW 进行归一化。

判别损失($L_{\text{adversarial}}$)通过生成器生成的图像欺骗判别器(D)来计算, 其中 z 是从噪声分布 p_z 中采样的随机向量, $G(z)$ 表示生成器生成的图像, $D(G(z))$ 表示判别器将生成图像判断为真实图像的概率, 模型对概率取负以最小化该损失。感知损失($L_{\text{perceptual}}$)则通过预训练的特征提取网络的多层特征图来计算, 其中 $\phi_l(X)$ 表示输入图像 X 在第 l 层网络信息提取的特征映射, 判别器损失函数主要包含对抗性损失和感知损失。其数学表达如下公式所示:

$$L_D = -\mathbb{E}_{x \sim p_{\text{data}}} [\log D(x)] - \mathbb{E}_{z \sim p_z} [\log(1 - D(G(z)))]$$

该损失函数由两部分组成: 第一部分 $-\mathbb{E}_{x \sim p_{\text{data}}} [\log D(x)]$ 鼓励判别器 D 将真实图像 x 正确识别为真实样本, 其中 x 是从真实数据分布 p_{data} 中采样的图像, $D(x)$ 表示判别器对真实图像的判别概率; 第二部分 $-\mathbb{E}_{z \sim p_z} [\log(1 - D(G(z)))]$ 鼓励判别器将生成图像 $G(z)$ 正确识别为去模糊图像, 其中 z 是从噪声分布 p_z 中采样的随机向量, $G(z)$ 是生成器基于噪声 z 恢复的去模糊图像, $D(G(z))$ 表示判别器对生成图像的判别概率。

4. 实验

本章节, 首先介绍所使用的数据集与具体实验设置细节, 其次通过本文模型与其他模型的对比实验与系统的消融实验, 验证本文提出模型的有效性。

4.1. 对比实验

本实验使用 GOPRO [17]数据集进行模型的训练, 其训练集包含 2103 对模糊(清晰)图像, 测试集包含 1111 对模糊(清晰)图像。为了保证实验可比性与快速验证新模型有效性, 本文沿用与基线模型一致的实验配置, 即采用 GoPro 训练集的一半作为训练集, 测试集的十分之一进行训练, 并在经典运动模糊数

数据集 HIDE [18]和真实模糊数据集(RealBlur) [19]上与其他主流模型架构(多尺度架构、注意力框架)与经典算法进行性能对比。

本文模型与其他去模糊框架在 GOPRO 测试集和 HIDE 数据集上的客观指标对比由表 1 呈现,从表 1 可知,在相同的实验配置下,新模型在 GOPRO 测试集上 PSNR 结果比基线模型提升 0.6861 dB,这一提升验证了模型中多尺度结构的有效性,新模型仍对不同尺度模糊核的适应能力。PSNR 结果比传统的注意力模型提升了 1.1299 dB,同时 SSIM 也有提升。由于传统注意力机制在处理图像时,仅简单计算所有像素区域间的依赖关系。这种方法可让的高频细节(如边缘、纹理)在全局的平滑化信息聚合中,从而导致复原图像在像素精度(PSNR)和结构保真度(SSIM)上均存在优化瓶颈。这表明新模型融合增强通道注意力后更合理的进行特征提取,使得去模糊效果得以保证。与之相比,新模型通过融合增强通道注意力机制,实现了更为合理的特征提取。该机制能够依据特征重要性自适应地校准通道响应,强化对细节敏感特征的输出。这种改进使得模型在复原过程中更好地保持高频内容,因而在去模糊效果上得到了可靠保证。在 HIDE 数据集上,新模型的 PSNR 相较于注意力模型大幅提升 3.1864 dB 的结果,凸显了自适应特征激励模块在去除真实场景运动模糊的独特价值。该模块能根据输入特征,自适应地校准并激励对恢复五官、轮廓等运动细节最为关键的特征响应,这使其在 HIDE 这类以真实运动模糊场景为评估核心的数据集上,优势被放大。

在真实昏暗场景数据集 Real Image 模糊去除的客观评估(表 2)中,新模型的 PSNR 与 SSIM 指标均达到了良好效果,这表明其在光照不足与运动模糊耦合的复杂退化条件下,仍能实现稳定的图像去模糊效果。在子数据集 Real Image R 上,本文模型的 PSNR 值为 30.7217 dB,较基线模型提升了 0.5337 dB;但未达到最优,由于该数据集模糊不仅来自相机抖动,还有物体运动、景深变化引起的散焦等。即图像背景可能因为抖动产生整体模糊,前景物体却因为快速运动出现局部拖影,不同区域的模糊类型都不一样。本文的模型主要针对的是 Gopro 那种合成数据,在处理这种复杂、多变的真实模糊时,其泛化能力有待加强,但其 SSIM 达到了最高的 0.9088,这说明本文模型对复合模糊背景下图像的结构恢复是有效。Real Image J 数据集以真实昏暗场景下的复合模糊为核心特点,新模型所采用的多尺度空间特征增强模块整合不同尺度的空间上下文,使模型能够协同处理低照度下的噪声增强、对比度衰减与动态模糊等多重退化问题;但其客观结果未达到最优,其主要原因未真实图像中的模糊很多是小尺度、局部性的,而本文模型中的增强的通道注意力机制更偏向提取深层、大范围的特征,其会忽略掉一些对恢复细节很重要的浅层特征。这个问题在 Gopro 合成数据上不太明显,因为合成模糊大多是全局性的;但一到真实图像上,这种细节丢失就被放大了,导致在纹理丰富的区域恢复效果不够好,从而拉低了整体的 PSNR 和 SSIM。这说明模型针对真实昏暗场景下的模糊去除需要优化浅层特征提取的能力,但本文模型仍与对比模型相近水平,因而在恢复昏暗场景中易丢失的细节与结构,本文模型仍具有一定的实用性。

Table 1. The comparison of results (PSNR/SSIM) for image denoising in mixed blur scenarios such as GoPro and HIDE
表 1. 在 GoPro 和 HIDE 等混合模糊场景下去除结果对比(PSNR/SSIM)

数据集	评估方法	mul [20]	Baseline [21]	our	Attention [22]	Deblur Gan [4]
HIDE	PSNR	16.5124	25.9518	26.5269	23.3405	24.8124
HIDE	SSIM	0.6802	0.8345	0.8597	0.7736	0.7671
GoPro	PSNR	16.7909	26.3403	27.0264	24.5758	26.0112
GoPro	SSIM	0.7221	0.8640	0.8791	0.7972	0.7886

Table 2. Comparison of objective results (PSNR/SSIM) for real-world image deblurring on the real image dataset
表 2. 在 real image 数据集上的真实场景模糊去除的客观结果(PSNR/SSIM)对比

数据集	评估方法	mul [20]	Baseline [21]	our	Attention [22]	Deblur Gan [4]
real image J	PSNR	21.5040	25.5982	25.5913	26.0613	26.3399
real image J	SSIM	0.7124	0.8060	0.7967	0.7821	0.7971
real inage R	PSNR	24.7256	30.1880	30.7217	31.8160	32.3532
real inage R	SSIM	0.5327	0.8872	0.9088	0.7548	0.9006



Figure 7. (1) Blurred image; (2) Sharp image; (3) Multi-scale model; (4) Baseline model; (5) Proposed model; (6) Attention model; (7) DeblurGAN model

图 7. (1) 模糊图像; (2) 清晰图像; (3) 多尺度模型; (4) 基线模型; (5) 本文模型; (6) 注意力模型; (7) DeblurGAN 模型

如图 7 所示, 在 GoPro 数据集上对比了多尺度框架、注意力框架以及 Deblur GAN 模型的去模糊结果。结果表明, 多尺度框架的去模糊图像出现了明显伪影, 而注意力框架并未在人脸区域恢复出清晰细节, 本文模型在多人物场景下不仅恢复更多的纹理细节且有效降低了伪影的产生; 同时在建筑背景商店商标模糊去除上, 本文模型恢复的清晰度优于基线模型。在人物背景的模糊去除上, 本文模型恢复的清晰度优于 Deblur GAN 模型; 以上效果得益于生成对抗网络的对抗学习机制, 其确保了图像模糊的有效。同时, 在金字塔框架中融合的多注意力机制, 进一步增强了细节信息的恢复能力。因此, 本文模型在真实场景图像去模糊任务中表现出了良好的有效性与鲁棒性。

4.2. 消融实验

本节在 GoPro、HIDE 与 Real Image 三个数据集上进行实验, 通过在基线模型中依次引入多尺度空间特征增强模块(Multi-scale Spatial Feature Enhancement Module)、自适应特征激励模块(Adaptive Feature Excitation Module)、增强通道注意力(Enhanced Channel Attention)以及复合损失函数, 验证了各组件对模型性能的有效性。

Table 3. Validation of whether MSFM and AFEM improve model performance, showing motion blur removal results on the HIDE and GOPRO datasets

表 3. MSFM 和 AFEM 是否存在对模型效果的验证, 在 HIDE 和 GOPRO 数据集上运动模糊去除效果

数据集	评估方法	Baseline [21]	无 MSFM 有 AFEM	有 MSFM 无 AFEM	有 MSFM 有 AFEM
HIDE	PSNR	25.9518	26.5076	26.6839	26.6775
HIDE	SSIM	0.8345	0.8587	0.8611	0.8526
GoPro	PSNR	26.3403	27.0620	26.7762	27.1772
GoPro	SSIM	0.8640	0.8830	0.8748	0.8825

Table 4. Validation of the effectiveness of MSFM and AFEM models on real-world image datasets for real-scene deblurring

表 4. MSFM 和 AFEM 是否存在对模型效果的验证, 在 Real image 数据集上的真实场景模糊去除效果

数据集	评估方法	Baseline [21]	无 MSFM 有 AFEM	有 MSFM 无 AFEM	有 MSFM 有 AFEM
real image J	PSNR	25.5982	25.3976	25.8126	25.3831
real image J	SSIM	0.8060	0.7874	0.7932	0.7965
real inage R	PSNR	30.1880	29.7164	30.7743	29.5419
real inage R	SSIM	0.8872	0.8694	0.8958	0.8465

从表 3 的结果可以看出, 基线模型加入 MSFM 和 AFEM 模块后, PSNR 和 SSIM 均有不同程度的提升, 说明这两个模块对模型的去模糊效果有提升。以下将对其作用进行具体分析, 首先, 基线模型在仅加入 MSFM 模块时对 HIDE 数据集去模糊的效果最好, 其 PSNR 和 SSIM 分别为 26.6839 dB 和 0.8611。

这一结果与 MSFM 模块的多尺度特征融合能力相关, 其通过提取并聚合不同感受野下的空间信息, 有效增强图像重建的质量。

当同时引入 MSFM 与 AFEM 模块时, 模型在 GoPro 测试集上的去模糊的 PSNR 提升至 27.1772 dB; 仅适用 AFEM 模块则将在 GOPRO 测试集上的客观指标 SSIM 提升到了 0.8830。由此可见, MSFM 模块

和 AFEM 模块对模型运动模糊去除并恢复图像细节有很大帮助。

表 4 对比了 AFEM 与 MSFM 模块在 Real Image 数据集上进行真实场景模糊去除的客观指标。结果显示, 在 Real image J 数据集上, AFEM 模块和 MSFM 模块的加入同样使 PSNR 和 SSIM 均达到了基线的相近水平, 并且有 MSFM 无 AFEM 情况下 PSNR 相较于基线提升了 0.2144 dB, 这表明 MSFM 模块在该数据集的去模糊任务中, 对多尺度特征信息提取有更明显的作用。在 Real image R 数据集中, AFEM 模块和 MSFM 模块的加入使得模型去模糊的客观指标达到了良好水平, 同时, 仅存在 MSFM 的情况下模型的去模糊效果提升更大, 其 PSNR 和 SSIM 分别达到了 30.7743 dB 和 0.8958。由此可得, MSFM 模块和 AFEM 模块在 Real Image 数据集上能取得较好的去模糊效果。

Table 5. Maintaining the presence of ECA and MSFM, validating the effectiveness of the AFEM module. Motion blur removal on the HIDE and GoPro datasets

表 5. 保持 ECA 和 MSFM 存在, 验证 AFEM 模块有效性。在 HIDE 和 GoPro 数据集上运动模糊去除

数据集	评估方法	Baseline [21]	our	无 AFEM
HIDE	PSNR	25.9518	26.5269	26.2058
HIDE	SSIM	0.8345	0.8537	0.8529
GoPro	PSNR	26.3403	27.0264	25.7463
GoPro	SSIM	0.8640	0.8791	0.8517

表 5 为在固定 MSFM 与 ECA 模块的基础上, 验证 AFEM 模块在 HIDE 和 GoPro 数据集上运动模糊去除效果的客观指标对比。根据该表结果。在 GOPRO 测试集和 HIDE 上, AFEM 模块的加入使得 PSNR 和 SSIM 均有不同程度的提升, 说明该模块对模型的去模糊重建能力提升有帮助; 其次, 在仅保留 ECA 模块和 MSFM 模块情况下, 模型对 GOPRO 测试集和 HIDE 数据集的去模糊效果仅达到基线相近水平, 其 PSNR 和 SSIM 分别为 25.7463 dB 和 0.8517, AFEM 模块凭借其自适应特征激励机制, 有效增强了模型在运动模糊去除过程中的关键信息捕获能力。

表 6 是保持 MSFM 和 ECA 模块存在, 验证 AFEM 模块在 Real Image 数据集上的真实场景模糊去除的有效性验证。

Table 6. Maintaining the presence of ECA and MSFM, validating the effectiveness of the AFEM module. Real-world scene blurring on the real image dataset

表 6. 保持 ECA 和 MSFM 存在, 验证 AFEM 模块有效性。在 real image 数据集上的真实场景模糊

数据集	评估方法	Baseline [21]	our	无 AFEM
real image J	PSNR	25.5982	25.5913	25.7566
real image J	SSIM	0.8060	0.7967	0.8100
real image R	PSNR	30.1880	30.7217	29.9372
real image R	SSIM	0.8872	0.9088	0.9267

根据表 6, 在 Real image J 数据集上, AFEM 模块使得 PSNR 和 SSIM 均达到同基线相近效果, 在仅保留 ECA 模块与 MSFM 模块, 却无 AFEM 的情况下, 模型对 Real image J 数据集的去模糊效果有大幅提升, PSNR 和 SSIM 分别为 25.7566 dB 和 0.8100。这一现象表明, 在该真实场景数据集中, AFEM 模块所具备的自适应特征激励机制, 与场景特性未完全匹配, 维持了模型的基础性能, 但未能带来预期

的提升效果；相反，ECA 与 MSFM 的组合更有利于该场景下图像结构与细节的稳定恢复。在 Real image R 数据集上，ECA 和 AFEM 模块的加入使得 PSNR 和 SSIM 均有不同程度的提升，说明 ECA 模块和 AFEM 模块对模型的去模糊重建有提升，综上，AFEM 模块对于模型在运动场景下模糊去除必不可少。

Table 7. Objective metrics for motion blur removal using models trained with different loss functions on the HIDE and GoPro datasets

表 7. 验证不同损失函数训练的模型，在 HIDE 和 GoPro 数据集上运动模糊去除的客观指标

数据集	评估方法	Baseline [21]	our	Losscontent	lossdis
HIDE	PSNR	25.9518	26.5269	26.2745	23.8425
HIDE	SSIM	0.8345	0.8537	0.8574	0.7693
GoPro	PSNR	26.3403	27.0264	27.1482	20.9830
GoPro	SSIM	0.8640	0.8791	0.8868	0.7334

根据表 7 在 HIDE 与 GoPro 数据集上进行的损失函数消融实验结果，可以看出不同损失项对模型性能的影响显著。当去除内容损失时，模型在测试集上的 PSNR 与 SSIM 均出现明显下降；而在仅使用内容损失、不引入判别损失的情况下，模型在 GoPro 数据集上仍能达到 27.1482 dB 的 PSNR 与 0.8868 的 SSIM。这一对比说明，内容损失在训练过程中起到了关键作用，其通过约束输出与清晰图像在特征层面的一致性，高效引导生成器恢复图像结构和细节信息，从而保障模型去模糊性能。相比之下，判别损失虽有助于增强生成图像的视觉真实感，但在以定量指标为主的评价体系下，单独使用该损失对 PSNR、SSIM 等数值提升的作用有限。因此，将内容损失与判别损失相结合的复合损失函数，能够更全面地引导模型优化，从而取得更稳定的训练效果。

Table 8. Verification of the real-world image deblurring performance of models trained with different loss functions on the real image dataset

表 8. 验证不同损失函数训练的模型，在 real image 数据集上的真实场景模糊去除效果

数据集	评估方法	Baseline [21]	our	losscontent	lossdis
real image J	PSNR	25.5982	25.5913	25.2957	24.9160
real image J	SSIM	0.8060	0.7967	0.7871	0.7508
real image R	PSNR	30.1880	30.7217	30.1144	29.3280
real image R	SSIM	0.8872	0.9088	0.8790	0.8386

表 8 展示了不同损失函数在 Real Image 数据集上的消融实验结果。可以看出，同时使用内容损失与判别损失的复合损失函数取得了最佳性能，这进一步印证了内容损失为图像重建提供了关键的结构约束，而判别损失则增强图像的视觉真实性，二者结合能更全面地促进模型优化。实验结果表明，在真实场景图像去模糊任务中，复合损失函数通过融合像素级保真度与对抗训练机制，有效保障并提升了模型的整体去模糊性能。

5. 结束语

在本实验中，我们提出融合注意力机制的生成对抗网络图像去模糊模型。实验结果所示，该模型在

去除图像运动模糊效果良好，可有效重建清晰的图像。我们使用了多种评价指标进行实验结果对比，包括峰值信噪比(PSNR)和结构相似性指数(SSIM)，结果表明新模型在这些指标上达到较好结果，进一步验证了在生成对抗网络中融合多注意力机制在图像去模糊任务中的有效性，未来将关注模型在目标检测等下游任务的应用。

参考文献

- [1] 李喆, 许佳路, 蔡志丹, 方明. 基于高斯曲率和加权图总变分正则化的遥感图像盲去模糊算法[J]. 吉林大学学报(工学版), 2023(9): 2649-2658.
- [2] 王珮, 朱宇, 闫庆森, 孙瑾秋, 张艳宁. 真实场景图像去模糊: 挑战与展望[J]. 中国图象图形学报, 2024, 29(12): 3501-3528.
- [3] 胡雪, 黄成泉, 冯润, 周丽华, 郑兰. 全变分极端通道先验的盲图像去噪去模糊[J]. 数据采集与处理, 2022, 37(3): 643-656.
- [4] Kupyn, O., Budzan, V., Mykhailych, M., Mishkin, D. and Matas, J. (2018). Deblurgan: Blind Motion Deblurring Using Conditional Adversarial Networks. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 8183-8192. <https://doi.org/10.1109/cvpr.2018.00854>.
- [5] Neji, H., Hamdani, T.M., Halima, M.B., *et al.* (2021) Blur2Sharp: A GAN-Based Model for Document Image Deblurring. Technical Report.
- [6] Zhao, S., Zhang, Z., Hong, R., Xu, M., Yang, Y. and Wang, M. (2022). FCL-GAN: A Lightweight and Real-Time Baseline for Unsupervised Blind Image Deblurring. *Proceedings of the 30th ACM International Conference on Multimedia*, Lisboa, 10-14 October 2022, 6220-6229. <https://doi.org/10.1145/3503161.3548113>
- [7] Fanous, M.J. and Popescu, G. (2022) Ganscan: Continuous Scanning Microscopy Using Deep Learning Deblurring. *Light: Science & Applications*, **11**, Article No. 265. <https://doi.org/10.1038/s41377-022-00952-z>
- [8] Chen, B., Jiang, L., Wu, P., Zheng, F., Li, K., Chen, T., *et al.* (2022). A Feature Attention Based Multi-Stage Network for Image Deblurring. 5th International Conference on Computer Information Science and Application Technology (CISAT 2022), Chongqing, 20 October 2022, 432-437. <https://doi.org/10.1117/12.2656814>
- [9] Cui, Y., Tao, Y., Ren, W. and Knoll, A. (2023) Dual-domain Attention for Image Deblurring. *Proceedings of the AAAI Conference on Artificial Intelligence*, **37**, 479-487. <https://doi.org/10.1609/aaai.v37i1.25122>
- [10] Sub Lee, H. and In Cho, S. (2024) Multi-Scale Attention in Attention Neural Network for Single Image Deblurring. *Displays*, **85**, Article ID: 102860.
- [11] Zhang, B., Sun, J., Sun, F., Wang, F. and Zhu, B. (2024) Image Deblurring Method Based on Self-Attention and Residual Wavelet Transform. *Expert Systems with Applications*, **244**, 123005. <https://doi.org/10.1016/j.eswa.2023.123005>
- [12] Chen, Z., Gu, J., Kong, L., Liu, D., Xia, B., Yuan, X., *et al.* (2023). Hierarchical Integration Diffusion Model for Realistic Image Deblurring. *Advances in Neural Information Processing Systems*, **36**, 29114-29125. <https://doi.org/10.52202/075280-1267>
- [13] Wu, J., Tsai, F., Peng, Y., Tsai, C., Lin, C. and Lin, Y. (2024). Id-blau: Image Deblurring by Implicit Diffusion-Based Reblurring Augmentation. 2024 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 16-22 June 2024, 25847-25856. <https://doi.org/10.1109/cvpr52733.2024.02442>
- [14] Feng, H., Zhou, H., Ye, T., Chen, S. and Zhu, L. (2025) Residual Diffusion Deblurring Model for Single Image Defocus Deblurring. *Proceedings of the AAAI Conference on Artificial Intelligence*, **39**, 2960-2968. <https://doi.org/10.1609/aaai.v39i3.32303>
- [15] Li, Y., Fang, H., Lei, X., Wang, Q., Hu, G., Dong, J., *et al.* (2025) Real-world Defocus Deblurring via Score-Based Diffusion Models. *Scientific Reports*, **15**, Article No. 22942. <https://doi.org/10.1038/s41598-025-07326-6>
- [16] Xie, X., Zhang, Q. and Zheng, W. (2025). Diffusion-based Event Generation for High-Quality Image Deblurring. 2025 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, 10-17 June 2025, 2194-2203. <https://doi.org/10.1109/cvpr52734.2025.00210>
- [17] Nah, S., Kim, T.H. and Lee, K.M. (2017). Deep Multi-Scale Convolutional Neural Network for Dynamic Scene Deblurring. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 257-265. <https://doi.org/10.1109/cvpr.2017.35>
- [18] Shen, Z., Wang, W., Lu, X., Shen, J., Ling, H., Xu, T., *et al.* (2019). Human-aware Motion Deblurring. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 5571-5580. <https://doi.org/10.1109/iccv.2019.00567>

-
- [19] Rim, J., Lee, H., Won, J. and Cho, S. (2020) Real-world Blur Dataset for Learning and Benchmarking Deblurring Algorithms. In: *Lecture Notes in Computer Science*, Springer International Publishing, 184-201. https://doi.org/10.1007/978-3-030-58595-2_12
- [20] Zhang, H. and Xu, M. (2018) Dual Pattern Learning Networks by Empirical Dual Prediction Risk Minimization.
- [21] Kupyn, O., Martyniuk, T., Wu, J. and Wang, Z. (2019). DeblurGAN-v2: Deblurring (Orders-of-Magnitude) Faster and Better. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 8877-8886. <https://doi.org/10.1109/iccv.2019.00897>
- [22] Hu, J., Shen, L., Albanie, S., Sun, G. and Wu, E. (2020) Squeeze-and-excitation Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **42**, 2011-2023. <https://doi.org/10.1109/tpami.2019.2913372>