

# 基于 GAN 的医学图像仿真数据集生成算法

孟 琰<sup>1</sup>, 钟健平<sup>1</sup>, 李 楠<sup>2</sup>  
(1. 东北大学 信息科学与工程学院, 辽宁 沈阳 110819; 2. 沈阳产品质量监督检验院, 辽宁 沈阳 110000)

**摘 要:** 基于生成对抗网络(generative adversarial networks, GAN), 提出了面向肝脏肿瘤 CT 图像仿真数据集生成深度学习算法. 首先, 将 CT 图像数据文件进行格式解析, 单独保存为 PNG 格式的图像文件; 然后, 将肝脏病变区域统一标注为白色, 并结合肝脏 CT 原图组成配对图片; 最后, 用生成对抗网络的 pix2pix 架构仿真生成病变肝脏图像. 为将生成图像与目标图像进行定量分析、比较, 本文采用了峰值信噪比和结构相似性作为模型的评价指标. 实验结果表明, 本文算法所生成的肝脏肿瘤 CT 仿真数据集的平均峰值信噪比为 64.72 dB, 平均结构相似性为 0.997 3, 证明了所生成的仿真图像数据有着非常高的真实度.

**关 键 词:** 生成对抗网络; 图像处理; 肝脏图像仿真; 参数调整; 数据增强

**中图分类号:** TP 391      **文献标志码:** A      **文章编号:** 1005-3026(2020)03-0332-05

## Generating Algorithm of Medical Image Simulation Data Sets Based on GAN

MENG Lu<sup>1</sup>, ZHONG Jian-ping<sup>1</sup>, LI Nan<sup>2</sup>  
(1. School of Information Science & Engineering, Northeastern University, Shenyang 110819, China;  
2. Shenyang Product Quality Supervision and Inspection Institute, Shenyang 110000, China. Corresponding author: MENG Lu, E-mail: menglu@ise.neu.edu.cn)

**Abstract:** Based on generative adversarial networks (GAN), a deep learning algorithm for generating diseased liver CT image data sets was proposed. Firstly, the CT image data file was formatted and saved as an image file in PNG format. Then the liver lesion area was uniformly marked as white, and the liver CT original image was combined to form a paired picture. Finally, diseased liver image was generated using a pix2pix architecture that created an anti-network. In order to quantitatively analyze and compare the generated image with the target image, the peak signal-to-noise ratio and structural similarity were used to evaluate the model. The results showed that the average peak signal-to-noise ratio of the simulated CT diseased liver image generated by the proposed algorithm is 64.72 dB and the average structural similarity is 0.997 3, thus proving these simulated image data have very high trueness.

**Key words:** generative adversarial networks (GAN); image processing; liver image simulation; parameter adjustment; data augmentation

对于深度学习来说,数据集的规模和数目将很大程度地影响最终训练效果的好坏<sup>[1]</sup>.但是在计算机辅助诊断领域中,由于涉及到患者隐私,研究者很难获取到大量的医学图像数据集.此外,医学图像的采集方式(如 CT, MRI 等)往往会伴随着有害人体健康的高辐射,同时成本极高<sup>[2]</sup>.因此,如何通过人工智能的方式,以仿真图像的形式扩充现有的医学图像数据集,成为了计算机辅助诊断领域亟待解决的问题.

目前,在医学图像仿真生成领域,国内外学者大多集中在图像风格类型转换的研究上,小部分则专注于类似视网膜眼底图像等组织结构简单的图像仿真生成中<sup>[3-4]</sup>,而在需求较大的肝脏图像合成领域却少有研究.

为解决这一问题,本文通过生成对抗网络<sup>[5]</sup>的方式来生成高真实度的仿真医学图像数据集.

生成的数据集能够提高医学图像分类或分割算法的准确率与精确度,也可以用于年轻医师的培训.

# 1 方 法

## 1.1 算法整体方案

算法的整体方案实现主要分为以下三步:

1) 对 CT 序列图像进行预处理,主要包括将三维图像数据解析为若干二维图像、图像格式转换、像素值归一化等.

2) 本文模型是通过 *pix2pixGAN*<sup>[6]</sup> 的改进来实现含有肿瘤的肝脏 CT 图像的仿真生成,如图 1 所示,该模型需要将 CT 肝脏原图与病变区域进行配对,并将该配对图像作为生成对抗网络的输入.因此,本文将肝脏 CT 图像中被标记为肿瘤的像素变为白色,称这个过程为“白化”,本文将原图与“白化”后的图像进行配对,得到“配对一”;然后将其送入生成器中,得到 1 张合成的图像,该合成图像与“白化”图像进行配对,得到“配对二”;再将配对一和配对二同时送入鉴别器中,由鉴别器判断两个配对的相似程度.生成器和鉴别器在模型训练的过程中反复对抗,最终实现对“白化”区域中像素仿照肝脏肿瘤的灰度、纹理、形状进行填充,从而得到合成的有病变的肝脏 CT 图像数据.图 1 中, $G$  表示生成器, $D$  表示鉴别器, $y$  表示原始图像, $x$  表示对原始图像进行“白化”操作后得到的结果, $G(x)$  表示生成器得到的结果.

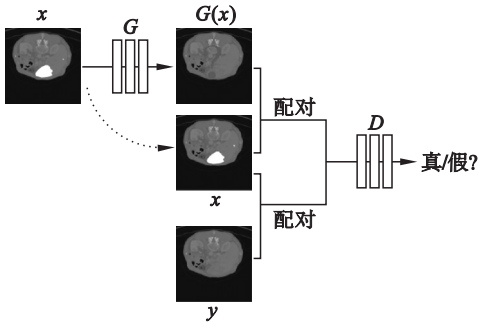


图 1 本文对 *pix2pixGAN* 模型的改进  
Fig. 1 Improved model based on *pix2pixGAN*

3) 将配对后的图片送入生成对抗网络中训练.生成对抗网络中的生成器输入病变白化的图片,然后使生成的图片中白化区域被智能填充而其他的区域则最大程度维持不变.生成的图片与真实的图片被送入鉴别器网络中进行真假鉴别,并且将损失反馈给生成器,两者不断对抗,相互优化,直至生成图片的质量达到最佳效果.

算法整体方案流程图如图 2 所示.

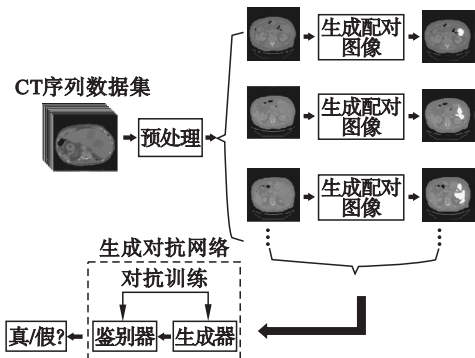


图 2 算法整体方案流程图  
Fig. 2 Flow chart of overall algorithm

## 1.2 搭建生成对抗网络

与普通图像相比,医学图像有着更高的分辨率、更多的纹理、细微结构、更强的相关性、更大的存储空间以及更宽的灰度范围<sup>[7]</sup>.由于医学图像直接涉及到临床医学应用,因此必须要确保医学图像的可靠性、严谨性和高质量.基于 GAN 的医学图像仿真算法正是基于这种理念,采用 GAN 的 *pix2pix* 架构<sup>[6]</sup> 可以对整体肝脏图片进行合成,并且最大程度降低图像细节的模糊度和减免不必要的噪声干扰.

*pix2pix* 架构的生成器采用的是 U-Net 结构<sup>[8]</sup>,这是一种全卷积的网络结构.U-Net 是在编码器-解码器网络的基础上,添加了“跳跃连接”,这样可以减小数据在层与层之间传递时的损耗与丢失.编码器-解码器网络是一个对称的网络,编码器通过下采样(即卷积操作)将图片的分辨率降低,然后再通过解码器上采样(即反卷积操作)还原图片的分辨率.

在图 3 中,卷积层是用来对图像进行特征提取的,反卷积层是进行特征还原的,ReLU 和 tanh 激活层是将输出进行非线性处理的,BN 层是进行归一化处理的.

将单通道、像素为  $512 \times 512$  的病变组织白化图片从编码器一端输入,每经过一个层块,图片的尺寸就会缩小一半.通道数会随着卷积核数目的变化而变化,总体趋势是增加的. $512 \times 512$  像素、通道数为 1 的白化图片,经过一系列卷积操作后,变成了  $1 \times 1$  像素、通道数为 1 024 的一维张量.随后一维张量输入解码器中进行反卷积解码操作,每经过 1 个层块,图片的尺寸就会放大 1 倍,而且通道数在不断减少,最终还原成  $512 \times 512$  像素、通道数为 1 的合成图片.在不断对抗迭代训练后,生成的图片会越来越逼真,清晰度越来越高,

病变白化区域也将会被智能填充。

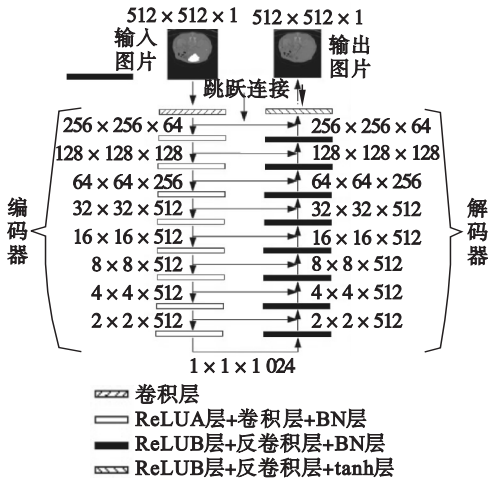


图 3 生成器网络

Fig. 3 Generator network

而在鉴别器中,“输入图像对”共有 2 对:一对为白化图像与目标图像,另一对为白化图像与合成图像.这 2 对图像分别在鉴别器中执行相应的卷积操作(见图 4).pix2pix 的鉴别器与其他 GAN 变体的鉴别器相比,大体相同,但有两处不一样.

1) 鉴别器的输入.白化图像与目标图像在平面法线方向进行拼接,之后形成一个  $512 \times 512 \times 2$  的三维矩阵  $A$ ;同样地,输入图像与合成图像在平面法线方向也进行拼接,形成一个  $512 \times 512 \times 2$  的三维矩阵  $B$ .矩阵  $A$  和矩阵  $B$  同时送入鉴别器网络进行概率输出.

2) 鉴别器采用了 PatchGAN 的思想.一般的鉴别器会根据一对输入图像输出 1 个 0 到 1 之间的 softmax 概率值,但是 pix2pix 的鉴别器是将输入图像划分为一个个区块.这些区块经过一层层卷积后,最后都做了 softmax 概率运算,然后总体输出 1 个大小为  $30 \times 30$ ,通道为 1 的概率矩阵,矩阵上每个点的数值对应着相应区块的 softmax 概率.最后,对大小为  $30 \times 30$  的概率矩阵中的元素求平均值,所得平均值即为鉴别器最终的鉴别概率.由于输入的是独立的区块,大大减少了输入参数量,运算速度比直接输入整张图片的情况还要快.

在 pix2pix 中,采用了 3 个损失函数,分别由式(1)、式(2)和式(3)表示:

$$L_{L_1} = E_{v, r \sim P_{data}(v, r)} [\|r - G(v)\|_1]; \quad (1)$$

$$L_{adv}(G) = E_{v, r \sim P_{data}(v, r)} [\lg D(v, r)_{fake}]; \quad (2)$$

$$L_{adv}(D) = E_{v, r \sim P_{data}(v, r)} [\lg D(v, r)_{real}] + E_{v \sim P_{data}(v)} [\lg (1 - D(v, G(v))_{fake})]. \quad (3)$$

其中: $L_{L_1}$  为  $L_1$  (范数) 损失,是为了减小输出图像和目标图像之间的差异所设定的,相当于粗调作用,是 pix2pixGAN 中必不可少的; $L_{adv}(G)$  为生成器损失,生成器力图减小它,而鉴别器力图增大它; $L_{adv}(D)$  为鉴别器损失,鉴别器力图减小它,而生成器力图增大它,这 2 个对抗损失是 GAN 的核心内容,它们有助于提高图像的细节度,相当于微调作用; $r$  代表真实图像分布; $v$  代表输入给生成器的图像分布; $G(v)$  代表生成器合成的图像分布; $D(v, r)$  代表鉴别器的鉴别概率.

总合成损失应为生成器损失和  $L_1$  损失的加权和,由式(4)表示:

$$L = \lambda_{adv} L_{adv}(G) + \lambda_{L_1} L_{L_1}. \quad (4)$$

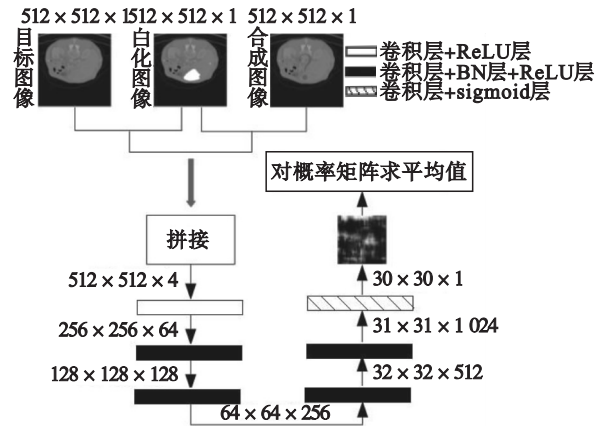


图 4 鉴别器网络

Fig. 4 Discriminator network

## 2 实 验

### 2.1 实验环境

1) 硬件环境. CPU: Intel Core i5 - 7500 @ 3.4 GHz; GPU: NVIDIA GeForce GTX 1080; 内存: 16 GB; 显存: 8 GB.

2) 软件环境. 操作系统: Ubuntu 16.04, 64 位; Python 2.7; tensorflow 1.1; matlab 2016b.

3) 数据集. 2017 年医学图像计算和计算机辅助干预会议中的肝脏肿瘤分割挑战赛 (liver tumor segmentation challenge, LiTS) 数据集.

### 2.2 参数设置

神经网络参数设置如表 1 所示.

### 2.3 训练结果与分析

由图 5 可见,训练的 4 条损失曲线均能够又快又稳地收敛,最终的稳定值也十分令人满意.

对测试集的合成图像与目标图像进行峰值信噪比 (peak signal-to-noise ratio, PSNR) 和结构相似性 (structural similarity index, SSIM)<sup>[9]</sup> 评估,并取平均值,评估结果如表 2 所示.



表 1 训练参数

| Table 1 The parameters of training |         |
|------------------------------------|---------|
| 参数                                 | 设定值     |
| Adam 初始学习率                         | 0.000 2 |
| Adam 动量项                           | 0.5     |
| $L_1$ 损失权重                         | 100     |
| 生成器损失权重                            | 1       |
| 滑动平均衰减率                            | 0.99    |
| batch_size                         | 1       |
| epoch                              | 10      |
| dropout                            | 0.5     |

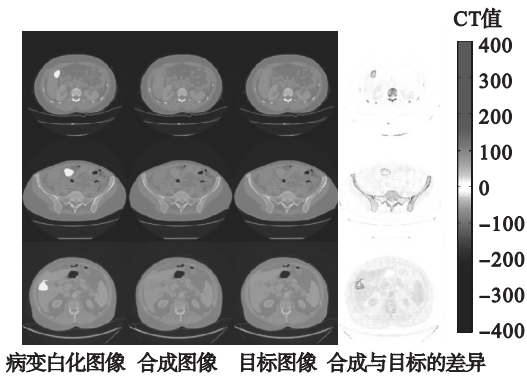


图 6 模型在测试集中的合成效果  
Fig. 6 Synthetic effect of the model in test set

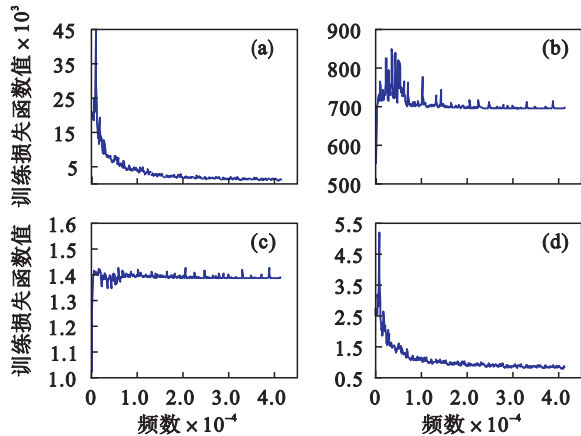


图 5 实验损失曲线  
Fig. 5 Experimental loss curve

(a)— $L_1$  损失曲线；(b)—生成器损失曲线；  
(c)—鉴别器损失曲线；(d)—总合成损失曲线。

表 2 评估结果  
Table 2 Assessment results

| 评估指标    | 评估值      |
|---------|----------|
| PSNR/dB | 64.723 9 |
| SSIM    | 0.997 3  |

在使用测试集生成的图片中,大部分图片的病变白化区域都得到了填充,而且填充效果非常好,完全不存在“棋盘效应”,只是或多或少的在灰度方面与目标图像存在些差异.病变白化区域填充完后与肝脏无缝衔接,之间没有明显的边缘线.病变白化区域的填充在灰度、纹理与风格上都有 GAN 独特的特色,而绝非简单的记忆.病变区域以外的其他区域都最大限度地保持了原状态不变,当然各个像素点之间的灰度值还是有差别的,只不过这种差别已经降到了最小,肉眼上分辨不出来.整张合成图片无论从颜色上还是结构上都显得特别逼真,很多器官组织的细节结构都体现了出来.因此可以说模型的训练是非常成功的,合成图片的质量也令人十分满意,见图 6.

模型训练好后,便可用于实际生产了.从现有的病变分割序列中将病变组织提取出来(也可用 DCGAN<sup>[10]</sup>生成病变组织),再将其放置到任何健康肝脏的任何合理部位,这种排列组合的搭配方式可以产生无穷的白化图片.这个世界上肝脏健康人群远超过不健康人群,而不健康人群的肝脏部位通常也不是从上到下都布满了病变组织,所以不带病变的肝脏横截面图片资源还是可观的.数不尽的病变形状加上数不尽的健康图片排列组合搭配,病变形状可以任意翻转、缩放、旋转,然后再嵌入到健康肝脏的任意合理部位,这样产生的白化图片是创造性的,而且数量足够用.

由图 7 可见,“嵌入病变式白化”的填充合成效果也是非常好,非常逼真,几乎不亚于图 6 测试集中的效果,实验证实了模型用于实际产生病变肝脏图片的可行性.

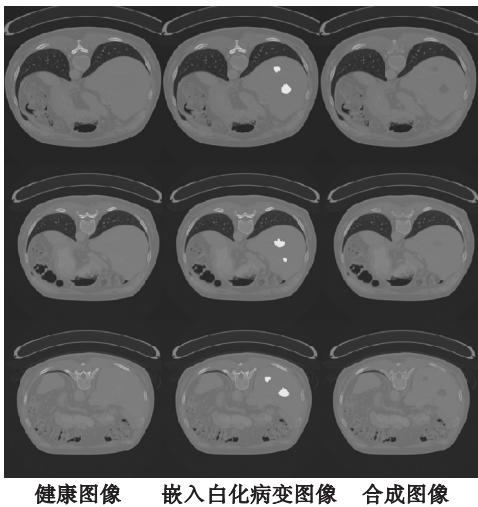


图 7 “嵌入病变式白化”的合成效果  
Fig. 7 Synthetic effect of “embedding pathological whitening”

## 2. 4 同类算法对比

本节将本文算法与其他 4 个同类的、广泛应

用的算法进行定量比较,这 4 个算法包括:基于地图的图像合成算法 (Atlas)<sup>[11]</sup>、基于稀疏表达的图像合成算法 (SR)、基于随机森林和上下文的图像合成算法 (SRF + )<sup>[12]</sup>、基于生成对抗网络的合成算法 (DCAN)<sup>[13]</sup>. 选择的定量比较指标是 PSNR,选择的比较数据集是 LiTS. 定量比较的结果如表 3 所示,通过比较可知本文算法取得了比较好的合成效果,在 PSNR 指标上,优于其他 4 个同类算法. 本文算法的优势主要在于使用了深度生成对抗网络,同时通过“白化”操作对肝部肿瘤进行重点的仿真生成,并且设计了专门的损失函数来保证迭代训练过程中生成的肝部 CT 图像逼真.

表 3 本文算法与同类算法的定量比较  
Table 3 The quantitative comparison between proposed method and four other approaches

| 指标           | 方法     |        |        |        | 本文算法   |
|--------------|--------|--------|--------|--------|--------|
|              | Atlas  | SR     | SRF +  | DCAN   |        |
| mean PSNR/dB | 45. 15 | 49. 77 | 55. 30 | 58. 26 | 64. 72 |

### 3 结 语

本文面向肿瘤肝脏 CT 图像特点,提出了一种基于生成对抗网络的病变肝脏图像仿真生成算法,通过该算法所生成的肝脏 CT 图像,其平均峰值信噪比为 64. 72 dB,平均结构相似性为 0. 997 3. 在实际问题的应用中,该算法表现出了良好的仿真效果,有很高的真实度,可以极大地扩充现有的肝脏 CT 图像数据集.

### 参考文献:

[ 1 ] 周志华. 机器学习 [ M ]. 北京:清华大学出版社,2016:23.  
( Zhou Zhi-hua. Machine learning [ M ]. Beijing: Tsinghua University Press,2016:23. )  
[ 2 ] 陈锟,乔沁,宋志坚. 生成对抗网络在医学图像处理中的应用 [ J ]. 生命科学仪器,2018,16:71 – 80,91.  
( Chen Kun, Qiao Qin, Song Zhi-jian. Applications of

generative adversarial nets in medical image processing [ J ]. *Life Science Instruments*,2018,16:71 – 80,91. )  
[ 3 ] 潘丹,贾龙飞,曾安,等. 生成式对抗网络在医学图像处理中的应用 [ J ]. 生物医学工程学杂志,2018,35 ( 6 ) :970 – 976.  
( Pan Dan, Jia Long-fei, Zeng An, et al. Applications of generative adversarial networks in medical image processing [ J ]. *Journal of Biomedical Engineering*,2018,35 ( 6 ) :970 – 976. )  
[ 4 ] 刘泽宇,郭炜婷. 计算机图像处理技术在医学影像中的进展与应用研究 [ J ]. 中国卫生标准管理,2018,9 ( 9 ) :116 – 118.  
( Liu Ze-yu, Guo Wei-ting. Progress and application of computer image processing technology in medical imaging [ J ]. *China Health Standard Management*, 2018, 9 ( 9 ) : 116 – 118. )  
[ 5 ] Goodfellow I J, Pouget A J, Mirza M, et al. Generative adversarial nets [ C ] // Proceedings of the 27th International Conference on Neural Information Processing Systems. Lake Tahoe, 2014:2672 – 2680.  
[ 6 ] Isola P, Zhu J Y, Zhou T H, et al. Image-to-image translation with conditional adversarial networks [ C ] // IEEE Conference on Computer Vision and Pattern Recognition. Hawaii, 2017: 5967 – 5976.  
[ 7 ] 冯杰,班彪华. 生成对抗网络模型的基本介绍和应用综述 [ J ]. 现代计算机,2019 ( 4 ) :34 – 39.  
( Feng Jie, Ban Biao-hua. Basic introduction and application survey about generating adversarial networks’ model [ J ]. *Modern Computer*,2019 ( 4 ) :34 – 39. )  
[ 8 ] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation [ C ] // International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich, 2015:234 – 241.  
[ 9 ] Zhou W, Alan C B, Hamid R S, et al. Image quality assessment: from error visibility to structural similarity [ J ]. *IEEE Transactions on Image Processing*, 2004, 13 ( 4 ) :1 – 14.  
[ 10 ] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks [ J ]. *arXiv Preprint arXiv*, 2015:1511. 06434.  
[ 11 ] Vercauteren T. Diffeomorphic demons: efficient non-parametric image registration [ J ]. *NeuroImage*, 2009, 45 ( 1 ) :S61 – S72.  
[ 12 ] Huynh T. Estimating CT image from MRI data using structured random forest and auto-context model [ J ]. *IEEE Transactions on Medical Imaging*, 2016, 35 ( 1 ) :174 – 183.  
[ 13 ] 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 ) : 2720 – 2730.