



计算机科学

COMPUTER SCIENCE

基于无监督显著性掩码引导的红外与可见光图像融合网络

李东阳, 聂仁灿, 潘琳娜, 李贺

引用本文

李东阳, 聂仁灿, 潘琳娜, 李贺. [基于无监督显著性掩码引导的红外与可见光图像融合网络](#)[J]. 计算机科学, 2024, 51(6A): 230600170-5.

LI Dongyang, NIE Rencan, PAN Linna, LI He. [UMGN:An Infrared and Visible Image Fusion Network Based on Unsupervised Significance MaskGuidance](#) [J]. Computer Science, 2024, 51(6A): 230600170-5.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[一种单阶段无监督可见光-红外跨模态行人重识别方法](#)

Single Stage Unsupervised Visible-infrared Person Re-identification

计算机科学, 2024, 51(6A): 230600138-7. <https://doi.org/10.11896/jsjcx.230600138>

[基于子空间的I-nice聚类算法](#)

Subspace-based I-nice Clustering Algorithm

计算机科学, 2024, 51(6): 153-160. <https://doi.org/10.11896/jsjcx.230800200>

[无监督单目深度估计研究综述](#)

Unsupervised Learning of Monocular Depth Estimation:A Survey

计算机科学, 2024, 51(2): 117-134. <https://doi.org/10.11896/jsjcx.230400197>

[基于配置语句树的网络设备配置异常检测算法](#)

Anomaly Detection Algorithm for Network Device Configuration Based on Configuration Statement Tree

计算机科学, 2023, 50(11A): 230200128-10. <https://doi.org/10.11896/jsjcx.230200128>

[双分支注意力网络的遥感图像融合](#)

Remote Sensing Image Fusion with Dual-branch Attention Network

计算机科学, 2023, 50(11A): 230200072-7. <https://doi.org/10.11896/jsjcx.230200072>

基于无监督显著性掩码引导的红外与可见光图像融合网络

李东阳 聂仁灿 潘琳娜 李贺

云南大学信息学院 昆明 650091

(dongyang_li@mail.ynu.edu.cn)

摘要 在具有挑战性的拍摄环境中,使用单张红外或可见光图像很难捕获清晰详细的纹理信息以及热辐射信息。然而,红外和可见光图像融合允许保存来自红外图像的热辐射信息和来自可见光图像的纹理细节。现有的许多方法在融合过程中直接生成融合图像,忽略了对源图像像素级权重贡献的估计,强调了不同源图像之间的学习。为此,提出了基于无监督显著性掩码引导的红外与可见光图像融合网络,利用密集结构在源图像中进行全面的特征提取。它产生一个权重估计概率来评估每个源图像对融合图像的贡献。此外,由于红外与可见光图像缺乏真实标签,难以使用有监督学习,UMGN 还引入了显著性掩码,便于网络集中学习红外图像的热辐射信息和可见光纹理信息。在训练过程中还引入了加权保真度项和梯度损失,以防止梯度退化。与大量其他最先进的方法进行对比实验,结果证明了所提出的 UMGN 方法的优越性和有效性。

关键词 无监督学习;显著性掩码;权重估计概率;红外与可见光图像融合

中图分类号 TP391

UMGN: An Infrared and Visible Image Fusion Network Based on Unsupervised Significance Mask Guidance

LI Dongyang, NIE Rencan, PAN Linna and LI He

School of Information Science and Engineering, Yunnan University, Kunming 650091, China

Abstract In challenging shooting environments, it is difficult to capture clear and detailed texture information and thermal radiation information using a single infrared or visible image. However, infrared and visible image fusion allows the preservation of thermal radiation information in infrared images and texture details in visible light images. Many existing methods directly generate fused images in the fusion process, ignore the estimation of pixel-level weight contribution of source images, and emphasize the learning between different source images. For this reason, an infrared and visible image fusion based on unsupervised significance mask guidance network is proposed, which uses DenseNet structure to extract comprehensive features from source images. It produces a weight estimation probability to evaluate the contribution of each source image to the fused image. Since infrared and visible images lack ground truth, it is difficult to use supervised learning. UMGN also introduces the significance mask to facilitate the network to focus on learning the thermal radiation information and visible light texture information of infrared images. A weighted fidelity term and gradient loss are also introduced in the training process to prevent gradient degradation. A large number of comparative experiments with other advanced methods prove the superiority and effectiveness of the proposed UMGN method.

Keywords Unsupervised learning, Significance mask, Weight estimation probability, Infrared and visible image fusion

1 引言

目前,图像融合在计算机视觉和图像处理领域越来越受关注。红外与可见光图像融合(Infrared and Visible Image Fusion, IVIF)旨在通过将红外传感器捕获的热辐射信息与可见光图像中的纹理细节相结合,生成一张新的图像,以尽可能保留源图像的信息。这种融合方法能够产生更清晰和显著的目标,提升视觉感知能力,为后续任务提供帮助^[1]。然而,红外图像与可见光图像捕获目标的重点不同,图 1 给出了这种

差异,红框内放大的灯在红外图像中更多地呈现像素亮度信息,而在可见光图像中更多地呈现纹理信息。因此,IVIF 的目标是提高融合图像的质量,以获得两者的优点,如图 1 所示。

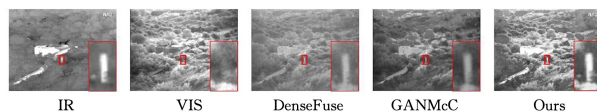


图 1 融合结果示例图

Fig. 1 Examples of fused results

传统的 IVIF 方法可以大致分为几类,分别是多尺度

基金项目:国家自然科学基金(61966037);云南省基础研究计划重点项目(202301AS070025, 202401AT070467);国家重点研发项目(2020YFA0714301);云南省科技厅项目基金(2012105AF150011)

This work was supported by the National Natural Science Foundation of China(61966037), Key Project of Yunnan Basic Research Program(202301AS070025, 202401AT070467), National Key Research and Development Program of China(2020YFA0714301) and Science and Technology Department of Yunnan Province Project Foundation(202105AF150011).

通信作者:聂仁灿(renie@ynu.edu.cn)

变换^[2]、稀疏表示^[3]和子空间学习^[4]。一般来说,这些融合方法都是将源图像分解为多层特征进行表示,随后通过某一融合规则将这些多层特征进行融合,最终解码获得融合图像。然而,这些人为设计的特征,在表示源图像时往往受到限制,并且对不同的输入十分敏感,即使是同一种方法,参数的调整和输入数据对变化都可能导致输出结果的差异。

近年来,深度学习广泛用于 IVIF,例如 Densefuse^[5], IPLF^[6], GANMcC^[7], CUFD^[8], SSL-WAEIE^[9] 和 PIAFusion^[10]等。与前文所提现存方法直接将图片进行编码解码进行直接融合不同的是,本文方法在像素级别对每个像素进行权重评估。此外,本文不像大多数方法直接通过红外图像生成掩码,而是使用显著性加强算法,增加对红外图像中热辐射信息的敏感性和可见光图像纹理细节的表示。最后,本文提出了一种加权损失函数,更关注像素级别的细节变化,减少图像的亮度退化。

2 相关工作

2.1 基于手工设计特征的图像融合

大部分基于传统方法的图像融合研究都是手工设计的特征,例如 LatIrr^[11]将源图像分解为低秩部分和显著部分后进行融合。源图像的低秩部分使用加权平均方法进行融合,而显著部分则直接进行逐像素融合。Zhou 等^[2]通过使用多尺度高斯

和双边滤波器,进一步将图像分解为多尺度纹理细节和边缘特征。类似地,GTF^[12]被提出保持梯度的变化来实现图像融合。

2.2 基于深度学习的图像融合

基于深度学习的图像融合方法可以分为两类:基于有监督的深度学习与基于无监督的深度学习。

正如前文提到的,由于 IVIF 缺乏标签,仅有少部分方法使用了有监督学习。例如 IPLF^[6],它通过生成的图像作为标签数据来指导融合模型的训练;因此,大部分的 IVIF 都是基于无监督学习的;Gao 等^[13]将源图像分为共有特征以及私有特征进行融合,CUFD^[8]则在此基础上,通过设计专门的网络结构更加细致地划分共同部分以及私有部分,但是两者都是直接生成的融合图像,忽略了源图像像素层面的贡献和对不同源图像的侧重;SOSMaskFusion^[14]虽然通过对红外图像目标进行目标检测来获取掩码,帮助网络学习红外图像的热辐射信息,但是,由红外图像直接生成的掩码,对可见光图像纹理细节的保留还不够完善。

3 网络介绍

3.1 网络框架

本文网络结构如图 2 所示,将源图像送入编码器。共享权重的卷积网络充分提取特征并生成源图像的权重估计策略 P ,通过将权重 P 与源图像进行点乘来获得融合图像。

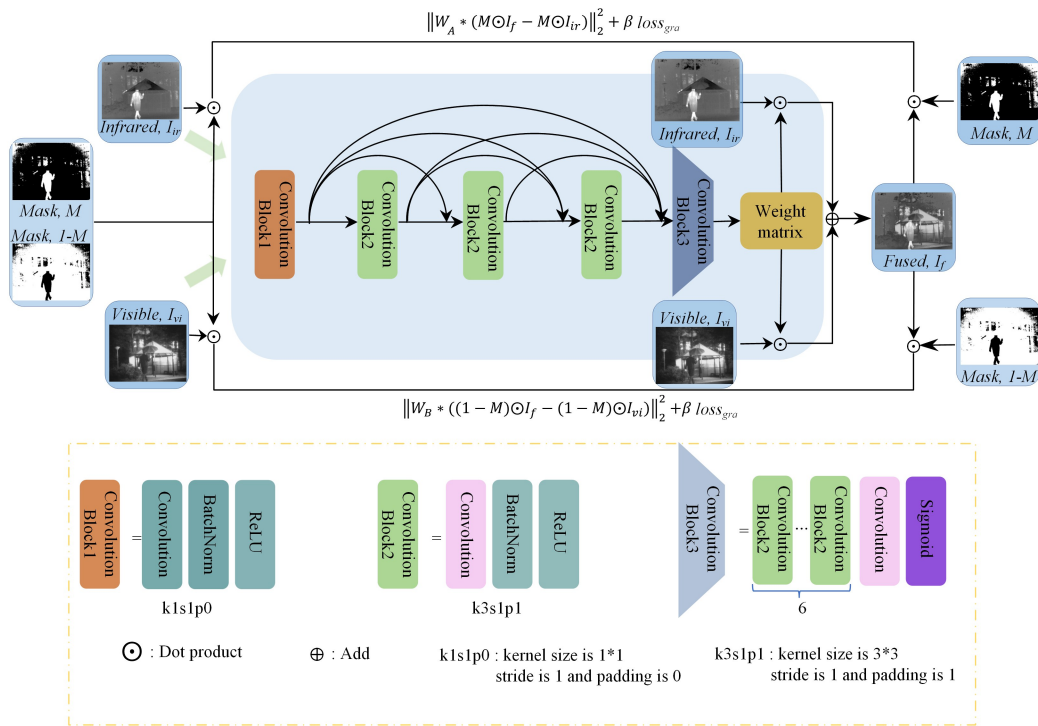


图 2 UMGN 网络框架

Fig. 2 Framework of UMGN

3.2 权重策略估计

不同于大部分方法直接生成融合图像,本文先生成权重概率分布 P ,计算红外与可见光每个像素的贡献,同时融合图像 I_f 通过式(1)生成:

$$I_f = I_{ir} \odot P + I_{vi} \odot (1 - P) \quad (1)$$

其中, \odot 表示点乘。UMGN 能够结合源图像的内容生成有丰富纹理细节的融合图像,另外式(1)也保证了融合图像在像素层面与源图像之间的连续性。

权重矩阵 P 的具体生成方式如图 2 所示,为了从源图像

获取丰富的细节信息,首先将源图像分别送入“Conv+BN+ReLU”的卷积块进行预卷积,其中卷积核、步长和填充分别设为 1,1 和 0,通道数为 16;为了防止信息在卷积的过程中丢失,我们在编码器中引入 DenseNet^[5] 结构,其中各卷积块结构一致,卷积核、步长和填充分别为 3,1 和 1,通道数则分别设为 32,64 和 256,最后将特征图送入由 7 个卷积块组成的解码器,特别是,在最后一个卷积块,进行卷积操作和 Sigmoid 激活函数,将输出归一化到 0~1 之间,获得了源图像的融合权重贡献估计 P 。

3.3 显著性掩码的生成

为了解决红外图像的热辐射信息不突出和可见光图像纹理不清晰的问题,本文采用文献[15]中的方法进行显著性增强。图3给出了增强前后的红外图像(见图3(a)、图3(b))和可见光图像(见图3(c)、图3(d))。可以观察到,经过显著性增强后的图像更加突出热辐射信息和背景纹理信息。但由于不同红外图像具有不同的重要信息,为了捕获有价值的红外图像信息,本文使用像素对比度作为衡量方法。掩码中每个元素的值的计算式如式(2)所示:

$$M_{ij} = \begin{cases} 0, & I_{rij} - \bar{I}_r > \lambda \\ 1, & \text{otherwise} \end{cases} \quad (2)$$

其中, M_{ij} 和 I_{rij} 代表掩码 M 和显著性图在第 i 行和第 j 列的元素; \bar{I}_r 是显著性图所有元素的平均值; λ 则是阈值,在UMGN中本文将 λ 设为10。

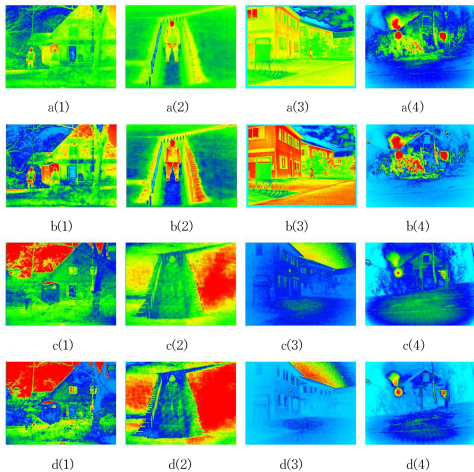


图3 伪彩处理后的显著性图与原图对比

Fig. 3 Comparison of significance graph after false color processing with the original graph

3.4 损失函数

在UMGN中,使用的损失函数由两部分组成,分别是加权保真度损失和梯度损失,如式(3)所示:

$$Loss = l_{w_f} + \beta l_{gra} \quad (3)$$

其中, l_{w_f} 为加权保真度损失, l_{gra} 为融合图像与源图像之间的梯度损失, β 为权重系数,本文将其设置为15;梯度大小可以衡量融合图像与源图像之间的梯度相似性。

3.3.1 加权保真度损失

加权保真度损失如式(4)所示:

$$l_{w_f} = \|W_A * (M \odot I_f - M \odot I_{ir})\|_2^2 + \|W_B * ((1-M) \odot I_f - (1-M) \odot I_{vi})\|_2^2 \quad (4)$$

其中, M 为掩码, W_A 和 W_B 分别是源图像与显著性图像之间的权重,两者和为1,具体定义为:

$$W_A = \frac{S_{ir}^\alpha \otimes I_{ir}^\gamma}{S_{ir}^\alpha \otimes I_{ir}^\gamma + S_{vi}^\alpha \otimes I_{vi}^\gamma} \quad (5)$$

$$W_B = \frac{S_{vi}^\alpha \otimes I_{vi}^\gamma}{S_{ir}^\alpha \otimes I_{ir}^\gamma + S_{vi}^\alpha \otimes I_{vi}^\gamma} \quad (6)$$

其中, S_{ir} 和 I_{ir} , S_{vi} 和 I_{vi} 分别是红外图像和可见光图像的显著性增强特征局部对比特征, α 和 γ 则是平衡损失函数中显著特征和局部对比特征的重要性,按照经验将它们设置为0.15和20。

3.3.2 梯度损失

为了保证融合图像具备源图像的梯度,本文选取1范数

来约束融合图像的梯度,如式(7)所示:

$$l_{gra} = \frac{1}{HW} \| |\nabla I_f| - \max(|\nabla I_A|, |\nabla I_B|) \|_1 \quad (7)$$

其中, H 和 W 为输出图像的高和宽, $|\nabla|$ 代表梯度。

4 实验

本章将对实验的具体细节、对比方法的选择和评价指标进行详细介绍。随后,本文将展示与几种最先进方法的比较结果,以证明UMGN网络的有效性。另外,通过参数分析和消融实验,本文将验证UMGN在参数选择和网络设计方面的有效性。

4.1 实验的设置与实现

4.1.1 实验设置

首先介绍本文使用的两个数据集,分别是TNO数据集以及Roadscene数据集。其中,TNO数据集包括来自不同红外传感器的真实场景图像,其包含多种类型的目标物体(如车辆、船只、人、建筑、桥梁等)。Roadscene数据集则包括了真实路面条件下复杂交通的场景图像,具有丰富的背景、车辆和行人信息。本文将UMGN与最先进的9种方法进行比较,例如GTF^[12],LatLRR^[11],DensFuse^[5],GANMcC^[7],IPLF^[6],PIA-Fusion^[10],SSL-WAEIE^[9],CUFD^[8]和SOSMaskFuse^[14]。

4.1.2 实验细节

本文使用PyTorch框架在NVIDIA 3090 GPU上实现了UMGN网络。在训练过程中,将批大小(batch size)设置为1,学习率设为 10×10^{-4} ,并使用Adam优化器对权重进行优化。参数 λ 和 β 在式(2)和式(3)中分别设为10和15。

由于红外与可见光融合的数据集较少,本文从TNO数据集中选择17对图片,以步长为50裁剪至大小为 256×256 图片,并选取1000对图片用于网络的训练,并选取43对图片(其中23对来自TNO,另外20对来自Roadscene数据集)用于测试训练后的模型,其中UMGN测试所用的图片均为原图,并未使用显著性增强后的图片,避免由于使用增强后的图片导致融合结果提高带来的影响。

4.2 融合指标

本文使用4个评估指标用于衡量融合结果,分别是归一化互信息(NMI)^[16]、结构相似性指数(QY)^[17]、基于梯度的融合度量(QABF)^[18]和像素级视觉信息保真度(VIFP)^[19]。NMI^[16]是一种改进的指标,用于衡量融合图像中包含的信息量,相比传统的互信息(MI),NMI^[16]更稳定且对噪声、缺失数据等具有鲁棒性。QY^[17]可以度量源图像和融合图像之间的结构相似性,能够反映出图像中结构信息的保持程度。QABF^[18]是一种基于梯度的融合度量方法,用于评估融合图像中梯度信息的保留情况,能够反映图像的清晰度和边缘保持能力。VIFP^[19]是基于像素级的视觉信息保真度指标,用于衡量融合图像与源图像之间的视觉信息的一致性和保持程度。

4.3 结果分析

TNO数据集以及Roadscene数据集部分图片的融合结果如图4和图5所示。为了更清晰地展示比较结果,背景纹理信息和伪彩处理后的热辐射信息在图片中被红框和绿框标记,并放大显示在图片右上角和右下角。另外,UMGN与其他对比方法的平均指标结果如表1所列,最佳和次佳的结果分别用粗体和斜体标识。不同方法在43对图片上的平均指标如图6所示。

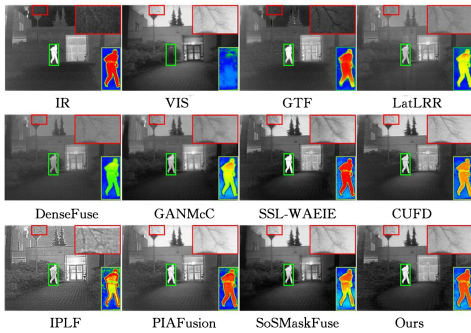


图 4 TNO 数据集的融合结果比较

Fig. 4 Comparison of fusion results of TNO dataset

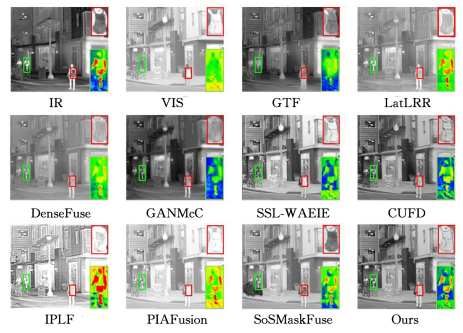


图 5 Roadscene 数据集的融合结果比较

Fig. 5 Comparison of fusion results of Roadscene dataset

表 1 不同方法的 TNO 与 Roadscene 数据的定量结果

Table 1 Averag quantitative results on TNO and Roadscene for different methods

Datasets	TNO datasets				Roadscene datasets			
	VIFP	Q _Y	NMI	Q _{ABF}	VIFP	Q _Y	NMI	Q _{ABF}
GTF	0.2567	0.6584	0.4322	0.3879	0.2484	0.6188	0.4390	0.3401
LatLRR	0.2841	0.7282	0.2742	0.4082	0.2747	0.7138	0.3833	0.4324
DenseFuse	0.2796	0.6747	0.3288	0.3435	0.3003	0.7026	0.4374	0.4038
GANMcC	0.2744	0.6287	0.3102	0.3143	0.2659	0.6075	0.4394	0.3625
IPLF	0.2047	0.6279	0.2751	0.4172	0.2470	0.6233	0.3170	0.3882
PIAFusion	0.3897	0.8604	0.4819	0.5183	0.3476	0.7959	0.5009	0.4524
SSL-WEIE	0.4512	0.8313	0.5878	0.5165	0.3804	0.7712	0.5523	0.5240
CUFD	0.3495	0.7598	0.4982	0.3990	0.2659	0.6075	0.4395	0.3625
SOSMaskFuse	0.4577	0.7887	0.5521	0.4562	0.3646	0.7940	0.4896	0.4048
ours	0.4692	0.8461	0.6011	0.5253	0.4184	0.7813	0.6558	0.5388

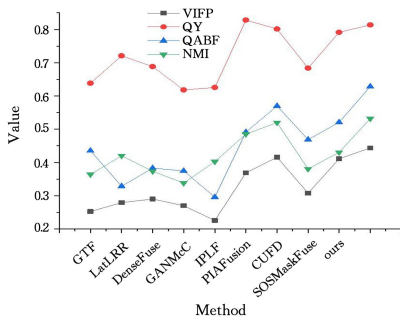


图 6 不同方法在 43 对图片上的平均指标

Fig. 6 Average metrix of different methods in 43 pairs of images

可以清晰地看出,除了 GTF^[12], SSL-WAEIE^[9], PIAFusion^[10], SoSMaskFuse^[14]和本文提出的 UMGN 网络外,其他方法在保留热辐射信息和亮度方面出现了明显的退化。由图 5 可以看出,GTF^[12]过分注重热辐射信息的保留,纹理信息几乎消失;GANMcC^[7]与其类似,几乎仅保留红外显著目标信息,缺乏纹理信息的保留;与之相反,PIAFusion^[10]则是过度强调亮度,在光照充足的情况下,热辐射信息与背景差别不大,难以察觉。从图 4 来看,CUFD 虽然在纹理细节上的保留不输 UMGN,但是对于热辐射信息的保留则不如 UMGN,出现了对比度退化;另外,SoSMaskFuse 由于是以红外图像进行显著目标检测获得掩码进行融合,而忽略了可见光图像对融合图像的贡献,因此在图 5 中可以看出可见光的纹理细节几乎不可见;SSL-WAEIE^[9]虽然不具备上述缺点,但是对热辐射信息的保留则不如 UMGN,发生了显著目标信息丢失的现象;相反,本文提出的 UMGN 方法在保留红外图像的显著目标的同时,也保留了与可见光图像几乎相同的纹理细节,如树枝、天空、灌木丛和背包等。

从表 1 可以看出,在 Roadscene 数据集和 TNO 数据集上,UMGN 在 VIFP^[19], NMI^[16]和 QABF^[18]指标上取得了

最佳结果。虽然在 QY^[17]指标上稍逊于 PIAFusion^[10]和 SoSMaskFuse^[14],但根据图 6 的结果显示,UMGN 在 VIFP^[19], QABF^[18]和 NMI^[16]这 3 个指标的平均值上表现最好,优于其他 9 种对比方法。

因此,本文方法能够保持最高的背景清晰度和边缘梯度。这些结果证明了本文方法在源图像上具有更好的融合效果。

4.4 参数分析

本小节中生成显著性掩码的阈值 λ 以及损失函数的 β 具体取值将通过实验找到最优解。另外,本小节的所有指标均来自 TNO 数据集。

4.4.1 显著性掩码的阈值分析

λ 不同的取值对融合结果的影响如表 2 所列。可以看到,4 个指标都在 λ 取值为 10 时达到最好,因此本文选取 λ 为 10。

表 2 不同阈值下的平均指标

Table 2 Average metrics of different threshold

Parameter	VIFP	Q _Y	NMI	Q _{ABF}
$\lambda=0$	0.4161	0.8267	0.4492	0.5079
$\lambda=5$	0.4446	0.8372	0.5568	0.5155
$\lambda=10$	0.4692	0.8461	0.6011	0.5253
$\lambda=15$	0.4342	0.8223	0.5865	0.5060
$\lambda=20$	0.4315	0.8137	0.5212	0.4851

4.4.2 损失函数的参数分析

β 是梯度损失 $Loss_{gra}$ 在融合损失中的权重,TNO 数据集 23 对图片在 β 不同取值下的平均指标详细信息则可以在表 3 中找到,可以看出 UMGN 在 β 取值为 15 时获得了最好的效果,可以清晰看出 VIFP^[19]以及 QABF^[18]在 β 大于 15 以及小于 15 时都开始轻微下降,表示融合图像的边缘信息在减少,只有在取值为 15 时,融合图像能保留源图像更多的边缘和结构信息,过大或过小的 β 都会导致融合图像含有过多的冗余梯度信息。

表3 不同参数的平均指标

Table 3 Average metrics of different parameters

Parameter	VIFP	Q _Y	NMI	Q _{ABF}
$\beta=5$	0.4319	0.8129	0.4844	0.4894
$\beta=10$	0.4434	0.8335	0.5919	0.5086
$\beta=15$	0.4692	0.8461	0.6011	0.5253
$\beta=20$	0.4506	0.8417	0.5258	0.5208
$\beta=25$	0.4089	0.8002	0.5242	0.4714

4.5 消融实验分析

表4第一行列出了基于红外图像生成掩码在TNO数据集的23张照片上的平均指标结果。表4第二行则是直接生成融合结果而不考虑源图像的贡献的平均指标结果。对加权损失函数的消融实验结果则如表4的第三行和第四行所列,具体来说,本文将对只有保真度损失(L_f)和只有加权保真度损失(L_{wof})分别进行测试,从表4可以看出,所有指标在降级版本都出现了明显下降,由此可以证明UMGN提出的几点都有利于提高融合质量。

表4 消融实验的平均指标

Table 4 Comparison of our method and its degraded versions

Parameter	VIFP	Q _Y	NMI	Q _{ABF}
No Mask _s	0.4531	0.8384	0.5661	0.5224
No weight	0.3354	0.7380	0.3359	0.4410
L_f	0.4138	0.8095	0.5961	0.4879
L_{wof}	0.4418	0.8383	0.5186	0.5143
UMGN	0.4692	0.8461	0.6011	0.5253

结束语 本文提出了一种基于无监督显著性掩码引导的红外与可见光图像融合网络,即UMGN。本文将DenseNet^[5]结构作为骨干网络用于提取源图像特征,并生成权重概率分布,从像素层面评估源图像的融合贡献。同时,引入显著性掩码帮助网络学习红外图像的热辐射信息和可见光图像的纹理信息。此外,加权保真度损失和梯度损失的引入确保了融合图像避免梯度退化现象的出现。通过在公共数据集上进行大量定性和定量实验,验证了所提方法的优越性。

参考文献

- [1] WEI Q,ZHAO J. Research Progress of Infrared and Visible Image Fusion Algorithms[J]. Computer Science,2023,50(2):11.
- [2] ZHOU Z,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.
- [3] ZHANG Q,LEVINE M D, et al. Robust Multi-Focus Image Fusion Using Multi-Task Sparse Representation and Spatial Context [J]. IEEE Transactions on Image Processing,2016,25(5):2045-2058.
- [4] FU Z,WANG X,XU J, et al. Infrared and visible images fusion based on RPCA and NSCT [J]. Infrared Physics & Technology,2016,77:114-123.
- [5] 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.
- [6] ZHU D,ZHAN W,JIANG Y, et al. IPLF: A novel image pair learning fusion network for infrared and visible image [J]. IEEE Sensors Journal,2022,22(9):8808-8817.
- [7] 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,2020,70:1-14.
- [8] XU H,GONG M,TIAN X, et al. CUFD: An encoder-decoder network for visible and infrared image fusion based on common and unique feature decomposition [J]. Computer Vision and Image Understanding,2022,218:103407.
- [9] ZHANG G,NIE R,CAO J. SSL-WAEIE: Self-Supervised Learning with Weighted Auto-Encoding and Information Exchange for Infrared and Visible Image Fusion [J]. IEEE/CAA Journal of Automatica Sinica,2022,9(9):1694-1697.
- [10] TANG L,YUAN J,ZHANG H, et al. PIAFusion: A progressive infrared and visible image fusion network based on illumination aware [J]. Information Fusion,2022,83:79-92.
- [11] LI H,WU X J. Infrared and visible image fusion using latent low-rank representation [J]. arXiv:180408992,2018.
- [12] MA J,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.
- [13] GAO Y H,LUO X Q,ZHANG Z C. Infrared and Visible Image Fusion Based on Feature Separation [J]. Computer Science,2022,49(5).
- [14] LI G,QIAN X,QU X. SOSMaskFuse: An Infrared and Visible Image Fusion Architecture Based on Salient Object Segmentation Mask [J]. IEEE Transactions on Intelligent Transportation Systems,2023,24(9):10118-10137.
- [15] ZHANG J,SCLAROFF S. Saliency detection: A boolean map approach[C]// Proceedings of the IEEE International Conference on Computer Vision,2013:153-160.
- [16] GUO X,NIE R,CAO J, et al. FuseGAN: Learning to fuse multi-focus image via conditional generative adversarial network [J]. IEEE Transactions on Multimedia,2019,21(8):1982-1996.
- [17] LI S,HONG R,WU X. A novel similarity based quality metric for image fusion [C]// Proceedings of the 2008 International Conference on Audio, Language and Image Processing. IEEE,2008.
- [18] XYDEAS C S,PETROVIC V. Objective image fusion performance measure [J]. Electronics letters,2000,36(4):308-309.
- [19] PIELLA G,HEIJMANS H. A new quality metric for image fusion [C]// Proceedings of the Proceedings 2003 international conference on image processing (Cat No 03CH37429). IEEE,2003.



LI Dongyang, born in 1997, postgraduate. His main research interests include deep learning and image fusion.



NIE Rencan, born in 1982, Ph.D, professor, doctoral supervisor. His main research interests include neural networks, image processing and machine learning.