基于特征解耦表征学习的无监督混合失真图像复原方法

符颖[1,2], 李卓遥[1], 朱欣宇[1], 龚敏学[1], 周激流[1,2]

[1] (成都信息工程大学计算机学院　成都　610225)
[2] (四川省 2011 图形图像与空间信息协同创新中心　成都　610225)
(fuying@cuit.edu.cn)

摘　要: 针对真实场景下多种混合失真组合的多任务图像复原, 结合生成对抗网络与编码器提出了一种基于无监督对偶学习的图像复原方法. 该方法引入特征解耦模块, 通过修正基于增益控制的归一化, 将不同退化机制的特征表示分配到不同的特征通道中, 使得不同退化特征表达相互独立, 实现了通道上的特征解耦. 同时, 为了进一步过滤掉退化机制的特征表示并保持原图像内容信息的细节, 利用通道注意力机制实现特征解耦后自适应地选择有用的特征表示, 使其适用于真实场景下混合失真组合的图像复原任务. 实验结果表明, 所提算法在单一退化类型 GoPro 数据集上对比基于尺度循环网络的算法(scale-recurrent network, SRN), 峰值信噪比和结构相似性 2 项指标值分别提高了 0.499 dB 和 0.044; 在混合退化类型 DIV2K 数据集上对比基于操作选择注意力网络的算法(operation-wise attention network, OWAN), 峰值信噪比和结构相似性 2 项指标值分别提高了 0.163 dB 和 0.015, 在实现复原图像的同时, 保留了图像的纹理和细节信息.

关键词: 混合失真图像; 图像复原; 特征解耦; 注意力机制
中图法分类号: TP391.41　　**DOI:** 10.3724/SP.J.1089.2024.19717

Unsupervised Hybrid-Distorted Image Restoration Method Based on Feature Disentangled Representation Learning

Fu Ying[1,2], Li Zhuoyao[1], Zhu Xinyu[1], Gong Minxue[1], and Zhou Jiliu[1,2]

[1] (*School of Computer Science and Technology, Chengdu University of Information Technology, Chengdu*　610225)
[2] (*Images and Spatial Information 2011 Collaborative Innovation Center of Sichuan Province, Chengdu*　610225)

Abstract: For the hybrid-distorted image restoration in real scenes, an image restoration algorithm based on unsupervised dual learning is proposed with the generative adversarial networks and the encoder. The algorithm introduces feature decoupling module where the different feature representations from different degradation mechanisms are assigned to different feature channels by modifying the normalization based on gain control so the feature disentanglement in channels is realized by independent feature representations. Meanwhile, the channel attention mechanism is used to realize the restoration of image content feature of clean images. Compared with the scale-recurrent network (SRN) algorithm, the proposed algorithm improves 0.499 dB and 0.044 in terms of peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) on GoPro dataset. Compared with the OWAN algorithm, the proposed algorithm improves 0.163 dB and 0.015 in terms of PSNR and SSIM on DIV2K dataset. The experiments also demonstrate that the detailed infor-

收稿日期: 2022-05-20; 修回日期: 2022-09-03. **基金项目**: 国家留学基金委资助项目(201908515022); 四川省重大科技专项(2019ZDZX0005); 四川省自然科学基金青年基金(2022NSFSC0907). **符颖**(1988—), 女, 博士, 副教授, 硕士生导师, 主要研究方向为信号与图像处理、非参数贝叶斯方法在图像中的应用; **李卓遥**(1995—), 女, 硕士研究生, 主要研究方向为计算机视觉、图像处理; **朱欣宇**(1999—), 女, 硕士研究生, 主要研究方向为计算机视觉、图像处理; **龚敏学**(1995—), 女, 硕士研究生, 主要研究方向为计算机视觉、图像处理; **周激流**(1963—), 男, 博士, 教授, 博士生导师, 主要研究方向为计算智能、图形图像处理及应用.

mation can be restored by the proposed algorithm.

Key words: hybrid-distorted image; image restoration; feature disentanglement; attention mechanism

传统的图像复原方法是以退化图像作为已知数据，根据造成图像退化的机制建立正确的关于图像退化的数学模型，从而复原出符合客观真实场景的原始清晰图像．近年来，生成对抗网络(generative adversarial networks, GAN)[1]因其强大的学习图像分布能力被学者广泛用于解决图像复原问题．Pathak 等[2]使用 GAN 来实现人脸图像修复，L_2 损失和对抗损失的结合可以实现更好的复原效果，但因该方法所复原区域的形状是固定的，其在实际应用中有所限制．空间规整网络(spatial regulation network, SRN)模型[3]将超分辨率重建与去卷积相结合，通过跨尺度权值共享方案以降低训练难度，多尺度循环渐进地生成清晰图像．

低层图像处理任务可以归结为图像源域到目标域的问题，网络可以学习到 2 个域之间的映射关系，以完成域之间的转换．同时，根据是否使用退化图像和与之对应的清晰图像对，图像任务分为监督和无监督两种方式．基于有监督的图像复原模型利用成对的样本进行训练，此类模型的训练需要大量成对的样本集，但清晰的真值图像(ground truth, GT)的获取相对困难且成本较高，并且强监督会造成过拟合和模型泛化能力较差．基于域迁移的无监督学习方法在图像转换任务中表现优异，如 CycleGAN[4]解决了在非成对样本下的图像翻译问题．传统的 GAN 是单向生成，而 CycleGAN 是互相生成，其网络为环形．基于无监督对偶学习的模型成为了学者的研究热点之一，出现了许多以该模型为基础的改进模型，例如，Lee 等[5]提出的 DRIT 模型，其中解耦的编码器进一步提高了 CycleGAN 的效果；又如 CoGAN[6], UNIT[7], MUNIT[8], Augmented CycleGAN[9], StarGAN[10]等，均取得了较好的效果．其中，Liu 等[7]提出 1 个前提假设，假定一对不同域的图像能够被映射到共享潜空间的同一个表示．MUNIT 在 UNIT 的基础上做了改进，实现了无监督 image-to-image 多模态翻译，设计了 1 个风格编码器和 1 个内容编码器，利用解码器来整合 2 个编码器以生成多样的图像．Lu 等[11]提出的 UID-GAN(unsupervised image deblurring via disentangled representations)在 CycleGAN 和 DRIT 的基础上进行了改进，同样由 2 个 GAN 构成了一个对偶网络，使循环一致性损失与对抗损失最小化，利用内容编码器和模糊信息编码器对模糊图像的内容信息与模糊特征分别进行编码，实现了基于无监督学习的特定领域图像的去模糊任务．然而，对于混合失真组合图像复原任务，当前存在的问题是混合失真组合中各种退化机制的相互影响，复原网络不能根据输入图像中存在的模糊、压缩、噪声等退化机制做出相应操作进行复原，可能出现退化机制相互纠缠的复原结果．目前，针对混合失真图像复原任务的算法较少，值得一提的算法有基于深度强化学习的复原算法(restoration by deep reinforcement learning, RL-Restore)[12]与基于操作选择注意力网络的算法(operation-wise attention network, OWAN)[13]．白亮等[14]在后者的基础上进行改进,充分地利用层级特征的相关特性提升了算法的效果．

针对上述问题，本文在 UIDGAN 的基础上提出了一种基于无监督对偶学习的图像复原方法，结合 GAN 与编码器进行混合失真组合图像的复原工作．本文引入了特征解耦模块(feature disentangled module, FDM)，实现了通道上的特征解耦，使得不同退化特征表达相互独立；为了进一步过滤掉退化机制的特征表示，并保持原图像内容信息的细节，本文利用通道注意机制在处理后的解耦特征表示中自适应地选择有用的特征表示，使其适用于真实场景下混合失真组合的图像复原任务．

1　相关工作

1.1　特征解耦表征学习

解耦表征学习是指在满足一定条件的数据集上，效仿人类的认知，从真实数据中对具有明确物理含义或者语义信息的变化因素(如类别、纹理、位置、外观等)分解出来，并给出其对应的潜在表示．这些变化因素相互独立，当单个变化因素发生变化时，只会使生成数据中的该因素发生变化，而其他因素保持不变．现有大多数表征学习网络通常基于 Bodin 等[15]提出的高斯过程潜变量模型(Gaussian process latent variable model, GP-LVM)和变分自编码(variational auto-encoders, VAE)模型相结合的方法以及 GAN. 2018 年，Li 等[16]提出了基于无监督学习的相似性限制的 GAN，通过在条件

与生成图像之间添加相似度约束来解耦出具有可解释性的表征，相似性约束也是作为指导生成网络来理解基于条件表征差异的一种监督. 2016 年，Higgins 等[17]提出的 β-VAE 也是一种无监督的解耦表征学习模型，其对 VAE 的目标函数上增加了一个超参 β，以促进潜在表征的更多因式分解，但这个修改降低了重建损失的权重，导致重建图像效果弱于 VAE. 2020 年，Patacchiola 等[18]提出的新模型 Y-AutoEncoder 实现了图像风格转换、内容与样式分离上效果的提升.

真实世界中的图像通常会受多重退化机制的影响，并且不同退化机制之间可能会相互干扰，而深度特征表示能够有效地表征各种图像退化机制，即自混合失真组合的图像中提取到的特征表示解耦后能够分别表征不同的退化机制，这一点由文献[19]可以证实. 另外，Schwartz 等[20]指出，多个滤波器和基于增益控制归一化能够对不同滤波器的响应进行分解，而卷积神经网络中的卷积层也是由一系列滤波器/卷积核构成.

1.2 注意力机制

人类视觉系统可以自然、高效地找到复杂场景中的重要的区域，受此启发，注意力机制被引入到计算机视觉系统中. 注意力机制已经在计算机视觉的各种任务(如图像生成、目标检测、图像识别、语义分割等)中取得了巨大的成功. 常见的注意力模型可以分为 3 类：基于空间域的注意力模型、基于通道域的注意力模型和基于混合域的注意力模型. 其中，基于空间域的注意力模型根据输入的特征计算得到保存关键空间信息的注意力掩码，将该掩码附加在原来的特征上，增大有用特征、弱化无用特征，从而达到特征筛选和增强的效果. 基于通道域的注意力模型单独为每个特征通道增加 1 个可学习的权重，其表示与关键信息的相关程度，从而加强对当前任务有用的通道；反之，则抑制. 空间注意力机制由于对每个通道均采取相同的处理方式，忽略了特征通道之间的相关性；通道注意力机制由于采用平均池化的方式压缩通道中的信息，忽略了每一个通道中的局部信息. 因此，有研究人员结合 2 种注意力机制的优点，提出了混合域的注意力机制，其中的典型代表为卷积块注意力模块 (convolutional block attention module, CBAM)[21]，该模块将空间注意力模块和通道注意力模块串行结合，并且通过融合原始输入特征与经过注意力模块处理后的特征，更好地凸显关键特征信息.

2 本文方法

2.1 本文方法框架

本文方法的框架如图 1 所示，以 UID-GAN 为基线网络，引入 FDM，使其适用于真实场景下混合失真组合的图像复原任务. 网络由 2 个 GAN，即 $\mathrm{GAN_C}=\{G_\mathrm{C}, D_\mathrm{C}\}$ 与 $\mathrm{GAN_H}=\{G_\mathrm{H}, D_\mathrm{H}\}$，2 个内容信息编码器 E_C 与 E_H，一个退化特征编码器 E_h 和 2 个 FDM 构成. 编码器的作用是对图像特征进行提取，其中，2 个内容信息编码器分别对清晰/退

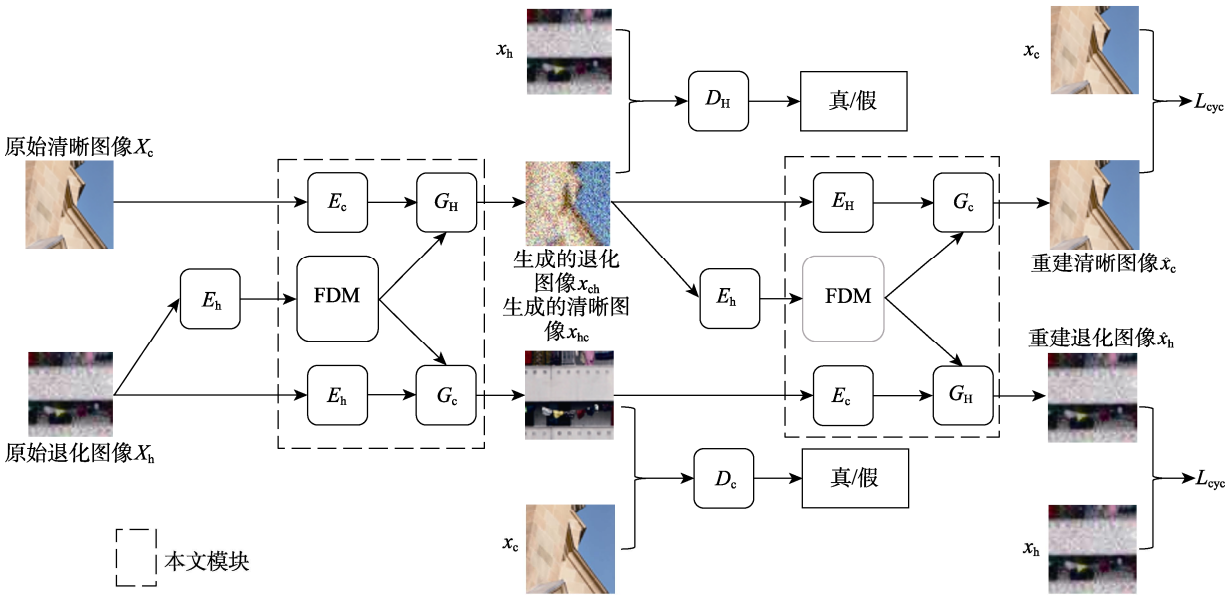


图 1 本文方法框架

化图像的内容信息进行编码，退化特征编码器对退化图像中的退化特征进行编码．首先，在UID-GAN的框架中，已经存在了第1次解耦，也就是输入原始退化图像后，退化特征编码器和内容信息编码器分别对其中的退化特征和内容信息进行编码．此时，为了提升退化图像内容编码器的性能，使清晰图像的内容信息编码器与之共享最后一层权重，指导其编码退化图像的内容信息(这里设定清晰图像不存在退化特征，因此，其内容编码器只能对清晰内容信息进行编码)．在退化特征编码器对原始退化图像进行编码之后，在网络中引入了FDM，这就是网络中的第2次解耦，目的是让自混合失真组合的图像中提取到的特征表示在解耦后能够分别表征不同的退化机制，使得这些不同退化特征表达相互独立．

受文献[22]的启发，本文在模型中添加了 1 个 FDM 来实现对不同通道的特征解耦，从而使不同退化机制的特征分布在不同的通道上，减少这些退化机制相互影响的可能性，图2给出了该模块工作原理示意图，利用其实现通道上的特征解耦，通过神经网络不同的层表达不同的退化机制．例如，第 1 层表示噪声特征，第 2 层表示模糊特征，第 3 层表示 JPEG 压缩．从经过处理的特征表示中收集有用的内容信息来重建图像是低层次图像复原的关键，因此，模型中还使用了基于通道的注意机制，自适应地选择有用的特征表示，以更好地复原原始图像的内容细节．最后，输入到清晰图像生成器 G_{C}，让其与相对应的清晰图像判别器 D_{C} 相互对抗，从而生成清晰图像．

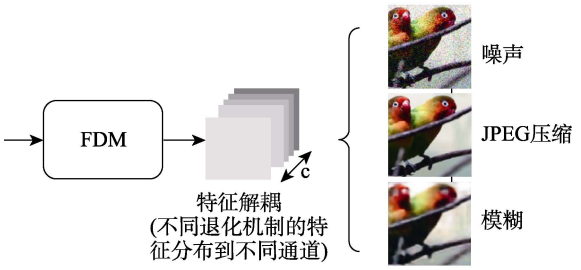


图 2　FDM 工作原理示意图

整个过程中使用了 4 个损失函数：(1) KL (Kullback-Leibler)散度 L_{KL}，用于规范退化特征编码器编码得到的退化特征的分布，使其接近正态分布; (2) 2 个 GAN 的对抗损失 $L_{D_{\mathrm{C}}}$ 与 $L_{D_{\mathrm{H}}}$，一个是使清晰图像转换为退化图像的退化网络，另一个是使退化图像转换为清晰图像的复原网络，两者分别将原始退化图像与原始清晰图像输入，并使退化图像判别器和清晰图像判别器与生成器进行对抗，从而生成新的退化图像和清晰图像; (3) 循环一致性损失 L_{cyc}，以确保在经过 2 个生成器后重建的图像能够近似于最开始输入的图像; (4) 特征正交损失 L_{F}，目的是使解耦后的每个通道能够表达足够的特征，让各个特征向量同样重要且相互独立．

2.2　损失函数

本文方法基于无监督对偶学习，使用非成对样本进行训练，退化图像没有与之对应的清晰图像．在这种情况下，从退化图像中分离出图像的内容信息并非易事．本文先假定清晰图像中只包含内容信息，不存在任何的退化特征，那么清晰图像的内容编码器 E_{C} 就能够从中提取出干净的内容信息．同时，为了使退化图像的内容编码器 E_{H} 能够更好地从退化图像中提取到内容信息，采用清晰/退化内容编码器共享其最后一层权重的方法，使前者能够指导后者有效地从退化图像中提取内容信息．通过内容编码器 E_{C} 得到清晰图像的内容信息，将其与退化特征编码器所编码的退化特征 z_{h} 一并送入退化图像生成器 G_{H} 中，以生成对应的退化图像 x_{ch}．为了使退化特征 z_{h} 的分布更加规律，添加KL散度损失使其接近于正态分布. KL散度损失公式为

$$\mathrm{KL}\left(q\left(z_{\mathrm{h}}\right)\middle\|p(z)\right)=-\int q\left(z_{\mathrm{h}}\right)\mathrm{lb}\frac{p(z)}{q\left(z_{\mathrm{h}}\right)}\mathrm{d}z \tag{1}$$

式(1)等价于

$$L_{\mathrm{KL}}=\frac{1}{2}\sum_{i=1}^{N}\left(\mu_i^2+\sigma_i^2-\mathrm{lb}\left(\sigma_i^2\right)-1\right) \tag{2}$$

其中，μ 和 σ 分别表示退化特征 z_{h} 的均值和标准差；N 表示 z_{h} 的维数．

本文方法包含2个网络：退化图像的复原网络和清晰图像的退化网络．前者将输入的原始退化图像 x_{h} 转换为清晰的图像 x_{hc}，利用判别器 D_{C} 判断生成的清晰图像 x_{hc} 的真假．该网络的对抗损失公式为

$$\begin{aligned}L_{D_{\mathrm{C}}}&=E_{\mathrm{c}\sim\mathrm{p(c)}}\left[\mathrm{lb}D_{\mathrm{C}}\left(x_{\mathrm{c}}\right)\right]+\\&E_{\mathrm{h}\sim\mathrm{p(h)}}\left[\mathrm{lb}\left(1-D_{\mathrm{C}}\left(G_{\mathrm{C}}\left(E_{\mathrm{H}}\left(x_{\mathrm{h}}\right),z_{\mathrm{h}}\right)\right)\right)\right]\end{aligned} \tag{3}$$

后者将清晰图像转换为退化图像，利用清晰图像内容编码器 E_{C} 与退化图像生成器 G_{H} 将输入网络的清晰图像 x_{c} 生成与之对应的退化图像 x_{ch}．该网络利用退化图像判别器 D_{H} 判断生成的退化

图像 x_{ch} 的真假，对抗损失定义公式为

$$L_{D_{\mathrm{H}}}=E_{\mathrm{h}\sim\mathrm{p(h)}}\left[\mathrm{lb}D_{\mathrm{H}}\left(x_{\mathrm{h}}\right)\right]+E_{\mathrm{c}\sim\mathrm{p(c)}}\left[\mathrm{lb}\left(1-D_{\mathrm{H}}\left(G_{\mathrm{H}}\left(E_{\mathrm{C}}\left(x_{\mathrm{c}}\right),z_{\mathrm{h}}\right)\right)\right)\right] \tag{4}$$

由于整个训练过程是无监督的，类似于CycleGAN，本文方法也利用了循环一致性损失，来确保源域 X 中的图像 x 经过一个生成器生成的图像 $G(x)$ 再作为另一个生成器的输入所生成的 $F(G(x))$ 能够尽可能接近原来的图像 x．也就是说，在本文模型训练过程中，需要使原来输入的退化图像 x_{h} 通过复原网络得到的图像 x_{hc}，在经过退化网络重建生成的 $\widehat{x_{\mathrm{h}}}$ 近似于原退化图像 x_{h}；同时，使输入的清晰图像 x_{c} 经过退化后再重建，也能够转换到原有的清晰图像的领域．具体过程定义公式为

$$\begin{cases}x_{\mathrm{hc}}=G_{\mathrm{C}}\left(E_{\mathrm{H}}\left(x_{\mathrm{h}}\right),E_{\mathrm{h}}\left(x_{\mathrm{h}}\right)\right)\\ x_{\mathrm{ch}}=G_{\mathrm{H}}\left(E_{\mathrm{C}}\left(x_{\mathrm{c}}\right),E_{\mathrm{h}}\left(x_{\mathrm{h}}\right)\right)\end{cases} \tag{5}$$

式(5)分别为原退化图像的输入 x_{h} 到复原图像 x_{hc}，与原清晰图像 x_{c} 到生成的退化图像 x_{ch} 的正向转换．复原图像 x_{hc} 经过退化网络生成的退化图像 $\widehat{x_{\mathrm{h}}}$，与退化图像 x_{ch} 到经过复原网络生成的复原图像 $\widehat{x_{\mathrm{c}}}$ 的反向转换，公式为

$$\begin{cases}\widehat{x_{\mathrm{h}}}=G_{\mathrm{H}}\left(E_{\mathrm{C}}\left(x_{\mathrm{hc}}\right),E_{\mathrm{h}}\left(x_{\mathrm{ch}}\right)\right)\\ \widehat{x_{\mathrm{c}}}=G_{\mathrm{C}}\left(E_{\mathrm{H}}\left(x_{\mathrm{ch}}\right),E_{\mathrm{h}}\left(x_{\mathrm{ch}}\right)\right)\end{cases} \tag{6}$$

则循环一致性损失定义为

$$L_{\mathrm{cyc}}=E_{x_{\mathrm{c}}\sim\mathrm{p}(x_{\mathrm{c}})}\left[\left\|x_{\mathrm{c}}-\widehat{x_{\mathrm{c}}}\right\|_1\right]+E_{x_{\mathrm{h}}\sim p(x_{\mathrm{h}})}\left[\left\|x_{\mathrm{h}}-\widehat{x_{\mathrm{h}}}\right\|_1\right] \tag{7}$$

另外，利用特征距离作为知觉相似性判断，在复原图像与原始退化图像之间添加了感知损失，定义公式为

$$L_{\mathrm{p}}=\left\|\xi_i\left(x_{\mathrm{hc}}\right)-\xi_i\left(x_{\mathrm{h}}\right)\right\|_2^2 \tag{8}$$

其中，$\xi_i(x)$ 表示预训练网络的第 i 层特征．

最后，本文方法的目标函数为上述式(2)~式(4)、式(7)和式(14)所有损失的加权和，即

$$L=\lambda_{\mathrm{adv}}L_{\mathrm{adv}}+\lambda_{\mathrm{KL}}L_{\mathrm{KL}}+\lambda_{\mathrm{cyc}}L_{\mathrm{cyc}}+\lambda_{\mathrm{p}}L_{\mathrm{p}}+\lambda_{\mathrm{F}}L_{\mathrm{F}} \tag{9}$$

包括 KL 散度损失 L_{KL}，2 个 GAN 的对抗损失 $L_{\mathrm{adv}}=L_{D_{\mathrm{C}}}+L_{D_{\mathrm{H}}}$，循环一致性损失 L_{cyc}，感知损失 L_{p}，以及第 2.3 节中介绍的特征正交性损失 L_{F}．

2.3 FDM

本节介绍了基于增益控制的信号分解，并分别阐述 FDM 的原理和作用．为了推导感觉处理的模型，Schwartz 等[20]提出了基于增益控制的归一化来实现自然信号的非线性分解．即给定一组信号 X，信号分解使用一组滤波器 $\{f_1,f_2,\cdots,f_n\}$ 从中提取信号的不同表征 $\{V_1,V_2,\cdots,V_n\}$，即

$$V_i=f_i(X),\ i=1,2,\cdots,n \tag{10}$$

然后，使用归一化进一步消除这些表征之间的依赖关系，根据适当的权重 ω_{ji} 和与输入 V_i 对应的偏移量 σ^2 来生成独立响应

$$R_i=\frac{V_i^2}{\sum_{j=1}^{n}\omega_{ji}V_j^2+\sigma^2} \tag{11}$$

输入信号 X 能够以其统计特性被分解为 $R_1,R_2,\cdots,R_n$．在此基础上，本文通过将信号分解扩展为通道特征解耦，降低多种退化机制之间的干扰．多个线性滤波器和分裂归一化的组合能够对不同滤波器的响应进行分解，而卷积神经网络中的卷积层也是由一系列滤波器/卷积核构成．因此，本文应用该算法实现自适应的通道特征分解．具体来说，在网络中使用卷积层 $(C_{\mathrm{in}}\times C_{\mathrm{out}}\times k\times k)$ 取代信号分解中的传统卷积核 $\{f_1,f_2,\cdots,f_n\}$，FDM 如图 3 所示．

从卷积层提取出来的结果分布在不同的通道中，输出特征中的各通道分量为

$$\left\{S_1,S_2,\cdots,S_{C_{\mathrm{out}}}\right\}=\mathrm{conv}\left(F_{\mathrm{in}}\right) \tag{12}$$

其中，输入通道数 C_{in} 表示输入特征 F_{in} 的通道维数，输出通道 C_{out} 表示输出特征 F_{out} 的通道维数；k 表示卷积核大小．

将基于增益控制的归一化引入到卷积神经网络中，再改进使其适用于本文的任务．一是将式

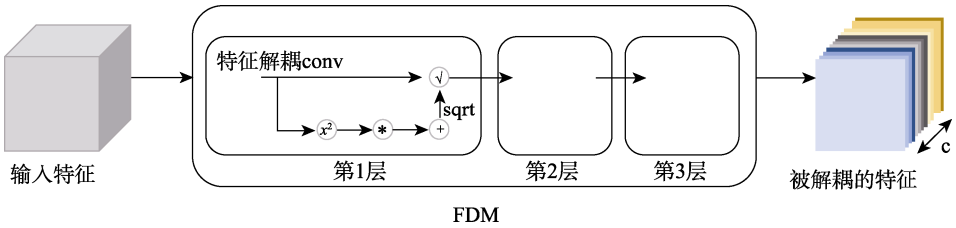


图 3 FDM

(11)中的分子分母改为平方根便于梯度传播的实现；二是用特征 S_i 的通道分量取代滤波器的响应 V_i，使其适用于特征的通道级解耦，则有

$$D_i = \frac{S_i}{\sqrt{\sum_{j=1}^{n} \omega_{ji} S_j^2 + b_i}} \tag{13}$$

其中，ω_{ji} 与 b_i 可以通过梯度下降来学习；S_i 与 D_i 分别表示增益控制前后特征的第 i 个通道分量. 为了降低特征之间的相关性，引入了特征正交性损失

$$L_{\mathrm{F}} = \left\| \lambda_1\left(\boldsymbol{F}\boldsymbol{F}^{\mathrm{T}}\right) - \lambda_2\left(\boldsymbol{F}\boldsymbol{F}^{\mathrm{T}}\right) \right\|_2^2 \tag{14}$$

其中，$\boldsymbol{F}$ 表示特征图；$\lambda_1\left(\boldsymbol{F}\boldsymbol{F}^{\mathrm{T}}\right)$ 与 $\lambda_2\left(\boldsymbol{F}\boldsymbol{F}^{\mathrm{T}}\right)$ 分别表示 $\boldsymbol{F}\boldsymbol{F}^{\mathrm{T}}$ 的最大特征值和最小特征值. 特征值表示的是特征分解到对应的特征向量的长度，不同的特征向量互相独立，设置该特征正交性损失 L_{F} 是为了使最大特征值和最小特征值一致. 在一个矩阵特征分解出来的特征值差异较大的情况下，舍弃掉特征值小的部分对于这个矩阵来说损失很小. 要求最大和最小特征值一致，即要求各个特征向量同样重要，不舍弃每个特征向量的信息. 因此，每个通道才能表达足够的信息，并且这些信息互相独立. 此处使用 $\boldsymbol{F}\boldsymbol{F}^{\mathrm{T}}$ 而非 $\boldsymbol{F}$ 矩阵的特征值，是因为只有方阵才可能具有特征值.

为了实现图像复原，进一步过滤掉退化机制的特征表示，并保持原图像内容信息的细节，利用通道注意机制在处理后的解耦特征表示中自适应地选择有用的特征表示，输出特征的第 i 个通道为

$$F_{\mathrm{pi}} = A\left(D_i\right) \tag{15}$$

其中，A 表示通道注意力；D_i 表示解耦特征的第 i 个通道.

在特征解耦的过程中，不同特征通道的互信息减少，可能会导致图像内容信息的丢失. 为了弥补特征解耦分支的不足，模型利用残差块增强图像信息的并行传输；同时，为了充分利用不同层次的特征信息，采用了多个模块堆叠的方式.

3　实验评估及分析

3.1　实验数据

本文方法在 DIV2K 数据集[23]上与其他经典的网络进行了比较. DIV2K 数据集有 800 幅高清大尺寸的图像，将其分为两部分，取 50 幅用于测试，其余用于模型训练，以 stride = 95 从这些图像中裁剪成 63×63 像素的小尺寸图像. 为了模拟真实复杂的退化场景，将高斯噪声、高斯模糊和 JPEG 压缩加入到图像中. 前两者的标准差分别从[0,50]和[0,5]的范围内随机选择，压缩质量则从[10,100]的范围内随机选择. 根据其不同退化程度，将生成的退化图像分为 3 类：轻度(mild)、中度(moderate)和重度(severe). 本节对于单一失真类型的图像复原实验，使用了包含 3 214 对模糊/清晰的图像的 GoPro 数据集[24]，测试集有 1 111 对图像，其余用于模型训练.

本节所使用的混合失真组合的图像数据集，除经过处理的 DIV2K 数据集之外，为了增加混合失真组合图像的复原难度，还构建了一个混合失真组合图像数据集. 使用雨纹数据集[25]作为基本图像集，训练集包含 861 对干净的场景图像与带有雨纹的退化版本，测试集包含 58 幅图像. 首先，从每幅图像中裁剪出包含雨纹的尺寸为 128×128 像素的小块图像，然后向其中添加高斯噪声、JPEG 压缩和运动模糊. 本文随机改变了每个 patch 中要添加的失真类型的数量，并分别从 [10,30] 和 [15,35]的范围内随机选择了高斯噪声的水平(标准差)和 JPEG 压缩的质量. 因此，该数据集包含噪声、模糊、压缩和雨纹 4 种失真类型，在对比实验和消融实验中用作测试，以此验证本文方法及其中各个模块的有效性. 最后，在模型鲁棒性实验中，对 DIV2K 数据集做了新的处理来验证模型的扩展性.

3.2　实验参数及配置

本文框架的实现基于 PyTorch，采用 NVIDIA 1080Ti GPU. 在训练过程中，采用了 Adam 优化器，学习率设置为 2×10^{-2}，训练时使用小批量梯度下降法，BatchSize 设置为 24. 一共训练 100 轮，模型在前期训练时学习率不变，后期逐渐衰减，使网络在最初快速收敛后能稳定收敛到最优点. 模型的超参数(即损失函数的权值)分别设置为 $\lambda_{\mathrm{adv}} = 1$，$\lambda_{\mathrm{KL}} = 0.01$，$\lambda_{\mathrm{cyc}} = 10$，$\lambda_{\mathrm{p}} = 0.1$，$\lambda_{\mathrm{F}} = 0.000\,01$.

3.3　对比实验

3.3.1　单一退化类型

考虑真实世界中的图像可能存在仅被一种退化机制影响的情况，为了验证方法的可扩展性与鲁棒性. 由于本文方法的框架基于无监督对偶学习特定域去模糊的方法，因此，本节使用包含模糊/干净的图像对的模糊图像数据集 GoPro 进行验证. 实验指标的平均结果如表 1 所示，本文方法在 PSNR 与 SSIM 上比基于域迁移的无监督学习方法 CycleGAN 与特定域去模糊算法 UIDGAN 有了明

表 1 单一退化类型-模糊图像复原实验结果

方法	PSNR/dB	SSIM
CycleGAN[4]	24.834	0.648
UIDGAN[11]	23.711	0.567
SRN[3]	25.566	0.672
本文	**26.015**	**0.716**

注. 粗体表示最优结果.

显的提高，相比尺度循环网络 SRN，本文算法的 PSNR 提高了 0.449 dB, SSIM 提高了 0.044.

图 4 展示了上述 3 种方法与本文方法在模糊数据集上的复原结果，截取了具有代表性的部分图像细节，复原字体、数字以及建筑物轮廓. 可以看出，CycleGAN 对于模糊图像的复原，虽然整体上色彩进行了还原，但字体、建筑物轮廓周围仍有伪影，第 3 个字母 A 的左边还出现了原图没有的颜色. UIDGAN 效果略好于前者，建筑图中前景与背景得到了良好区分，第 3 幅图比 CycleGAN 图像锐利一些，但也没有还原出字母. SRN 复原效果整体相对较好，复原了建筑物的纹理，但对于字母的复原效果不如本文方法清晰. 总体而言，本文方法能够生成丰富纹理，说明了在图像特征提取方面，结合了图像全局和局部的特征，对单一退化的模糊图像复原效果良好.

a. 输入图像 b. CycleGAN[4] c. UIDGAN[11] d. SRN[3] e. 本文

图 4 模糊图像复原实验结果

3.3.2 混合失真组合图像类型

在本节中，将本文方法在 DIV2K 数据集上与其他经典方法进行比较，包括前文提到的基于域迁移的无监督学习方法 CycleGAN、特定域去模糊算法 UIDGAN，与 RL-Restore, OWAN 这 2 种混合失真图像复原算法进行比较，表 2~表 3 展示了 5 种方法在 PSNR 和 SSIM 指标上的表现. 可以看到，本文方法的 PSNR 和 SSIM 上比 CycleGAN 与 UIDGAN 有了明显的提高，但对于 RL-Restore 与 OWAN 这 2 种使用成对样本进行训练基于监督学习方法，本文方法仍有提升：轻度退化图像复原结果的 PSNR 分别提高了 0.157 dB 和 0.163 dB, SSIM 分别提高了 0.030 和 0.015；中度退化图像复原结果的 PSNR 分别提高了 0.715 dB 和 0.277 dB, SSIM 分别提高了 0.042 和 0.019；重度退化图像复原结果的 PSNR 分别提高了 0.522 dB 和 0.049 dB, SSIM

表 2 5 种混合失真图像复原方法的 PSNR 对比 dB

方法	图像退化		
	轻度	中度	重度
CycleGAN[4]	22.547	21.080	20.412
UIDGAN[11]	26.686	24.197	23.684
RL-Restore[12]	28.047	26.456	25.203
OWAN[13]	28.041	26.894	25.676
本文	**28.204**	**27.171**	**25.725**

注. 粗体表示最优结果.

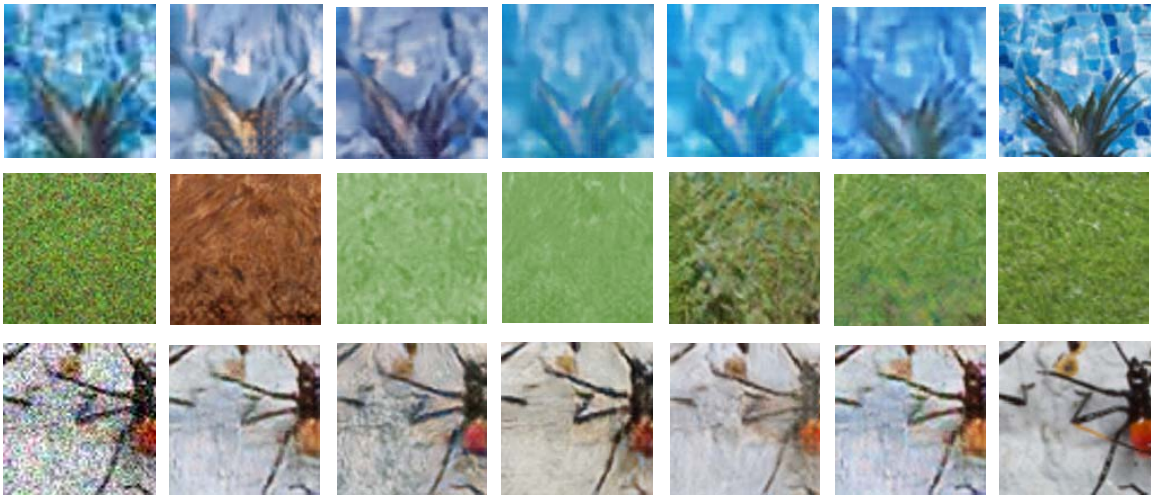
表 3 5 种混合失真图像复原方法的 SSIM 对比

方法	图像退化		
	轻度	中度	重度
CycleGAN[4]	0.490	0.408	0.330
UIDGAN[11]	0.576	0.531	0.476
RL-Restore[12]	0.731	0.656	0.592
OWAN[13]	0.746	0.679	0.617
本文	**0.761**	**0.698**	**0.628**

注. 粗体表示最优结果.

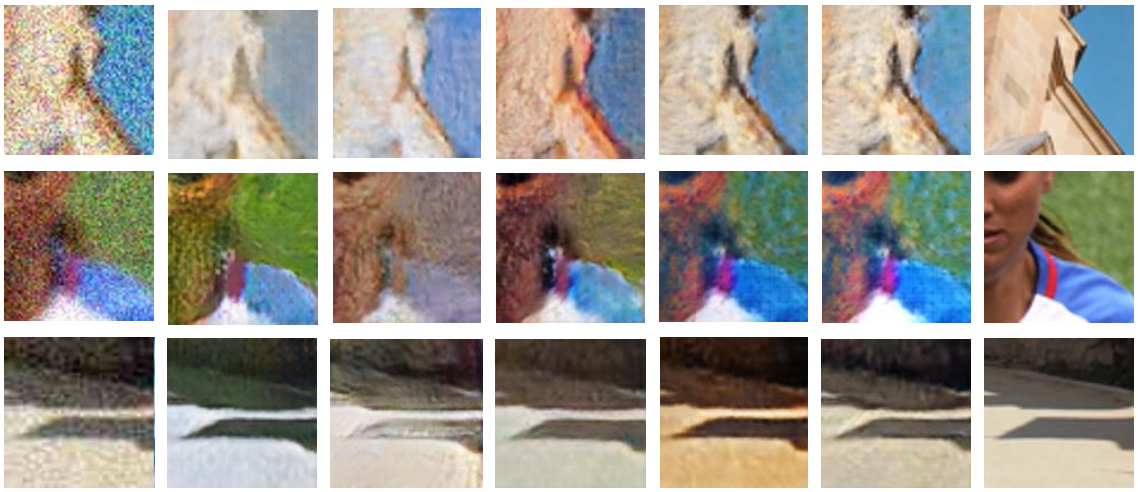
分别提高了 0.036 和 0.011. 这说明了本文方法中解耦模块确实能够有效地减少各种退化机制之间的相互影响，使其表达独立，从而更好地针对混合失真组合的图像进行复原.

图 5~图 7 所示分别为以上几种方法在不同退化程度的图像复原结果. 可以看出，在图 5 中，对于 mild 类型，RL-Restore 重建的图像前景背景没有区分开，对于纹理不清晰的输入图像，其倾向于把整个图像平滑化，忽略了原图像中的一些内容信息. 在基于域迁移的无监督学习方法 CycleGAN 和 UIDGAN 的复原结果中，可以明显看出退化机制之间存在相互干扰，复原图像平滑中又带一点压缩的伪影. 但经过本文方法提出的解耦模块后就能够减少这种退化机制互相影响的可能性，加上通道注意力对经过处理的解耦特征赋予权重，再直接输入到清晰图像生成器，利用其和清晰图像判别器相互对抗来生成清晰图像. OWAN 复原图像仍保留压缩的伪影，效果不理想. 图 6 中，对于 moderate 类型，输入图像纹理大致可见，但在 UIDGAN 复原的图像中，前景和背景几乎融为一体，而 OWAN 较好地区分了前、背景，但整体与清晰图像仍有一点偏差. 相较而言，本文方法在此类型的还原效果最佳. 在图 7 中，对于 severe 类型，由于输入图像受损严重，复原较为困难. 在 CycleGAN, UIDGAN 和 RL-Restore 复原的图像中，芒果已经糊成了一团，OWAN 虽然还原了图像大致结构，但背景与原图仍有差异；本文方法也存在同样的问题，但对输入图像中压缩和模糊的体现，例如，芒果个体之间的阴影以及最后一幅图中管道颜色的交错，处理效果相对较好.



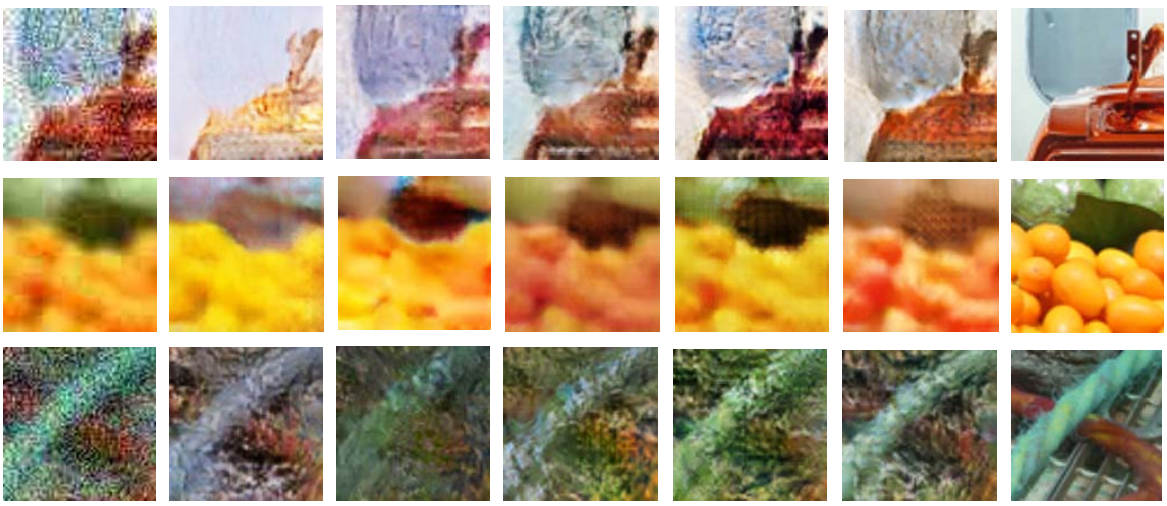
a. 输入图像　b. CycleGAN[4]　c. UIDGAN[11]　d. RL-Restore[12]　e. OWAN[13]　f. 本文　g. 清晰图像

图 5　不同方法对轻度退化图像 mild 复原实验结果



a. 输入图像　b. CycleGAN[4]　c. UIDGAN[11]　d. RL-Restore[12]　e. OWAN[13]　f. 本文　g. 清晰图像

图 6　不同方法对中度退化图像 moderate 复原实验结果



a. 输入图像 b. CycleGAN[4] c. UIDGAN[11] d. RL-Restore[12] e. OWAN[13] f. 本文 g. 清晰图像

图 7 不同方法对重度退化图像 severe 复原实验结果

然后，使用包含雨、压缩、噪声和模糊 4 种退化机制的数据集进行实验，本文模型在经过重新训练后与其他方法再次进行了比较，对比实验的客观评价数据如表 4 所示.

表 4 不同方法对混合失真组合图像复原对比实验结果

方法	PSNR/dB	SSIM
CycleGAN[4]	21.220	0.435
UIDGAN[11]	24.062	0.548
RL-Restore[12]	26.881	0.708
OWAN[13]	26.113	0.712
本文	**27.121**	**0.723**

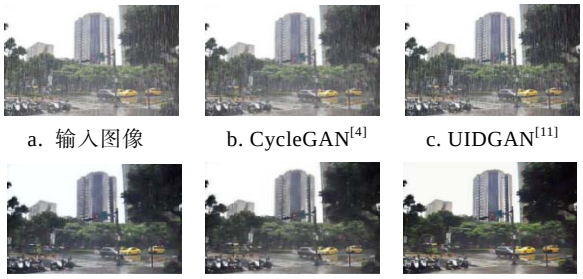
注. 粗体表示最优结果.

图 8 所示为包含 4 种混合失真组合图像数据集上的实验结果. 可以看出，CycleGAN 与 UIDGAN 重建的图像没有成功去除掉雨纹；RL-Restore 虽去除了雨纹，但对于建筑物上的雨纹，该方法的视觉效果如同直接进行了平滑处理，并且图像整体色彩加深，与输入图像出现了偏差. OWAN 有与之类似的问题：大楼上的纹理没有处理到位，但该方法成功地还原了图上细节部分，如左下角摩托车与交通信号灯. 本文方法对细节和纹理的恢复处理效果良好，还原了原图的色彩. 这说明 RL-Restore 与 OWAN 还不能较好地处理复杂、大尺寸的混合失真. 本文方法在客观评价指标与主观视觉质量上均获得了较好的结果.

3.4 消融实验

3.4.1 模块适应性消融实验

本文提出的 FDM 主要是为了降低不同退化机制之间的相互影响，通道注意力的作用则是在处理后的解耦特征表示中自适应地选择有用的特征表示. 为了探究本文所提出的 FDM 和注意力机制各自的效果，分别进行了消融实验. 本节使用 DIV2K 数据集中 moderate 类型的图像进行实验. 针对注意力机制的具体做法是将本文算法中的注意力权重设置为一个固定值，在相同环境下重新训练网络模型并进行测试.



a. 输入图像 b. CycleGAN[4] c. UIDGAN[11] d. RL-Restore[12] e. OWAN[13] f. 本文

图 8 4 种混合失真组合图像复原对比实验结果

消融实验包括：(1) 不使用含 FDM 与注意力机制的基线模型 UIDGAN；(2) 只使用含注意力机制的模型；(3) 只使用含有 FDM(固定注意力权重)的模型；(4) 本文方法. 表 5 所示为在 moderate 类型的图像数据集上进行的消融实验对比结果. 通过对比可以看出，在相同条件下，使用注意力机制的算法比不使用注意力机制的算法在中度退化程度的测试集上取得了更好的 PSNR 和 SSIM 指标，同样地，FDM 也能提升算法的复原效果，但单独使用 2 个模块对算法的提升效果弱于本文方法，即同时使用 2 个模块对图像进行复原的效果更佳.

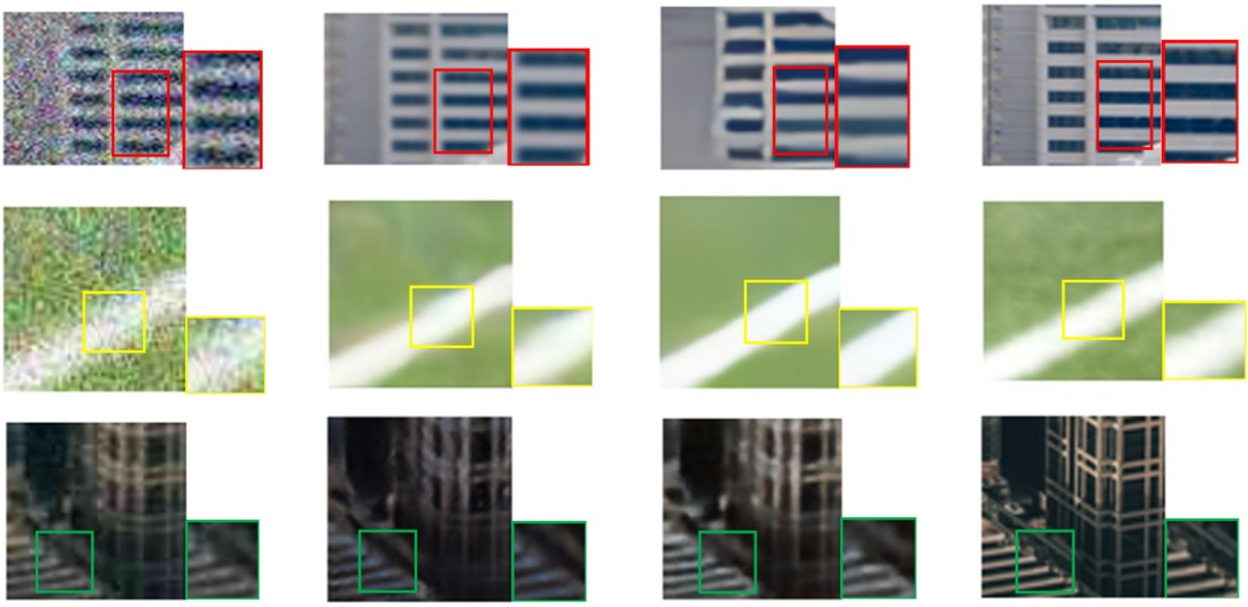
3.4.2 特征正交性损失消融实验

为了验证特征正交性损失对本文算法复原结

表 5　moderate 类型消融实验结果

FDM	通道注意力	PSNR/dB	SSIM
		24.197	0.531
	√	24.263	0.545
√		25.863	0.617
√	√	27.171	0.698

果的影响，分别在相同的实验环境下进行使用特征正交性损失和不使用特征正交性损失的消融实验．将使用对抗损失、循环一致性损失和感知损失函数训练的复原结果，与同时使用这 3 种损失和特征正交性损失训练的复原结果进行对比，如图 9 所示．



a. 输入图像　b. 不使用特征正交性损失的复原结果　c. 使用特征正交性损失的复原结果　d. 清晰图像

图 9　本文方法使用和不使用特征正交性损失的复原结果对比

其中，图 9b 为只使用对抗损失、循环一致性损失和感知损失函数函数训练的复原结果；图 9c 为使用对抗损失、循环一致性损失、感知损失和特征正交性损失函数训练的复原结果．可以看出，图 9c 第 1 行使用特征正交性损失的复原结果在建筑物横向处明显比图 9b 第 1 行不使用特征正交性损失的复原结果具有更清晰的细节结构；图 9c 第 2 行的复原图像比图 9b 第 2 行的复原图像在直线突变处更锐利；图 9c 第 3 行的复原图像比图 9b 第 3 行对比度更强，并且更接近图 9d 第 3 行的清晰图像．综上，复原图像的视觉效果验证特征正交性损失的有效性．

3.4.3　不同特征通道数的实验结果

表 6 所示为对于 moderate 退化类型特征解耦后设置不同特征通道数的复原图像的 PSNR 和 SSIM 指标值．特征通道数可设置为 16，32，48 和 64，实验发现，当通道数设置为 32 时，评价指标值最优．

表 6　不同特征通道数实验结果

通道数	PSNR/dB	SSIM
16	26.934	0.587
32	27.171	0.698
48	27.093	0.677
64	27.098	0.682

3.5　鲁棒性分析

为了在更广泛的失真范围上测试本文方法，评估其在与训练集不同的失真强度上的性能，同时验证本文提出的 FDM 在模型中的效果，证明其重要性．本文基于 DIV2K 数据集创建了一个新的训练集：按照范围为[0,20]的标准差随机采样高斯噪声、范围为[60,100]的质量随机采样 JPEG 压缩、从[10,40]的范围中随机采样运动模糊的最大轨迹长度，将采样到的 3 种退化机制添加到原始数据集图像中．然后，将训练好的模型再应用于具有不同失真参数范围的测试集：采样高斯噪声[20,40]，模糊轨迹长度[40,80]，JPEG 压缩[15,60]．实验结果如表 7 所示，这表明本文方法优于基线方法 UIDGAN．

表 7　鲁棒性分析实验结果

方法	PSNR/dB	SSIM
UIDGAN[11]	19.401	0.434
本文	**22.167**	**0.561**

注．粗体表示最优结果．

4 结 语

为了解决混合失真图像中各种退化机制相互影响、无法充分解耦的情况，本文提出了一种基于无监督对偶学习的图像复原方法，结合 GAN 与编码器进行混合失真组合图像的复原工作，引入了 FDM，实现了通道上的特征解耦，使得不同退化特征表达相互独立，从而提升了复原效果. 在实验中，对实验结果客观数据及主观图像进行了评估和分析，并用消融实验验证了模块的作用，充分说明了本文方法对于混合失真图像复原任务的有效性.

考虑基于无监督学习的混合失真图像复原网络结构比较复杂，并且真实场景复杂多变，未来将进一步丰富实验数据类型，以提高网络泛化性能，同时改进模型结构，减少参数量和模型训练时间，降低时间成本.

参考文献(References):

[1] Goodfellow I J, Pouget-Abadie J, Mirza M, *et al*. Generative adversarial nets[C] //Proceedings of the 27th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2014: 2672-2680

[2] Pathak D, Krähenbühl P, Donahue J, *et al*. Context encoders: feature learning by inpainting[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2016: 2536-2544

[3] Tao X, Gao H Y, Shen X Y, *et al*. Scale-recurrent network for deep image deblurring[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2018: 8174-8182

[4] Zhu J Y, Park T, Isola P, *et al*. Unpaired image-to-image translation using cycle-consistent adversarial networks[C] //Proceedings of the IEEE International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2017: 2242-2251

[5] Lee H Y, Tseng H Y, Huang J B, *et al*. Diverse image-to-image translation via disentangled representations[C] //Proceedings of the European Conference on Computer Vision. Heidelberg: Springer, 2018: 36-52

[6] Liu M Y, Tuzel O. Coupled generative adversarial networks[C] //Proceedings of the 30th International Conference on Neural Information Processing Systems. New York: ACM, 2016: 469-477

[7] Liu M Y, Breuel T, Kautz J. Unsupervised image-to-image translation networks[C] //Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: ACM, 2017: 700-708

[8] Huang X, Liu M Y, Belongie S, *et al*. Multimodal unsupervised image-to-image translation[C] //Proceedings of the European Conference on Computer Vision. Heidelberg: Springer, 2018: 179-196

[9] Almahairi A, Rajeswar S, Sordoni A, *et al*. Augmented CycleGAN: learning many-to-many mappings from unpaired data[C]//Proceedings of the 35th International Conference on Machine Learning. New York: PMLR, 2018: 195-204

[10] Choi Y, Choi M, Kim M, *et al*. StarGAN: unified generative adversarial networks for multi-domain image-to-image translation[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2018: 8789-8797

[11] Lu B Y, Chen J C, Chellappa R. UID-GAN: unsupervised image deblurring via disentangled representations[J]. IEEE Transactions on Biometrics, Behavior, and Identity Science, 2020, 2(1): 26-39

[12] Yu K, Dong C, Lin L, *et al*. Crafting a toolchain for image restoration by deep reinforcement learning[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2018: 2443-2452

[13] Suganuma M, Liu X, Okatani T. Attention-based adaptive selection of operations for image restoration in the presence of unknown combined distortions[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2019: 9031-9040

[14] Bai Liang, Liu Hui, Shang Zhenhong. Mixed degraded image restoration algorithm based on adaptive fusion of hierarchical features[J]. Journal of Computer-Aided Design & Computer Graphics, 2021, 33(2): 215-222(in Chinese)
(白亮，刘辉，尚振宏. 自适应融合层级特征的混合退化图像复原算法[J]. 计算机辅助设计与图形学学报，2021，33(2): 215-222)

[15] Bodin E, Malik I, Ek C H, *et al*. Nonparametric inference for auto-encoding variational Bayes[OL]. [2022-05-20]. https://api.semanticscholar.org/CorpusID:4661887

[16] Li X Q, Chen L B, Wang L, *et al*. SCGAN: disentangled representation learning by adding similarity constraint on generative adversarial nets[J]. IEEE Access, 2018, 7: 147928-147938

[17] Higgins I, Matthey L, Pal A, *et al*. Beta-vae: learning basic visual concepts with a constrained variational framework[OL]. [2022-05-20]. https://api.semanticscholar.org/CorpusID: 46798026

[18] Patacchiola M, Fox-Roberts P, Rosten E. Y-Autoencoders: disentangling latent representations via sequential encoding[J]. Pattern Recognition Letters, 2020, 140: 59-65

[19] Bianco S, Celona L, Napoletano P. Disentangling image distortions in deep feature space[J]. Pattern Recognition Letters, 2021, 148: 128-135

[20] Schwartz O, Simoncelli E P. Natural signal statistics and sensory gain control[J]. Nature Neuroscience, 2001, 4(8): 819-825

[21] Woo S, Park J, Lee J Y, *et al*. CBAM: convolutional block attention module[C] //Proceedings of the European Conference on Computer Vision. Heidelberg: Springer, 2018: 3-19

[22] Li X, Jin X, Lin J X, *et al*. Learning disentangled feature representation for hybrid-distorted image restoration[C] //Proceedings of European Conference on Computer Vision. Heidelberg: Springer, 2020: 313-329

[23] Timofte R, Agustsson E, Van Gool L, *et al*. Ntire 2017 challenge on single image super-resolution: methods and results[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Workshops. Los Alamitos: IEEE Computer Society Press, 2017: 1110-1121

[24] Sun J, Cao W F, Xu Z B, *et al*. Learning a convolutional neural network for non-uniform motion blur removal[OL]. [2022-05-20]. https://arxiv.org/abs/1503.00593

[25] Qian R, Tan R T, Yang W H, *et al*. Attentive generative adversarial network for raindrop removal from a single image[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2018: 2482-2491