

基于残差密集融合对抗生成网络的 PET-MRI 图像融合

刘尚旺, 杨荔涵

(河南师范大学 计算机与信息工程学院; 智慧商务与物联网技术河南工程实验室, 河南 新乡 453007)

摘要: 为了增强核磁共振与正电子发射断层扫描图像融合的纹理细节, 摆脱人工设计融合规则对先验知识的依赖, 提出了自适应的残差密集生成对抗网络(adaptive dense residual generative adversarial network, ADRGAN)来融合两种模态的医学图像. ADRGAN 设计了区域残差学习模块与输出级联生成器, 在加深网络结构的同时避免特征丢失; 然后, 设计了基于自适应模块的内容损失函数, 强化输出融合图像的内容信息; 最后, 通过源图像的联合梯度图与融合图像的梯度图构建对抗性博弈来高效训练生成器与鉴别器. 实验结果表明, ADRGAN 在哈佛医学院 MRI/PET 数据集的测试中峰值信噪比和结构相似度分别达到 55.212 4 和 0.469 7, 均优于目前最先进的算法; 所构建的模型具有端到端和无监督两特性, 无需人工干预, 也不需要真实数据作为标签.

关键词: 深度学习; 对抗生成网络; 多模态图像融合; 密集残差网络

中图分类号: TP391

文献标志码: A

文章编号: 1000-2367(2024)01-0074-10

近年来随着医学成像技术与计算机科学的发展, 医学影像呈现出越来越多的模态, 而单模态医学图像所提供的信息有限, 不能反映相关组织的所有细节信息. 其中正电子发射断层扫描(positron emission computed tomography, PET)选用正电子核素标记葡萄糖等代谢物作为显像剂, 可以为临床诊断提供生物代谢信息, 但其灰度分辨率较低, 常使用伪彩色增强进行标识. 核磁共振图像(magnetic resonance imaging, MRI)虽不包含代谢信息, 但其软组织分辨率高于 PET, 可以更好地提供解剖学精细信息. 因此, 通过图像融合技术融合两种模态的医学图像的信息, 实现优势互补, 为医生提供高质量的临床诊断依据^[1].

在融合过程中, 需要将 MRI 数据中的组织结构信息引入 PET 中, 得到同时具有 MRI 图像的空间信息以及 PET 图像的代谢信息的融合结果^[2], 超越单模态医学图像的局限, 提高图像在诊断和评估医学问题上的临床适用性.

基于深度学习的研究在过去几年成为图像融合领域的一个活跃话题. 许多基于深度学习的融合方法相继被提出, 并逐渐形成了一个重要的分支. 尽管这些方法已经取得了不错成果, 但大部分融合规则的设计仍然是人工的. 因此, 整个方法不能摆脱传统融合方法的局限性. 利用深度学习进行图像融合最大障碍是缺失真实标签数据, 直接获取真实的 PET-MRI 融合图像是不可能的, 本文只能采用自监督的方式构建逻辑闭环训练模型, GAN 作为最有前景的生成式模型之一, 通过生成器与鉴别器的对抗性博弈, 能生成和辨别更真实的样本^[3-4].

受此启发, 本文设计了自适应的残差密集生成对抗网络(ADRGAN)来实现高质量的两种模态医学图像融合. 为了避免颜色失真, 首先将 RGB 通道的 PET 图像转化为 $YCbCr$ 模式, 以分离亮度通道(Y)与颜色通道(C_b, C_r). 将 Y 通道亮度图像与 MRI 图像输入到 ADRGAN 进行融合, 输出融合后的亮度图像 I_{fused} , 之后,

收稿日期: 2022-08-11; **修回日期:** 2022-08-30.

基金项目: 国家自然科学基金(U1304607); 河南省高等学校重点科研项目基础研究计划(21A520022).

作者简介(通信作者): 刘尚旺(1973-), 男, 河南新乡人, 河南师范大学副教授, 博士, 研究方向为计算机视觉、生物图像处理、深度学习, E-mail: shwl2012@hotmail.com.

引用本文: 刘尚旺, 杨荔涵. 基于残差密集融合对抗生成网络的 PET-MRI 图像融合[J]. 河南师范大学学报(自然科学版), 2024, 52(1): 74-83. (Liu Shangwang, Yang Lihan. Fusion of PET and MRI images using adaptive dense residual generative adversarial network[J]. Journal of Henan Normal University(Natural Science Edition), 2024, 52(1): 74-83. DOI:10.16366/j.cnki.1000-2367.2022.08.11.0003.)

将 I_{fused} 与 C_b 与 C_r 分量经过 RGB 分解的逆变换,生成 RGB 图像,随后与 (C_b, C_r) 反变换得到 RGB 通道的输出结果.生成器中包含了一个自适应决策块,用于判断像素是否被聚焦.另外本模型为筛选映射专门设计了与之对应的损失函数,从而约束生成器生成一个与清晰的源图像一致的融合图像.本文还建立了生成器和鉴别器之间的对抗博弈从而进一步增强融合图像的纹理细节.经过连续的对抗学习,融合图像的梯度映射与 PET-MRI 联合梯度映射逐渐相似,融合图像的纹理细节逐渐清晰.值得一提的是,由于小卷积核的使用以及特征通道的控制,网络参数被限制在一定范围内,能够更快速地完成融合任务.对比以往方法,本模型拥有以下优点:1)提出了一种用于医学图像融合的基于联合梯度的自适应生成对抗网络,使用端对端框架解决问题,不需要真实数据作为标签,不需要人为规则设计与后续处理,提高了模型鲁棒性.2)使用包含两幅源图像特征的联合梯度图与融合图像梯度图进行对抗博弈,解决了以往生成对抗网络丢失源图像特征的问题.3)基于医学图像融合对纹理细节的需求,设计了结构性内容损失函数以及对抗性损失函数来保护高频信息.融合图像轮廓清晰,更易于寻找病灶.

1 相关工作

传统医学图像融合领域应用最广泛的技术是像素级融合^[5].该技术可分为空间域与变换域两类.其中,前者融合规则直接作用于像素,规则简单,但融合效果较差.如,文献[6]中将图像转换为亮度、色度、饱和度(intensity, hue, saturation, IHS)通道,而 IHS 变换会造成颜色与空间失真^[7].基于变换域的图像融合技术大多采用多尺度变换(multi-scale transform, MST)技术,它分为分解、融合、重构 3 个过程.源图像首先变换到频域根据某种规则进行融合,然后使用融合后的系数与变换基进行图像重构;其很好地保护了源图像的细节信息,但忽视了空间一致性,导致融合图像亮度与颜色失真.传统融合方法的规则需要人为地设计与选择, MST 中选择不同的滤波器参数,得到的融合效果差异很大.但是,由于特征提取的多样性和融合规则的复杂性,人工设计融合方法变得困难,使得模型鲁棒性降低.随着近年来深度学习的崛起,神经网络被用来解决上述问题.

现有基于深度学习的图像融合大多使用卷积神经网络(convolutional neural network, CNN)构建模型.图像融合领域基于深度学习的研究在过去几年逐渐活跃^[8].学者们相继提出了众多融合方法,逐渐形成了一个重要的分支.在一些方法中,深度学习框架被用来以端到端的方式提取图像特征以进行重建.具有代表性的是,文献[9]将卷积稀疏表示(CSR)应用于图像融合,提取多层特征,并利用这些特征生成融合图像. IFC-NN 将卷积神经网络加入变换域图像融合算法^[10];文献[11]基于稀疏表示和孪生卷积神经网络提出了用于医学图像的融合方法,文献[12]选择在传统图像融合方案中加入深度学习技术, CNN 框架被用于融合低频系数与高频系数. Densefuse 包括卷积层、融合层密集块,编码器负责为网络提供输入,网络得到特征图后,通过解码器重建融合图像^[13];文献[14]通过组合 CNN 和 RNN 提取特征,然后融合进行分割^[15]首次将生成对抗网络引入到红外与可见光图像融合中. DDcGan 构筑了双鉴别器生成对抗网络^[16];文献[17]提出通过生成性对抗网络融合生物图像; PMGI 利用图像梯度与对比度提取信息,并在同一条路径上进行特征重用^[18].

虽然这些工作取得了很好的效果,但仍然存在一些不足:1)深度学习框架仅用来弥补传统融合方法的某些缺陷,例如特征提取,整个融合方法的设计仍旧是基于传统方法的^[19-20].2)由于标签数据的缺失,导致依赖损失函数设计的解决方案不全面.3)基于传统生成对抗网络的解决办法只能使结果与一幅源图像相似,另一幅源图像包含的部分信息丢失.

本文设计出一个无监督残差密集生成对抗网络 ADRGAN 来解决上述问题,该网络基于梯度约束且拥有自适应能力.

2 本文方法

2.1 预处理

将 PET 图像作为彩色图像处理,其颜色可以反映代谢功能.为了使融合图像拥有与 PET 图像相同的颜色信息.对 PET 图像进行预处理,将图像从 RGB 通道转换为 YC_bC_r 通道,分离色差与亮度,以避免颜色

失真.

YC_bC_r 是一个色彩空间,具有红、绿、蓝 RGB通道彩色图像可以转换为 y 、 C_b 和 C_r 分量,其中 y 是亮度分量, C_b 和 C_r 分别是蓝色和红色的色差分量.

对 Y 通道单独融合可不影响颜色空间中的色彩,可以用来解决色彩畸变问题. YC_bC_r 是处理颜色和知觉一致性的合适近似值,它将相应的红、绿、蓝三原色^[21]处理成有意义和可感知的信息,比其他颜色空间更好地保留了亮度分量的详细信息^[22].这种变换可以用于多传感器图像的融合. YC_bC_r 与 RGB 变换过程公式化如下:

$$\begin{pmatrix} Y \\ C_b \\ C_r \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{pmatrix} \cdot \begin{pmatrix} R \\ G \\ B \end{pmatrix}. \quad (1)$$

逆变换方程

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 1.000 & 0.587 & 1.402 \\ 1.000 & -0.344 & -0.714 \\ 1.000 & 1.772 & 0.000 \end{pmatrix} \cdot \begin{pmatrix} Y \\ C_b \\ C_r \end{pmatrix}. \quad (2)$$

经过变换后的图像融合迭代,只需要处理 MRI 图像和 PET 图像中 Y 通道分量的灰度信息.

2.2 网络架构

为了避免图像融合过程中空间信息的丢失,保护 MRI 与 PET 图像所具有的空纹理结构,本文提出了基于生成对抗网络的 ADRGAN 架构.下面详细介绍生成对抗网络以及 ADRGAN 的总览、生成器、鉴别器和区域残差学习模块.

2.2.1 生成对抗网络

生成对抗网络(generative adversarial network, GAN)是古德费罗等人提出的一种创新的生成模式^[23],它可以在不依赖任何先验假设的情况下估计目标分布,并生成与其匹配的数据.基于此特性,GAN 被广泛应用于各种视觉任务,包括图像融合^[24],并取得了良好的性能.GAN 通过同时训练两个模型(生成模型 G 和鉴别器模型 D ^[25])的对抗过程来估计生成模型.

G 生成器捕获分布式数据,并将其发送到生成网络,生成新的样本.新样本可表示为 $x = G(Z)$,其中 Z 代表了从数据空间中采集的噪声,其作为生成器的输入.由于生成网络多样性,样本形成的概率分布 $P_g(X)$ 可以很复杂.对于双输入, G 的训练目标是使 $P_G(x, y)$ 和 $P_{\text{data}}(x, y)$ 尽可能接近,优化公式

$$G^* = \arg \min_G \text{Var}(P_G(x, y), P_{\text{data}}(x, y)), \quad (3)$$

其中, $\text{Var}()$ 代表了两个分布之间的差异,然而由于其他项都是未知的,无法计算差异值.这时,就需要鉴别器 G 来解决这个问题,因为他的样本标签来自输入图像而不是 G, D 的目标函数

$$D^* = \arg \max_D V(G, D), \quad (4)$$

其中,

$$V(G, D) = E_{x, y \sim P_{\text{data}}(x, y)} [\lg D(x, y)] + E_{x, y \sim P_G(x, y)} [\lg(1 - D(x, y))]. \quad (5)$$

目标值越大, Jensen-Shannon 发散度越大,鉴别器越容易区分出真假数据.因此, G 的最优化公式可以进一步转换

$$D^* = \arg \min_G \max_D V(G, D). \quad (6)$$

在训练对抗生成网络时时, G 和 D 的对抗过程就是一个最大-最小博弈. G 生成不断接近于源图像的假数据用来迷惑 D , D 被一直训练从而更好地辨别真假数据.在训练过程中,区分虚假数据和真实数据变得越来越困难.直到鉴别器无法分辨真假数据.此时,形成的生成器是所需的生成器,生成的数据可以视为实际数据.

2.2.2 网络总体架构

ADRGAN 架构可以将 PET 图像的低分辨率 Y 分量(PET_Y)与具有更高空间分辨率的灰度图像 MRI 融合,获得包括丰富的结构信息和更高的空间分辨率的融合图像.为了同时保存高分辨率图像的纹理与细节

信息以及低分辨率图像的结构信息,模型使用调整损失函数的机制来进行预测结果的优化.并且,图像融合任务会出现平滑效果从而损失图像的细节,而细节的缺失在医学图像中是不能接受的.基于以上考虑,本文设计了一个基于联合梯度约束的自适应生成性对抗网络,它的总体架构如图 1 所示.

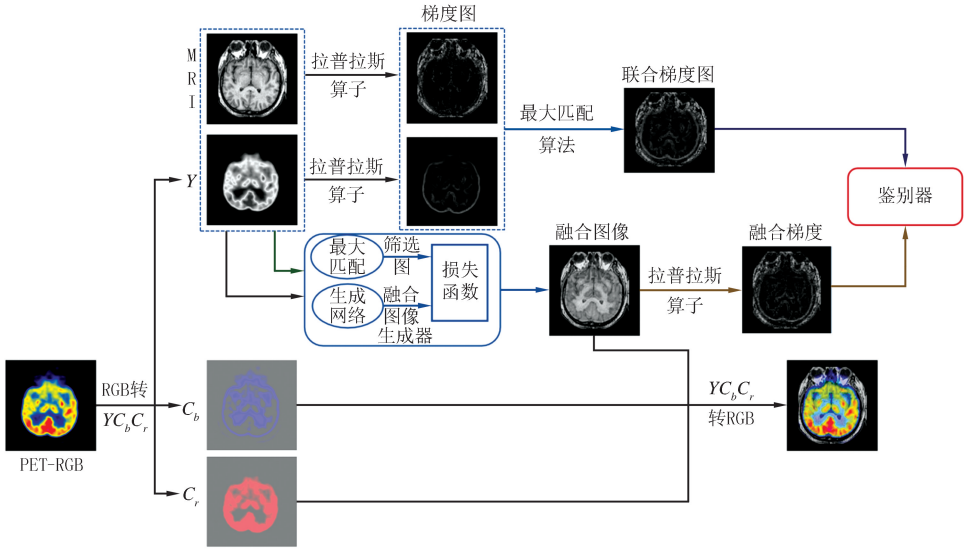


图1 ADRGAN总体架构

Fig.1 ADRGAN general architecture

网络的自适应来源于如图 1 所示的决策块.它可以引导融合结果逼近源图像的亮度与梯度分布.其原理是评估每个像素的清晰度,从而生成一个具有有效信息位置的筛选图.区别于传统深度学习的固定损失函数,模型使用筛选图动态构造损失函数.在优化过程中,只有那些被判决块判定为有效的像素才能参与损失函数计算.

为了在自适应的基础上增强图像的纹理细节,基于梯度图在生成器与鉴别器之间建立对抗性博弈.首先,使用拉普拉斯算子 $\nabla(\cdot)$ 与 maximum 原则计算两幅输入图像的联合梯度图,鉴别器会将联合梯度图定义为真实数据,与被定义为伪数据的融合图像梯度图进行持续性的对抗学习.因此,可以获得含有更多纹理细节的高质量图像.因此,GAN 的目标函数

$$\min_G \max_D V_{GAN}(G, D) = E_{x, y \sim P_{data}(x, y)} [\lg D(\nabla x, \nabla y)] + E_{a, b \sim p_z(a, b)} [\lg(1 - D(\nabla G(a, b)))]. \quad (7)$$

2.2.3 生成器架构

生成器的结构,如图 2 所示.

模型将 MRI 与 PET_Y 分成两条路径分别处理.在基于 CNN 的网络中简单的每条路径都使用了 3 个卷积层与一个 LRLP 来提取特征,其中 LRLP 模块会在 2.2.4 中详细介绍.在准备模块中,第一卷积层用来提取浅特征 F_{pre} .

$$F_{pre} = H_{conv}(I), \quad (8)$$

其中, H_{conv} 表示浅特征提取层中卷积核为 5×5 的卷积运算.随后将提取出的浅特征送入提取模块的 LRLP 层中,LRLP 层的输出可以通过

$$F_1 = H_{LRLP}(F_{pre}), \quad (9)$$

得到,其中 H_{LRLP} 是 LRLP 层运算的复合函数; F_1 是充分利用块内每一层卷积产生的,可以将其视作局部特征.之后进一步卷积实现深层特征提取,为了避免卷积过程中丢失特征,基于残差连接的思想,在随后的卷积层中,输入都是之前所有层以及 RDB 模块的输出级联.同时,为了将多层次的信息充分融合,在两条路径之间也设置了参数共享.第三层输出

$$F_2 = H_{conv}(H_{RL}(\lambda_1 \cdot F_{pre}, \lambda_2 \cdot F_1)), \quad (10)$$

其中, H_{RL} 代表残差连接; λ 为残差融合时的权重.第四层与第三层同理,其输入基于前三层的输出级联.第四层输出

$$F_3 = H_{\text{conv}}(H_{LR}[\lambda_1 F_{\text{pre}}, \lambda_2 F_1, \lambda_3 F_2]), \quad (11)$$

其中, H_{concat} 代表了特征图拼接运算. 提取模块最后一层被设置为特征图拼接后加入 1×1 的卷积, 输出公式

$$F_{\text{ext}} = H_{\text{conv}}(H_{\text{concat}}(W(F_{\text{pre}}, F_1, F_2, F_3))), \quad (12)$$

其中, W 为融合提取模块前四层的权重矩阵. 之后两条路径的输出 $F_{\text{ext},1}, F_{\text{ext},2}$ 进入融合模块, 经过融合运算后, 最终融合图像, 如式(13)所示:

$$I_{\text{fuse}} = H_{\text{fuse}}(F_{\text{ext},1}, F_{\text{ext},2}), \quad (13)$$

其中, H_{fuse} 代表融合模块的复合运算.

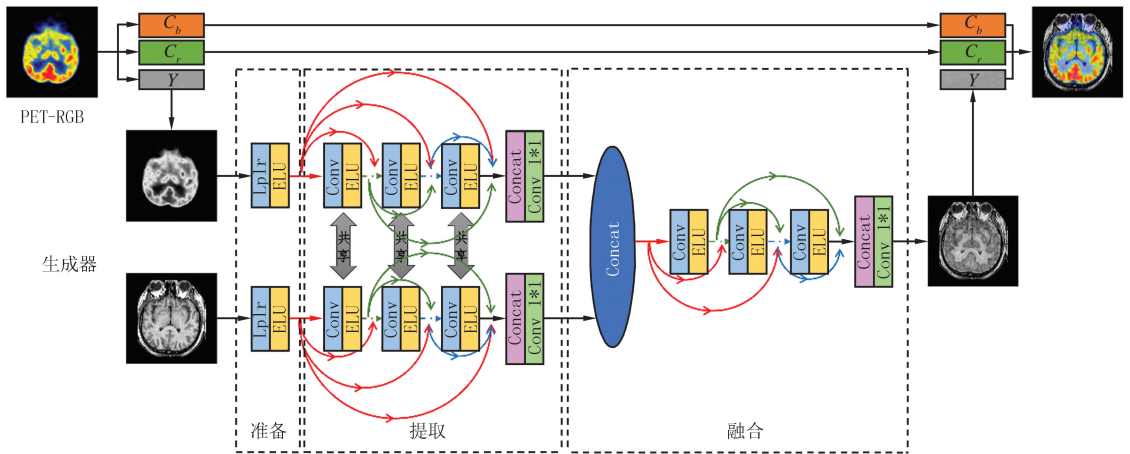


图2 生成器

Fig.2 Generator

2.2.4 鉴别器架构

鉴别器的结构, 如图 3 所示. 其输入有两个来源, 输入的两张图像通过拉普拉斯算子计算梯度图, 随后经过 maximum 函数得到的联合梯度图以及融合图像通过拉普拉斯算子计算得来的梯度图作为鉴别器的两个输入. 4 个卷积层与一个线性层组成了模型的鉴别器. 卷积层的卷积核大小均设置为 3×3 , 步长设置为 4, 使用 relu 作为激活函数. 最后一层是线性层, 用来计算概率, 从而判断生成数据的真假.

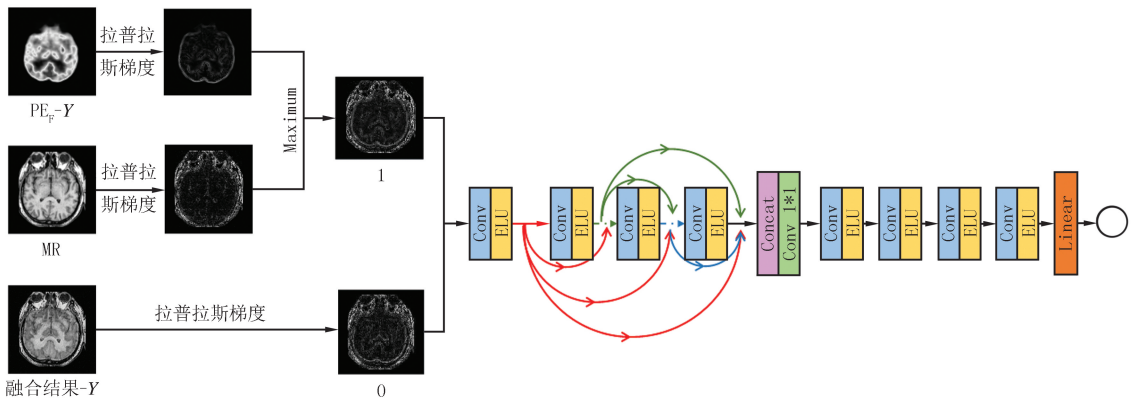


图3 鉴别器

Fig.3 Discriminator

2.2.5 区域残差学习模块(LRLP)

在卷积神经网络的前向传输过程中, 随着网络深度加大, 卷积得到的特征图蕴含信息会逐渐减少, 该结论可以由 DPI(数据处理不等式)证明. 为了解决上述问题, 本文使用了区域残差学习模块, 通过不同层之间信息的直接映射, 尽可能保存每一层蕴含的特征. 其中 LRLP 架构如图 2 所示. 假设 LRLP 有 c 个卷积层, 那

么它的最终输出,如下,

$$F = H_{\text{active}}(W(H_{\text{concat}}[F_1, F_2, \dots, F_c])), \quad (14)$$

其中, H_{active} 表示对拼接后的数据进行 ELU 激活; W 是一组代表了拼接时各卷积层权重的联合函数; H_{concat} 代表了特征图的拼接函数; F_c 为第 c 个卷积层的输出. 在 LRLP 块中, 前面每一层的输出都作为下一层的输入, 既保留了前馈特性, 又提高了对输入数据的利用程度.

2.3 损失函数

为了在 GAN 中实现对抗性学习, 本文提出了一组可区分的半监督损失函数. 其中包括生成器的损失函数与鉴别器的损失函数.

2.3.1 生成器损失

生成器的损失函数基于对抗性损失, 像素级欧几里得损失以及纹理损失, 损失函数

$$\zeta = \zeta_{\text{Gan}} + \lambda_1 \zeta_{\text{pixel}} + \lambda_2 \zeta_{\text{texture}}, \quad (15)$$

其中, ζ_{Gan} 是来自生成器-鉴别器网络的对抗性损失; ζ_{pixel} 是利用筛选图优化的像素级欧几里得损失; ζ_{texture} 表示基于梯度图的纹理损失; λ_1 和 λ_2 分别是像素级损失和纹理损失的权重, 用来保证 3 种损失函数重要性相同.

2.3.2 对抗性损失

为了生成器生成的图像更加接近于理想的融合图像, 需要在生成器与鉴别器之间建立损失. 传统的对抗损失将最大-最小问题简化为 $\lg(1 - D(G(1)))$, 但训练开始阶段, $\lg(1 - D(1))$ 就可能饱和, 所以使用 $D(G(1))$ 最大化来训练生成器网络. 为了提供更强的梯度, 在最大化操作基础上加入了平方操作. ζ_{Gan} 定义如下,

$$\zeta_{\text{Gan}} = \frac{1}{M} \sum_{m=1}^M (D(\nabla G(M_m, Y_m)) - c)^2, \quad (16)$$

其中, M 是训练期间一个 batch 的图像数量; c 是鉴别器鉴定真假图像的概率标签; 本模型取 ∇ 表示使用拉普拉斯算子来进行梯度图的计算; M, Y 表示输入的 MRI 图像与 PET 图像 Y 通道.

2.3.3 像素级欧几里得损失

模型利用融合图像与原图像像素之间的欧几里得距离来约束融合图像与原图像在清晰区域的强度分布.

$$\zeta_{\text{pixel}} = \frac{1}{H \times W} \sum_{h=1}^H \sum_{w=1}^W \text{Map}_1(G(M_{h,w}, Y_{h,w}) - M_{h,w})^2 + \text{Map}_2(G(M_{h,w}, Y_{h,w}) - Y_{h,w})^2, \quad (17)$$

其中, h, w 表示第 h 行第 w 列的像素值; H, W 分别是图像的高度与宽度; $\text{Map}_1, \text{Map}_2$ 代表了判决块基于两个输入图像所产生的筛选图.

2.3.4 纹理损失

图像的梯度可以部分表征纹理细节, 对比尖锐的 MRI 图像更是如此, 因此要求融合图像与输入图像具有相似的梯度. 结合筛选图, 纹理损失可以公式化如下,

$$\zeta_{\text{texture}} = \frac{1}{H \times W} \sum_{h=1}^H \sum_{w=1}^W \text{Map}_1(\nabla G(M_{h,w}, Y_{h,w}) - \nabla M_{h,w})^2 + \text{Map}_2(\nabla G(M_{h,w}, Y_{h,w}) - \nabla Y_{h,w})^2, \quad (18)$$

2.3.5 鉴别器损失

不仅生成器需要损失函数优化融合图像的质量, 鉴别器也需要损失函数准确地识别出融合图像的真假. 本文为鉴别器设计了基于梯度图的损失函数, 其中“假数据”是融合图像的梯度图, 可以公式化为

$$\text{Grad}_{\text{fused}} = \text{abs}(\nabla G(W, Y)). \quad (19)$$

鉴别器需要的“真数据”来自 MRI 与 PET_y 构造的联合梯度图, 公式化为式(20)所示:

$$\text{Grad}_{\text{union}} = \text{maximum}(\text{abs}(\nabla M), \text{abs}(\nabla Y)), \quad (20)$$

其中, abs 表示绝对值函数; maximum 代表了最大化函数.

基于以上两个梯度图, 损失函数

$$\zeta_D = \frac{1}{M} \sum_{m=1}^M (D(\text{Grad}_{\text{fused}}^m) - a)^2 + (D(\text{Grad}_{\text{union}}^m) - b)^2, \quad (21)$$

其中, a 为“假数据”的标签, 设置为 0; b 为“真数据”的标签, 设置为 1. 从而使鉴别器将图像的联合梯度图视为真数据, 融合图像的梯度图视为假数据. 此约束可以引导生成器基于 $\text{Grad}_{\text{union}}$ 来调整 $\text{Grad}_{\text{fused}}$, 在对抗中增强融合图像的纹理.

3 实验结果与分析

在本节, 为了验证 ADRGAN 对于 PET-MRI 图像融合的优越性和鲁棒性, 本文基于公开数据集将其与现有的 5 种方法进行定性与定量比较.

3.1 训练细节

本实验所使用的 PET, MRI 图像来自哈佛大学医学院网站的公开数据集 (<http://www.med.harvard.edu/AANLIB/home.html>). 其中, MRI 图像为大小 256×256 的单通道图像; PET 图像为大小 $256 \times 256 \times 3$ 的伪彩色图像. 需注意的是, 本模型不需要标签数据.

根据对抗性过程迭代地训练生成器和鉴别器. 批处理大小设置为 b , 训练一个迭代需要 k 步骤. 鉴别器训练数与生成器训练数之比为 p , 共训练 M 次. 在实验中, 经过多次试验得到: 设置 $b = 32$, $p = 2$, $M = 300$. ADRGAN 中的参数由 AdamOptimizer 更新. 为使 GAN 的训练更加稳定, 对损失项参数使用软标签: 对于应该设置为 1 的标签, 将其设置为从 0.7 到 1.2 的随机数.

图像经过预处理从 RGB 通道到 $YCbCr$ 颜色空间. 因为 Y 通道(亮度通道)可以表示结构细节和亮度变化, 只需要融合 Y 通道. 使用基于颜色空间方法融合 C_b 和 C_r 通道. 然后将融合分量逆变换为 RGB 通道. 本模型的实验环境为: Windows 10, CPU AMD R5 5600X, 内存 16 G GPU RTX-3060(6 G). 软件环境为 Python 3.7.6 和 Pytorch 1.10.0. 数据集的训练集, 验证集和测试集按照 7 : 2 : 1 划分.

3.2 定量评价指标

本文使用 5 个评价指标客观评价本方法与对比方法, 5 种指标分别为 Q_{abf} , Q_{cv} , PSNR, SSIM 以及 RMSE.

Q_{abf} 算法使用局部度量来估计融合图像中输入的重要信息的性能. Q_{abf} 值越高, 融合图像的质量越好,

$$Q(A, B, F) = \frac{1}{|W|} \sum_{\omega \in W} (\lambda(\omega) Q_0(A, F | \omega) + (1 - \lambda(\omega)) Q_0(B, F | \omega)), \quad (22)$$

其中, W 用来划分局部区域; $\lambda(\omega)$ 代表局部区域权值; A, B, F 分别是两个输入图像与融合图像.

Q_{cv} 通过计算融合后的区域图像和源区域图像的加权差图像的均方误差, 得到局部区域图像的质量. 最后, 融合后的图像质量是局部区域图像质量测度的加权和^[26]. 公式化如下,

$$Q_{cv}(A, B, F) = \frac{\sum_{\omega \in W} (\lambda(A_{\omega}) D(\omega_1, F | \omega) + \lambda(B_{\omega}) D(B_{\omega}, F | \omega))}{\sum_{\omega \in W} (\lambda(A_{\omega}) + \lambda(B_{\omega}))}, \quad (23)$$

其中, D 为局部区域相似性度量函数.

波峰信噪比(PSNR)是融合图像中波峰功率与噪声功率的比值, 反映了融合图像的失真情况. 计算方法如式(24)~(27)所示:

$$MSE_{AF} = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (A(i, j) - F(i, j))^2, \quad (24)$$

$$MSE_{BF} = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (B(i, j) - F(i, j))^2, \quad (25)$$

$$MSE = \frac{MSE_{AF} + MSE_{BF}}{2}, \quad (26)$$

$$PSNR = 10 \ln \left(\frac{r^2}{MSE} \right), \quad (27)$$

其中, MSE 为均方误差, 表示图像中 i 行 j 列的像素; r 表示融合图像的峰值, 峰值信噪比越大, 融合图像越接近源图像。

结构相似度(SSIM)用来模拟图像的丢失和失真. 该指数由 3 部分组成, 分别是相关损失、对比损失和亮度损失. 这 3 个分量的乘积是融合图像的评估结果, 定义如下:

$$SSIM_{x,f} = \sum_{x,f} \frac{2u_x u_f + C_1}{u_x^2 + u_f^2 + C_1} \cdot \frac{2\sigma_x \sigma_f + C_2}{\sigma_x^2 + \sigma_f^2 + C_2} \cdot \frac{\sigma_{x_f} + C_3}{\sigma_x \sigma_f + C_3}, \quad (28)$$

其中, x 和 f 分别表示源图像与融合图像中的一块; σ_{x_f} 为两个块之间的协方差; σ_x, σ_f 表示标准差(SD); u_x, u_y 表示两个块之间的平均值. 另外加入 C_1, C_2, C_3 使损失函数更稳定.

均方根误差(RMSE)基于 MSE , 通过计算源图像和结果图像的均方误差, 完成对源图像与融合图像之间差异的定量化描述. 计算如下,

$$RMSE = \frac{1}{2} (\sqrt{MSE_{AF}} + \sqrt{MSE_{BF}}). \quad (29)$$

3.3 定量与定性对比结果

为了验证本模型对于 PET-MRI 图像融合的效果并验证其鲁棒性. 实验选取了 DDcGan, Densefuse, GCF, IFCNN, PMGI 等 5 种方法与本文方法比较. 以上方法均在以往的医学图像融合中取得了很好的效果.

6 种相关方法的视觉实验效果, 如图 S1(见附录)所示. 其中, DDcGan 的结果(第 3 列)存在颜色失真的问题, 且边缘较本模型模糊; Densefuse(第 4 列)损失了 PET 图像中颜色的强度, 损失部分功能信息, 加大寻找病灶的难度; GCF(第 5 列)颜色保存较好, 但多幅图像出现大面积噪声块, 结构信息直接丢失, 其信息丢失会误导临床判断, 鲁棒性差. IFCNN(第 6 列)在边界线附近会模糊损失细节, 纹理密集处不够清晰; PMGI(第 7 列)在融合时, 虽然颜色强度高, 功能信息保存完整, 但背景虚化, 高频信息丢失, 无纹理细节. 本文方法的融合图像无以上问题, 结构与功能信息保存完好, 细节清晰且对比度高. 尤其边缘处对比明显, 纹理密集处细节清晰, 包含足以满足临床诊断所需图像信息. 由于这些方法大多尝试通过直接的目标增强和梯度来锐化边缘, 因此融合后的图像的自然性和真实度差异仍然很大. 此外, 几乎所有这些方法都依赖于大型数据集. 而本文提出的结构性内容损失函数以及对抗性损失函数分别来保护高频信息和内容信息, 通过分别的非线性损失约束提升融合图像的效果. 仅仅用人眼的主观感觉定性评估融合效果具有很大局限性, 为了客观验证本文方法的优越性, 本文选择客观评价方式对本模型实验结果进行定量评估, 其结果如表 1 所示.

表 1 定量化对比表

Tab. 1 Quantitative comparison table

实验方法	PSNR	SSIM	RMSE	Q_{abf}	Q_{cv}
DDcGan	54.816 2	0.300 0	0.214 6	0.160 2	2 534.607
DenseFuse	55.183 0	0.362 8	0.198 6	0.136 8	2 242.367
GCF	54.416 3	0.334 7	0.236 7	0.340 1	2 521.672
IFCNN	54.416 3	0.416 0	0.208 3	0.351 6	2 226.219
PMGI	54.015 1	0.102 2	0.258 1	0.046 0	3 469.525
本文	55.212 4	0.469 7	0.196 8	0.363 5	2 009.348

可以看出, 本文方法在 5 个指标上均分别优于其余 5 个对比方法. Q_{cv} 指标基于人体视觉系统(HVS)与区域均方误差, 得益于自适应模块, 模型可以自适应地判断像素权重, 从而提高区域相似度. 本方法相比于 DDcGan 降低了 20.7%, 证明了在人眼感知意义上强于其余方法且区域相似度更高; 对抗性博弈使模型拥有优秀的去噪能力, 从而使 PSNR 提升, 证明相比于其他方法, 本模型噪声更少, 干扰信息少; SSIM 指标倾向于对结构信息的验证, 与最优异的 IFCNN 相比增加了 11.4%, 该指标更高说明本文方法纹理结构保存完整, 模糊区域少, 其中 PMGI 结构相似性仅为本方法的 21.0%, 结构保存差, 与定性对比结果; 模型采用低像素尺度控制策略, 像素之间欧几里得距离控制良好, 像素级融合指标 $Q(A, B, F)$ 更高, 视觉信息完善, 融合图像与源图像像素层面差距小; 同时更小的 RMSE 表明本模型融合后的图像具有更少的误差和失真.

4 结 论

本文通过构造自适应的残差密集生成对抗网络(ADRGAN),结合基于 YC_bC_r 的颜色空间方法,提出了一种新的基于联合梯度图像融合方法。(1)采用区域残差学习模块与输出级联加深生成网络;(2)设计自适应判决块动态引导生成器产生与源图像分布相同的融合图像;(3)在融合图像梯度图与输入图像联合梯度图之间进行对抗性博弈,从而得到细节丰富且纹理清晰的融合图像.该算法不需要真实数据作为标签进行训练,可以在不引入传统框架的情况下融合不同分辨率的图像.此外,为了进一步增强融合结果的边缘锐度,引入了自适应的对抗性损失函数.经过测试,本文提出的方法在哈佛医学院 MRI/PET 数据集的测试中达到 $PSNR=55.2124$, $SSIM=0.4697$, $RMSE=0.1968$, $Q_{abf}=0.3635$ 和 $Q_{cv}=2009.348$,均优于目前最先进的算法,且图像具有更多的细节及内容信息更有助于临床应用诊断.在此基础上,未来将针对使用注意力机制更好的过滤和融合高维特征进行更深入研究,致力于提供更高质量的融合模型.

附录见电子版(DOI:10.16366/j.cnki.1000-2367.2022.08.11.0003).

参 考 文 献

- [1] HERMESSI H, MOURALI O, ZAGROUBA E. Multimodal medical image fusion review: theoretical background and recent advances[J]. *Signal Processing*, 2021, 183: 108036.
- [2] DANESHVAR S, GHASSEMIAN H. MRI and PET image fusion by combining IHS and retina-inspired models[J]. *Information Fusion*, 2010, 11(2): 114-123.
- [3] 刘尚旺, 赵欣莹, 杨磊. 基于全局和局部信息融合的显著性检测[J]. 河南师范大学学报(自然科学版), 2020, 48(3): 26-33.
LIU S W, ZHAO X Y, YANG L. Saliency detection via fused global and local information[J]. *Journal of Henan Normal University(Natural Science Edition)*, 2020, 48(3): 26-33.
- [4] 陈向东, 邓江洪, 安国成. 融合算法在高空间分辨率影像中的应用研究[J]. 河南师范大学学报(自然科学版), 2019, 47(4): 31-36.
CHEN X D, DENG J H, AN G C. Application of fusion algorithm in high spatial resolution images[J]. *Journal of Henan Normal University(Natural Science Edition)*, 2019, 47(4): 31-36.
- [5] YANG D S, HU S H, LIU S Q, et al. Multi-focus image fusion based on block matching in 3D transform domain[J]. *Journal of Systems Engineering and Electronics*, 2018, 29(2): 415-428.
- [6] HE C T, LIU Q X, LI H L, et al. Multimodal medical image fusion based on IHS and PCA[J]. *Procedia Engineering*, 2010, 7: 280-285.
- [7] LIU Y, ZHOU D, NIE R, et al. Construction of high dynamic range image based on gradient information transformation[J]. *IET Image Processing*, 2020, 14(7): 1327-1338.
- [8] LIU Y, CHEN X, WANG Z F, et al. Deep learning for pixel-level image fusion: recent advances and future prospects[J]. *Information Fusion*, 2018, 42: 158-173.
- [9] LIU Y, CHEN X, WARD R K, et al. Image fusion with convolutional sparse representation[J]. *IEEE Signal Processing Letters*, 2016, 23(12): 1882-1886.
- [10] 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.
- [11] SABEEH Y A, OMAR Z, ULLAH S U. An improved approach for medical image fusion using sparse representation and Siamese convolutional neural network[J]. *Biomedical Signal Processing and Control*, 2022, 72: 103357.
- [12] HOU R C, ZHOU D M, NIE R C, et al. Brain CT and MRI medical image fusion using convolutional neural networks and a dual-channel spiking cortical model[J]. *Medical & Biological Engineering & Computing*, 2019, 57(4): 887-900.
- [13] LI H, WU X J. DenseFuse: a fusion approach to infrared and visible images[J]. *IEEE Transactions on Image Processing: a Publication of the IEEE Signal Processing Society*, 2018, 28(5): 2614-2623.
- [14] CHEN M J, ZHENG H, LU C S, et al. A spatio-temporal fully convolutional network for breast lesion segmentation in DCE-MRI[C]//International Conference on Neural Information Processing. Cham: Springer, 2018: 358-368.
- [15] 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.
- [16] 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: a Publication of the IEEE Signal Processing Society*, 2020, 29(3): 4980-4995.
- [17] TANG W, LIU Y, ZHANG C, et al. Green fluorescent protein and phase-contrast image fusion via generative adversarial networks[J].

- Computational and Mathematical Methods in Medicine, 2019, 2019: 5450373.
- [18] ZHANG H, XU H, XIAO Y, et al. Rethinking the image fusion: a fast unified image fusion network based on proportional maintenance of gradient and intensity[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12797-12804.
- [19] XIA K J, YIN H S, WANG J Q. A novel improved deep convolutional neural network model for medical image fusion[J]. Cluster Computing, 2019, 22(1): 1515-1527.
- [20] HOU R C, ZHOU D M, NIE R C, et al. Brain CT and MRI medical image fusion using convolutional neural networks and a dual-channel spiking cortical model[J]. Medical & Biological Engineering & Computing, 2019, 57(4): 887-900.
- [21] ZHANG H, LE Z L, SHAO Z F, et al. MFF-GAN: an unsupervised generative adversarial network with adaptive and gradient joint constraints for multi-focus image fusion[J]. Information Fusion, 2021, 66: 40-53.
- [22] IRAJI M S, YAVARI A. Skin color segmentation in fuzzy YCBCR color space with the mamdani inference[J]. American Journal of Scientific Research, 2011(7): 131-137.
- [23] KANG B, JEON C, HAN D K, et al. Adaptive height-modified histogram equalization and chroma correction in YCbCr color space for fast backlight image compensation[J]. Image and Vision Computing, 2011, 29(8): 557-568.
- [24] MASI G, COZZOLINO D, VERDOLIVA L, et al. Pansharpening by convolutional neural networks[J]. Remote Sensing, 2016, 8(7): 594.
- [25] 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. New York: ACM, 2014: 2672-2680.
- [26] CHEN H, VARSHNEY P K. A human perception inspired quality metric for image fusion based on regional information[J]. Information Fusion, 2007, 8(2): 193-207.

Fusion of PET and MRI images using adaptive dense residual generative adversarial network

Liu Shangwang, Yang Lihan

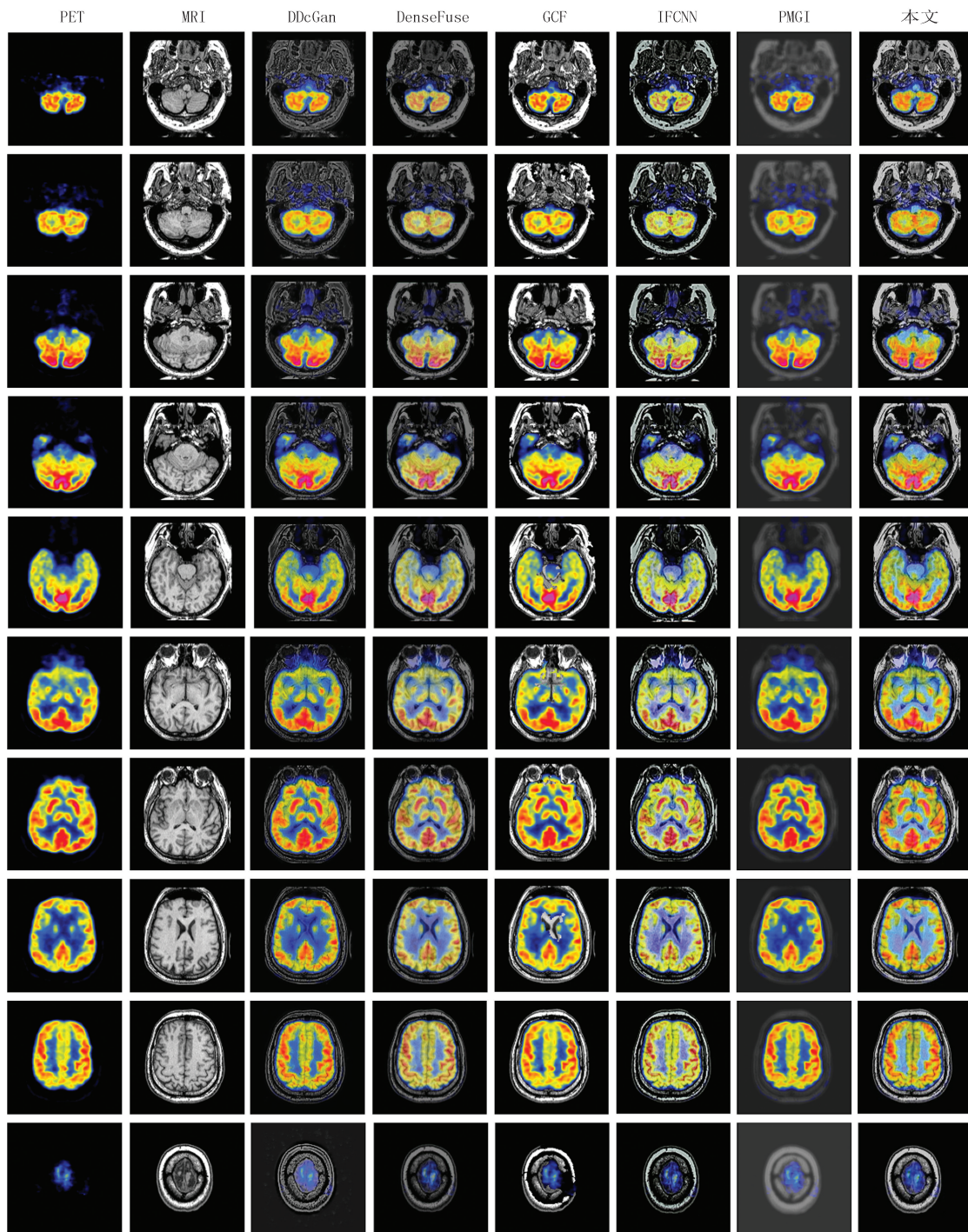
(College of Computer and Information Engineering; Engineering Lab of Intelligence Business and Internet of Things,
Henan Normal University, Xinxiang 453007, China)

Abstract: In this paper, in order to enhance the texture details of fusion of magnetic resonance imaging(MRI) and positron emission computed tomography imaging(PET) images and get rid of the dependence of artificial design fusion rules on a priori knowledge. An adaptive dense residual generative adversarial network(ADRGAN) is proposed to fuse two modes of medical images. ADRGAN designs a regional residual learning module and output cascade generator to deepen the network structure and avoid feature loss. Then, a content loss function based on adaptive module is designed to enhance the content information of the output fused image. Meanwhile, an adversarial game is constructed through the joint gradient map of the source image and the gradient map of the fused image, and an efficient training generator and discriminator are constructed. The experimental results show that the peak signal-to-noise ratio and structural similarity of ADRGAN proposed in this paper are 55.212 4 and 0.469 7 respectively, which are better than the most advanced algorithms. Moreover, the model in this paper has two characteristics: end-to-end and unsupervised, without manual intervention and real data as labels.

Keywords: deep learning; generative adversarial network; multi-model image fusion; dense residual network

[责任编辑 陈留院 赵晓华]

附 录



图S1 融合效果对比

Fig.S1 Fusion effect comparison