

<http://bhxb.buaa.edu.cn> jbuaa@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2024.0072

视觉显著性增强的双鉴别器红外与可见光图像融合

陈永^{1,2,*}, 周方春¹, 董珂¹

(1. 兰州交通大学电子与信息工程学院, 兰州 730070; 2. 甘肃省人工智能与图形图像处理工程研究中心, 兰州 730070)

摘要: 针对红外与可见光图像融合中边缘不清晰、细节缺失等问题, 提出了一种视觉显著性增强的双鉴别器融合方法。采用局部自适应对可见光图像进行增强, 并采用各向异性扩散对红外与可见光图像分解; 通过视觉显著性检测对分解后的细节层图像和基础层图像进行视觉增强; 设计密集连接 DenseNet 生成器模型对视觉增强后图像进行特征学习; 通过与双鉴别器博弈对抗得到融合结果。在公开数据集中与 10 种融合方法进行对比, 实验结果表明: 所提方法具有更清晰的细节信息, 在主客观评估上均优于对比方法, 客观评价指标较 FusionGAN 方法在信息熵、空间频率、结构相似性和标准偏差上分别提高了 7.4%、58.8%、25.5% 和 35.7%。

关键词: 红外与可见光图像融合; 视觉显著性增强; 各向异性扩散; 双鉴别器; 生成对抗网络
中图分类号: TP391

文献标志码: A

文章编号: 1001-5965(2026)04-1107-09

红外辐射是不可见光, 红外热成像利用热辐射原理形成, 受天气、光照影响小, 但红外图像存在清晰度较低、边缘结构模糊等问题^[1]。而可见光图像通过反射可见光成像, 虽细节纹理较为丰富, 但在光照较弱条件下存在辨识度低的问题^[2]。融合图像能够吸收源图像中的互补特征, 并具有更好的视觉感知效果^[3], 可广泛应用于目标检测^[4]、遥感测量^[5]、医学分析^[6]、军事^[7]等领域。

目前, 红外与可见光图像融合方法分为传统方法和基于深度学习的融合方法。传统方法通过人工设计规则融合, 如多尺度变换^[8]、稀疏表示^[9]、各向异性扩散融合 (anisotropic diffusion fusion, ADF)^[10]、视觉显著映射和加权最小二乘 (visual saliency map and weighted least square, VSMWLS)^[11]、多尺度引导图像与视频融合 (multi-scale guided image and video fusion, MGF)^[12]、引导滤波增强 (guided filter context enhancement, GFCE)^[13] 等, 但该类方法固有的数学变换无法有效提取目标特征, 未能充分考虑源图像

之间的模态差异, 易出现边缘模糊、对比度不足等问题^[14]。

深度学习的融合方法^[15-17]通过特征学习进行融合增强。文献 [18] 提出了一种基于密集连接网络的融合方法来提高融合质量, 但该方法通过单尺度卷积提出精细特征能力有限, 融合后纹理细节信息缺失比较严重。文献 [19] 提出了一种生成对抗融合方法, 但该方法仅使用可见光进行博弈对抗, 忽略了红外图像所包含的目标信息, 导致融合后的图像目标边缘信息模糊。文献 [20] 提出了一种具有多分类约束的生成对抗网络, 但该方法采用特定的内容丢失来限制生成器对源图像特征的提取和处理, 导致融合图像特征提取不足, 融合后细节对比度低。文献 [21] 提出了一种多分辨率生成对抗融合网络, 但该方法存在模型训练易崩塌的问题, 导致最终的融合图像细节纹理易丢失。文献 [22] 提出了一种具有满量程跳跃连接和双马尔可夫鉴别器的生成对抗网络, 通过全尺寸跳跃连接生成器

收稿日期: 2024-01-31; 录用日期: 2024-05-28; 网络出版时间: 2024-08-29 09:37

网络出版地址: link.cnki.net/urlid/11.2625.V.20240828.1354.001

基金项目: 国家自然科学基金 (62462043, 61963023); 甘肃省自然科学基金 (26JRRA589)

* 通信作者. E-mail: edukeylab@126.com

引用格式: 陈永, 周方春, 董珂. 视觉显著性增强的双鉴别器红外与可见光图像融合 [J]. 北京航空航天大学学报, 2026, 52 (4): 1107-1115.
CHEN Y, ZHOU F C, DONG K. Dual discriminator fusion of infrared and visible light images for visual saliency enhancement [J]. Journal of Beijing University of Aeronautics and Astronautics, 2026, 52 (4): 1107-1115 (in Chinese).

与双马尔可夫鉴别器建立对抗博弈来实现图像融合,但马尔可夫鉴别器关注于图像局部信息,难以提取完整的全局信息,融合结果存在信息缺失的问题。

综上,针对融合后边缘不清晰、细节缺失等问题,本文提出了一种视觉显著性增强的双鉴别器融合方法。通过各向异性扩散分解和视觉显著性检测^[23]来对图像特征进行增强,并通过密集网络和跳跃连接级联的方式进行特征提取,设计双鉴别器博弈对抗完成融合输出,最终利用公开数据集与10种有代表性的方法比较,从主观、客观角度分析并验证本文方法的融合效果。

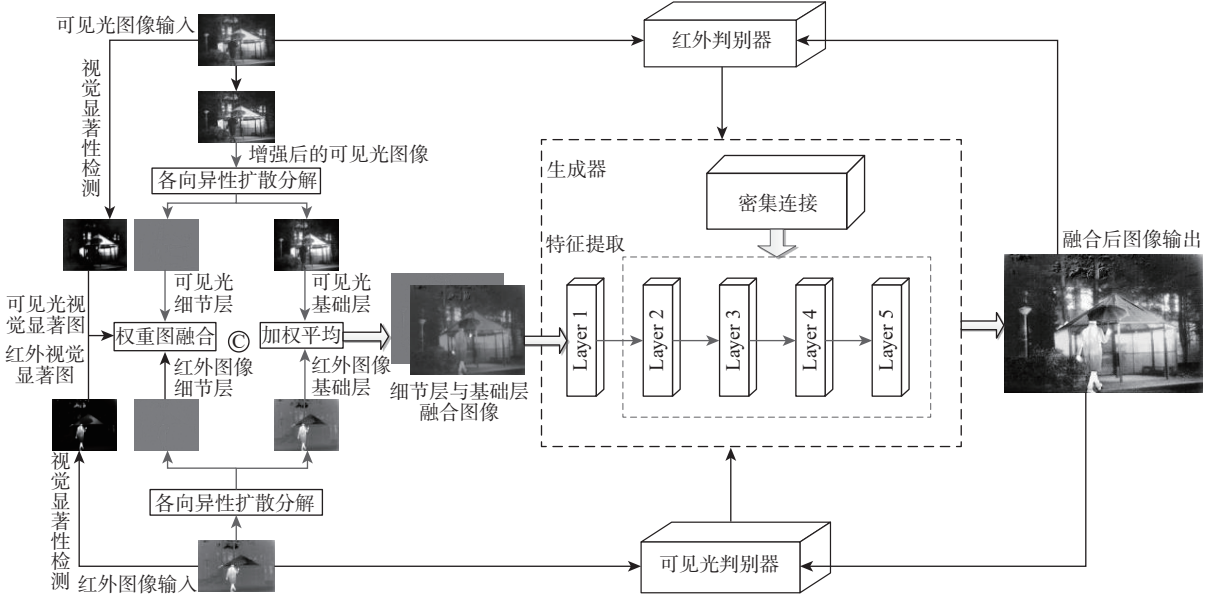


图1 网络总体框架

Fig. 1 Overall network framework

1.1 红外与可见光图像分解

1.1.1 可见光图像增强

为增强可见光图像边缘轮廓细节信息,对可见光图像采用自适应对比度增强 (adaptive contrast enhancement, ACE) 算法^[24]进行可见光图像增强。对可见光图像分别计算像素局部均值与局部方差:

$$M_{\text{vis}}(i, j) = \frac{1}{(2n+1)(2m+1)} \sum_{s=i-n}^{i+n} \sum_{k=j-m}^{j+m} f_{\text{vis}}(s, k) \quad (1)$$

$$\sigma(i, j) = \frac{1}{(2n+1)(2m+1)} \sum_{s=i-n}^{i+n} \sum_{k=j-m}^{j+m} (f_{\text{vis}}(s, k) - M_{\text{vis}}(i, j))^2 \quad (2)$$

式中: $f_{\text{vis}}(s, k)$ 为坐标 (s, k) 的点像素值; $M_{\text{vis}}(i, j)$ 表示以点 (i, j) 为中心、大小为 $(2n+1) \times (2m+1)$ 区域的局部均值; $\sigma(i, j)$ 为局部的方差。

在求得局部均值与标准差后,就可以进行图像

1 本文方法

为有效提升红外与可见光图像特征提取性能,获得更好的图像融合效果,本文方法基于生成对抗的思想进行图像融合网络的构建,如图1所示。本文融合模型由以下工作组成:首先,对可见光图像采用自适应对比度方法增强可见光图像边缘轮廓信息,并通过各向异性扩散将待融合图像分解为细节层和基础层;其次,利用显著性检测对分解后图像进行视觉增强;然后,设计生成器,并设计基于密集连接 DenseNet 的生成器对增强后图像进行特征提取;最后,通过与双鉴别器进行博弈对抗,输出最终的融合后图像。

增强,如下:

$$F_{\text{vis-ACE}}(i, j) = M_{\text{vis}}(i, j) + G(f_{\text{vis}}(i, j) - M_{\text{vis}}(i, j)) \quad (3)$$

$$G = \alpha \frac{\overline{M}_{\text{vis}}}{\sigma(i, j)} \quad 0 < \alpha < 1 \quad (4)$$

式中: $F_{\text{vis-ACE}}(i, j)$ 为增强后像素值; $\overline{M}_{\text{vis}}$ 为 $M_{\text{vis}}(i, j)$ 在整个图像上的平均值,即全局均值; α 为系数。

1.1.2 各向异性扩散分解

在完成可见光图像增强的基础上,为避免红外与可见光图像未区分细节层和基础层直接融合,造成融合后出现边缘问题,本文采用各向异性扩散的方法,对待融合的红外与可见光图像分解后分别进行融合。各向异性扩散方程使用通量函数来控制原始图像的扩散,可以表示为

$$f_t = c(x, y, t) \Delta f + \nabla c \cdot \nabla f \quad (5)$$

式中: Δ 为拉普拉斯操作符; ∇ 为梯度操作符; t 为迭

代次数; $c(x, y, t)$ 为通量函数, 也称为扩散系数, 其控制扩散的程度; f 为图像灰度强度函数。

选择的非线性扩散系数为

$$c(|\nabla f|) = 1 / \left[1 + \left(\frac{|\nabla f|}{r} \right)^2 \right] \quad (6)$$

式中: r 为常数, 可以调节扩散系数的大小, 从而控制各向异性扩散方程对于图像的平滑程度, r 值越大, 得到的图像越平滑, 但同时图像中的细节也更加模糊。

对式 (5) 进行求解的过程为

$$f_{i,j}^{t+1} = f_{i,j}^t + \lambda [c_N \bar{\nabla}_N f_{i,j}^t + c_S \bar{\nabla}_S f_{i,j}^t + c_E \bar{\nabla}_E f_{i,j}^t + c_W \bar{\nabla}_W f_{i,j}^t] \quad (7)$$

式中: $f_{i,j}^{t+1}$ 为较粗糙的图像; λ 为平滑系数; $\bar{\nabla}_N$ 、 $\bar{\nabla}_S$ 、 $\bar{\nabla}_E$ 、 $\bar{\nabla}_W$ 分别为北、南、东、西 4 个方向的最近邻差异; c_N 、 c_S 、 c_E 、 c_W 分别为北、南、东、西 4 个方向的扩散系数。

沿 4 个方向的局部图像梯度表示为

$$\bar{\nabla}_N f_{i,j} \equiv f_{i-1,j} - f_{i,j} \quad (8)$$

$$\bar{\nabla}_S f_{i,j} \equiv f_{i+1,j} - f_{i,j} \quad (9)$$

$$\bar{\nabla}_E f_{i,j} \equiv f_{i,j+1} - f_{i,j} \quad (10)$$

$$\bar{\nabla}_W f_{i,j} \equiv f_{i,j-1} - f_{i,j} \quad (11)$$

4 个方向扩散系数或通量函数可以表示为

$$c'_{N_{i,j}} = g(\|(\nabla f)_{i+1/2,j}^t\|) = g(\|\bar{\nabla}_N f_{i,j}^t\|) \quad (12)$$

$$c'_{S_{i,j}} = g(\|(\nabla f)_{i-1/2,j}^t\|) = g(\|\bar{\nabla}_S f_{i,j}^t\|) \quad (13)$$

$$c'_{E_{i,j}} = g(\|(\nabla f)_{i,j+1/2}^t\|) = g(\|\bar{\nabla}_E f_{i,j}^t\|) \quad (14)$$

$$c'_{W_{i,j}} = g(\|(\nabla f)_{i,j-1/2}^t\|) = g(\|\bar{\nabla}_W f_{i,j}^t\|) \quad (15)$$

式中: $g(\cdot)$ 为单调递减函数, 其为平滑和边缘保留之间提供权衡, 表达式为

$$g(|\nabla f|) = \begin{cases} \left| 1 - \frac{(\|\nabla f\|)^2}{K^2} \right|^2 & (\|\nabla f\|)^2 \leq K^2 \\ 0 & (\|\nabla f\|)^2 > K^2 \end{cases} \quad (16)$$

其中: K 为边缘导热系数。

上述各向异性扩散方法可以用 $\text{aniso}(f_n(x, y))$ 来表示, 通过边缘保持平滑的各向异性扩散过程来获得基础层 B_n 后, 细节层 D_n 的计算方法如下:

$$f_{B_n}(x, y) = \text{aniso}(f_n(x, y)) \quad (17)$$

$$f_{D_n}(x, y) = f_n(x, y) - f_{B_n}(x, y) \quad (18)$$

设原始红外图像与增强后的可见光图像分布为 f_{ir} 和 $F_{\text{vis-ACE}}$, 经过各向异性扩散分解, 则可以获得对应的基础层 $F_{\text{ir}}^B(x, y)$ 和 $F_{\text{vis}}^B(x, y)$ 如下:

$$\begin{cases} F_{\text{ir}}^B(x, y) = \text{aniso}(f_{\text{ir}}(x, y)) \\ F_{\text{vis}}^B(x, y) = \text{aniso}(F_{\text{vis-ACE}}(x, y)) \end{cases} \quad (19)$$

分解后对应的细节层图像计算如下:

$$\begin{cases} F_{\text{ir}}^D(x, y) = f_{\text{ir}} - F_{\text{ir}}^B(x, y) \\ F_{\text{vis}}^D(x, y) = F_{\text{vis-ACE}} - F_{\text{vis}}^B(x, y) \end{cases} \quad (20)$$

1.2 图像融合策略

经过上述分解操作后, 采用视觉显著性检测方法进行视觉显著性增强并进行融合操作。

1.2.1 改进视觉显著性检测模型

视觉显著性检测频率调谐 (frequency-tuned, FT) 方法是一种从频率角度出发, 利用颜色和亮度特征输出具有清晰对象边界的方法^[25]。FT 显著性检测为了消除高频噪声干扰, 使用 5×5 高斯核, 获取显著图, 计算公式为

$$S(x, y) = \|I_{\mu} - I_{\text{whc}}(x, y)\| \quad (21)$$

式中: I_{μ} 为算术平均特征向量; $I_{\text{whc}}(x, y)$ 为高斯模糊后像素值。

但原始 FT 方法利用高斯模糊滤波器, 易出现模糊的问题。为克服上述不足, 本文采用引导滤波器对原始 FT 显著性检测进行改进。引导滤波器是线性移变滤波器, 单个像素经过滤波输出的加权平均值表示为

$$q_i = \sum_j W_{ij}(\mathbf{I}) p_j \quad (22)$$

式中: i 和 j 为像素位置; W_{ij} 为滤波核函数; \mathbf{I} 为引导图像; p 、 q 分别为输入、输出图像。

滤波核函数的表达式为

$$W_{ij}(\mathbf{I}) = \frac{1}{|\omega|^2} \sum_{h(i,j) \in \omega_k} \left[1 + \frac{(I_i - \mu_k)(I_j - \mu_k)}{\sigma_k^2 + \varepsilon} \right] \quad (23)$$

式中: ω 为窗口内的像素个数; ω_k 为第 k 个核函数窗口; μ_k 和 σ_k^2 分别为引导图像 \mathbf{I} 在窗口内的均值和方差; ε 为平滑因子。

本文利用引导滤波对其进行改进:

$$S(x, y) = \|I_{\mu} - I_{\text{Guided}}(x, y)\| \quad (24)$$

式中: $I_{\text{Guided}}(x, y)$ 为引导滤波后的输出。

1.2.2 细节层和基础层融合

为突出融合图像的细节信息, 本文对权重图进行归一化处理, 其构建方法如下:

$$C_{\text{vis}}^D(x, y) = \frac{S_{\text{vis-ACE}}(x, y)}{S_{\text{vis-ACE}}(x, y) + S_{\text{ir}}(x, y)} \quad (25)$$

$$C_{\text{ir}}^D(x, y) = \frac{S_{\text{ir}}(x, y)}{S_{\text{vis-ACE}}(x, y) + S_{\text{ir}}(x, y)} \quad (26)$$

式中: C_{vis}^D 和 C_{ir}^D 分别为 F_{vis}^D 和 F_{ir}^D 所对应的权重图; $S_{\text{vis-ACE}}$ 和 S_{ir} 分别对应增强后的可见光和红外显著图。

利用权重图进行加权融合得到细节图 F_{fused}^D :

$$F_{\text{fused}}^D = C_{\text{ir}}^D F_{\text{ir}}^D + C_{\text{vis}}^D F_{\text{vis}}^D \quad (27)$$

采用平均相加融合规则得到融合图像的基础层 F_{fused}^B , 计算如下:

$$F_{\text{fused}}^B = \frac{1}{2} (F_{\text{vis}}^B + F_{\text{ir}}^B) \quad (28)$$

1.2.3 生成器的网络模型

完成细节层和基础层融合后, 通过生成对抗网络完成图像融合。本文设计的生成器如图2所示。

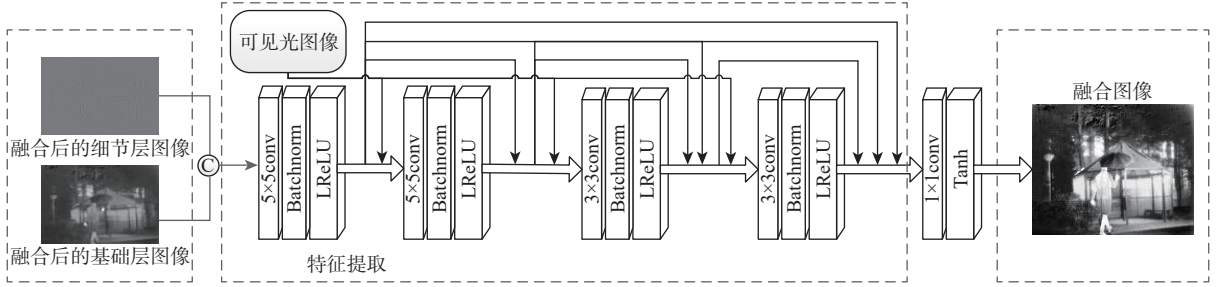


图2 生成器结构框架

Fig. 2 Network structure of generator

1.2.4 鉴别器的网络架构

在完成生成器设计后, 采用红外与可见光双鉴别器, 其均分别由4个卷积层级联1个全连接层组成。在前4层均使用LReLU函数, 最终通过线性层输出一个标签, 以获得分类结果, 表示输入图像是真实图像或融合图像。

1.2.5 损失函数

损失函数 L_G 包括多对抗性损失 L_{adv} 和内容损失 L_{con} , 可表示为

$$L_G = L_{\text{adv}} + \nu L_{\text{con}} \quad (29)$$

式中: ν 为平衡系数。

$$L_{\text{adv}} = \frac{1}{N} \sum_{i=1}^2 \sum_{n=1}^N (D_i(F_{\text{fused}}^n) - a)^2 \quad (30)$$

式中: N 为融合图像的数量; F_{fused}^n 为融合图像; $D_i(\cdot)$ 表示鉴别器分类结果; a 为生成器希望鉴别器能相信的虚假数据值。

式(29)中内容损失函数^[22]定义为

$$L_{\text{con}} = \frac{1}{HW} (\|F_{\text{fused}} - f_{\text{ir}}\|_F^2 + \xi \|\nabla F_{\text{fused}} - \nabla f_{\text{vis}}\|_F^2) \quad (31)$$

式中: H 为图像的高; W 为图像的宽; F_{fused} 为融合图像; ξ 为平衡参数。

本文设计的双鉴别器损失函数可表示为

$$L_{D-\text{ir}} = \frac{1}{N} \sum_{N=1}^N (D_{\text{ir}}(f_{\text{ir}}) - b)^2 + \frac{1}{N} \sum_{N=1}^N (D_{\text{ir}}(F_{\text{fused}}) - d)^2 \quad (32)$$

示。其中, 生成器是具有密集块的5层卷积神经网络, 对视觉显著性增强后的细节层和基础层图像进行特征提取。第1层和第2层均使用 5×5 卷积, 第3层和第4层使用 3×3 卷积减少网络参数, 第5层采用 1×1 卷积核将串联特征降维, 实现特征融合。第1~5层形成了一个密集块卷积层, 每层中的步幅长度设置为1。为保留图像更多的有效信息, 引入密集网^[21]的思想构造生成器模型。通过连接和卷积的方法, 将前一层的信息输出连接在一起, 作为下一个卷积层的输入。

$$L_{D-\text{vis}} = \frac{1}{N} \sum_{N=1}^N (D_{\text{vis}}(f_{\text{vis}}) - c)^2 + \frac{1}{N} \sum_{N=1}^N (D_{\text{vis}}(F_{\text{fused}}) - d)^2 \quad (33)$$

式中: b 、 c 、 d 分别为红外图像、可见光图像、融合图像的真值数据。

2 实验结果与分析

2.1 实验设置与环境

本节进行不同方法的对比实验, 选取公共红外与可见光融合 TNO 数据集, 实验的仿真平台配置 Inter(R) Core i7, 3.6 GHz, 16 GB, NVIDIA GeForce GTX 1660, 操作系统 Windows10, 软件环境为 python3.6, 网络训练环境 Tensorflow1.12。参数优化器选择 Adam, 学习率设置为 0.001, epoch 设置为 100, 损失函数参数设置为: $\lambda_1=0.1$ 、 $\xi_1=2$ 、 $\xi_2=5$ 。

为验证本文方法的有效性, 将其与10种融合方法比较。对比方法选择的主要依据如下: 因本文方法主要涉及红外与可见光图像分解、视觉显著性增强及生成对抗深度学习融合3种主要技术, 因而从上述3个角度选择了10种同类比较方法。其中, 在图像分解层面选择了 ADF^[10] 方法, 该方法通过使用各向异性扩散将原始图像进行分解生成基础层和细节层, 来提高图像融合质量。在视觉显著性增强角度, 主要选择 VSMWLS^[11]、MGF^[12]、二维经验模态分解 (bidimensional empirical mode decomposition, BEMD)^[23] 3种显著性增强方法, 方法均使用视

觉显著性方法来增强融合图像的信息获取能力。在生成对抗深度学习角度, 选择包括 GFCE 方法^[13]、多分类约束的 GAN 网络 (generative adversarial network with multiclassification constraints, GANMcC)^[20]、双鉴别器 GAN 融合 (dual-discriminator conditional generative adversarial network, DDcGAN)^[21]、具有满量程跳跃连接和双马尔可夫生成对抗融合 (generative adversarial network with full-scale skip connection and dual Markovian discriminator, GAN-FM)^[22]、密集连接网络 DenseFuse^[18] 和基于生成对抗网络的图像融合 FusionGAN^[19] 等 6 种方法, 其主要以生成对抗网络和密集连接网络为主完成融合。在对比实验中, 通过与上述代表性的图像分解、视觉显著性增强、生成对抗深度学习方法进行对比, 更好地对本文方法进行主客观评价分析。

2.2 融合图像主观评价

为进行不同融合方法的主观评价, 从 TNO 数据集中选取 6 组不同场景进行对比实验, 并深入验证其在不同关键技术路径上的有效性, 本文设计了

2 组不同的对比实验。第 1 组实验选取图像分解层面和视觉显著性增强角度与本文方法进行比较, 如图 3 所示。第 2 组实验以生成对抗与深度学习核心维度, 与所选的代表性方法进行针对性比对。如图 4 所示。不仅能够展示本文方法在融合结果上的整体优越性, 更能清晰剖析其性能增益的具体技术来源, 从而提供全面而深入的性能评估。

为进行不同融合方法的主观评价, 从 TNO 数据集中选取 6 组不同场景进行对比实验, 图 3 和图 4 给出了 6 组图像的主观对比结果。为能更加直观地观察, 对图像中信息进行标注, 从图 3 和图 4 中可以看出, DDcGAN 方法融合图像出现红外目标模糊、细节纹理信息保留较差且有严重的噪声现象。ADF、GFCE、GAN-FM 和 DenseFuse 方法结果相比于 DDcGAN 结果包含了较为清晰的轮廓信息, 但仍存在部分细节不清晰的问题。如图 3 中红框所示, ADF 方法融合结果中, 第 1 组图像的人物纹理模糊和第 5 组的窗户轮廓不清晰。GFCE 方法第 3 组汽车轮胎细节信息缺失, 在图 4 中, GAN-FM 方



图 3 图像分解层面和视觉显著性增强角度融合方法比较

Fig. 3 Comparison of image decomposition and visual saliency enhancement fusion methods



图4 生成对抗与深度学习融合方法比较

Fig. 4 Comparison of generative adversarial and deep learning fusion methods

法融合结果中第3组结果中树木和第5组结果中存在警示牌丢失部分细节、纹理信息的现象。在图3中, MGF方法融合结果中第2组图像右侧边缘出现了细节信息丢失的问题, 整体画面对比度较弱。文献[23]的BEMD方法相对于MGF方法融合效果有所提升, 但在融合后轮廓信息存在清晰度低的问题, 如第3组图像中窗户边缘不够清晰。在图4中, DenseFuse方法融合结果中第1组图像中草丛和第5组图像中窗户轮廓细节缺失。在图3中, VSMWLS方法也出现了部分细节纹理缺失的问题, 如第6组中草丛与地面的轮廓不清晰, 对比度不明显。在图4中, GANMcC方法和FusionGAN方法通过生成对抗网络进行融合, 其结果主要体现了可见光的信息, 但对于红外特征存在融合后信息缺失较多的问题。与比较方法相比, 本文方法既保留了红外目标的显著性信息, 又包含了丰富的可见光细节特征, 具有较好的融合效果。

2.3 融合图像客观评价

为增强实验的准确性和客观性, 选取信息熵、互信息、结构相似性(structural similarity index measure, SSIM)、标准偏差、空间频率、视觉保真度等6种客

观评价指标对融合效果进行对比分析。指标数值越大, 融合图像质量越好。对TNO实验进行定量分析, 如表1所示。可以发现, 信息熵、SSIM、标准偏差和空间频率这4个客观评价指标均优于其他方法。本文方法的空间频率指标相对于FusionGAN方法提高了58.8%, 图像中边缘和纹理信息更加清晰。同时, 信息熵、SSIM和标准偏差指标也优于FusionGAN方法, 分别提高了7.4%、25.5%和35.7%。对于图像的互信息和视觉保真度, 虽然在所有对比方法中无法达到最佳, 但相对于FusionGAN方法, 都有不同程度的提高。与比较方法相比, 本文融合主客观评价均优于比较方法。

2.4 消融实验

本节进行消融实验, 分别为: ①仅以生成对抗网络模型进行融合; ②ACE算法增强生成对抗网络模型(ACE+GAN); ③增强图像各向异性分解生成对抗网络模型(AD+ACE+GAN); ④增强图像各向异性分割密集连接生成双对抗网络模型(AD+ACE+DenseGAN)。以TNO数据集图像“Street”为例进行主观对比, 如图5所示。可以看出, 仅以GAN模型实验结果模糊不清, 融合结果较差; ACE

表 1 TNO 客观对比值

Table 1 Objective comparison value of TNO

方法	信息熵	互信息	SSIM	标准偏差	空间频率	视觉保真度
ADF ^[10]	7.161 7	0.688 3	0.503 0	38.450 6	0.031 8	0.707 5
VSMWLS ^[11]	7.174 8	0.659 0	0.624 8	40.063 5	0.032 3	0.740 1
GFCE ^[13]	7.161 3	0.688 3	0.503 0	36.403 6	0.043 6	0.724 1
MGF ^[12]	7.255 3	0.693 2	0.612 3	41.847	0.031 1	0.492 4
BEMD ^[23]	7.324 1	0.762 3	0.653 8	48.032 4	0.035 39	0.562 4
GANMcC ^[20]	7.396 0	0.772 5	0.605 3	42.572 9	0.039 5	0.585 3
DDcGAN ^[21]	6.632 4	0.567 4	0.442 3	31.436 7	0.025 6	0.464 0
GAN-FM ^[22]	7.665 4	0.452 5	0.436 5	52.546 6	0.038 3	0.720 6
DenseFuse ^[18]	7.136 0	0.677 0	0.547 4	35.75	0.030 9	0.652 1
FusionGAN ^[19]	7.337 4	0.642 4	0.537 0	40.354 6	0.031 1	0.530 5
本文	7.878 7	0.693 8	0.674 0	54.757 3	0.049 4	0.735 8

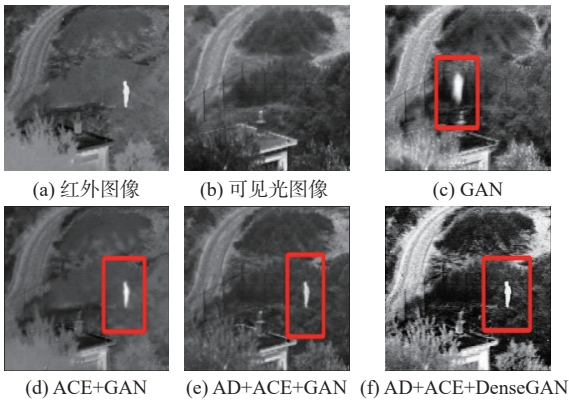


图 5 消融实验结果

Fig. 5 Results of ablation experiments

算法增强提高了清晰度;增强图像各向异性分解生成对抗网络模型融合结果改善了上述问题,但融合后结果依然存在模糊的现象;本文方法融合结果纹理信息更丰富。

选取 SSIM 和平均梯度进行量化评价,结果如表 2 所示。可以看出,各模块对最终融合性能提升均有所贡献,表明了提出的视觉显著性增强的双鉴别器红外与可见光图像融合方法的有效性。

表 2 消融实验客观指标

Table 2 Objective indicators of ablation experiment

模块	SSIM	平均梯度
GAN	0.694 4	2.084
ACE+GAN	0.757 2	2.719
AD+ACE+GAN	0.814 0	3.104
AD+ACE+DenseGAN	0.838 6	3.592

2.5 复杂度及效率分析

对模型复杂度和效率进行比较分析。与 ADF^[10]、VSMWLS^[11]、DenseFuse^[18]、FusionGAN^[19] 进行比较,结果如表 3 所示。分析得知,ADF 方法通过对图像

进行分解、融合,网络简单,因此,其计算量和运行时间在 5 种方法中最小,但从表 1 主客观评价可知,该方法融合效果较差。VSMWLS 方法采用迭代引导滤波算法对原始图像进行迭代多次,利用视觉显著映射算法,计算权重,最终进行图像融合,相比 ADF 方法能获取更多的图像细节、纹理信息,其计算开销较小。从表 3 可以看出,本文方法计算量及运行时间相较于 DenseFuse 和 FusionGAN 方法均有所减少,但其计算量和运行时间都高于 VSMWLS 方法。结合表 1 和表 3 可以得出,本文融合主客观评价虽然相较于其他 10 种对比方法取得了更好的融合效果,但从表 3 可知,本文方法在实时性方面还存在一定的不足。红外与可见光融合系统在实际应用时,会涉及具体的硬件设备和融合软件系统的部署,其对实时性、小型化及功耗都有相应的要求。本文方法目前仅从显著性增强及双鉴别器对抗技术方面来提高融合质量,在未来融合系统部署时,为了满足融合系统对实时性的要求,需要对本文方法进一步优化。主要优化提升措施有:模型轻量化与硬件并行化加速。在模型轻量化方面,对于本文模型,可以采用通道剪枝的策略,对生成网络冗余的权重参数进行剪枝,达到模型轻量化的目的,并在剪枝后采用参数微调的方法维持性能。此外,还可以采用知识蒸馏迁移学习的策略,通过利

表 3 计算量和运算时间

Table 3 Computational amount and processing time

方法	浮点运算速度/ $10^9 s^{-1}$	运行时间/s
ADF ^[10]	2.874	0.786
VSMWLS ^[11]	3.024	0.793
DenseFuse ^[18]	3.592	0.881
FusionGAN ^[19]	3.648	0.867
本文	3.153	0.853

用教师-学生知识蒸馏机制,重构本文深度学习模型,以便轻量化后更适合硬件部署场景。在硬件并行加速方面,后续研究中也将会进一步结合硬件加速技术来提高融合系统的实时性。例如,采用边缘计算技术,将嵌入式设备边缘计算推理任务在本地完成,而红外与可见光融合操作在边缘端完成来提高融合的实时性。利用并行计算技术,将融合处理任务采用分治法的算法思想,分解成多个融合子任务,通过对FPGA加速融合处理来提高实时性。

3 结论

1) 本文提出了一种视觉显著性增强的双鉴别器红外与可见光图像融合方法,通过各向异性扩散分解图像,并利用视觉显著性检测方法分别增强细节层和基础层。

2) 通过设计密集连接DenseNet生成器进行特征提取,并通过双鉴别器博弈对抗得到最好的融合结果。

3) 实验结果表明,较其他对比方法,本文方法客观评价指标都有所提高。较FusionGAN方法,信息熵、空间频率、结构相似性和标准偏差分别提高了7.4%、58.8%、25.5%和35.7%,结果表明,本文方法融合后细节信息更丰富,有效提高了融合性能。

参考文献 (References)

- [1] 陈永,王镇,卢晨涛,等. 红外弱光下多特征融合与注意力增强铁路异物检测[J]. 北京航空航天大学学报, 2023, 49(8): 1884-1894.
CHEN Y, WANG Z, LU C T, et al. Detection of railway object intrusion under infrared low light based on multi-feature and attention enhancement network[J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49(8): 1884-1894(in Chinese).
- [2] GONZÁLEZ J R, DAMIÃO C, MORAN M, et al. A computational study on the role of parameters for identification of thyroid nodules by infrared images (and comparison with real data)[J]. Sensors, 2021, 21(13): 4459.
- [3] MA J Y, MA Y, LI C. Infrared and visible image fusion methods and applications: a survey[J]. Information Fusion, 2019, 45: 153-178.
- [4] TU Z Z, LI Z, LI C L, et al. Multi-interactive dual-decoder for RGB-thermal salient object detection[J]. IEEE Transactions on Image Processing, 2021, 30: 5678-5691.
- [5] XIAO W X, ZHANG Y F, WANG H B, et al. Heterogeneous knowledge distillation for simultaneous infrared-visible image fusion and super-resolution[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 5004015.
- [6] DINH P H. Combining Gabor energy with equilibrium optimizer algorithm for multi-modality medical image fusion[J]. Biomedical Signal Processing and Control, 2021, 68: 102696.
- [7] 陈永,张娇娇,王镇. 多尺度密集连接注意力的红外与可见光图像融合[J]. 光学精密工程, 2022, 30(18): 2253-2266.
CHEN Y, ZHANG J J, WANG Z. Infrared and visible image fusion based on multi-scale dense attention connection network[J]. Optics and Precision Engineering, 2022, 30(18): 2253-2266(in Chinese).
- [8] WANG Z S, XU J W, JIANG X L, et al. Infrared and visible image fusion via hybrid decomposition of NSCT and morphological sequential toggle operator[J]. Optik, 2020, 201: 163497.
- [9] 娄熙承,冯鑫. 潜在低秩表示框架下基于卷积神经网络结合引导滤波的红外与可见光图像融合[J]. 光子学报, 2021, 50(3): 0310004.
LOU X C, FENG X. Infrared and visible image fusion in latent low rank representation framework based on convolution neural network and guided filtering[J]. Acta Photonica Sinica, 2021, 50(3): 0310004(in Chinese).
- [10] VASU G T, PALANISAMY P. Multi-focus image fusion using anisotropic diffusion filter[J]. Soft Computing, 2022, 26(24): 14029-14040.
- [11] 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.
- [12] BAVIRISSETTI D P, XIAO G, ZHAO J H, et al. Multi-scale guided image and video fusion: a fast and efficient approach[J]. Circuits, Systems, and Signal Processing, 2019, 38(12): 5576-5605.
- [13] ZHOU Z Q, DONG M J, XIE X Z, et al. Fusion of infrared and visible images for night-vision context enhancement[J]. Applied Optics, 2016, 55(23): 6480-6490.
- [14] SUN C Q, ZHANG C, XIONG N X. Infrared and visible image fusion techniques based on deep learning: a review[J]. Electronics, 2020, 9(12): 2162.
- [15] LI J, ZHU J M, LI C, et al. CGTF: convolution-guided Transformer for infrared and visible image fusion[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 5012314.
- [16] WANG Z S, WANG J Y, WU Y Y, et al. UNFusion: a unified multi-scale densely connected network for infrared and visible image fusion[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2022, 32(6): 3360-3374.
- [17] LI H, WU X J, KITTLER J. RFN-Nest: an end-to-end residual fusion network for infrared and visible images[J]. Information Fusion, 2021, 73: 72-86.
- [18] 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.
- [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] 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: 5005014.
- [21] 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.
- [22] ZHANG H, YUAN J T, TIAN X, et al. GAN-FM: infrared and visible image fusion using GAN with full-scale skip connection and dual Markovian discriminators[J]. IEEE Transactions on Computa-

- tional Imaging, 2021, 7: 1134-1147.
- [23] 崔晓荣, 沈涛, 黄建鲁, 等. 基于 BEMD 改进的视觉显著性红外和可见光图像融合[J]. 红外技术, 2020, 42(11): 1061-1071.
- CUI X R, SHEN T, HUANG J L, et al. Infrared and visible image fusion based on BEMD and improved visual saliency[J]. Infrared Technology, 2022, 42(11): 1061-1071(in Chinese).
- [24] YE X Y, GAO S L, LI F M. ACE-STDN: an infrared small target detection network with adaptive contrast enhancement[J]. Journal of Infrared Millim Waves, 2023, 42(5): 701-708.
- [25] JOO M G, LEE K T, SANG P, et al. Laser-generated focused ultrasound transmitters with frequency-tuned outputs over sub-10-MHz range[J]. Applied Physics Letters, 2019, 115(15): 154103.

Dual discriminator fusion of infrared and visible light images for visual saliency enhancement

CHEN Yong^{1, 2, *}, ZHOU Fangchun¹, DONG Ke¹

(1. School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China;

2. Gansu Provincial Engineering Research Center for Artificial Intelligence and Graphics & Image Processing, Lanzhou 730070, China)

Abstract: In order to solve the problem of unclear edges and missing details in infrared and visible light image fusion, a saliency enhanced dual discriminator generation adversarial infrared and visible light image fusion method is proposed. First, infrared and visible light images are broken down using anisotropic diffusion, while visible light images are improved using local adaptation. Then, visual saliency detection is used to visually enhance the decomposed detail layer image and the base layer image. Next, a dense connected DenseNet generator model is designed to perform feature learning on visually enhanced images. Finally, the fusion result is obtained by competing with the dual discriminator game. Experimental results demonstrate that the suggested approach has more precise information and performs better than the comparison algorithm in both subjective and objective assessments when compared to ten fusion techniques in a public dataset. Compared with the FusionGAN algorithm, the proposed method has improved objective evaluation indicators such as information entropy, spatial frequency, structural similarity, and standard deviation by 7.4%, 58.8%, 25.5%, and 35.7%, respectively.

Keywords: infrared and visible light image fusion; visual saliency enhancement; anisotropic diffusion; dual-discriminator; generate adversarial network

Received: 2024-01-31; Accepted: 2024-05-28; Published Online: 2024-08-29 09:37

URL: link.cnki.net/urlid/11.2625.V.20240828.1354.001

Foundation items: National Natural Science Foundation of China (62462043,61963023); Gansu Provincial Nature Science Foundation (26JRRA589)

* Corresponding author. E-mail: edukeylab@126.com