文章编号:1001-5078(2024)07-1157-07　　· 图像与信号处理 ·

基于改进 pix2pix 的红外图像转换技术

叶明亮[1],史春景[1],郝永平[2],李大伟[1]
(1. 沈阳理工大学机械工程学院,辽宁 沈阳 110159;2. 沈阳理工大学装备工程学院,辽宁 沈阳 110159)

摘　要:针对不同波段图像获取代价不同的问题,提出一种基于 pix2pix 的图像转换方法并进行改进。主要针对生成器和鉴别器两方面进行改进。生成器方面,使用残差结构的生成器替换原来的 U-Net 生成器以缓解梯度消失问题;引入可变形卷积,提高目标边缘和小目标的生成效果;引入 BAM 注意力机制,提高了算法对图像中主要目标的特征提取能力以提升生成图像的效果。鉴别器方面:改变 PatchGAN 中卷积层的层数(原 PatchGAN 为 3 层卷积),设置对照实验找到转换效果最好的卷积层数。以可见光图像和红外图像之间的转换为例进行实验。实验结果表明,改进后的算法在生成图像上的均方根误差(MSE)下降了 31.4%、结构相似性(SSIM)提高了 11.2%,可以更好的实现红外图像和可见光图像之间的转换。

关键词:生成对抗网络;pix2pix;图像转换;残差结构
中图分类号:TP391.41;TN29　　**文献标识码:**A　　**DOI:**10.3969/j.issn.1001-5078.2024.07.024

Infrared image conversion technology based on improved pix2pix

YE Ming-liang[1], SHI Chun-jing[1], HAO Yong-ping[2], LI Da-Wei[1]
(1. School of Mechanical Engineering, Shenyang Ligong University, Shenyang 110159, China;
2. School of Equipment Engineering, Shenyang Ligong University, Shenyang 110159, China)

Abstract: In order to solve the problem of different cost of image acquisition in different light segments, an image conversion method based on pix2pix was proposed. It mainly focuses on the generator and discriminator. In terms of generators, the residual structures generator was used instead of the original U-Net generator to alleviate the gradient vanishing problem. Deformable convolution is introduced to improve the generation effect of target edges and small targets. The BAM attention mechanism is introduced to improve the feature extraction ability of the algorithm for the main target in the image to improve the image generation effect. In terms of discriminators: change the number of convolutional layers in PatchGAN (the original PatchGAN is 3-layer convolution), and set up a control experiment to find the convolutional layer with the best conversion effect. Some KAIST datasets are selected for training and testing. The experimental results show that the Root Mean Square Error (MSE) of the improved algorithm is reduced by 31.4% and the Structural Similarity (SSIM) is increased by 11.2%, which can better realize the conversion between infrared and visible images.

Keywords: generative adversarial network; pix2pix; image conversion; residual structures

作者简介:叶明亮(2000-),男,满族,硕士研究生,研究方向为图像处理。E-mail:1528216489@qq.com
通讯作者:郝永平(1960-),男,博士,教授,研究方向为光电装备智能探测。E-mail:yphsit@126.com
收稿日期:2023-11-15;**修订日期:**2023-12-26

1 引 言

图像转换(Image-to-Image Translation)是指建立从输入到输出图像的映射,将一张输入图像经过特定的变换得到相应的输出图像的过程[1]。传统意义上的图像转换只研究图片纹理生成,主要是用的一些统计模型和复杂的数学公式来描述和生成图像局部纹理特征[2]。由于这种图像转换方法只能提取图像的底层特征,而非高层抽象特征,在处理颜色和纹理较复杂的图像时,生成的图像效果粗糙。2014 年 Goodfellow 等人提出了一种通过对抗训练生成图像模型的新框架,即生成对抗网络(GAN),该网络包括生成器模型和鉴别器模型两个相互对抗的模型[3]。生成器模型用于拟合样本数据分布,鉴别器模型用于估计输入样本是否是真实的训练数据。GAN 的提出显著提高了计算机绘制图像的真实性,相关理论迅速发展[4]。

目前,实现可见光图像与红外图像的转换通常是利用反演的方式,也就是通过寻找同一目标可见光图像与红外图像的映射关系,得到同样条件下的反演关系[5]。其中典型的就是图像转换任务使用的统一框架 pix2pix[6]。pix2pix 采用 U 型网络(U-Net)作为生成器,可以有效地结合底层和高层信息;提出了马尔科夫鉴别器 PatchGAN,可以分块的对图像进行鉴别。然而,随着计算机视觉领域的发展,pix2pix 的转换效果已经难以达到要求。本文旨在提出一种改进的 pix2pix 以提升图像转换的效果。主要是对 pix2pix 的生成器和鉴别器两部分进行改进。

2 生成器的改进

2.1 生成器替换为 ResNet 生成器

pix2pix 的生成器部分采用的是 U 型网络(U-Net)。如图 1 所示,U-Net 左侧网络为特征提取网络,使用卷积和池化操作。右侧网络为特征融合网络,使用上采样生成特征图,并与左侧特征图进行拼接操作。原作者之所以会选择 U-Net 生成器主要是因为以下两点,首先,U-Net 的浅层网络关注图像的纹理特征,深层网络关注图像的本质特征,故而,U-Net 生成器能够有效结合底层和高层的信息[3];另外,U-Net 生成器可以通过特征的拼接,实现边缘特征的找回。不过 U-Net 生成器也有一些缺点,第一,冗余太大,这导致网络训练的很慢;第二,感受野和定位精度不可兼得。U-Net 生成器的架构如图 1 所示。

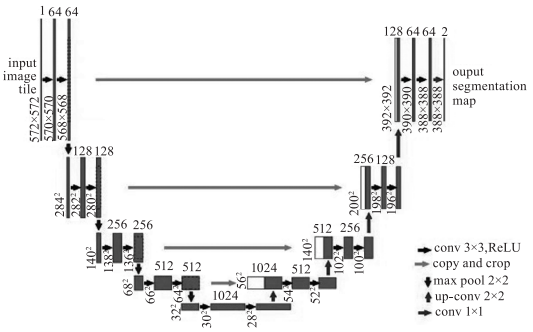

图 1 U-Net 架构

Fig. 1 U-Net architecture

由于本文针对的是图像转换任务,需要生成的图像与真实图像越接近越好,对结构相似性(SSIM)的要求较高。所以 U-Net 生成器并不能很好的满足本文需求,而 ResNet 结构的生成器能够在深度的卷积中缓解梯度消失问题,生成图像的结构相似性(SSIM)较高,能够更好的完成图像转换任务,故而生成器部分采用 ResNet 生成器。

在 ResNet 被提出之前,人们认为神经网络的层数越深,学习的效果就越好。但是在后来的实验中发现随着网络层数的加深,会出现梯度消失、梯度爆炸和退化等问题[7]。ResNet 结构通过隔层相连的方法弱化每层之间的强联系,也就是残差结构(residual 结构)来减轻退化问题。residual 结构如图 2 所示。

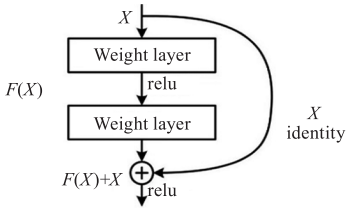

图 2 Residual 结构

Fig. 2 Residual structure

ResNet 生成器具有以下优点:可以训练非常深的神经网络,避免了梯度消失问题,提高了模型的表达能力和性能;使用残差连接,可以保留原始特征。并且 ResNet 生成器生成图片的结构相似性(SSIM)更高,更适合本文图像转换任务。故而将 pix2pix 中的 U-Net 生成器替换为 ResNet 生成器。本文的 ResNet 生成器使用 9 个残差块。

将 U-Net 生成器替换为 ResNet 生成器,对于图

像转换的效果提升如图 3 所示。

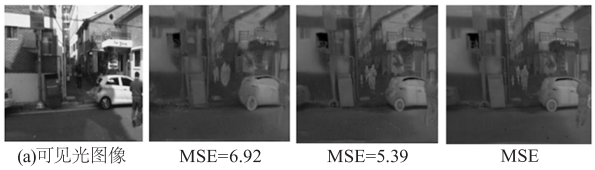

图 3 转换效果提升展示

Fig. 3 Conversion effect enhancement is displayed

如图 3 所示,将生成器替换为 ResNet 生成器后。从主观上来看,图中的一些重要目标,比如人的轮廓更加清晰,车体和车轮的畸变现象也得到改善,生成的红外图片更接近真实红外图片。从客观上来看。图像的均方根误差(MSE)有明显缩小,结构相似性(SSIM)有明显提升。所以,对于图像转换任务,ResNet 生成器明显优于 U-Net 生成器。

2.2 引入可变形卷积

将 pix2pix 的生成器更换为 ResNet 生成器后,依然存在着对小目标生成效果差,重要目标的边缘轮廓不清晰、不准确等问题。这是由于 ResNet 生成器使用的传统卷积操作采用的卷积核形状是固定的,采样过程中极易发生核内背景像素占比大于特征像素占比的现象。这会造成对重要目标的边缘轮廓和小目标的生成效果差的问题。而可变形卷积的采样位置是可以变化的,因此可变形卷积可以更好的生成图像中的小目标和重要目标[8]。

可变形卷积的原理是基于一个网络学习 offset(偏移),使得卷积核在输入特征图的采样点发生偏移,集中于我们感兴趣的区域或者重要目标[9]。它和传统卷积的采样方式如图 4 所示。

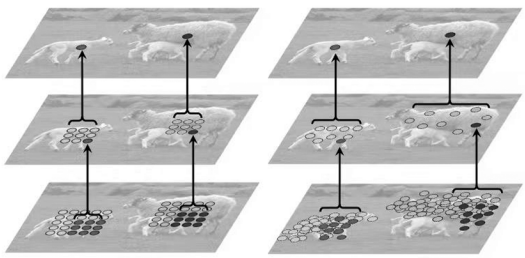

图 4 传统卷积与可变形卷积的采样方式

Fig. 4 Sampling methods of traditional convolution and deformable convolution

通过图 4 的左右对比可以明显的看出,可变形卷积的采样位置更符合物体本身的形状和尺寸,而标准卷积的形式却不能做到这一点。可变形卷积顶层特征图中最终的特征点学习了物体的整体特征,这个特征只针对于物体本身,相比原始的卷积它更能排除背景噪声的干扰,得到更有用的信息。

以普通的 3 × 3 卷积为例,卷积核的 9 个位置为:

$$R = \{(-1,-1),(-1,0),\cdots,(0,1),(1,1)\} \tag{1}$$

所以传统卷积的输出就是:

$$y(P_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n) \tag{2}$$

式中,x 为输入特征图;w 为加权的采样值总和;p_0 为输出特征图上的每个位置;p_n 为卷积核 R 上的每个位置。可变形卷积中,对规则的网格 R 引入了偏移量,式(2) 变成:

$$y(P_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n + \Delta p_n) \tag{3}$$

式中,Δp_n 为偏移量。

偏移量是由输入特征图与一个额外使用的卷积生成的,这个额外使用的卷积要与后面要进行的可变形卷积具有相同的尺寸、步长和扩张率。用这个额外使用的卷积分别学习 x 方向和 y 方向的偏移量。在基于这个偏移量进行可变形卷积操作。使用此方法可以根据不同的输入特征图得到不同的偏移量。

采用可变形卷积的方法可有效的提升小物体像素在卷积核中的占比,进而提高网络对目标边缘和小目标的生成能力,但是由于在卷积核中引入了偏移量,若将网络内部卷积核全部替换为可变形卷积核,将极大的增加网络计算量。通过多次实验,最终采用在 ResNet 生成器的下采样层之前的第一层卷积和上采样层之后的输出层引入可变形卷积 DCN。在上采样层、下采样层以及残差块中依旧使用传统的卷积操作。

引入用可变形卷积之后,相对于普通的固定卷积对于图像转换的效果提升如图 5 所示。

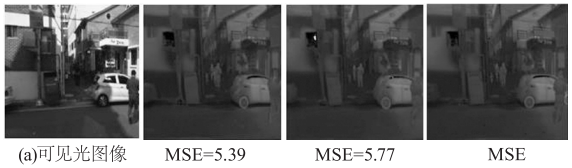

图 5 转换效果提升展示

Fig. 5 Conversion effect enhancement is displayed

如图 5 所示,引入可变形卷积后,虽然增加了一些噪音点,使生成图像的均方根误差(MSE)大了一些,不过图像的结构相似性(SSIM)也得到进一步提升。图中人的轮廓更接近真实图像,对于小目标,如图中的小孩的生成效果也更好。由于本文研究的是图像转换任务,最重要的就是生成图像的一些重要目标(如图中的车和人)的生成效果更好,所以在图像的背景区域引进一些噪音点对重要目标的生成并无影响。综上所述,引入可变形卷积后对图像的生成效果是有提升的。

2.3 引入 BAM 注意力机制

为了进一步提高网络对图像中重要目标的特征提取能力和生成效果,本文在 pix2pix 的 ResNet 生成器部分引入注意力机制。

注意力机制(Attention Mechanism,AM)是机器学习中的一种数据处理方法,通过模拟人的注意力的方式,让卷积神经网络去注意特征图中应该注意的地方而不是什么都关注[10]。BAM(Bottleneck Attention Module)是一种可以在通道和空间维度上进行应用的轻量化注意力模块,可以在不降低计算效率的情况下,更好的生成关键目标。

BAM 注意力机制分为两个部分,第一部分为通道上的注意力模块 CAM(Channel attention module),第二部分为空间上的注意力模块 SAM(Spatial attention module)[11]。CAM 模块与 SE(Squeeze and Excitation)的结构基本一样。SAM 模块在通道维度进行池化(pool),然后用了两次 3×3 的空洞卷积,最后在用一次 1×1 的卷积得到 Spatial Attention 的矩阵。最后 Channel Attention 和 Spatial Attention 矩阵进行相加(这里用到了广播机制),并进行归一化,这样一来,就得到了空间和通道结合的注意力(attention)矩阵[12]。BAM 注意力机制的原理如图 6 所示。

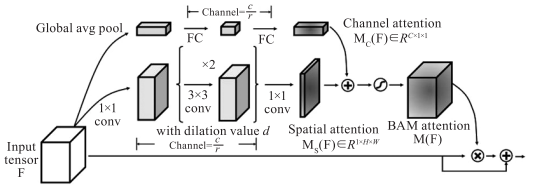

图 6 BAM 注意力机制

Fig. 6 BAM attention mechanism

经过多次实验后,本文在 ResNet 生成器的下采样层之后和残差块之后分别引入 BAM 注意力机制。可以使网络对重要特征的提取能力和重要目标的生成能力有一个显著的提升。

引入基于 BAM 的注意力机制之后,对于图像转换的效果提升如图 7 所示。

图 7 转换效果提升展示

Fig. 7 Conversion effect enhancement is displayed

在图 7 中可以直观的看出,引入 BAM 注意力机制后,网络会更加关注图中的车、人、房子、栅栏等主要目标,对这些重要目标的生成效果更好。从评价指标是来看,生成的图像在保证结构相似性(SSIM)不下降的前提下,大幅降低均方根误差(MSE)。有效缓解了引入可变形卷积带来的误差。

在生成器中引入 BAM 注意力机制,能够进一步提升图像的生成效果。图 8 是本文改进后生成器的网络结构图。

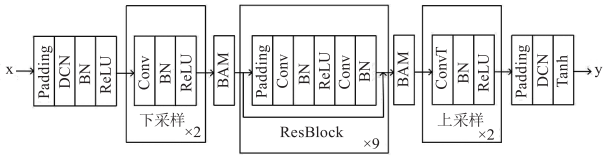

图 8 生成器网络结构

Fig. 8 Generator network structure

3 鉴别器的改进

3.1 改变 PatchGAN 中卷积层的层数

由于对抗生成网络生成图像是靠生成器和鉴别器之间博弈式的训练,来不断提高各自的能力,这种博弈是交替进行的[13]。若是生成器很强,而鉴别器很弱,在下次训练生成器时,鉴别器的输出结果接近于真,就无法给出足够的误差来训练生成器。反之,当鉴别器很强,而生成器很弱,在下次训练生成器时,由于给出的误差太大,而让生成器无法准确感知能力差距在哪,从而也不能很有效的提升自己。这就需要生成器和鉴别器效果相当,使这种博弈式的训练能够达到纳什均衡点,鉴别器分辨不出真实图像和生成器生成的假图像,即训练 200 个周期后,鉴别器鉴别生成器生成的假图像为真和为假的概率都接近 0.5。上文主要是针对生成器做出的改进。下面针对鉴别器做出一些改进。

在 pix2pix 的原始论文中,为了能更好得对图像

的局部做判断,鉴别器部分采用 PatchGAN 结构。在那以前的鉴别器都是把整张图像作为输入,在输出一个是否为真的概率。而 PatchGAN 鉴别器是把图像等分成多个固定大小的 Patch,每个 Patch 的大小为 70 ×70,分别判断每个 Patch 的真假,最后再取平均值作为最后的输出[14]。

原始 PatchGAN 鉴别器的卷积层数为 3 层,本文通过改变 PatchGAN 中卷积层的层数,拟找出最适合改进后生成器的鉴别器卷积层数。本文把 PatchGAN 的卷积层数进行增加和减少,并依次进行训练和测试。对生成图片分别计算均方根误差(MSE)、结构相似性(SSIM)、

峰值信噪比(PSNR)等参数,实验结果如表 1 所示。

表 1　使用不同卷积层数的效果对比

Tab. 1 Comparison of the effects of using different convolution layers

卷积层数	MSE ↓	SSIM ↑	PSNR ↑
1 层	5. 14	0. 907	33. 8
2 层	**4. 75**	**0. 915**	**34. 6**
3 层(原)	5. 21	0. 912	33. 9
4 层	5. 39	0. 902	33. 4
5 层	5. 81	0. 889	32. 8
6 层	6. 01	0. 873	32. 5

从表 1 中不难看出,在依次增加 PatchGAN 卷积层数的过程中,均方根误差(MSE)先降后升,在使用 2 层卷积的时候达到最小值;结构相似性(SSIM)和峰值信噪比(PSNR)先升后降,在使用 2 层卷积的时候达到最大值。综上所述,将 PatchGAN 鉴别器的鉴别器层数设置为 2 层,比原 PatchGAN(3 层卷积)更适合本文改进后的生成器,能够更好的完成本文图像转换任务。

本文改进后的 2 层卷积的 PatchGAN 鉴别器的网络结构如图 9 所示。

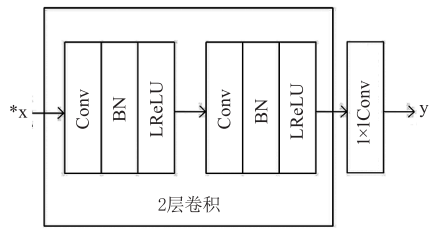

图 9　鉴别器网络结构

Fig. 9 Discriminator network structure

4　实验及结果分析

4. 1　实验过程

本实验使用开源红外数据集 KAIST。在其中选取 1422 对用作训练,16 对用作测试。训练 200 个周期(epochs)。本实验采取的设备是 I5 - 8300H 处理器、NVIDIA GTX-1050 显卡。训练过程中,输入图片大小设置为 256 ×256,为了验证本文对模型进行优化后的效果的变化,每次实验初始学习率(learning rate)均为 0. 0002,训练前 100 个周期使用这个初始学习率,后 100 个周期学习率开始线性衰减至 0。

对图像转换效果的评价方法有很多,本文主要通过均方根误差(MSE)、结构相似性(SSIM)、峰值信噪比(PSNR)等参数来评价。其中均方根误差(MSE)反映了转换图像和目标图像之间的差异,值越小代表两幅图像的差异越小。结构相似性(SSIM)是衡量两幅图像结构相似程度的重要指标,值越大代表两幅图像的相似度越高[3]。峰值信噪比(PSNR)是表示信号最大可能功率和影响它的表示精度的破坏性噪声功率的比值。值越大代表信息图像的信息相对越多、噪音相对越少,图片的质量也就越好。

4. 2　实验结果与分析

为了验证改进 pix2pix 算法的图像生成效果,采用了对比的方式进行实验。为了验证将 U-Net 生成器替换为 ResNet 生成器对生成效果的提升,设置实验一与实验二进行生成效果对比。实验一:利用原 pix2pix 网络(U-Net 生成器、PatchGAN 鉴别器)对数据集进行训练和测试。实验二:将 U-Net 生成器替换为 ResNet 生成器,对数据集进行训练和测试。如表 2 所示,使用 U-Net 生成器生成图像的均方根误差(MSE)为 6. 92、结构相似性(SSIM)为 0. 823,使用 ResNet 生成器生成图像的均方根误差(MSE)为 5. 39、结构相似性(SSIM)为 0. 902,相较于改进前误差下降了 22. 1 %、结构相似性上升了 9. 6 %。生成图像的效果有显著提升。为了验证可变形卷积对网络的提升,设置实验三:在实验二的基础上,对 ResNet 生成器的下采样层之前的第一层卷积和上采样层之后的输出层引入可变形卷积,并进行训练和测试。得到的结果如表 2 所示,采用实验三中的算法生成图像的 MSE 为 5. 77、SSIM 为 0. 913,相对于实验二 MSE 上升了 7. 1 %、SSIM 上升了 1. 2 %。虽

然引入了一些噪音点,增大了误差,不过对于图像转换任务来说,结构相似性的提升更为重要。所以,引入可变形卷积对图像生成效果是提升的。为了验证加入 BAM 的注意力机制对算法性能的提升,设置实验四:在实验三的基础上在 ResNet 生成器的下采样层之后和残差块之后分别加入一个 BAM 注意力机制,并进行训练和测试。如表 2 所示,采用实验四中的算法生成图像的 MSE 为 5.21、SSIM 为 0.912,相对于实验三 MSE 下降了 9.7 %、SSIM 几乎不变。在结构相似性不变的前提下,大大降低了 MSE,中和了上一步加入可变形卷积引入的误差。因此实验四对算法性能是提升的。在设置实验五:在实验四的基础上,对鉴别器进改进。原 pix2pix 使用的 PatchGAN 鉴别器的卷积层为 3 层,根据前文可知,将卷积层修改为 2 层更适合优化后的生成器,能更好的完成图像转换任务。如表 2 所示,采用实验五中的算法生成图像的 MSE 为 4.75、SSIM 为 0.915,相对于实验四 MSE 下降了 8.8 %、SSIM 提高了 3.3 %。进一步提升了图像转换效果。以上结果表明本文提出的方法能够更好的完成图像转换任务。

表 2 各算法实验结果对比

Tab. 2 Comparison of experimental results of each algorithm

实验	ResNet	可变形卷积	BAM	2_layers_D	MSE ↓	SSIM ↑	PSNR ↑
一	否	否	否	否	6.92	0.823	31.3
二	是	否	否	否	5.39	0.902	33.5
三	是	是	否	否	5.77	0.913	33.7
四	是	是	是	否	5.21	0.912	33.9
五	是	是	是	是	**4.75**	**0.915**	**34.6**

从表 2 中可以看出,在依次加入本文的改进方法后,生成图像的均方根误差(MSE)整体呈下降趋势,结构相似性(SSIM)和峰值信噪比(PSNR)整体呈上升趋势。生成图像的质量得到有效提升。

5 结 论

本文通过对 GAN 的衍生模型 pix2pix 进行改进来实现可见光图像与红外图像的转换任务。主要从生成器和鉴别器两个方面进行改进。对于生成器,通过将 U-Net 生成器替换为残差结构的生成器来缓解梯度消失问题;通过引入可变形卷积,提高目标边缘轮廓和小目标在特征图上像素的占比;通过引入 BAM 注意力机制,提高了算法对图像中主要目标的特征提取能力以提升图像生成效果。对于鉴别器,通过调整 PatchGAN 中卷积层的层数,来适应改进后的生成器。通过对比得出,鉴别器使用 2 层卷积时,图像的生成效果最好。并在公开数据集 KAIST 中选取部分数据集进行了对比实验。实验结果表明,本文提出的改进方法在红外生成图像上的评价指标有显著提升。仿真结果和真实图像之间的差别已经很小,人的肉眼已经很难辨别真伪,所以此方法有较高的实用性和可行性。

参考文献:

[1] Li Guowei, Shi Zhiguang, Zhang Yan. Image conversion technology based on generative adversarial network[J]. Journal of Terahertz Science and Electronic Information Technology, 2021, 19(4): 724 – 727, 732. (in Chinese)
李国威,石志广,张焱. 基于生成对抗网络的图像转换技术[J]. 太赫兹科学与电子信息学报,2021,19(4):724 – 727,732.

[2] Yang Xiaoli, Lin Shuzhen, Lu Xiaofei, et al. Multimodal image fusion based on generative adversarial networks [J]. Laser and Optoelectronics Progress, 2019, 56 (16): 525 – 536. (in Chinese)
杨晓莉,蔺素珍,禄晓飞,等. 基于生成对抗网络的多模态图像融合[J]. 激光与光电子学进展,2019,56(16):525 – 536.

[3] Yu Peilun, Shi Yan, Wang Han. Infrared-visible image conversion of parallel generation networks[J]. Chinese Journal of Image and Graphics, 2021, 26(10): 2346 – 2356. (in Chinese)
余佩伦,施佺,王晗. 并行生成网络的红外 – 可见光图像转换[J]. 中国图象图形学报,2021,26(10):2346 – 2356.

[4] Engin D, Genç A, Ekenel H K. Cycle-dehaze: enhanced

cycleGAN for single Image dehazing[C]//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPR). Salt Lake City, USA: IEEE, 2018: 825 – 833.

[5] Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative adversarial-nets[C]//Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, Canada: ACM, 2014: 2672 – 2680.

[6] Guo Lijun, Zhao Jieyu, Shi Zhongzi. Image classification method based on the fusion of generation model and discriminant method[J]. Electronica Sinica, 2010, 38(5): 1141 – 1145. (in Chinese)
郭立君,赵杰煜,史忠植. 生成模型与判别方法相融合的图像分类方法[J]. 电子学报, 2010, 38(5): 1141 – 1145.

[7] Wu Guojun, Bai Tingzhu, Bai Funing. Research on infrared image inversion basedon visible image[J]. Infrared Technique, 2011, 33(10): 574 – 579. (in Chinese)
武国军,白廷柱,白茯宁. 基于可见光图的红外图像反演研究[J]. 红外技术, 2011, 33(10): 574 – 579.

[8] Li Xingji, Liu Chaoming, Yang Jianqun. Synergistic effect of ionization and displacement damage in NPN transistors caused by protons with various energies[J]. IEEE Transaction on Nuclear Science, 2015, 62(3): 1375 – 1382.

[9] Bao Jun, Liu Hongzhe. Fisheye image object detection based on deformable convolutional network[J]. Computer Engineering, 2021, 47(4): 248 – 255. (in Chinese)
包俊,刘宏哲. 融合可变形卷积网络的鱼眼图像目标检测[J]. 计算机工程, 2021, 47(4): 248 – 255.

[10] Barnaby H J, Smith S K, Schrimpf R D, et al. Analytical model for proton radiation effects in bipolar devices[J]. IEEE Transaction on Nuclear Science, 2002, 49(6): 2643 – 2649.

[11] Li Jiaxin, Hou Jin, Sheng Boying, et al. Remote sensing small object detection network based on improved YOLOv5[J]. Computer Engineering, 2023, 49(9): 256 – 264. (in Chinese)
李嘉新,侯进,盛博莹,等. 基于改进 YOLOv5 的遥感小目标检测网络[J]. 计算机工程, 2023, 49(9): 256 – 264.

[12] Huang G, Liu Z, van der Maaten L, et al. Densely connected convolutional networks[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE, 2017: 12.

[13] Liu Yang, Yu Jianhua, et al. Research and application of an imprived enhanced image processing algorithm[J]. Laser & Infrared, 2019, 49(3): 381 – 384. (in Chinese)
刘洋,余建华,等. 一种改进型增强图像处理算法研究与应用[J]. 激光与红外, 2019, 49(3): 381 – 384.

[14] Huang Huang, Tao Haijin, Wang Haifeng. Low-illumination image enhancement using a conditional generative adversarial network[J]. Journal of Image and Graphics, 2019, 24(12): 2149 – 2158. (in Chinese)
黄鎕,陶海军,王海峰. 条件生成对抗网络的低照度图像增强方法[J]. 中国图象图形学报, 2019, 24(12): 2149 – 2158.