



# 智能系统学报

CAAI TRANSACTIONS ON INTELLIGENT SYSTEMS

## 结合亮度感知与密集卷积的红外与可见光图像融合

曲海成, 王宇萍, 谢梦婷, 肖苇

引用本文:

曲海成,王宇萍,谢梦婷,肖苇. 结合亮度感知与密集卷积的红外与可见光图像融合[J]. 智能系统学报, 2022, 17(3): 643–652.

QU Haicheng,WANG Yuping,XIE Mengting,XIAO Wei. Infrared and visible image fusion combined with brightness perception and dense convolution[J]. *CAAI Transactions on Intelligent Systems*, 2022, 17(3): 643–652.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202104004>

## 您可能感兴趣的其他文章

### 基于生成式对抗网络的道路交通模糊图像增强

Enhancement of blurred road-traffic images based on generative adversarial network

智能系统学报. 2020, 15(3): 491–498 <https://dx.doi.org/10.11992/tis.201903041>

### 融合整体与局部信息的武夷岩茶叶片分类方法

Classification of Wuyi rock tealeaves by integrating global and local information

智能系统学报. 2020, 15(5): 919–924 <https://dx.doi.org/10.11992/tis.202003018>

### 基于生成对抗网络的机载遥感图像超分辨率重建

Super-resolution reconstruction of airborne remote sensing images based on the generative adversarial networks

智能系统学报. 2020, 15(1): 74–83 <https://dx.doi.org/10.11992/tis.202002002>

### 基于注意力融合的图像描述生成方法

An image caption generation method based on attention fusion

智能系统学报. 2020, 15(4): 740–749 <https://dx.doi.org/10.11992/tis.201910039>

### 基于改进的稀疏表示和PCNN的图像融合算法研究

Image fusion based on the improved sparse representation and PCNN

智能系统学报. 2019, 14(5): 922–928 <https://dx.doi.org/10.11992/tis.201805045>

### 基于卷积神经网络的遥感图像分类研究

Classification of remote-sensing image based on convolutional neural network

智能系统学报. 2018, 13(4): 550–556 <https://dx.doi.org/10.11992/tis.201706078>



微信公众平台



期刊网址

DOI: 10.11992/tis.202104004

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1538.tp.20220419.0945.004.html>

# 结合亮度感知与密集卷积的红外与可见光图像融合

曲海成<sup>1</sup>, 王宇萍<sup>1</sup>, 谢梦婷<sup>1</sup>, 肖苇<sup>2</sup>

(1. 辽宁工程技术大学 软件学院, 辽宁 葫芦岛 125105; 2. 中国铁路沈阳局集团有限公司阜新车务段, 辽宁 阜新 123100)

**摘要:** 为解决弱光照条件下红外与可见光图像融合质量差的问题, 提出一种结合亮度感知与密集卷积的红外与可见光图像融合方法 (brightness perception and dense convolution, BPD-Fusion)。首先, 对可见光图像进行亮度计算, 得到亮度权重并对其暗区域进行亮度增强; 然后, 将增强的可见光图像与红外图像级联输入生成器, 在其 Conv1 阶段后嵌入密集卷积以获取更丰富的图像特征; 最后, 为了达到较强的图像重构与生成能力, 建立多损失函数构建端到端的图像融合过程。在 TNO 和 KAIST 数据集上进行融合质量测评: 主观评价上, 提出的方法视觉效果良好; 客观评价上, 差异相关和、信息熵、互信息和平均梯度指标均优于对比方法。

**关键词:** 图像融合; 亮度感知; 密集卷积网络; 对抗生成网络; 红外与可见光图像; 信息熵; 互信息; 差异相关和  
**中图分类号:** TP391    **文献标志码:** A    **文章编号:** 1673-4785(2022)03-0643-10

中文引用格式: 曲海成, 王宇萍, 谢梦婷, 等. 结合亮度感知与密集卷积的红外与可见光图像融合 [J]. 智能系统学报, 2022, 17(3): 643-652.

英文引用格式: QU Haicheng, WANG Yuping, XIE Mengting, et al. Infrared and visible image fusion combined with brightness perception and dense convolution[J]. CAAI transactions on intelligent systems, 2022, 17(3): 643-652.

## Infrared and visible image fusion combined with brightness perception and dense convolution

QU Haicheng<sup>1</sup>, WANG Yuping<sup>1</sup>, XIE Mengting<sup>1</sup>, XIAO Wei<sup>2</sup>

(1. School of Software, Liaoning Technical University, Huludao 125105, China; 2. Fuxin Depot of China Railway Shenyang Bureau Group Co., Ltd, Fuxin 123100, China)

**Abstract:** To solve the problem of poor fusion quality of infrared and visible images under weak illumination, we propose an infrared and visible image fusion method (BPD-Fusion) combining brightness perception and DenseNet. First, the brightness of the visible image is evaluated to obtain its weight and enhance the brightness of the dark area. Then, the enhanced visible and infrared images are input into the generator after concatenating, and DenseNet is embedded after the Conv1 stage to obtain more abundant source image features. Finally, to achieve stronger image reconstruction and generation ability, a multi-loss function is established to construct the end-to-end image fusion process. The fusion quality is evaluated on the TNO dataset and the challenging KAIST dataset. In subjective evaluation, a good visual effect is observed in the proposed method. In objective evaluation, the difference correlation sum, information entropy, mutual information, and average gradient of our method are better than those of the contrast method.

**Keywords:** image fusion; brightness perception; dense convolution network; GAN; infrared image and visible image; information entropy; mutual information; sum of difference and correlation

图像融合是一种图像增强技术, 其目的是融合不同传感器提供的图像, 生成信息更丰富的图

像, 以便于后续处理。其中红外与可见光图像融合是计算机视觉领域中非常重要的部分。首先, 红外图像与可见光图像的信号来自不同的模式, 因此可以从不同方面提供场景信息; 其次, 二者几乎呈现了物体的所有特性, 并且具有互补性。

收稿日期: 2021-04-02. 网络出版日期: 2022-04-20.

基金项目: 辽宁工程技术大学学科创新团队资助项目 (LNTU20TD-23); 辽宁省教育厅一般项目 (LJ2019JL010).

通信作者: 曲海成. E-mail: [quhaicheng@lntu.edu.cn](mailto:quhaicheng@lntu.edu.cn).

可见光图像捕获反射光,通常具有较高的空间分辨率和可观的细节和明暗度,但很容易受到恶劣条件如光照差、烟雾等的影响,而捕获物体热辐射的红外图像能够抵抗这些干扰,但通常分辨率低,纹理较差。根据二者的成像特点进行图像融合,得到的互补融合图像可以广泛应用到遥感图像分类<sup>[1-2]</sup>、目标检测<sup>[3-4]</sup>和视频监控<sup>[5]</sup>等领域。

图像融合分为像素级融合、特征级融合和决策级融合。在过去的几十年里,国内外学者提出了多种红外与可见光图像融合方法。在像素级融合方法中,具有代表性的方法有变换域的 ADF<sup>[6]</sup>、CBF<sup>[7]</sup>和 WLS<sup>[8]</sup>和稀疏域的 ASR<sup>[9]</sup>、LP<sup>[10]</sup>等方法。变换域中基于多尺度变换<sup>[11]</sup>的方法最为活跃,将源图像分解成多个层次,用特定的规则融合相应的层次,再重建目标图像,但由于其忽略了融合过程中的空间一致性,容易在融合图像中引入光晕。稀疏域的方法<sup>[12]</sup>需要构建一个过完备字典,需要多次迭代,非常耗时。图像融合中合适的信息提取方法和有效的融合原理是保证融合性能的关键<sup>[13]</sup>。近年来随着深度学习的兴起,国内外学者提出了许多基于深度学习的融合方法。2017年,Prabhakar等<sup>[14]</sup>针对多曝光融合问题提出了基于卷积神经网络的图像融合方法(DeepFuse),该方法网络结构过于简单,并且只使用编码网络中最后一层计算的结果,中间层获得的有用信息会丢失。2019年, Ma等<sup>[15]</sup>提出了一种基于生成对抗网络的融合方法(a GAN network for image fusion, FusionGAN),在保留热辐射信息的同时保留细节纹理信息,达到了较好的性能。但生成器与判别器网络结构相对简单,损失函数只考虑梯度来衡量纹理信息,从而导致一些重要信息的丢失。2018年, Li等<sup>[16]</sup>提出了一种基于深度网络的红外与可见光图像融合方法,这种方法将源图像通过导向滤波的方式进行分解,得到图像的基础部分和细节内容,对基础部分和细节内容分别融合。该方法耗费大量计算资源。2019年, Li等<sup>[17]</sup>提出了带有密集连接的红外与可见光图像融合方法(DenseFuse),网络由编码网络、融合层和解码网络组成。该方法采用相同的编码网络对红外图像与可见光图像进行特征提取,会存在特征丢失问题。2020年, Xu等<sup>[18]</sup>提出了一种端到端的统一无监督学习图像融合方法(U2Fusion),是一种通用图像融合方法,但训练起来非常复杂,占用大量资源。

上述研究主要针对光照条件较好的场景图像进行融合。由于光照条件好时,可见光图像的纹

理细节特征清晰,而光照条件差时,可见光图像的伪影和噪声会影响融合质量。于是,为消除光照条件对融合图像性能的影响,提高暗条件下融合图像目标的清晰度,提出一种结合亮度感知与密集卷积的红外与可见光图像融合方法(BPD-Fusion)。本文的主要工作:1)建立可见光图像暗区域增强模型,消除光照条件对融合性能的影响;2)将密集卷积嵌入特征提取网络,提取图像深层特征同时保留浅层特征,避免梯度消失问题;3)由多损失函数构建端到端的无监督学习图像融合过程;4)将提出的方法在 TNO 和 KAIST 数据集上进行对比实验。

## 1 相关研究

### 1.1 FusionGAN 方法

Ma等<sup>[15]</sup>提出的 FusionGAN 方法首次将对抗生成网络应用于图像融合领域。FusionGAN 方法整体框架如图 1 所示。生成器是整个模型的主要网络,影响最终融合图像的质量。判别器的目的是将可见光图像的细节信息加入到融合图像。训练阶段,生成器网络与判别器网络协同、交替更新网络参数;测试时直接利用训练好的生成器生成融合图片。通过生成器与判别器对抗训练,使得最终的融合结果中既包含红外热辐射信息,又包含可见光图像的细节特征。

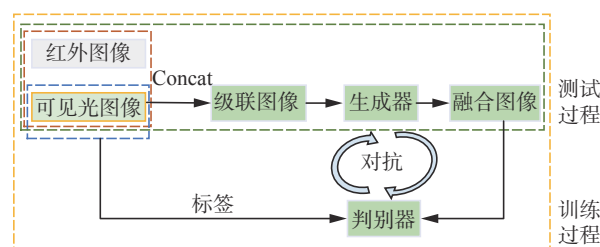


图 1 FusionGAN 的整体框架  
Fig. 1 Entire process of FusionGAN

FusionGAN 方法存在的问题:1)未考虑暗条件下可见光图像暗区域会给融合图像引入干扰信息;2)生成器与判别器网络结构相对简单,突出特征提取不当;3)损失函数只考虑可见光梯度和红外强度来衡量融合图像与源图像的关系,会导致一些重要信息的丢失。

### 1.2 密集卷积网络

为解决特征提取过程中特征丢失问题,对密集卷积网络进行了调研。Huang等<sup>[19]</sup>为了使网络中所有层之间的信息流最大,将网络中的所有层两两进行连接,使得网络中每一层都接受它前面所有层的特征作为输入,并将这种网络结构称为 DenseNet。密集卷积网络实现了信息流动最大

化,通过多个前层与后层的连接,保证前层信息得到充分利用,使所有中间层信息得以保留。为了进行特征复用,在跨层连接时使用的是在特征维度上的 Concatenate 操作,每经过一个单元模块,下一层的特征维度数量就会增加  $k$ ,  $k$  值越大意味着在网络中流通的信息越大,但是整个模型的尺寸和计算量也会增加。为了避免随着网络层数的增加,后面层的特征维度增长过快,本文方法在每个阶段之后进行下采样时,首先通过一个卷积层将特征维度压缩至当前输入一个适当值,然后再进行 Pooling 操作,这样能适当降低网络复杂度和计算量。

### 1.3 多损失函数

在无监督学习图像融合方法中,损失函数在融合图像生成过程中发挥重要作用。FusionGAN 的生成器损失主要由对抗损失、红外强度损失及可见光梯度损失组成,没有考虑到红外图像的热辐射信息和可见光图像的纹理信息不能单纯靠强度梯度损失来约束,忽略了源图像与融合图像整体信息关联性。结构相似性是衡量两幅图像相似

度的指标,本文方法提出多损失函数模型来提高网络的学习能力,在内容损失中增加结构相似性损失和像素损失,更具针对性地保留源图像特征信息。

## 2 BPD-Fusion

由于亮度对红外与可见光图像融合性能有较大影响,本文提出结合亮度感知与密集卷积的红外与可见光图像融合方法(BPD-Fusion)。BPD-Fusion 方法整体框架如图2所示,训练时,对于生成器,首先将可见光图像输入亮度感知模块,得到暗区域增强的可见光图像;然后,将可见光图像与红外图像 Concat,通过密集卷积模块提取丰富的图像特征,将特征图经特征重构成融合图像。对于判别器,输入可见光图像或生成器输出的初步融合图像,由卷积神经网络提取图像特征,输出对两类图像的判别值,以便于将可见光图像的细节信息不断地加入到融合图像中。整个过程以多损失函数约束融合图像与源图像的关系,增强模型图像生成与重构能力。

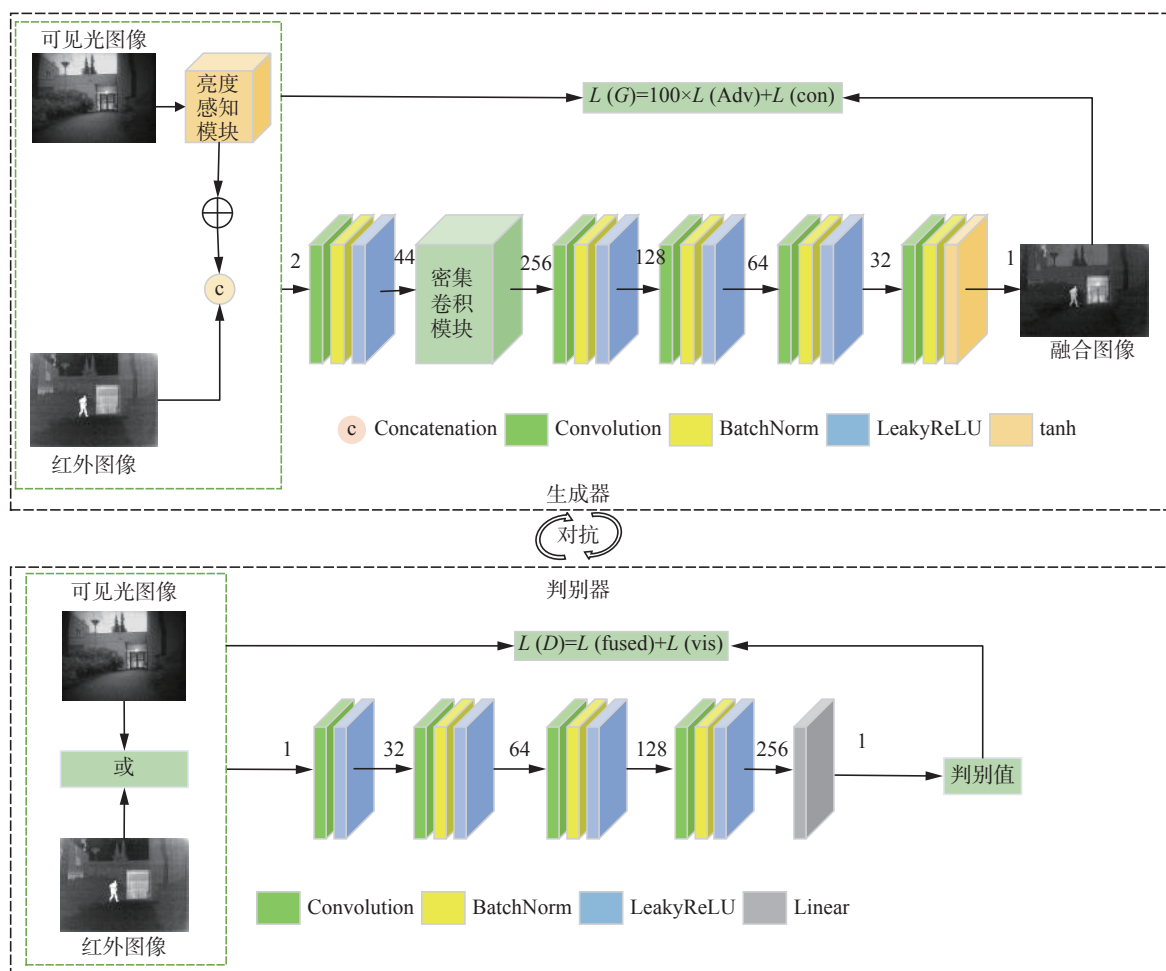


图2 BPD-Fusion 整体框架

Fig. 2 Overall framework of BPD-Fusion



生成器包含了亮度感知模块、输入层、特征提取模块、特征重构模块和输出层 5 个部分。将可见光图像输入亮度感知模块,得到暗区域增强的可见光图像,然后将可见光图像与红外图像 Concat 作为 2 通道输入图像,进入密集卷积模块进行特征提取,得到 264 通道特征图,通过特征重构得到单通道融合图像。其具体网络结构参数见表 1。Conv1 采用  $5 \times 5$  大小的卷积核,以较大感受野获取源图像特征。

判别器的目的是将可见光图像的细节信息加入到融合图像中。若对可见光图像的判别结果越大,则判别能力越强<sup>[20]</sup>。判别器包含 4 个卷积层和 1 个线性层,卷积层以  $3 \times 3$  大小的卷积核进行特征提取,填充方式均为 VALID,激活函数为 LeakyReLU,除第一个卷积层外,均使用 Batch-Norm 进行参数批归一化处理,提高模型收敛速度。线性层输出对输入图像的判别结果。判别器的网络结构及参数信息见表 2。

表 1 亮度感知机制网络结构

Table 1 Brightness sensing mechanism and network structure

名称	网络层	$k$	$s$	$n_1$	输入	$n_2$	输出	填充	激活函数
亮度感知	Conv_l1	$3 \times 3$	1	1	VIS	64	I1	VALID	ReLU
	Conv_l2	$3 \times 3$	1	64	I1	256	I2	VALID	ReLU
	gap	$3 \times 3$	1	256	I2	1	$w \times \text{VIS}$	VALID	ReLU
	Add	—	—	—	VIS	—	VIS*	VALID	ReLU
输入层	Conv1	$5 \times 5$	1	2	Concat(IR, VIS)	44	Net1	VALID	LeakyReLU
特征提取	Dense1	$3 \times 3$	1	44	Net1	44	Net2	SAME	LeakyReLU
	Dense2	$3 \times 3$	1	88	Concat(Net1, Net2)	44	Net3	SAME	LeakyReLU
	Dense3	$3 \times 3$	1	132	Concat(Net1, Net2, Net3)	44	Net4	SAME	LeakyReLU
	Dense4	$3 \times 3$	1	176	Concat(Net1, Net2, Net3, Net4)	44	Net5	SAME	LeakyReLU
	Dense5	$3 \times 3$	1	220	Concat(Net1, Net2, Net3, Net4, Net5)	44	Net6	SAME	LeakyReLU
特征重构	Conv2	$5 \times 5$	1	264	Concat(Net1, Net2, Net3, Net4, Net5, Net6)	128	Net7	VALID	LeakyReLU
	Conv3	$3 \times 3$	1	128	Net7	64	Net8	VALID	LeakyReLU
	Conv4	$3 \times 3$	1	64	Net8	32	Net9	VALID	LeakyReLU
输出层	Conv5	$1 \times 1$	1	32	Net9	1	融合图像	VALID	tanh

注: VIS代表可见光图像;IR代表红外图像; $w$ 为亮度权重; $k$ 为卷积核大小; $s$ 为移动步长; $n_1$ 为输入通道数; $n_2$ 为输出通道数;VIS\*为预处理后的可见光图像

表 2 判别器网络整体结构

Table 2 Overall structure of the discriminator network

名称	网络层	$k$	$s$	$n_1$	输入	$n_2$	输出	填充	激活函数/操作
特征提取	Conv2_1	$3 \times 3$	1	1	VIS/ IR	32	Net1	VALID	LeakyReLU
	Conv2_2	$3 \times 3$	1	32	Net1	64	Net2	VALID	LeakyReLU
	Conv2_3	$3 \times 3$	1	64	Net2	128	Net3	VALID	LeakyReLU
	Conv2_4	$3 \times 3$	1	128	Net3	256	Net4	VALID	LeakyReLU
分类	Line_5	—	—	256	Linear(Net4)	1	判别值	—	matmul

注: VIS代表可见光图像;IR代表红外图像; $k$ 为卷积核大小; $s$ 为移动步长; $n_1$ 为输入通道数; $n_2$ 为输出通道数;matmul为矩阵相乘

## 2.1 亮度感知与暗区域亮度增强

为了提高融合图像质量,消除光照环境对融合性能的影响,设计了亮度感知模块。该模块对可见光图像进行亮度计算,按亮度权重为图像暗

区域进行亮度增强,避免光照条件差时可见光图像向融合图像引入较多噪声。该模块对输入的多通道可见光图像以 YCbCr 模式读取。首先,取第一个通道 Y 即亮度通道将其转为灰度图像,此时

图像的灰度值即亮度值;然后,通过卷积操作得到特征图,使用平均池化计算平均亮度;最后,使用 Sigmoid 操作将图像亮度转换到 0 到 1 之间作为亮度权重,在通道维对图像进行亮度加权,在保持较亮区域亮度同时,增强暗区域的亮度,并与源图像做残差处理,避免特征丢失问题。预处理完毕后,得到增强的可见光图像作为下一个模块的输入。该模块整体结构如图 3 所示。

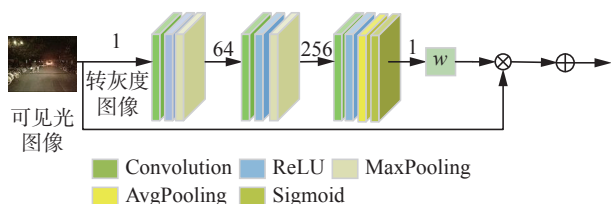


图 3 亮度感知模块整体结构图

Fig. 3 Structural diagram of the brightness sensing module

模块的输入为多通道彩色或灰度可见光图像,然后将其转换成单通道灰度图像输入卷积神经网络。网络的第一层 Conv\_11 和第二层 Conv\_12 均由 Conv+ReLU+MaxPooling 组成,第三层 gap 由 Conv+ReLU+Avg\_Pooling+Sigmoid 组成。图 3 中数字代表图像通道数。

网络的参数设置及功能如下:

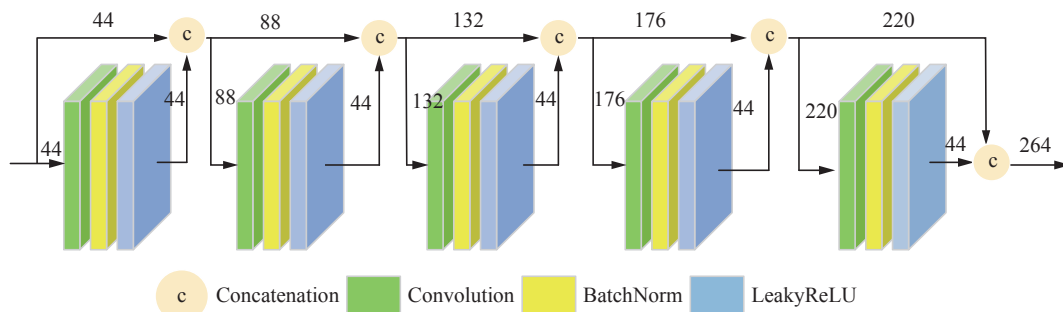


图 4 密集卷积模块结构图

Fig. 4 Structural diagram of the dense convolution module

网络的参数设置及功能如下:

1) 卷积核大小设置。密集卷积模块为了提取深层特征,5 个卷积层均使用  $3 \times 3$  大小的内核进行内部特征提取,将特征进行通道维级联可避免特征丢失,使融合图像内容更丰富。

2) 步长与填充设置。每个卷积的步长均为 1,为保证特征图尺寸不变,卷积时使用 SAME 方式进行填充,减少了中间层的特征丢失,层与层之间的特征关联更强,使得融合图像与源图像具有高度相关性。

3) 批归一化与激活函数。BatchNorm 操作对卷积层输出进行批归一化处理,可加快模型收敛速度,前向传播时每一层直接与密集块输入相

1) 卷积核与激活函数。Conv\_11、Conv\_12 和 gap 这 3 个卷积层均使用  $3 \times 3$  大小的卷积核进行内部特征提取,激活函数使用 ReLU。

2) 步长与填充设置。每个卷积的步长均为 1,在特征通道数增加的同时,在 Conv\_11 和 Conv\_12 阶段后,使用最大池化减小特征图尺寸,卷积时不需填充,因此填充方式为 VALID。

3) 平均池化与 Sigmoid 输出。在 gap 层,使用平均池化 reduce\_mean 操作得到平均亮度,然后,通过 Sigmoid 函数将图像的平均亮度值转换成 (0,1) 之间的值得到亮度权重,对图像进行暗区域亮度增强,得到增强的可见光图像。

## 2.2 密集卷积结构

为保证融合图像包含更多源图像信息,将密集卷积网络应用于图像融合。从 Conv1 阶段得到的 44 通道特征图输入密集卷积模块,对于块内每一层,其输入特征图为前面所有层输出特征通道的级联。增长率设置为 44,为提高模型计算速度,使内部每个卷积块均输出 44 通道特征图,在保留浅层特征的同时,又增加深层特征。最后整个密集块的输出为 264 通道特征图。具体网络结构设计如图 4 所示。

连,反向传播时误差可以更早地传播到前层,可有效避免梯度消失问题,有利于提高模型训练的稳定性<sup>[15]</sup>。每层使用 LeakyReLU 激活函数。

## 2.3 多损失函数

BPD-Fusion 的损失函数包括生成器损失和判别器损失两部分:生成器损失计算融合图像与源图像之间的差别;判别器以可见光图像作为标签,训练辨别可见光图像与融合图像的能力。

### 1) 生成器损失函数

生成器损失 ( $L_G$ ) 包含对抗损失 ( $L_{Adv}$ ) 及内容损失 ( $L_{con}$ ), 其定义为

$$L_G = 100L_{Adv} + L_{con}$$

内容损失 ( $L_{con}$ ) 是生成器的主要损失函数,决

定了融合图像与源图像间的相关性及重点提取红外图像与可见光图像的哪些特征。内容损失越小,融合图像中源图像的特征越多。本文方法在内容损失中加入结构相似性损失和像素损失,更具针对性地保留源图像特征信息。BPD-Fusion 的内容损失由强度梯度损失 ( $L_{\text{Grad}}$ )、结构相似度损失 ( $L_{\text{SSIM}}$ ) 和像素损失 ( $L_{\text{Pixel}}$ ) 组成,其定义为

$$L_{\text{con}} = L_{\text{Grad}} + 100 \times L_{\text{SSIM}} + 100 \times L_{\text{Pixel}}$$

$$L_{\text{Grad}} = \frac{1}{HW} (\|F - I\|_F^2 + 800 \times \|\nabla F - \nabla I\|_F^2)$$

$$L_{\text{SSIM}} = 1 - 0.5 \times (S(I, F) + S(V, F))$$

$$L_{\text{Pixel}} = 0.5 \times \|F - I\|_2 + 0.5 \times \|F - V\|_2$$

对抗损失 ( $L_{\text{Adv}}$ ) 是生成器与判别器交互的损失函数,其定义为

$$L_{\text{Adv}} = \frac{1}{N} \sum_{n=1}^N (D(F) - a)^2$$

$D(F)$  为判别器对融合图像的判别结果,在  $(0, 1)$  之间,  $a$  为软标签(取值不确定的标签)。 $a$  值越大,意味着融合图像正在向可见光图像靠近,又要保证损失最小,根据文献 [15] 将软标签  $a$  的取值范围设置为  $(0.7, 1.2)$ 。

### 2) 判别器损失函数

判别器损失函数 ( $L_D$ ) 的目的是约束判别器网络反向传播参数更新,训练判别器对可见光图像和融合图像的判别能力,其定义为

$$L_D = \frac{1}{N} \sum_{n=1}^N (D(V) - b)^2 + \frac{1}{N} \sum_{n=1}^N (D(F) - c)^2$$

式中:  $D(V)$  为判别器对可见光图像的判别结果;  $D(F)$  为判别器对融合图像的判别结果;  $b$ 、 $c$  为软标签(取值不确定的标签),根据文献 [15] 将  $b$  的范围设置为  $(0.7, 1.2)$ , 将  $c$  的范围设置为  $(0, 0.3)$ 。

## 3 实验结果与分析

实验在 Windows10 操作系统和 NVIDIA GTX 1080Ti GPU 上完成,采用深度学习框架 TensorFlow 1.12.0。训练过程学习率设为 0.000 01, batch\_size 设置为 32, 训练迭代次数设置为 20, 其中每次判别器训练次数设置为 2。BPD-Fusion 方法主要目的是解决暗条件红外与可见光图像融合效果差的问题,本节采用 TNO<sup>[21]</sup> 数据集的 35 对白天图像和 KAIST<sup>[22]</sup> 数据集的 21 对夜间图像评估提出方法的性能。选择 ADF<sup>[6]</sup>、CBF<sup>[7]</sup>、WLS<sup>[8]</sup>、Deepfuse<sup>[14]</sup>、deeplearning<sup>[16]</sup> 和 Densefuse<sup>[17]</sup> 这 6 种方法与本文方法作对比实验,对融合图像进行主观评价和客观评价。对比实验图片由 PyCharm 和 Matlab 程序生成,客观评价指标计算通过 MATLAB 软件实现。

### 3.1 主观评价

BPD-Fusion 与对比方法在 TNO 数据集中傍晚的“房屋与车辆”实验结果如图 5 所示。

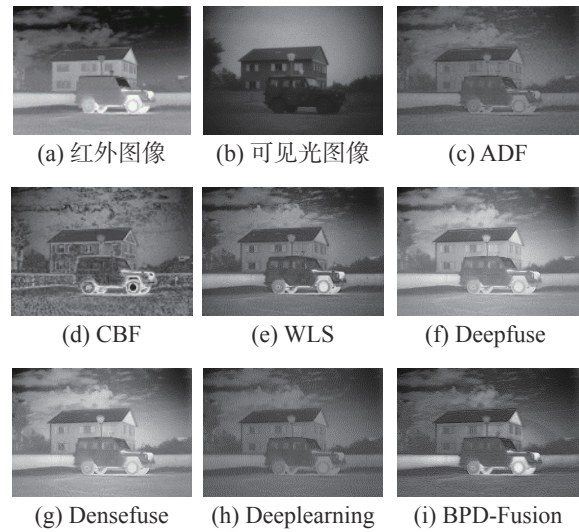


图 5 TNO 数据集中“房屋与车辆”融合结果

Fig. 5 Fusion results of “house and vehicle” in TNO dataset

从图 5 的实验结果可以看出:图(c)中,房屋左侧的树轮廓模糊,整幅图像清晰度较低;图(d)中,整体噪声大,视觉效果较差;图(e)中,天空中的云边缘不清晰,整体纹理细节信息少;图(f)、(g)中,红外目标较亮,可见光图像纹理信息部分丢失,房屋前的路灯周围有光晕;图(h)中,整体背景偏暗,有轻微伪影;图(i)中,整体背景较亮且清晰,目标突出,能清楚地识别地砖、树木、云、房屋、路灯和车辆等。

BPD-Fusion 与对比方法在 KAIST 数据集中路灯下的“行人与自行车”实验结果如图 6 所示。

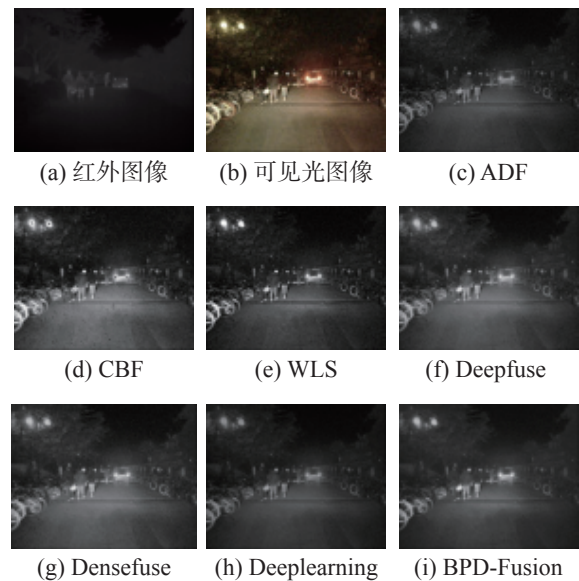


图 6 KAIST 数据集中“行人与车辆”融合结果

Fig. 6 Fusion results of “pedestrian and vehicles” in KAIST dataset



从图 6 的实验结果可以看出: 图(c)、(e)中, 左侧树旁边的暗处细节不清晰; 图(d)中, 图像整体上有明显的黑块, 不够平滑; 图(f)、(g)中, 人与汽车的轮廓都比较模糊; 图(h)中, 右边一排暗处的自行车边缘不清晰; 图(i)中, 自行车与路中间的人轮廓清晰可见, 图像整体平滑, 细节信息与热辐射信息都能较好地体现。

BPD-Fusion 与对比方法在 KAIST 数据集中夜晚照明条件差的“行人”实验结果如图 7 所示。

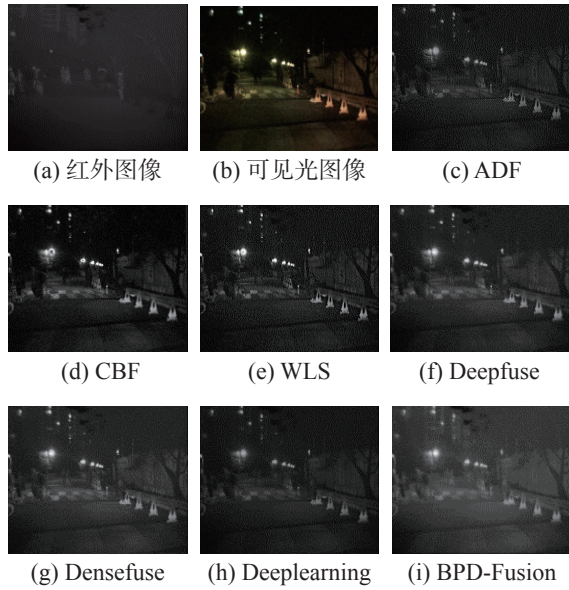


图 7 KAIST 数据集中“行人”融合结果

Fig. 7 Fusion results of “pedestrian” in the KAIST dataset

从图 7 的实验结果可以看出: 图(c)中, 背景较暗, 道路中间的两个行人模糊; 图(d)中, 路灯中心有黑块, 整体噪声较大; 图(e), 图像整体不够平滑, 暗处物体特征丢失; 图(f)、(g)中, 人的轮廓信息不清楚, 丢失了可见光图像的细节信息; 图(h)中, 整体偏暗, 目标与背景没有很好地区分开; 图(i)中, 整体较亮, 可以看到每层楼的窗户轮廓, 道路中间和左边的行人轮廓清晰, 墙壁、树、路障的细节信息丰富。

### 3.2 客观评价

为了进一步评价融合图像的融合效果, 使用客观评价方法进行评估。选取信息熵 (information entropy, EN)<sup>[23]</sup>、互信息 (mutual information, MI)<sup>[24]</sup>、差异相关和 (sum of correlation and difference, SCD)<sup>[25]</sup> 和平均梯度 (average gradient, AG) 为测评指标, 从图像信息丰富度、互补性和纹理细节等不同角度对融合图像进行评价。以下为测评指标的计算公式。

1) 信息熵 (EN)。其值大小能够体现融合图像中平均信息量的多少, 其定义为

$$EN = -\sum_{i=0}^L p_i \log_2 p_i$$

式中:  $p_i$  为灰度值为  $i$  的像素数与图像总像素数之比, 信息熵越大, 表明图像所含信息越丰富。

2) 互信息 (MI)。该指标的大小表示融合图像中源图像信息量的多少, 用于衡量二者之间的信息相关性, 互信息值越大, 代表融合图像包含源图像的信息越多, 图像融合效果越好, 计算方法为

$$\begin{cases} MI_{A,F} = \sum_{i=1}^{L-1} \sum_{j=1}^{L-1} h_{A,F}(i,j) \cdot \log_2 \frac{h_{A,F}(i,j)}{h_A(i) + h_F(j)} \\ MI_{B,F} = \sum_{i=1}^{L-1} \sum_{j=1}^{L-1} h_{B,F}(i,j) \cdot \log_2 \frac{h_{B,F}(i,j)}{h_B(i) + h_F(j)} \\ MI_{AB,F} = \frac{MI_{A,F} + MI_{B,F}}{2} \end{cases}$$

式中:  $A$ 、 $B$  为源图像;  $F$  为融合图像。

3) 差异相关和。该指标衡量两源图像之间信息的互补性, 差异之和越大, 说明融合图像中互补信息越多, 图像融合效果越好, 其定义为

$$\begin{aligned} SCD &= r_{D_A,A} + r_{D_B,B} \\ D_A &= F - B \\ D_B &= F - A \end{aligned}$$

式中:  $F$  为融合图像;  $r$  为计算  $D_A$  和  $D_B$  与两张源图像  $A$  与  $B$  的相关系数;  $D_A$  和  $D_B$  分别定义融合图像与源图像之间的差异。

4) 平均梯度。该指标反映了图像中细节反差与纹理变化特征及图像的清晰度, 其值越大, 图像越清晰, 其定义为

$$AG = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \sqrt{\frac{\Delta_x f(i,j)^2 + \Delta_y f(i,j)^2}{2}}$$

式中:  $\Delta_x f(i,j)$ 、 $\Delta_y f(i,j)$  分别为像素  $(i,j)$  在  $x$ 、 $y$  方向上的一阶差分。

首先, 在 TNO 数据集中选取了 35 对照明条件一般的红外与可见光图像, 由 BPD-Fusion 与 6 种对比方法分别生成融合图像进行测评, 测评结果见表 3。表 3 中数值均为每种方法生成的 35 张融合图像在 EN、MI 和 SCD 指标上测评结果的平均值, 所有指标的值越大越好。

由表 3 可以看出, BPD-Fusion 在 EN、MI 两个指标中取得最大值, 相比其他方法有明显的提高, 说明得到的融合图像信息更丰富。本组源图像光线稍好, 保留了更多的可见光图像有利细节信息, 互补性不是最大, 差异相关和较最优值稍低一点, 但主观评价和客观评价上仍然达到了较好的效果。

然后, 为了验证 BPD-Fusion 在暗条件下融



合的优越性,在 KAIST 数据集中选取 21 对照明条件较差的红外与可见光图像,与 6 种对比方法的测评结果见表 4。表 4 中数值均为每种方法生成的 21 张融合图像在 EN、MI 和 SCD 指标上测评结果的平均值,其值越大越好,计算时间越小越好。

表 3 TNO 数据集 35 组融合图像客观评价指标均值  
Table 3 Mean value of the objective evaluation of the fusion images of 35 groups in the TNO dataset

方法名称	EN	MI	SCD
WLS	6.803 5	13.607 1	1.778 7
ADF	6.402 2	12.825 7	1.639 5
CBF	6.906 9	13.914 5	1.391 5
Densefuse	6.836 4	13.689 1	1.831 1
Deepfuse	6.868 3	13.736 9	<b>1.842 2</b>
Deeplearning	6.301 2	12.315 6	1.356 9
BPD-Fusion	<b>6.990 5</b>	<b>13.981 0</b>	1.757 3

注:加粗字体为最大值

由表 4 可见,BPD-Fusion 在 KAIST 数据集上的融合图像其 EN、MI 和 SCD 这 3 个方面都是最优的。KAIST 数据集中图像偏暗,BPD-Fusion 中的亮度感知机制很好地感知了亮度条件,提取源

图像中有用的信息,抑制干扰信息,使融合图像测评指标值提高。

表 4 KAIST 数据集 21 组融合图像客观评价指标均值  
Table 4 Mean values of the objective evaluation of the fusion images of 21 groups in KAIST dataset

方法名称	EN	MI	SCD	计算时间/s
WLS	6.314 0	12.628 0	1.301 2	1.762
ADF	6.130 7	12.261 3	1.170 7	0.526
CBF	6.754 0	13.507 9	0.979 9	22.903
Densefuse	6.580 3	13.160 7	1.632 8	0.509
Deepfuse	6.592 7	13.185 4	1.639 4	0.421
Deeplearning	6.085 1	12.170 2	1.736 1	8.092
BPD-Fusion	<b>6.827 3</b>	<b>13.654 6</b>	<b>1.655 6</b>	<b>0.195</b>

注:加粗字体为最优值

最后,为验证 BPD-Fusion 中亮度感知暗区域增强模型、密集卷积模块及多损失函数的作用,使用 KAIST 数据集中的 21 对典型红外与可见光图像做消融实验,采用客观评价对所有方法生成的融合图像进行测评,测评结果见表 5。表 5 中数值均为每种方法生成的 21 张融合图像在 EN、MI、SCD 和 AG 指标上测评结果的平均值,所有指标的值越大越好。

表 5 KAIST 数据集消融实验客观评价  
Table 5 Objective evaluation of the ablation experiment in KAIST data set

方法编号	FusionGAN	多损失函数	密集卷积	亮度感知	EN	MI	SCD	AG
1	√				5.909 0	11.818 1	1.194 4	0.301 2
2	√	√			6.455 5	12.911 0	1.520 7	0.340 6
3	√	√	√		6.585 2	13.170 4	1.619 6	0.367 5
4	√	√	√	√	<b>6.827 3</b>	<b>13.654 6</b>	<b>1.655 6</b>	<b>0.501 6</b>

注:加粗字体为最大值

由表 5 可以看出,方法 2 在方法 1 基础上采用多损失函数,3 个指标较方法 1 均有提升,在生成器损失函数中加入结构相似度损失和像素损失,更加全面地优化了融合图像。方法 3 在方法 2 基础上,在生成器的 Conv1 阶段后嵌入密集卷积模块,指标值较方法 2 均有提升,密集卷积提取深层特征同时保留浅层特征,优化时避免了梯度消失问题,使融合图像中的互补信息增多。方法 4 在方法 3 的基础上加入亮度感知模块,对可见光图像的暗区域进行亮度增强,充分利用了可见

光图像的细节特征与红外图像的热辐射特征,使测评结果大幅提高。方法 4 即 BPD-Fusion 生成的融合图像如图 5、图 6 和图 7 所示,图像中的信息更丰富,视觉效果良好。

## 4 结束语

针对弱光照条件下红外与可见光图像融合效果较差的问题,本文提出一种结合亮度感知与密集卷积的红外与可见光图像融合方法。该方法使用亮度感知机制,对可见光图像暗区域进行亮度

增强处理, 向融合图像中加入有利的可见光图像细节信息, 抑制不利信息的引入。同时, 在生成器特征提取阶段嵌入密集卷积, 减少了中间层的特征丢失, 同时解决了梯度消失问题, 使得融合图像与源图像具有高度相关性。此外, 使用多损失函数构建端到端的无监督学习图像融合过程, 增强模型学习能力。主观评价和客观评价结果表明, 提出的方法在图像的视觉观察、目标清晰度和信息互补性等方面具有良好的性能。本文方法可应用于目标检测和目标追踪等领域, 进一步探索如何保留色彩信息是今后的工作方向。

## 参考文献:

- [1] 刘帅, 李士进, 冯钧. 多特征融合的遥感图像分类 [J]. *数据采集与处理*, 2014, 29(1): 108–115.  
LIU Shuai, LI Shijin, FENG Jun. Remote sensing image classification based on adaptive fusion of multiple features[J]. *Journal of data acquisition and processing*, 2014, 29(1): 108–115.
- [2] 孙洁, 黄承宁, 王玉祥. 基于多模态聚类及决策融合的 SAR 图像分类方法 [J]. *现代雷达*, 2020, 42(12): 66–71.  
SUN Jie, HUANG Chengning, WANG Yuxiang. Target classification of SAR images based on clustering of multiple modes and decision fusion[J]. *Modern radar*, 2020, 42(12): 66–71.
- [3] 白玉, 侯志强, 刘晓义, 等. 基于可见光图像和红外图像决策级融合的目标检测算法 [J]. *空军工程大学学报(自然科学版)*, 2020, 21(6): 53–59, 100.  
BAI Yu, HOU Zhiqiang, LIU Xiaoyi, et al. An object detection algorithm based on decision-level fusion of visible light image and infrared image[J]. *Journal of Air Force Engineering University (natural science edition)*, 2020, 21(6): 53–59, 100.
- [4] 李舒涵, 许宏科, 武治宇. 基于红外与可见光图像融合的交通标志检测 [J]. *现代电子技术*, 2020, 43(3): 45–49.  
LI Shuhan, XU Hongke, WU Zhiyu. Traffic sign detection based on infrared and visible image fusion[J]. *Modern electronics technique*, 2020, 43(3): 45–49.
- [5] 马瞳宇, 崔静, 储鼎. 基于 WebGL 的实景三维场景与视频监控图像融合技术研究 [J]. *测绘与空间地理信息*, 2020, 43(S1): 80–83.  
MA Tongyu, CUI Jing, CHU Ding. Research on fusion technology of real 3D scene and video surveillance image based on WebGL[J]. *Geomatics & spatial information technology*, 2020, 43(S1): 80–83.
- [6] BAVIRISETTI D P, DHULI R. Fusion of infrared and visible sensor images based on anisotropic diffusion and Karhunen-loeve transform[J]. *IEEE sensors journal*, 2016, 16(1): 203–209.
- [7] Kumar B K S. Image fusion based on pixel significance using cross bilateral filter[J]. *Signal, image and video processing*, 2015, 9(5): 1193–1204.
- [8] MA Jinlei, ZHOU Zhiqiang, WANG Bo, 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.
- [9] LIU Yu, WANG Zengfu. Simultaneous image fusion and denoising with adaptive sparse representation[J]. *IET image processing*, 2015, 9(5): 347–357.
- [10] Burt P, Adelson E. The laplacian pyramid as a compact image code[J]. *IEEE transactions on communications*, 1983, 31(4): 532–540.
- [11] LI Shutao, YANG Bin, HU Jianwen. Performance comparison of different multi-resolution transforms for image fusion[J]. *Information fusion*, 2011, 12(2): 74–84.
- [12] LI Shutao, YIN Haitao, FANG Leyuan. Group-sparse representation with dictionary learning for medical image denoising and fusion[J]. *IEEE transactions on bio-medical engineering*, 2012, 59(12): 3450–3459.
- [13] MA Jiayi, MA Yong, LI Chang. Infrared and visible image fusion methods and applications: a survey[J]. *Information fusion*, 2019, 45: 153–178.
- [14] 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 2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2017: 4724–4732.
- [15] MA Jiayi, YU Wei, LIANG Pengwei, et al. FusionGAN: a generative adversarial network for infrared and visible image fusion[J]. *Information fusion*, 2019, 48: 11–26.
- [16] LI Hui, WU Xiaojun, KITTLER J. Infrared and visible image fusion using a deep learning framework[C]//Proceedings of 24th International Conference on Pattern Recognition (ICPR). Beijing: IEEE, 2018: 2705–2710.
- [17] LI Hui, WU Xiaojun. DenseFuse: A fusion approach to infrared and visible images[J]. *IEEE transactions on image processing*, 2019, 28(5): 2614–2623.
- [18] XU Han, MA Jiayi, JIANG Junjun, et al. U2Fusion: a unified unsupervised image fusion network[J]. *IEEE transactions on pattern analysis and machine intelli-*

- gence, 2022, 44(1): 502–518.
- [19] HUANG Gao, LIU Zhuang, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Honolulu: IEEE, 2017: 2261–2269.
- [20] 刘万军, 佟畅, 曲海成. 空洞卷积与注意力融合的对抗式图像阴影去除算法 [J]. 智能系统学报, 2021, 16(6): 1081–1089.
- LIU Wanjun, TONG Chang, QU Haicheng. An antagonistic image shadow removal algorithm based on dilated convolution and attention mechanism[J]. CAAI transactions on intelligent systems, 2021, 16(6): 1081–1089.
- [21] Alexander T. TNO image fusion dataset[EB/OL]. (2018-09-15)[2021-04-02]<https://flgshare.com/articles/TNO-Image-Fusion-Dataset/1008029>.
- [22] HWANG S, PARK J, KIM N, et al. Multispectral pedestrian detection: benchmark dataset and baseline[C]//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston: IEEE, 2015: 1037–1045.
- [23] ROBERTS J W, VAN AARDT J, 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.
- [24] QU Guihong, ZHANG Dali, YAN Pingfan. Information measure for performance of image fusion[J]. *Electronics letters*, 2002, 38(7): 313–315.
- [25] 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.

#### 作者简介:



曲海成, 副教授, 博士, 辽宁工程技术大学软件学院副院长, 主要研究方向为图像与智能信息处理。主持省自然科学基金项目 1 项、省教育厅面上项目 2 项。发表学术论文 60 余篇。



王宇萍, 硕士研究生, 主要研究方向为图像与智能信息处理。



谢梦婷, 本科生, 主要研究方向为图像处理技术研究。