

基于细节增强的多能 DR 图像融合网络



刘 祎^{1,2}, 刘宇航^{1,3}, 颜溶标^{1,2}, 桂志国^{1,2*}

(1. 中北大学省部共建动态测试技术国家重点实验室 太原 030051; 2. 中北大学信息与通信工程学院 太原 030051;
3. 中北大学计算机科学与技术学院 太原 030051)

【摘要】针对单能 X 射线无法对复杂工件各部位同时曝光成像的问题, 提出一种基于细节增强的多能 DR 融合网络-双编码巢式连接融合网络。该网络以 Inception 模块作为基础卷积层, 在双编码器的辅支路设计了可训练的 LOG 卷积模块, 用来提取多尺度边缘特征, 并将其补充到主支路以增强全局特征。在训练阶段, 提出一种基于图像块的局部能量一致性损失函数, 以减少输入、输出的局部性误差。融合时, 采用通道和空间注意力机制作为融合策略, 对双编码提取的多尺度增强特征进行融合, 并将融合后的多尺度特征输入嵌套连接的解码器进行重构。结果表明, 该融合网络具有细节增强效果, 能够完整地再现复杂工件的内部结构及缺陷。

关键词 注意力机制融合策略; 复杂工件; DR 图像融合; Inception 模块; LOG 卷积模块; 巢式连接
中图分类号 TP391 文献标志码 A doi:10.12178/1001-0548.2022127

Detail Enhancement-Based Fusion Network for Multi-Energy Digital Radiography Images

LIU Yi^{1,2}, LIU Yuhang^{1,3}, YAN Rongbiao^{1,2}, and GUI Zhiguo^{1,2*}

(1. State Key Laboratory of Dynamic Testing Technology, North University of China Taiyuan 030051;
2. School of Information and Communication Engineering, North University of China Taiyuan 030051;
3. School of Computer Science and Technology, North University of China Taiyuan 030051)

Abstract To solve the problem that X-ray with single energy could not simultaneously expose each part of a complex workpiece, a multi-energy digital radiography (DR) fusion network based on detail enhancement, namely dual-encoder nest connection-based fusion network, is proposed. In this network, the inception module is used as the basic convolution layer and a trainable LOG (Laplacian of Gaussian) convolution module is designed in the auxiliary branch of the dual-encoder to extract multi-scale edge features and add them to the main branch to enhance the global features. In the training stage, a local energy consistency loss function based on image block is proposed to reduce the local errors of input and output. In the fusion process, channel and spatial attention mechanisms are used as the fusion strategy to fuse the multi-scale enhanced features extracted from dual-encoder, and the fused multi-scale features are input into the nest connected decoder for reconstruction. Experimental results show that the proposed fusion network has the effect of detail enhancement and can reproduce the internal structure and defects of complex workpiece completely and clearly.

Key words attention module-based fusion strategy; complex workpiece; DR image fusion; Inception module; LOG convolution module; nest connection

X 射线数字成像技术 (digital radiography, DR) 已广泛应用于工业无损检测领域, 实现工件内部缺陷的检测^[1-3]。然而, 对于一些结构不规则、厚度变化范围大的复杂工件, 单一能量的 X 射线无法实现对整个工件同时曝光成像, 容易在有效厚度厚的

地方呈现曝光不足, 在有效厚度薄的地方出现曝光过度的现象, 导致被测工件的 DR 图像质量差、空间分辨率低、结构信息严重缺失^[4]。

为了解决单能 DR 图像信息量不足的问题, 国内外研究者进行了多能 DR 图像融合技术研究, 将

收稿日期: 2022-04-29; 修回日期: 2022-06-21

基金项目: 国家自然科学基金 (61801438); 山西省高等学校科技创新项目 (2020L0282)

作者简介: 刘祎 (1987-), 女, 博士, 副教授, 主要从事无损检测、图像处理与重建方面的研究。

*通信作者: 桂志国, E-mail: guizhiguo@nuc.edu.cn

多能 DR 图像中的有效信息结合起来,完整清晰地再现复杂工件的内部结构。如文献 [5] 提出了基于贝叶斯理论的数据融合技术,对双能 DR 图像进行融合。文献 [6-7] 针对不同电流得到的高、低能量 DR 图像,通过极大似然估计的方法获得融合图像。为解决双能 DR 成像无法满足等效厚度差异特别大的工件检测问题,文献 [8] 提出了一种变电压扫描的 X 射线图像融合方法,借助灰度-电压-厚度关系模型对多个能量下的 DR 图像序列进行加权融合,但此方法所确定的权值不够准确,需要进一步动态调整权值。在此基础上,文献 [9] 采用主成分分析方法对多能 DR 图像进行融合,在不需要任何先验知识的情况下,扩展了图像的动态范围。文献 [10] 提出了一种基于加权最小二乘滤波和引导滤波的铸件 DR 图像融合方法,通过加权求和的方式融合不同窗宽、窗位的 DR 图像。除了基于空间的融合方法,多尺度变换被用于图像融合中^[11-14]。这类方法首先通过变换提取多尺度特征,然后采用适当的融合策略对多尺度特征进行融合,最后对融合后的特征图像通过逆变换进行重构。如文献 [15] 针对不同电压下的 DR 图像利用小波变换进行多尺度分解,然后在小波域内对高、低频分量分别采用不同的规则进行融合,实现了图像动态范围的扩展。然而,多尺度变换方法的融合性能高度依赖于所使用的特征提取方法(变换方法)。基于字典学习的稀疏表示方法在图像融合中取得了较好的效果,这类方法通过提取图像块进行字典训练、稀疏编码,将图像融合问题转化为稀疏系数融合问题。如文献 [16] 提出了一种基于分类图像块稀疏表示的医学图像融合方法,文献 [17] 基于稀疏系数得到源图像的全局和局部显著性映射,进而指导红外与可见光图像融合。文献 [18] 将多尺度变换方法与稀疏表示方法相结合,提出了一种多尺度变换和稀疏表示相结合的通用图像融合框架。对于多能 DR 图像融合,文献 [19] 提出了一种基于模糊熵和稀疏表示的图像融合方法,利用模糊熵设计融合权重对稀疏系数融合,通过调节参数可以获得对比度较高的融合效果,但该方法会在灰度陡变处引入严重的黑色晕影。

随着深度学习技术的发展,许多基于神经网络的图像融合方法被提出。一种较为简单的方法是利用预训练的卷积神经网络进行特征提取,以获得特征权重图,然后以加权融合方式得到最终结果^[20-22]。但是这些方法本身没有训练阶段,采用

的预训练模型可能并不适合融合任务,相比之下,通过训练合适的图像融合模型可以获得更好的融合性能。如文献 [23] 针对红外与可见光融合任务,提出了一种基于稠密块和自编码器结构的融合框架 DenseFuse。文献 [24-26] 基于生成抵抗网络 (generative adversarial network, GAN) 分别提出了 FusionGAN、FusionGANv2 和 DDcGAN 网络,对红外与可见光具有良好的融合效果。文献 [27] 在 Unet++ 巢式连接结构的基础上,提出了一种基于注意力机制的融合网络 (NestFuse)^[28],在细节信息保存方面有很好的性能,紧接着在 NestFuse 基础上提出了基于残差融合网络的融合训练策略^[29],通过两阶段的训练策略来训练融合网络。

然而上述基于神经网络的融合方法多是针对红外/可见光图像融合任务、多焦点融合任务提出的,针对多能 DR 图像的融合网络较少,若直接将上述网络用于多能 DR 图像融合,并不能得到清晰的融合结果。这是由于 DR 图像本身具有对比度低、边缘不清晰等特点,即使能够提取多尺度深度特征,也无法改变源图像对比度低、细节不清的事实。针对上述问题,本文提出了一种基于细节增强的双编码巢式连接融合网络。该网络采用 Inception 结构^[30] 作为基础卷积块,在巢式连接架构的基础上,设计了密集连接的双编码特征提取网络。为增强融合图像的细节,在双编码器的辅支路中设计了一个可训练的边缘检测 LOG 卷积模块,用于提取边缘和细节特征,并对主支路提取的图像特征进行补充和增强。在训练阶段,为提高网络性能,提出了一种基于图像块的局部能量一致性损失函数,以减少编解码网络输出与输入的局部性误差。融合时,采用通道和空间注意力机制作为融合策略,对双编码提取的多尺度增强特征进行融合,并将融合后的多尺度特征输入嵌套连接的解码器进行重构。该网络可以对任意多幅 DR 图像进行融合,融合后的 DR 图像具有细节增强效果,能够完整清晰地再现复杂非等厚工件的内部结构。

1 相关工作

1.1 巢式连接架构

巢式连接架构最初被用于医学图像分割任务^[27]。在深度学习网络中,跨越连接可以保留更多信息,然而当跨越连接的跨度较长时,语义鸿沟会造成不理想的结果。因此,巢式连接采用上采样和短跨越

的连接方式代替长跨越连接, 如图 1 所示。

巢式结构利用密集的跨越连接可以限制特征图之间语义鸿沟的负面影响, 同时让更多的信息被保留, 从而获得更好的融合结果。

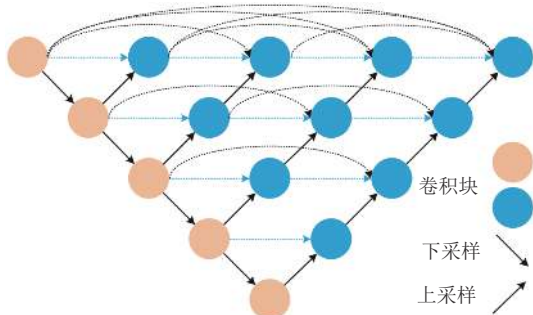


图 1 巢式连接架构

1.2 Inception 模块

Inception 模块通过多个卷积核提取图像不同尺度的信息进行融合, 得到图像更好的表征。如图 2 所示, Inception 模块具有 4 条分支, 其中每个卷积块均由卷积层、激活函数和批量归一化层组成。每条支路都使用 1×1 卷积核来降低通道数, 减少参数量, 前 3 条支路具有 3 种不同大小的感受野, 即 1×1、3×3 和 5×5, 其中第 3 条支路用两个 3×3 卷积核代替 5×5 卷积核, 在达到相同作用的同时进一步减少参数量。对于第 4 条支路, 虽然与第 2 条支路的感受野相同, 获得相同维数的输出, 但输出的内容却不同, 增加了特征融合多样性。最终, 将 4 条支路上提取的不同尺度特征累加作为 Inception

模块的输出。Inception 模块使得网络的深度和宽度也都得到了增加, 提升了网络的泛化能力。

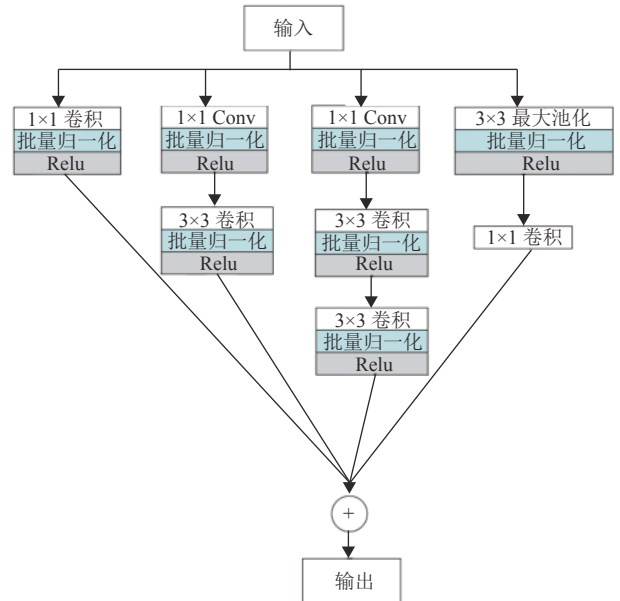


图 2 Inception 模块

2 本文方法

2.1 融合网络架构

本文提出的融合网络采用巢式连接架构, 网络结构框架如图 3 所示, 包含 3 个部分: 编码器、融合策略以及解码器。为了在提取特征时能够捕获更大的感受野, 使提取特征的语义更加丰富, 编码器和解码器中的卷积块均采用 Inception 模块作为卷积块, 实现多尺度特征提取和重构。

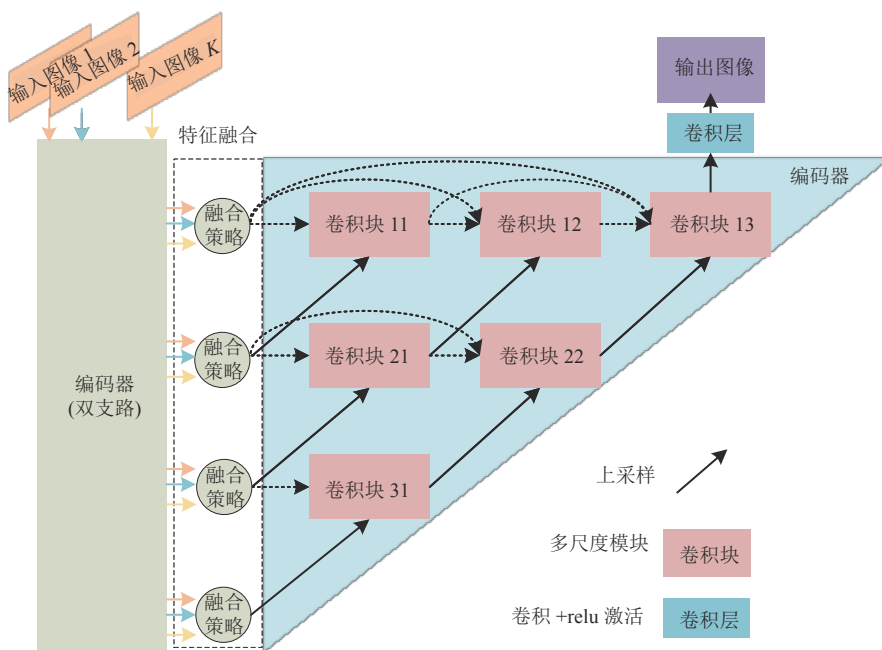


图 3 融合网络结构框架示意图

编码器是由特征提取的主支路和边缘特征提取的辅支路组成的双支路网络，如图 4 所示。编码器中右侧的主支路先用一个 1×1 卷积层增加输入图像的特征通道数，然后用 4 个卷积块提取图像的全局特征，每个卷积块的输出通过最大值池化进行下采样，以获取网络传输方向上的多尺度特征。为增强特征复用能力，编码器在主支路上添加密集的旁路连接。左侧的辅支路用于提取边缘和细节特征，在该支路中，本文在 5×5 的 LOG 边缘检测算子的基础上，通过逐元素乘以一个学习因子 α ，设计了可训练的 LOG 卷积层，如图 5 所示。在 LOG 算子的基础上设计可训练边缘提取层，是因为 LOG 算子在检测边缘时可实现对噪声的抑制，防止噪声对边缘特征的干扰。 α 初始值设为 1，通过网络反向传播机制，其数值可随网络训练自适应地调整。LOG 卷积在单通道上享有同一可学习因子 α ，但不同通道上的 α 经训练后不再相同。LOG 卷积逐通道作用于输入图像，由于每个通道的 α 不同，输出包含不同强度边缘信息的特征图。边缘特征提取辅支路首先经过 LOG 层提取浅层边缘特征，然后经过 4 个卷积块并采用密集的旁路连接提取深层多尺度边缘细节信息。辅支路结构与主支路结构对称，因此对应深度的特征图具有相同维度。为增强图像细节特征，本文将辅支路各层的输出特征图都添加到主支路的对应层的特征图上，对主支路提取的图像特征进行补充和增强，得到多尺度增强特征图。最后，添加了边缘特征信息的各层特征图进入解码器重构图像。

器中浅层的特征图相结合，避免了编码器和解码器之间的语义鸿沟，保证编码器网络能够在不同尺度上提取更多的深度特征，保留更多的信息，以获得更好的融合结果。

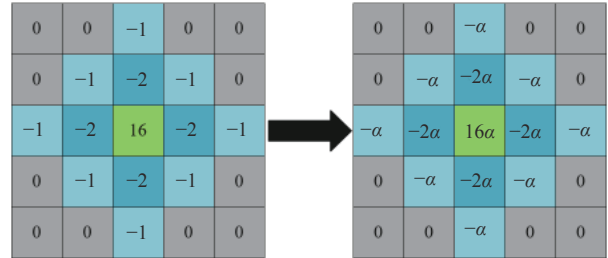


图 5 可学习的 LOG 卷积

2.2 网络训练

本文提出的融合方法为两阶段：网络训练阶段和融合阶段。在训练阶段，图 3 中编解码器间的融合策略被移除，网络只包含编解码器，通过训练得到一个自动编解码器网络，训练框架如图 6 所示。在这个网络中，编码器能够提取多尺度的、边缘增强的深度特征，解码器从这些特征中重建输入图像，网络参数设置如表 1 所示。

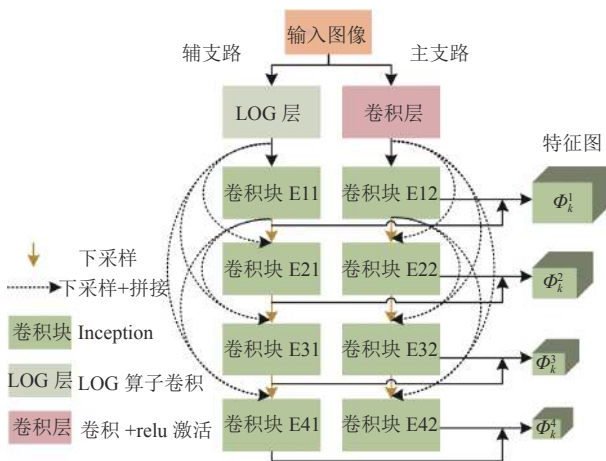


图 4 编码器示意图

解码器网络由 6 个卷积块和一个卷积层组成，上采样采用插值方法，并采用巢式连接结构将 6 个卷积块连接起来，让解码器中深层的特征图和编码

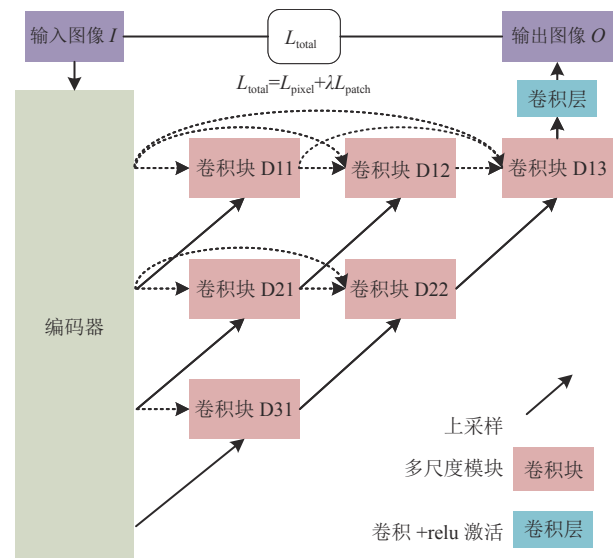


图 6 训练过程框架图

在训练阶段，为了提高编解码器网络的性能，本文提出了一种新的基于图像块的损失函数，称为局部能量一致性损失 L_{patch} ，其定义为输入图像和输出图像“局部能量图”的 L_1 范数值，具体为：

$$L_{patch} = \|E_{local}(I) - E_{local}(O)\|_{L_1} \quad (1)$$

$$E_{local}(I(x,y)) = \frac{1}{8} \sum_{s,t \in N_{x,y}} (I(x,y) - I(s,t))^2 \quad (2)$$

式中, E_{local} 为局部能量函数; $E_{\text{local}}(\mathbf{I})$ 和 $E_{\text{local}}(\mathbf{O})$ 分别为输入图像 \mathbf{I} 和输出图像 \mathbf{O} 的“局部能量图”; $E_{\text{local}}(\mathbf{I}(x,y))$ 为图像 \mathbf{I} 中像素点 (x,y) 的局部能量; $N_{x,y}$ 为像素点 (x,y) 的邻域, 能量图中的每一个点由图像 \mathbf{I} 中该点像素与其 8 邻域像素的差值平方和计算得到。

表 1 编解码器的网络参数设置

器件	卷积块	大小	步长	输入通道数	输出通道数
编码器	LOG层	卷积层	3	1	64
	卷积块E11	卷积块E12	-	-	64
	卷积块E21	卷积块E22	-	-	128
	卷积块E31	卷积块E32	-	-	192
	卷积块E41	卷积块E42	-	-	256
解码器	卷积块D31	-	-	128	64
	卷积块D21	-	-	128	64
	卷积块D22	-	-	192	64
	卷积块D11	-	-	128	64
	卷积块D12	-	-	192	64
	卷积块D13	-	-	256	64
	卷积层	3	1	64	1

由于局部能量一致性损失 L_{patch} 是基于图像块计算得到的, 它体现了网络输出与输入的局部性误差。本文将其与基于像素一致性的损失函数一起指导网络训练, 复合损失函数定义为:

$$L_{\text{total}} = L_{\text{pixel}} + \lambda L_{\text{patch}} \quad (3)$$

式中, λ 为权重参数; L_{pixel} 为基于像素水平的全局一致性损失函数, 其定义为:

$$L_{\text{pixel}} = \|\mathbf{O} - \mathbf{I}\|_{\text{F}}^2 \quad (4)$$

L_{pixel} 计算了 \mathbf{I} 和 \mathbf{O} 的 Frobenius 范数, 可以保证编解码网络的重构图像在像素级水平上更接近输入图像。 L_{pixel} 体现了网络输出与输入的全局性误差, 而局部能量一致性损失 L_{patch} 体现了网络输出与输入的局部性误差。因此, 复合损失函数不仅可以在全局一致性方面指导网络训练, 也可以在局部一致性方面进行约束。由于 L_{pixel} 和 L_{patch} 之间的数量级不同, 利用 λ 权衡 L_{pixel} 和 L_{patch} 之间的比例。

网络训练时, 为了能够训练出更适用于 DR 图像的编码器网络和解码器网络, 采集了 8 000 张 1 024×1 024 的 DR 图像作为训练数据集, 这些图像来自 7 种不规则工件的多能 DR 图像, 由于图像尺寸较大, DR 图像先被缩放到 256×256 大小, 然后再输入网络进行训练。本文使用 SGD 算法对损失函数进行了优化, 学习率初始化为 10^{-4} , 每 2 个 epoch 下降为上一次学习率的 1/3, 复合损失函数

的权重参数 λ 设为 10, 训练总次数设为 10 epoch。参数 λ 的数值通过多次实验后设定, 实验时设置了不同数量级的 $\lambda(1, 10, 100, 1\ 000, 10\ 000)$, 实验结果证实: 当参数 $\lambda=10$ 时, 融合结果在人眼视觉和质量评价参数方面整体最优。本文网络参数存储空间为 13.79 MB。

实验平台采用配置为 Intel Core i9-10900K CPU, 主频 3.7 GHz, 运行内存 64 GB 以及 NVIDIA RTX3090 显卡, 搭载 64 位 Windows10 系统的台式机, 编程环境为 Python 3.9.2, 深度学习工具包为 Pytorch 1.8.1。

2.3 融合策略

训练阶段的目的是获得自动编解码器网络, 并没有对多能 DR 图像进行融合。融合时, 在训练好的编码器和解码器之间添加融合策略, 对双编码提取的多尺度增强特征进行融合, 并将融合后的多尺度特征输入到解码器进行重构。

本文采用基于通道注意力和空间注意力的融合策略。空间注意力告诉网络“融合哪里”, 而通道注意力告诉网络“融合什么”, 两种融合机制互补, 从而抑制图像中曝光不足、过曝光等不良区域, 同时突出待融合图像中良好的特征区域。由于基于注意力机制的融合策略不需要学习参数, 因此在训练阶段不需要对融合策略进行训练。

本文提出的融合网络可对多张 DR 图像进行融合, 假设将待融合的 K 张不同能量的 DR 图像输入模型, 由于输入的图像数量是 K 张, 编码器的每一层多尺度 Inception 结构都会输出 K 组特征图, 利用融合策略对每一层的 K 组特征图进行融合, 即融合策略模块的输入为 K 组特征图。令 $\Phi_k^m \in R^{W \times H \times C}$ ($m \in \{1, 2, 3, 4\}, k \in \{1, 2, \dots, K\}$) 表示第 k 组第 m 层的特征图, W 、 H 和 C 分别为该层特征图的长、宽及通道数。对每一层的 K 组特征图分别使用通道注意力模块和空间注意力模块进行融合, 并将两个融合模块的输出取均值, 作为特征图融合结果, 融合过程如图 7a 所示。

通道注意力模块使用平均池化对 K 组特征图 $\Phi_k^m \in R^{W \times H \times C}$ 计算各自的通道注意力系数 $\hat{\beta}_k^m \in R^{1 \times 1 \times C}$, 如式 (5) 所示, 然后用当前特征图的注意力系数除以所有注意力系数之和得到归一化的通道注意力系数 $\beta_k^m \in R^{1 \times 1 \times C}$, 并以此作为通道融合的权重, 如式 (6) 所示。再将特征图与归一化的通道注意力系数相乘、累加, 得到通道方向上的特征融合结果 $\tilde{\Phi}^m \in R^{W \times H \times C}$, 如式 (7) 所示。

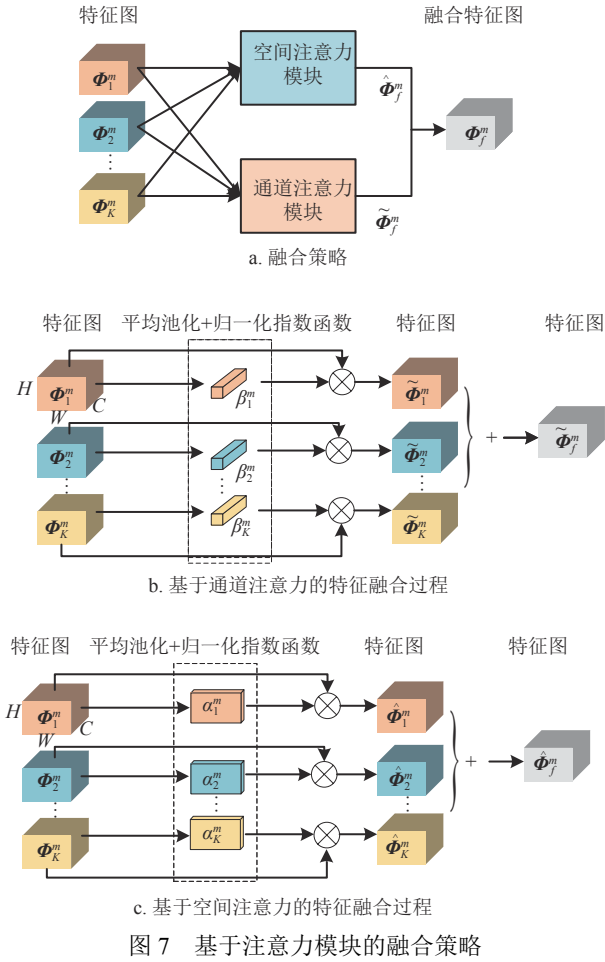


图7 基于注意力模块的融合策略

$$\tilde{\beta}_k^m(c) = P(\Phi_k^m(c)) \quad (5)$$

$$\beta_k^m = \tilde{\beta}_k^m / \sum_1^K \tilde{\beta}_k^m \quad (6)$$

$$\tilde{\Phi}_f^m = \sum_{k=1}^K \beta_k^m(c) \times \Phi_k^m(c) \quad (7)$$

式中, $P(\cdot)$ 是全局池化操作; $c \in \{1, 2, \dots, c\}$ 为特征图 Φ_k^m 和注意力系数 β_k^m 的第 c 个通道。通道方向上的融合过程如图 7b 所示。

除了通道方向上的融合外, 基于空间注意力模块进行多尺度特征图融合。类似地, 首先对 K 组特征图 $\Phi_k^m \in R^{W \times H \times C}$ 计算各自的空间注意力系数 $\tilde{h}_k^m \in R^{W \times H \times 1}$, 如式 (8) 所示, 然后用当前特征图的空间注意力系数除以所有空间注意力系数得到归一化的空间注意力系数 $h_k^m \in R^{W \times H \times 1}$, 如式 (9) 所示, 并将其复制 C 个通道以匹配特征图的维度, 输出的空间注意力系数为 $h_k^m \in R^{W \times H \times C}$ 。再将特征图与对应的空间注意力系数相乘、累加, 得到空间注意

力融合策略的结果 $\hat{\Phi}_f^m \in R^{W \times H \times C}$, 如式 (10) 所示, 融合过程如图 7c 所示。

$$\tilde{h}_k^m = \|\Phi_k^m(x, y)\|_1 \quad (8)$$

$$h_k^m = \tilde{h}_k^m / \sum_1^K \tilde{h}_k^m \quad (9)$$

$$\hat{\Phi}_f^m = \sum_{k=1}^K h_k^m \times \Phi_k^m \quad (10)$$

式中, $\|\cdot\|_1$ 指 L1 范数; (x, y) 指多尺度特征的对应位置。

然后将两个融合模块的输出取均值, 作为该层最终的特征融合结果 Φ_f^m :

$$\Phi_f^m = (\tilde{\Phi}_f^m + \hat{\Phi}_f^m) / 2 \quad (11)$$

3 实验对比及分析

3.1 实验设置

为验证本文所提网络的性能, 对没有参与训练的 3 种工件, 在不同 X 射线电压下分别采集 7 个照射角, 共 21 组 DR 图像进行测试, 采集参数如表 2 所示。图 8 显示了来自 3 种测试工件的 3 组待融合图像, 可以看出, 当电压较小时, DR 图像中较厚的区域较黑, 难以感知该部分的细节信息; 当电压增大时, 厚区域的细节能够依次显现出来, 但是薄区域处产生了过曝光。为了验证算法的可行性与有效性, 将提出的融合网络与现有的 9 种融合方法进行对比, 这些对比方法包括: 基于结构块分解的多曝光图像融合^[31]、基于多尺度结构块分解的多曝光图像融合^[32]、基于自适应权值的多曝光图像融合^[33]、基于密集 SIFT 和引导滤波的多曝光图像融合^[34]、多尺度变换和稀疏表示的图像融合^[18]、基于模糊熵和稀疏表示的图像融合方法^[19]、IFCNN 网络^[35]、DenseFuse 网络^[23] 以及 NestFuse 网络^[27]。上述这些融合方法均在公开的代码基础上实现, 前 6 种方法为传统方法, 相关参数根据能达到的最优视觉效果进行设置, 后 3 种为基于深度学习的融合方法, 均在本文的 DR 数据集上进行训练得到融合网络模型。

表2 测试数据扫描参数设置

工件	能量个数	图像大小	电压/Kvp	电流/mA
工件1	3	1 024×1 024	150、190、240	1
工件2	3	1 024×1 024	120、140、180	1
工件3	2	1 024×1 024	120、140	1

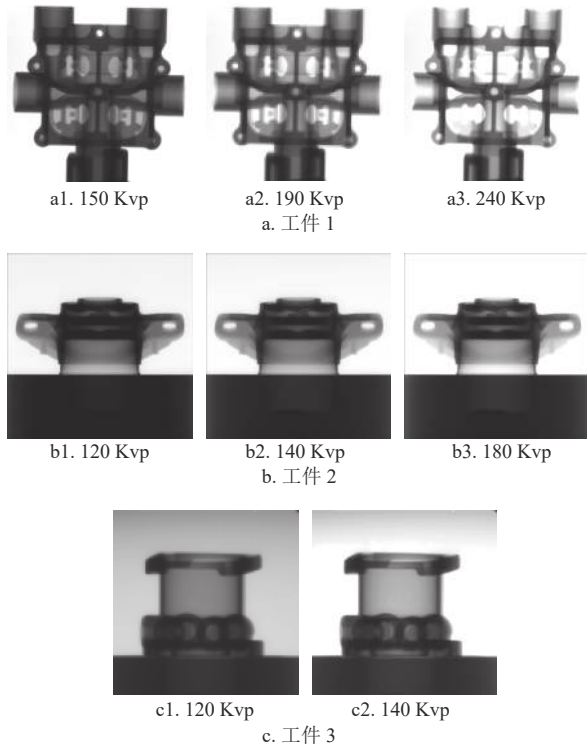


图8 3组待融合的DR图像

3.2 算法对比

图9~图11显示了不同融合方法对图8中3组待融合DR图像的融合结果。为了更好地观察

融合结果, 截取感兴趣区域 (region of interest, ROI) 并放大显示在图中的顶部或底部, ROI 由矩形标记。可以看出, 基于结构块的方法 (文献 [31] 和文献 [32]) 得到的融合图像具有较好的对比度和完整性, 但细节对比度有待提高, 如图 9a、10a、11a 和图 9b、10b、11b 所示。自适应权值方法 (文献 [33]) 的融合图像整体灰度值较低, 工件中厚的区域较暗, 边界比较模糊, 如图 9c、10c、11c 所示。文献 [34] 方法的融合图像整体较亮, 工件中较厚区域比较清晰, 但是工件较薄区域过于明亮, 不能很好地显示细节部分, 如图 9d 所示。相比之下, 两种基于字典学习和稀疏表示的融合方法 (文献 [18] 和文献 [19]) 在结构完整性方面比前几种方法好, 能够保留一些细小的结构, 如图 9d 和图 9f 中箭头指向的部位。特别是文献 [19] 的方法, 在对比度方面有了很大的提升, 不足的是该方法在灰度陡变处产生了黑色晕影, 如图 9f 中椭圆所圈之处。IFCNN、DenseFuse 和 NestFuse 这 3 种融合网络的整体融合效果相近, 其中 IFCNN 的对比度稍好, 但仍然无法清晰地显示较厚区域的内部结构。图 9j、10j、11j 为利用本文提出的融合网络得到的融合图像, 不仅工件的结构信息完整, 而且对比度较高, 图像完整清晰地显示复杂工件的信息。

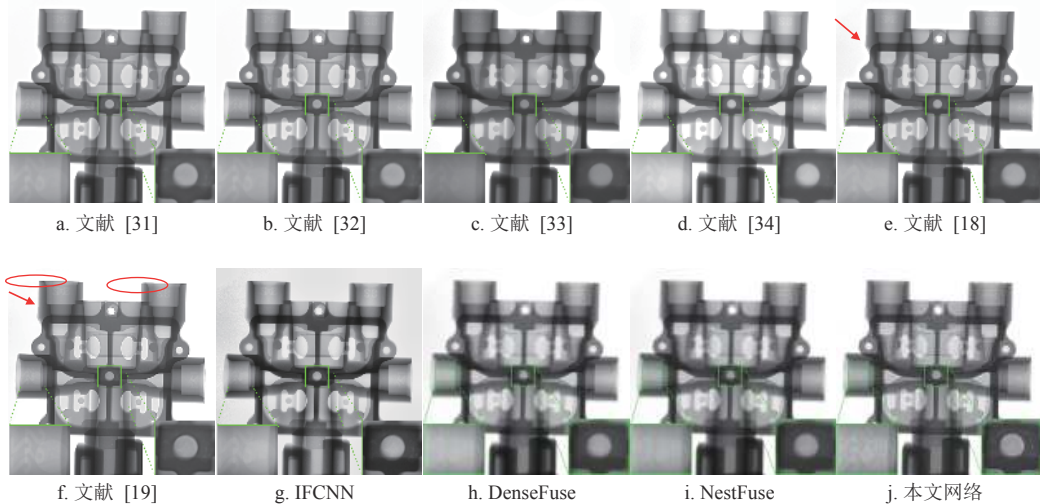


图9 工件1的融合结果

为了进一步验证算法的可行性与有效性, 本文采用多种量化指标对不同方法的融合结果进行定量分析, 包括: 熵 (Entropy)^[36]、平均梯度 (average gradient, AG)^[37]、无参考结构清晰度 (no-reference structural sharpness, NRSS)^[38]、互信息 (mutual information, MI)^[39] 以及基于视觉感知一致

性的视觉信息保真度 (visual information fidelity, VIF)^[40]。熵值体现了图像中的信息量, 熵越大表示图中包含的信息越多。AG 体现了图像中的边缘锐度, AG 越大表明边缘越清晰。NRSS 利用低通滤波器为待评估图像构造参考图像, 进而通过评价参考图像与待评估图像的结构相似度来评价图像

的清晰度, NRSS 越大表明图像结构越清晰。AG、NRSS 和 Entropy 值均为无参考图像的质量评价指标, 而 MI 和 VIF 为需要参考图像的质量评价

指标, 分别表示融合图像与源图之间的相关性以及保真度, MI 和 VIF 越大表明融合图像保留了源图像中更多的特征信息。

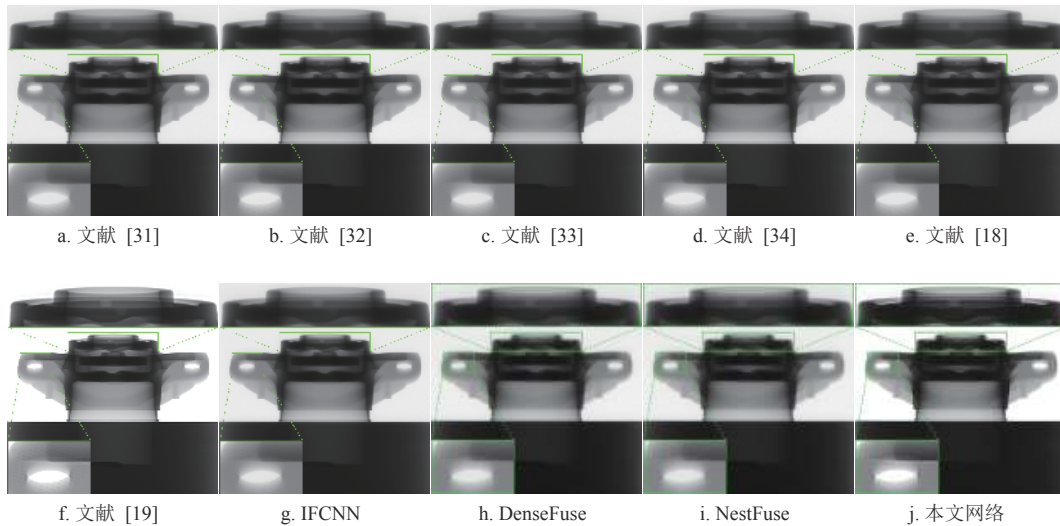


图 10 工件 2 的融合结果

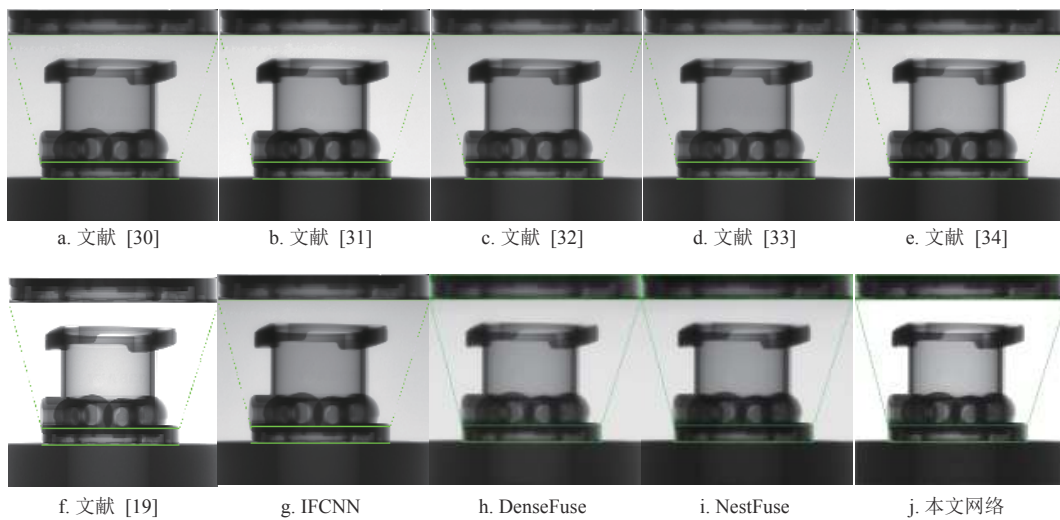


图 11 工件 3 的融合结果

表 3~表 5 列出了图 9~图 11 中不同方法融合结果的质量评价参数, 最好的指标值用粗体标记。可以看出, 在每一组实验中, 本文提出的融合网络都有最多的最优值和次优值, 表明提出的融合网络能够在融合后的图像中保留更多的细节信息和特征信息。

本文方法是基于深度学习的融合网络, 因此对几种网络方法进行了融合时间对比: 当融合两张图时, 本文方法为 1.80s, Densefuse 为 0.96s, NestFusion 为 1.10 s, IFCNN 为 1.00 s。由于本文采用了双路编码器, 因此融合时间比其他 3 种网络方法长。

表 3 不同算法融合工件 1 的质量评价参数

融合方法	Entropy	AG	NRSS	MI	VIF
文献[31]	4.949 3	1.433 8	0.740 5	17.959 8	0.952 9
文献[32]	5.603 6	1.569 6	0.812 9	16.819 0	1.006 5
文献[33]	5.815 1	1.318 1	0.823 0	18.656 2	0.760 9
文献[34]	5.447 5	1.401 5	0.822 8	16.479 5	0.845 8
文献[18]	5.094 7	1.559 7	0.800 1	18.448 6	0.948 1
文献[19]	5.885 5	1.728 2	0.806 3	17.656 5	1.072 9
IFCNN	5.835 9	2.210 8	0.795 9	20.818 6	1.056 2
DenseFuse	6.092 2	1.472 2	0.790 1	18.276 5	0.852 3
NestFuse	6.050 5	1.440 1	0.798 1	18.151 5	0.847 3
本文网络	6.551 7	2.609 2	0.805 9	19.655 0	1.124 2

表 4 不同算法融合工件 2 的质量评价参数

融合方法	Entropy	AG	NRSS	MI	VIF
文献[31]	6.277 9	0.924 5	0.806 4	18.833 6	0.926 5
文献[32]	6.424 4	0.978 0	0.770 3	19.273 3	0.943 6
文献[33]	6.274 2	0.907 6	0.791 9	18.822 6	0.909 7
文献[34]	6.377 9	0.921 6	0.791 9	19.133 6	0.934 2
文献[18]	6.377 5	0.976 3	0.780 0	19.132 5	0.934 3
文献[19]	5.280 6	<i>2.143 1</i>	0.848 5	15.841 9	<i>1.385 1</i>
IFCNN	6.273 3	0.909 9	0.802 8	18.820 0	0.907 7
DenseFuse	6.499 3	0.898 6	0.793 6	19.497 9	0.988 4
NestFuse	6.608 2	0.876 1	0.791 0	19.824 6	0.966 6
本文网络	6.544 0	2.385 5	0.821 4	19.632 0	1.559 3

表 5 不同算法融合工件 3 的质量评价参数

融合方法	Entropy	AG	NRSS	MI	VIF
文献[31]	7.161 4	1.101 7	0.746 0	14.322 8	1.085 2
文献[32]	7.368 4	1.144 8	0.733 1	14.736 9	1.170 1
文献[33]	7.293 7	0.988 0	0.729 5	14.587 5	1.004 1
文献[34]	7.236 0	1.074 6	0.739 4	14.472 0	1.126 6
文献[18]	7.341 7	1.216 2	0.771 6	14.695 7	1.189 3
文献[19]	6.984 9	2.493 1	0.794 9	13.977 9	1.170 2
IFCNN	7.315 2	1.004 1	0.690 7	14.630 4	0.977 3
DenseFuse	7.090 7	0.901 0	0.732 3	14.181 4	0.998 1
NestFuse	7.099 5	0.964 4	0.752 1	14.199 0	1.091 6
本文网络	7.315 0	3.110 3	0.851 0	14.630 0	1.905 9

4 消融实验

本文提出的双编码巢式连接融合网络与 NestFuse 均是在巢式连接架构的基础上提出来的, 与 NestFuse 相比, 在编解码器中采用多尺度 Inception 结构作为基础卷积块, 代替了 NestFuse 网络中的两层卷积操作, 并在编码器中通过设计 LOG 卷积模块增加了边缘特征提取支路。为了研究不同模块对网络性能的影响, 本文进行了一组消融实验。记 *A* 表示在本文提出的损失函数指导下进行训练的 NestFuse 网络, *B* 表示在本文网络的基础上去掉编码器中的边缘提取支路网络, *C* 表示本文提出的融合网络。将 *A*、*B* 两个网络在 8 000 张 DR 图像组成的数据集上进行训练, 并从量化指标与视觉效果两方面比较其与本文网络的测试结果。

本文利用 5 个质量指标对消融网络的融合性能进行评价, 表 6 给出了消融网络及本文网络对 21 组 DR 图像融合结果的平均质量评价指标, 指标的最好值用粗体标出。从表 6 中可以看出, *B* 与 *A* 的指标接近, 除了 VIF 外, 其他指标都略高于 *A*, 说明 Inception 结构可以在一定程度上提高网络

的融合性能。相比之下, *C* 的编码器中增加了边缘特征提取支路, 除了 Entropy 小幅度提高, 其他融合指标都有明显的提升, 说明边缘特征提取支路在融合图像中的对比度体现方面起着决定性作用。

表 6 网络结构调整对融合结果的影响

网络	Incep	LOG	Entropy	AG	NRSS	MI	VIF
<i>A</i>	×	×	6.371 1	1.027 0	0.778 6	17.688 3	0.942 2
<i>B</i>	√	×	6.481 6	1.030 6	0.793 4	18.014 7	0.940 2
<i>C</i>	√	√	6.558 1	2.285 4	0.848 2	18.211 2	1.314 0

消融网络融合图像的视觉效果如图 12 所示, 显示了 *A*、*B*、*C* 网络对图 8 中的 3 组 DR 图像的融合效果, 可以看出 *A* 网络和 *B* 网络的融合结果相似, 但都不具备增强效果, 而本文网络降噪图像更清晰、对比度更强, 说明本文网络具有更好的融合性能。

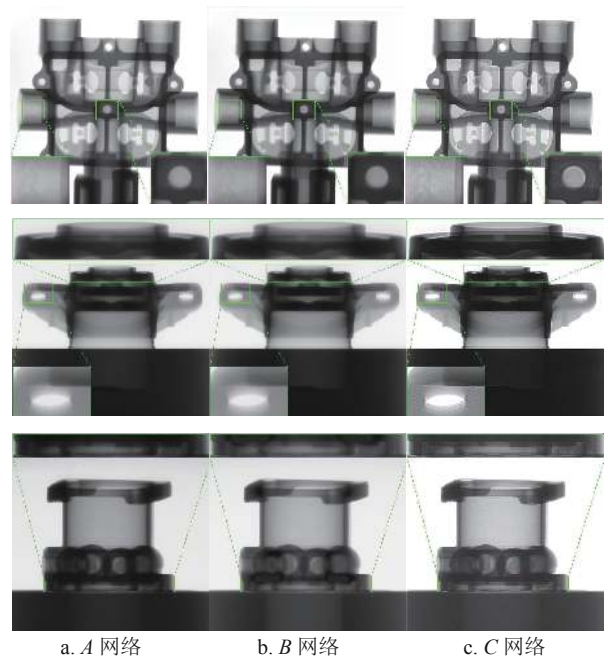


图 12 消融实验结果图

5 结束语

本文提出了一种基于细节增强的多能 DR 融合网络。该融合网络不仅能够完整再现复杂工件的内部结构, 而且增强了融合图像的细节, 便于发现工件内部的缺陷。

然而, 本文提出的融合网络是通过两阶段实现的, 即网络训练和融合过程是分开进行的, 并没有对融合策略模块进行训练, 未来需要改进融合策

略, 加入能够训练的网络模块, 实现端到端的融合网络。

参 考 文 献

- [1] HAMIDEEN M S, SHARAF J, AL-SALEH K A, et al. Description of a transmission X-Ray computed tomography scanner[J]. *Radiation Physics and Chemistry*, 2011, 80(11): 1162-1165.
- [2] ZARB F, RAINFORD L, MCENTEE M F. Image quality assessment tools for optimization of CT images[J]. *Radiography*, 2010, 16(2): 147-153.
- [3] LIU Y, ZHANG P C, GUI Z G. An enhancement framework based on gradient domain tone mapping and fuzzy logical for X-Ray image of complex workpiece[J]. *NDT and E International*, 2021, 121: 102455.
- [4] BUTLER A P H, BUTZER J, SCHLEICH N, et al. Processing of spectral X-Ray data with principal components analysis[J]. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2011, 633(2): S140-S142.
- [5] DROMIGNY A, ZHU Y M. Improving the dynamic range of real-time X-Ray imaging systems via Bayesian fusion[J]. *Journal of Nondestructive Evaluation*, 1997, 16(3): 147-160.
- [6] SISNIEGA A, VAQUERO J J, ABELLA M, et al. Automated dual-exposure technique to extend the dynamic range of flat-panel detectors used in small-animal cone-beam micro-CT[C]//2009 IEEE Nuclear Science Symposium Conference Record, VOLS 1-5. Orlando, FL: IEEE, 2009: 2948-2950.
- [7] SISNIEGA A, ABELLA M, DESCO M, et al. Dual-Exposure technique for extending the dynamic range of X-Ray flat panel detectors[J]. *Physics in Medicine and Biology*, 2014, 59(2): 421-439.
- [8] 段彦杰, 杨明, 刘宾, 等. 基于变剂量扫描的 X 射线图像增强技术[J]. *无损检测*, 2011, 33(2): 13-15.
DUAN Y J, YANG M, LIU B, et al. X-Ray image enhancement technology based on variable dose[J]. *NDT*, 2011, 33(2): 13-15.
- [9] 魏交统, 陈平, 潘晋孝. 基于主成分分析的递变能量 X 射线图像融合[J]. *中国体视学与图像分析*, 2013, 18(2): 103-108.
WEI J T, CHEN P, PAN J X. Gradient-Energy digital radiography image fusion based on principle component analysis[J]. *Chinese Journal of Stereology and Image Analysis*, 2013, 18(2): 103-108.
- [10] 羊肇俊, 曾理. 基于加权最小二乘滤波和引导滤波的铸件 DR 图像融合[J]. *仪器仪表学报*, 2021, 42(6): 211-220.
YANG Z J, ZENG L. Casting DR image fusion based on weighted least squares filter and guided filter[J]. *Chinese Journal of Scientific Instrument*, 2021, 42(6): 211-220.
- [11] LUO X Q, ZHANG Z C, ZHANG B C, et al. Image fusion with contextual statistical similarity and nonsubsampling shearlet transform[J]. *IEEE Sensor Journal*, 2017, 17(6): 1760-1771.
- [12] LI X X, GUO X P, HAN P F, et al. Laplacian re-decomposition for multimodal medical image fusion[J]. *IEEE Transaction on Instrumentation and Measure*, 2020, 69(9): 6880-6890.
- [13] YANG S Y, WANG M, JIAO L C, et al. Image fusion based on a new contourlet packet[J]. *Information Fusion*, 2010, 11(2): 78-84.
- [14] LI S T, KANG X D, HU J W. Image fusion with guided filtering[J]. *IEEE Transactions on Image Processing*, 2013, 22(7): 2864-2875.
- [15] 杨氏, 吴美金, 魏东波, 等. 双能透照模式下涡轮叶片 DR 图像融合方法[J]. *北京航空航天大学学报*, 2011, 37(12): 1494-1497.
YANG M, WU M J, WEI D B, et al. Turbine blade DR images fusion based on bi-energy X-ray radiography[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2011, 37(12): 1494-1497.
- [16] ZONG J J, QIU T S. Medical image fusion based on sparse representation of classified image patches[J]. *Biomedical Signal Processing and Control*, 2017, 34: 195-205.
- [17] 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.
- [18] LIU Y, LIU S, WANG Z. A general framework for image fusion based on multi-scale transform and sparse representation[J]. *Information Fusion*, 2015, 24: 147-164.
- [19] ZHAO R G, LIU Y, ZHAO Z, et al. Multi-Energy X-ray images fusion method based on fuzzy entropy and sparse representation for complex castings[J]. *NDT and E International*, 2021, 124(102535): 1-9.
- [20] 马旗, 朱斌, 张宏伟. 基于 VGG 网络的双波段图像融合方法[J]. *激光与红外*, 2019, 49(11): 1374-1380.
MA Q, ZHU B, ZHANG H W. Dual-Band image fusion method based on VGGNet[J]. *Laser and Infrared*, 2019, 49(11): 1374-1380.
- [21] LI H, WU X J, DURRANI T S. Infrared and visible image fusion with ResNet and zero-phase component analysis[EB/OL]. [2022-03-12]. <https://doi.org/10.48550/arXiv.1806.07119>.
- [22] LI H, WU X J, KITTLER J. Infrared and visible image fusion using a deep learning framework[C]//24th International Conference on Pattern Recognition (ICPR). [S. l.]: IEEE, 2018: 2705-2710.
- [23] 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.
- [24] 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.
- [25] MA J Y, LIANG P W, YU W, et al. Infrared and visible image fusion via detail preserving adversarial learning[J]. *Information Fusion*, 2019, 54: 85-98.
- [26] 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.
- [27] ZHOU Z, SIDDIQUEE M, TAJBAKSH N, et al. UNet++: A nested U-Net architecture for medical image segmentation[C]//*Deep Learning in Medical Analysis and Multimodal for Clinical Decision Support: The 4th International Workshop and the 8th International Workshop*. Granada: Springer, 2018: 3-11.
- [28] 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.
- [29] 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.
- [30] CHRISTIAN S, VINCENT V, LOFFE S, et al. Rethinking the inception architecture for computer vision[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016: 2818-2826.
- [31] MA K, LI H, YONG H, et al. Robust multi-exposure image fusion: A structural patch decomposition approach[J]. *IEEE Transactions on Image Processing*, 2017, 26(5): 2519-2532.
- [32] LI H, MA K, YONG H, et al. Fast multi-scale structural patch decomposition for multi-exposure image fusion[J]. *IEEE Transactions on Image Processing*, 2020, 29: 5805-5816.
- [33] LEE S H, PARK J S, CHO N I. A multi-exposure image fusion based on the adaptive weights reflecting the relative pixel intensity and global gradient[C]//*25th IEEE International Conference on Image Processing (ICIP)*. Athens: IEEE, 2018: 1737-1741.
- [34] HAYAT N, IMRAN M. Ghost-Free multi exposure image fusion technique using dense SIFT descriptor and guided filter[J]. *Journal of Visual Communication and Image Representation*, 2019, 62: 295-308.
- [35] ZHANG Y, LIU Y, SUN P, et al. IFCNN: A general image fusion framework based on convolutional neural network[J]. *Information Fusion*, 2020, 54: 99-118.
- [36] ROBERTS J W, AARDT J V, AHMED F B. Assessment of image fusion procedures using entropy, image quality, and multispectral classification[J]. *Journal of Applied Remote Sensing*, 2008, 2(1): 1-28.
- [37] ZHANG F, XIE W, MA G, et al. High dynamic range compression and detail enhancement of infrared images in the gradient domain[J]. *Infrared Physics&Technology*, 2014, 67: 441-454.
- [38] 谢小甫, 周进, 吴钦章. 一种针对图像模糊的无参考质量评价指标[J]. *计算机应用*, 2010, 30(4): 921-924.
- XIE X F, ZHOU J, WU Q Z. No-Reference quality index for image blur[J]. *Journal of Computer Applications*, 2010, 30(4): 921-924.
- [39] PENG H, LONG F, DING C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2005, 27(8): 1226-1238.
- [40] HAN Y, CAI Y, CAO Y, et al. A new image fusion performance metric based on visual information fidelity[J]. *Information Fusion*, 2013, 14(2): 127-135.

编辑 税红