

# 基于无监督深度学习的红外图像与 可见光图像融合算法

张宇苏<sup>1,2</sup>, 吴小俊<sup>1,2</sup>, 李 辉<sup>1,2</sup>, 徐天阳<sup>1,2</sup>

(1.江南大学人工智能与计算机学院, 江苏 无锡 214122)

(2.江南大学江苏省模式识别与计算智能工程实验室, 江苏 无锡 214122)

[摘要] 红外和可见光图像表征了互补的场景信息。现有的基于深度学习的融合方法大多通过独立提取网络分别提取两个源图像特征, 从而丢失了源图像之间的深度特征联系。基于此, 提出了一种新的基于无监督深度学习的红外图像与可见光图像融合算法, 针对不同模态的特点采用不同的编码方式提取图像特征, 利用一个模态的信息补充另一个模态的信息, 并对提取到的特征进行融合, 最后根据融合特征重建融合图像。该算法可在两个模态的特征提取路径之间建立交互, 不仅可预融合梯度信息和强度信息, 且能增强后续处理的信息。同时设计了损失函数, 引导模型保留可见光的细节纹理, 并保持红外的强度分布。将所提算法与多种融合算法在公开数据集上进行对比实验, 结果表明, 所提算法获得了良好的视觉效果, 客观指标评价方面对比现有的优秀算法也有一定的提升。

[关键词] 图像融合, 可见光图像, 红外图像, 无监督学习, 卷积神经网络

[中图分类号] TP391.4 [文献标志码] A [文章编号] 1672-1292(2023)01-0001-09

## Infrared Image and Visible Image Fusion Algorithm Based on Unsupervised Deep Learning

Zhang Yusu<sup>1,2</sup>, Wu Xiaojun<sup>1,2</sup>, Li Hui<sup>1,2</sup>, Xu Tianyang<sup>1,2</sup>

(1.School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi 214122, China)

(2.Jiangsu Provincial Engineering Laboratory of Pattern Recognition and Computational Intelligence, Jiangnan University, Wuxi 214122, China)

**Abstract:** Infrared and visible images represent complementary scene information. Most of the existing deep learning-based fusion methods extract the feature of the two source images through independent extraction networks, which leads to the loss of deep feature relationships between source images. To solve this problem, a new infrared and visible image fusion algorithm based on unsupervised deep learning is proposed. Specifically, the proposed algorithm adopts different encoding approaches to extract image features according to the characteristics of different modalities, and uses the information of one modality to supplement that of another one. Then, the extracted features are fused, and finally the fused image is reconstructed according to the fused features. The algorithm can establish an interaction between the feature extraction paths of the two modalities, which can not only pre-fuse gradient information and intensity information, but also enhance the information for subsequent processing. A loss function is designed to guide the model to preserve the detailed texture of visible image and retain the intensity distribution of infrared image. The proposed algorithm is compared with a variety of fusion algorithms on the public dataset. The experimental results show that the proposed algorithm has achieved good visual effects, and that the objective evaluation is also improved compared with the existing excellent algorithms.

**Key words:** image fusion, visible image, infrared image, unsupervised learning, convolutional neural network

图像融合作为一种图像增强技术, 是指从不同的源图像中提取并聚合有意义的信息, 目的是生成信息量更大且有利于后续应用的单个图像<sup>[1-2]</sup>。红外和可见光图像融合在许多方面具有优势, 是图像融合研究的一个热门领域。红外图像能捕捉热辐射信息并具有显著的对比度, 可在背景中突出目标, 但通常具有低分辨率和较差的纹理。可见光图像包含丰富的细节信息并具有较高的分辨率, 更符合人类视觉感知, 但易

收稿日期: 2022-09-15.

基金项目: 国家自然科学基金资助项目(62020106012、U1836218、62106089)、教育部 111 项目(B12018)。

通讯作者: 吴小俊, 博士, 教授, 博士生导师, 研究方向: 模式识别、计算智能、计算机视觉和信息融合。E-mail: wu\_xiaojun@jiangnan.edu.cn

受天气、照明等因素的影响. 结合两种不同模态图像的优点进行融合可获得具有高对比度和丰富纹理的融合图像. 红外和可见光图像融合在军事侦察、安防监控等许多领域发挥着重要作用<sup>[3-6]</sup>.

现有的图像融合方法可分为两大类:传统的图像融合方法和基于深度学习的图像融合方法. 传统的图像融合方法主要包含基于多尺度变换的方法<sup>[7-9]</sup>、基于稀疏表示的方法<sup>[10-11]</sup>、基于子空间的方法<sup>[12-13]</sup>和基于显著性的方法<sup>[14-15]</sup>. 然而,大多数传统方法依赖于大量的手工设计,耗时长、效率低.

深度学习强大的特征表示能力极大地推动了图像融合领域的发展<sup>[16]</sup>. Liu 等<sup>[17]</sup>首次将 CNN 引入到多聚焦图像融合任务中,相比传统算法有了显著的改进,但仍需进行有监督的训练,且红外图像和可见光图像无可参考的融合图像. Li 等<sup>[18]</sup>提出了一种具有密集块的融合架构,使用 MS\_COCO 数据集来训练网络以实现图像特征的提取与重建,并将中间层提取的特征连接到解码网络中. 上述方法均需利用预先训练好的模型提取图像特征,而非端到端的方法,无法避免人工设计融合规则,因而影响了最终融合图像的性能. 生成式对抗网络可作为无监督图像融合的有效手段. Ma 等<sup>[19]</sup>首先应用 GAN 实现了红外与可见光图像融合,利用生成器和鉴别器之间的对抗博弈迫使融合图像保留显著的红外目标和更多细节纹理信息.

虽然现有的基于深度学习的方法可在融合任务中取得更好的结果,但仍存在一些不足之处. 首先,卷积操作存在信息丢失的缺陷,许多方法忽略了中间卷积提取层获得的大量有用信息,影响了最终的融合性能. 其次,未考虑源图像固有的不同特性,而是提出相同的特征变换或表示模型来无差别地提取特征,某些情况下并不合适. 第三,大多数方法采用的源图像对的特征提取网络是相互独立的,在特征提取过程中会丢失两个源图像之间的深度特征关联.

为克服以上问题,本文提出一种基于无监督的深度学习红外图像与可见光图像融合网络. 首先,该网络是一种端到端模型,以源图像为输入,以融合图像为输出,中间无任何手工干预. 其次,对于不同模态的图像,并不采用相同的特征编码方法,而是根据各个模态的特点选择合适的特征提取方式. DenseFuse 保留可见光细节信息优于红外热辐射信息<sup>[18]</sup>. 受此启发,所提方法对可见光图像采用密集块进行特征提取,对红外图像利用特征金字塔提取多个尺度的深层显著特征. 网络在对源图像的特征提取过程中利用一个模态的特征信息来补充另一个模态的特征信息,在两个独立的特征提取路径之间建立交互,从而达到预融合的目的. 所提算法基于图像的梯度信息、强度分布和结构相似性设计损失函数. 实验结果表明,所提算法融合结果无论在主观评价还是客观指标评价方面都具有一定的竞争力.

## 1 所提算法模型

### 1.1 网络模型

所提网络模型结构如图 1 所示,主要包含编码器、融合层和解码器三部分. 网络的输入为一对配准好的可见光图像  $I_v$  和红外图像  $I_r$ ,对网络进行端到端训练以生成融合图像  $I_f$ .

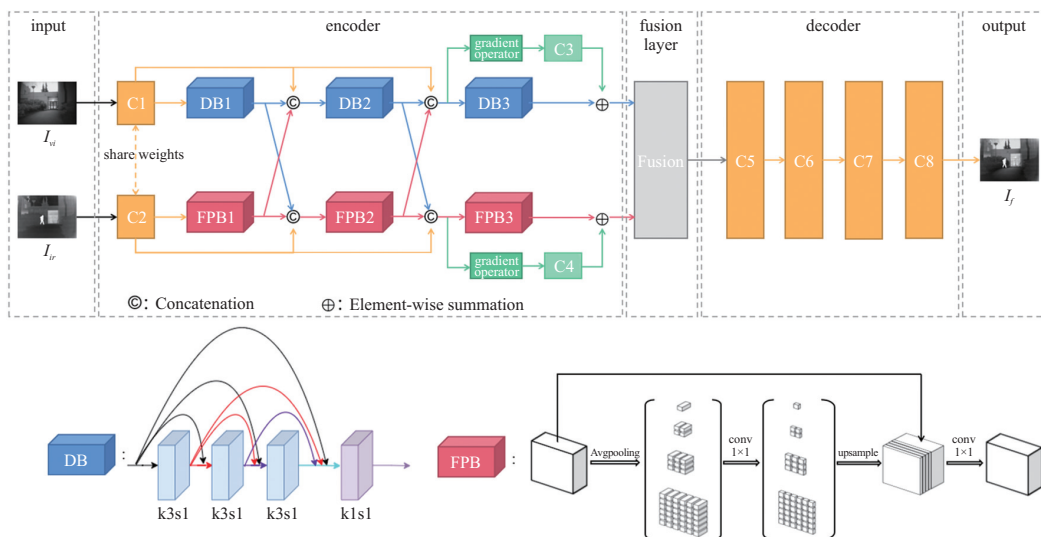


图 1 所提网络结构

Fig. 1 Structure of the proposed network

编码器的目的是提取输入源图像的特征. 由于红外和可见光图像中的信息非常不同,故对应两个特征提取路径采用不同的结构. 密集结构有利于在融合结果中保留可见光的细节信息,这在之前的研究(DenseFuse)中已有验证. 因此,将密集块(DBs)嵌入到可见光图像特征提取分支中. 在该分支中,首先采用卷积核尺寸为 $3\times 3$ 、激活函数为 ReLU 的卷积层 C1 提取可见光图像的浅层特征,而后通过级联 3 个 DB 提取可见光图像的细粒度特征. 对于红外图像,更多地关注边缘对应的特征有利于保留其显著性信息,故采用特征金字塔结构(FPBs)帮助网络提取红外图像中的多尺度特征. 类似地,先使用卷积核尺寸为 $3\times 3$ 、激活函数为 ReLU 的卷积层 C2 提取红外图像的浅层特征. 随后,级联 3 个 FPB 来提取红外图像的深层特征. 为了能使两个模态的信息保持潜在相关性,让两个特征提取路径进行信息交互,不仅能够预融合可见光梯度和红外强度信息,同时也可在下一次提取之前对信息进行补充和增强. 因此,第二个 DB/FPB 与第三个 DB/FPB 的输入不仅取决于先前所有卷积层的输出,还取决于另一个特征提取路径中卷积层的输出. 此外,利用跳跃连接将浅层特征连接至后面的卷积层以实现特征重用,减少卷积导致的信息丢失,提高特征利用率. 借鉴 Resblock<sup>[20]</sup>的思想,在每一个特征提取路径的最后一个模块,添加了梯度运算补充边缘信息. 例如在可见图像特征路径中,最后一个 DB 模块即可看作 Resblock 的变体,其主流仍为密集块,残差流采用梯度运算来计算特征的梯度幅度,并采用 $1\times 1$  卷积层 C3 来消除通道维数差异,最后通过元素加法聚合主要密集流和残差梯度流的输出.

密集块(DB)和特征金字塔模块(FPB)的结构如图 1 中所示. DB 利用三层核尺寸大小为 $3\times 3$ 的卷积提取层次特征,其中每层的输出级联为下一层的输入. FPB 的结构借鉴了 Zhao 等<sup>[21]</sup>的工作,将输入的特征图通过池化变成 4 个不同的特征图,通过 $1\times 1$  卷积对 4 个特征图进行降维,最后将经过上采样的 4 个特征图与输入特征图进行拼接. DB 和 FPB 模块最后都引入了 $1\times 1$  卷积用于降维.

融合层的目的是将编码器提取到的可见光图像和红外图像的特征进行融合. 本文使用了两种融合方式,一是直接将编码器提取到的两个模态的特征级联进行融合,但该方式对显著特征选择过于粗糙,会对融合结果产生影响. 另一种方式是基于空间的融合策略,如图 2 所示.  $\Phi_{vi}$ 和 $\Phi_{ir}$ 分别表示编码器提取的可见光图像特征和红外图像特征, $\beta_{vi}$ 和 $\beta_{ir}$ 分别表示由 $l_1$ 范数和归一化函数根据图像特征 $\Phi_{vi}$ 和 $\Phi_{ir}$ 计算得到的权重映射图. $\beta_{vi}$ 和 $\beta_{ir}$ 的计算公式为:

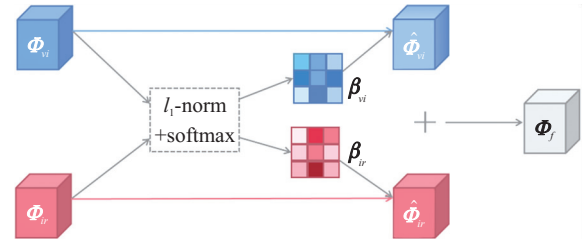


图 2 基于空间的融合策略图

Fig. 2 Diagram of space-based fusion strategy

$$\beta_{vi}(x,y) = \frac{\|\Phi_{vi}(x,y)\|_1}{\|\Phi_{ir}(x,y)\|_1 + \|\Phi_{vi}(x,y)\|_1}, \quad (1)$$

$$\beta_{ir}(x,y) = \frac{\|\Phi_{ir}(x,y)\|_1}{\|\Phi_{ir}(x,y)\|_1 + \|\Phi_{vi}(x,y)\|_1}, \quad (2)$$

式中, $(x,y)$ 表示图像特征 $\Phi_{vi}$ 、 $\Phi_{ir}$ 和权重映射图 $\beta_{vi}$ 、 $\beta_{ir}$ 中的相应位置,每个位置表示图像特征中的一个 $C$ 维向量.

$\hat{\Phi}_{vi}$ 和 $\hat{\Phi}_{ir}$ 分别表示经 $\beta_{vi}$ 和 $\beta_{ir}$ 加权的增强图像特征. 增强图像特征的计算公式为:

$$\hat{\Phi}_{vi}(x,y) = \beta_{vi}(x,y) \times \Phi_{vi}(x,y), \quad (3)$$

$$\hat{\Phi}_{ir}(x,y) = \beta_{ir}(x,y) \times \Phi_{ir}(x,y). \quad (4)$$

最后将增强的图像特征相加得到融合特征 $\Phi_f$ ,计算公式为:

$$\Phi_f = \hat{\Phi}_{vi}(x,y) + \hat{\Phi}_{ir}(x,y). \quad (5)$$

解码器的作用是对融合后的特征进行重构得到期望的融合图像. 解码器由 4 个卷积层(C5、C6、C7、C8)组成,卷积核大小均为 $3\times 3$ ,卷积步长为 1,激活函数为 ReLU.

## 1.2 损失函数

本文所提算法采用无监督的方式训练网络,对融合图像和输入源图像对之间的信息相似性进行计算. 为了减少信息损失,最大限度地保留红外图像的热辐射信息和可见光图像的纹理信息,在 $L_{total}$ 中采用



了多种损失,分别包括结构相似性损失、强度损失和梯度损失,三者的联合约束可使网络获得满意的融合结果. 网络的总体损失可表示为:

$$L_{\text{total}} = L_{\text{ssim}} + \alpha L_{\text{int}} + \beta L_{\text{grad}}, \quad (6)$$

式中,  $L_{\text{ssim}}$  表示结构相似性损失;  $L_{\text{int}}$  表示强度损失;  $L_{\text{grad}}$  表示梯度损失;  $\alpha, \beta$  是超参数,用于平衡 3 个损失的权重.

结构相似度量 SSIM<sup>[22]</sup> 根据亮度、对比度和结构信息的相似性来模拟失真,本文选用其来约束输入源图像对  $I_{vi}$ 、 $I_{ir}$  和融合图像  $I_f$  之间的结构相似性. 结构相似性损失  $L_{\text{ssim}}$  的计算公式为:

$$L_{\text{ssim}} = ((1 - \text{SSIM}(I_f, I_{ir})) + (1 - \text{SSIM}(I_f, I_{vi}))) \times 0.5. \quad (7)$$

强度损失约束融合图像保持与源图像相似的强度分布. 强度损失  $L_{\text{int}}$  的计算公式为:

$$L_{\text{int}} = \lambda_{\text{int}} L_{\text{int}}(I_f, I_{ir}) + (1 - \lambda_{\text{int}}) L_{\text{int}}(I_f, I_{vi}), \quad (8)$$

$$L_{\text{int}}(\mathbf{x}, \mathbf{y}) = \frac{1}{HW} \|\mathbf{x} - \mathbf{y}\|_2^2, \quad (9)$$

式中,  $H$  和  $W$  分别表示图像的高度和宽度;  $L_{\text{int}}(\mathbf{x}, \mathbf{y})$  表示图像  $\mathbf{x}$  和  $\mathbf{y}$  之间的像素误差;  $\lambda_{\text{int}}$  是控制两项之间权衡的一个正参数.

梯度损失迫使融合图像包含丰富的纹理细节信息. 梯度损失  $L_{\text{grad}}$  的计算公式为:

$$L_{\text{grad}} = \lambda_{\text{grad}} L_{\text{grad}}(I_f, I_{ir}) + (1 - \lambda_{\text{grad}}) L_{\text{grad}}(I_f, I_{vi}), \quad (10)$$

$$L_{\text{grad}}(\mathbf{x}, \mathbf{y}) = \frac{1}{HW} \|\nabla \mathbf{x} - \nabla \mathbf{y}\|_2^2, \quad (11)$$

式中,  $H$  和  $W$  分别表示图像的高度和宽度;  $\nabla$  表示 Sobel 梯度算子;  $L_{\text{grad}}(\mathbf{x}, \mathbf{y})$  表示图像  $\mathbf{x}$  和  $\mathbf{y}$  之间的边缘信息损失;  $\lambda_{\text{grad}}$  用于调整融合图像中不同模态梯度信息的比例.

### 1.3 实验执行细节及环境

本文从 TNO<sup>[23]</sup> 数据集中选择了 45 对不同场景的红外图像和可见光图像作为训练数据. 45 对训练数据并不足以训练一个好的模型. 因此,通过裁剪原始图像进行数据增强处理,设置步幅为 14,每张子图像的大小为 120×120,同时,将裁剪后的图像对称展开. 最后一共得到 44 385 对图像作为训练数据集.

本文所提方法的代码是使用带有 NVIDIA 2080Ti GPU 的 PyTorch 框架实现的. 网络模型使用 Adam ( $\beta_1=0.9, \beta_2=0.999$ ) 优化器进行优化,初始的学习率为  $1 \times 10^{-3}$ ,权重衰减率为  $1 \times 10^{-4}$ . 网络一共训练 50 个周期,批量池大小为 8. 此外,损失函数中的超参数分别设置为  $\alpha=1, \beta=10$ .

## 2 实验及结果分析

从文献[23]中随机选出 20 对红外图像和可见光图像作为测试集. 该测试集中的图像均为配准好的红外图像和可见光图像,其中选取的许多图像也被现有许多红外图像和可见光图像融合算法用作测试数据,具有一定代表性. 本文首先进行消融实验来探究所提算法的合理性,然后将所提算法融合结果与现有的一些典型和最新的算法的融合结果进行对比,包括基于梯度转移和总变差最小化 (GTF) 的融合算法<sup>[24]</sup>、基于多尺度分解 (MSD) 的融合算法<sup>[25]</sup>、基于密集连接的自编码网络 (DenseFuse) 融合算法<sup>[18]</sup>、基于生成对抗网络 (FusionGAN) 的融合算法<sup>[19]</sup>、基于压缩分解网络 (SDNet) 的实时融合算法<sup>[26]</sup>、基于梯度和强度比例保持 (PMGI) 的融合算法<sup>[27]</sup>、基于模型的卷积神经网络 (AUIF) 融合算法<sup>[28]</sup>、基于端到端残差融合网络 (RFN-Nest) 的融合算法<sup>[29]</sup>、基于光照感知网络 (PIAFusion) 的渐进式融合算法<sup>[30]</sup> 和基于通用融合 (U2Fusion) 的融合算法<sup>[31]</sup>.

本文选择从主观评价和客观评价两方面来评价图像质量,引入 6 种常用的图像客观评价指标来衡量融合图像的质量,分别为标准差 SD<sup>[32]</sup>、峰值信噪比 PSNR<sup>[33]</sup>、均方误差 MSE<sup>[34]</sup>、相关系数 CC<sup>[35]</sup>、差异相关总和 SCD<sup>[36]</sup> 及信息熵 EN<sup>[37]</sup>.

### 2.1 消融实验

对于红外和可见光图像融合,期望在融合结果中对可见光图像的梯度信息和红外图像的强度信息进行主要保留,对红外图像的梯度信息和可见光图像的强度信息进行次要保留. 为达到上述期望,设置  $\lambda_{\text{int}}=0.7, \lambda_{\text{grad}}=0.3$ . 为验证上述参数的合理性,设置了实验进行论证,将强度和梯度信息的权重参数均设置为 0.5 来进

行比较. 从表 1 可见,损失函数参数的设置是合理的,能够有效保留更多可见光图像的纹理信息和红外图像的热辐射信息.

表 1 损失函数中采用不同参数的实验结果

Table 1 Experimental results using different weights in the loss function

Experiment	SD $\uparrow$	PSNR $\uparrow$	MSE $\downarrow$	CC $\uparrow$	SCD $\uparrow$	EN $\uparrow$
$\lambda_{\text{int}} = 0.5, \lambda_{\text{grad}} = 0.5$	9.280 8	62.763 0	0.037 8	0.477 5	1.681 3	6.758 0
$\lambda_{\text{int}} = 0.7, \lambda_{\text{grad}} = 0.3$	9.373 2	63.393 6	0.021 7	0.524 4	1.682 2	6.767 3

本文共设置了 4 个消融实验来验证所提网络模型的合理性. 首先,在编码器网络中,利用不同模态之间的信息交互进行预融合,同时利用跳跃连接实现特征复用. 为了验证其有效性,设置了实验 1(图 3(a)),该实验基于双独立特征提取网络对源图像进行特征提取. 其次,编码器网络中的另一个关键组件是在最后一个特征提取模块中加入残差梯度提取,增强网络对特征的描述能力. 为验证其有效性,设置了实验 2(图 3(b)),实验 2 在实验 1 的基础上加入了残差梯度流. 然后,设置实验 3(图 3(c))对进行信息交互的特征提取模块个数进行合理性验证,实验 3 采用 1 个模块进行信息交互. 最后,实验 4(图 3(d))分别在两个特征提取路径使用 2 个信息交互的特征提取模块,即本文所提方法. 上述所有实验均采用基于空间的融合策略.

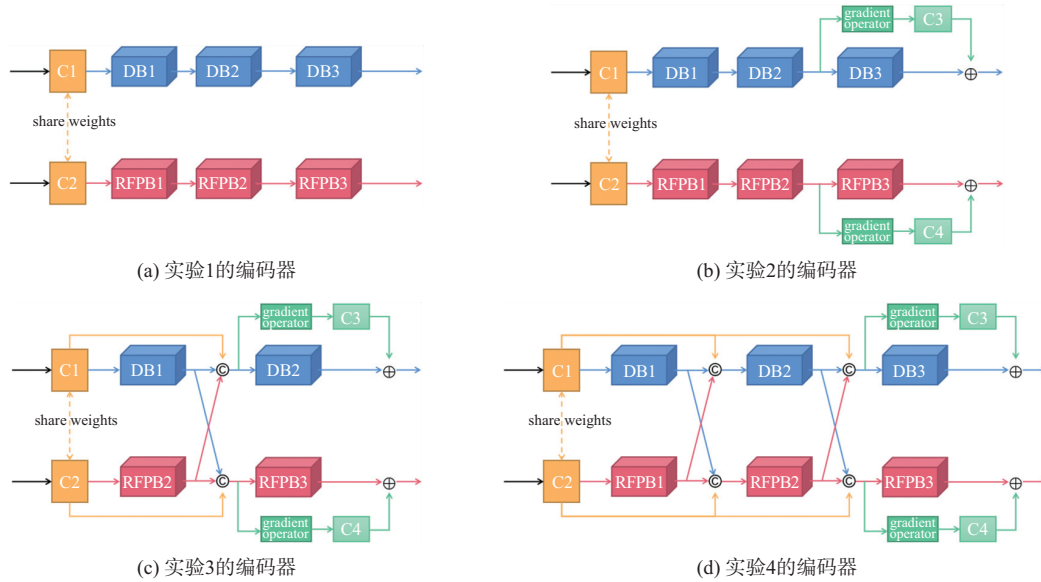


图 3 消融研究

Fig. 3 Ablation study

主观评价方面,选取了 1 对具有代表性的图像展示上述 4 个实验的结果,如图 4 所示,图中用红框标注了差异区域. 从图 4 可以看出,实验 2 相比实验 1 保留了更多可见光图像的信息. 实验 3 相比实验 2,融合结果更为清晰. 相比于前 3 个实验,本文所提算法保留的纹理信息最为丰富,清晰度也最高.

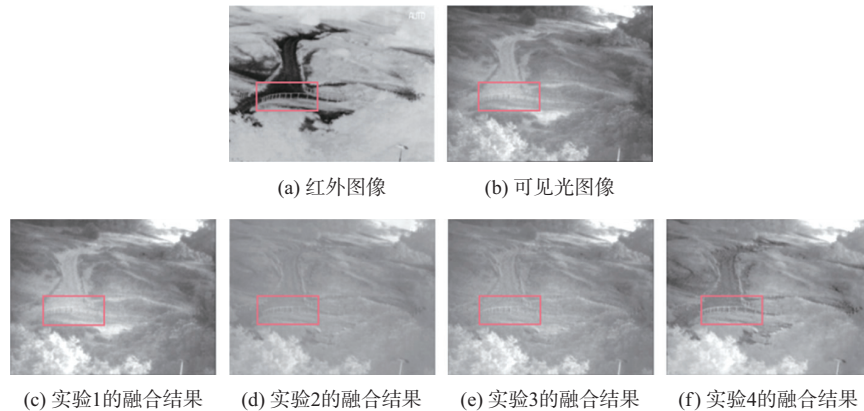


图 4 消融实验结果

Fig. 4 Ablation experiment results

客观评价方面,表 2 是实验 1、2、3 与本文所提方法 20 张融合图像的 6 个客观指标的均值. 从表 2 可以看出,实验 2 相比于实验 1 在峰值信噪比、均方误差、相关系数和信息熵 4 个指标上都有所提升,本文所提方法与实验 3 相比,在 6 个指标上都有显著提升. 此外,本文所提方法在所有指标上都取得了最优的结果.

表 2 不同实验中 20 张融合图像评价指标均值

Table 2 Average quality metrics of 20 fused images in different experiments

Experiment	SD $\uparrow$	PSNR $\uparrow$	MSE $\downarrow$	CC $\uparrow$	SCD $\uparrow$	EN $\uparrow$
Experiment1	8.952 0	62.271 1	0.047 5	0.425 2	1.637 3	6.000 3
Experiment2	7.793 1	62.493 5	0.042 8	0.515 2	1.353 2	6.057 0
Experiment3	9.059 2	62.733 7	0.039 5	0.498 2	1.653 2	6.722 0
Experiment4	9.373 2	63.393 6	0.021 7	0.524 4	1.682 2	6.767 3

## 2.2 对比实验

为了全面评估所提算法的融合性能,本文将所提算法与现有 10 种融合算法进行了对比.

图 5 和图 6 分别是不同融合算法在两组图像上的可视化结果. 为方便比较分析,图中用红色矩形框标记了亮度特征,用绿色矩形框标记了纹理特征. 图 5 是不同的融合算法在“soldier\_behind\_smoke”图像上的融合结果. 从图 5 可以看出,融合算法 GTF、FusionGAN 能够突出显著的红外热辐射信息,但不能保留细节纹理信息,清晰度较差,例如几乎未保留可见光图像中的烟雾信息(绿色矩形框),背景树木也较暗. PIAFusion、AUIF、RFN-Nest、U2Fusion 的融合结果中无法突出红外热辐射信息,例如图像中的士兵(红色矩形框). DenseFuse、MSD、SDNet、PMGI 对细节刻画不足. 本文所提算法的融合结果相对较好.

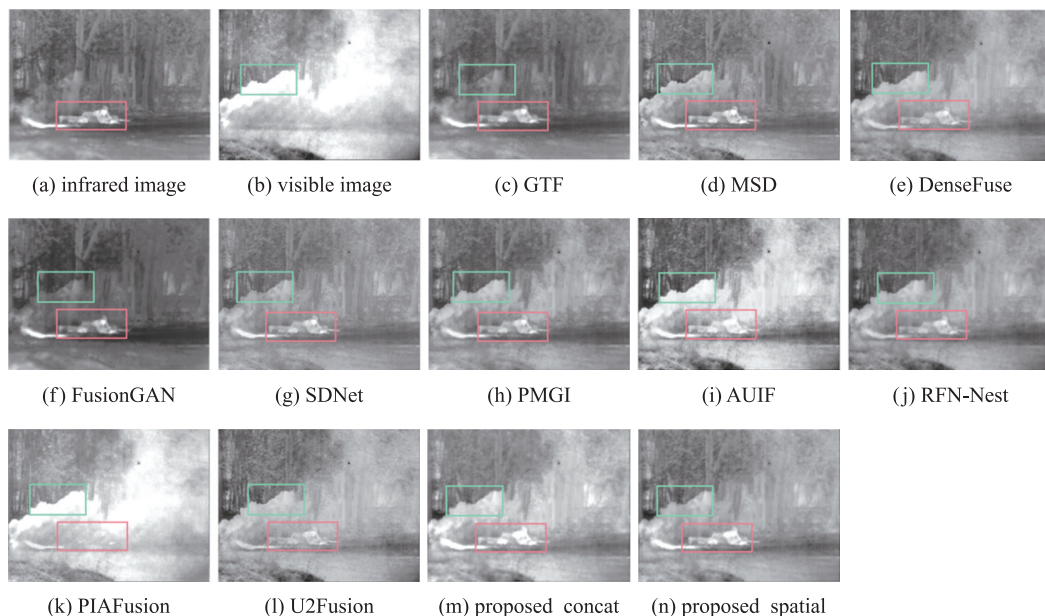


图 5 不同算法在“soldier\_behind\_smoke”图像上的融合结果对比

Fig. 5 Comparison of fusion results of different algorithms on “soldier\_behind\_smoke” images

图 6 是不同的融合算法在“Kaptein\_1123”图像上的融合结果. 如图 6 所示,GTF、FusionGAN、SDNet 丢失了大量纹理细节信息,热辐射信息比较显著. DenseFuse、MSD、PMGI、AUIF 对纹理刻画不足,例如树的轮廓不够清晰(绿色矩形框). RFN-Nest、U2Fusion 相较其他几种融合算法而言亮度偏低,热辐射信息不够突出. 本文所提算法基于空间的策略相比直接级联融合更多细节信息,也保留了一定的红外热辐射信息.

为了进一步验证所提算法的泛化能力,从测试集中随机选取 5 对图像,用不同算法进行融合,融合结果如图 7 所示. 从图 7 可以看出,所提融合算法能够对源图像中的纹理细节信息和热辐射信息有较好的保留,视觉效果良好.

表 3 是不同的融合算法在 TNO 数据集上的客观评价指标. 从表 3 可以看出,级联策略的结果在 SCD 中获得最佳值,在 PSNR、CC、EN 中获得次优值,这表明融合图像与源图像的相关性较好,结果包含来自源



图像的大量信息. 基于空间策略的结果在 PSNR 和 MSE 中获得最佳值,这意味着图像失真较少,与源图像差异不大. 在 SD 中具有次优值,这意味着融合结果对比度较高,视觉感知较好.

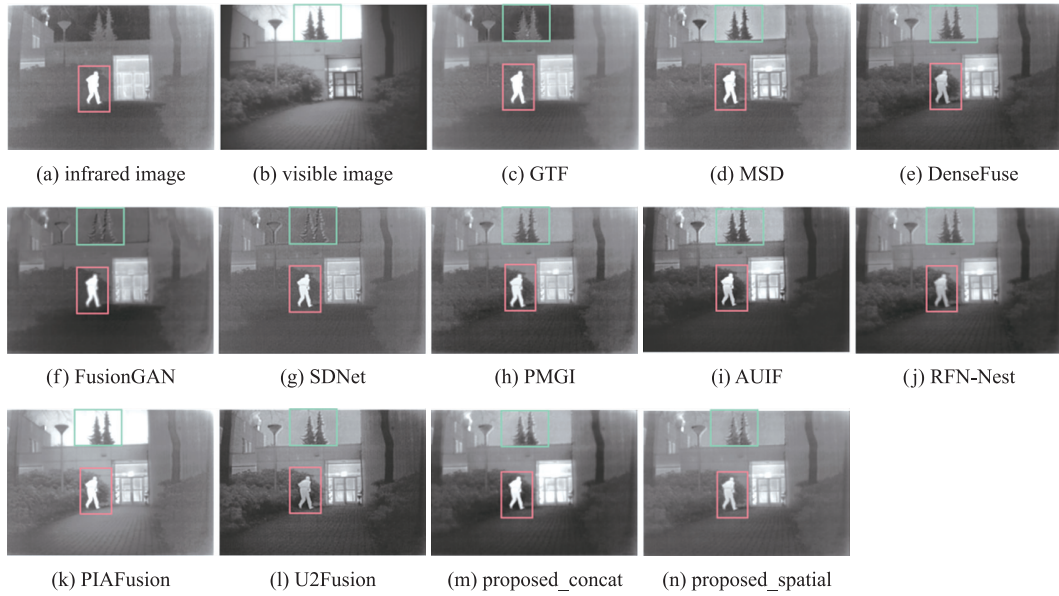


图 6 不同算法在“Kaptein\_1123”图像上的融合结果对比

Fig. 6 Comparison of fusion results of different algorithms on “Kaptein\_1123” images

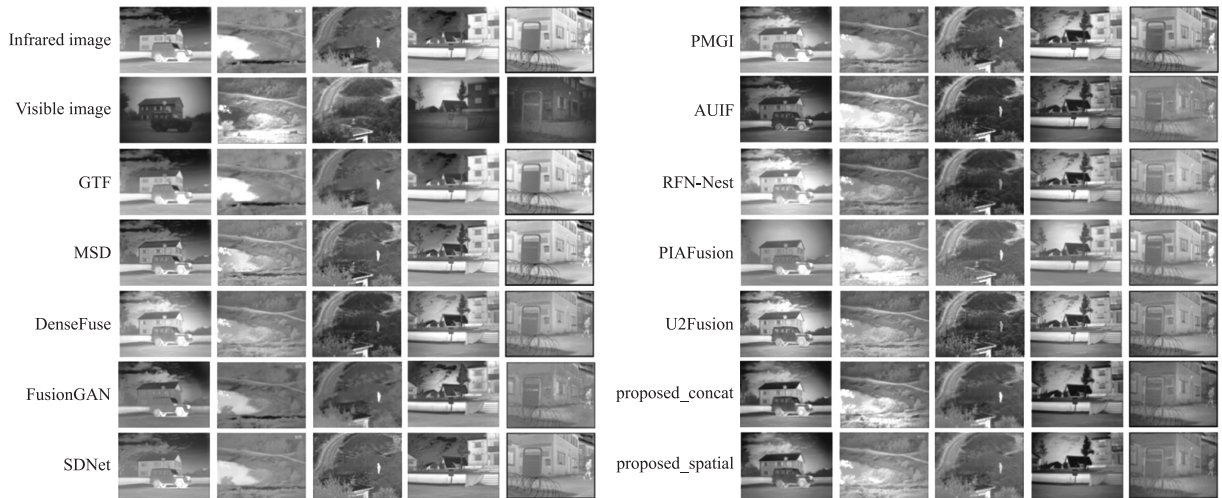


图 7 不同融合算法在 5 对图像上的实验结果

Fig. 7 Experiments results of different fusion algorithms on 5 pairs of images

表 3 20 张融合图像的平均指标

Table 3 Average metrics of 20 fused images

Algorithms	SD $\uparrow$	PSNR $\uparrow$	MSE $\downarrow$	CC $\uparrow$	SCD $\uparrow$	EN $\uparrow$
GTF	9.320 7	62.153 2	0.053 1	0.355 6	0.969 6	6.676 8
MSD	9.321 3	62.548 3	0.039 6	0.470 0	1.594 0	6.843 3
DenseFuse	9.133 9	63.234 9	0.037 7	0.543 0	1.764 6	6.788 9
FusionGAN	8.437 3	61.338 6	0.054 8	0.444 7	1.352 7	6.498 4
SDNet	8.917 5	62.771 2	0.041 9	0.474 3	1.548 5	6.638 3
PMGI	9.265 0	62.683 2	0.041 3	0.524 8	1.696 5	6.795 0
AUIF	9.005 6	62.595 8	0.040 1	0.521 4	1.617 8	6.670 4
RFN-Nest	9.281 6	62.683 1	0.042 7	0.534 9	1.771 6	6.829 8
PIAFusion	9.008 8	62.638 7	0.045 2	0.456 0	1.563 7	6.839 9
U2Fusion	9.462 1	63.362 3	0.035 0	0.522 9	1.781 5	7.038 6
Proposed_concat	9.276 8	63.366 1	0.037 7	0.537 0	1.796 7	6.869 5
Proposed_spatial	9.373 2	63.393 6	0.021 7	0.524 4	1.682 2	6.767 3

### 3 结论

本文提出了一种基于无监督深度学习的红外图像与可见光图像融合算法. 该算法由编码器、融合层、解码器 3 部分组成. 编码器通过密集块挖掘可见光图像丰富的纹理信息, 引入特征金字塔提取红外图像的多尺度特征, 在两个特征提取路径间建立信息交互, 采用跳跃连接减少卷积信息丢失. 融合层采用级联或基于空间的策略将源图像对的特征集成到一个特征图中. 最后通过解码器重构特征得到融合图像. 实验结果表明, 本文所提算法能够对源图像的信息进行有效地提取和选择, 融合结果与现有算法相比均有一定的提升.

#### [参考文献](References)

- [1] LI S T, KANG X D, FANG L Y, et al. Pixel-level image fusion: a survey of the state of the art[J]. *Information Fusion*, 2017, 33: 100-112.
- [2] 汪前进, 朱斌, 李存华. 基于特征点的图像拼接方法的研究与应用[J]. *南京师范大学学报(工程技术版)*, 2016, 16(3): 48-53.
- [3] ZHANG H, XU H, TIAN X, et al. Image fusion meets deep learning: a survey and perspective[J]. *Information Fusion*, 2021, 76: 323-336.
- [4] 唐聪, 凌永顺, 杨华, 等. 基于深度学习的红外与可见光决策级融合跟踪[J]. *激光与光电子学进展*, 2019, 56(7): 209-216.
- [5] MA J Y, MA Y, LI C. Infrared and visible image fusion methods and applications: a survey[J]. *Information Fusion*, 2019, 45: 153-178.
- [6] 游梓童, 吴福明, 赵森, 等. 融合高阶信息增强模块的复杂背景植物叶片图像分类[J]. *南京师范大学学报(工程技术版)*, 2022, 22(3): 45-52.
- [7] 李莹, 朱文艳, 袁飞, 等. 基于形态学和平均梯度的小波图像融合算法[J]. *南京师范大学学报(工程技术版)*, 2013, 13(4): 76-81.
- [8] CHEN J, LI X J, LUO L B, et al. Infrared and visible image fusion based on target-enhanced multiscale transform decomposition[J]. *Information Sciences*, 2020, 508: 64-78.
- [9] LI H, WU X J, KITTLER J. MDLatLRR: a novel decomposition method for infrared and visible image fusion[J]. *IEEE Transactions on Image Processing*, 2020, 29: 4733-4746.
- [10] LIU Y, CHEN X, WARD R K, et al. Image fusion with convolutional sparse representation[J]. *IEEE Signal Processing Letters*, 2016, 23(12): 1882-1886.
- [11] ZHANG Q, LIU Y, BLUM R S, et al. Sparse representation based multi-sensor image fusion for multi-focus and multi-modality images: a review[J]. *Information Fusion*, 2018, 40: 57-75.
- [12] LI H, LIU L, HUANG W, et al. An improved fusion algorithm for infrared and visible images based on multi-scale transform[J]. *Infrared Physics & Technology*, 2016, 74: 28-37.
- [13] BAVIRISETTI D P, XIAO G, LIU G. Multi-sensor image fusion based on fourth order partial differential equations[C]// *Proceedings of the 20th International Conference on Information Fusion(Fusion)*. Xi'an, China: IEEE, 2017.
- [14] NAIDU A R, BHAVANA D, REVANTH P, et al. Fusion of visible and infrared images via saliency detection using two-scale image decomposition[J]. *International Journal of Speech Technology*, 2020, 23(4): 815-824.
- [15] MA J L, ZHOU Z Q, WANG B, et al. Infrared and visible image fusion based on visual saliency map and weighted least square optimization[J]. *Infrared Physics & Technology*, 2017, 82: 8-17.
- [16] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]// *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*. Honolulu, USA: IEEE, 2017.
- [17] LIU Y, CHEN X, PENG H, et al. Multi-focus image fusion with a deep convolutional neural network[J]. *Information Fusion*, 2017, 36: 191-207.
- [18] 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.
- [19] MA J Y, YU W, LIANG P W, et al. FusionGAN: a generative adversarial network for infrared and visible image Fusion[J].



- Information Fusion,2019,48:11–26.
- [20] HE K M,ZHANG X Y,REN S Q,et al. Deep residual learning for image recognition[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition( CVPR). Las Vegas,USA:IEEE,2016.
- [21] ZHAO H S,SHI J P,QI X J,et al. Pyramid scene parsing network[ C]//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition( CVPR). Honolulu,USA:IEEE,2017.
- [22] WANG Z,BOVIK A C,SHEIKH H R,et al. Image quality assessment:from error visibility to structural similarity[J]. IEEE Transactions on Image Processing,2004,13(4):600–612.
- [23] TOET A. TNO image fusion dataset,2014[ EB/OL]. [ 2021–02–20]. [https://figshare.com/articles/TNO\\_Image\\_Fusion\\_Dataset/1008029](https://figshare.com/articles/TNO_Image_Fusion_Dataset/1008029).
- [24] MA J Y,CHEN C,LI C,et al. Infrared and visible image fusion via gradient transfer and total variation minimization[J]. Information Fusion,2016,31:100–109.
- [25] ZHOU Z Q,WANG B,LI S,et al. Perceptual fusion of infrared and visible images through a hybrid multi-scale decomposition with Gaussian and bilateral filters[J]. Information Fusion,2016,30:15–26.
- [26] ZHANG H,MA J Y. SDNet: a versatile squeeze-and-decomposition network for real-time image fusion[ J]. International Journal of Computer Vision,2021,129(10):2761–2785.
- [27] 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[J]. Proceedings of the AAAI Conference on Artificial Intelligence,2020,34(7):12797–12804.
- [28] ZHAO Z X,XU S,ZHANG J S,et al. Efficient and model-based infrared and visible image fusion via algorithm unrolling[J]. IEEE Transactions on Circuits and Systems for Video Technology,2021,32(3):1186–1196.
- [29] LI H,WU X J,KITTLER J. RFN-Nest: an end-to-end residual fusion network for infrared and visible images[J]. Information Fusion,2022,73:72–86.
- [30] TANG L F,YUAN J T,ZHANG H,et al. PIAFusion: a progressive infrared and visible image fusion network based on illumination aware[J]. Information Fusion,2022,83/84:79–92.
- [31] XU H,MA J Y,JIANG J J,et al. U2Fusion: a unified unsupervised image fusion network[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2022,44(1):502–518.
- [32] RAO Y J. In-fibre Bragg grating sensors[J]. Measurement Science and Technology,1997,8(4):355–375.
- [33] HUYNH-THU Q,GHANBARI M. Scope of validity of PSNR in image/video quality assessment[ J]. Electronics Letters,2008,44(13):800–801.
- [34] SARA U,AKTER M,UDDIN M S. Image quality assessment through FSIM,SSIM,MSE and PSNR—a comparative study[J]. Journal of Computer and Communications,2019,7(3):8–18.
- [35] LARSON E C,CHANDLER D M. Most apparent distortion:full-reference image quality assessment and the role of strategy[J]. Journal of Electronic Imaging,2010,19(1):011006.
- [36] ASLANTAS V,BENDES E. A new image quality metric for image fusion: the sum of the correlations of differences[ J]. AEU—International Journal of Electronics and Communications,2015,69(12):1890–1896.
- [37] ROBERTS J W,VAN AARDT J A,AHMED F B. Assessment of image fusion procedures using entropy, image quality, and multispectral classification[J]. Journal of Applied Remote Sensing,2008,2(1):023522.

[ 责任编辑:严海琳]