

基于 RCM-cycleGAN 的跨模态眼底图像重建

王聪茹¹ 史铭宇² 李婷雪¹ 李 妍¹ 张艳珠¹

WANG Congru SHI Mingyu LI Tingxue LI Yan ZHANG Yanzhu

摘 要

眼部多普勒彩超和眼底 OCT 图像常用于眼部临床医学诊断。由于图像包含的信息和设备等因素的影响,获得眼底 OCT 的途径会受到限制。针对这一限制,提出一种针对 cycleGAN 网络改进的多尺度残差生成对抗网络(RCM-cycleGAN)模型,使用眼部多普勒彩超图片重建眼底 OCT。通过采集眼部多普勒彩超和眼底 OCT 制作相应的数据集,首先在 cycleGAN 网络的生成器中加入残差网络结构,降低 GAN 网络生成器梯度爆炸;其次在生成器结构中引入通道注意力,提升网络在眼底照相视盘和血管特征的关注;最后在网络中引入多尺度特征提取,进一步优化了重建效果。使用改进后的 RCM-cycleGAN 算法与 cycleGAN、RegGAN、DC-cycleGAN、DualGAN 等对眼部彩超和眼底 OCT 之间的重建进行训练,根据重建图像与测试集数据对比的效果表明所设计的 RCM-cycleGAN 可以更好地实现眼部彩超对眼底照相的重建。

关键词

跨模态;生成对抗网络;医学图像重建;RCM-cycleGAN

doi: 10.3969/j.issn.1672-9528.2025.03.030

0 引言

彩色多普勒超声和眼底照相(眼底 OCT)在眼科检查中非常重要,广泛用于眼部等各种疾病的诊断以及放疗治疗计划。一方面,彩色多普勒超声图像在捕捉对比度的多样性和准确定位病灶和血流方面非常出色软组织的结构图像^[1]。另一方面,眼底 OCT 图像能够显示病变的细节并提供电子密度信息^[2]。由于可以从每种模式中获得独特的信息,因此使用患者的彩色多普勒超声和眼底 OCT 图像可以显著改善临床诊断和治疗。这些局限性,包括成本、时间、辐射剂量以及准确彩色多普勒超声/眼底 OCT 配准的必要性,可能会限制两种图像的获取。可以通过从另一种方式合成医学图像来避免,即从彩色多普勒超声图像生成眼底图像,是一种跨模态图像重建^[3]。

跨模态医学图像重建可以定义为目标图像和源图像之间的映射。现有方法分为特征映射和深度学习(DL)方法。特征映射方法通常需要许多预处理/后处理步骤以及调整转换参数的先验知识。

基于深度学习的方法直接学习从源域到目标域的复杂

非线性映射^[4-5]。卷积神经网络(CNNs)^[6]是最受欢迎的基于深度学习的医学图像合成方法之一,因其具有学习图像结构的能力。Nie 等人^[7]采用三维全卷积神经网络(FCN)从 MR 图像合成 CT 图像。Nie 等人^[8]使用 U-Net 结构从 MR 图像合成 CT 图像。此外,基于生成的方法,包括生成对抗网络(GAN)^[9]和变分自动编码器(VAE)^[10],在合成一种模态的医学图像方面有显著效果,以另一种模态为条件。Li 等人^[11]提出一个通用的图像转换框架 pixel2style2pixel(pSp),直接生成一系列风格向量,然后将它们输入到预训练的 StyleGAN 生成器中扩展空间。

上述方法依赖于大量的配对图像,此类图像很难获得,为了克服此类问题,Zhu 等人^[12]提出了 cycleGAN 以无监督的方式从未配对的数据中合成图像。尽管 cycleGAN 在重建与真实输入相同的图像(例如 CT-to-CT)方面显示出显著的结果,但它不能直接用于合成来自另一种模式(例如 MR-to-CT)的样本,因为真实源图像和合成图像之间没有直接限制^[13]。Santini 等人^[14]通过额外的损失来强制生成的图像与真实图像相同,从而缓解了这个问题。

对 cycleGAN 进行了改进,构建了一种 RCM-cycleGAN 学习模型。首先在 cycleGAN 网络的生成器中加入残差网络,构建了一个残差网络生成器,降低 GAN 网络生成器梯度爆炸;其次在生成器结构中引入通道注意力,提升网络训练对眼底 OCT 视盘和血管特征的关注;最后在网络中引入多

1. 沈阳理工大学自动化与电气工程学院 辽宁沈阳 110159

2. 中国医科大学附属第四医院 辽宁沈阳 110005

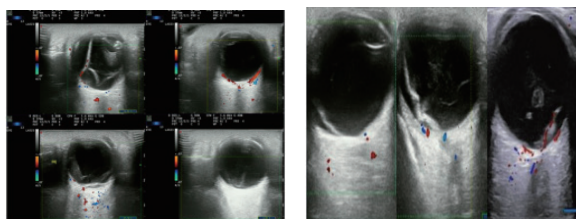
[基金项目]全自动智能眼科超声机器人关键技术(LJ212410144052)

尺度特征提取,进一步优化了重建效果更进一步地提取眼底 OCT 的特征。

1 相关工作

在图像合成的研究领域中,数据、算法与算力三者缺一不可。在中国医科大学第四附属医院的支持下,获得了用于眼底 OCT 生成研究的数据集。这一数据集的构建严格遵循了研究要求,确保了数据的准确性和可靠性。

在数据处理阶段,首先筛选出彩色超声图像中能够映射到眼底 OCT 的有效部分,如图 1 所示,确保后续分析的精确性,统一将图像中包含眼部信息的部分裁剪,确保所有图像在尺度上的一致性。面对给定的矩形图像,采用简单的缩放方法,确保图像包含完整信息,同时避免了复杂的预处理步骤,以最大程度地保留图像的原始信息。



(a) 原图 (b) 预处理

图 1 眼部彩色多普勒超声图像

2 方法改进

对抗式神经网络 (generative adversarial network, GAN) 的是两个网络的组合,可以理解为一个网络生成模拟数据 (生成网络 Generator), 另一个网络判断生成的数据是真实的还是模拟的 (判别网络 Discriminator)。如图 2 所示。

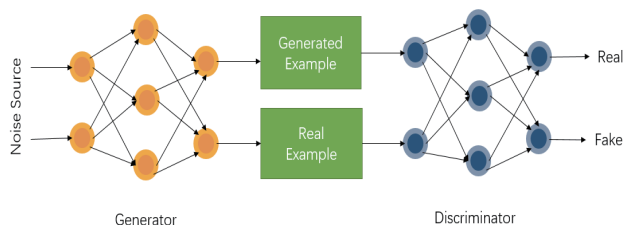


图 2 GAN 的基本网络结构

2.1 Resnet50 生成器

残差网络 ResNet50 是一种包含 50 层的卷积神经网络 (CNN), 其核心组成部分是所谓的“残差块”或“shortcut 连接”。每个残差块通常包含两个或三个卷积层, 并且有一个跳跃连接 (shortcut) 绕过这些卷积层直接将输入与输出相加。这种结构使得信息和梯度可以更直接地向前和向后传播, 从而有助于训练更深的网络。其架构通过引入残差块解决了深层神经网络中的梯度消失问题, 使得网络能够训练得更深, 从而在许多计算机视觉任务上取得了成功。

将 ResNet50 用作生成器并不是一个常见的选择, ResNet50 生成器的设计初衷是为了特征提取和分类, 其架构也可以被调整来用作生成器的一部分, 如图 3 所示。在这样的场景下, ResNet50 的残差块可以用于保持细节和解决训练过程中的梯度问题。

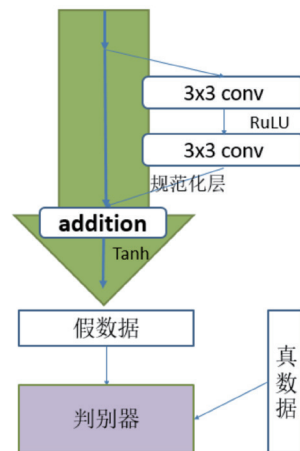


图 3 改进后的 ResNet50 生成器结构

2.2 通道注意力

为利用不同通道间的特征依赖性, 通道注意力已广泛应用于计算机视觉领域。在图像处理领域中, 更侧重于提取序列特征, 如目标图像特征, 不同通道更细节的信息等。相较于空间注意力, 通道注意力涵盖更多不同特征的信息。使用通道注意力 (channel attention, CA) 使得模型在忽略眼底外区域的同时, 专注于眼底图像的视盘特征, 从而提高重建细节。在生成器中, 通道注意力用于实现捕捉不同通道特征的上下文的依赖, 有助于更好地重构提取特征。

在 RCM 模型中, 引入了 CAM 模块。通过将算力集中在对生成结果更细节的特征上, 增加数据的类间方差, 从而提高分类精度, 对数据的通道特征赋予相对应的权重。ResNet 网络默认特征图的每个通道同等重要, 而在本文中眼底 OCT 中的轮廓、病灶、视盘和血管特征所占重要性并不相同, 因此, 研究通过对生成器模型加入通道注意力机制模块重新校准不同特征所对应的权重。改进后的网络更加能关注眼底 OCT 视盘和血管的特征。

SE 模块主要包括 3 部分, 首先, 对输入的特征图进行挤压 (Squeeze) 操作, 得到一个通道的全局特征 Squeeze 操作通过使用全局平均池化得到 $1 \times 1 \times C$ 的特征图, 图 4 中 X 表示输入注意力机制之前的特征图; F 表示卷积算子; U_c 表示特征图; F_{sq} 表示 Squeeze 压缩变换函数; F_{ex} 表示 Excitation 操作; H 和 W 表示每个特征图的高和宽将每一个通道上的特征重新编码成为一个全局特征, 代替一般的卷积操作, 以获得对彩超图像信息更大的感受视野。计算公式为:

$$Y_e = F_{sq} = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W U_e(i, j) \tag{1}$$

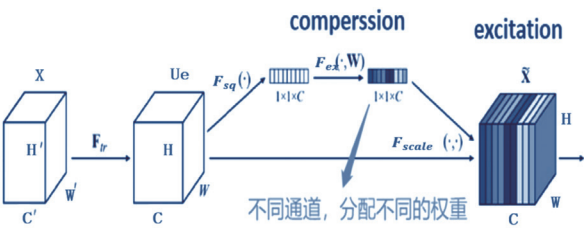


图 4 通道注意力

2.3 多尺度特征提取生成器

每个卷积层只能处理一个特殊的尺度，尺度变化同样是深度学习任务中的挑战之一。U-Net 在复杂医疗场景下的成功运用，本文设计了多尺度特征感知模块作为 GAN 的判别器，来实现捕获和提取多尺度信息来表征对象尺度的变化。

为了解决在复杂场景中面临的诸多问题，提取到更好的眼底特征信息，在引入 cycleGAN 的基础上引入多尺度卷积，改进网络得到一种新的生成器，以提供更加丰富的特征信息。

对 cycleGAN 网络改进的重点是融合来自不同的特征级别的信息，而不是简单的特征映射。这种改进使得该网络结构可以在更好地融合特征信息的基础上适应眼底 OCT 的重建，如图 5 所示。本文在多尺度特征提取中运用了 3×3 卷积。

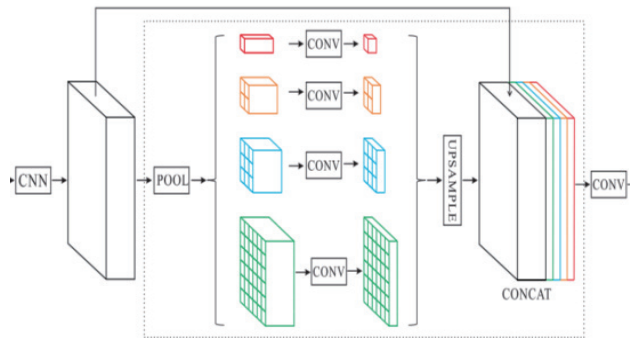


图 5 多尺度特征提取

改进后的结构如图 6 所示。

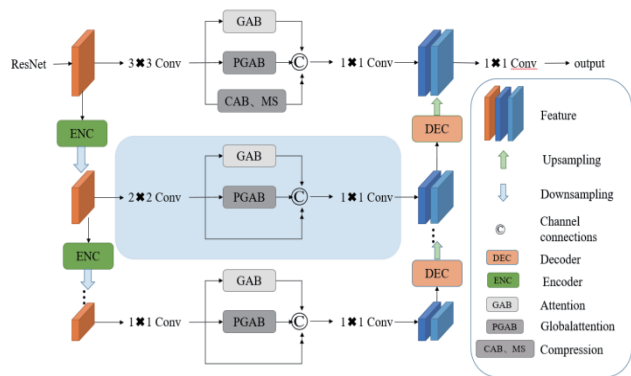


图 6 改进生成器内部结构

3 实验设置

3.1 对比试验

研究选取了 4 种对比算法，这些方法包括传统的 DC-cycleGAN^[15]、RegGAN^[16]、DualGAN^[17] 和 cycleGAN。其发布的代码用于使用相应参考中的相同参数生成结果。通过表 1 可以看出，生成眼底 OCT 图像的各种方法的定量准确度。RCM-cycleGAN 在 MAE、PNSR 和 SSIM 方面优于其他方法。这主要是由于集成到 RCM-cycleGAN 结构中的对比度损失。

表 1 跨模态重建后指标结果

	PSNR	SSIM	MAE
RegGAN	25.67	0.681 3	34.42
cycleGAN	24.28	0.510 3	35.47
DualGAN	25.77	0.734 7	33.25
DC-cycleGAN	27.15	0.878 0	32.74
本文方法	28.23	0.928 0	33.36

图 7 给出了合成的眼底 OCT 图像沿着以及通过不同方法的真实的图像和合成图像之间的误差。可以看出，与其他方法相比，RCM-cycleGAN 合成的图像更接近于真实的图像。这表明 SSIM 和 CE 沿着对比度在合成图像中的有效性。此外，与其他方法相比，RCM-cycleGAN 合成的眼底 OCT 图像与欧堡实况之间的误差相对较小。

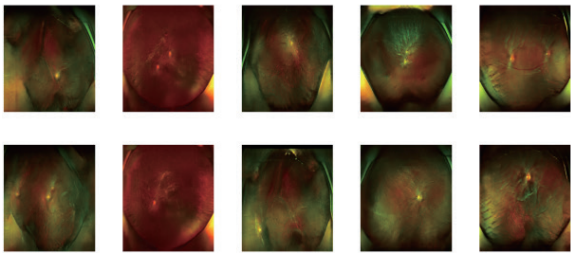


图 7 不同方法重建的图像对比

3.2 消融实验

为了验证 ResNet、CA、MS 的有效性，进行了消融实验。其中，RCM 为本文改进网络，采用了 3 个模块，M1、M2、M3 分别为不添加 ResNet、CA、MS 的情况。为保证消融实验的有效性，M1~M4 使用一致的数据集和训练参数。

表 2 消融实验结果

	PSNR	SSIM	MAE
M1	26.28	0.510 3	35.80
M2	25.77	0.734 7	34.58
M3	26.95	0.878 0	31.42
本文方法	28.23	0.928 0	33.36

由表 2 可以看出：与 RCM-GAN 相比，由于无法将图像特征充分重建，M1 的 PSNR 下降了 1.95 dB；由于无法充分

融合深度和强度特征,进而无法引导恢复重建,M2的PSNR下降了2.46 dB;由于ResNet和CA被移除,M3的PSNR下降了1.28 dB。因此,每一个模块都有助于提高网络性能,RCM-cycleGAN重建效果具有更为明显的提升作用。

4 结论

本文提出的RCM-cycleGAN模型可用于无监督医学图像合成,将学习样本映射到远离源域的图像。RCM-cycleGAN结构集成了SSIM和CE损失中,增加对样本亮度和结构的关注。改进后的RCM-cycleGAN模型,首先将残差网络引入生成器结构中,降低GAN网络生成器梯度爆炸。其次引入通道注意力,对生成器结构中的不同通道特征进行不同的权重提取。最后在生成器结构中加入多尺度提取的改进,更大程度的生成眼底欧堡的特征。对眼部彩超重建眼底图像的结果进行了评估,RCM-cycleGAN与cycleGAN、RegGAN和DualGAN等先进的方法相比,结果表明RCM-cycleGAN能够在准确性和视觉质量方面产生更好的眼底图像。

参考文献:

- [1] 吴德正,马红婕,马静琳.200°超广角眼底像图谱[M].北京:人民卫生出版社,2017.
- [2] 杨文利.临床超声诊断学[M].北京:科学技术文献出版社,2019.
- [3] ISOLA P, ZHU J Y, ZHOU T H, et al. Image-to-image translation with conditional adversarial networks[C/OL]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).Piscataway:IEEE,2017[2024-04-05].<https://ieeexplore.ieee.org/document/8100115>.
- [4] JEONG S, KIM Y J, LEE E, et al. Memory-guided unsupervised image-to-image translation[DB/OL].(2021-04-12)[2024-06-11].<https://doi.org/10.48550/arXiv.2104.05170>.
- [5] ZHOU X L, LIU H, POURPANAH F, et al. A survey on epistemic (model) uncertainty in supervised learning: recent advances and applications[J]. Neurocomputing, 2022, 489: 449-465.
- [6] MINAR M R, NAHER J. Recent advances in deep learning[DB/OL].(2018-07-21)[2024-07-05].<https://doi.org/10.48550/arXiv.1807.08169>.
- [7] DONG N, CAO X H, GAO Y Z, et al. Estimating CT image from MRI data using 3D fully convolutional networks[J]. Deep learning and data labeling for medical applications, 2016, 10008: 170-178.
- [8] NIE D, TRULLO R, LIAN J, et al. Medical image synthesis with deep convolutional adversarial networks[J]. IEEE transactions on biomedical engineering, 2018, 65(12): 65, 2720-2730.
- [9] GOODFELLOW L J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]//NIPS'14: Proceedings of the 28th International Conference on Neural Information Processing Systems. New York: ACM, 2014: 2672-2680.
- [10] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[DB/OL]. (2015-05-18)[2024-03-12].<https://doi.org/10.48550/arXiv.1505.04597>.
- [11] LI X Y, ZHANG S C, HU J, et al. Image-to-image translation via hierarchical style disentanglement[C/OL]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2021[2024-05-29].<https://ieeexplore.ieee.org/document/9577868>.
- [12] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C/OL]//2017 IEEE International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2017[2024-07-06].<https://ieeexplore.ieee.org/document/8237506>.
- [13] YANG H R, SUN J, CARASS A, et al. Unsupervised mr-to-ct synthesis using structure-constrained cycleGAN[J]. IEEE transactions on medical imaging, 2020, 39(12): 4249-4261.
- [14] SANTINI G, FOURCADE C, MOREAU N, et al. Unpaired pet/ct image synthesis of liver region using cycleGAN[J/OL]. Computer science, 2020[2024-05-19].<https://www.semanticscholar.org/paper/Unpaired-PET-CT-image-synthesis-of-liver-region-Santini-Fourcade/0e6218bd9b373cf954a803c68cc9e2a4ed7dfbaf>.
- [15] WANG J Y, JONATHAN WU Q M, POURPANAH F. Dc-cycleGAN: bidirectional CT-to-MR synthesis from unpaired data[J]. Computerized medical imaging and graphics, 2023, 108: 102249.
- [16] KONG L K, LIAN C Y, HUANG D T, et al. Breaking the dilemma of medical image-to-image translation[DB/OL]. (2021-11-11)[2024-04-19].<https://doi.org/10.48550/arXiv.2110.06465>.
- [17] YI Z L, ZHANG H, TAN P, et al. DualGAN: unsupervised dual learning for image-to-image translation[DB/OL]. (2018-10-09)[2024-06-12].<https://doi.org/10.48550/arXiv.1704.02510>.

【作者简介】

王聪茹(2000—),女,河南安阳人,硕士研究生,研究方向:模式识别与智能系统。

(收稿日期:2024-11-07)