

DOI: 10.3969/j.issn.1673-6141.2023.05.007

基于空间多尺度残差网络的红外与可见光 图像融合

张亦孟¹, 林伟国^{2*}

(1 北京机电工程总体设计部, 北京 100005;
2 北京化工大学信息科学与技术学院, 北京 100029)

摘要: 针对如何充分提取和融合红外与可见光图像典型特征的问题, 提出一种基于空间多尺度残差网络的图像融合算法。首先, 将源图像输入基于空间多尺度残差模块组成的编码器网络, 通过源图像重建任务, 训练编码器自动获取重要特征信息的能力; 然后, 引入特征金字塔结构, 设计了特征通道自注意力机制, 编码器输出的基础层和细节层进行融合, 减小尺度噪声, 并由解码器重构出融合图像; 最后, 利用公开数据集进行定性和定量实验, 证明了改进算法在突出红外图像目标和保留可见光图像纹理细节两方面的优势, 相比于DDcGAN算法, 新算法的标准差和平均梯度分别提升了12.91%和47.41%。

关键词: 图像融合; 自动编码器; 空间多尺度残差模块; 通道自注意力

中图分类号: TP391.41

文献标识码: A

文章编号: 1673-6141(2023)05-469-010

Infrared and visible images fusion with spatial multiscale residual networks

ZHANG Yimen¹, LIN Weigu^{2*}

(1 Beijing System Design Institute of Electro Mechanic Engineering, Beijing 100005, China;

2 College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China)

Abstract: To fully extract and fuse typical features of infrared and visible images, an image fusion algorithm based on spatial multi-scale residual network is proposed. Firstly, the source image is input into an encoder network composed of spatial multi-scale residual modules, and through the task of image reconstruction, the encoder network is trained to automatically obtain important features. Then, a feature pyramid and a channel self-attention are introduced, the output of basic layer and detail layer by the endoder are fused to reduce scale noise, and the fused image is reconstructed by the decoder. Finally, qualitative and quantitative experiments on public datasets are carried out, and it is demonstrated that the imporved algorithm outperforms the alternatives on highlighting infrared image targets and preserving

基金项目: 辽宁省应用基础研究计划项目 (2023JH2, 101300239)

作者简介: 张亦孟 (1990-), 北京人, 工程师, 主要从事图像处理、运载器系统电气综合与健康管理方面的研究。

E-mail: zhangyimeng1990@vip.sina.com

导师简介: 林伟国 (1968-), 浙江宁波人, 博士, 教授, 硕士生导师, 主要从事图像处理、智能检测与故障诊断方面的研究。

E-mail: linwg@mail.buct.edu.cn

收稿日期: 2022-02-17; 修改日期: 2022-04-08

*通信作者。

visible image texture details. Compared with the DDcGAN algorithm, the standard deviation and average gradient of the proposed algorithm have been improved by 12.91% and 47.41%, respectively.

Key words: image fusion; auto-encoder; spatial multi-scale residual module; channel self-attention

0 引言

由于硬件设备的限制,单个传感器获取的信息往往受到限制,难以提取到实际场景内的全部有效信息。为了解决此问题,可以使用多个传感器提取互补的多个图像信息,而后使用图像融合方式对其进行融合,生成视觉效果良好的单幅融合图像,为研究人员提供全面的场景描述,便于应用到后续的计算机视觉任务,如视频监控、场景理解和目标识别等^[1]。

现如今可见光与红外光图像融合是图像融合技术中一个热门的研究方向,原因在于两种图像互补性强、获取成本低。其中红外图像能够反映目标发出的热辐射量,能够克服雨、雾和尘埃等恶劣天气条件进行全天候工作,但图像存在细节性差、分辨率低的缺点;可见光图像的波段范围与人类视觉系统一致,能够提供丰富的纹理细节,易于场景理解,但是对环境变化非常敏感,受影响之后图像质量会明显下降。因此,红外与可见光图像融合的目的是从红外图像获取显著目标、从可见光图像获取丰富的纹理细节,形成高质量的单幅图像。

随着近些年芯片技术的不断进步,基于深度学习的图像融合方法相比于传统方法展现出更优异的性能,经过大量数据训练的神经网络能够保证特征提取的高精度与高效率,是多尺度变换^[2]、稀疏表示和显著性检测等需要大量手动设计的传统图像融合算法难以超越的^[3]。

深度学习框架下的图像融合技术可大致分为基于两种网络结构:卷积神经网络(CNN)与生成对抗网络(GAN)^[4]。Li和Wu^[5]基于密集连接的卷积神经网络(DenseNet)^[6]提出DenseFuse算法,首先训练编码器-解码器的特征提取和重构能力,之后加入手动设计的融合规则实现了特征融合,DenseNet在卷积层间建立的短连接提升了算法内部的特征传递效率、缓解了网络退化问题。Li等^[7]基于嵌套连接^[8]改进编码器结构,同时为了得到具有更高对比度的融合图像,其采用了空间注意力机制。An和Wang^[9]将现有文献的融合结果作为标签数据训练模型重构源图像特征的精度,弥补了可见光与红外光图像融合模型训练中缺少真实数据的缺陷。但以上算法的融合策略是不可学习的,而且采用手动设计的方式增加了算法设计的复杂性。

Ma等^[10]提出端到端的图像融合算法(FusionGAN),首次将红外与可见光图像的融合问题转换为GAN网络内部生成器与判别器的对抗问题,输出融合图像的特征分布能够不断接近两种源图像。Xu等^[11]基于残差网络(ResNet)^[12]改进了FusionGAN的基础结构,增强了特征提取能力。Ma等^[13]后续又提出了双判别器的图像融合算法(DDcGAN),解决了FusionGAN算法使用单判别器造成融合图像纹理细节丢失的问题。但以上基于GAN的图像融合算法,因GAN自身存在的难以稳定训练问题使得图像融合结果具有不确定性,且容易产生视觉伪影。

针对上述问题,本文提出一种基于空间多尺度残差网络(SMSRN)的图像融合模型。首先,使用多尺度残差模块构建自动编码器,将源图像分解为基础层和细节层;其次,引入基于空间金字塔的多尺度模块,提高模型的空间自适应性;然后,采用混合损失函数约束源图像和融合图像之间在像素强度与梯度两方面的接近程度;最后,引入通道自注意机制,融合红外和可见光图像的通道维度特征,并由解码器重构出融合图像。

1 相关工作

1.1 多尺度残差网络

在计算机视觉任务中, 感知不同尺度特征信息对于理解目标和背景的局部和全局信息至关重要。U 型网络 (U-Net)^[14] 利用跳跃连接, 有效组合了降采样提取的细节信息和上采样提取的语义信息; 特征金字塔方法 (FPN)^[15] 通过在图像的特征空间构造金字塔, 将高层的强语义信息传播到低层, 同时实现了特征融合和特征预测。这些方法在下采样扩大感受野的过程中都引入了池化层, 丢失了图像的部分原始信息。

针对上述问题, 刘鹏飞等^[16] 提出改进残差网络 (I-ResNet), 与原始 ResNet 的对比如图 1 所示。其中图 1 (a) 为原始残差网络, 网络中包括批量归一化 (BN)。图 1 (b) 是改进后的残差网络, 其包含两个步幅卷积模块, 步长为 1/2。

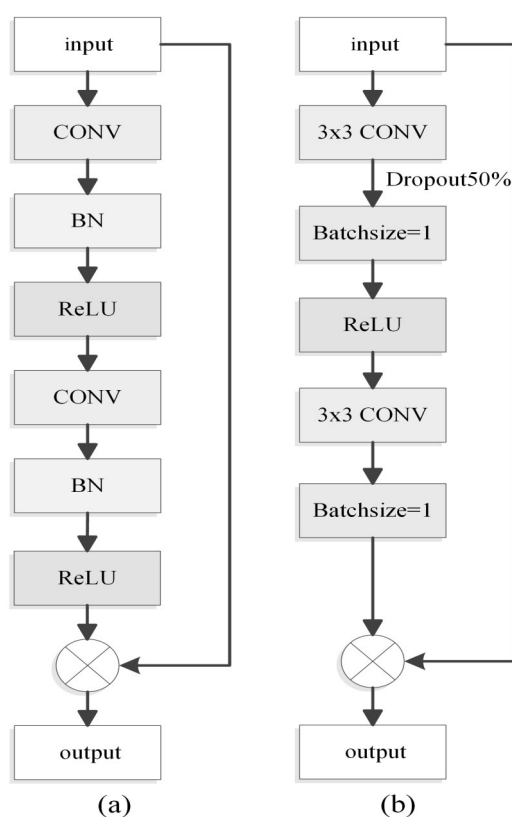


图 1 改进前后残差网络结构对比。(a) 原始残差网络结构; (b) 改进残差网络结构

Fig. 1 Comparison of residual network structure before and after improvement. (a) Original residual network structure; (b) improve residual network structure

在特征提取前首先使用 3×3 卷积扩大感受野; 然后添加概率为 50% 的随机失活 (Dropout), 可以提高网络的泛化能力, 防止模型过拟合; 最后通过 3×3 的卷积核, 再次融合多尺度特征, 每批次大小 (Batchsize) 为 1。这种先拆分后融合的策略, 可以使卷积层更加高效地处理特征。

1.2 注意力机制

注意力机制能够通过引导网络关注重要特征并抑制不必要特征, 在深层次的 CNN 架构上可以获得明显

的性能增益。压缩-激励网络 (SE-Net)^[17]首次提出了以全局平均池化计算通道注意力,精确地推断出通道维度的显著特征;卷积注意力模块 (CBAM)^[18]使用平均池化和最大值池化进一步增强通道维度的特征聚合,并通过串行方式与二维卷积得出的空间注意力相结合,从空间和通道两个维度建立了注意力机制。以上模型为避免计算量暴增、显存爆炸,强力压缩了通道维度,不可避免地造成了信息丢失,在细粒度的像素级任务中难以满足精度需求。

对于可见光与红外图像融合任务,本研究采用特征通道注意力机制,解决了图像在不同尺度条件下存在的语义差异,通过建立的图像金字塔结构,并结合特征通道注意力机制,根据感受野大小融合局部和全局特征,减小了尺度噪声,消除了深度学习用于融合任务时常出现的伪影现象。

2 图像融合模型

2.1 网络结构设计

网络为了将红外与可见光图像的特征分解、融合重构,而使用编码器-解码器框架。在训练阶段,输入红外与可见光图像,目的是训练编码器、解码器的特征提取和重构能力。网络总体结构如图2所示,是基于DenseNet网络设计的端到端的图像融合模型,输入是通道维度连接后的红外图像 I_i 和可见光图像 I_v 。

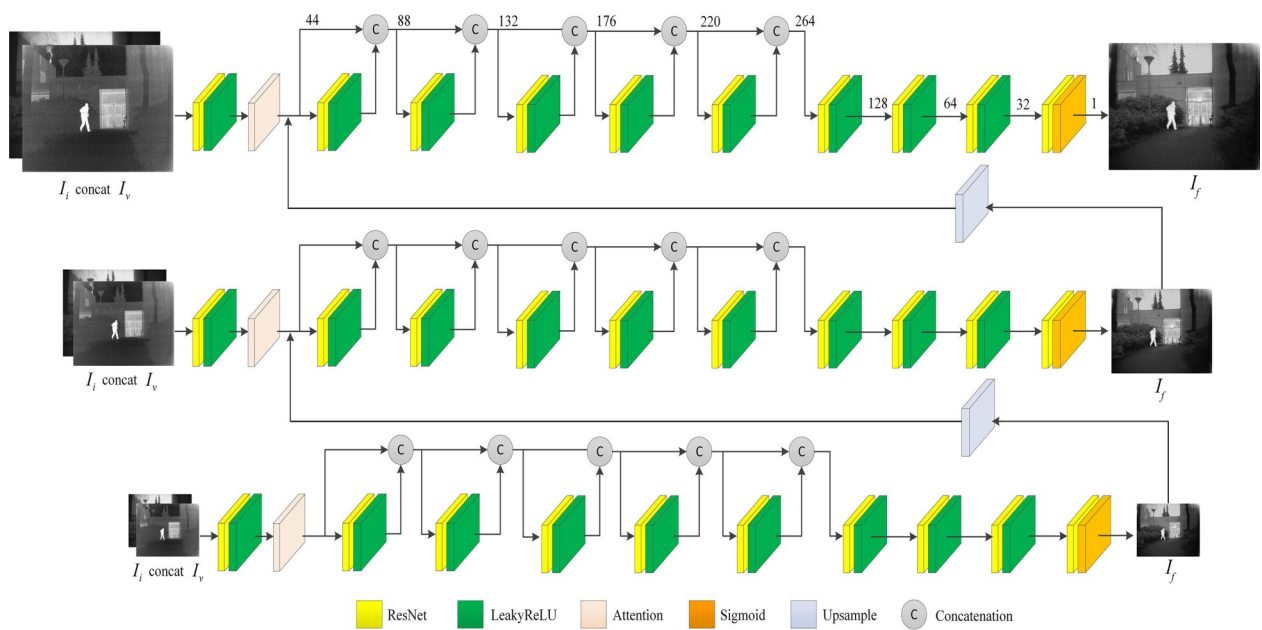


图2 网络总体结构图

Fig. 2 Overall structure of the network

首先由单层的 3×3 残差模块完成浅层特征提取,之后将上一层的融合图像与浅层特征进行融合,通过设计好的空间多尺度注意力网络,完成对源图像重要特征的增强和非显著性特征的削弱,便于后续的特征提取。算法不需要额外设计融合规则,直接输出融合图像 I_f 。网络结构包含10层,每一层均由改进ResNet模块和激活函数组成,其中前7层作为特征提取网络基于DenseNet结构实现。改进ResNet模块结合密集连接实现特征重用,能够在缓解梯度消失和爆炸的同时,更加全面地保留源图像特征信息;再经过3层卷积核为 3×3 的普通残差后,最后一层使用Sigmoid激活函数将通道数降为1,输出重构的融合图像。

为减少特征提取和重构过程中出现的特征信息丢失, 网络结构不引入池化层, 并且保证融合图像与输入图像分辨率相同, 便于算法性能的客观分析。同时引入 Leaky ReLU 激活函数加速网络收敛。

2.2 空间多尺度注意力网络

算法设计了空间多尺度网络, 用于红外图像的特征提取, 如图 3 所示。对于可见光图像使用相同的网络架构, 但二者参数相互独立。深层次的残差网络可以有效增大图像感受野, 提取图像深层特征。因此, 模型建立包含粗粒度层、中间层和细节层的逐级精细的多层高斯金字塔结构, 输入每层高斯金字塔的图像块分辨率分别为 256×256 , 128×128 和 64×64 , 输入与输出之间的尺度比例为 $1/2$ 。

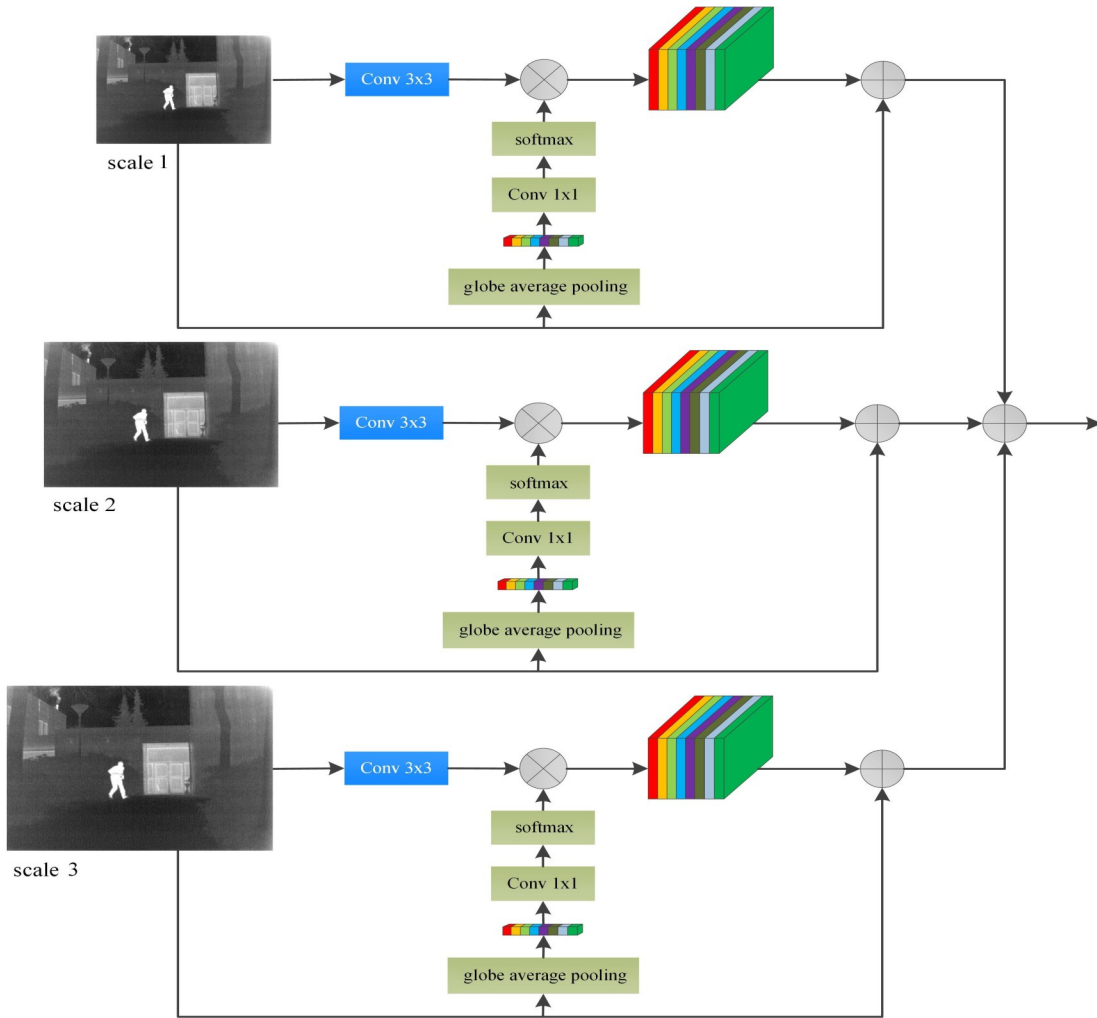


图 3 基于红外图像的空间多尺度注意力网络结构

Fig. 3 Spatial multi-scale attention network structure based on infrared images

注意力网络使用空洞卷积作为基本模块提取特征, 通过超参数-膨胀率控制卷积核的大小, 相比于标准卷积, 其能够获取不同卷积核下对应的感受野特征^[19]; 同时摒弃了池化层, 可以在避免信息丢失的同时扩大感受野。然后通过赋予权重系数的形式, 增强显著特征, 抑制无用特征的影响。最后, 网络融合了尺度注意力模块, 可以确定源图像在哪个尺度上值得关注, 从而滤除尺度噪声的干扰。

依据不同输入尺度的源图像, 在三层空间金字塔结构中使用了通道注意力模块, 在三个图像多尺度下融

合,将图像特征经过卷积和激活操作后可以增强有用信息,经过权值的拟合,增大强特征的比重,减小弱特征或无用信号的干扰,能够提升融合图像的空间纹理细节。最后,红外与可见光图像对应相同大小的多尺度通道连接融合并输入解码器,经历特征重构,输出最终融合图像。

2.3 损失函数

为了更精准地获取源图像的多尺度信息,构建包含结构相似性损失 L_{SSIM} 、均方误差损失 L_{MSE} 和梯度损失 $L_{Gradient}$ 的混合损失函数 L ,计算公式为

$$L = \lambda_1 L_{SSIM} + \lambda_2 L_{MSE} + \lambda_3 L_{Gradient}, \quad (1)$$

式中 λ_1 、 λ_2 和 λ_3 代表损失函数的权重系数。

L_{SSIM} 能够从亮度、对比度和结构方面对源图像和融合图像进行相似性度量,其计算公式为

$$L_{SSIM} = \frac{(2\mu_x\mu_f + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_f^2 + c_1)(\sigma_x^2 + \sigma_f^2 + c_2)}, \quad (2)$$

式中 μ_x 和 μ_f 代表源图像和融合图像的标准差, σ_{xy} 代表源图像和融合图像的协方差, c_1 和 c_2 是常数, σ_x^2 和 σ_f^2 代表源图像和融合图像的方差。

但 L_{SSIM} 缺少红外图像的像素强度分布约束,故引入均方误差构建强度损失,计算公式为

$$L_{MSE} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [x(i,j) - f(i,j)]^2, \quad (3)$$

式中 m 和 n 代表源图像和融合图像的长和宽, $x(i,j)$ 和 $f(i,j)$ 代表源图像和融合图像的像素值。

使用均方误差构建损失函数,会引入噪声^[20],掩盖融合图像的部分纹理细节,需额外建立梯度损失,计算公式为

$$L_{Gradient} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [\|\nabla I_f - \nabla I_v\|_F^2], \quad (4)$$

式中 ∇ 代表梯度算子, $\|\cdot\|_F$ 代表Frobenius范数, I_f 和 I_v 分别代表融合图像和可见光图像。

3 实验与结果分析

3.1 实验设置

数据集方面选择了一TNO数据集,搭建算法的训练集和测试集。其中训练集由选取的42对红外与可见光图像设置步长为32裁剪而来,能够保证网络模型获得充分训练,测试集则由目标和背景特征差异明显的10对红外与可见光图像构成。本文所提算法基于Tensorflow深度学习框架,实验平台配置为:CPU Intel I7-10700K和GPU NVIDIA GeForce RTX3070Ti。网络参数的优化器选择Adam,Epoch为600,学习率设置为0.001。权重系数 $\lambda_1=5$ 、 $\lambda_2=0.1$ 、 $\lambda_3=5$ 。

3.2 对比实验

为充分验证所提算法的优异性能,从基于CNN图像融合方法和基于GAN图像融合方法各选取2种算法进行对比实验,分别是:DenseFuse算法、双分支网络(DBN)算法^[21]、FusionGAN算法和DDcGAN,实验结果

如图 4 所示。由实验结果可知, DenseFuse 模型的输出在目标区域保留了过多的纹理细节信息, 图像对比度较低, 原因在于手工设计的融合策略是不可学习的, 无法充分融合两种源图像特征, 主观效果较差; DBN 算法重点围绕可见光图像建立基于局部细节和全局结构的双特征提取分支, 但缺少损失函数项约束融合图像与红外图像的像素强度分布相似性, 导致人物和建筑的亮度较低; FusionGAN 算法仅建立单判别器-生成器的对抗网络, 无法同时保留两种源图像特征, 融合图像以红外图像的像素强度信息为主, 丢失了草丛和地面的纹理细节信息, 无法正确判断出目标所处环境; DDcGAN 算法虽通过构建双判别器结构, 克服了 FusionGAN 算法在融合纹理细节方面的能力缺陷, 但没有稳定双重对抗网络的措施, 反而进一步加剧了 GAN 训练不稳定的问题, 使融合图像出现伪影, 而且第一组实验结果中人物和树木等背景边缘模糊; SMSRN 算法通过空间多尺度金字塔网络结构, 有效保留了两种源图像的典型特征, 使融合图像趋近于具有突出红外目标的可见光图像, 在图像融合精度和视觉效果两方面均优于其他对比算法。

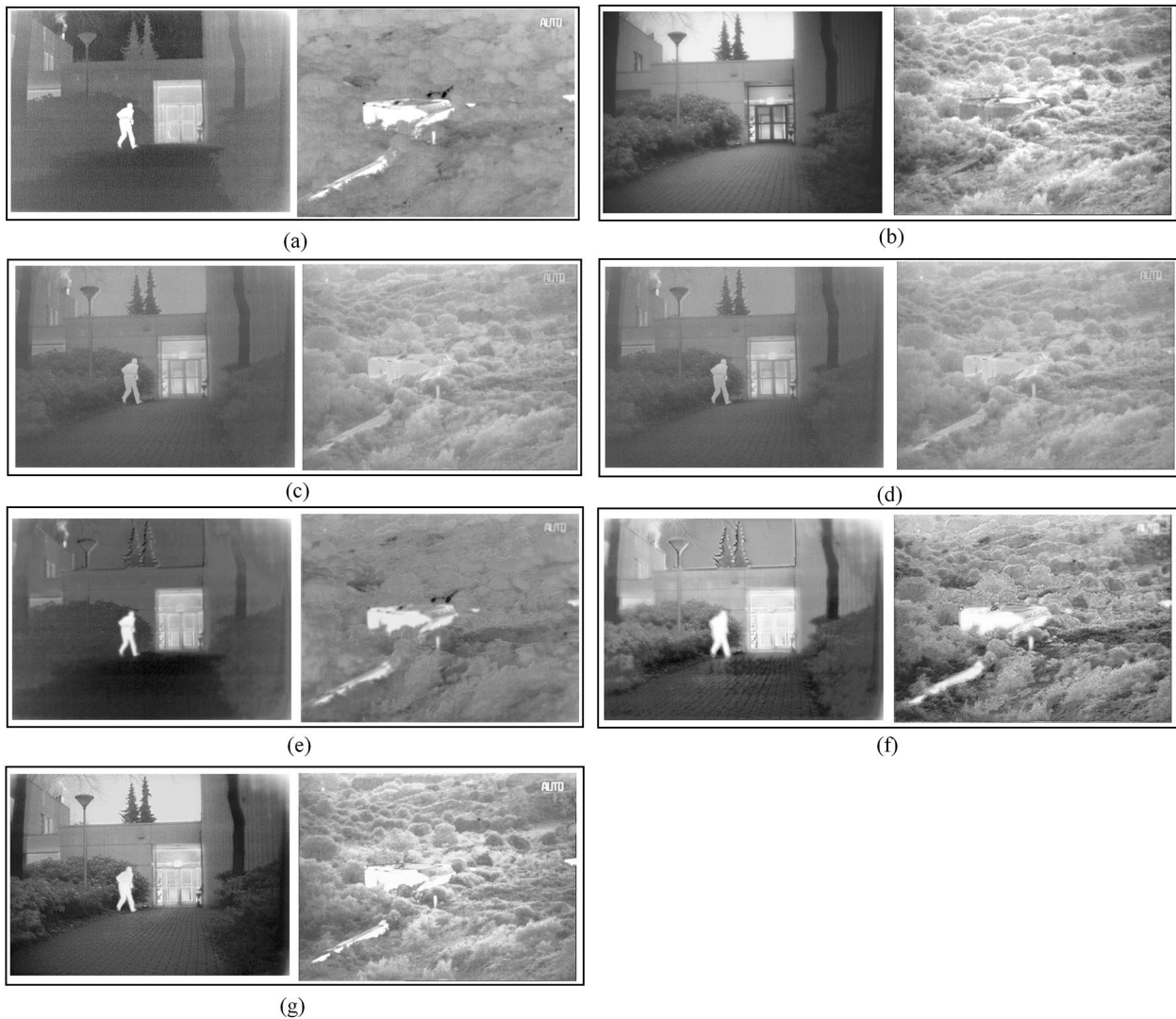


图 4 对比实验主观结果。(a) 红外图像; (b) 可见光图像; (c) DenseFuse 算法; (d) DBN 算法;
(e) FusionGAN 算法; (f) DDcGAN 算法; (g) SMSRN 算法

Fig. 4 Contrast experiment with subjective results. (a) Infrared image; (b) visible image; (c) the algorithm of DenseFuse; (d) the algorithm of DBN; (e) the algorithm of FusionGAN; (f) the algorithm of DDcGAN; (g) the algorithm of SMSRN

选取6种客观指标,定量分析上述算法在保留目标对比度和纹理细节、融合图像质量、视觉效果和效率等多方面的性能^[22],主要包括:标准差(SD)、平均梯度(AG)、结构相似性(SSIM)、熵(EN)和差异相关性总量(SCD)和平均图像融合时间。评价结果如表1所示。由表可知,DenseFuse和DBN算法的SD指标最差,在提升对比度和突出目标方面需要做出改进;FusionGAN算法的SD指标优于DenseFuse和DBN算法,但AG指标最低,说明缺失重要的纹理细节,因为单判别器的对抗形式仅能保留单一的源图像特征;DDcGAN算法的SSIM数值最低,丢失过多的源图像特征,需要改进网络结构稳定训练,减少伪影对融合图像的影响;SMSRN算法在前5项指标中均取得最优结果,特别是SD与AG两项指标相比于其他算法优势十分明显。相比于DDcGAN算法,标准差提升了12.91%,平均梯度则提升了47.41%。依据指标的定义,这两项评价指标优秀的定量分析结果证明了算法能够高效提取和融合红外图像的目标对比度信息、可见光图像的纹理细节信息,与主观评价结果相吻合;在效率方面,SMSRN算法为充分提取源图像特征,设计了相对复杂的网络结构,因此指标未能超越DenseFuse与DBN算法,仅取得较优结果;但与传统方法动辄数十秒的图像融合时间相比,SMSRN算法的效率已能够满足实际应用的需求。

表1 对比实验客观评价结果

Table 1 Objective evaluation results of comparative experiments

Index	DenseFuse	DBN	FusionGAN	DDcGAN	SMSRN
SD	24.40	26.34	31.47	45.77	51.68
AG	2.97	3.04	2.67	5.59	8.24
SSIM	0.63	0.71	0.62	0.57	0.73
EN	6.34	6.57	6.64	7.35	7.69
SCD	1.33	1.59	1.16	1.51	1.76
Time/s	0.26	0.14	0.51	0.82	0.32

3.3 损失函数实验分析

算法需选取合理的损失函数权重系数,保证在迭代优化中获得高质量的融合图像。结构相似性损失 L_{SSIM} 重点关注对比度和结构的变化,对强度分布差异的约束较弱;红外图像的热辐射信息主要通过像素强度表征,因此引入均方误差构建强度分布约束 L_{MSE} ,但同时过量的 L_{MSE} 会平滑图像特征、降低清晰度^[23],因此权重系数取值小于 L_{SSIM} ;为进一步弥补 L_{MSE} 的缺陷,同时考虑可见光图像的纹理细节主要通过梯度表征,额外增加梯度损失 $L_{Gradient}$ 约束融合图像具有与可见光图像相似的特征分布。综上,当 L_{SSIM} 、 L_{MSE} 和 $L_{Gradient}$ 的权重系数依次取值为5、0.1和5时,对应的损失函数曲线以及总损失函数曲线在算法经充分的训练迭代后均能获得稳定的收敛(图5),此时算法可以输出最优的图像融合结果,证明权重取值合理。

4 结论

提出一种用于融合红外光与可见光图像的改进的自动编码器模型。在训练阶段,基于空间金字塔多尺度残差网络的编码器能够深度挖掘源图像的特征信息,并由解码器重构出源图像;在测试阶段,由设计的通道注意力机制融合源图像的基础和细节信息,再通过解码器输出融合图像。最后,在公开图像融合数据集上

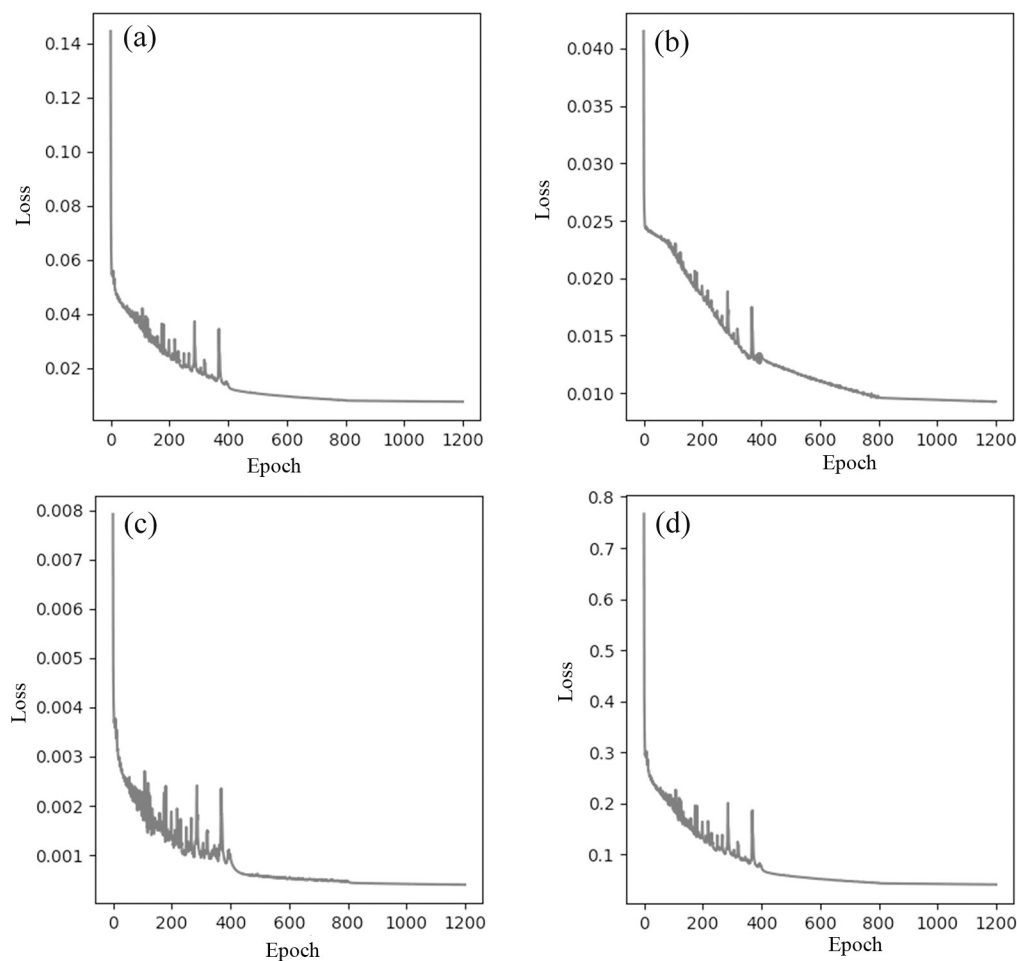


图 5 损失函数曲线。(a) 结构相似性损失; (b) 均方误差损失; (c) 梯度损失; (d) 总损失

Fig. 5 Loss function curve. (a) Loss of SSIM; (b) loss of MSE; (c) loss of gradient; (d) total loss

进行对比和消融实验, 证明了算法在提升融合图像质量和主观视觉效果方面的优势。但所提算法在多曝光图像融合等其他类型的图像融合任务中性能欠佳。在之后的研究中, 将把算法进一步改进成通用的图像融合模型, 拓展应用到上述领域。

参考文献:

- [1] Castillo J C, Fernández-Caballero A, Serrano-Cuerda J, *et al.* Smart environment architecture for robust people detection by infrared and visible video fusion [J]. *Journal of Ambient Intelligence and Humanized Computing*, 2017, 8(2): 223-237.
- [2] Tian Y, Yang W J, Wang J. Image fusion using a multi-level image decomposition and fusion method [J]. *Applied Optics*, 2021, 60(24): 7466-7479.
- [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] Goodfellow I, Pouget-Abadie J, Mirza M, *et al.* Generative adversarial networks [J]. *Communications of the ACM*, 2020, 63(11): 139-144.
- [5] 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.
- [6] Huang G, Liu Z, Van Der Maaten L, *et al.* Densely connected convolutional networks [C]. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 2261-2269.
- [7] 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.
- [8] Zhou Z W, Rahman Siddiquee M M, Tajbakhsh N, *et al.* UNet++: A Nested U-Net Architecture for Medical Image Segmentation [M]. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Cham: Springer International Publishing, 2018: 3-11.
- [9] An W B, Wang H M. Infrared and visible image fusion with supervised convolutional neural network [J]. *Optik*, 2020, 219: 165120.
- [10] Ma J Y, Yu W, Liang P, *et al.* FusionGAN: A generative adversarial network for infrared and visible image fusion [J]. *Information Fusion*, 2019, 48: 11-26.
- [11] Xu D D, Wang Y C, Xu S Y, *et al.* Infrared and visible image fusion with a generative adversarial network and a residual network [J]. *Applied Sciences*, 2020, 10(2): 554.
- [12] He K M, Zhang X Y, Ren S Q, *et al.* Deep residual learning for image recognition [C]. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 770-778.
- [13] 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.
- [14] Ronneberger O, Fischer P, Brox T. U-net: Convolutional Networks for Biomedical Image Segmentation [M]. *Lecture Notes in Computer Science*. Cham: Springer International Publishing, 2015: 234-241.
- [15] Lin T Y, Dollár P, Girshick R, *et al.* Feature pyramid networks for object detection [C]. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 936-944.
- [16] Liu P F, Zhao H C, Cao F D. Blind deblurring of noisy and blurry images of multi-scale convolutional neural network [J]. *Infrared and Laser Engineering*, 2019, 48(4): 300-308.
刘鹏飞, 赵怀慈, 曹飞道. 多尺度卷积神经网络的噪声模糊图像盲复原 [J]. *红外与激光工程*, 2019, 48(4): 300-308.
- [17] Hu J, Shen L, Sun G. Squeeze-and-excitation networks [C]. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. June 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 7132-7141.
- [18] Woo S, Park J, Lee J Y, *et al.* CBAM: Convolutional Block Attention Module [M]. *Computer Vision-ECCV 2018*. Cham: Springer International Publishing, 2018: 3-19.
- [19] Liu P F, Zhao H C, Li P X. Hyperspectral images reconstruction using adversarial networks from single RGB image [J]. *Infrared and Laser Engineering*, 2020, 49(S1): 143-150.
刘鹏飞, 赵怀慈, 李培玄. 对抗网络实现单幅RGB重建高光谱图像 [J]. *红外与激光工程*, 2020, 49(S1): 143-150.
- [20] Xu H, Ma J Y, Le Z L, *et al.* FusionDN: A unified densely connected network for image fusion [J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, 34(7): 12484-12491.
- [21] Fu Y, Wu X J. A dual-branch network for infrared and visible image fusion [C]. *2020 25th International Conference on Pattern Recognition (ICPR)*. January 10-15, 2021, Milan, Italy. IEEE, 2021: 10675-10680.
- [22] Zhang X C, Ye P, Xiao G. VIFB: A visible and infrared image fusion benchmark [C]. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. June 14-19, 2020, Seattle, WA, USA. IEEE, 2020: 104-105.
- [23] Tian S W, Lin S Z, Lei H W, *et al.* Multi-band image synchronous super-resolution and fusion method based on improved WGAN-GP [J]. *Acta Optica Sinica*, 2020, 40(20): 2010001.
田嵩旺, 蔺素珍, 雷海卫, 等. 基于改进WGAN-GP的多波段图像同步超分与融合方法 [J]. *光学学报*, 2020, 40(20): 2010001.