

基于自注意力增强DCGAN的肺结节影像样本生成技术研究

任涵煜, 张致铭, 陈 萌*

宁波工程学院网络空间安全学院, 浙江 宁波

收稿日期: 2026年2月27日; 录用日期: 2026年4月14日; 发布日期: 2026年4月23日

摘 要

基于肺结节医疗影像的自动化检测中训练样本不足的问题, 重点探索了基于Attentional DCGAN (DCGAN-Attn), 一种在DCGAN架构中深度融合注意力机制的新型生成对抗网络的肺结节数据增扩技术。运用改良后的生成对抗网络技术对初始肺结节影像数据集实施数据扩充, 并且在LIDC-IDRI数据集上予以验证。实验结果显示, 与原始GAN相比, DCGAN-Attn技术在生成肺结节影像方面效果更好, FID指标均值达到87.19。并且通过多种生成图像质量评估指标综合测试, DCGAN-Attn网络生成的肺结节生成图像质量更高。

关键词

深度卷积生成对抗网络(DCGAN), 自注意力机制, 医疗影像, 数据增扩

Research on Lung Nodule Image Sample Generation Based on Self-Attention Enhanced DCGAN

Hanyu Ren, Zhiming Zhang, Meng Chen*

School of Cyber Science and Engineering, Ningbo University of Technology, Ningbo Zhejiang

Received: February 27, 2026; accepted: April 14, 2026; published: April 23, 2026

Abstract

Addressing the issue of insufficient training samples in automated detection based on lung

*通讯作者。

文章引用: 任涵煜, 张致铭, 陈萌. 基于自注意力增强 DCGAN 的肺结节影像样本生成技术研究[J]. 软件工程与应用, 2026, 15(2): 267-273. DOI: 10.12677/sea.2026.152025

nodule medical images, this study focuses on a data augmentation approach using Attentional DCGAN (DCGAN-Attn), a novel generative adversarial network that deeply integrates an attention mechanism into the DCGAN architecture. The improved GAN technique is applied to augment the initial lung nodule image dataset, and validation is conducted on the LIDC-IDRI dataset. Experimental results demonstrate that, compared to the original GAN, the DCGAN-Attn method achieves superior performance in generating lung nodule images, with a mean FID score of 87.19. Furthermore, comprehensive evaluations using multiple image quality assessment metrics confirm that the lung nodule images generated by the DCGAN-Attn network exhibit higher quality.

Keywords

Deep Convolutional Generative Adversarial Network (DCGAN), Self-Attention Mechanism, Medical Imaging, Data Augmentation

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

肺癌是全球范围内发病率和死亡率极高的恶性肿瘤之一，其早期发现和诊断对于提高患者生存率至关重要。近年来，随着医学影像技术的快速发展，特别是计算机断层扫描(CT)在肺部疾病诊断中的广泛应用，使得肺结节的检测成为可能。传统的肺结节检测方法主要依赖于放射科医生的经验和视觉判断，这种方法不仅耗时耗力，而且容易受到主观因素的影响，导致漏诊或误诊。为了提高肺结节检测的准确性和效率，研究人员开始探索利用深度学习技术进行自动化检测[1]-[4]。深度学习模型通常需要大量的标注数据进行训练，而高质量的医学图像数据往往难以获取，尤其是在肺结节这种罕见病例的情况下[5][6]。因此，如何有效地扩充肺结节数据集成为了一个亟待解决的问题。

深度卷积生成对抗网络(Deep Convolutional Generative Adversarial Networks, DCGAN)作为一种强大的生成模型，近年来在图像处理领域取得了显著的成果，为解决肺结节数据不足的问题提供了新的思路。本文提出 Attentional DCGAN (DCGAN-Attn)，一种在 DCGAN 架构中深度融合注意力机制的新型生成对抗网络，旨在克服 DCGAN 在生成肺结节影像这种精细结构和复杂场景的图像时，容易产生模糊的细节生成瓶颈，以生成更逼真、细节更丰富的图像的肺结节影像数据增扩技术，以期对肺结节的自动检测和诊断提供更加丰富的数据支持。

2. 研究背景

2.1. 国内外研究现状

目前，国内外学者已经在肺结节检测和生成对抗网络方面开展了大量的研究工作。在肺结节检测方面，许多研究集中在使用卷积神经网络(Convolutional Neural Network, CNN)和其他深度学习算法来提高检测精度[7]-[9]。例如，一些研究通过改进网络结构、优化损失函数或采用多阶段检测策略来提升模型的性能[10]。然而，这些方法往往受限于训练数据的质量和数量，难以达到理想的效果。深度卷积生成对抗网络(Deep Convolutional Generative Adversarial Networks, DCGAN)是由 Radford 等人于 2015 年提出的一种里程碑式的生成模型架构。它标志着生成对抗网络(GAN)在图像合成领域取得的关键突破。DCGAN 系统地将卷积神经网络(CNN)深度集成到 GAN 的生成器(Generator)和判别器(Discriminator)中，取代了之

前 GAN 架构中常用的全连接层。生成器利用转置卷积层(Transposed Convolutional Layers)将低维随机噪声向量逐步上采样,生成具有空间结构的逼真图像;而判别器则采用标准的卷积层(Convolutional Layers)对输入图像进行下采样和特征提取,以判别其真伪。这种设计赋予了 DCGAN 强大的空间特征学习能力,使其能够生成比早期 GAN 模型质量更高、细节更丰富的图像样本。一些研究将 DCGAN 应用于医学图像生成、图像修复等领域,展示了其在处理复杂图像数据方面的潜力。尽管如此,将 DCGAN 应用于肺结节影像数据增扩的研究相对较少,且仍面临诸多挑战,如生成图像的质量、多样性以及与实际应用的结合等。

2.2. 研究内容与创新点

针对现有研究的不足,本文提出了一种基于生成对抗网络(GAN)技术,与卷积神经网络(CNN)深度结合,并引入注意力机制,实现了对肺结节影像数据的有效增扩。具体来说,本研究的主要内容包括:构建了一个生成对抗网络模型,用于生成具有多样性的肺结节图像;引入注意力机制增强模型对关键特征的捕捉能力,对比分析了 DCGAN-Attn 这种在 DCGAN 架构中深度融合注意力机制的新型生成对抗网络与原始的 DCGAN 网络模型在肺结节生成任务中的性能差异。并且验证了基于生成对抗网络的数据生成方法在实际肺结节训练任务中的有效性,为后续相关研究提供了有价值的参考。

3. 相关理论与技术基础

3.1. 肺结节的定义与临床重要性

肺结节是指肺部组织中出现的小块固体或半固体异常组织团块,其直径通常小于 3 厘米。根据大小不同,肺结节可分为微小结节(直径 < 5 mm)、小结节(直径在 5~10 mm 之间)和大结节(直径 > 10 mm)。此外,按照性质的不同,肺结节还可以分为良性结节和恶性结节两大类。良性结节主要包括肉芽肿、错构瘤、纤维瘤等,而恶性结节则多为肺癌的早期表现。早期发现和诊断肺结节对于肺癌的治疗至关重要。研究表明,早期发现的肺癌患者五年生存率明显高于晚期发现者。因此,及时准确地检测出肺结节并进行良恶性评估,有助于医生制定合理的治疗方案,提高患者的治愈率和生活质量。然而,由于肺结节在 CT 图像中的表现较为模糊,且与周围组织对比度低,使得其自动检测成为一项具有挑战性的任务。

3.2. 生成对抗网络原理

生成对抗网络(Generative Adversarial Network, GAN)是一种由生成器(Generator)和判别器(Discriminator)组成的深度学习模型[11] [12]。生成器负责生成尽可能逼真的数据样本,而判别器则负责判断输入的数据是真实的还是由生成器产生的。两者通过对抗训练的方式共同进化,最终使得生成器能够产生高度逼真的数据样本。

GAN 的工作原理可以简单描述为一个动态的博弈过程。在每次迭代中,生成器尝试生成能够欺骗判别器的假数据,而判别器则努力区分真假数据。具体来说,生成器接收随机噪声作为输入,并通过多层神经网络映射到数据空间中,生成伪造的数据样本 $G(z)$ 。同时,判别器接收真实数据样本 $x \sim p_{\text{data}}(x)$ 和生成数据样本 $G(z)$,并输出它们属于真实数据的概率 $D(x)$ 和 $D(G(z))$ 。生成器的目标是最大化 $D(G(z))$,即让判别器难以区分真假数据;而判别器的目标是最小化 $\log(1 - D(G(z))) + \log(D(x))$,即准确区分真假数据。通过不断地优化这个过程,生成器逐渐学会生成越来越逼真的数据样本。

3.3. 本文改进的 DCGAN-Attn

本文提出在 DCGAN 的生成器与判别器中集成自注意力机制(Self-Attention Mechanism)。该创新使模

型能够在所有图像位置之间直接建立远程依赖关系，从而在基于局部卷积特征和全局语义上下文的基础上合成高度一致且细节精细的图像。图 1 为 DCGAN-Attn 的网络结构图。

在 LIDC-IDRI (Lung Image Database Consortium)公共数据集上的实验结果表明，与 DCGAN 相比[13][14]，本文所提出的 DCGAN-Attn 将自注意力模块作为 DCGAN 生成器和判别器的一个有效补充组件。在 DCGAN 的中间卷积层之后插入自注意力模块。该模块以该层的特征图为输入，通过计算查询(Query)、键(Key)、值(Value)矩阵，生成一个注意力图(Attention Map)。该注意力图明确地编码了特征图所有位置之间的相关性权重。对卷积输出的增强。模型既保留了卷积层提取局部纹理和细节的优势，又获得了自注意力层提供的全局结构建模能力。这是一种协同与互补的架构设计。这使得模型能够自适应地关注与当前生成任务最相关的图像区域。不仅在生成器(G)，也在判别器(D)中引入注意力模块。增强后的判别器不再局限于比较图像的局部 patches，而是能够从全局上下文判断图像的真实性。它能更好地识别出因远程依赖关系错误而导致的虚假图像(如不对称、不符合物理规律的物体)，从而为生成器提供更具指导性的梯度反馈，驱动生成质量的整体提升。

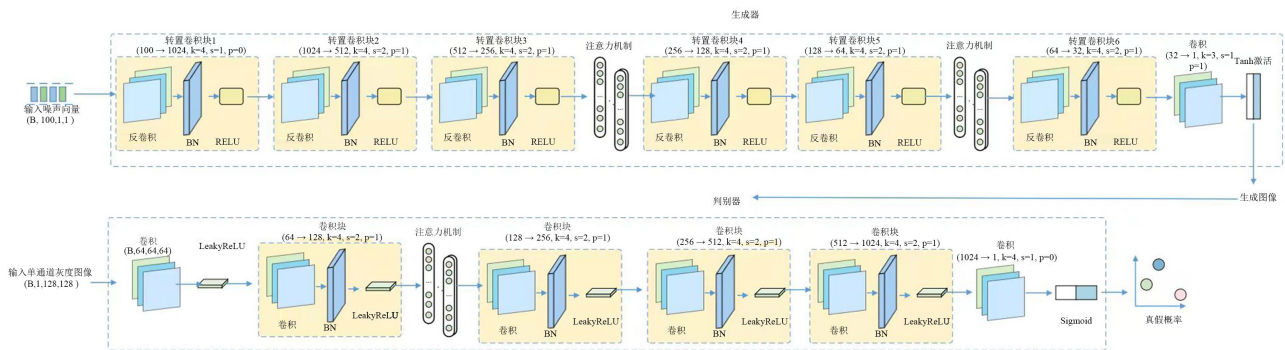


Figure 1. Architecture of the DCGAN-Attn network
图 1. DCGAN-Attn 网络结构图

4. 实验结果

4.1. 生成实验

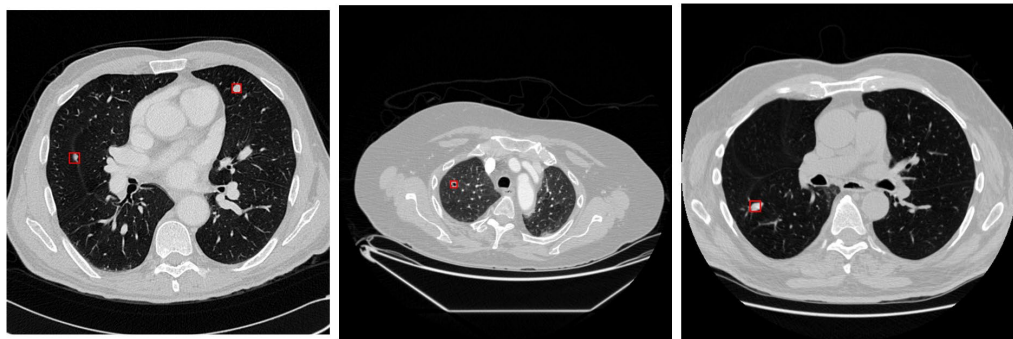


Figure 2. Original lung nodules images from the LIDC dataset
图 2. LIDC 数据集原始肺结节图像

主标本文实验数据来源于 LIDC-IDRI (Lung Image Database Consortium)公共数据库，共纳入 10,158 例病例样本。所有影像均经尺寸标准化处理为 128×128 像素，并将像素值归一化至 $[-1, 1]$ 的区间内。在模型训练阶段，批大小(Batchsize)设为 16，训练周期(Epoch)为 80 轮。将预处理后的图像输入集成自注意

力机制(Self-Attention)的深度卷积生成对抗网络(DCGAN)中进行训练。训练过程中同步监测生成器与判别器的损失变化,并选用结构相似性指数(SSIM)、峰值信噪比(PSNR)、弗雷歇距离(FID)和初始分数(Inception Score)等多类评价指标对生成图像质量进行量化评估。训练结束后保存生成器与判别器权重参数。原始 CT 影像示例如图 2 所示,而经由 DCGAN 网络生成的肺结节图像示例如图 3 所示。

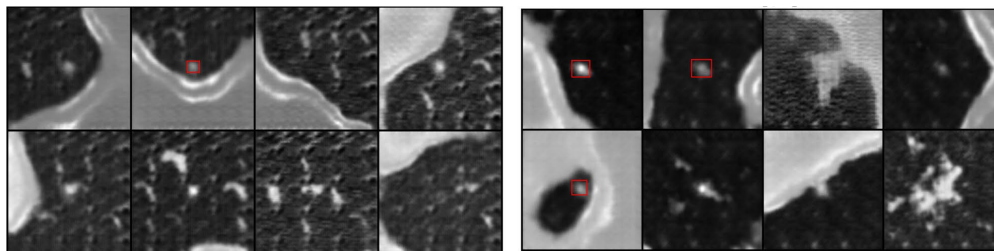


Figure 3. Lung nodules images generated by the DCGAN network
图 3. DCGAN-Attn 网络生成肺结节图像结果

4.2. 实验结果分析

除了主观评价外,我们还采用了一系列客观评价指标来量化生成图像的质量。表展示了所提方法与其他几种常见数据增强方法以及原始图像在峰值信噪比(PSNR)、结构相似性(SSIM)等指标上的对比结果。

在医学图像处理领域,特别是肺部 CT 扫描中肺结节的检测与分析,高质量标注数据的稀缺性是一个长期存在的挑战。本研究旨在通过结合注意力机制和深度卷积生成对抗网络(DCGAN),实现对肺结节影像数据的有效增扩。进一步地,我们引入了注意力机制来提升生成图像的质量与真实性。实验结果表明,该方法能够显著提高肺结节检测模型的性能,为临床诊断提供了有力支持。如表 1 和表 2 所示,实验结果对比了引入注意力机制前后生成图片指标的变化。

Figure 1. Comparison of experimental results

表 1. 实验结果对比

样本	IS	FID	SSIM	PSNR
增扩数据集(1)	1.05	82.56	0.5214	25.40
增扩数据集(2)	1.05	91.82	0.4789	23.22
原始数据集(1)	1.05	106.81	0.3935	19.25
原始数据集(2)	1.07	104.03	0.3620	20.78

另外,我们将原始数据集与经 DCGAN-Attn 增扩后的数据集分别送入主流检测网络(YOLOv8),对比两者的召回率、准确率等指标,以实质性数据证明该增扩技术对临床辅助诊断的增益效果(见表 2)。

Figure 2. Comparison of other indicators

表 2. 其他指标对比

样本	Box (P)	R	mAP50	mAP50-95
增扩数据集(1)	0.813	0.586	0.658	0.326
增扩数据集(2)	0.811	0.574	0.650	0.348
原始数据集(1)	0.566	0.390	0.406	0.152
原始数据集(2)	0.517	0.384	0.387	0.150

为全面评估本文提出的 Attention-GAN 方法在肺结节图像生成中的有效性,实验从图像质量与下游任务适用性两个维度展开分析。图像质量方面,选取了初始分数(IS)、弗雷歇距离(FID)、结构相似性指数(SSIM)以及峰值信噪比(PSNR)作为客观评价指标。由表 1 可知,经注意力机制增强的生成数据集在各项指标上均优于原始数据集。具体而言,相较于原始数据集(1)与原始数据集(2),增扩数据集(1)在 FID 指标上分别降低了 22.7%和 20.6%,SSIM 分别提升了 32.5%和 44.0%,PSNR 分别提升了 31.9%和 22.2%。上述结果表明,注意力机制的引入使生成器更关注结节边缘与纹理等关键结构,显著提升了生成图像的保真度与结构相似性。

为进一步验证生成图像在临床诊断中的实用价值,本文将原始数据集与经 Attention-GAN 增强后的数据集分别输入 YOLOv8 目标检测网络,评估其对肺结节的检测性能。从表 2 可以看出,使用增扩数据集训练的模型在各项指标上均显著优于仅使用原始数据集的模型。其中,增扩数据集(1)的 Box 精确率(P)达到 0.813,召回率(R)达到 0.586,mAP50 与 mAP50-95 分别达到 0.658 和 0.326,相较于原始数据集(1)分别提升了 43.6%、50.3%、62.1%和 114.5%。该结果充分说明,本文方法生成的高质量肺结节图像能够有效缓解原始数据类别不平衡与样本量不足的问题,显著提升检测模型的泛化能力与鲁棒性。

综上所述,本文提出的融合注意力机制的生成对抗网络在肺结节图像生成任务中表现出明显优势,不仅在图像质量客观指标上优于传统方法,且在实际检测任务中展现出良好的临床辅助诊断潜力。

5. 结语

本研究提出的一种注意力机制和 DCGAN 相结合的数据增扩方法,通过创新性地引入注意力机制,验证了基于生成对抗网络的数据生成方法在实际肺结节训练任务中的有效性,为后续相关研究提供了有价值的参考,并有效提升了生成图像的质量和真实性。实验结果表明,该方法能够显著提高肺结节检测模型的性能指标,为临床医学诊断提供了强有力的技术支持。未来将继续优化和完善相关技术体系,推动其在更广泛的医疗应用领域发挥作用。

基金项目

浙江省大学生科技创新活动计划(新苗人才计划 2024R428A024)。

参考文献

- [1] 陈静聪,冉凤伟,章浩伟,等. 基于 DCGAN 的脑膜瘤与听神经瘤检测模型优化方法研究[J]. 波谱学杂志, 2025, 42(2): 117-129.
- [2] 朱勤海. 用于职业性尘肺病分类的生成对抗网络图像增强方法[D]: [硕士学位论文]. 徐州: 中国矿业大学, 2024.
- [3] 董雪莲,付雨菲,蒋凤银,等. 生成对抗网络在放射影像领域的研究进展[J]. 临床放射学杂志, 2026, 45(2): 247-251.
- [4] 邢雄飞,樊星佳,李睿. 基于生成对抗网络的细胞图像超分辨重建[J]. 中国医学物理学杂志, 2026, 43(1): 52-59.
- [5] 魏光松,颜吉平,彭昌勇. CT 影像特征及影像组学在肺部淋巴瘤与肺浸润性黏液腺癌诊断中的应用[J]. 现代医学影像学, 2025, 34(11): 2009-2011, 2015.
- [6] 王淑君,邵圣淇. 肺结节患者 CT 征象分析及肺部影像报告和数据系统分级的诊断价值[J]. 实用医学影像杂志, 2025, 26(4): 311-314.
- [7] 唐潇. 基于深度学习的肺炎影像分析方法研究[D]: [硕士学位论文]. 锦州: 辽宁工业大学, 2024.
- [8] 包强强. 基于深度学习肺部 CT 图像的肺结节检测算法研究[D]: [硕士学位论文]. 包头: 内蒙古科技大学, 2025.
- [9] 魏海月. 基于深度学习的肺结节检测与分类方法研究[D]: [硕士学位论文]. 哈尔滨: 哈尔滨理工大学, 2025.
- [10] 葛侯蒙. 基于深度学习的肺结节检测和分割方法研究[D]: [硕士学位论文]. 沈阳: 沈阳工业大学, 2025.
- [11] 张凤荔,周志远,王瑞锦,等. 基于生成对抗网络的无线电数据增扩与分类[J]. 信息安全学报, 2023, 8(5): 47-60.

- [12] 金强国. 基于深度学习的医疗影像分割模型研究[D]: [硕士学位论文]. 天津: 天津大学, 2021.
- [13] 李杰. 基于深度卷积神经网络的动态手势识别[D]: [硕士学位论文]. 济南: 山东大学, 2019.
- [14] 赵彦杰. 基于深度卷积神经网络的肺结节自动检测方法研究[D]: [硕士学位论文]. 厦门: 厦门大学, 2018.