

基于显著性的双鉴别器GAN图像融合算法

谢一博 刘卫国 周顺 李梦晗

Saliency-based dual discriminator GAN image fusion algorithm

XIE Yibo, LIU Weiguo, ZHOU Shun, LI Menghan

引用本文:

谢一博, 刘卫国, 周顺, 等. 基于显著性的双鉴别器GAN图像融合算法[J]. 应用光学, 2024, 45(1): 107–117. DOI: 10.5768/JAO202445.0102005

XIE Yibo, LIU Weiguo, ZHOU Shun, et al. Saliency-based dual discriminator GAN image fusion algorithm[J]. Journal of Applied Optics, 2024, 45(1): 107–117. DOI: 10.5768/JAO202445.0102005

在线阅读 View online: <https://doi.org/10.5768/JAO202445.0102005>

您可能感兴趣的其他文章

Articles you may be interested in

改进VGG网络的多聚焦图像的融合方法

Multi-focus image fusion method based on improved VGG network

应用光学. 2020, 41(3): 500–507 <https://doi.org/10.5768/JAO202041.0302003>

基于对偶学习的图像去雾网络

Image defogging network based on dual learning

应用光学. 2020, 41(1): 94–99 <https://doi.org/10.5768/JAO202041.0102005>

基于无人机载的光电与SAR图像融合技术研究

Fusion method of optical image and SAR based on UAV

应用光学. 2017, 38(2): 174–179 <https://doi.org/10.5768/JAO201738.0201004>

一种基于GAN和自适应迁移学习的样本生成方法

Sample generation method based on GAN and adaptive transfer learning

应用光学. 2020, 41(1): 120–126 <https://doi.org/10.5768/JAO202041.0102009>

单光子探测三维点云与可见光图像融合处理算法研究

Fusion processing algorithm of single-photon detection for three-dimensional point cloud and visible light image

应用光学. 2021, 42(6): 1034–1039 <https://doi.org/10.5768/JAO202142.0602004>

结合深度学习的非下采样剪切波遥感图像融合

Remote sensing image fusion based on deep learning non-subsampled shearlet

应用光学. 2018, 39(5): 655–666 <https://doi.org/10.5768/JAO201839.0502001>



关注微信公众号, 获得更多资讯信息

文章编号: 1002-2082 (2024) 01-0107-11

基于显著性的双鉴别器 GAN 图像融合算法

谢一博, 刘卫国, 周 顺, 李梦晗

(西安工业大学 光电工程学院, 陕西 西安 710021)

摘 要: 针对红外图像与可见光图像在不同场景的特征表达不同的问题, 提出一种基于显著性的双鉴别器生成对抗网络方法, 将红外与可见光的特征信息相融合。区别于传统的生成对抗网络, 该算法采用双鉴别器方式分别鉴别源图像与融合图像中的显著性区域, 以两幅源图像的显著性区域作为鉴别器的输入, 使融合图像保留更多的显著特征; 并将梯度约束引入其损失函数中, 使显著对比度和丰富纹理信息保留在融合图像中。实验结果表明: 本文方法在熵值 (entropy, EN)、平均梯度 (mean gradient, MG)、空间频率 (spatial frequency, SF) 及边缘强度 (edge intensity, EI) 4 个评价指标中均优于其他对比算法。该研究实现了红外图像与可见光图像高效融合, 有望在目标识别等领域中获得应用。

关键词: 图像处理; 生成对抗网络; 图像融合; 显著性区域; 目标识别

中图分类号: TN219; TP391

文献标志码: A

DOI: 10.5768/JAO202445.0102005

Saliency-based dual discriminator GAN image fusion algorithm

XIE Yibo, LIU Weiguo, ZHOU Shun, LI Menghan

(School of Photoelectric Engineering, Xi'an Technological University, Xi'an 710021, China)

Abstract: To address the problem that infrared images and visible images have different feature expressions in different scenes, an saliency-based dual discriminator generative adversarial network method was proposed to fuse the infrared and visible feature information. Different from the traditional generative adversarial network, a dual discriminator approach was adopted to discriminate the saliency regions in the source images and the fusion images respectively in this algorithm, and the saliency regions of the two source images were used as the input of the discriminator so that the fusion image retained more salient features. The gradient constraint was introduced into its loss function so that the salient contrast and rich texture information could retain in the fusion image. The experimental results show that the proposed method outperforms other comparison algorithms in four evaluation indexes: entropy (EN), mean gradient (MG), spatial frequency (SF) and edge intensity (EI). This study achieves efficient fusion of infrared images and visible images, which is expected to gain applications in fields such as target recognition.

Key words: image processing; generative adversarial networks; image fusion; salient regions; target recognition

引言

随着现代科技的快速发展, 传感器技术不断进步与完善, 应用不断复杂化。为了更好地满足不

同领域的需求, 各种类型的传感器不断涌现并应用于不同的环境和场景, 以获取更为全面且准确的信息。红外传感器得到的红外图像通过区分物

收稿日期: 2023-04-03; 修回日期: 2023-07-06

基金项目: 陕西省教育厅科研计划重点项目 (21JY017); 陕西省自然科学基金基础研究计划 (2022JQ-676); 中国博士后科学基金 (2022M712493)

作者简介: 谢一博 (1997—), 男, 博士研究生, 主要从事光电子技术及图像处理研究。E-mail: 13319215096@163.com

通信作者: 刘卫国 (1964—), 男, 博士, 教授, 主要从事光电子技术及电子材料研究。E-mail: wgliu@163.com

体之间的辐射差异来进行目标和背景的区分,在夜间环境下能有较好的工作效果;而可见光传感器获取的可见光图像则可以提供清晰度的纹理细节信息和高分辨率。单一类型的传感器所采集到的信息可能很难对某一问题做出准确的判断。因此将红外图像和可见光图像进行融合处理,可以构造出包含热辐射源和多光谱信息的图像,既解决单一红外传感器背景模糊、细节缺失的问题,又克服了可见光传感器对环境条件的依赖性。由于红外和可见光图像的融合可以大大提高图像信息的解读能力,其已成为图像融合领域的重要组成部分。该融合技术最早应用于军事领域,在 1999 年,国外研发出第一个红外与可见光图像融合的夜视系统,用于在夜晚探测隐藏目标^[1]。随后,各国相继研制出其他彩色图像融合系统、多功能望远镜、夜视镜等设备。随着红外与可见光融合技术的进步,当前该技术在夜间辅助驾驶、社区安防监控、目标检测等实际应用中具有实际价值^[2]。

典型的传统融合方法包括基于多尺度变换方法^[3]、基于稀疏表示方法^[4]、基于子空间方法^[5]、基于显著性方法^[6]、基于总变异方法^[7]等。然而,这些方法的局限性十分明显,传统方法没有考虑不同传感器图像的特征差异,而是采用相同的变换对不同的源图像提取特征,导致提取的特征表达性较差,因此基于深度学习的图像融合方法应运而生并成为学者们研究的热点^[8]。在红外与可见光图像融合方面,基于深度学习的融合方法主要分为三大类:基于自编码器的图像融合^[9]、基于卷积神经网络的图像融合^[10]以及基于生成对抗网络(generative adversarial networks, GAN)的图像融合^[11]。2015 年, HUANG W 等^[12]通过构建堆栈自编码网络保留了全色图像的细节特征和多光谱图像的光谱特性。后来, LIU 等提出了一个基于卷积稀疏表示的融合框架,该方法有效利用图像的稀疏性特征,便于表征融合图像。随着研究的深入, Li 等^[1]基于 VGG-19 构建深度学习框架对图像分解后的细节部分进行特征提取与融合。

目前红外与可见光图像融合中最常用的方法为基于 GAN 网络的方法。2019 年, MA J Y 首次将 GAN 网络应用于图像融合,提出了基于 GAN 网络的无监督融合模型(FusionGAN)^[13],但存在细节丢失和目标边缘模糊的问题。2020 年, MA J Y 团队又提出了一种基于双鉴别器生成对抗网络的红外与可见光图像融合方法(DDcGAN)^[14],该算法

将红外与可见光源图像分别输入两个鉴别器,生成器与鉴别器对抗博弈能更好地保留红外和可见光图像的信息,使得到的融合图像具有更好的纹理细节和目标热辐射信息,但双鉴别器会忽略掉高层语义信息。2021 年 HOU J L 等人^[15]提出了基于语义分割的红外与可见光图像融合方法(SSGAN),该网络既考虑了红外和可见光图像的底层特征,又考虑了红外和可见光图像的高层语义信息。

由于基于 GAN 的红外与可见光图像融合方法在对抗训练过程中会丢失部分细节信息,因此,本文对 GAN 的网络结构以及损失函数进行优化,提出了更加适用于红外与可见光图像融合的网络模型。通过以两幅源图像的显著性区域作为鉴别器输入判定融合图像,从而使得融合图像保留原始图像更多显著特征;并借助添加梯度约束这一损失函数,保存显著对比度和丰富纹理,防止高对比度目标区域边缘扩散造成的背景纹理污染。

1 生成对抗网络模型

2014 年, GOODFELLOW I 等学者提出了由生成模型 G 和鉴别模型 D 组成的生成对抗网络 GAN^[16]。该模型是两个网络相互对抗博弈的过程,通过不断地反向传播更新参数,能提高生成器的生成能力和鉴别器的判别准确度,从而在数据中学习并生成新的数据。基于 GAN 的图像融合框架如图 1 所示。

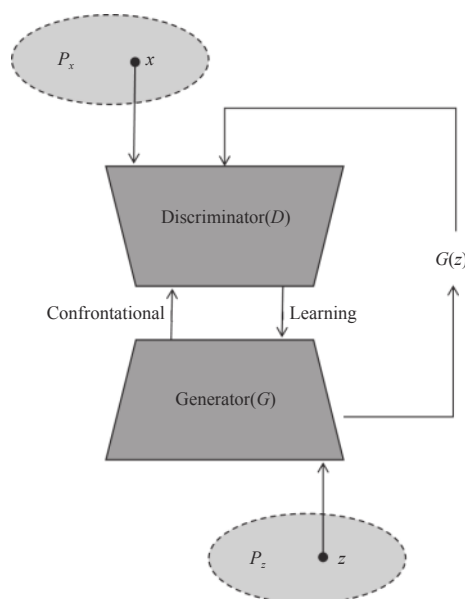


图 1 基于生成对抗网络的结构框架图

Fig. 1 Structural framework diagram based on generative adversarial networks

从数学的角度来看, 为了学习生成器在数据 x 上的分布 P_G , 首先对输入噪声变量定义一个先验, 然后对数据空间的映射表示为 $G(z; \theta_g)$, 其中 G 是一个可微分的函数, 由参数为 θ_g 的多层感知器表示。再定义第 2 个多层感知器 $D(x; \theta_d)$, 输出一个标量, 表明数据 x 来自于输入数据而非 P_G 。通过训练 D , 以最大限度地提高为训练实例和样本分配正确标签的概率; 同时训练 G , 使其最小化 $\log(1-D(G(z)))^{[16]}$ 。生成器 G 和鉴别器 D 之间进行一场对抗性训练, 此过程表示如下:

$$\min_G \max_D V_{\text{GAN}}(G, D) = E_{x \sim P_{\text{data}}(x)} [\log D(x)] + E_{z \sim P_z(z)} [\log(1 - D(G(z)))] \quad (1)$$

基于生成对抗网络的强大生成能力, 一种基于 GAN 网络的无监督融合模型 (FusionGAN)^[13] 被提出。在该模型中, 红外与可见光图像在生成器的指导下融合成一幅图像, 鉴别器判定融合图像来自真实图像的概率, 直到鉴别器无法区分融合图像与真实图像时博弈结束, 此时融合规则下得到的融合图像为最佳图像。

FusionGAN 网络模型首次运用 GAN 去解决图像融合问题, 尽管该网络可以避免手动设计融合策略, 通过引入 GAN 并生成具有丰富细节信息的融合图像, 但仍然存在细节丢失和目标边缘模糊

的问题。为了进一步提高融合图像的质量, 该问题在当前亟须解决。

2 本文方法

2.1 网络总体结构

针对单鉴别器无法同时获取两幅源图像信息的问题, 本文设计了一个基于显著性的双鉴别器 GAN 红外与可见光图像融合方法, 模型框架如图 2 所示。在内容损失和感知损失的指导下, 生成器将生成一张信息丰富且细节完整的融合图像。鉴别器的作用是不断迫使生成器学习, 直到能够生成与鉴别器输入相似的图像为止, 因此, 为了使融合图像保留两幅源图像的显著区域, 将构建两幅源图像的视觉显著图 (即 IR_SM、VIS_SM) 作为两个鉴别器的输入。鉴别器分别鉴别源图像视觉显著图和融合图像, 指导生成器生成高质量的包含源图像的细节信息和显著区域的融合图像。生成器与鉴别器进行不断对抗性训练, 当鉴别器无法区分融合图像与可见光图像时, 判定融合图像中的细节信息分布与可见光图像中的细节信息分布近似, 最终获得一幅包含红外图像辐射信息及可见光图像纹理细节的融合图像。

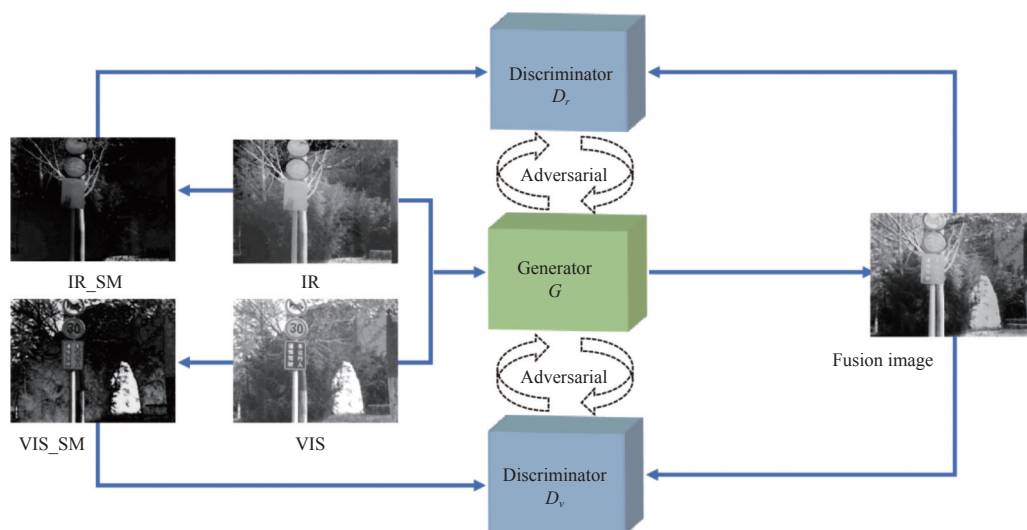


图 2 本文算法模型结构框架图

Fig. 2 Structural framework diagram of proposed algorithm model

本文方法的主要核心包括 3 部分: 1) 显著性视图的构造, 通过均值滤波与中值滤波的结果作差来求解两种图像的显著图, 并将其作为鉴别器的输入图像; 2) 网络结构, 两个鉴别网络模型与一个生成网络模型的结构; 3) 损失函数, 在生成损失

函数中引入梯度约束以保留更多的细节信息, 及两个鉴别损失函数的数学表示。

2.2 视觉显著图的构造

视觉显著图 (saliency map, SM) 能够反映图像中像素点的显著性程度, 具体流程如图 3 所示。首

先采用均值滤波处理红外与可见光两幅源图像,再采用中值滤波,并将两者输出结果作差处理,即获得相应的高通图像^[17-18]。

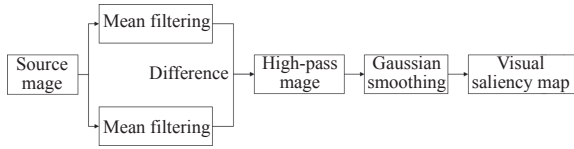


图3 视觉显著性提取流程图

Fig. 3 Flow chart of visual saliency extraction

红外图像与可见光图像分别由 I_1 、 I_2 表示,对应的高通图像分别由 H_1 、 H_2 表示,其过程由式(2)表示:

$$H_n(x, y) = \varphi_{\mu_n}(x, y) - \varphi_{\rho_n}(x, y), n = 1, 2 \quad (2)$$

式中: (x, y) 表示图像像素坐标; $\varphi_{\mu 1}$ 和 $\varphi_{\mu 2}$ 为两幅源图像均值滤波输出; $\varphi_{\rho 1}$ 和 $\varphi_{\rho 2}$ 为两幅源图像的中值滤波输出。

之后,对两幅高通图像 H_1 和 H_2 进行高斯滤波平滑处理,即获得两幅图像的显著值图像 S_1 与 S_2 。由式(3)表示:

$$S_n = |H_n| * g_{r_g, \sigma_g} \quad (3)$$

式中: $*$ 表示图像卷积操作; g_{r_g, σ_g} 是高斯滤波器,其窗口大小为 $(2r_g+1)(2r_g+1)$; r_g 、 σ_g 分别为滤波半径与标准差。

对可见光的灰度图像进行显著性区域提取,结果如图4所示。通过该方法提取到图像的显著性区域,其中石头与标牌为图像的主要信息。

2.3 网络结构

本文算法采用了基于卷积神经网络的生成器和鉴别器两个部分组成整个网络结构。

生成器 G 的网络结构如图5所示,红外与可

见光图像作为生成器的输入,而最终输出为 IR-RGB 融合图像。该生成器由 5 个公共卷积层组成:对于每个卷积层, padding 设置为 SAME, 步长设置为 1, 特征图的大小将不会改变。为了解决数据初始化的敏感性以及避免消失梯度,所有卷积层中均采用批量标准化 (BatchNorm, BN) 模块和激活函数。如图5中的箭头所示,在前4层中应用来自密集连接卷积网络的密集连接层。以前反馈方式在每层和所有层之间建立连接,以实现特征重用,防止在卷积过程中删除一些重要信息,选择的输入和输出通道是 2 : 16、16 : 16、32 : 16、48 : 16 和 64 : 1。



(a) 可见光图像



(b) 提取的显著图像

图4 可见光显著性区域提取结果

Fig. 4 Extraction results of visible-light salient area

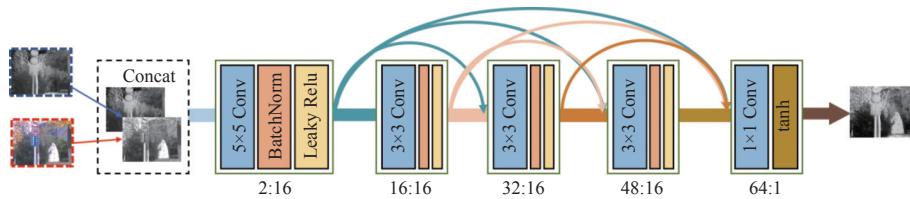
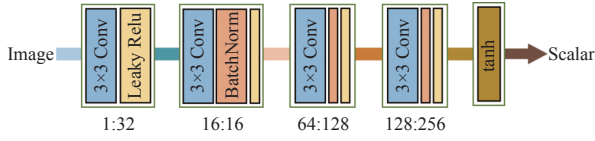


图5 生成器 G 的网络结构图

Fig. 5 Network structure diagram of generator G

本文设计算法的网络中有两个鉴别器 D_r 和 D_v , 它们具有相同架构,如图6所示。其作用均为分类器,生成一个输入图像来自真实数据而不是 G 的概率。该鉴别器采用 4 个卷积层和 1 个全连接层,

步长和 padding 分别设置为 2 和 VALID, 因此不需要池化层。该网络的前 4 层卷积核大小为 3×3, 卷积核个数分别为 32、64、128 和 256, 因此输入和输出通道分别为 1 : 32、32 : 64、64 : 128 和 128 : 256。

图6 鉴别器 D 的网络结构图Fig. 6 Network structure diagram of discriminator D

2.4 损失函数

在对网络进行训练时,需要以预测值和真实值之间的误差为基础对网络进行惩罚,这种惩罚通过损失函数来计算。由于损失函数种类繁多,因此需要根据不同的目标任务选择最适合的损失函数,以达到最佳的训练效果,从而使网络能够成功实现我们所期望的功能。生成对抗网络的损失函数主要包括生成器损失函数和鉴别器损失函数。

2.4.1 生成器损失函数

根据网络中生成器的目的,需要通过训练来欺骗鉴别器,并且要求约束内容中生成图像与源图像之间的相似性。故生成器包含生成器与鉴别器之间的对抗损失 L_{adv} 和生成图像与真实图像之间的内容损失 $L_{content}$,生成器损失函数定义如下:

$$L_G = L_{adv} + \alpha \cdot L_{content} \quad (4)$$

式中: α 为平衡系数; L_{adv} 和 $L_{content}$ 定义如式(5)、式(6)。

$$L_{adv} = -\frac{1}{N} \sum_{n=1}^N \{(1 - D_r(I_{rs}^n))^2 + \lambda(1 - D_v(I_{vsf}^n))^2\} \quad (5)$$

式中: I_f^n 表示在总数为 N 的同批次中的第 n 幅融合图像; $D(\cdot)$ 表示鉴别器的输出。

$$L_{content} = L_{grad} + \frac{\beta}{\alpha} L_{in} \quad (6)$$

图像的梯度信息能够反映边缘比较锐利的清晰度,且图像梯度值较大,则反映相应的图像边缘更加锐化。通过在损失函数中引入梯度约束来增强图像的细节信息。其中,损失 L_{grad} 表示梯度约束,强制融合的图像 I_f 具有类似于 I_r 和 I_v 的梯度,具体定义如下:

$$L_{grad} = \frac{1}{WH} (\|\nabla I_f - \nabla I_r\|_2^2 + \mu \|\nabla I_f - \nabla I_v\|_2^2) \quad (7)$$

式中: H 和 W 代表输入图像的尺寸; ∇ 为梯度算子; μ 为平衡参数。

强度损失 L_{in} 约束融合图像保持与源图像相似的强度分布,以便保留重要的对比度信息,强度损失定义为

$$L_{in} = \xi \|I_f - I_r\|_F^2 + (1 - \xi) \|I_f - I_v\|_F^2 \quad (8)$$

式中: $\|\cdot\|_F$ 表示矩阵 Frobenius 范数; ξ 是控制两项

之间权衡的参数。

2.4.2 鉴别器损失函数

鉴别器的作用是通过训练输出一个标量将真实数据与虚假数据区分开来,在提出的图像结构中,使用两个独立的鉴别器(即 D_r 、 D_v) 来约束生成器分别捕获更多的对比度和纹理信息。 D_r 、 D_v 旨在分别将融合图像与红外和可见光图像区分,需要同时考虑生成器与鉴别器之间的对抗关系和 D_r 与 D_v 的平衡。鉴别器的对抗性损失能够计算出分布之间的散度,从而识别出强度或纹理信息是否虚假,以尽可能获取真实信息^[14]。将相应的损失函数 L_{Dr} 和 L_{Dv} 定义为

$$L_{Dr} = \frac{1}{N} \sum_{n=1}^N (1 - D_{rs}(I_{rs}^n))^2 + (D_{rs}(I_{rsf}^n))^2 \quad (9)$$

$$L_{Dv} = \frac{1}{N} \sum_{n=1}^N (1 - D_{vs}(I_{vs}^n))^2 + (D_{vs}(I_{vsf}^n))^2 \quad (10)$$

鉴别器则是通过提取可见光与红外图像中的显著性特征来帮助区分融合图像和可见光图像。网络中采用最小平方损失函数,并基于最小化皮尔逊 χ^2 散度的原则,使网络训练过程更加稳定,从而使得鉴别器的损失函数能快速收敛。

依据两者图像不同的特性,可见光图像中主要存在丰富的纹理结构信息,而红外图像中主要存在显著的对比度信息。然而,在红外图像中也包含了一些结构信息;同样地,可见光图像也包含一些对比度信息。通过网络连续对抗的方式,本文方法可以生成具有良好视觉效果的融合图像,且以双鉴别器的方式在保留红外和可见光的固有信息方面取得了良好的平衡。

3 实验结果与分析

本文的实验图像均来自于自主搭建的红外-可见光成像系统,红外相机与可见光相机的分辨率、像素尺寸、帧率分别为 1280×1024 像素和 1280×1024 像素、 $12 \mu\text{m}$ 和 $4.8 \mu\text{m}$ 、 25 frame/s 和 211 frame/s 。两相机相对位置固定,通过软件触发控制采集图像。由于两相机之间具有不同的外部及内部参数,故在图像融合之前需对图像进行配准裁剪,最终进行融合实验的图像分辨率均为 800×600 像素。为了突出本文算法的优越性,本文选用几种算法进行定性和定量比较,包括:基于强度色调饱和度(intensity hue saturation, IHS)^[19]、主成分分析(principal component analysis, PCA)^[19]、拉普拉斯金

字塔(Laplacian pyramid, LP)^[20]、离散小波变换(discrete wavelet transform, DWT)^[21]、双尺度融合算法(two-scale image fusion, TIF)^[22]、FusionGAN^[13]及GANMcC^[23]、DDcGAN^[14]与SSGAN^[15]。

定量评价依赖于一些现有的统计指标,从不同方面评估融合结果的质量。为了客观地评价各种融合方法的结果,选用信息熵(EN)^[24]、平均梯度(MG)^[25]、空间频率(SF)^[26]和边缘强度(EI)^[27]作为评价指标来验证本文方法的有效性。不同指标所表示的意义不同,EN表示图像信息丰富度,MG用于衡量图像的纹理和细节的表征能力,SF衡量图像在空间中的活跃程度,EI表示图像的边缘细节

信息。本文选用的评价指标均与图像质量的关系呈正比例。

3.1 融合结果主观评价

不同场景下不同算法进行红外与可见光图像融合的结果如图7所示。场景1(Img 1)为道路地面场景,Img 1中红外图像的车位线几乎全黑,在对道路环境不明的情况下很难分辨出其为停车线,容易误判,在可见光图像中井盖与道路“融为一体”,而融合图像较好地还原了道路环境并能分辨出井盖;LP和DWT的融合效果更加偏向可见光图像,其结果有丰富的纹理细节,但目标和背景之间的对比并不明显;IHS、PCA、FusionGAN、GAN-

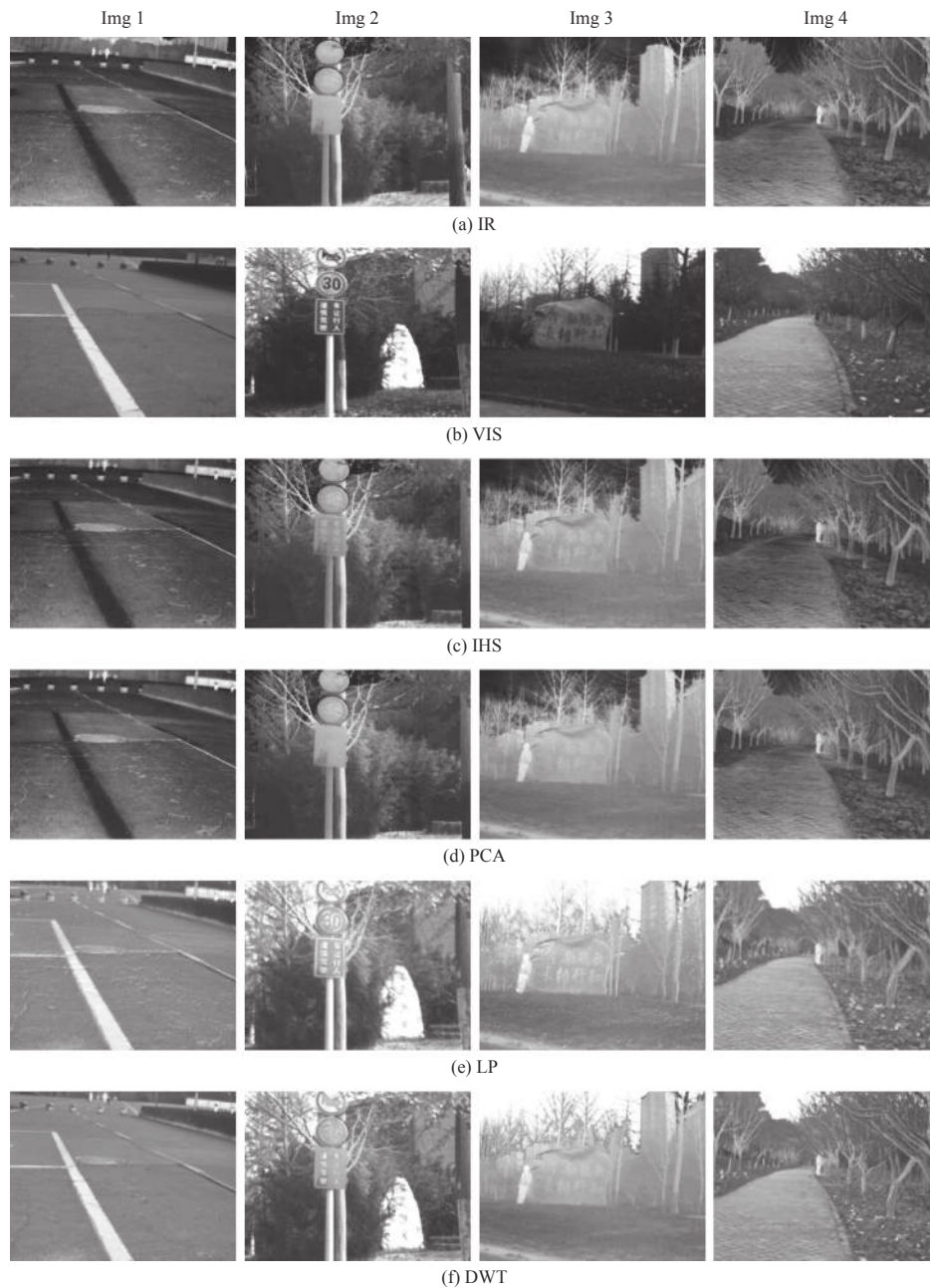




图7 不同算法在不同场景下进行红外与可见光融合

Fig. 7 Infrared and visible light fusion in different scenes by different algorithms

McC 及 DDcGAN 的融合效果更加偏向红外图像, 具有良好的热辐射信息, 但纹理细节不足; SSGAN 的细节信息较明显, 但其主观视觉效果不符合真实场景; 而 TIF 和本文方法更好地融合了红外和可见光的原始特征, 相较于 TIF, 本文方法的融合图像将道路上的石头路障和人更好地显示出来。场景 2(Img 2) 的红外图像无法识别到石头目标, 采用不同算法将红外图像与可见光图像的信息进行融合以突显目标; IHS、PCA 与 FusionGAN 的目标同样无法识别; LP、DWT 的结果过亮, 主观视觉效果

不佳; TIF、GANMcC 与 DDcGAN 的结果良好, 路牌的文字信息与场景的信息都能够观察; SSGAN 方法的融合图像细节信息丰富, 但图像中并未蕴含热辐射信息; 而本文所提出的算法的融合结果不仅有明显的对比度, 而且有丰富的纹理细节。场景 3(Img 3) 为包含天空背景下的人物与石碑场景, 在 IHS、PCA、FusionGAN、GANMcC 与 DDcGAN 的融合图像中, 天空背景是更加偏向红外图像的黑色, 与主观视觉特征不符合; SSGAN 的结果特征细节过于突出, 但背景信息不真实; TIF 的结果并

没有凸显显著目标,图像整体较为模糊,且人物目标不明显;LP、DWT与本文方法在凸显显著目标的同时保留了天空背景的视觉特征,但LP与DWT对草地表征过亮,整体图像效果不佳;本文方法不仅突出了人物目标,周围的环境信息也得以较好地保留。在场景4(Img 4)中,IHS、PCA、FusionGAN、GANMcC与DDcGAN、SSGAN的融合图像中,天空背景偏向红外图像的黑色,且整体图像主管视觉偏暗,不符合拍摄的光照场景;TIF的结果未凸显人物目标;LP、DWT与本文方法既凸显了目标,也较好保留了天空背景的视觉特征,周围的树木也较明显;本文方法生成的图像中人物(即目标)和树木(即背景)之间的对比较强,且树木的纹理细节丰富,融合图像在两幅源图像的信息之间取得了较好的平衡。

总结4组实验结果,IHS、PCA方法的融合效

果接近红外图像,拥有较强的对比度,但缺少纹理细节;LP、DWT方法的图像接近可见光图像,保留了丰富的纹理细节,但其对比度不明显;TIF方法的图像目标与背景区分不明显,且主观视觉效果不佳;FusionGAN、GANMcC与DDcGAN、SSGAN方法的融合图像虽然同时保留了红外的热辐射信息与可见光图像的纹理信息,但是对实际环境背景的描述不恰当,且整体效果偏暗;本文方法的显著性特征明显,且目标信息突出,对环境描述符合实际情况。

3.2 融合结果客观评价

除了通过主观视觉对不同算法进行定性描述之外,本文选择了信息熵、平均梯度、空间频率以及边缘强度4种评价指标,对10种图像融合算法在不同场景下的图像进行定量评价。具体的客观评价结果如表1所示。

表1 不同方法对比实验客观评价结果

Table 1 Objective evaluation results for different methods of comparison experiments

Indices	Image	IHS	PCA	LP	DWT	TIF	FusionGAN	GAN-McC	DDc-GAN	SS-GAN	Ours
EN	Img 1	7.1385	7.1396	6.0798	5.9014	6.4340	6.6317	6.6732	6.9024	6.7654	6.6439
	Img 2	7.3526	7.3047	7.4338	7.4658	7.3544	6.8938	7.3268	7.4759	7.1954	7.7750
	Img 3	7.1899	7.2188	6.8083	6.8806	6.8213	6.8766	6.8131	6.9745	6.8953	7.3277
	Img 4	7.0535	6.9670	6.8570	6.8859	6.7813	6.5858	6.8968	7.1750	6.8662	7.3797
MG	Img 1	3.0475	3.0168	2.2450	1.8258	2.5076	1.8915	1.7213	2.6036	2.9483	2.5333
	Img 2	4.4226	4.2808	3.9696	4.3262	4.7368	2.5705	3.5501	4.2402	4.6658	4.7938
	Img 3	2.9760	2.8248	2.5365	2.3250	3.7478	2.3401	2.0830	3.2244	3.6167	3.3216
	Img 4	4.6270	4.2898	3.6627	3.9604	4.2811	2.6069	2.9144	4.4677	4.4814	5.2256
SF	Img 1	7.1119	7.0532	5.2430	4.6424	6.3809	4.3924	4.0784	5.8936	6.6952	5.9644
	Img 2	11.2139	11.0484	8.9492	9.6821	11.5608	5.7327	7.8452	9.4811	10.2115	11.7453
	Img 3	7.9201	7.6780	5.8432	6.0149	8.8624	5.1718	4.5104	7.2417	7.8375	7.7968
	Img 4	9.8239	9.1511	7.5521	8.4333	9.7530	5.2895	5.9182	9.0139	9.0211	10.8283
EI	Img 1	32.8634	32.5167	23.9207	19.5275	26.3026	20.6171	18.8558	28.0324	31.9833	26.9535
	Img 2	47.2782	45.6847	42.7457	46.7484	50.3630	27.9883	38.5317	45.3338	50.6896	51.6590
	Img 3	31.7773	30.0927	26.8156	24.9121	39.7449	25.2312	22.5273	34.0047	38.7299	34.9828
	Img 4	49.7003	45.8882	39.2003	42.5217	44.7360	28.3632	31.6111	47.8448	48.6068	56.1800

分析表1中不同方法评价指标的结果,就EN而言,相较于其他融合算法,本文方法的数据结果最佳,其熵值最大表明了其所蕴含的信息量越丰富,说明本文方法较好地实现了两种图像不同信息的融合。就MG而言,本文方法相较于其他融合算法表现效果更好,TIF算法次之,本文方法的图像结果平均梯度综合最大,说明该图像所能体现的层次更多,细节表现力更强,图像更清

晰。就SF而言,TIF算法与本文方法的空间频率数据指标为最高;在不同的实验结果中,除去第1组实验结果较低,其他几组实验结果证明本文改进的算法更胜一筹,指标值约为8.0~11.0;因此,本文方法所得图像的边缘信息更丰富,对比度更高,图像的融合质量最佳。就EI而言,本文改进的算法表现更好,在4组数据中有2组数据值最高,本文方法的融合图像边缘强度最大,而采用IHS

与 TIF 算法结果的边缘强度数据次之。可以得出结论, 本文方法具有更丰富的细节信息, 保留了红外图像与可见光图像中显著的对比度和丰富的纹理。

为了更直观地观察不同方法的指标对比结果, 将表 1 可视化表示, 如图 8 中 (a)、(b)、(c)、(d) 子图分别表示信息熵、平均梯度、空间频率以及边缘强度的折线堆积图。

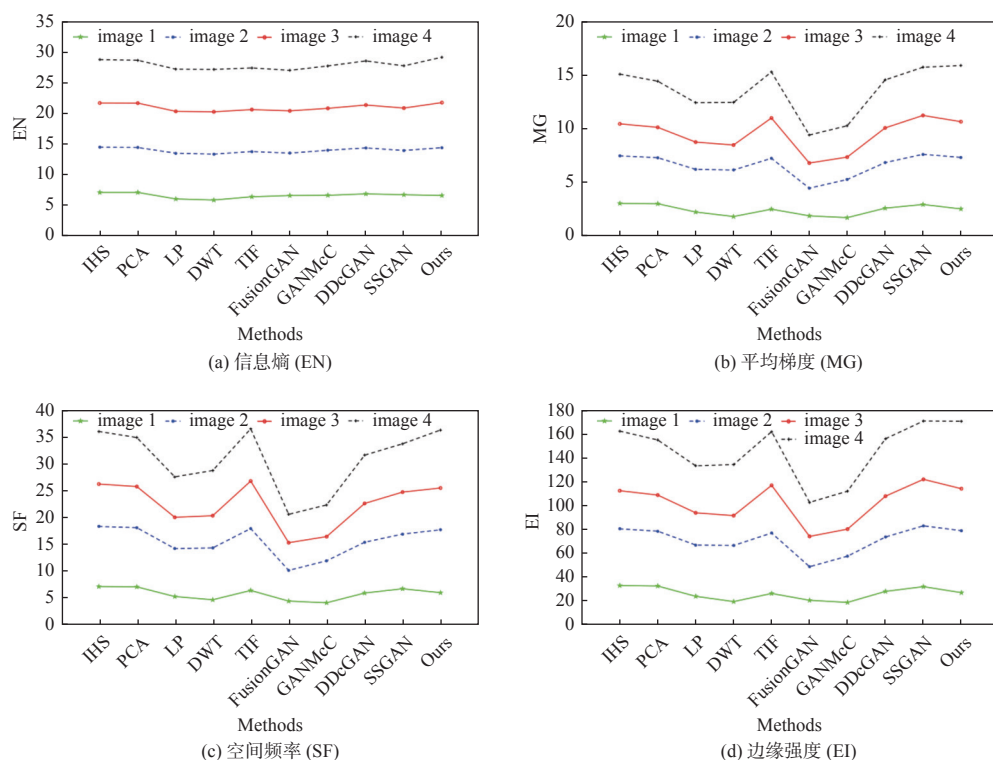


图 8 不同方法的评价指标折线堆积图

Fig. 8 Folded-line stacking diagram of evaluation indexes of different methods

观察图 8 不同方法的评价指标折线堆积图, 在所有指标中, 本文方法与 TIF 算法、SSGAN 算法的数据结果较好。综合几组实验及不同指标, 本文方法的表现最佳, 表明该算法能够从源图像获得更全面且有价值的信息。融合图像既包含了丰富的纹理细节, 也较好地保留了源图像的边缘细节特征, 其结果具有良好的视觉体验。通过对主观视觉效果图的观察及客观数据的描述, 最终得出本文方法在融合图像效果上优于其余算法。

4 结论

针对红外与可见光图像在不同场景下对目标识别能力的强弱问题, 本文将红外与可见光图像进行融合, 提出了一种基于显著性的双鉴别器生成对抗网络红外与可见光图像融合方法, 创新性地以两幅源图像的显著性区域作为鉴别器输入, 从而使得融合图像保留原始图像更多显著特征; 该网络的生成损失函数中添加了梯度约束, 以保

留更多的细节信息, 防止了高对比度目标区域边缘扩散造成的背景纹理污染。结果表明, 本文方法不仅能够保留红外图像强对比的目标信息, 而且能够保留可见光图像的主观视觉优势。综合评价验证了本文方法的可行性及优越性。随着红外与可见光图像融合技术广泛应用于工程技术和科学研究中, 本文所优化的图像融合方法实现了红外图像和可见光图像的有效融合, 提高了融合图像的质量, 在目标识别等领域中具有十分重要的意义。

参考文献:

- [1] LI H, WU X J, KITTLER J. Infrared and visible image fusion using a deep learning framework[C]// International Conference on Pattern Recognition, 2018. New York: IEEE, 2018.
- [2] 蔡伟, 徐佩伟, 杨志勇, 等. 复杂背景下红外图像弱小目标检测[J]. 应用光学, 2021, 42(4): 643-650.

- CAI Wei, XU Peiwei, YANG Zhiyong, et al. Dim-small targets detection of infrared images in complex backgrounds[J]. *Journal of Applied Optics*, 2021, 42(4): 643-650.
- [3] LIU Y, LIU S P, WANG Z F. A general framework for image fusion based on multi-scale transform and sparse representation[J]. *Information Fusion*, 2015, 24(24): 147-164.
- [4] YANG B, LI S T. Multifocus image fusion and restoration with sparse representation[J]. *IEEE Transactions on Instrumentation & Measurement*, 2010, 59(4): 884-892.
- [5] HARSANYI J C, CHANG C I. Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection approach[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 1994, 32(4): 779-785.
- [6] HAN J G, PAUWELS E J, ZEEUW P D. Fast saliency-aware multi-modality image fusion[J]. *Neurocomputing*, 2013, 111: 70-80.
- [7] 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.
- [8] 陈清江, 汪泽百, 柴昱洲. 改进VGG网络的多聚焦图像的融合方法[J]. *应用光学*, 2020, 41(3): 500-507.
- CHEN Qingjiang, WANG Zebai, CHAI Yuzhou. Multifocus image fusion method based on improved VGG network[J]. *Journal of Applied Optics*, 2020, 41(3): 500-507.
- [9] 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.
- [10] MA J Y, TANG L F, XU M L, et al. STDFusionNet: an infrared and visible image fusion network based on salient target detection[J]. *IEEE Transactions on Instrumentation and Measurement*, 2021, 70: 1-13.
- [11] LI J, HUO H T, LI C, et al. Multi-grained attention network for infrared and visible image fusion[J]. *IEEE Transactions on Instrumentation and Measurement*, 2020, 70: 1-12.
- [12] HUANG W, XIAO L, WEI Z, et al. A new pan-sharpening method with deep neural networks[J]. *IEEE Geoscience and Remote Sensing Letters*, 2015, 12(5): 1037-1041.
- [13] MA J Y, WEI Y, LIANG P W, et al. FusionGAN: a generative adversarial network for infrared and visible image fusion[J]. *Information Fusion*, 2019, 48: 11-26.
- [14] MA J Y, XU H, JIANG J 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.
- [15] HOU J L, ZHANG D Z, WU W, et al. A generative adversarial network for infrared and visible image fusion based on semantic segmentation[J]. *Entropy*, 2021, 23(3): 376.
- [16] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]// *Neural Information processing Systems*. Massachusetts: MIT Press, 2014.
- [17] 霍星, 邹韵, 陈影, 等. 双尺度分解和显著性分析相结合的红外与可见光图像融合[J]. *中国图象图形学报*, 2021, 26(12): 13.
- HUO Xing, ZOU Yun, CHEN Ying, et al. Dual-scale decomposition and saliency analysis based infrared and visible image fusion[J]. *Journal of Image and Graphics*, 2021, 26(12): 13.
- [18] 谢一博. 分布孔径成像阵列的图像处理与系统研究[D]. 西安: 西安工业大学, 2022.
- XIE Yibo. Image processing and system research of distributed aperture imaging array[D]. Xi'an: Xi'an Technological University, 2022.
- [19] 王瀛, 余岚旭, 王春喜, 等. 基于改进的IHS、PCA和小波变换的遥感图像融合算法[J]. *计算机与数字工程*, 2021, 49(4): 797-803.
- WANG Ying, YU Lanxu, WANG Chunxi, et al. Remote sensing image fusion algorithm based on improved IHS, PCA and wavelet transform[J]. *Computer & Digital Engineering*, 2021, 49(4): 797-803.
- [20] 卢越. 基于Laplacian图像金字塔的超分辨率算法研究[D]. 北京: 北京邮电大学, 2020.
- LU Yue. Research on image super-resolution algorithm based on Laplacian pyramid structure[D]. Beijing: Beijing University of Posts and Telecommunications, 2020.
- [21] 邵婷婷, 白宗文, 周美丽. 基于离散小波变换的信号分解与重构[J]. *计算机技术与发展*, 2014, 24(11): 159-161.
- SHAO Tingting, BAI Zongwen, ZHOU Meili. Decomposition and reconstruction of signal based on DWT[J]. *Computer Technology and Development*, 2014, 24(11): 159-161.
- [22] BAVIRISETTI D P, DHULI R. Two-scale image fusion of visible and infrared images using saliency detection[J].

- [Infrared physics and technology](#), 2016, 76: 52-64.
- [23] MA J Y, ZHANG H, SHAO Z F, 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(99): 1-1.
- [24] 谢一博, 徐乃涛, 周顺, 等. 分布式红外阵列相机的超分辨图像重建研究 [J]. *激光与光电子学进展*, 2022, 59(16): 254-261.
XIE Yibo, XU Naitao, ZHOU Shun, et al. Research on super-resolution image reconstruction of distributed infrared array camera[J]. *Laser & Optoelectronics Progress*, 2022, 59(16): 254-261.
- [25] LIU Z, SONG Y Q, SHENG V S, et al. MRI and PET image fusion using the nonparametric density model and the theory of variable-weight[J]. *Computer Methods & Programs in Biomedicine*, 2019, 175: 73-82.
- [26] 邸敬, 王国栋, 马帅, 等. 基于MSPCNN与FCM的红外与可见光图像融合[J]. *红外技术*, 2023, 45(1): 69-76.
DI Jing, WANG Guodong, MA Shuai, et al. Infrared and visible image fusion based on MSPCNN and FCM[J]. *Infrared Technology*, 2023, 45(1): 69-76.
- [27] 张小利, 李雄飞, 李军. 融合图像质量评价指标的相关性分析及性能评估[J]. *自动化学报*, 2014, 40(2): 306-315.
ZHANG Xiaoli, LI Xiongfei, LI Jun. Validation and correlation analysis of metrics for evaluating performance of image fusion[J]. *Acta Automatica Sinica*, 2014, 40(2): 306-315.