

DOI:10.19651/j.cnki.emt.2108051

基于深层次多分类生成对抗网络的 红外与可见光图像融合方法*

徐慧娴¹ 黄坤超² 陈明举^{1,3} 熊兴中¹ 田洋川¹(1. 四川轻化工大学人工智能四川省重点实验室 宜宾 644000; 2. 成都西测防务科技有限公司 成都 610000;
3. 四川轻化工大学企业信息化与物联网测控技术四川省高校重点实验室 宜宾 644000)

摘要: 为了实现红外和可见光图像信息的良好平衡,利用生成对抗网络技术,提出了一种深层次多分类的生成对抗网络红外与可见光图像融合方法。该方法将主辅思想引入到生成器的梯度和强度信息提取中,并提高了生成器卷积层的深度及浅层网络信息提取能力。在鉴别器中使用多分类器同时估计可见光和红外区域的分布。经过连续的对峙学习,使融合结果中具有显著的对比度和丰富的纹理细节信息。实验获得的信息熵及香农熵值为 6.86、互信息值为 13.72、标准差值为 34.82、结构相似性值为 0.71。对比实验结果表明,在主客观评价中,本文提出的方法获得更好的红外与可见光图像融合性能。

关键词: 生成对抗网络;红外图像;可见光图像;多分类

中图分类号: TP391.7 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Infrared and visible image fusion method based on deep multi-classification generative adversarial network

Xu Huixian¹ Huang Kunchao² Chen Mingju^{1,3} Xiong Xingzhong¹ Tian Yangchuan¹(1. Artificial Intelligence Key Laboratory of Sichuan Province, Sichuan University of Science and Engineering, Yibin 644000, China;
2. Chengdu Xice Defence Technology Co., Ltd., Chengdu 610000, China; 3. Key Laboratory of Higher Education of Sichuan Province for Enterprise Informationalization and Internet of Things, Sichuan University of Science and Engineering, Yibin 644000, China)

Abstract: In order to achieve a good balance between infrared and visible image information, proposes a deep multi-classification method of infrared and visible image fusion based on generative adversarial network technology. In this method, the idea of principal and auxiliary is introduced into the gradient and intensity information extraction of generator, and the depth and shallow network information extraction ability of generator convolution layer is improved. In the discriminator, multiple classifiers are used to estimate the distribution of visible and infrared regions simultaneously. After continuous face-off learning, the fusion results have remarkable contrast and rich texture details. The obtained information entropy and Shannon entropy value are 6.86, mutual information value is 13.72, standard deviation value is 34.82 and structural similarity value is 0.71. The experimental results show that the proposed method achieves better performance of infrared and visible image fusion in subjective and objective evaluation.

Keywords: generative adversarial network; infrared image; visible image; multi-classification

0 引言

为了提取和重建最有意义的信息,提高对场景描述的完整性和准确性,通常对不同传感器获得的红外与可见光图像进行融合。可见光图像由捕获反射光的可见光传感器

生成。其特点是纹理细节信息丰富,符合人眼观察规律。红外传感器可以感知红外波段,将热辐射信息转换成灰度图像。红外图像具有强烈的对比度,即使在夜间和恶劣天气下也能有效区分背景和目标。红外和可见光图像融合将这两种特征结合在一起,生成对比度显著、纹理细节丰富的

收稿日期:2021-10-10

* 基金项目:企业信息化与物联网测控技术四川省高校重点实验室开放基金(2021WYY01)、人工智能四川省重点实验室(2020RZY02)、四川轻化工大学研究生课程建设项目(KA202030)、四川省科技厅项目(2020JDJQ0075,2020YFSY0027)资助

图像,在军事监视、目标检测和车辆夜间导航等领域具有良好的应用前景^[1,2]。

传统方法在空间域或变换域测量像素或区域的活动水平,并根据特定的融合规则实现图像融合。Arif等^[3]利用曲波变换(curvelet transform, CVT)和遗传算法提出一种多模态医学图像融合的新方法,保持图像完美的数据信息和颜色一致性。张贵仓等^[4]利用双树复小波变换(dual-tree complex wavelet transform, DTCWT)得到低频子带图像和高频子带图像,再分别使用基于直觉模糊集和信息反差对比度的融合规则进行处理。王园园等^[5]提出一种基于Canny算子边缘检测的小波变换图像融合方法,对两张待融合图像使用小波变换进行分层,并使用不同的融合策略对低频和高频系数进行处理。传统算法的融合规则通常需要人工设计,这无法在各种类型的融合任务中实现鲁棒性。

基于深度学习的方法利用神经网络强大的非线性拟合能力,使融合图像具有期望的分布,进一步提高了融合图像的性能。卷积神经网络(convolutional neural network, CNN)通过对自然样本的学习训练,建立深层的结构神经网络,实现更深层次图像特征的提取,以实现融合图像结构信息完整与细节的保持^[6]。CNN在多焦距图像中获得较好的性能,但其运算耗时,融合的性能依赖于训练样本的特性。Ma等^[7]创新性地将GAN引入图像融合,从而更好地引导网络通过对抗性学习和特定内容丢失,在无监督的情况下保留显著的对比度和纹理细节。随后, Ma等^[8]基于FusionGAN使用多分类约束和专门设计的内容丢失,考虑源图像中包含的二次信息,生成具有良好视觉效果的融合图像。上述方法通过特定的内容丢失来提取源图像的信息,在网络结构的设计中对浅层网络的信息提取考虑不足,融合性能有待进一步提高。若能提高生成对抗网络在训练过程中对信息的利用程度,实现在保持红外和可见光图像信息的良好平衡,势必进一步提高图像融合的性能。

鉴于此,本文利用生成对抗网络技术兼顾考虑对比度信息和纹理细节的优点,建立一种深层次多分类的生成对抗网络(deep multi-classification generative adversarial networks, DMcGAN)红外与可见光图像融合方法。该方法将生成器的输入分为梯度路径和强度路径,构造了主辅信息损失函数,并提高了生成器卷积层的深度及浅层网络信息提取能力。在鉴别器中使用多分类器,它可以确定输入是红外图像和可见光图像的概率。通过约束这两个概率,经过连续的对峙学习,生成器可以同时拟合红外图像和可见光图像的概率分布,从而生成对比度显著且纹理细节丰富的结果。通过对比实验证明,提出的DMcGAN方法相比其他类似融合方法具有更好的融合性能。

1 生成对抗网络基本原理

原始生成对抗网络通过两个模块之间的相互博弈实现无监督的分布估计^[9]。网络中涉及到生成器 G 和鉴别器 D

两个模块,生成器专门用来产生可以愚弄鉴别器的假数据,而鉴别器则是用来区分生成器产生的假数据和真实数据。

假设输入网络的训练数据为 $\mathbf{X} = \{x_1, x_2, \dots, x_n\}$,服从特定的分布。生成器 G 估计 \mathbf{X} 的分布,并尽最大努力产生服从这种特定分布的假数据 $\mathbf{G}(\mathbf{X})$ 。那么鉴别器 D 需要学习区分真实训练数据 \mathbf{X} 和假数据 $\mathbf{G}(\mathbf{X})$ 。总之,生成对抗网络的目的是使假数据的分布 P_G 逐渐接近真实数据的分布 P_{data} ,通过式(1)中目标函数来实现。

$$\min_D \max_G E_{x \sim P_{data}} [\log \mathbf{D}(x)] + E_{x \sim P_G} [\log(1 - \mathbf{D}(\mathbf{G}(x)))] \quad (1)$$

作为对抗关系,生成器和鉴别器在不断的迭代训练中相互促进,不断提高自己的伪造或鉴别能力。当这两个分布之间的距离足够小时,鉴别器无法区分真实数据和虚假数据。那么可以说生成器已经成功地估计了训练数据的分布。

后续研究发现,原始生成对抗网络的训练过程非常不稳定,生成的图像质量不高。为了改善这一现象, Mao等^[10]提出用最小二乘损失函数代替交叉熵损失函数来指导生成对抗网络的优化。损失函数定义如下:

$$\min_D V_{LSGAN}(\mathbf{D}) = \frac{1}{2} E_{x \sim P_{data}} [(\mathbf{D}(\mathbf{X}) - a)^2] + \frac{1}{2} E_{x \sim P_G} [(\mathbf{D}(\mathbf{G}(\mathbf{X})) - b)^2] \quad (2)$$

$$\min_G V_{LSGAN}(\mathbf{G}) = \frac{1}{2} E_{x \sim P_G} [(\mathbf{D}(\mathbf{G}(\mathbf{X})) - c)^2] \quad (3)$$

其中, a 和 b 是指导鉴别器优化的概率标签。具体来说, a 是真实数据对应的概率标签, b 是生成器产生的假数据对应的概率标签。另外, c 是指导生成器优化的概率标签,也就是说 c 是生成器期望鉴别器判定假数据的标签。显然, b 应该尽可能接近0。相反, a 和 c 应该尽可能大,接近1。

最小二乘生成对抗网络所提出的内容丢失函数使用了主要信息和辅助信息的概念,可以从源图像中提取更充分的强度和梯度信息。其次,采用多分类器作为鉴别器,同时估计可见光和红外两个不同区域的分布。由于概率分布的一致性会使融合结果具有目标分布的最显著特征,因此生成器可以生成同时具有红外和可见光特征的融合结果,即显著的对比度和丰富的纹理细节。若应用于红外与可见光图像的融合中,势必获得更好的融合性能。

2 基于深层次多分类生成对抗网络的红外与可见光图像融合

2.1 深层次多分类生成对抗网络总体框架

为了同时保留红外图像中显著的对比度信息和丰富的可见光图像细节纹理信息,提出了一种新的红外与可见光图像融合的深层次多分类约束端到端GAN模型,整个算法融合框架如图1所示。

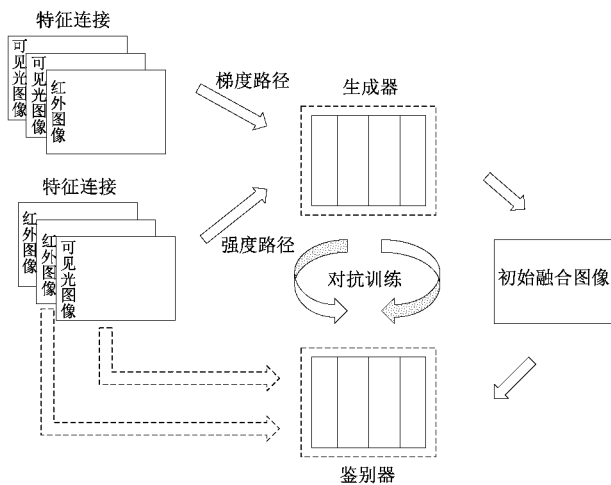


图 1 算法框架图

针对可见光图像也有显著的对比度,红外图像中也包含一些结构信息的现象,提出了主轴信息的思想。一方面,设置相应的内容损失来充分提取图像中有价值的信息,其中纹理信息由梯度表示,对比度信息由强度表示。另一方面,将生成器的输入分为梯度路径和强度路径。对于梯度路径,沿着通道维度连接两个可见图像和一个红外图像作为输入;对于强度路径,沿着通道维度连接两个红外图像和一个可见图像作为输入。在这种特定的内容损失和网络设计下,可以驱动网络不均衡地提取梯度和强度信息,生成器可以获得来自可见光图像的主梯度和次强度信息以及来自红外图像的主强度和次梯度信息。以上信息可以相互补充。

其次,由于融合任务需要红外和可见光两种图像特征之间的博弈,在鉴别器中使用多分类器。鉴别器的输出是 1×2 的概率向量,表示输入图像是可见图像的概率 P_{vis} 和输入图像是红外图像的概率 P_{ir} 。鉴别器判别融合图像

时,生成器希望两个概率都大,也就是说让鉴别器判定融合图像既是可见光图像又是红外图像。而鉴别器的任务就是将融合图像精确地确定为假数据,也就是说,使两个概率同时变小。这样,在生成器和鉴别器之间建立了一个对抗博弈。当鉴别器判定融合图像的概率 P_{vis} 和 P_{ir} 都较小时,即可得到信息均衡的融合图像。

通过以上设计,本文方法可以生成良好的融合结果,不仅具有显著的对对比度,而且包含丰富的纹理细节。

2.2 深层次多分类生成对抗网络结构

整体的网络结构由生成器网络和鉴别器网络构成。

在生成器的网络结构中,分为梯度路径和强度路径进行信息提取。由于纹理信息主要包含在在可见光图像中,其次是红外图像。对于梯度路径,使用两幅可见光图像和一幅红外图像沿着通道连接作为输入,主要负责提取纹理细节信息,即高频特征。同样,强度路径主要负责对比度信息的提取,而对比度信息主要包含在红外图像中,其次是可见图像。因此使用两幅红外图像和一幅可见光图像沿着通道连接作为输入。

为了提高模型训练的准确性,提高了 GAN 网络卷积层的深度,在每个信息提取路径中,采用 8 个卷积层进行特征提取。卷积层采用 3×3 卷积核 ($3 \times 3 \times 3 \times 128$ 代表输入通道数为 3,输出通道数为 128, 3×3 为卷积核的大小),并且每一层都具有一个 Batch-Norm 层和 Leaky-ReLU 激活函数层。值得注意的是,卷积层第 2 层的输出与第 3 层的输出进行特征拼接后再作为第 4 层的输入,第 4 层的输出与第 5 层的输出进行特征拼接后再作为第 6 层的输入。这一结构可以尽可能多地保留浅层网络提取地特征,有效减少只用一层输出结果可能导致的信息丢失。最后,为了完全合并信息,沿着通道交叉连接两个特征图。并在最后一层使用 1×1 的卷积核和 Tanh 激活函数输出网络结果,步幅设置为 1。生成器网络结构如图 2 所示。

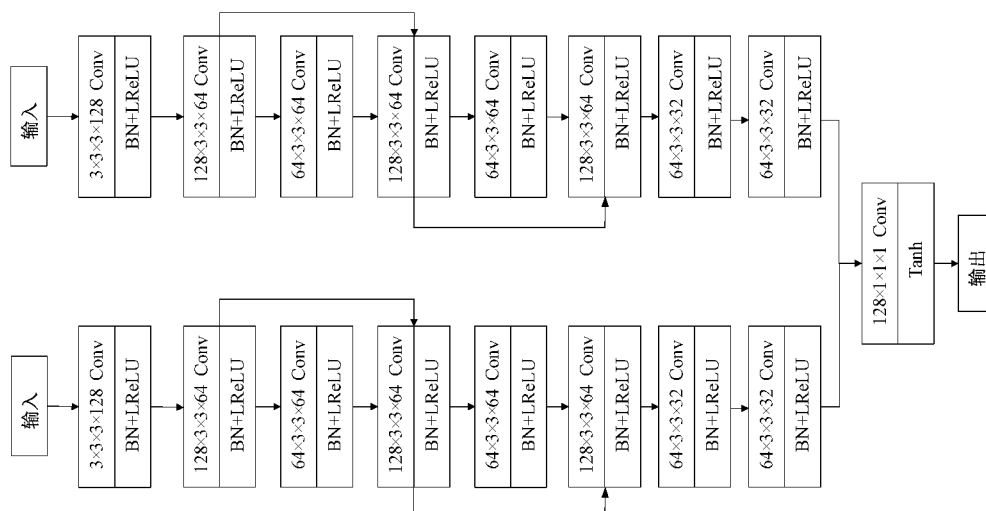


图 2 生成器网络结构

鉴别器网络结构如图3所示,鉴别器本质上是一个多分类器,它可以估计输入图像的每个类别的概率。它的输出是大小为 1×2 的概率向量。该鉴别器由4个卷积层和1个线性层组成。4个卷积层使用 3×3 卷积核和Leaky-ReLU激活函数,后3个也添加了Batch-Norm层进行批处理归一化,所有卷积层的步幅为2。最后一个线性层根据前4个卷积层提取的特征鉴别输入,输出分类概率。

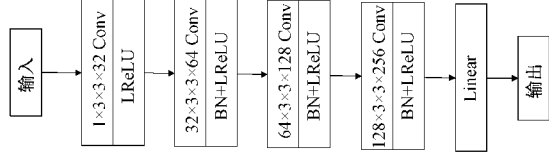


图3 鉴别器网络结构

2.3 损失函数

本文的损失函数主要包括生成器损失函数 L_G 及鉴别器损失函数 L_D 。下面将分别对这两部分进行介绍。

1)生成器的损失函数:生成器损失函数由两部分组成,即约束信息提取的内容损失 $L_{content}$ 和约束信息平衡的对抗损失 L_{GAN} 。具体公式如式(4)所示。

$$L_G = L_{GAN} + \gamma L_{content} \quad (4)$$

式中: γ 是负责维持两项之间平衡的正则化参数,设置为100。

对于红外图像,其主要特点是反映场景的对比度信息,能够从背景中突出目标。因此,主要信息是其强度分布,主要强度损失定义为:

$$L_{intensity_{main}} = \| \mathbf{I}_{fused} - \mathbf{I}_{ir} \|_F^2 \quad (5)$$

式中: \mathbf{I}_{fused} 为融合图像, \mathbf{I}_{ir} 为红外源图像。至于可见光图像,包含丰富的纹理细节,符合人眼的观察习惯。因此,从可见光图像获得的主要信息是其梯度信息,主要梯度损失定义为:

$$L_{grad_{main}} = \| \nabla \mathbf{I}_{fused} - \nabla \mathbf{I}_{vis} \|_F^2 \quad (6)$$

式中: ∇ 为二阶梯度算子, \mathbf{I}_{vis} 为可见光图像。如上所述,红外图像也有一些纹理细节,可见光图像也包含对比度信息。因此提出了辅助损失的概念。也就是说,在融合图像和可见光图像之间构造一个辅助强度损失 $L_{intensity_{aux}}$,如式(7)所示;在融合图像和红外图像之间构造一个辅助梯度损失 $L_{grad_{aux}}$,如式(8)所示。

$$L_{intensity_{aux}} = \| \mathbf{I}_{fused} - \mathbf{I}_{vis} \|_F^2 \quad (7)$$

$$L_{grad_{aux}} = \| \nabla \mathbf{I}_{fused} - \nabla \mathbf{I}_{ir} \|_F^2 \quad (8)$$

综上所述,内容损失由4部分组成,即主强度损失、主梯度损失、辅助梯度损失和辅助强度损失,如式(9)所示。

$$L_{content} = L_{intensity_{main}} + L_{grad_{main}} + L_{intensity_{aux}} + L_{grad_{aux}} = \alpha_1 \| \mathbf{I}_{fused} - \mathbf{I}_{ir} \|_F^2 + \alpha_2 \| \nabla \mathbf{I}_{fused} - \nabla \mathbf{I}_{vis} \|_F^2 + \alpha_3 \| \mathbf{I}_{fused} - \mathbf{I}_{vis} \|_F^2 + \alpha_4 \| \nabla \mathbf{I}_{fused} - \nabla \mathbf{I}_{ir} \|_F^2 \quad (9)$$

式中: $\alpha_{(,)}$ 是一个常数,对其进行调整,可以实现这些项之间的主次关系。此外,梯度损失项在一般情况下小于强度损失项,需要调整 $\alpha_{(,)}$,使它们在优化过程中同等重要。因

此, $\alpha_{(,)}$ 的设置规则可以概括为:

$$\alpha_1 > \alpha_3, \alpha_2 > \alpha_4, \{\alpha_2 > \alpha_4\} > \{\alpha_1 > \alpha_3\} \quad (10)$$

式中: α_1 设置为1, α_2 设置为5, α_3 设置为0.3, α_4 设置为4。

为了实现各种信息之间的平衡,在生成器的损失函数中引入带有鉴别器的对抗性损失,可以定义为:

$$L_{GAN} = (\mathbf{D}(\mathbf{I}_{fused}^n)[1] - d)^2 + \mathbf{D}(\mathbf{I}_{fused}^n)[2] - d)^2 \quad (11)$$

式中: d 是鉴别器判定融合图像的概率标记。鉴别器是一个输出 1×2 概率向量的多分类器。因此, $\mathbf{D}(\cdot)[1]$ 表示向量的第1项,即融合图像是可见光图像的概率。 $\mathbf{D}(\cdot)[2]$ 表示向量的第2项,即融合图像是红外图像的概率。值得注意的是,对两种概率使用相同的标签 d ,因此鉴别器判定融合图像是红外图像还是可见光图像的概率相同。 d 设置为1,因为生成器预期鉴别器不能区分融合图像和真实数据。

2)鉴别器的损失函数:鉴别器是一个多分类器,其损失函数必须不断提高其鉴别能力,使鉴别器能够有效地识别红外和可见光图像。鉴别器的损失函数 L_D 由3部分组成,即可见光图像、红外图像和融合图像的损失,分别表示为 L_{vis} 、 L_{ir} 和 L_{fused} ,如式(12)所示。

$$L_D = L_{vis} + L_{ir} + L_{fused} \quad (12)$$

考虑鉴别器输出的是 1×2 矢量,有 $P_{vis} = \mathbf{D}(x)[1]$ 和 $P_{ir} = \mathbf{D}(x)[2]$ 。当输入图像是可见光图像时,预计 P_{vis} 应该接近1, P_{ir} 接近0。相应的损失定义为:

$$L_{vis} = \frac{1}{2N} \sum_{i=1}^N ((P_{vis}(\mathbf{I}_{vis}^n) - \beta_1)^2 + (P_{ir}(\mathbf{I}_{vis}^n) - \beta_2)^2) \quad (13)$$

式中: β_1 和 β_2 是概率标签, β_1 设置为1, β_2 设置为0。也就是说,当输入可见光图像时,鉴别器想要判断是可见光图像的概率大,红外图像的概率小。

类似地,红外图像损失项定义为:

$$L_{ir} = \frac{1}{2N} \sum_{i=1}^N ((P_{vis}(\mathbf{I}_{ir}^n) - \delta_1)^2 + (P_{ir}(\mathbf{I}_{ir}^n) - \delta_2)^2) \quad (14)$$

式中: δ_1 设置为0, δ_2 设置为1。

最后,当输入图像是融合图像时,损失函数定义为:

$$L_{fused} = \frac{1}{2N} \sum_{i=1}^N ((P_{vis}(\mathbf{I}_{fused}^n) - \epsilon)^2 + (P_{ir}(\mathbf{I}_{fused}^n) - \epsilon)^2) \quad (15)$$

式中: ϵ 是鉴别器判定融合图像的概率标签,设置为0。同样,我们也对两种概率使用相同的标签 ϵ 来实现平衡。也就是说,在鉴别器看来,融合后的图像既是伪可见光图像,也是伪红外图像。

3 实验结果分析

3.1 实验设置

本文方法的实验运行在PyCharm 2019及MATLAB

R2018b 上,硬件环境为 Intel-Xeon CPU/2.10 GHz/256 GB。将图像裁剪成多个 120×120 的图像块,然后填充到 132×132 。生成器的输入是 132×132 个图像块,每次选择 32 对红外与可见光图像对送入网络训练,训练数的比值为 1/2,训练迭代次数设置为 10,初始学习率设置为 0.000 1,并采用 Adam 作为优化器训练网络。实验选用公开 TNO 数据集^[11]中红外与可见光图像进行实验。

客观评价选取信息熵 (entropy, EN)^[12]、互信息 (mutual information, MI)^[13]、差异相关系数的总和 (the sum of the correlations of differences, SCD)^[14]、香农熵 (shannon entropy, SEN)^[15]、标准差 (standard deviation,

STD)^[16]、结构相似性 (structural similarity, SSIM)^[17] 6 个质量指标。各项指标主要用于衡量融合图像的清晰程度、包含信息量的丰富程度以及边缘或视觉信息的保持度,值越大,说明融合图像的融合性能越好。

3.2 TNO 数据集上的融合结果及对比分析

为证明本文 DMcGAN 算法的优越性,选择与本文算法相近的 5 种代表性图像融合算法进行对比,它们分别是低通金字塔 (radio of low-pass pyramid, RP)^[18]、曲波变换 (CVT)、双树复小波变换 (DTCWT)、FusionGAN、GANMcC。给出 TNO 数据库中 21 对源图像融合后的 6 个质量指标的平均值如表 1 所示。

表 1 融合图像的 6 个质量指标的平均值

方法	EN	MI	SCD	SEN	STD	SSIM
RP	6.451 7	12.903 4	1.425 5	6.451 6	33.51	0.666 4
CVT	6.589 2	13.180 0	1.552 8	6.589 9	32.51	0.703 3
DTCWT	6.537 8	13.075 6	1.552 5	6.537 9	31.95	0.710 7
FusionGAN	6.580 1	13.160 0	1.266 5	6.580 0	33.91	0.620 9
GANMcC	6.730 6	13.460 0	1.628 3	6.730 6	33.03	0.703 5
DMcGAN	6.862 4	13.724 9	1.618 8	6.862 4	34.82	0.714 1

表 1 中数据表明,本文提出的 DMcGAN 在 6 个客观质量评价指标中获得了 5 个最高值,1 个次最高值。其中,虽然 SCD 值仅是次最高值,但可以看出,已明显超过其他方法尤其是 FusionGAN 的 SCD 值,这说明本文方法较于 FusionGAN 可以更好地保持红外和可见光图像信息的良好平衡。另外,本文方法获得的 STD 值远大于所有其他方法,从而表明本文方法可以更好地融合红外与可见光图像的对比度及纹理细节信息,融合后的图像信息更丰富。

不同方法实现红外与可见光图像融合前后的结果如

图 4 所示,并在图中给出局部放大图。从图 4 中可以看出,RP、CVT、DTCWT 在图像的对比度显示上有很大的缺陷,图中的人物均被虚化,边缘存在一定程度的伪影,并且 RP 融合方法噪声比较明显,清晰度较差。FusionGAN 融合方法获得的图像尽管对比度信息丰富,但纹理细节信息丢失严重。在局部放大图中,只有本文提出的 DMcGAN 不仅可以保留丰富的纹理细节,如路灯周围的树枝、旁边建筑物的 3 个窗口,并且可以保留丰富的对比度信息。总体来看,本文方法清晰度与对比度有明显的优势,可以实现红外和可见光图像信息的良好平衡。

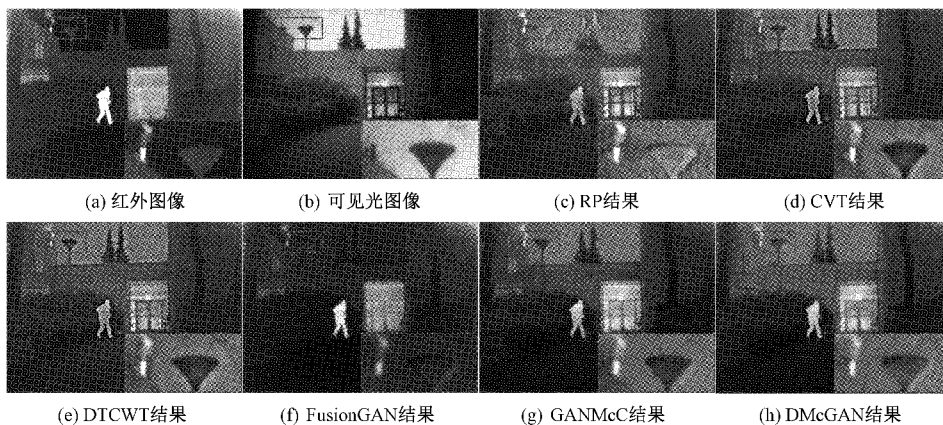


图 4 “路灯”图像集各融合结果

3.3 RoadScene 数据集上的融合结果及对比分析

为了进一步证明本文 DMcGAN 算法的优越性,使用 RoadScene 数据集^[19]进行泛化实验,并与相关 5 种图像融合算法进行对比,他们分别是梯度金字塔 (gradient

pyramid, GP)、低通金字塔 (RP)、小波变换 (Wavelet)^[20]、FusionGAN、GANMcC。给出 RoadScene 数据库中 10 对源图像融合后的 6 个质量指标的平均值如表 2 所示。

表2 融合图像的6个质量指标的平均值

方法	EN	MI	SCD	SEN	STD	SSIM
GP	6.770 7	13.541 4	1.623 0	6.771	30.030	0.615 0
RP	6.781 3	13.562 7	1.537 0	6.781	32.190	0.664 3
Wavelet	6.642 0	13.284 0	1.611 3	6.642	28.398	0.731 9
FusionGAN	7.114 9	14.227 6	1.495 4	7.110	39.110	0.607 6
GANMcC	7.058 0	14.116 2	1.789 5	7.058	38.110	0.669 6
DMcGAN	7.125 8	14.251 6	1.722 4	7.126	39.240	0.679 5

表2中数据表明,本文提出的DMcGAN仍然在EN、MI、SEN和STD上实现了最佳性能,至于SSIM和SCD,我们的方法排名第2。综合数据来看,本文方法获得更好的融合效果,可以保持红外和可见光图像信息的良好平衡。

RoadScene数据集中红外与可见光图像融合前后的结果如图5所示,并在图中给出局部放大图。RP融合方法失真严重,局部放大图中的山峰已经很难辨认,而GP、Wavelet对比度较低,并没有体现出显著目标的特征。

FusionGAN丢失了较多纹理细节,且边界伪影较为严重,使得山峰及云的边界范围扩大,整体较为模糊。从图5中可以看出,DMcGAN获得图像的对比度明显优于GANMcC,图中的屋顶及车子等目标更为显著。通过使用RoadScene数据集进行泛化实验,可以进一步证明本文方法在保留图像细节纹理及对比度信息时较其它几种算法有比较明显的优势,可以实现红外和可见光图像信息的良好平衡。

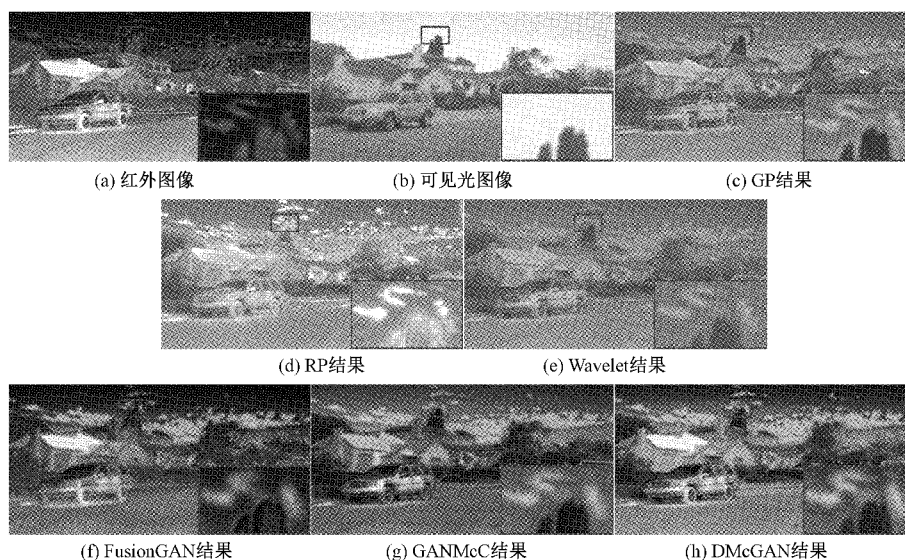


图5 “房屋”图像集各融合结果

3.4 消融实验

为了验证损失函数的有效性,进行相关的消融实验,包括内容损失项和对抗性损失项。具体来说,首先仅使用内容丢失来指导生成器的优化。然后,消除内容丢失,并仅通过对抗性丢失来训练网络。最后,将对抗性损失和内容损失结合起来,共同约束网络的优化。实验结果如图6所示。

从图6中可以看出,当仅存在内容丢失时,所生成融合图像的纹理细节不够丰富,如花丛的树叶、路灯周围的树枝均丢失了部分纹理细节。当仅存在对抗性丢失时,融

合结果显示了红外和可见光图像两种分布之间的博弈。尽管融合结果具有一定的对比度和纹理细节,但它们与源图像不匹配,并且与预期相差甚远。将内容丢失和对抗性丢失相结合,生成的结果不仅对比明显,纹理细节丰富,而且很好地符合源图像的特征。如图中人物对比度明显,非常类似于红外图像,而地面及花丛等地方与可见光图像中的纹理细节信息大体一致。由以上分析可知,设计的内容损失和对抗性损失是互补的,它们共同工作,能产生更好的融合结果。

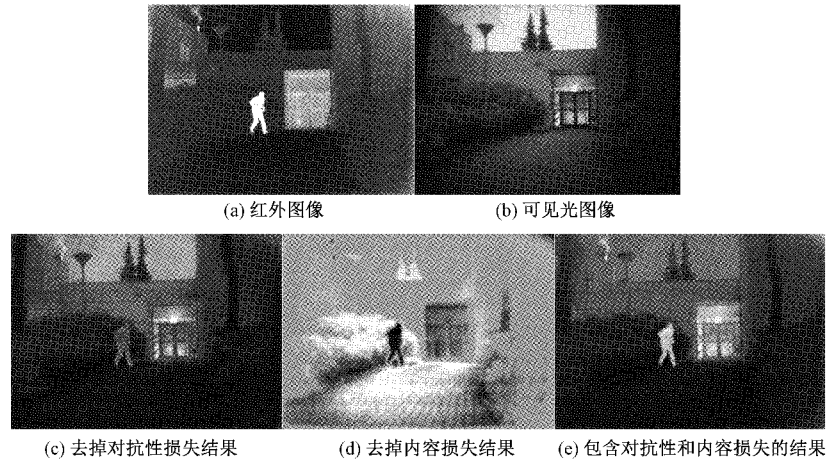


图 6 消融实验结果

4 结 论

本文利用生成对抗网络技术兼顾考虑对比度信息和纹理细节的优点,提出了一种深层次多分类的生成对抗网络红外与可见光图像融合方法。该方法在运用生成对抗网络进行模型训练的基础上,将主辅思想引入到生成器的梯度和强度信息提取中,进一步提高了生成器卷积层的深度及浅层网络信息提取能力。在鉴别器中使用多分类器,同时估计可见光和红外区域的分布,经过连续的对峙学习,使融合结果中具有显著的对比度和丰富的纹理细节信息。通过对比实验证明提出的 DMcGAN 更好地融合图像的对比度和纹理细节信息,具有更好的红外与可见光图像融合性能,研究的结论具有较大的应用价值。此外,提高本文提出方法在其余数据集上的适应性,是后续的研究目标。

参考文献

- [1] MA J, MA Y, LI C. Infrared and visible image fusion methods and applications: A survey[J]. Information Fusion, 2019;153-178.
- [2] 蒋永光,张倩武. 基于 OpenCV2 的红外热源和可见光图片的图像处理应用[J]. 电子测量技术, 2019, 42(14):131-136.
- [3] ARIF M, WANG G. Fast curvelet transform through genetic algorithm for multimodal medical image fusion[J]. Soft Computing, 2020, 24(2):1-22.
- [4] 张贵仓,苏金凤,拓明秀. DTCWT 域的红外与可见光图像融合算法[J]. 计算机工程与科学, 2020, 42(7):1226-1233.
- [5] 王园园,张娜,韩美林. 基于 Canny 算子边缘检测的小波变换多聚焦图像融合方法[J]. 电子测量技术, 2019, 42(17):148-151.
- [6] LIU Y, CHEN X, PENG H, et al. Multi-focus image fusion with a deep convolutional neural network[J]. Information Fusion, 2017, 36:191-207.
- [7] MA J, WEI Y, LIANG P, et al. FusionGAN: A generative adversarial network for infrared and visible image fusion[J]. Information Fusion, 2019, 48:11-26.
- [8] MA J, ZHANG H, SHAO Z, et al. GANMcC: A generative adversarial network with multiclassification constraints for infrared and visible image fusion[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70:1-14.
- [9] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[J]. Advances in Neural Information Processing Systems, 2014, 3:2672-2680.
- [10] MAO X, LI Q, XIE H, et al. Least squares generative adversarial networks [C]. 2017 IEEE International Conference on Computer Vision (ICCV), IEEE, 2017.
- [11] TOET A. TNO Image fusion dataset [DB/OL]. Figshare, 2014, [https://figshare.com/articles/TNO/Image/Fusion/Data-setYang](https://figshare.com/articles/TNO_Image/Fusion/Data-setYang).
- [12] 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.
- [13] KINNEY J B, ATWAL G S. Equitability, mutual information, and the maximal information coefficient[J]. Proceedings of the National Academy of Sciences, 2014, 111(9):3354-3359.
- [14] 童林,官铮,杨文韬,等. 潜在低秩表示下 VSM 联合 PCNN 的红外与可见光图像融合[J]. 国外电子测量技术, 2021, 40(10):84-90.
- [15] XIE W, SCHVITZ M D, LISTER R, et al. Epigenomic analysis of multilineage differentiation of human embryonic stem cells[J]. Cell, 2013, 153(5):

- 1134-1148.
- [16] 刘峰,沈同圣,马新星. 交叉双边滤波和视觉权重信息的图像融合[J]. 仪器仪表学报, 2017, 38(4): 1005-1013.
- [17] 李春艳,孙韬,谢俊峰. EMF 深度学习可见光/红外图像融合算法[J]. 国外电子测量技术, 2020, 39(10): 25-32.
- [18] 陈荣元,谢伟,王四春,等. 数据同化框架下红外与可见光图像的可控融合[J]. 光电子. 激光, 2010, 21(4): 610-613.
- [19] XU H, MA J, JIANG J, et al. U2Fusion: A unified unsupervised image fusion network [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 44(1): 502-518.
- [20] 刘进辉,王峰,王宏武,等. 基于小波变换的像素级图像融

合实例研究[J]. 电脑知识与技术, 2019, 15(17): 201-203.

作者简介

徐慧娴, 硕士, 主要从事图像融合方面的研究。

E-mail: 571001172@qq.com

黄坤超, 正高级工程师, 主要研究方向为测量技术与仪器。

E-mail: 12347259@qq.com

陈明举, 博士, 主要从事智能信息处理、图像处理和机器学习方面的研究。

E-mail: chenmingju@livc.cn

熊兴中, 博士, 主要从事人工智能、设备状态监测及健康管理、物联网技术、无线通信技术、VLSI 设计等方面的教学及科研工作。

E-mail: 2554438@qq.com

田洋川, 硕士, 主要从事多聚焦图像融合方面的研究。

E-mail: 917031283@qq.com