

多路径生成对抗网络的红外与可见光图像融合^{*}

许光宇 陈浩宇 张 杰

(安徽理工大学计算机科学与工程学院 淮南 232001)

摘要: 生成对抗网络在红外与可见光图像融合领域受到广泛关注,但单路径进行融合容易丢失浅层信息、分支路特征提取融合能力有限。提出一种基于多路径生成对抗网络的红外与可见光图像融合方法。在生成器端,利用源图像与导向滤波结果构建3条输入路径提取更多源图像特征信息,以获得细节更丰富的融合图像;然后,卷积层加入掩码注意力机制模块,提升显著信息的提取效率,引入密集连接和残差连接,在提升特征传递效率的同时可获取更多源图像重要特征信息。在鉴别器端,采用双鉴别器估计红外与可见光图像的区域分布,避免单鉴别器网络丢失对比度信息的模态失衡问题。在TNO数据集上进行了实验,实验结果表明,所提算法在5个客观评估指标上4项取得了最好结果,优于多数主流算法,在主观评估方面,所提算法保留了更多的纹理细节信息,具有更好的视觉效果。

关键词: 图像融合;生成对抗网络;浅层特征提取;导向图像滤波;双鉴别器

中图分类号: TP391.41 **文献标识码:** A **国家标准学科分类代码:** 520.4

Multipath generative adversarial network for infrared and visible image fusion

Xu Guangyu Chen Haoyu Zhang Jie

(School of Computer Science and Engineering, Anhui University of Science and Technology, Huainan 232001, China)

Abstract: Generative adversarial network has received widespread attention in the field of infrared and visible image fusion, but single-path fusion is prone to lose shallow information and the ability to branch feature extraction is limited. This paper proposes a fusion method for infrared and visible images based on multi-path generative adversarial networks. In the generator, three input paths are constructed using the source images and the results of guided image filter to extract more source image feature information to obtain detailed and rich fused images. Then, the convolutional layer adds an extract mask attention module to improve the efficiency of extracting significant information. In addition, dense connections and residual connections are introduced for improving the efficiency of feature transmission and obtaining more important feature information of the source image. In the discriminator, to avoid the modal imbalance problem of losing contrast information in a single discriminator network, dual discriminators are used to estimate the regional distribution of infrared and visible light images. Experiments are performed on the TNO dataset, and the experimental results show that the proposed algorithm achieves the best results in four of the five objective evaluation indicators and outperforms most mainstream algorithms. In terms of subjective evaluation, the proposed algorithm retained more texture detail information and has better visual effects.

Keywords: image fusion; GAN; shallow feature extraction; guided image filtering; dual discriminators

0 引言

图像融合是指将不同从传感器获得的图像信息进行

互补综合,满足人们对更全面信息量的需求。图像融合技术常用于成像方式各有不同的红外与可见光图像、遥感图像和医学图像的融合。红外图像由场景热辐射成像得到,

收稿日期:2023-09-23

^{*} 基金项目:国家自然科学基金(61471004)、安徽理工大学博士基金(ZX942)、安徽理工大学研究生创新基金(2022CX2125)项目资助

受天气影响较小,热目标较为显著,但由于传感器的限制,背景模糊,纹理缺失。可见光图像其纹理细节较为丰富,边缘信息较为清晰,但会受到大气和光照等因素影响图像质量。红外与可见光图像融合是将红外图像中的主要热目标信息与可见光图像的纹理边缘信息进行互补,提高图像的清晰度与信息量,生成热目标显著、纹理细节丰富的融合图像。红外和可见光图像融合技术被广泛应用于军事监视、目标检测、夜间导航等领域^[1]。

红外与可见光图像融合技术发展多年,主要分为传统方法和基于深度学习的方法。传统的图像融合方法可分为通过直接对象图像像素值进行操作的空域方法^[2-6]、将图像先分解为高频低频系数再设计融合规则的变换域方法^[7-11]、使用稀疏编码机制进行融合的方法^[12-13]、多种方法相结合的混合模型方法^[14-17]。然而,传统方法依赖于固定的数学模型,需要根据手工设计融合规则,缺乏自适应能力。

随着深度学习技术的快速发展,研究者提出了各种各样基于深度学习的红外与可见光图像融合方法^[18-21]。Prabhakar等^[22]首次提出一基于无监督的深度学习图像融合框架(DeepFuse)。Ma等^[23]在红外和可见光图像融合生成对抗网络(FusionGAN)中首次将图像融合问题定义为生成器与鉴别器之间的对抗博弈。后续, Ma等^[24]又使用红外图像鉴别器和可见光图像鉴别器提高融合图像中的两幅特征信息量。文献^[25]引入注意力机制,突出红外与可见光图像中的显著区域。Long等^[26]提出结合残差网络和密集网络结构优势,通过综合特征提取和组合,自动估计对应源图像的信息保存程度,实现有效融合。

目前,生成对抗网络(generative adversarial networks, GAN)中用于红外与可见光图像融合的方法中,生成器的输入主要分两种,一种方法是对于输入的红外图像与可见光图像,分支路单独进行特征提取,可以较为全面的提取到各自源图像的有用信息,降低了冗余信息的干扰。然而这种分离特征提取与融合阶段的方法,用于融合特征的层级较少,无法得到更好的融合结果。另外一种方法是在输入端将两幅源图像进行拼接作为输入,同时提取两幅源图像对的特征信息,并进行特征融合,这种方法会导致网络模型丢失源图像浅层特征信息。为了更充分利用图像的特征信息,本文提出一种基于多路径的生成对抗网络体系结构,通过在TNO数据集上的主客观评价分析以及消融实验表明,所提的红外与可见光图像融合方法取得较好的效果,优于相比较的融合方法。

在生成器中,通过导向滤波处理的源图像作为多路径输入的一部分,通过两条浅层特征提取路径,与一条深层特征提取融合路径,有效的保留红外图像中的微小纹理信息和红外图像中的显著信息。引入密集网络提高特征重用率,避免过拟合。加入残差网络保留更多的浅层特征信息。在各路径中加入掩码注意力机制模块,使网络模型突

出图像中的显著目标。在鉴别器中,使用双鉴别器与生成器进行对抗训练,使融合图像保留更多红外与可见光源图像中的信息,防止融合结果偏向单一的红外图像或可见光图像。

1 相关算法

1.1 FusionGAN

FusionGAN算法将红外与可见光图像的融合定义为保持红外热辐射信息与保持可见光纹理细节信息的对抗过程,具体过程如图1所示。

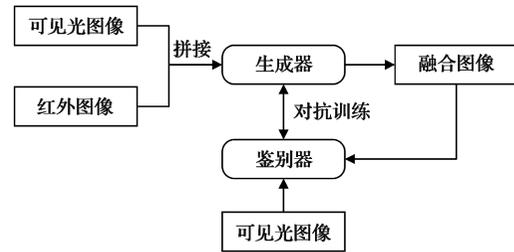


图1 FusionGAN网络结构

红外图像和可见光图像拼接输入到生成器中,生成器进行特征提取、融合、降维,得到融合图像。通过可见光图像对鉴别器进行训练,使得鉴别器获得判定可见光图像的能力。融合图像传入鉴别器,鉴别器对其进行判别,并将结果反馈给生成器。为了生成鉴别器无法判别的融合图像,生成器不断优化,与鉴别器之间产生了对抗。经过大量训练之后,生成器生成的融合图像会包含越来越多的源图像信息。然而FusionGAN生成器网络过于简单导致特征提取不充分,融合结果信息丢失严重,且只使用可见光鉴别器使得融合图像的信息分布偏向于可见光图像而导致模态失衡。

1.2 导向滤波

导向滤波^[27]是一种基于局部线性模型的边缘保持滤波器,利用导向图指导滤波过程,可以实现图像的边缘保持和结构传递功能。给定输入图像 P ,导向滤波将输出图像 O 假定为导向图像 I 的线性变换,即:

$$O_j = a_i I_j + b_i, \forall j \in \omega_i \quad (1)$$

式中: I_j, O_j 分别为输入、输出图像的像素, j 为像素索引; a_i 和 b_i 是当窗口位于中心时该线性函数的系数。 ω_i 是以导向图像 I 中一点 i 为中心,大小为 $(2r+1) \times (2r+1)$ 的局部窗口。像素点 i 包含在相邻不同的窗口 ω_i 内,故 O_j 的值会根据窗口 ω_i 的变化而改变。为了求解线性系数 a_i 和 b_i ,并满足输入和输出图像之间差别最小,转化为最优化问题求解。

$$E(a_i, b_i) = \sum_{j \in \omega_i} ((a_i I_j + b_i - P_j)^2 + \epsilon a_i^2) \quad (2)$$

式中: ϵ 为归一化因子。使用线性回归求解公式的解:

$$a_i = \frac{\frac{1}{|\omega|} \sum_{j \in \omega_i} I_j P_j - \mu_i \bar{P}_i}{\sigma_i^2 + \epsilon_1} \quad (3)$$

$$b_i = \bar{P}_i - a_i \mu_i \quad (4)$$

式中： μ_i 为向导图 I 均值； σ_i^2 表示 I 的方差； \bar{P}_i 表示 P 在窗口 ω_i 中的均值； $|\omega_i|$ 表示窗口内的像素数。

通过导向滤波的源图像如图 2 所示,GF_IR 为红外图像为输入图像,可见光图像为引导图像,经过导向滤波处理后的图像,GF_VI 为可见光图像作为为输入图像,红外图像作为引导图像,经过导向滤波处理后的图像。GF_IR 可以保留可见光源图像的微小纹理细节,GF_VI 可以保留红外源图像中人物的显著信息。

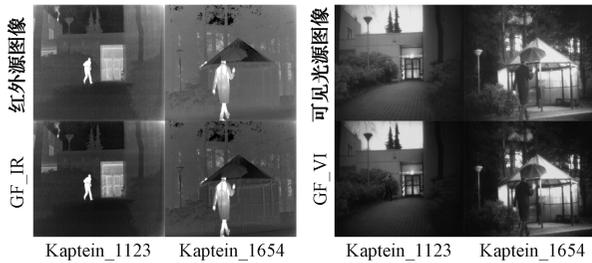


图 2 源图像与导向滤波图像

2 本文算法

2.1 融合框架

为了更有效提取红外图像 IR 中的对比度信息和可见光图像 VI 中丰富的纹理细节信息,同时增加用于融合的层级,得到更好的融合结果。本文提出一种基于多路径的生成对抗网络体系结构,算法流程如图 3 所示。通过将导向滤波处理后的源图像 GF_IR 和 GF_VI 作为输入的一部分,分别与源图像 IR 和 VI 进行拼接联结输入,可以充分的提取通常被忽略的特征信息。本文采用双鉴别器同时估计可见光和红外光两个不同域的概率分布。概率分布的一致性会使融合结果具有目标分布中最显著的特征,生成器更有可能捕获关键特征并增强它们。多路径图像输入通过生成器,得到融合图像。将融合图像传入鉴别器,两个鉴别器分别判断融合图像是红外或可见光图像或是融合图像的概率。两个鉴别器分别将概率反馈给生成器,生成器不断优化生成的融合图像。通过生成器与两个鉴别

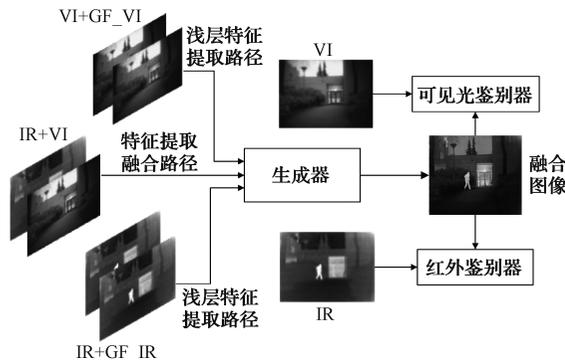


图 3 整体融合框架

器之间的不断训练,最终生成器可以生成被红外鉴别器判断为红外图像,被可见光鉴别器判定为可见光图像的融合图像。融合图像对比度显著、纹理细节丰富,包含了红外与可见光图像中的有用信息。

2.2 生成器结构

本文的生成器结构如图 4 所示,输入分为可见光图像浅层特征提取路径,红外图像浅层特征提取路径和特征提取融合路径。两条浅层特征提取路径主要用于提取纹理细节和边缘特征信息。特征提取融合路径,在保留图像的浅层和深层特征信息的同时可以使得源图像的特征提取与融合阶段相统一。在两个浅层特征提取路径中,采用 2 个 5×5 的卷积核以及 2 个 3×3 的卷积核进行特征提取,并且进行归一化批处理(batch normalization, BN)和带漏线性整流函数(leaky rectified linear unit, LeakyReLU)进行激活。特征提取融合路径中,采用 2 个 5×5 的卷积核以及 3 个 3×3 的卷积核进行特征提取,通过密集连接,使得每一层都与通道维度中的所有之前层连接起来,作为下一层的输入,以提取更多的有效特征信息。为了更好的保留图像的浅层特征。在 3 条路径的第 2 层加入掩码注意力机制(extract mask attention module, EM),对输入的特征图进行自适应特征细化,保留更多显著特征信息。各条路径使用 3×3 的卷积核进行残差连接,保留源图像中的浅层信息。最后,3 条路径特征图通过 concat 拼接,最后使用 1×1 的卷积核,通过双曲正切函数(Tanh)激活,将连接的特征降维为单通道图像,实现特征融合。生成器结构中所有步长均设置为 1。

受文献[28-29]启发,本文将掩码注意力网络,用于图像融合任务中。通过掩码注意力机制,提取红外和可见光图像中的高频信息。在掩码注意力模块中,输入特征图经过激励与挤压网络(squeeze and excitation, SE),以加权特征图上的重要通道信息。再将加权的特征信息输入到特征提取模块和掩码模块,得到的特征图逐点相乘,再与通过 SE 模块加权的特征图逐点相加,得到最终结果。其结构如图 5 所示。

特征提取模块如图 6 所示,通过卷积与反卷积相结合进行特征提取,可以得到细节信息丰富的特征图。输入特征图 F , 首先通过 4×4 的反卷积层,得到 F^{0}_{trans} , 再经过 6×6 的卷积层得到 F^{0}_{conv} :

$$F^{0}_{trans} = f_{trans_conv}(F) \quad (5)$$

$$F^{0}_{conv} = f_{conv}(F^{0}_{trans}) \quad (6)$$

通过残差连接, F 与 F^{0}_{conv} 逐点相减再通过反卷积层,得到 F^{1}_{trans} 。最后与 F^{0}_{trans} 逐点相加得到输出 F_{out} :

$$F^{1}_{trans} = f_{trans_conv}(F - F^{0}_{conv}) \quad (7)$$

$$F_{out} = f_{conv}(F^{0}_{trans} + F^{1}_{trans}) \quad (8)$$

卷积与反卷积相结合进行特征提取,可以使得普通卷积层提取不到的特征信息得到保留。经过 2 次反卷积与卷积后,红外与可见光图像的高频细节信息被最大限度提取。

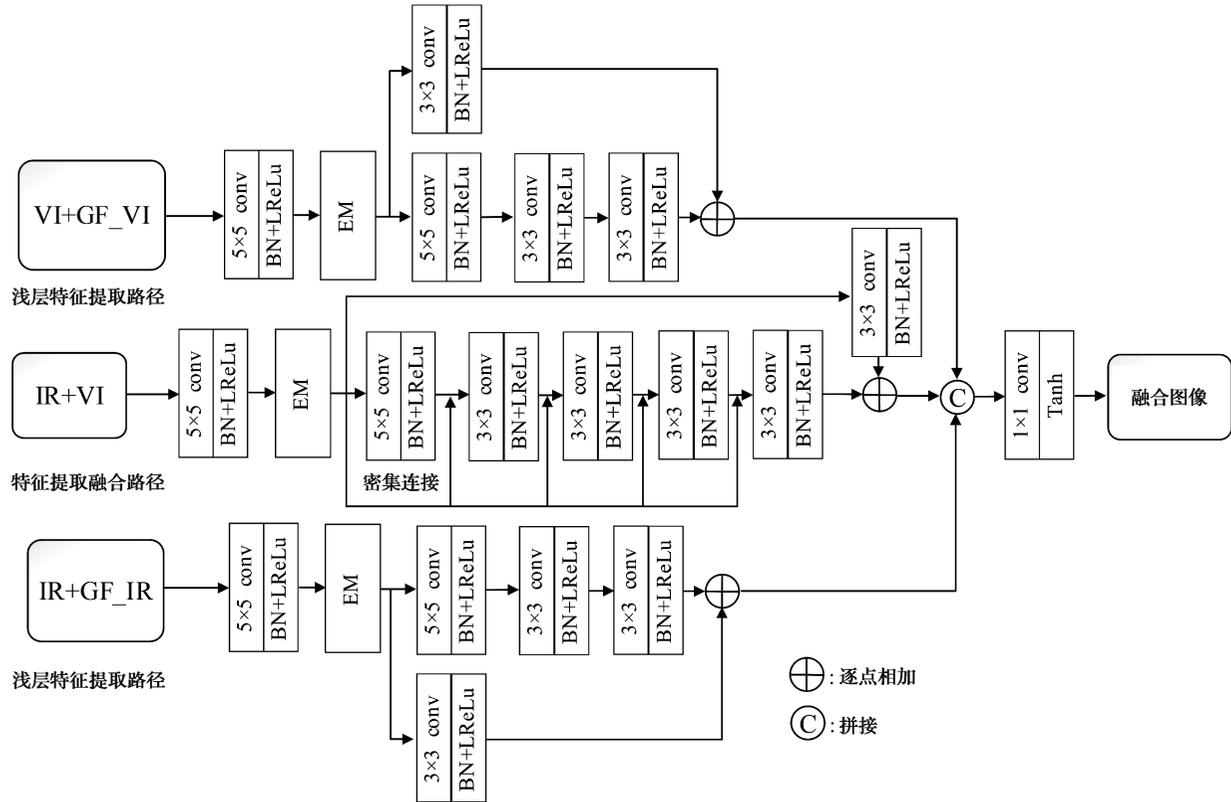


图4 生成器结构

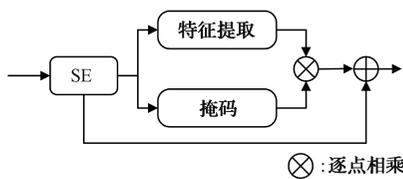


图5 掩码注意力机制

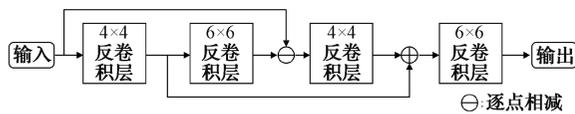


图6 特征提取模块

掩码模块如图7所示,输入的特征通过 3×3 卷积和参数线性整流函数(parametric rectified linear unit, PReLU)激活后,再经过平均池化对通道信息进行压缩。之后通过 1×1 的卷积层和PReLU激活函数。最后通过 1×1

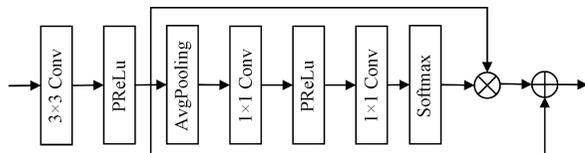


图7 掩码模块

的卷积层,以及归一化指数函数(soft maximum, Soft-max)激活加权特征通道。

2.3 鉴别器结构

本文提出的网络结构拥有两个鉴别器,红外光鉴别器 D_{ir} 和可见光鉴别器 D_{vi} , D_{ir} 用于区分传入图像是红外图像或是融合图像, D_{vi} 用于区别传入图像是可见光图像或是融合图像,它们有着相同的结构,如图8所示。图像传入鉴别器,生成一个概率标量,分别估计输入图像是红外图像或可见光图像或是融合图像的概率。将鉴别器中使用4个 3×3 卷积核,卷积后进行BN处理,使用LeakyReLU函数作为激活函数。最后一层使用全连接层进行分类,最终生成一个概率标量。步长均设置为1。

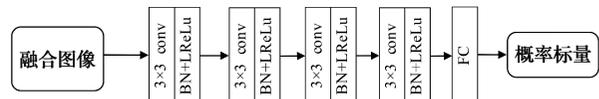


图8 鉴别器结构

2.4 损失函数

本文网络模型的损失函数分别为生成器损失函数 L_G 、红外图像鉴别器损失函数 L_{Dir} 和可见光图像鉴别器损失函数 L_{Dvi} ,红外图像鉴别器 D_{ir} 和可见光图像鉴别器 D_{vi} 与生成器G进行对抗,两个鉴别器分别将红外图像 I_{ir} 和可见光图像 I_{vi} 视为真实数据,融合图像 I_{fused} 作为 D_{ir}

和 D_{vi} 的假数据, 本文将两幅源图像的导向滤波图像 I_{gf_ir} 和 I_{gf_vi} 作为辅助输入以提高融合图像质量, 因此本文引入了导向滤波图像的强度损失和梯度损失作为辅助损失。

1) 生成器损失函数

生成器损失函数由生成器内容损失 L_{con} 和生成器对抗损失 L_{adv} 两部分组成, L_{con} 用于约束信息提取, 内容损失越小, 融合图像中的源图像细节信息越丰富。 L_{adv} 则用于约束信息平衡。

$$L_G = \gamma L_{con} + L_{adv} \quad (9)$$

其中, γ 是正则化参数, 为了维持内容损失与对抗损失之间的平衡关系, 本文设置为 100。内容损失 L_{con} 由梯度损失、强度损失和结构相似性误差损失构成:

$$L_{con} = L_{grad} + L_{Intensity} + L_{ssim} \quad (10)$$

梯度信息作为从可见光图像获得的主要信息, 考虑红外图像也包含部分纹理细节, 故梯度损失 L_{grad} 定义如下:

$$L_{grad} = \alpha_1 \|\nabla I_{fused} - \nabla I_{vi}\|_F^2 + \alpha_2 \|\nabla I_{fused} - \nabla I_{ir}\|_F^2 + \|\nabla I_{fused} - \nabla I_{gf_vi}\|_F^2 + \|\nabla I_{fused} - \nabla I_{gf_ir}\|_F^2 \quad (11)$$

式中: ∇ 为二阶梯度算子; $\|\cdot\|_F$ 为 F 范数; α_1, α_2 为常数, 用来调整红外图像与可见光图像主次要梯度信息的参数。

强度分布作为从红外图像中获得的主要信息, 可见光图像同样也包含部分对比度信息, 故强度损失定义为:

$$L_{Intensity} = \beta_1 \|I_{fused} - I_{ir}\|_F^2 + \beta_2 \|I_{fused} - I_{vi}\|_F^2 + \|I_{fused} - I_{gf_ir}\|_F^2 + \|I_{fused} - I_{gf_vi}\|_F^2 \quad (12)$$

式中: β_1 和 β_2 是常数, 对它们进行调整实现各项之间的主次关系。

考虑融合后图像与原始图像的结构相似性, 引入了结构相似性(structural similarity, SSIM) 损失 L_{ssim} , 计算图像 x 和图像 y 之间的亮度、对比度和结构差异的值, 计算公式如下:

$$L_{ssim} = 1 - \frac{SSIM(I_{fused}, I_{vi}) + SSIM(I_{fused}, I_{ir})}{2} \quad (13)$$

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (14)$$

式中: μ 代表均值; σ 代表标准差; c_1 和 c_2 是为了防止分母为 0 的情况。SSIM 越大, 两者之间的结构相似性越高。

为了达到各种信息之间的平衡, 引入生成器与两个鉴别器之间的对抗损失 L_{adv} , 其定义为:

$$L_{adv} = (1 - D_{vi}(I_{fused}^n))^2 + (1 - D_{ir}(I_{fused}^n))^2 \quad (15)$$

式中: $D_{vi}(I_{fused}^n)$ 表示融合后的图像判定为可见光图像的概率; $D_{ir}(I_{fused}^n)$ 表示融合后的图像为判定红外图像的概率; 常数 1 代表生成器期望生成无法被鉴别器的假数据。

2) 鉴别器损失函数

对鉴别器进行训练, 损失函数需要不断提高其鉴别能力, 以区分真实数据和虚假数据, 并输出一个标量, 估计输

入来自真实数据而不是源自生成器的概率。红外图像鉴别器 L_{Dir} 与可见光图像鉴别器 L_{Dvi} 的损失定义如下:

$$L_{Dir} = \frac{1}{N} \sum_{n=1}^N ((D_{ir}(I_{ir}^n) - \alpha_1) + (D_{ir}(I_{fused}^n) - \epsilon_1)) \quad (16)$$

$$L_{Dvi} = \frac{1}{N} \sum_{n=1}^N ((D_{vi}(I_{vi}^n) - \alpha_2) + (D_{vi}(I_{fused}^n) - \epsilon_2)) \quad (17)$$

式中: $D_{ir}(I_{ir}^n)$ 和 $D_{ir}(I_{fused}^n)$ 分别表示红外鉴别器对红外图像和融合图像的预测; $D_{vi}(I_{vi}^n)$ 和 $D_{vi}(I_{fused}^n)$ 表示可见光鉴别器 D_{vi} 对可见光图像和融合图像的预测。为了避免损失, D_{ir} 中红外图像样本尽量判成 1, 融合图像样本尽量判断成 0。同理, D_{vi} 中可见光图像样本尽量判成 1, 融合图像样本尽量判断成 0。故 α_1, α_2 设置为 1, ϵ_1, ϵ_2 设置为 0。

3 实验结果分析

3.1 实验设置

为了验证所提算法的有效性, 选用校准好的 TNO 标准数据集进行测试。TNO 数据集是红外和可见图像融合常用数据集之一, 其包含 60 对军事相关场景的红外和可见光图像, 适用于复杂场景下的图像融合算法研究。本文从 TNO 数据集中共选取 40 对红外与可见光图像组成训练集, 分别选取 10 对图像作为测试集。选取 8 种有代表性的融合方法进行比较, 包括离散小波变换(discrete wavelet transform, DWT)^[30]、基于分类显著性的红外与可见光图像融合(CSF)^[31]、基于密集连接网络的红外与可见光图像融合(Densefuse)^[32]、FusionGAN^[19]、多分类约束生成对抗网络(GANMcC)^[33]、双鉴别器语义监督生成对抗网络(SDDGAN)^[34]、基于梯度和强度比例保持的快速统一图像融合网络(PMGI)^[35]。并且分别通过主观评价和客观评价指标进行对比分析。实验环境为 Windows10、Python3.6 和 Tensorflow1.15.0。硬件配置环境为 CPU Intel i7-12700F, GPU NVIDIA GeForce RTX2080。所有实验均在同一环境下进行。

3.2 TNO 数据集上的实验结果分析

本文选取 TNO 数据集测试结果中的五幅图像进行主观评价分析, 分别为 2_men、Soldier_in_smoke、SandPath、Kaptein_1123、Kaptein_1654。源图像如图 9 所示, 本文算法及其他算法结果如图 10 所示。

1) TNO 数据集实验结果主观评价

融合结果 2_men 中, 本文算法很好的保留了纹理细节与边缘信息, 其余算法天空与屋顶边缘模糊, 且天空与左边树枝末端均产生伪影。融合结果 Soldier_in_smoke 中, FusionGAN 算法丢失可见光图像中的烟雾信息, CSF、DenseFuse、PMGI 算法士兵热目标信息不显著。在图像 SandPath 融合结果中, DWT、FusionGAN、GANMcC、PMGI 算法只保留了红外源图像中的人物丢失了树枝的

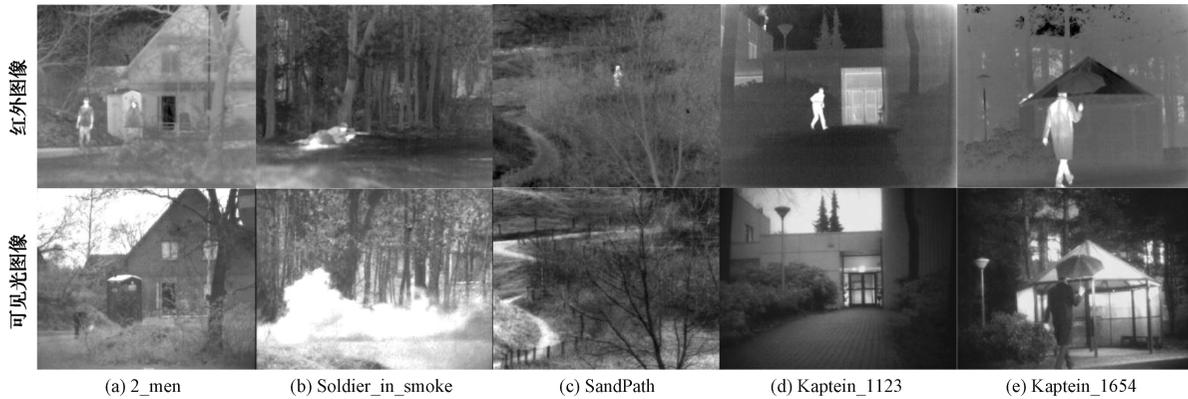


图9 TNO数据集测试源图像

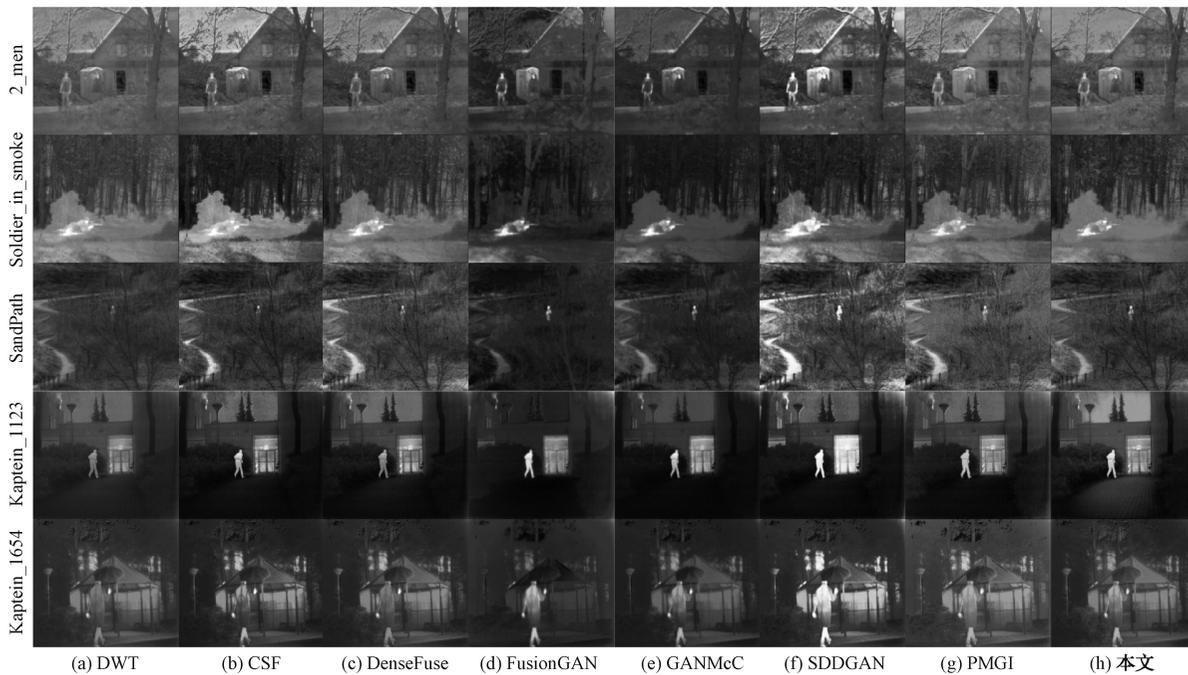


图10 TNO数据集各算法融合结果

纹理细节。融合结果 Kaptein_1123 中, DWT、FusionGAN、GANMcC、SDDGAN、PMGI 算法对草丛纹理信息的提取较差, CSF 与 DenseFuse 算法融合图像中, 人物的热目标不够突出。本文提出算法可以较好的保留草丛的纹理细节并且天空与树木边缘具有较好的区分度, 在很好的保留了人物的红外特征情况下, 依然具有良好的视觉效果。融合结果 Kaptein_1654 中, 除 CSF、DenseFuse 和本文算法外, 其他算法左下角草丛纹理细节丢失严重, 路灯底部与草丛边缘混淆, 右下角的凳子除本文算法外均较为模糊。与上述 7 种方法相比, 本文算法取得了最优的融合效果。导向滤波图像的加入可以有效的表达出红外与可见光图像中的微小细节信息, 两条浅层特征提取路径分别提取边缘细节纹理, 使得本文算法生成的融合图像纹理边缘较为清晰, 生成的融合图像视觉效果较好。掩码注意力

机制的加入使得融合图像的人物热目标较为突出。特征提取融合路径使得融合图像保留了较多源图像的有用信息。

2) TNO 数据集实验结果客观评价

为了更好地分析融合图像的质量, 选取平均梯度 (average gradient, AG)、空间频率 (spatial frequency, SF)、信息熵 (information entropy, EN)、互信息 (mutual information, MI)、SSIM 和峰值信噪比 (peak signal-to-noise ratio, PSNR) 共 6 种客观评价指标对融合图像质量进行分析, 六项指标均为正向指标。分别选取 TNO 数据集 10 组图像的各项指标进行定量比较, 结果如图 11 所示。AG 可以用于衡量融合图像的清晰程度, AG 越大, 图像清晰度越高。SF 反映图像灰度的变化率, SF 越大表示图像越清晰。EN 主要是度量图像包含信息量多少的一个客观量, EN

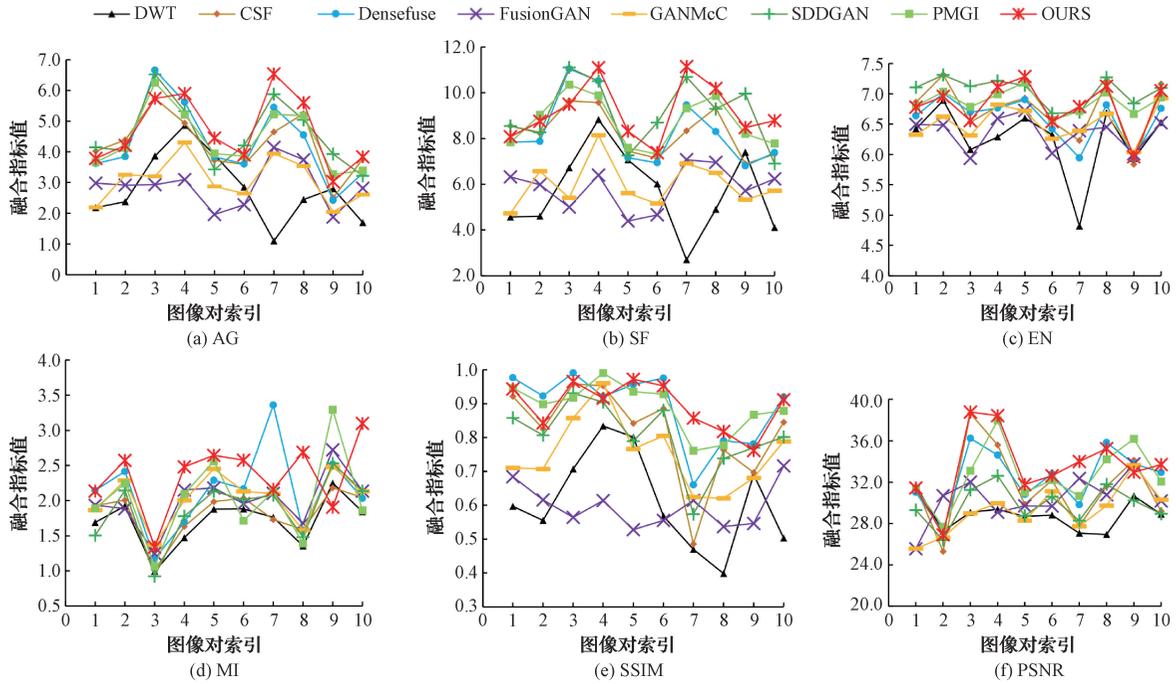


图 11 TNO 数据集 10 对图像 6 项指标定量比较

值越大,表明融合图像中的信息越丰富。MI 可度量两幅图像之间的相似程度,即融合图像获取了源图像信息量的多少。互信息越大,表示融合图像保留更多源图像信息。SSIM 测量两个图像之间的相似性。人类对边缘和纹理的位置非常敏感,SSIM 通过关注边缘和纹理相似性来模仿人类感知,SSIM 值越高代表图像结构越相似。PSNR 用于表示信号的最大可能功率与影响其表示的保真度的破坏噪声的功率之间的比率,PSNR 指标值越高说明图像质量越好。

TNO 数据集各算法 10 组实验数据平均值如表 1 所

示。本文算法在 10 组实验对比中,大多数客观评价指标优于其他对比算法。其中 AG、SF、MI、SSIM 和 PSNR 指标相比其余算法为最优,本文算法 EN 值为次优,略低于 SDDGAN 和 PGMI 算法。与主观评价结果一致,说明本文所提出算法的融合图像也包含有较多的信息量。与主观评价结果一致,AG、SF、MI 取得最优值,表示本文算法融合图像纹理边缘较为清晰,体现了导向滤波输入和浅层特征提取路径的有效性,SSIM 和 PSNR 取得最优,说明融合阶段的层数增加使得网络模型很好的对源图像中的结构信息进行了保留。

表 1 TNO 数据集 10 组图像数据平均值

算法	AG	SF	EN	MI	SSIM	PSNR
DWT	2.810 2	5.686 4	6.256 9	1.704 6	0.612 6	30.441 7
CSF	4.258 0	8.244 8	6.737 8	1.814 1	0.817 5	31.580 5
DenseFuse	4.600 1	8.881 6	6.689 3	1.988 4	0.839 8	31.034 5
FusionGAN	2.976 8	5.870 2	6.320 8	2.022 0	0.597 5	30.383 2
GANMcC	3.060 3	6.004 7	6.507 4	2.040 2	0.752 1	30.597 1
SDDGAN	4.574 4	9.119 2	7.043 4	1.872 5	0.805 1	29.916 4
PMGI	4.414 3	8.718 2	6.868 8	2.024 2	0.890 2	32.649 6
本文	4.700 2	9.168 2	6.815 5	2.359 7	0.894 1	33.579 9

注:加粗字体表示最优值

3.3 消融实验

为了进一步验证本文方法的有效性,将与无导向滤波图像输入的模型(No_GF)、无掩码注意力机制的模型(No_EM)、导向滤波图像输入和掩码注意力机制均没有的模

型结构(No_Both)、只有一条拼接特征提取融合路径的模型(Input_concat)和只有分支路输入浅层特征提取路径的模型(Input_branch)进行比较。消融实验结果如图 12 所示,实验数据如表 2 所示。

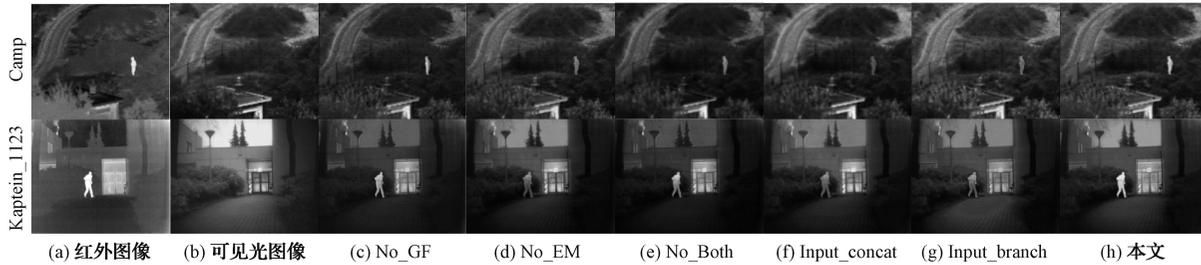


图 12 消融实验结果比较

表 2 消融实验图像数据

算法	AG	SF	EN	MI	SSIM	PSNR
No_GF	4.226 7	8.310 9	6.617 1	2.047 6	0.892 6	33.329 6
No_EM	4.508 6	8.753 1	6.677 4	1.955 0	0.890 7	33.377 0
No_Both	3.005 5	5.999 1	6.581 6	1.823 8	0.795 7	31.530 3
Input_concat	3.202 9	6.236 7	6.792 0	2.101 3	0.810 8	32.897 1
Input_each	4.312 0	8.410 9	6.800 0	2.117 0	0.856 1	31.806 1
本文	4.700 2	9.168 2	6.81 55	2.359 7	0.894 1	33.579 9

注:加粗字体表示最优值

从实验结果可以看出, No_GF 模型由于少了 GF_IR 和 GF_VI 的输入, 原本存在于 GF_IR 中的少量纹理细节和 GF_VI 中的对比度信息未被提取, 导致结果中, 微小细节未被保留。No_EM 模型缺少了掩码注意力机制模块, 高频信息提取不充分, 人物热目标区域不够突出。No_both 模型对红外与可见光的信息保留较差, 背景偏暗, 边缘混滑, 图像较为模糊, No_GF 和 No_both 模型的融合结果视觉效果优于 No_Both 模型。Input_concat 模型缺失两条浅层提取路径提供的浅层信息, 融合结果清晰度较差且热目标不突出。Input_branch 模型用于融合的层级较少, 融合较差, 融合结果 Kaptein_1123 中地面出现较为明显的灰度差。然而 Input_concat 和 Input_branch 模型组合构成的本文算法, 其融合效果却显著优于上述两个模型。

从客观评估结果来看, No_GF 和 No_EM 模型对源图像的信息提取不够充分, 导致各项略低于本文算法。No_Both 各项指标均为最低, 说明引入导向滤波输入和掩码注意力机制对融合结果客观评估指标提升的有效性。Input_branch 模型除 PSNR 值以外各项指标均高于 Input_concat 模型, 表明浅层特征信息对于融合图像各项客观评估指标较为重要, 且将上述两个模型结合的本文算法可以平衡两模型中的不足之处, 显著提升融合图像的指标。综上所述, 客观评估结果与主观评估结果保持一致, 证明本文算法中各结构的有效性, 表明本文算法在融合时可以保留更多源图像信息的同时对于热目标信息的提取也较为突出, 融合结果的对比度较为优异, 视觉效果较好。

4 结 论

本文基于 FusionGAN 算法, 提出一种多路径生成对

抗网络的红外与可见光图像融合的算法, 通过设计一条特征提取融合路径, 增加网络融合的层级。增加两条浅层提取路径, 提取在网络模型中容易丢失的浅层信息。引入导向滤波处理的源图像, 获取源图像中被忽略的信息, 使用掩码注意力机制模块突出显著特征, 利用密集连接提升网络模型的特征传递效率, 加入残差连接保留浅层信息。使用双鉴别器分别对生成器生成的图像进行鉴别, 避免单鉴别器网络中只保留可见光图像纹理, 而丢失红外图像信息的问题。本文所提出算法的融合结果保留了红外与可见光图像中更多的特征信息。通过 TNO 数据集和消融实验的对比表明, 相较于其他算法, 本文算法主观评价均明显优于其余对比算法, 本文提出的算法能够有效减少伪影, 使得边缘信息更加清晰, 具有更高的对比度和更好的纹理细节。客观评价指标多数有所提高, 证明本文算法的有效性。

尽管本文算法在主观视觉和多数评价指标相较于其余主流算法有所提升, 但本文算法由于多路径的存在使得过程较为复杂, 需要更长的运行时间。因此, 下一步的研究工作着重于优化算法的运行效率, 进一步提高算法的性能。

参 考 文 献

- [1] PIELLA G. A general framework for multiresolution image fusion: From pixels to regions[J]. Information Fusion, 2003, 4(4): 259-280.
- [2] WAN T, ZHU C, QIN Z. Multifocus image fusion based on robust principal component analysis [J]. Pattern Recognition Letters, 2013, 34(9): 1001-1008.

- [3] MITIANOUDIS N, STATHAKI T. Pixel-based and region-based image fusion schemes using ICA bases[J]. *Information Fusion*, 2007, 8(2): 131-142.
- [4] JIANG Y, WANG M H. Image fusion with morphological component analysis[J]. *Information Fusion*, 2014, 18(1): 107-118.
- [5] ZHU J, JIN W, LI L. Multiscale infrared and visible image fusion using gradient domain guided image filtering[J]. *Infrared Physics & Technology*, 2018, 89: 8-19.
- [6] 刘明葳, 王任华, 李静, 等. 各向异性导向滤波的红外与可见光图像融合[J]. *中国图象图形学报*, 2021, 26(10): 2421-2432.
- [7] PETROVIĆ V S, XYDEAS C S. Gradient-based multiresolution image fusion[J]. *IEEE Transactions on Image Processing*, 2004, 13(2): 228-237.
- [8] IOANNIDOU S, KARATHANASSI V. Investigation of the dual-tree complex and shift-invariant discrete wavelet transforms on quickbird image fusion[J]. *IEEE Geoscience & Remote Sensing Letters*, 2007, 4: 166-170.
- [9] BHATNAGAR G, WU Q, ZHENG L. Directive contrast based multimodal medical image fusion in NSCT domain[J]. *IEEE Transactions on Multimedia*, 2014, 9(5): 1014-1024.
- [10] GANASALA P, KUMAR V. Multimodality medical image fusion based on new feature in NSST domain[J]. *Biomedical Engineering Letters*, 2015, 4(4): 414-424.
- [11] 羊肇俊, 曾理. 基于加权最小二乘滤波和引导滤波的铸件 DR 图像融合[J]. *仪器仪表学报*, 2021, 42(6): 211-220.
- [12] ZHANG Q, LIU Y. Sparse representation based multi-sensor image fusion for multi-focus and multimodality images: A review[J]. *Information Fusion*, 2018, 40: 57-75.
- [13] YU X, GAO G, XU J, et al. Remote sensing image fusion based on sparse representation[J]. *Acta Optica Sinica*, 2013, 33(4): 259-266.
- [14] LIU Y, LIU S, WANG Z. A general framework for image fusion based on multi-scale transform and sparse representation[J]. *Information Fusion*, 2015, DOI:10.1016/j.inffus.2014.09.004.
- [15] OUYANG N, ZHENG X Y, YUAN H. Multi-focus image fusion based on NSCT and sparse representation[J]. *Computer Engineering and Design*, 2017, 38: 177-182.
- [16] 童林, 官铮, 杨文韬, 等. 潜在低秩表示下 VSM 联合 PCNN 的红外与可见光图像融合[J]. *国外电子测量技术*, 2021, 40(10): 84-90.
- [17] 贺腾飞, 贺利乐, 高党国. 基于 NSST 的改进最大最小滤波与 DCT-LSF 的多聚焦图像融合[J]. *电子测量技术*, 2022, 45(22): 99-105.
- [18] 陈永, 张娇娇, 王镇. 多尺度密集连接注意力的红外与可见光图像融合[J]. *光学精密工程*, 2022, 30(18): 2253-2266.
- [19] 杨莘, 田立凡, 梁佳明. 改进双路径生成对抗网络的红外与可见光图像融合[J]. *电子与信息学报*, 2023, 45(8): 3012-3021.
- [20] 徐慧娴, 黄坤超, 陈明举, 等. 基于深层次多分类生成对抗网络的红外与可见光图像融合方法[J]. *电子测量技术*, 2022, 45(4): 126-133.
- [21] 陈广秋, 温奇璋, 尹文卿, 等. 用于红外与可见光图像融合的注意力残差密集融合网络[J]. *电子测量与仪器学报*, 2023, 37(8): 182-193.
- [22] PRABHAKAR K R, SRIKAR V S, BABU R V. DeepFuse: A deep unsupervised approach for exposure fusion with extreme exposure image pairs[C]. *Proceedings of IEEE International Conference on Computer Vision*. Los Alamitos: IEEE Computer Society Press, 2017: 4724-4732.
- [23] MA J, WEI Y, LIANG P, et al. FusionGAN: A generative adversarial network for infrared and visible image fusion [J]. *Information Fusion*, 2019, 48: 11-26.
- [24] MA J, XU H, JIANG J, et al. DDcGAN: A dual-discriminator conditional generative adversarial network for multi-resolution image fusion[J]. *IEEE Transactions on Image Processing*, 2020, 29: 4980-4995.
- [25] LI J, HUO H T, LI C, et al. AttentionFGAN: Infrared and visible image fusion using attention-based generative adversarial networks[J]. *IEEE Transactions on Multimedia*, 2020, 23: 1383-1396.
- [26] LONG Y, JIA H, ZHONG Y, et al. RXDNFuse: A aggregated residual dense network for infrared and visible image fusion[J]. *Information Fusion*, 2021, 69(1): 128-141.
- [27] HE K M, SUN J, TANG X O, et al. Guided image filtering[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(6): 1397-1409.
- [28] 陈一鸣, 周登文. 基于自适应级联的注意力网络的超分辨率重建[J]. *自动化学报*, 2022, 48(8): 1950-1960.
- [29] 李晨, 侯进, 李金彪, 等. 基于注意力与残差级联的红外与可见光图像融合方法[J]. *计算机工程*, 2022, 48(7): 234-240.
- [30] 晁锐, 张科, 李言俊. 一种基于小波变换的图像融合算法[J]. *电子学报*, 2004, 32(5): 750-753.
- [31] XU H, ZHANG H, MA J. Classification saliency-

- based rule for visible and infrared image fusion[J]. IEEE Transactions on Computational Imaging, 2021, 7: 824-836.
- [32] LI H, WU X J. DenseFuse: A fusion approach to infrared and visible images[J]. IEEE Transactions on Image Processing, 2018, 28(5): 2614-2623.
- [33] MA J, ZHANG H, SHAO Z, et al. GANMcC: A generative adversarial network with multiclassification constraints for infrared and visible image fusion[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-14.
- [34] ZHANG H, WU W, ZHANG Y, et al. Semantic-supervised infrared and visible image fusion via a dual-discriminator generative adversarial network [J]. IEEE Transactions on Multimedia, 2021, 25: 635-648.
- [35] ZHANG H, XU H, XIAO Y, et al. Rethinking the

image fusion: A fast unified image fusion network based on proportional maintenance of gradient and intensity[C]. Proceedings of the AAAI Conference on Artificial Intelligence, Menlo Park: AAAI Press, 2020, 34(7): 12797-12804.

作者简介

许光宇, 博士, 副教授, 硕士生导师, 主要研究方向为数字图像处理、机器学习。

E-mail: gyxu@aust.edu.cn

陈浩宇, 硕士研究生, 主要研究方向为图像融合、深度学习。

E-mail: 858114506@qq.com

张杰, 硕士研究生, 主要研究方向为图像融合、深度学习。