

基于注意力密集网络的伪彩色红外与 可见光图像融合^{*}

漆建环 倪波 周晓彦 倪海彬 杨凌升 常建华
(南京信息工程大学电子与信息工程学院 南京 210044)

摘要:针对现有红外与可见光图像融合算法中存在融合图像的纹理细节不清晰,红外信息和纹理细节的显示不平衡等问题,提出了一种基于注意力密集网络的伪彩色红外与可见光图像融合方法。首先对灰度的红外图像进行伪彩色处理再与彩色的可见光图像组成多通道数据输入融合网络。其次,设计了一种由卷积层和带有注意力模块的密集连接块组成的生成器网络结构,关注源图像的关键信息,增强网络提取源图像信息的能力。最后,利用红外像素、可见光像素、可见光梯度和红外梯度构建内容损失函数,以保持融合图像中红外目标和纹理细节的平衡。与5种具有代表性的融合方法进行定性和定量比较。结果表明,该方法所获得融合图像的峰值信噪比、信息熵、平均梯度和互信息取得最优值,分别为31.684 1、6.558 1、6.009 6、3.096 0。定量以及定性结果证明所提融合方法使融合图像具有更为丰富的纹理细节以及良好的视觉效果。

关键词:红外与可见光图像;图像融合;注意力模块;密集连接块

中图分类号: TP391 **文献标识码:** A **国家标准学科分类代码:** 520.2060

Pseudo-color infrared and visible image fusion based on attention-dense network

Qi Jianhuan Ni Bo Zhou Xiaoyan Ni Haibin Yang Lingsheng Chang Jianhua

(School of Electronic and Information Engineering, Nanjing University of Information Science and Technology,
Nanjing 210044, China)

Abstract: Current infrared and visible image fusion algorithms often suffer from issues such as unclear texture details in the fused image and an unbalanced display of infrared information and texture details. In this paper, we propose an image fusion method of pseudo-color infrared and visible images based on attention-dense network. The greyscale infrared image is first processed in pseudo-color and then combined with the colored visible image to form a multi-channel data input fusion network. Secondly, a generator network structure consisting of convolutional layers and densely connected blocks with attention modules is designed to focus on the key information of the source image and enhance the ability of the network to extract information from the source image. Finally, the content loss function is constructed by using infrared pixels, visible pixels, visible gradient and infrared gradient to keep the stability of infrared target and texture details in the fused image. Qualitative and quantitative comparisons are made with five representative fusion methods. The results show that the peak signal-to-noise ratio, information entropy, average gradient, and mutual information of the fused images obtained by this method achieve the optimal values of 31.684 1, 6.558 1, 6.009 6, and 3.096 0, respectively. The quantitative and qualitative results demonstrate that the proposed fusion method results in a fused image with richer texture details and good visual effects.

Keywords: infrared and visible images; image fusion; attention module; dense connected block

收稿日期:2023-11-09

^{*} 基金项目:国家自然科学基金(62175114, 61875089)项目资助

0 引言

红外图像通常反映物体的热辐射信息,具有较强的适应性和抗干扰能力,如在弱光、风、沙、烟等环境下对目标进行成像^[1]。但由于红外传感器的局限性,红外图像不能有效反映图像纹理信息^[2]。可见光图像具有较高的分辨率,可以提供高质量的图像和纹理细节,但容易受到光线和气候等外部因素的影响^[3]。由于红外和可见光图像各有不同的特点和优势,红外与可见光图像融合可以生成同时包含热红外信息和纹理细节信息的高质量、高分辨率图像^[4],在军事探测、医学诊断、遥感等领域^[5]具有广泛的应用前景。

目前,用于红外与可见光图像融合的方法主要有传统方法和深度学习方法两种。传统的图像融合方法通常涉及对图像进行像素级的测量,包括在空间域或在变换域,然后手动设计融合规则,以确保不同图像来源的信息被合理结合在一起,从而实现图像融合的目的。经典的传统图像融合框架主要包括基于多尺度变换的融合框架^[6]、基于稀疏表示的融合框架^[7]、基于子空间的融合框架^[8]、基于显著性的融合框架^[9]、基于变分模型的融合框架^[10]等。虽然现有的传统融合方法在一定程度上取得了良好的融合效果,但是融合过程中需要手工设计复杂的融合规则,阻碍了传统融合方法的进一步发展。

深度学习凭其强大的特征提取能力在计算机视觉领域取得显著优势,其中包括图像融合任务。基于深度学习的图像融合算法可分为基于自编码器(autoencoder, AE)^[11]的融合方法、基于卷积神经网络(convolutional neural network, CNN)^[12]的融合方法和基于生成对抗网络(generative adversarial network, GAN)^[13]的融合方法。Zhang等^[14]提出了一种基于CNN的端到端融合框架(PMGI),该框架通过强度和梯度路径维持梯度和强度的比例,避免了手动设计融合规则。其方法对一般损失函数进行建模,并针对不同的图像融合任务调整了超参数。为了促使网络有目的地提取和合并有意义的特征, Ma等^[15]通过引入显著目标掩模定义了红外和可见光图像融合所需的信息。其网络可以同时融合红外和可见光图像并实现显著目标检测,但其更加关注源图像的显著目标特征而忽视了纹理细节特征的提取。Li等^[16]首次将AE框架应用于图像融合(DenseFuse),该融合框架由编码层、融合层和解码组成,编码过程中利用密集连接块(DenseNet)提取更多的图像特征,但由于AE结构的特征提取能力有限,最后的融合图像失真。Li等^[17]进一步提出一种基于蜂巢连接的融合方法(NestFuse),蜂巢连接的引入使网络可以从源图像中提取多尺度特征,但由于手动设计的融合规则比较简单,无法将编码网络提取到的特征充分传递给解码器,导致融合图像部分区域纹理细节模糊。冯鑫等^[3]提出一种双通道残差密集融合网络,同时利用L1范数和注意力机制分别对残差通道和密集通道设计融合规则,将更多

的特征传递给解码器,但是仅在测试阶段使用注意力机制,设计的融合规则无法适用于全部图像场景。Ma等^[13]首次将GAN框架应用于红外与可见光图像融合,将融合过程看成是生成器与判别器的博弈。然而单判别器容易在特征提取过程中造成模态失衡问题,导致融合图像缺失热红外目标信息或纹理细节信息,融合图像模糊。Ma等^[18]进一步提出双判别器条件GAN(DDcGAN)融合算法来维持不同模态间的信息平衡,但是,具有双判别器的生成对抗网络在训练过程中很难达到纳什均衡,融合图像依旧缺失部分细节信息。

为解决融合算法中纹理细节信息缺失,红外目标和纹理细节显示不平衡的问题,提出一种基于注意力密集网络的伪彩色红外与可见光图像融合算法。首先,将伪彩色处理后的红外图像与彩色可见光图像组成多通道数据输入融合网络,融合网络通过从多通道数据中提取互补信息,更多地保留源图像中的纹理细节信息和红外信息。其次,利用注意力密集连接保证在关注各层重要信息的同时充分利用各层信息,使得每一层的输出信息都可以被充分利用,从而达到更加理想的融合效果。最后,考虑到红外图像中不仅包含大量的目标信息,还具有少量纹理细节信息;可见光图像中不仅具有丰富纹理细节,也包含少量目标信息。设计内容损失函数使融合图像具有更多的纹理细节和目标信息,同时使融合图像显示平衡,进一步提升融合性能。

1 本文方法

1.1 网络框架

在以往的红外与可见光图像融合研究中,都是对彩色可见光图像进行灰度处理,再与灰度红外图像融合,由于不是所有的通道都呈现在输入数据中,因此很难构造多通道分布并提取多通道互补信息,从而导致融合图像中纹理细节模糊。本文首先通过通道复制的方法将灰度红外图像转化三通道伪彩色图像,然后与彩色可见光图像共同组成多通道数据输入融合网络,过程如图1所示。

训练过程如图1(a)所示,首先对灰度红外图像进行伪彩色处理,然后与彩色可见光图像根据通道进行拼接。它们被当作生成器的输入来输出融合图像。判别器由一个细节判别器和一个目标判别器组成。细节判别器对作为输入的融合图像或可见图像进行评估,以判断融合图像是来自生成器还是真实图像。目标判别器以融合图像或红外图像作为输入,判断融合图像是来自于生成器还是真实图像。通过生成器和判别器的对抗博弈,源图像中的可见纹理细节和热红外信息不断被补充到融合图像中。

测试过程如图1(b)所示,对红外图像进行伪彩色处理,然后将红外和可见光图像输入到已训练好的生成器。对生成的图像进行灰度化处理,得到最终的灰度融合图像。

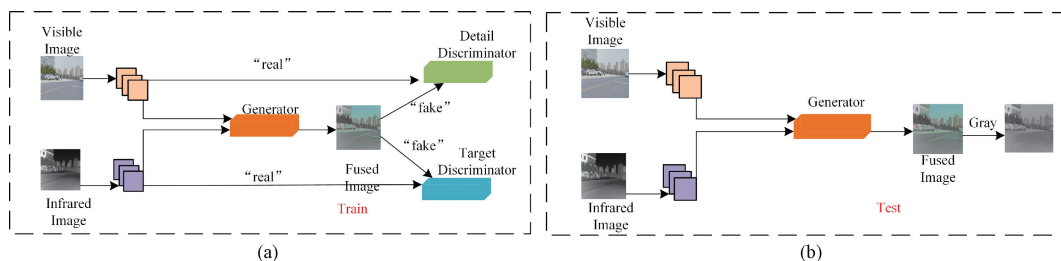


图1 融合框架

Fig. 1 Fusion framework

1) 生成器网络架构

生成器网络架构由编码器(Encoder)和解码器(Decoder)组成。编码器由4个卷积层(Conv)和卷积注意力模块(CBAM)构成。为在特征提取过程中保留更多的信息,编码器的后3个卷积层中添加了密集连接操作。解码器由4个串联的卷积层组成。网络参数如图2所示,其中 k 表示卷积核的大小, n 表示输出通道数量, s 代表卷积操作移动的步幅。由于进行下采样操作会造成源图像信息

的丢失,因此根据卷积层输出特征尺寸计算方法将 k 设置为3, s 设置为1,特征提取过程中将不进行下采样,从而使输入卷积层的特征尺寸和卷积层输出的特征尺寸一致,减少特征提取过程中的信息丢失。编码器和解码器前3个卷积层均使用批标准化(BatchNorm)和修正线性单元(ReLU)激活函数,最后一个卷积层使用激活函数Tanh来输出融合图像。

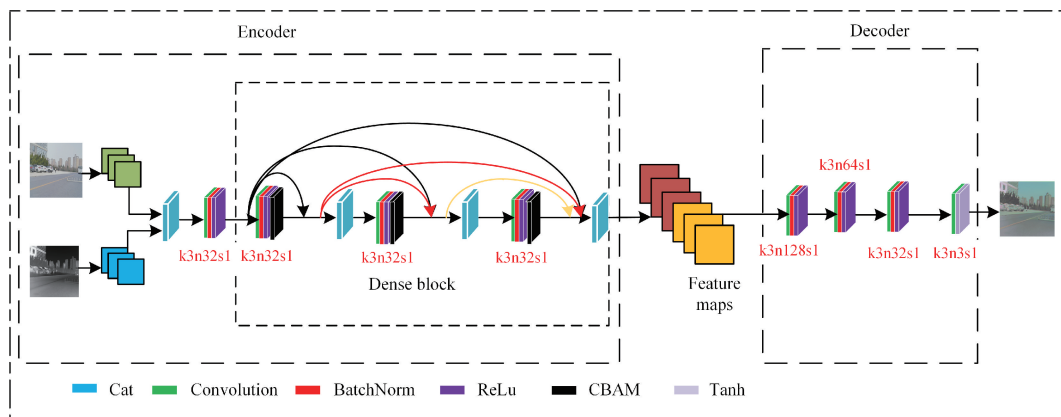


图2 生成器网络架构

Fig. 2 Generator network architecture

2) 判别器网络架构

在生成对抗网络中有两个判别器,分别为细节判别器和目标判别器。与只有一个判别器的生成对抗网络相比,具有两个判别器的生成对抗网络有利于融合图像在保留更多纹理细节信息的同时保留尽可能多的热红外目标信息。细节判别器和目标判别器具有相同的网络结构,如图3所示,判别器由4个卷积层和一个全连接层(FC)组成。其网络参数如图3中所示,4个卷积层的卷积核大小均为 4×4 ,移动步长均设置为2,每个卷积层中都是用实例标准化(InstanceNorm2d)和LeakyReLU激活函数。

3) 注意力模块

本文注意力模块采用的是CBAM,该模块是由通道注意力模块(channel attention module)和空间注意力模块(spatial attention module)组成,如图4所示。相比于

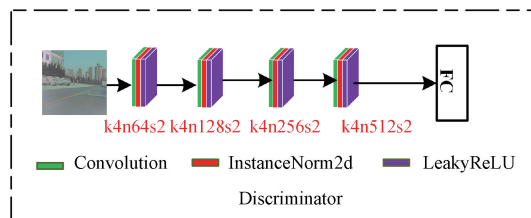


图3 判别器网络架构

Fig. 3 Discriminator network architecture

SENet模块只关注于通道的注意力机制,CBAM会对输入进来的特征层依次进行通道注意力机制的处理和空间注意力机制的处理,可以取得更好的结果。

1.2 损失函数设计

本文生成器损失函数由两部分组成,分别为对抗损失

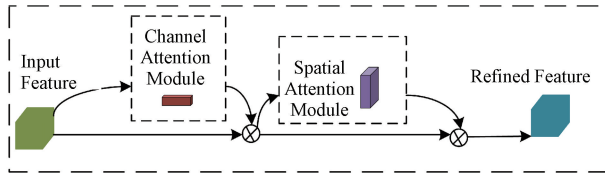


图4 卷积注意力模块

Fig. 4 Convolutional block attention module

和内容损失,其函数如下:

$$L_G = L_{adv} + \lambda L_{con} \quad (1)$$

内容损失比例系数用 λ 表示,对抗损失的表达式为:

$$L_{adv} = \frac{1}{N} \sum_{n=1}^N [\log(1 - D_D(I_f^n))] + \frac{1}{N} \sum_{n=1}^N [\log(1 - D_T(I_f^n))] \quad (2)$$

式中: $D_D(I_f)$ 表示细节判别器对融合图像的判别结果; $D_T(I_f)$ 表示目标判别器对融合图像的判别结果; N 表示融合图像的个数。

内容损失表达式如下:

$$L_{con} = \frac{1}{HWC} (\alpha_1 \|I_f - I_r\|_F^2 + \alpha_2 \|\nabla I_f - \nabla I_v\|_F^2 + \alpha_3 \|I_f - I_v\|_F^2 + \alpha_4 \|\nabla I_f - \nabla I_r\|_F^2) \quad (3)$$

式中: I_f 表示融合后的图像; I_r 表示红外图像; I_v 表示可见光图像; H, W, C 分别表示图像的高度、宽度和通道数; $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 分别表示各种损失的比例系数。 ∇ 为图像矩阵的梯度运算。

梯度损失约束融合图像包含丰富的纹理细节,其中, α_2 和 α_4 分别表示可见光梯度和红外梯度损失项的权重系数。像素损失约束融合图像保持与源图像相似的图像分布,其中, α_1 和 α_3 分别表示红外像素和可见光像素损失权重因子。

通过调整权重 α 的大小来决定每种信息的保留程度。在红外与可见光图像的融合中,期望得到的融合图像保留了可见光图像的纹理信息(梯度信息细节)和红外图像的热红外目标信息(像素信息),即可见光梯度信息和红外像素信息是主要信息,红外梯度信息和可见光像素信息是辅助信息。主要信息的权重应该大于辅助信息的权重,此外,梯度损失项一般也小于像素损失项,故为使它们在优化过程中同等重要,梯度损失的权重要大于像素损失的权重,所以4个权重值的设置应该遵循以下规则:

$$\alpha_1 > \alpha_3, \alpha_2 > \alpha_4, \{\alpha_2, \alpha_4 > \alpha_1, \alpha_3\} \quad (4)$$

如此设置权重系数是为使融合图像具有丰富的纹理细节和显著目标的同时保持融合图像的显示平衡,其权重数值通过消融实验获得。

本文中判别器的损失函数只由对抗损失组成,其表达式如下:

$$L_{D_D} = \frac{1}{N} \sum_{n=1}^N [-\log(D_D(I_v^n))] + \frac{1}{N} \sum_{n=1}^N [-\log(1 - D_D(I_f^n))] \quad (5)$$

$$L_{D_T} = \frac{1}{N} \sum_{n=1}^N [-\log(D_T(I_v^n))] + \frac{1}{N} \sum_{n=1}^N [-\log(1 - D_T(I_f^n))] \quad (6)$$

$$L_D = L_{D_D} + L_{D_T} \quad (7)$$

式中: $D_D(I_v)$ 表示为细节判别器对可见光图像的判别结果; $D_T(I_r)$ 表示为目标判别器对红外图像的判别结果; $D_D(I_f)$ 表示为细节判别器对融合图像的判别结果; $D_T(I_f)$ 表示为目标判别器对融合图像的判别结果。

2 实验与结果

2.1 实验设计

实验使用了M³FD数据集^[19],该数据集是由大连理工大学软件学院在2022年所公布,该数据集包含不同场景下不同对象类型的高分辨率红外和可见光图像。本文选取了4 000对红外和可见光图像作为训练集,10对具有代表性的红外和可见光图像作为测试集。

在生成器损失函数中,根据多次实验,设置内容损失项的权重系数为 $\lambda = 100$ 。像素损失和梯度损失项的权重系数设为 $\alpha_1 = 1, \alpha_2 = 5, \alpha_3 = 1, \alpha_4 = 3$ 。实验环境基于PyTorch框架,CUDA版本11.6,使用的显卡是NVIDIA RTX 4060,训练轮次设置为400轮,批次大小(batch_size)设为4,优化器选用Adam,学习率设置为 1×10^{-4} 。

客观评价可以提供定量的评价指标,而主观评价更能反映人的视觉感知。因此,在评价图像融合效果时,可以将客观评价与主观评价相结合,综合考虑评价结果,得到更加准确、全面的评价。为了客观评价融合图像的质量,本文选取结构相似度(SSIM)^[20]、峰值信噪比(PSNR)^[21]、信息熵(EN)^[22]、空间频率(SF)^[23]、平均梯度(AG)^[24]、互信息(MI)^[25]等作为评价指标。

2.2 对比实验

为了验证提出方法的有效性,与另外5种融合方法进行了比较分析,包括FusionGAN^[13]、DenseFuse^[16]、DDcGAN^[18]、AttentionFGAN^[26]、TarDal^[19]。

将上述5种方法应用于10对随机选择的具有代表性的红外和可见光源图像。不同方法对6组源图像的融合结果如图5所示,源图像大小均设置为 256×256 。

各种融合方法的融合图像如图5所示。使用红框标记热红外目标信息,用蓝框标记可见纹理信息,以便进行清晰的比较。基于主观分析,FusionGAN方法得到的融合图像具有较为清晰的纹理细节,但是由于该方法只有一个将可见光图像作为输入的判别器,导致其融合图像中红外信息缺失,不能有效融合红外图像信息;比如图5B和C中的行人轮廓模糊,对比度低,但本文方法将行人轮廓清晰呈现,且与背景的对比度明显。DenseFuse方法通过编码网络有效地从源图像中提取突出特征,使融合后的图像具有良好的目标显著性。然而,由于在自编码器训练过程

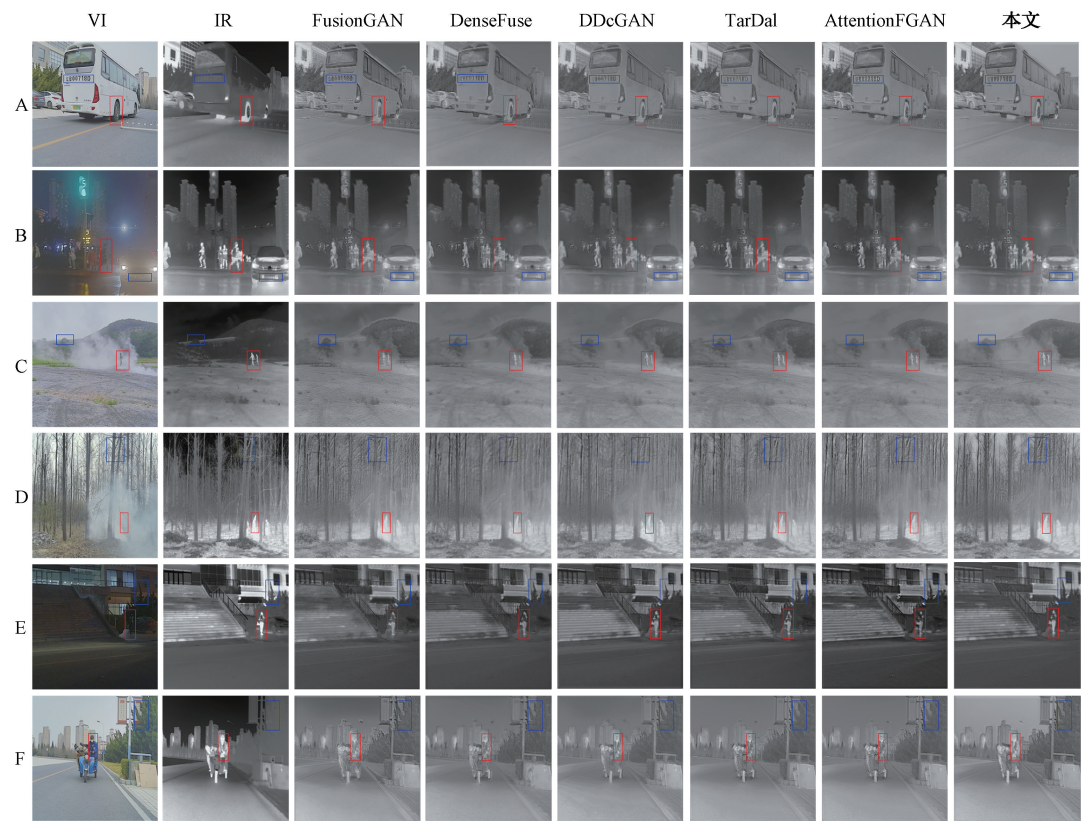


图 5 不同算法融合结果比较
Fig. 5 Comparison of fusion results of different algorithms

中对源图像信息的利用不足,降低了源图像之间互补信息的保存,导致融合图像缺失纹理细节信息。比如图 5A 中的车牌号部分信息缺失,导致车牌号不清晰,而所提方法充分保留了车牌号的信息,使融合图像具有清晰的车牌号。DDcGAN 方法中使用了双判别器,得到的融合图像比 FusionGAN 方法得到的融合图像具有相对清晰的红外目标,但是由于双判别器难以达到纳什均衡,导致图 5D 中树枝重叠处出现融合失真,树枝的细节模糊。本文提出方法中树枝的细节清晰呈现且几乎不失真。AttentionF-GAN 方法在 DDcGAN 的基础上往生成器和判别器都添加了注意力模块,得到图 5C 中清晰的车牌号信息,但同样由于训练双判别器的困难,导致图 5B 中人物目标不清晰,

出现伪影等现象。本文所提方法不仅能保留清晰的车牌号信息,还能呈现对比度高的人物目标。TarDal 方法将图像融合与目标检测进行联合优化,检测精度高,但融合图像却出现了部分噪声,导致视觉效果不想理想。综上所述,本文提出的融合方法在同时保留纹理细节信息和突出红外目标上展现了良好的性能,获得较好的融合结果。

不同算法在不同评价指标下生成的融合图像的平均值如表 1 所示。表 1 为 6 种方法在 6 个指标下的比较结果,其中黑体为最优值,下划线为次优值。在本文提出的算法中,PSNR 取得最优值说明融合过程中失真最小;EN 均取得最优值说明融合图像中包含的信息更丰富;MI 取得最优值说明两个源图像给融合图像传递了更可观的信

表 1 不同融合算法的图像度量平均值
Table 1 Mean values of image metrics for different fusion algorithms

算法	SSIM	PSNR	EN	SF	AG	MI
FusionGAN	0.817 5	29.802 4	6.053 9	15.747 8	5.684 1	2.530 4
DenseFuse	0.852 4	30.717 6	6.388 8	11.880 5	5.229 0	2.792 9
DDcGAN	0.841 2	30.739 9	6.395 4	11.436 5	4.962 0	2.812 5
AttentionFGAN	0.856 5	30.648 4	6.361 4	12.996 4	5.318 5	2.804 9
TarDal	0.855 9	30.761 7	6.442 4	11.092 8	4.968 0	2.663 5
本文	0.830 5	31.684 1	6.558 1	13.894 8	6.009 6	3.096 0

息;AG 和 SF 分别取得最优值和次优值表明融合图像中拥有更加丰富的纹理细节信息。总之,本文提出的融合方法在最大程度上保留了源图像的有用信息,尤其是纹理信息。

上述 6 种算法融合结果的 SSIM、PSNR、EN、SF、AG、MI 等指标的对比如图 6 所示。本文提出的方法具有最佳的融合性能。

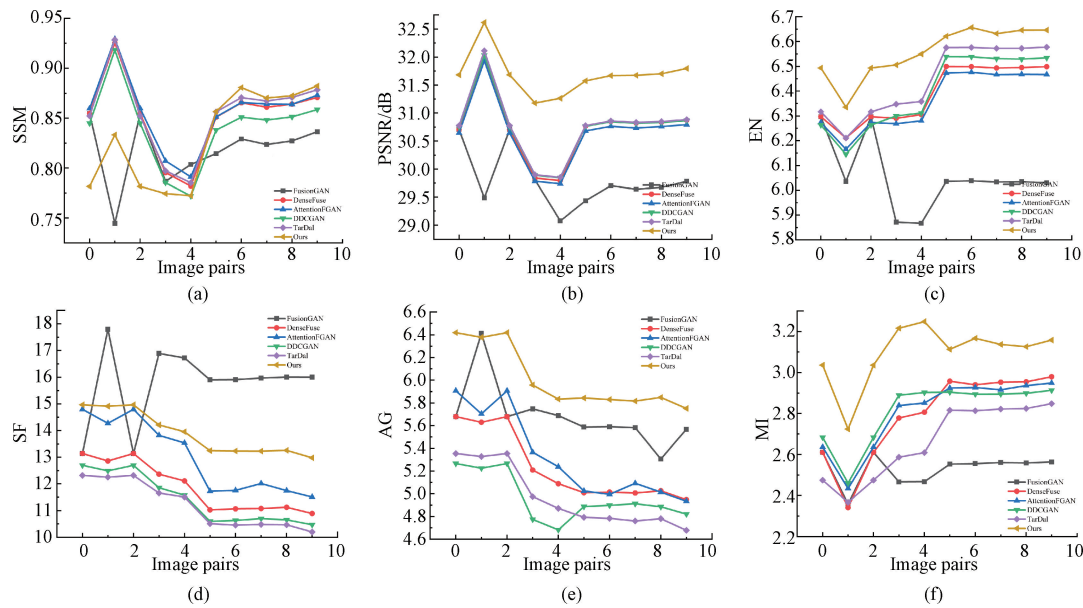


图 6 客观评价融合效果
Fig. 6 Objective evaluation fusion effect

为了验证本文提出的将红外图像进行伪彩色处理这一图像处理方式的有效性,本文与以往的图像处理方法进行了比较分析。在实验分为两组实验,第 1 组实验是将可见光图像灰度处理,再用本文提出的融合框架进行图像融合,第 2 组实验是按照本文提出的融合方法进行图像融合。

将上述两组融合实验应用于 10 对随机选择的红外和可见光源图像。两种不同图像处理方法对一组源图像的融合结果如图 7 所示。

由图 7 可知,相比较于以往的图像处理方式,在本文提出的图像处理方式下所得到的融合图像具有更为清晰的可见光纹理细节,同时,得到的融合图像保留更为清晰的红外



图 7 两种不同图像处理的结果比较
Fig. 7 Comparison of the results of two different image processing

信息。本文提出的图像处理方式使融合图像具有更高的融合质量和更好的视觉效果。两种不同图像处理方式下得到的融合图像在不同评价指标下的平均值如表 2 所示。

表 2 两种图像处理方式下的图像度量平均值

Table 2 Mean values of image metrics for the two image processing methods

方法	SSIM	PSNR	EN	SF	AG	MI
Original	0.855 6	30.193 9	6.351 8	11.600 9	5.203 8	2.866 5
本文	0.830 5	31.684 1	6.558 1	13.894 8	6.009 6	3.096 0

由表 2 可知,在本文提出的图像处理方式在 5 个关键指标(PSNR、EN、SF、AG、MI)上优于以往的图像处理方式。尤其是 SF 和 AG 这两个指标提升的幅度最大,这表明本文提出的图像处理方式有利于网络从源图像中提取纹理信息,使融合图像具有更加丰富的纹理细节。

2.3 消融实验

本文的网络结构中包括两个关键设计,即在生成器中引入了注意力模块和密集连接块。为了验证这两种设计

对融合性能的影响,本文进行了消融实验,实验分为 3 组,与提出的融合方法进行对比分析。

实验 A 在网络训练时,在生成器网络架构中引入密集连接块而不引入注意力模块。实验 B 在网络训练时,在生成器网络架构中引入了注意力模块而不引入密集连接块。实验 C 在网络训练过程中使用了生成器中的常规卷积层。3 个实验皆使用 M³FD 数据集进行训练和测试,且对 3 组实验以及本文提出方法的融合结果进行汇总及

评价。
表 3 比较了融合图像的平均评价指标。结果表明,本文的融合方法在 5 个关键指标上 (PSNR、EN、SF、AG、MI) 优于其他方法。这表明引入密集连接块和注意力模块可以有效增强网络从源图像中提取特征信息的能力,进一步提高融合图像的信息量。

表 3 不同融合算法的图像度量平均值
Table 3 Mean values of image metrics for different fusion algorithms

实验	SSIM	PSNR	EN	SF	AG	MI
A	0.845 1	30.617 3	6.387 8	12.811 3	5.774 2	2.580 1
B	0.837 5	30.553 3	6.415 3	12.588 9	5.514 0	2.630 0
C	0.782 6	30.672 1	6.253 9	9.846 8	4.353 1	2.805 5
本文	0.830 5	31.684 1	6.558 1	13.894 8	6.009 6	3.096 0

3 结 论

本文提出了一种基于注意力密集网络的伪彩色红外与可见光图像融合方法。与现有融合方法最大的不同之处在于本文的融合方法将红外图像转化成伪彩色图像再与可见光图像进行融合,使融合图像保留了更多的纹理细节。同时,为有效地从源图像中提取互补信息且得到的特征映射集中在源图像的关键信息上以进行增强融合性能,引入了带有注意力模块的密集连接网络。此外,本文设计了一种内容损失函数来保持融合图像中红外信息与纹理细节之间的平衡。在 M³FD 数据集上,从定量和定性层次分析,与其他几种融合方法相比,提出的方法在保留红外目标信息和纹理细节方面有一定的提升。同时,通过两组实验的对比(灰度的可见光图像与红外图像融合实验,伪彩色的红外与可见光图像融合实验)可知 SF 和 AG 提升的幅度最大,这验证了将红外图像转换成伪彩色图像再与可见光图像融合这一处理方式可使融合图像具有更丰富的纹理细节。

参 考 文 献

[1] 陈广秋, 温奇璋, 尹文卿, 等. 用于红外与可见光图像融合的注意力残差密集融合网络[J]. 电子测量与仪器学报, 2023, 37(8): 182-193.
CHEN G Q, WEN Q ZH, YIN W Q, et al. Attentive residual dense fusion network for infrared and visible image fusion[J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(8): 182-193.

[2] LIU Y, DONG L, JI Y, et al. Infrared and visible image fusion through details preservation [J]. Sensors, 2019, 19(20): 45-56.

[3] 冯鑫, 杨杰铭, 张鸿德, 等. 基于双通道残差密集网络的红外与可见光图像融合[J]. 光子学报, 2023, 52(11): 285-296.
FENG X, YANG J M, ZHANG H D, et al. Infrared and visible image fusion based on two-channel residual dense network [J]. Journal of Photonics, 2023, 52(11): 285-296.

[4] 王天元, 罗晓清, 张战成. 自注意力引导的红外与可见光图像融合算法[J]. 红外技术, 2023, 45(2): 171-177.
WANG T Y, LUO X Q, ZHANG ZH CH. Self-attention guided fusion algorithm for infrared and visible images [J]. Infrared Technology, 2023, 45(2): 171-177.

[5] ZHANG H, MA J. SDNet: A versatile squeeze-and-decomposition network for real-time image fusion[J]. International Journal of Computer Vision, 2021, 129: 2761-2785.

[6] 付涵, 严华. 基于多尺度变换和 VGG 网络的红外与可见光图像融合[J]. 现代计算机, 2021(16): 112-117.
FU H, YAN H. Infrared and visible image fusion based on multi-scale transform and VGG network[J]. Modern Computer, 2021(16): 112-117.

[7] LIU C H, QI Y, DING W R. Infrared and visible image fusion method based on saliency detection in sparse domain[J]. Infrared Physics & Technology, 2017, 83: 94-102.

[8] WU Z, LI J, XU J, et al. Subspace-based self-weighted multiview fusion for instance retrieval[J]. Information Sciences, 2022, 592: 261-276.

[9] 江兆银, 王磊. 基于显著性检测与权重映射的可见光与红外图像融合算法[J]. 电子测量与仪器学报, 2021, 35(1): 174-182.
JIANG ZH Y, WANG L. Visible and infrared image fusion algorithm based on significance detection and weight mapping [J]. Journal of Electronic Measurement and Instrumentation, 2021, 35(1): 174-182.

[10] 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.

- [11] XU H, ZHANG H, MA J. Classification saliency-based rule for visible and infrared image fusion[J]. IEEE Transactions on Computational Imaging, 2021, 7: 824-836.
- [12] LI H, CEN Y, LIU Y, et al. Different input resolutions and arbitrary output resolution: A meta learning-based deep framework for infrared and visible image fusion [J]. IEEE Transactions on Image Processing, 2021, 30: 4070-4083.
- [13] MA J, YU W, LIANG P, et al. FusionGAN: A generative adversarial network for infrared and visible image fusion[J]. Information Fusion, 2019, 48: 11-26.
- [14] ZHANG H, XU H, YANG X, et al. Rethinking the image fusion: A fast unified image fusion network based on proportional maintenance of gradient and intensity[C]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12797-12804.
- [15] MA J, TANG L, XU M, 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.
- [16] 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.
- [17] LI H, WU X J, DURRANI T. NestFuse: An infrared and visible image fusion architecture based on nest connection and spatial/channel attention models[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(12): 9645-9656.
- [18] MA J, XU H, JIANG 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.
- [19] LIU J, FAN X, HUANG Z, et al. Target-aware dual adversarial learning and a multi-scenario multi-modality benchmark to fuse infrared and visible for object detection [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 5802-5811.
- [20] REN L, PAN Z, CAO J, et al. Infrared and visible image fusion based on variational auto-encoder and infrared feature compensation[J]. Infrared Physics & Technology, 2021, 117: 103839.
- [21] XU X, SHEN Y, HAN S. Dense-FG: A fusion gan model by using densely connected blocks to fuse infrared and visible images [J]. Applied Sciences, 2023, 13(8): 4684.
- [22] 崔宝京, 程晓荣. 基于深度学习的紫外成像仪的图像配准与融合 [J]. 计算机仿真, 2023, 40(9): 151-157, 201.
- CUI B J, CHENG X R. Image alignment and fusion for UV imager based on deep learning[J]. Computer Simulation, 2023, 40(9): 151-157, 201.
- [23] WANG J X, XI X L, LI D M, et al. FusionGRAM: An infrared and visible image fusion framework based on gradient residual and attention mechanism [J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 1-12.
- [24] LI K, QI M, ZHUANG S, et al. TIPFNet: A transformer-based infrared polarization image fusion network[J]. Optics Letters, 2022, 47(16): 4255-4258.
- [25] 张薇, 汪雷丹, 刘晓亮, 等. 基于深度特征提取的红外与可见光图像融合[J]. 软件工程, 2023, 26(11): 52-58.
- ZHANG W, WANG L D, LIU X L, et al. Infrared and visible image fusion based on depth feature extraction[J]. Software Engineering, 2023, 26(11): 52-58.
- [26] LI J, HUO H, LI C, et al. AttentionFGAN: Infrared and visible image fusion using attention-based generative adversarial networks [J]. IEEE Transactions on Multimedia, 2020, 23: 1383-1396.

作者简介

漆建环, 硕士研究生, 主要研究方向为深度学习、图像处理。

E-mail: 202212490712@nuist.edu.cn

倪波(通信作者), 博士, 研究生导师, 主要研究方向为微纳光子学、图像处理、机器学习。

E-mail: bni@nuist.edu.cn

周晓彦, 博士, 副教授, 研究生导师, 主要研究方向为模式识别、图像处理、机器学习等。

E-mail: xiaoyan_zhou@nuist.edu.cn

倪海彬, 博士, 副教授, 研究生导师, 主要研究方向为图像处理、微结构光纤传感、纳米光电传感等。

E-mail: nihaibin@nuist.edu.cn

杨凌升, 博士, 教授, 研究生导师, 主要研究方向为图像处理、微波工程与天线、无人驾驶汽车雷达系统等。

E-mail: ylsinchina@163.com

常建华, 博士, 教授, 博士生导师, 主要研究方向为图像处理、光子学与光器件、全固态激光器、光电传感器等。