

# 基于改进 WGAN-GP 的多波段图像 同步超分与融合方法

田嵩旺, 蔺素珍\*, 雷海卫\*\*, 李大威, 王丽芳

中北大学大数据学院, 山西 太原 030051

**摘要** 针对低分辨率源图像的融合结果质量低下不利于后续目标提取的问题, 提出一种基于梯度惩罚 Wasserstein 生成对抗网络(WGAN-GP)的多波段图像同步超分与融合方法。首先, 基于双三次插值法将多波段低分辨率源图像分别放大至目标尺寸; 其次, 将该放大结果输入特征提取(编码)网络分别提取特征并在高层特征空间进行组合; 然后, 通过解码网络重构出初步融合图像; 最后, 经过生成器和判别器的动态博弈得到高分辨率的融合图像。实验结果表明, 所提方法不仅可以同步实现多波段图像的超分和融合, 而且融合图像的信息量、清晰度和视觉质量明显高于其他代表性方法。

**关键词** 图像处理; 图像融合; 图像超分辨率; 多波段图像; 生成对抗网络

中图分类号 TP391

文献标志码 A

doi: 10.3788/AOS202040.2010001

## Multi-Band Image Synchronous Super-Resolution and Fusion Method Based on Improved WGAN-GP

Tian Songwang, Lin Suzhen\*, Lei Haiwei\*\*, Li Dawei, Wang Lifang

College of Data Science and Technology, North University of China, Taiyuan, Shanxi 030051, China

**Abstract** Aiming at the problem that the fused results of low resolution source images are not good for the subsequent target extraction, a multi-band image synchronous super-resolution and fusion method based on Wasserstein generative adversarial network with gradient penalty (WGAN-GP) is proposed. Firstly, the multi-band low-resolution source images are enlarged to the target size respectively based on the bicubic interpolation method. Secondly, the enlarged results are input to a feature extraction (encoding) network to extract features respectively and combine them in a high-level feature space. Then, the initial fused images are reconstructed by decoding network. Finally, a high-resolution fused image is obtained through a dynamic game between the generator and the discriminator. The experimental results show that the proposed method can not only achieve multi-band images super-resolution and fusion simultaneously, but also the information amount, clarity, and visual quality of the fused images are significantly higher than other representative methods.

**Key words** image processing; image fusion; image super-resolution; multi-band image; generative adversarial network

**OCIS codes** 100.2960; 350.2660; 100.6640; 200.4260

## 1 引 言

多波段图像融合旨在通过提取同一场景不同波段图像中的互补信息组合成一幅信息更丰富、更全面的图像, 用于后续目标提取和决策制定<sup>[1]</sup>。长期以来, 基于多尺度变换<sup>[2]</sup>、稀疏表示<sup>[3]</sup>和显著特征提

取<sup>[4]</sup>等的图像融合方法已取得丰硕成果, 但此类方法大多计算复杂且需要根据先验知识确定分解方法和融合规则, 具有较大局限性。新近基于数据驱动的图像融合方法取得较大进展, 典型的有文献[5]利用卷积神经网络(CNN)获得权重图整合可见光图像和红外图像的像素活动信息, 取得较好效果; 文献

收稿日期: 2020-04-09; 修回日期: 2020-05-17; 录用日期: 2020-07-03

基金项目: 山西省应用基础研究项目(201701D121062, 201901D111151)、山西省研究生创新项目(2019SY446, 2019BY115, 2020SY382)、中北大学第十六届研究生科技立项(20191636)

\* E-mail: lsz@nuc.edu.cn; \*\* E-mail: lhw0312@nuc.edu.cn

[6]基于生成对抗网络(GAN)融合红外和可见光图像,结果可同时具有红外图像热辐射信息和可见光图像纹理细节信息;文献[7]在编码网络中设计了卷积层、融合层和密集块,使得每层输出都与其他层相连,使得融合结果中保留更多源图像信息。鉴于这些方法只能融合双幅图像,不能满足多波段/模态探测需要(“多”通常指三幅及以上),本课题组基于深度堆叠卷积神经网络和集成模型-数据驱动方法分别实现了多波段图像自适应融合<sup>[8-9]</sup>,虽然效果显著,但在低分辨率成像时,效果不尽人意。

目前,对低分辨率成像往往通过超分辨率(SR)重建来提升像质<sup>[10]</sup>。当前基于人工神经网络的SR方法可分为单图像超分辨率(SISR)和多对比度超分辨率(MCSR)。在SISR中,常用CNN<sup>[11]</sup>、残差密集网络(RDN)<sup>[12]</sup>和GAN进行图像超分辨率重建<sup>[13-14]</sup>。MCSR多用于医学图像处理,通常是通过提取高分辨率参考图像中相关信息指导低分辨率图像中感兴趣细节的恢复,如文献[15]通过建立不同对比度图像之间的梯度关系模型,从输入的低分辨率图像恢复出高分辨率图像;文献[16]用不同对比度图像之间的局部权重相似度实现图像超分辨率。在此基础上,图像融合领域主要通过先对低分辨率源图像SR再对其融合;或先融合低分辨率源图像再对结果进行SR两种方式来处理。不过,这两种方式会丢失细节信息或引入噪声<sup>[17]</sup>。

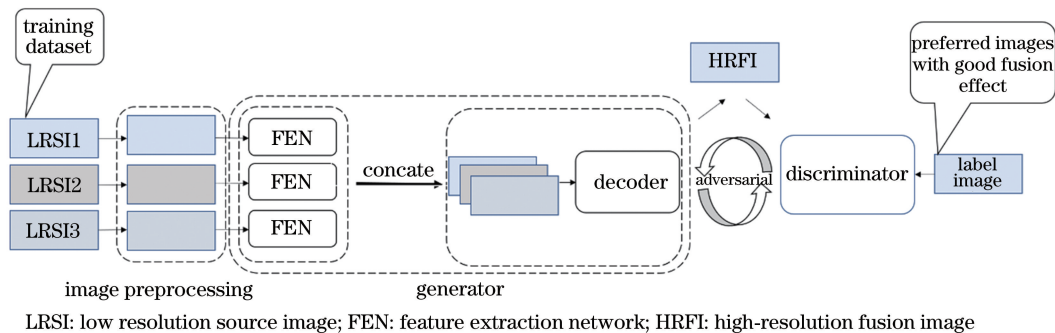


图1 本文方法的整体框架图

Fig. 1 Overall framework of the proposed method

## 2.2 图像同步超分与融合网络实现

### 2.2.1 生成器网络结构及实现

充分利用编码-解码网络能够改善图像质量的优点<sup>[18-19]</sup>,通过编码网络提取特征、解码网络重构图像。其中,编码-解码网络采用文献[10]的结构,生成器架构设计如图2所示。

本特征提取(编码)网络由8个连续的卷积层构成,每个卷积层后跟随修正线性单元(ReLU),卷积

本文提出基于梯度惩罚的Wasserstein生成对抗网络(WGAN-GP)多波段图像同步超分与融合方法,利用WGAN-GP训练稳定、收敛速度快以及可避免梯度爆炸的优势,设计合适的生成器、判别器网络架构和损失函数,训练一个将低分辨率多波段源图像生成高分辨率融合结果的模型,使得图像融合和图像超分辨率同步实现。本文的主要创新点有:1)提出一种同步实现多波段图像融合和图像超分辨率的端到端网络结构,模型训练好后仅需在输入端输入待融合图像,即可得到超分融合结果,不需要传统方法中分步处理;2)提出在高层特征空间组合多波段图像特征信息得到较好融合增强结果;3)复合使用对抗损失、内容损失和感知损失,确保生成图像可以在实现超分的同时达到较好融合效果。

## 2 基于WGAN-GP的多波段图像超分融合方法

### 2.1 整体框架

本文首先用双三次插值法将源图像放大至目标尺寸,将放大后的低分辨率图像输入特征提取(编码)网络,将提取的多波段图像特征在高层特征空间组合;然后通过解码网络重构融合图像;最后将生成器输出的融合图像和标签图像送入判别器分类识别,经过生成器和判别器的动态博弈和迭代更新,得到超分融合结果。总体框架如图1所示。

滤波器的数量分别为32、32、64、64、128、128、256和256,每个卷积层使用 $3 \times 3$ 卷积核,步长为1,填充为0。特征提取网络分别提取各波段特征后,利用合并连接层将所提取的各波段信息在高层特征空间组合,送入解码网络逐层扩展信息并将特征图转为图像。解码网络同样由8个连续卷积层构成,卷积滤波器的数量分别为256、128、128、64、64、32、32和1,每个卷积层使用 $3 \times 3$ 卷积核,步长为1,填充为0。为

保存更多源图像信息,在特征提取网络和解码网络间分别设计三个跳连接(见图2中9条黑色线条)。

### 2.2.2 判别器网络结构及实现

判别网络如图3所示,包括6个卷积层、3个最大池化层和两个全连接层,每个卷积层后跟随Leaky ReLU。卷积滤波器个数分别为64、64、128、

128、256和256,每个卷积层用 $3 \times 3$ 卷积核,步长为1,填充为1。在两个全连接层中,神经元数量分别为128和1。

### 2.2.3 损失函数构建

本文在原有对抗损失 $L_{adv}$ 的基础上,增加内容损失 $L_{con}$ 和感知损失 $L_{per}$ 。

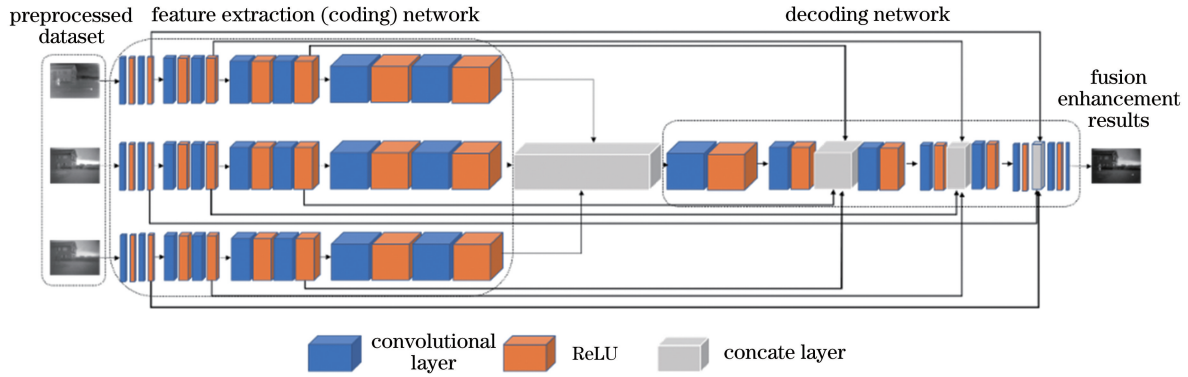


图2 生成器网络架构

Fig. 2 Generator network architecture

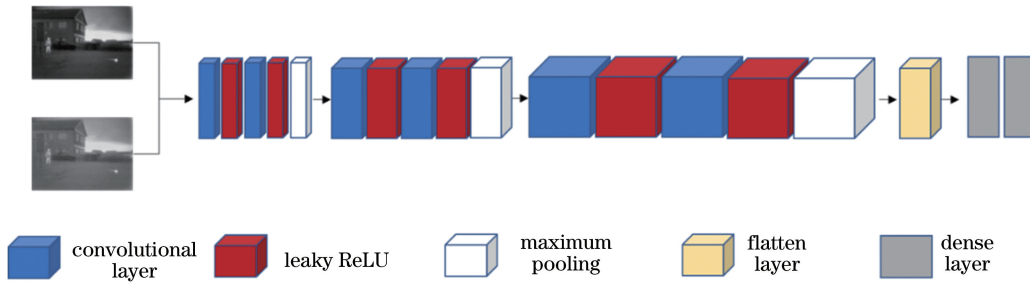


图3 判别器网络架构

Fig. 3 Discriminator network architecture

对抗损失可表示为<sup>[20]</sup>

$$L_{adv} = -E_{x \sim I_{LR}} \{D[G(x)]\}, \quad (1)$$

式中: $G$ 为生成器; $D$ 为判别器; $E$ 为期望; $x$ 为输入样本; $I_{LR}$ 为低分辨率源图像。

考虑到在融合结果中既要保留可见光图像细节信息也要保留红外图像热目标,细节信息可以用梯度表示,而热辐射信息可以用像素强度表示,所以,本文希望融合图像 $I_{SRF}$ 与标签图像 $I_{HRF}$ 有相似的梯度和像素强度。均方误差(MSE)可以表征像素的强度损失,这里与梯度损失一起来评估内容损失。内容损失 $L_{con}$ 定义为

$$L_{con} = E_{x \sim I_{LR}, y \sim I_{HRF}} [\|G(x) - y\|^2 + \lambda \|\nabla G(x) - \nabla y\|^2], \quad (2)$$

式中: $\|\cdot\|$ 为F范数; $y$ 为输入的标签样本; $I_{HRF}$ 为标签图像; $\nabla$ 为梯度; $\lambda$ 为权重系数。

为防止内容损失造成生成器输出图像过于平滑和清晰度低,在损失函数中增加感知损失 $L_{per}$ 来评

判高层特征空间中测量图像与标签图像的相似性。使用预训练的VGG19模型第2、4、8和12个卷积层的特征图(每个卷积层等价贡献)来计算感知损失。 $L_{per}$ 由生成器输出和标签图像之间特征图的均方误差获得,表示为

$$L_{per} = E_{x \sim I_{LR}, y \sim I_{HRF}} \|\varphi[G(x)] - \varphi(y)\|^2, \quad (3)$$

式中: $\varphi$ 为特征提取器VGG19。

综上,生成器的目标损失函数 $L_G$ 为

$$\min_{\theta_G} L_G = L_{adv} + \lambda_1 L_{con} + \lambda_2 L_{per}, \quad (4)$$

式中: $\theta_G$ 为生成器的训练参数; $\lambda_1$ 、 $\lambda_2$ 分别为内容损失 $L_{con}$ 和感知损失 $L_{per}$ 的权重。

判别器的损失函数 $L_D$ 与原WGAN-GP模型相同,可表示为

$$\max_{\theta_D} L_D = E_{x \sim I_{LR}} \{D[G(x)]\} + E_{\hat{x} \sim I_{HRF}} [D(\hat{x})] + \lambda_3 E_{\tilde{x} \sim (I_{LR}, I_{HRF})} \|\nabla_{\tilde{x}} [D(\tilde{x})] - 1\|^2, \quad (5)$$

式中: $\tilde{x}$ 为从标签数据分布 $\hat{x}$ 和生成器输入数据分

布  $x$  采样的成对点之间的直线上的随机采样,即  $\tilde{x} = \alpha x + (1-\alpha)\hat{x}$ , 其中  $\alpha \in [0, 1]$ ;  $\theta_D$  为判别器的训练参数,前两项执行 Wasserstein 距离估计,最后一项是网络正则化的梯度惩罚项;  $\lambda_3$  为梯度惩罚项的权重。

### 3 实验结果与分析

#### 3.1 实验参数设置

由于图像融合尚无金标准,根据文献[21],本文将公开 TNO 图像融合数据集 (<https://doi.org/10.6084/m9.figshare.1008029.v1>) 中 20 组图像分块处理并默认为高分辨率源图像,并将其采用现有性能良好的多种方法进行融合,然后将融合结果利用多个客观指标进行评价,优选出高质量融合结果作为标签图像。为保证训练图像为低分辨率,将分块后的图像长和宽各缩小为原图像的 1/4,作为训练集。将其中的 80% 用于训练,20% 用于验证。测试图像由 TNO 图像融合数据集剩余多波段图像构成(长和宽各缩小为原图像的 1/4)。

本实验所选择的学习率为 0.0002, batch 为 16, 训练过程持续 100 轮。经过实验,生成器损失函数中参数  $\lambda$ 、 $\lambda_1$  和  $\lambda_2$  分别设置为 5、1 和 0.1。判别器损失函数中参数  $\lambda_3$  参考文献[20], 设置为 10。选用 Adam 优化器训练神经网络。本实验所有深度学习方法均在 NVIDIA TITAN X GPU 服务器上使用 tensorflow 1.12.0 和 python 3.6.2 实现,其余对比实验在 Intel(R) Core(TM) i5-9300H CPU 上使用 Matlab R2018a 实现。

#### 3.2 主要影响因素分析

##### 3.2.1 损失函数影响

为分析不同损失函数的影响,本节对“对抗损失+内容损失”、“对抗损失+感知损失”和“对抗损失+感知损失+内容损失”三种情况进行比较,从图 4 可以看出,采用“对抗损失+内容损失”时图像较模糊,整体偏暗,虽然对比度较高,但热目标不明显;使用“对抗损失+感知损失”时整体图像呈灰度,对比度较低但清晰度明显提高;“对抗损失+感知损失+内容损失”图像更符合人眼视觉特性,边缘清晰且细节丰富。

##### 3.2.2 网络结构影响

通常文献在低层图像空间合并信息<sup>[6]</sup>,其结果如图 5(a)所示。本文提出在高层特征空间组合信息,结果如图 5(b)所示,二者其余网络结构相同。可以看出,图 5(b)结果更清晰,边缘更明显。说明



图 4 损失函数对结果的影响。(a)对抗损失+内容损失;(b)对抗损失+感知损失;(c)对抗损失+感知损失+内容损失

Fig. 4 Effect of loss function on results. (a) Adversarial loss + content loss; (b) adversarial loss + perceptual loss; (c) adversarial loss + content loss + perceptual loss

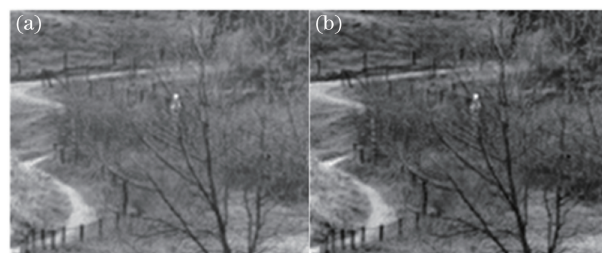


图 5 网络结构对结果的影响。(a)低层图像空间组合信息结果;(b)高层特征空间组合信息结果

Fig. 5 Influence of network structure on results. (a) Low-level image spatial combination information result; (b) high-level feature space combined information results

在高层特征空间组合信息可以有效改善图像的清晰度。

#### 3.3 融合结果分析

由于尚未检索到图像融合和超分辨率同步实现的报道,本节使用文献中常用的“先超分再融合”和“先融合再超分”两种方式进行实验,选择二者中较好结果与本文方法进行对比。其中,对比实验中图像超分辨率方法选用较好的残差密集网络(RDN)<sup>[12,22-23]</sup>。融合方法选用效果较好的 CNN<sup>[5]</sup>、GAN<sup>[6]</sup>、Densefuse<sup>[7]</sup> 和深度门卷积神经网络(DGCNN)<sup>[8]</sup>。由于 CNN、Densefuse 和 GAN 均用于双波段图像融合,这里通过两两序贯相融方式转换为三波段图像融合。

图 6 从测试图像中选择 6 组典型的融合图像进行比较。为便于观察,图 6 将低分辨率图像放大至与结果图像相同尺寸。可以看出,所有方法都得到了较好的超分融合结果。其中,GAN 能较好保留目标,图像整体对比度较高,但图像失真严重、缺乏纹理细节,边缘较模糊;CNN、Densefuse 和 DGCNN 达到较好的融合效果,但 CNN 结果的某些场景存

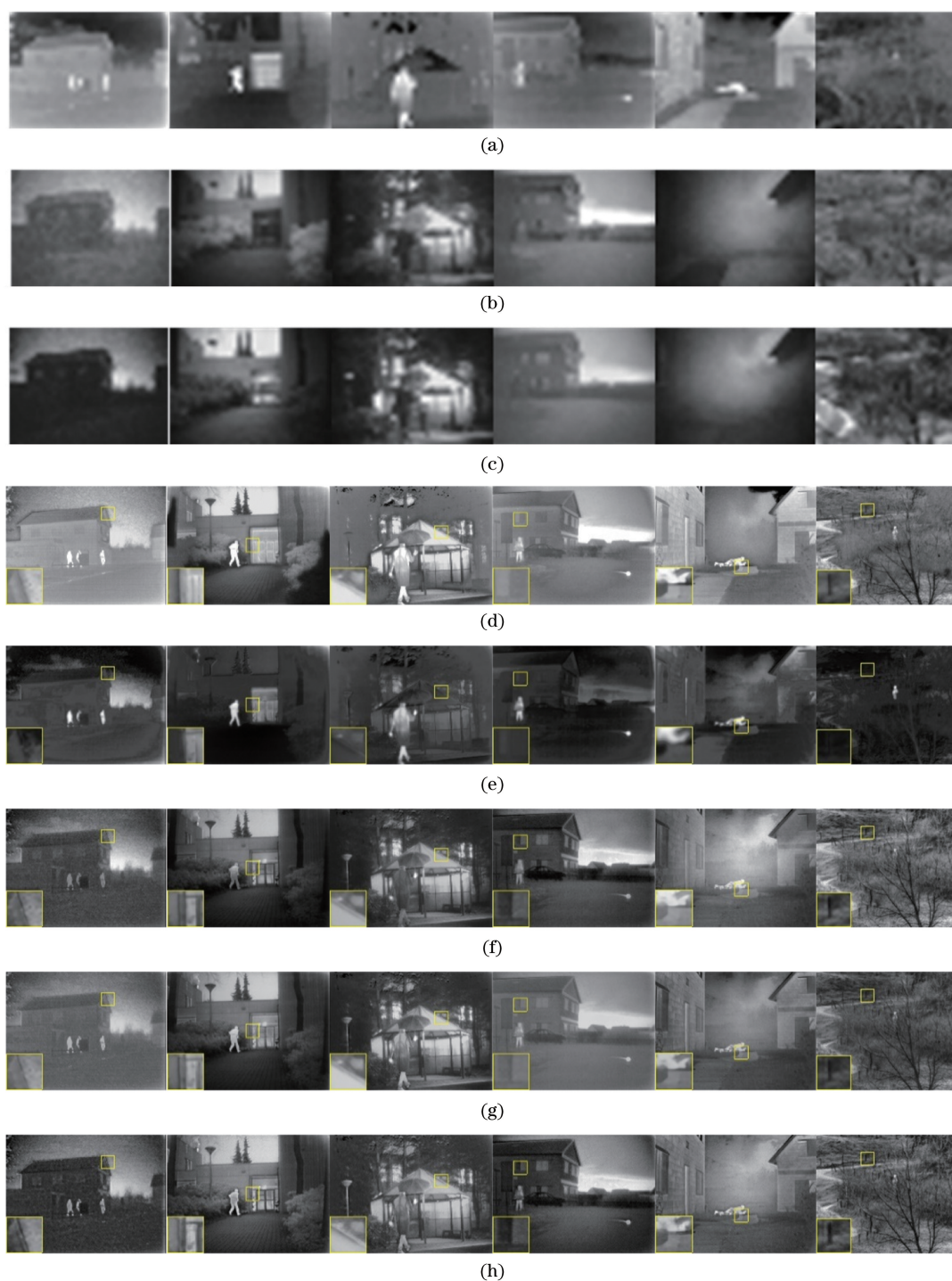


图 6 融合结果。(a)低分辨率红外长波图像;(b)低分辨率红外短波图像;(c)低分辨率可见光图像;(d) CNN;(e) GAN;(f) Densefuse;(g) DGCNN;(h)本文方法

Fig. 6 Fusion results. (a) Low resolution infrared long wave images; (b) low-resolution infrared short wave images; (c) low resolution visible images; (d) CNN; (e) GAN; (f) Densefuse; (g) DGCNN; (h) proposed method

在缺失部分,DGCNN 和 Densefuse 图像较平滑,清晰度不高;本文方法更加清晰且边缘明显,更符合人的视觉特性(图中左下角为小方框局部放大图)。

由于像质评价尚无金标准,因此本文用信息熵(IE)、互信息(MI)、平均梯度(AG)、对比度(C)、峰

值信噪比(PSNR)、相关系数(CC)、结构相似性(SSIM)、视觉信息保真度(VIFF)、空间频率(SF)和边缘信息保留值(EIPV)这 10 种指标对源图像和最终融合结果进行客观评价,归一化每组图像所得指标值并用折线图表示(图 7)。

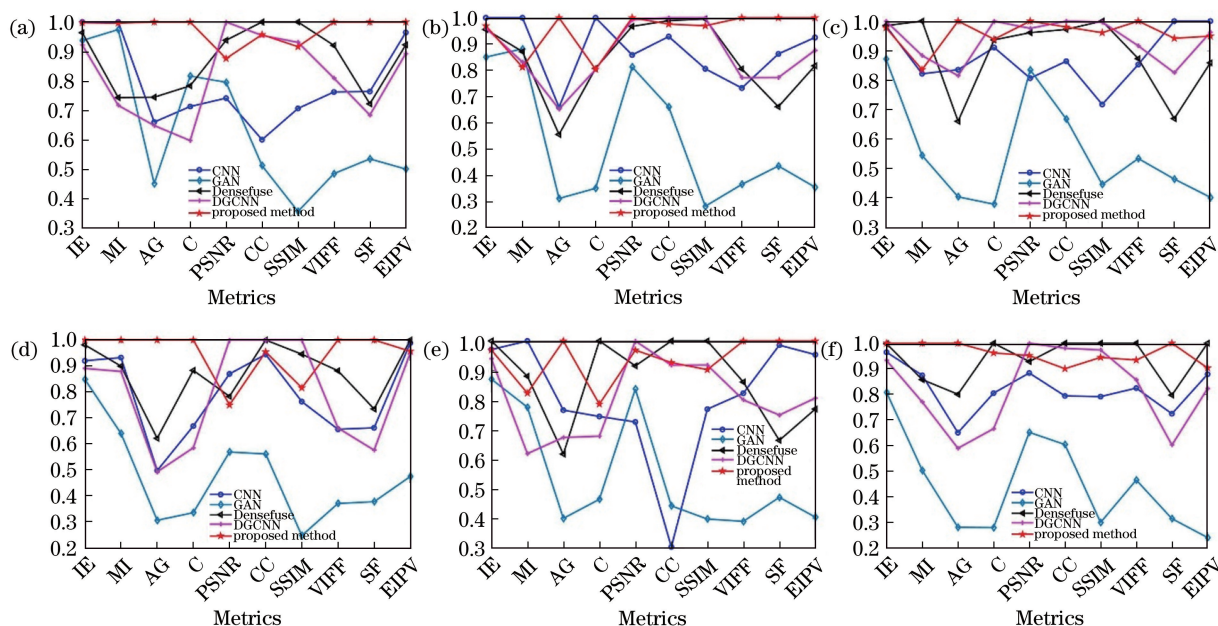


图 7 客观评价结果。(a)第一组结果的评价指标;(b)第二组结果的评价指标;(c)第三组结果的评价指标;  
(d)第四组结果的评价指标;(e)第五组结果的评价指标;(f)第六组结果的评价指标

Fig. 7 Objective evaluation results. (a) Evaluation metrics for the first fused results; (b) evaluation metrics for the second fused results; (c) evaluation metrics for the third fused results; (d) evaluation metrics for the fourth fused results; (e) evaluation metrics for the fifth fused results; (f) evaluation metrics for the sixth fused results

由图 7 可知, GAN 的指标位于折线图底部, 处于较低水平; CNN 和 Densfuse 的 C 值较高, 但 CNN 的 AG、PSNR、CC 和 SSIM 值较小; DGCNN 和 Densfuse 较为稳定, 但 AG 和 EIPV 值较小。本文方法指标均在折线图中上位置, AG、VIFF、SF 和 EIPV 值相对较高, 且 AG、VIFF 和 SF 值高出较多, 说明本文方法具有较好的梯度、边缘保留较好、图像整体活跃度较高、更加符合人眼视觉特性等, 能够有效改善结果清晰度。

### 3.4 附加实验

使用 TRICLOBS 动态多波段图像数据集<sup>[24]</sup>进行测试, 此数据集包括三个视频图像序列(即 TRI\_A、TRI\_B 和 TRI\_C), 每个视频包含配准好的长波红外(LWIR, 8~14 μm)、近红外(NIR, 700~1000 nm)和可见光(VIS, 400~700 nm)序列, 每隔 3 s 捕获一帧, 其中捕获图像大小为 212×156。在所有捕获图像上进行测试, 鉴于篇幅关系这里随机选择一组图像进行展示, 如图 8 所示(源图像大小为

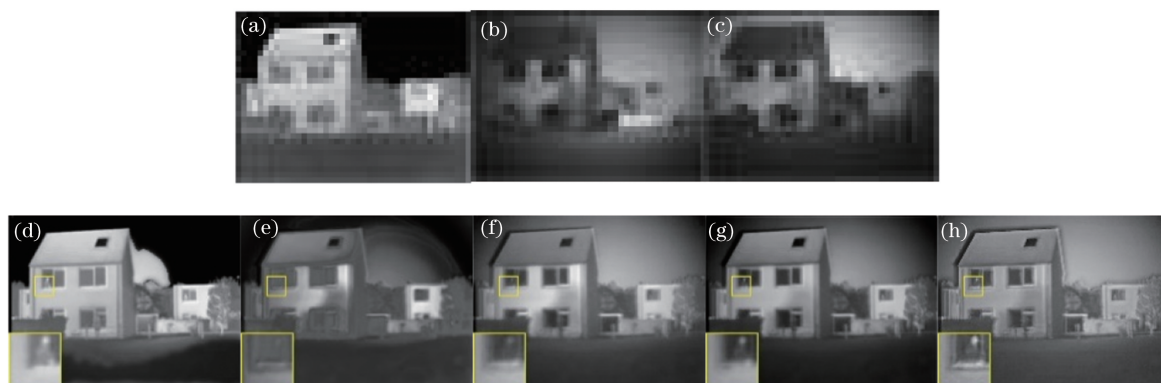


图 8 融合结果展示。(a)低分辨率红外长波图像;(b)低分辨率红外短波图像;(c)低分辨率可见光图像;(d) CNN;  
(e) GAN;(f) Densfuse;(g) DGCNN;(h)本文方法

Fig. 8 Fusion results. (a) Low resolution infrared long wave image; (b) low-resolution infrared short wave image; (c) low resolution visible image; (d) CNN; (e) GAN; (f) Densfuse; (g) DGCNN; (h) proposed method

212×156,增强融合结果图像大小为 848×624)。

由图 8 可知,CNN 目标较清晰,整体效果较好,但融合结果明显存在缺失;GAN 失真较严重且清晰度较低;Densefuse、DGCNN 和本文方法取得较好的融合结果,但 Densefuse 和 DGCNN 图像过于平滑,Densefuse 图像整体偏亮,DGCNN 整体偏暗,本

文方法图像亮度介于二者之间且图像边缘更清晰。

由于该数据集缺乏高分辨率源图像,选用 IE、AG、C、SF、边缘强度(EI)、图像清晰度(IC)和灰度均值(GM)7 个无参考图像指标客观评价,对所有指标取平均值,结果如表 1 所示。可以看出,除 IE 和 C 之外,本文方法的指标均最好,优势显而易见。

表 1 客观评价指标值

Table 1 Objective evaluation index values

Method	IE	AG	C	SF	EI	IC	GM
CNN	7.01	3.03	<b>70.16</b>	9.25	32.64	3.02	98.98
GAN	7.09	1.73	27.85	4.63	17.71	1.72	86.73
Densefuse	<b>7.59</b>	2.13	43.37	5.82	22.42	2.12	97.45
DGCNN	7.37	2.52	49.31	6.94	25.25	2.51	77.43
Proposed method	7.44	<b>4.24</b>	37.02	<b>10.34</b>	<b>38.95</b>	<b>4.22</b>	<b>107.52</b>

## 4 结 论

本文提出一种能将多波段图像融合和超分辨率同步实现的多波段图像超分辨率融合方法。针对图像融合过程中图像分辨率低导致融合结果质量差的问题,设计了一种基于 WGAN-GP 的网络结构,将多波段图像特征在高层特征空间组合进行超分;此外,复合使用对抗损失、内容损失和感知损失,提高网络性能。经过实验对比,证明本文方法获得的图像清晰度更高、边缘强度更大且更符合人眼视觉特性,整体效果更好,对提高低分辨率源图像的融合效果有重要作用。需要指出的是,本文方法结果中对比度指标未达到理想效果,下一步将重点改进此问题。

## 参 考 文 献

- [1] Zhou X L, Jiang Z T. Infrared and visible image fusion combining pulse-coupled neural network and guided filtering [J]. Acta Optica Sinica, 2019, 39 (11): 1110003.  
周晓玲, 江泽涛. 结合脉冲耦合神经网络与引导滤波的红外与可见光图像融合 [J]. 光学学报, 2019, 39 (11): 1110003.
- [2] Yan H B, Li Z M. Infrared and visual image fusion based on multi-scale feature decomposition [J]. Optik, 2020, 203: 163900.
- [3] Wang W Q, Gao Y D, Liu H, et al. Infrared and visible image sequence fusion via low-rank and sparse representation [J]. Journal of Xi'an University of Technology, 2019, 35(3): 276-283.  
王文卿, 高钰迪, 刘涵, 等. 基于低秩稀疏表示的红外与可见光图像序列融合方法 [J]. 西安理工大学学
- [4] Zhang L B, Zhang J. A new adaptive fusion method based on saliency analysis for remote sensing images [J]. Chinese Journal of Lasers, 2015, 42 (1): 0114001.  
张立保, 章珏. 基于显著性分析的自适应遥感图像融合 [J]. 中国激光, 2015, 42(1): 0114001.
- [5] Liu Y, Chen X, Cheng J, et al. Infrared and visible image fusion with convolutional neural networks [J]. International Journal of Wavelets, Multiresolution and Information Processing, 2017, 16(3): 1850018.
- [6] 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.
- [7] Li H, Wu X J. DenseFuse: a fusion approach to infrared and visible images [J]. IEEE Transactions on Image Processing, 2019, 28(5): 2614-2623.
- [8] Lin S Z, Han Z. Images fusion based on deep stack convolutional neural network [J]. Chinese Journal of Computers, 2017, 40(11): 2506-2518.  
蔺素珍, 韩泽. 基于深度堆叠卷积神经网络的图像融合 [J]. 计算机学报, 2017, 40(11): 2506-2518.
- [9] Lin S Z, Han Z, Li D W, et al. Integrating model- and data-driven methods for synchronous adaptive multi-band image fusion [J]. Information Fusion, 2020, 54: 145-160.
- [10] Lyu Q, Shan H M, Wang G. Multi-contrast super-resolution MRI through a progressive network [EB/OL]. (2019-08-05) [2020-05-14]. <http://doi-org-s.vpn.nuc1941.top:8118/10.1109/TMI.2020.2974858>.
- [11] Dong C, Loy C C, He K M, et al. Image super-resolution using deep convolutional networks [J].

- IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(2): 295-307.
- [12] Zhang Y L, Tian Y P, Kong Y, et al. Residual dense network for image super-resolution[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 2472-2481.
- [13] Sajjadi M S M, Schölkopf B, Hirsch M. EnhanceNet: single image super-resolution through automated texture synthesis [C]//2017 IEEE International Conference on Computer Vision (ICCV). October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 4501-4510.
- [14] Ledig C, Theis L, Huszár F, et al. Photo-realistic single image super-resolution using a generative adversarial network [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 105-114.
- [15] Zheng H, Zeng K, Guo D, et al. Multi-contrast brain MRI image super-resolution with gradient-guided edge enhancement [J]. IEEE Access, 2018, 6: 57856-57867.
- [16] Zheng H, Qu X B, Bai Z J, et al. Multi-contrast brain magnetic resonance image super-resolution using the local weight similarity [J]. BMC Medical Imaging, 2017, 17(1): 1-13.
- [17] Zhong J Y, Yang B, Li Y H, et al. Image fusion and super-resolution with convolutional neural network [M]. Singapore: Springer Singapore, 2016: 78-88.
- [18] Chen L C, Zhu Y K, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [M]. Cham: Springer International Publishing, 2018: 833-851.
- [19] Lei B Y, Huang S, Li R, et al. Segmentation of breast anatomy for automated whole breast ultrasound images with boundary regularized convolutional encoder-decoder network [J]. Neurocomputing, 2018, 321: 178-186.
- [20] Gulrajani I, Ahmed F, Arjovsky M, et al. Improved training of Wasserstein GANs [EB/OL]. (2017-12-25) [2020-05-14]. <https://arxiv.org/abs/1704.00028>.
- [21] Han Z. Research on multi-band image adaptive fusion method based on deep learning [D]. Taiyuan: North University of China, 2018.  
韩泽. 基于深度学习的多波段图像自适应融合方法研究 [D]. 太原: 中北大学, 2018.
- [22] Sun Y W, Li L T, Cong P, et al. Super-resolution method for radiation image based on deep learning [J]. Atomic Energy Science and Technology, 2017, 51(5): 890-895.  
孙跃文, 李立涛, 丛鹏, 等. 基于深度学习的辐射图像超分辨率重建方法 [J]. 原子能科学技术, 2017, 51(5): 890-895.
- [23] Wang W, Zhang T, Wang X. Survey of single image super-resolution based on deep learning [J]. Journal of Chinese Computer Systems, 2019, 40(9): 1891-1896.  
王威, 张彤, 王新. 用于图像超分辨率重构的深度学习方法综述 [J]. 小型微型计算机系统, 2019, 40(9): 1891-1896.
- [24] Toet A, Hogervorst M A, Pinkus A R. The TRICLOBS dynamic multi-band image data set for the development and evaluation of image fusion methods [J]. Plos One, 2016, 11(12): e0165016.