岩石矿物智能识别中图像样本的智能增量方法研究*

宋飞雪[1,2],孙林[1,2,3],张艳博[2,3],姚旭龙[2,3],陶志刚[4,5],来有邦[6],曹冲[7]

(1.华北理工大学 人工智能学院, 河北 唐山市 063210;
2.河北省矿山绿色智能开采技术创新中心, 河北 唐山 063210;
3.华北理工大学 矿业工程学院, 河北 唐山市 063210;
4.深部岩土力学与地下工程国家重点实验室, 北京 100083;
5.中国矿业大学(北京) 力学与土木工程学院, 北京 100083;
6.河北钢铁集团司家营研山铁矿有限公司, 河北 唐山市 063210;
7.华北理工大学 应急管理与安全工程学院, 河北 唐山市 063210)

摘要:样本稀缺问题严重制约了深度学习技术在岩石矿物智能识别技术领域的推广应用,基于此,提出了一种基于改进的深度卷积生成对抗网络(DCGAN)的岩矿图像智能增量方法。通过引入自注意力机制增强生成图像的细节与真实性,采用带有梯度惩罚机制的损失函数(WGAN-GP)提升训练稳定性,扩展网络结构以增强模型的表现力,并结合多种数据增强技术扩充样本多样性与质量,有效缓解了样本稀缺问题。实验结果表明,所提方法在峰值信噪比、Fréchet初始距离和结构相似性指数等评价指标上均优于对比模型;损失函数曲线显示模型训练过程稳定,具备良好的收敛性能。聚类分析结果表明,模型生成图像在复杂性与多样性方面优于原始图像;问卷调查结果显示,模型能够生成高质量和高真实性的岩矿图像。研究有效缓解了岩石矿物图像样本稀缺问题,为岩石矿物图像生成的智能化发展和实际应用奠定了基础。

关键词:矿物智能识别;深度学习;生成对抗网络;图像生成;数据增强

中图分类号:TP391.8;P579 **文献标识码**:A **文章编号**:1005-2763(2025)12-0271-12

Research on Intelligent Incremental Method for Image Samples in Intelligent Recognition of Rock and Mineral

SONG Feixue[1,2], *SUN Lin*[1,2,3], *ZHANG Yanbo*[2,3], *YAO Xulong*[2,3], *TAO Zhigang*[4,5], *LAI Youbang*[6], *CAO Chong*[7]

(1.College of Artificial Intelligence, North China University of Science and Technology, Tangshan, Hebei 063210, China;
2.Green Intelligent Mining Technology Innovation Center of Hebei Province, Tangshan, Hebei 063210, China;
3.College of Mining Engineering, North China University of Science and Technology, Tangshan, Hebei 063210, China;
4.State Key Laboratory for Geomechanics and Deep Underground Engineering, Beijing 100083, China;
5.School of Mechanics and Civil Engineering, China University of Mining and Technology-Beijing, Beijing 100083, China;
6.Sijaying Yanshan Iron Mine Co., Ltd., Hebei Iron and Steel Group, Tangshan, Hebei 063210, China;
7.School of Emergency Management and Safety Engineering, North China University of Science and Technology, Tangshan, Hebei 063210, China)

Abstract: The scarcity of samples seriously restricts the promotion and application of deep learning technology in the field of intelligent recognition of rock and mineral. Therefore, an intelligent incremental method for image samples of rock and mineral based on an improved deep convolutional generative adversarial network (DCGAN) was proposed. By introducing self-attention mechanism to enhance the details and authenticity of generated images, using loss function with gradient penalty mechanism (WGAN-GP) to improve training stability, expanding network structure to enhance model performance, and combining various data augmentation techniques to expand sample diversity and quality, the problem of sample scarcity has been effectively alleviated. The experimental results show that the proposed method outperforms comparative models in evaluation indicators such as peak signal-to-noise ratio, Fréchet initial distance, and structural similarity index. The loss function curves show that

* **收稿日期**:2024-11-09

基金项目:国家自然科学基金项目(52074123);河北省自然科学基金项目(E2022209143,E2023209087);河北省创新能力提升计划项目(23564201D);唐山市科技计划项目(22130224H)

作者简介:宋飞雪(1999—),女,河北唐山人,硕士研究生,主要从事岩石矿物图像处理方面的研究。E-mail:1792801116@qq.com

通信作者:孙林(1981—),男,河北唐山人,博士,副教授,主要从事数字信号采集、处理相关研究工作。E-mail:sunlin@ncst.edu.cn

the model training process is stable and has a good convergence performance. The clustering analysis results indicate that images generated by the model are superior to the original images in terms of complexity and diversity. The survey results show that the model can generate high-quality and highly realistic images of rock and mineral. The research effectively alleviated the problem of scarcity of rock and mineral image samples, laying the foundation for the intelligent development and practical application of rock and mineral image generation.

Key words: Intelligent identification of minerals, Deep learning, Generative adversarial network, Image generation, Data enhancement

0 引言

岩石矿物的智能识别技术在地质学和矿物学领域扮演着至关重要的角色,在矿产勘探[1]、矿产资源开发[2]、环境保护以及地质灾害防治[3]中发挥着重要作用。传统的矿物识别方法依赖专业人员将岩石制备成薄片,并借助偏光显微镜等精密仪器进行多角度观察,分析岩石薄片的颜色、纹理、形状等微观特征[4-6]。此外,X 射线衍射、扫描电子显微镜、拉曼光谱等技术也被应用于矿物分析[7]。然而,这些方法不仅耗时耗力,还对操作人员的技术水平要求极高,极大地限制了其推广应用和自动化程度。

近年来,深度学习技术,特别是卷积神经网络(Convolutional Neural Network,CNN)已成为推动岩石矿物智能识别领域发展的关键驱动力。通过学习岩石矿物图像的复杂模式和特征,深度学习技术实现了自动化的岩石矿物识别[8]。与传统的基于规则的识别方法相比,深度学习方法能够处理海量数据并自动提取特征,从而降低对专业知识的依赖。此外,深度学习技术能够从图像中学习到岩石矿物的微观结构和纹理特征。尽管深度学习技术在岩石矿物识别领域展现出巨大潜力,其应用效果仍高度依赖于训练样本的质量和数量[9]。样本的稀缺性不仅限制了模型的泛化能力,还影响其在实际应用中的准确性和可靠性。在地质学领域,获取充足且高质量的岩矿样本用于训练深度学习模型,往往面临诸多挑战,包括高昂的成本、复杂的样本制备过程以及对专业技术的严格要求[10]。因此,扩充岩石矿物样本库并提高样本的多样性和质量,已成为当前研究的关键问题。

为了解决样本稀缺问题,研究者们广泛采用数据增强技术扩充样本库。传统数据增强方法如几何变换和像素级操作等[11-14],通过旋转、翻转、增加噪声和调节亮度等方式改变图像的外观和内容,增加样本的多样性。然而,这些方法可能会引起图像变形和失真,也会导致增强后的样本与原始图像高度相似[15],从而增加了过拟合的风险,限制了深度学习算法的效能。GOODFELLOW 等[16]在 2014 年首次提出生成对抗网络(Generative Adversarial Network,GAN)概念,该技术基于博弈论原理,通过模仿真实数据分布来生成样本图像,为数据增强领域带来了全新的视角,其能够学习样本的高维特征并生成高质量、逼真的图像,为解决岩石矿物样本稀缺问题提供了一种高效的策略[17]。与传统的数据增强技术相比,GAN 在生成多样化岩石矿物图像方面展现出巨大潜力。尽管在地质学领域的应用尚处于起步阶段,但 GAN 已经在自动化[18]、医学[19-20]等领域证明了其强大的图像生成能力。然而,传统的 GAN 在处理不平衡数据集时可能会遇到模式崩溃[21],特别是在岩石矿物图像的纹理和结构信息差异显著的情况下,容易导致生成的图像质量下降,从而影响模型的整体性能。为解决这一问题,RADFORD 等[22]提出了深度卷积生成对抗网络(Deep Convolutional GAN,DCGAN),该模型结合了卷积神经网络的特征提取能力,稳定性和生成图像的能力优于传统 GAN。HUSSAIN 等[23]研究表明,DCGAN 在生成逼真图像方面具有显著优势,尤其适用于地质学领域对图像质量的高标准要求。王琳等[24]采用 DCGAN 对数据集中矿物图像较少的矿物类别进行数据增强,结果表明,增强方法使模型精度提升了 3.12%。赵高昌等[25]通过 DCGAN 进行数据增强,提高了岩石图像的分类精度,然而 DCGAN 在生成复杂结构的大尺寸矿物图像时,仍存在一定的稳定性问题,影响了生成图像的细节质量与分辨率。

针对现有技术的局限性,本研究提出了一种创新的改进型深度卷积生成对抗网络方法,旨在实现岩石矿物薄片图像数据样本的增量生成与质量提升。核心改进措施包括:首先,引入自注意力机制至生成器架构中,显著优化网络对岩矿图像关键区域的识别能力,提高生成图像的细节表现和整体质量;其次,采用带有梯度惩罚的 Wasserstein 距离解决传统 GAN 训练过程中可能出现的模式崩溃与梯度消失问题,显著提升模型的稳定性与收敛效率;此外,在生成器和判别器中增设三层转置卷积层,以增

强模型生成大尺寸、高分辨率图像的能力,并提高对图像真伪的判别精度;最后,采用统一剪裁、伽马变换、镜像旋转和高斯噪声变换等多种数据增强方法,进一步提升模型对岩矿图像的鲁棒性与泛化能力。试验结果表明,改进型 DCGAN 在生成岩石矿物图像的多样性和质量方面表现优异。改进后的模型不仅减少了图像生成过程中可能出现的过拟合现象,还显著提升了生成图像的纹理清晰度与视觉真实感。试验中,模型生成了高分辨率、类型多样的岩矿图像,为数据集的扩展和多样性增添了实质性内容。本研究提出的创新方法不仅有效解决了岩矿样本稀缺的问题,还为地质学领域的岩石矿物智能识别与分类任务提供了丰富的数据支持,推动了矿业智能化发展的进程。

1 基于 DCGAN 的岩石矿物样本智能生成方法

1.1 GAN 基本原理

生成对抗网络是一种先进的深度学习生成模型,通过引入对抗学习机制,实现了基于博弈论的机器学习方法。GAN 由两个核心组件组成:生成器(Generator)和判别器(Discriminator)。如图 1 所示,生成器以随机噪声向量为输入,目标是生成一组伪造数据 $G(z)$,使其在分布上尽可能接近真实样本,以此试图"欺骗"判别器。判别器的任务是从特征区分真实图像与生成器产生的伪造图像,并不断提升其判别能力。在训练过程中,生成器不断优化生成的伪造数据,使其愈加逼近真实样本,而判别器则致力于提高对真实数据和伪造数据的区分准确性。这一过程是动态的,两者相互竞争,相互优化,直至达到一种平衡状态,即纳什均衡(Nash Equilibrium)。在此状态下,判别器无法有效区分真实数据与生成数据,生成器生成的样本质量达到最佳。

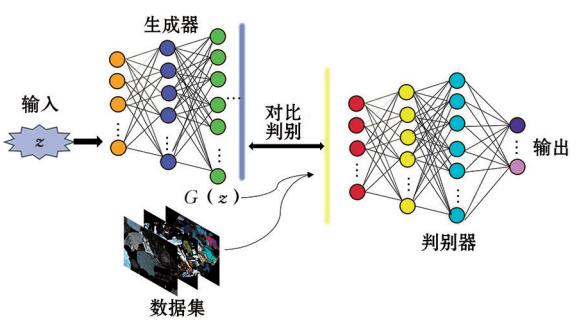


图 1 岩矿图像生成对抗网络原理

Fig.1 Principle of generation adversarial network for rock and mineral images

生成器的工作原理是接收随机噪声向量 z 作为输入,并通过一系列转置卷积层和批量归一化层的神经网络将其转化为具有高逼真度的视觉特征图像样本 $G(z)$。判别器则接收图像样本 x,通过由卷积层和批量归一化层构成的神经网络对输入图像是否为真实数据进行判断,其概率为 $D(x)$。

生成器 G 和判别器 D 的训练目标是通过对抗学习优化以下损失函数:

$$\min_G \max_D V(D,G)=E_{x-P_{\text{data}}}[\log D(x)]+E_{z-p_z}[\log(1-D(G(z)))] \quad (1)$$

式中:$D(x)$表示判别器 D 对真实样本 x 判别为真实的概率;$G(z)$表示生成器 G 基于随机噪声 z 生成的图像;$E_{x-P_{\text{data}}}$ 表示真实数据分布 P_{data} 的期望值;E_{z-p_z} 表示噪声分布 p_z 的期望值。

在这个优化目标中,生成器的损失函数旨在最小化生成样本与真实样本之间的差异,而判别器的损失函数则旨在最大化区分真假样本的能力。两者之间对抗性训练不断推动网络向生成高质量、高逼真度的图像样本方向发展。

1.2 DCGAN 核心原理

深度卷积生成对抗网络 DCGAN 是一种先进的 GAN 变体,结合了卷积神经网络的强大特征提取能力,对图像生成任务进行了优化。作为一种高效的前馈神经网络,CNN 的卷积结构通过局部感受野机制显著减少了深层网络的内存占用,并有效降低了参数数量。这种设计不仅缓解了过拟合问题,还增强了 DCGAN 的特征提取能力,使其成为图像生成领域的优选算法。

DCGAN 的网络结构与传统 GAN 相似,均由生成器和判别器组成,但为提高模型的性能和稳定性,DCGAN 进行了以下关键改进[26]。

(1) 去除池化层。DCGAN 舍弃了传统的池化层设计,转而在生成器中使用反卷积层(转置卷积层)逐步增加图像的分辨率,并在判别器中通过卷积层直接提取特征。这种方法避免了池化操作可能导致的细节损失,更有助于生成高质量图像。

(2) 去除全连接层。为了保留网络的局部连接特性,生成器和判别器均去除了全连接层。这一设计减少了网络参数数量,提高了模型的计算效率,同时避免了全连接层可能引起的信息损失。

(3) 批量归一化。生成器和判别器的所有卷积层均统一采用批量归一化,以稳定训练过程、缓解梯

度消失或梯度爆炸现象,并加速模型收敛。

(4) 激活函数的选择。生成器中除最后一层使用 Tanh 激活函数以标准化生成的像素值外,其他层均采用 ReLU 激活函数,确保生成器输出的非线性表达能力。判别器中除最后一层采用 Sigmoid 激活函数以输出图像真假概率外,其他层均使用 LeakyReLU 激活函数,允许输入为负值时保留小梯度,从而避免因输入负值而导致的梯度消失问题。

上述优化措施显著提升了 DCGAN 在生成高质量、高分辨率图像样本时的表现,尤其适用于处理纹理特征复杂的岩石矿物图像样本。

1.3 基于改进的 DCGAN 岩矿图像智能增量模型

岩矿图像的纹理特征结构复杂多样且种类繁多,传统深度卷积生成对抗网络模型在捕捉这些细微特征方面存在局限,导致生成图像时容易忽略小目标。同时,网络在图像生成过程中容易出现模式崩溃现象。针对上述问题,本研究提出了一种改进的 DCGAN 网络模型,作为岩矿图像智能增量试验的初步建模基础。改进措施包括以下几个关键方面。

(1) 引入自注意力机制[27]增强模型对岩矿图像关键区域的聚焦能力,从而提升生成图像的清晰度和真实性,有效避免图像模糊和模式崩溃问题。

(2) 优化损失函数解决 DCGAN 训练过程中可能出现的模式崩溃和梯度消失问题,确保模型训练的稳定性与高效性。

(3) 扩展判别器和生成器的网络结构,提高生成器在生成高分辨率、高真实感岩矿图像方面的能力,并增强判别器对图像真伪的识别精度。

(4) 结合多种数据增强技术,包括统一剪裁、伽马增强、镜像旋转和高斯噪声变换等,以增强模型对岩矿图像的鲁棒性和泛化能力。

1.3.1 引入自注意力机制

DCGAN 生成器结构主要基于卷积操作,其本质是对输入岩矿图像逐步进行尺寸扩展与通道数量的减少,但每次的操作覆盖范围有限[28],导致多次卷积操作之后出现部分关键像素丢失的情况。为解决上述问题,本研究在 DCGAN 的生成器中引入了自注意力机制。该机制通过计算图像内部特征的关联性,使得模型能够识别并强调矿物的关键微观特征。自注意力机制原理如图 2 所示。

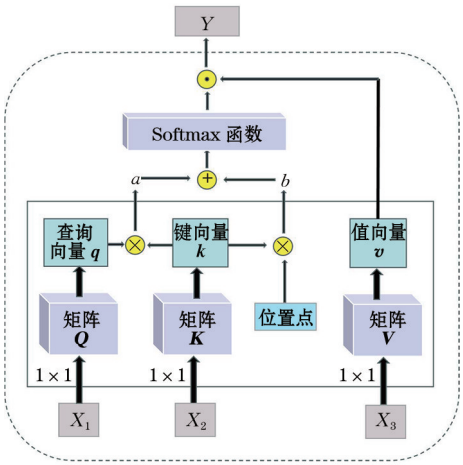


图 2 自注意力机制原理示意

Fig.2 Schematic of the self-attention mechanism

自注意力机制原理公式如下:

$$A(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})=\mathrm{Softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d_{\mathrm{k}}}}\right)\boldsymbol{V} \tag{2}$$

式中:$\boldsymbol{Q}$ 为查询(Query)矩阵,用于表示当前位置的重要程度;$\boldsymbol{K}$ 为键(Key)矩阵,用于表示其他位置对当前位置的重要程度;$\boldsymbol{V}$ 为数值(Value)矩阵,用于表示当前位置的特征;d_{k} 为键值映射的维度;Softmax 函数用于计算注意力权重,确保所有注意力权重的总和为 1。

1.3.2 优化损失函数

损失函数在模型训练中起着关键性的作用,其核心作用是在模型进行训练时指导模型中的参数更新和优化。传统 GAN 训练过程常常表现出不稳定性,这一问题在 DCGAN 网络中同样显著,甚至可能导致模型训练崩溃。为解决这一问题,本研究采用带有梯度惩罚机制的 WGAN-GP[29]替代 DCGAN 原有的损失函数。WGAN-GP 使用 Wasserstein 距离来度量生成器和判别器之间的分布差异,具有更高的稳定性。此外,WGAN-GP 引入的梯度惩罚机制通过约束生成器和判别器的参数更新,确保两者之间的训练保持平衡,有效防止梯度爆炸或梯度消失问题。计算公式如下:

$$L(G,D)=E[D(y)]-E[D(G(x))]+\lambda E[(\|\nabla_{\hat{x}}D(\hat{x})\|_2-1)^2] \tag{3}$$

式中:$\hat{x}$ 为生成图像与目标图像对应点之间直线上的样本,单位为像素值;λ 为梯度惩罚的权重;x、y 分别表示整体网络的输入图像和输出图像。该方法的计算思路是放弃对整个样本空间的 $D(\hat{x})$ 求导,

只需要对真实样本空间与生成样本空间的中间值求导,从而稳定训练过程。

1.3.3 扩展网络架构

在对 DCGAN 网络进行改进的基础上,原始生成器新增了三层卷积转置层,显著提高了生成图像的复杂性和细节的精细度,使生成的岩矿图像在视觉表现上更加丰富和自然。同时,在判别器中新增三层卷积层,进一步增强了其对生成图像与真实图像之间微妙差异的捕捉能力。

这种双向的结构优化不仅提升了模型的整体表现力,还有效优化了梯度流动,显著增强了模型训练的稳定性和收敛效果。改进的 DCGAN 模型(Self-GP-DCGAN)架构如图 3 所示,展现了生成器与判别器在结构优化后的协同工作机制。

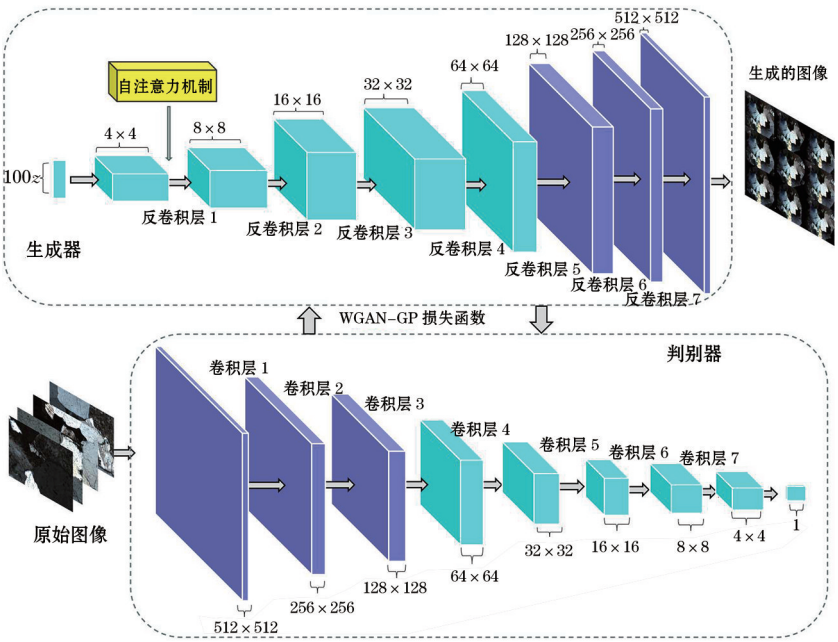


图 3 Self-GP-DCGAN 模型架构

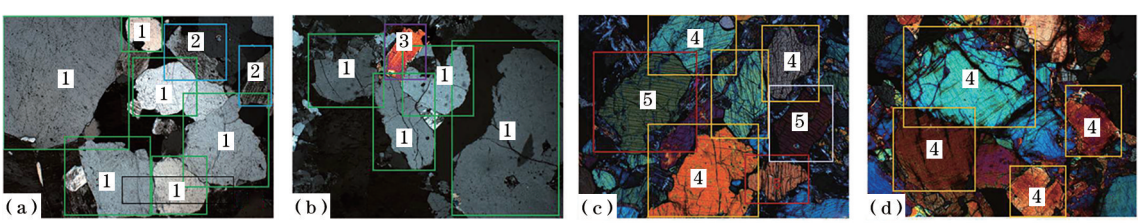
Fig.3 Architecture of Self-GP-DCGAN model

2 岩矿图像智能增量模型应用与验证

2.1 岩矿图像样本数据集

研究试验数据集来自华北理工大学矿业工程专业的岩矿薄片教学样本,样本通过偏光数码显微镜在实验室条件下拍摄,获取了单偏光图像和正交偏光图像。为了全面捕捉不同偏振方向和光强条件下的岩矿特征,每张薄片均进行了多角度和多次拍摄。试验中采用的偏光显微倍数为 50 倍,共采集了 840 组图像,每张图像的分辨率为 1 292 像素×968 像素。

基于岩石中矿物成分的多样性,试验将采集到的薄片图像分为两大类别:石英闪长岩和辉橄榄岩。石英闪长岩主要包含石英、长石和角闪石,而辉橄榄岩则以辉石和橄榄石为主。各类矿物的具体形态如图 4 所示,其中图 4(a)、图 4(b)为石英闪长岩样本,图 4(c)、图 4(d)为辉橄榄岩样本。



1—石英;2—长石;3—角闪石;4—橄榄石;5—辉石

图 4 岩矿图像样本

Fig.4 Samples of rock and mineral images

石英闪长岩中石英呈半透明状,长石以均匀的灰白色纹理分布,角闪石则为鲜艳的橙黄色;辉橄榄岩中橄榄石为绿色或黄绿色的四面体或六面体结构,辉石则呈深绿色或黑色。

这些岩矿样本经过拍摄、选样、去重和去暗等一系列处理操作,能够作为试验数据使用。为丰富数据集以满足模型训练需求,对样本进行了数据增强处理,具体包括以下步骤。

(1) 剪裁。去除图像边缘模糊部分,将样本统一剪裁为 800 像素×1 000 像素(高×宽),以规范图像尺寸。

(2) 伽马变换。修正图像中因过曝或过暗引起的亮度不均问题[30],针对拍摄样本普遍偏暗的现象,通过试验优化,确定伽马值 $\gamma=1.3$ 为最佳亮度矫正参数。

(3) 镜像和旋转。对剪裁后的样本进行镜像变换以及 30°、60°、90°、150°、180°的旋转变换处理。由于这些变换不改变图像的物理性质,因此在扩充数据集的同时能够保持样本的原生形态。

(4) 高斯噪声处理。在原始图像中添加高斯噪声,通过改变像素点分布有效增加样本多样性,从而减少模型训练过程中可能出现的过拟合现象。

图 5 为岩矿样本图像数据增强处理流程。经过上述增强步骤,基于数据库原始的 1 093 张样本最终生成了石英闪长岩样本 3 880 张,辉橄榄岩样本 2 208 张,为试验提供了充足且多样化的数据基础。

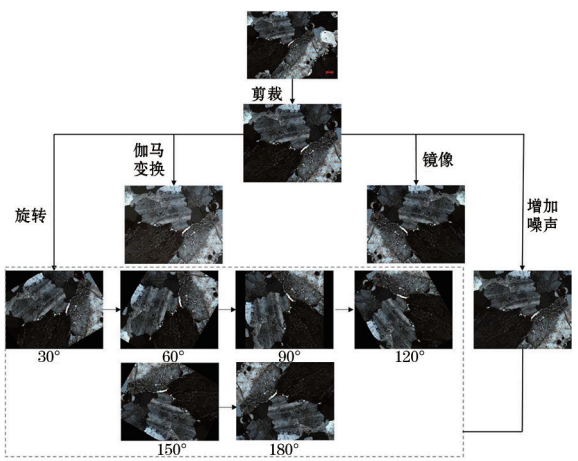


图 5 岩矿图像数据增强处理流程

Fig.5 Flow of rock and mineral image data enhancement processing

2.2 模型构建及评价指标

为实现模型训练,首先收集并整理了所需的岩矿图像数据集,确保其涵盖了研究所需的不同岩矿类别和特征,同时具备足够的样本量。在实施过程中,对两组试验样本进行了预处理,将图像尺寸统一调整至 512 像素×512 像素,以满足模型结构的要求。岩矿图像数据智能增量模型环境配置见表 1。

表 1 岩矿图像数据智能增量模型环境配置

Table 1 Intelligent incremental model environment configuration for rock and mineral image data

软/硬件	配置
GPU	NVIDIA A100
内存	80G
Python 版本	Python 3.7.0
深度学习框架	Tensorflow 1.15.0
可视化软件包	NumPy 1.21.6

生成对抗网络通常使用峰值信噪比 PSNR(R_{PSN})、Fréchet 初始距离 FID(D_{FI})、结构相似性指数 SSIM(I_{SS})[31]作为评估模型好坏的评价指标。

PSNR 是用于评估两张图像之间的像素级差异,适用于岩矿图像结构不同的情况,因此可以直观地评估生成器的性能。计算公式如下:

$$R_{PSN}=10\times\log_{10}\frac{I_{max}^2}{E_{MS}} \tag{4}$$

式中:I_{max} 为图像的最大像素值;E_{MS} 为均方误差。PSNR 的单位为 dB。PSNR 值越大,表示两张图像的差异越小,即生成的图像质量越好。

FID 通过计算两个分布之间的 Fréchet 距离来衡量生成模型和真实数据分布之间的差异。该指标考虑了两个分布的均值和协方差矩阵的高层次的特征信息,可以更好地描述两个分布之间的差异,其值越小,表示生成图像和真实图像特征更加逼近。计算公式如下:

$$D_{IF}=\|\mu_r-\mu_g\|^2+Tr(C_r+C_g-2(C_rC_g)^{\frac{1}{2}}) \tag{5}$$

式中:μ_r、μ_g 分别为真实样本和生成样本的高斯分布特征均值;C_r、C_g 分别为真实样本和生成样本的高斯分布特征的协方差矩阵;Tr 为矩阵的迹,即矩阵对角线元素的和。

SSIM 通过亮度、对比度、结构三个指标进行评估,该指标总体以人眼视觉的主观感知角度出发,以上述三种指标进行组合计算,其值介于−1 和 1 之间,值越大表示两幅图像越相似。表达式如下:

$$I_{SS}=\frac{(2\mu_r\mu_g+c_1)(2\sigma_{rg}+c_2)}{(\mu_r^2+\mu_g^2+c_1)(\sigma_r^2+\sigma_g^2+c_2)} \tag{6}$$

式中:σ_{rg} 为真实样本和生成样本在局部窗口内的协方差,表示图像的结构相似性;μ_r^2、μ_g^2 分别为真实样本和生成样本在局部窗口内的方差,表示图像的对比度;σ_r、σ_g 则表示图像的亮度水平;c_1、c_2 为大于 0

的常数。

本研究综合采用 PSNR、FID 和 SSIM 指标,从不同角度全面评估岩矿图像智能增量模型生成图像的质量。

3 图像生成结果与质量评价

3.1 图像生成示例

岩矿图像智能增量模型成功生成了两组岩石矿物图像,结果如图 6 和图 7 所示。这些图像是从模型生成的图像集中随机选取,作为具有代表性的样本组合。通过观察可以发现,模型生成的图像纹理清晰可见,其形态特征在一定程度上区别于原始岩矿图像,同时仍保留了岩矿本身的典型特征。



图 6 石英闪长岩图像

Fig.6 Images of quartz diorite

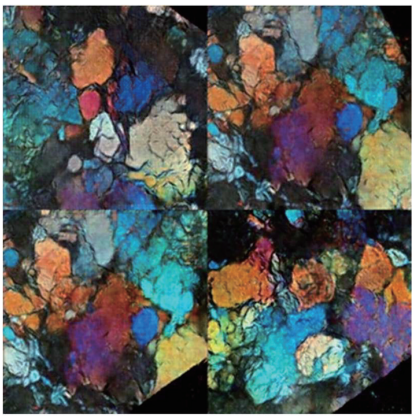


图 7 辉橄榄岩图像

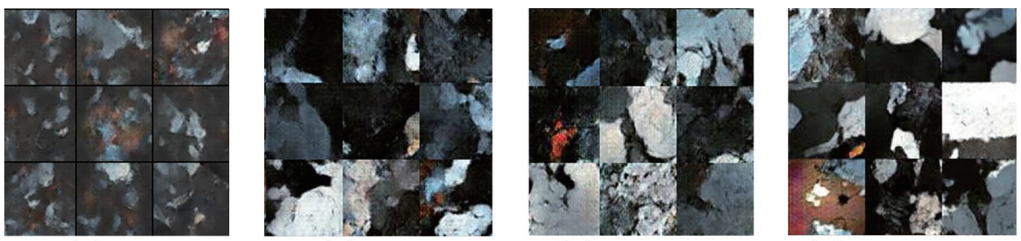
Fig.7 Images of peridotite

试验结果表明,所构建的岩矿图像智能增量模型不仅能够生成具有多样性的岩石矿物图像,还能有效反映岩石矿物在自然界中可能呈现的不同形态。模型生成的图像为岩石矿物的识别与分类任务提供了丰富的多样性样本,进一步验证了模型在数据扩充方面的实用价值,体现了其在大规模生成高质量、多样化岩矿样本方面的优势,不仅为深度学习模型的训练提供了充足的数据支持,也为进一步提升模型精度开辟了新的技术路径。

3.2 图像质量评价

3.2.1 定量评价

将本次试验模型与 GAN、DCGAN、引入自注意力机制的 Self-DCGAN 等具有代表性的模型进行对比试验。图 8 为各模型在石英闪长岩岩矿图像数据集上的结果对比(颜色标识见电子版)。



(a) GAN (b) DCGAN (c) Self-DCGAN (d) Self-GP-DCGAN

图 8 各模型生成岩矿图像对比

Fig.8 Comparison of rock and mineral images generated by each model

从图 8 可以看出,GAN 模型生成的岩矿图像存在较大的缺陷,表现为色彩保真度较低、纹理边缘模糊且严重失真。相比之下,DCGAN 模型可以较好地保留图像色彩信息,但图像的纹理结构仍然不够清晰。引入自注意力机制的 Self-DCGAN 模型生成的岩矿图像在纹理细节方面有了显著提升,生成的岩矿图像纹理更加清晰,边缘失真问题也得到了有效缓解。在此基础上,进一步结合 WGAN-GP 梯度惩罚机制的 Self-GP-DCGAN 模型不仅生成的岩矿图像纹理更为清晰,同时模型在训练过程中的稳定性也显著增强。

为定量评估不同模型生成图像的质量,采用峰

值信噪比 PSNR、Fréchet 初始距离 FID 和结构相似性指数 SSIM 等评价指标,结果见表 3。结果显示,本文所提出的 Self-GP-DCGAN 模型在生成高质量图像方面具有显著优势。

表 2 不同模型的定量评价结果

Table 2 Quantitative evaluation results of different models

方法模型	PSNR/dB	FID	SSIM
GAN	27.84	271.70	1.5
DCGAN	27.93	246.62	1.6
Self-DCGAN	28.02	237.05	1.9
Self-GP-DCGAN	28.13	232.20	2.1

由表 2 数据可知,在 PSNR 方面,改进后的 Self-GP-DCGAN 模型的 PSNR 值相较于 GAN 模型和 DCGAN 模型仅存在微小提升;但 Self-GP-DCGAN 模型得 FID 值较 GAN 模型降低了14.5%,较 DCGAN 模型降低了 5.8%,说明改进后的模型生成图像的质量和分布相似性有明显提升。此外,因生成图像样本的多样性较高,各模型的 SSIM 值均偏低,但改进模型的 SSIM 值仍达到 2.1,较 GAN 模型提高 40%,较 DCGAN 模型提高 31%,反映出改进后模型生成图像在图像纹理清晰度和结构相似性方面的显著进步。

综上所述,引入自注意力机制与 WGAN-GP 损失函数的 Self-GP-DCGAN 模型在生成图像的视觉质量、分布相似性和结构完整性方面均优于对比模型,展现出良好的综合性能。

3.2.2 损失函数曲线

损失函数是评估基于 DCGAN 的岩矿图像智能增量模型训练稳定性的关键指标之一。通过观察模型训练的损失函数曲线,可以分析生成器和判别器的训练表现及其优化效果。模型训练过程中,判别器的损失函数逐渐减小,表明其区分真实图像与生成图像的能力不断增强;生成器的损失函数逐渐减小,表明生成器能够生成更加逼真的图像,并逐渐接近真实图像分布。

为进一步探索和比较不同模型的性能,本研究设置了三种模型架构,并记录了其在相同岩矿图像样本数据集上的损失函数的变化趋势,如图 9 所示。

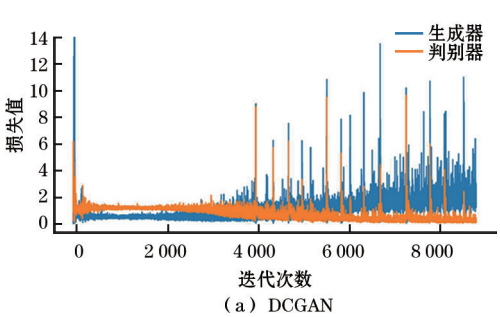


图 9 岩矿图像智能增量模型损失函数对比

Fig.9 Comparison of loss functions of intelligent incremental models for rock and mineral image

对 3 种模型的损失函数曲线进行分析,总结得到以下规律。

(1) DCGAN 模型:模型训练过程中,生成器和判别器的损失函数曲线存在较大波动,这种波动可能源于原始损失函数在训练过程中的不稳定性,导致模型难以有效收敛。

(2) Self-DCGAN 模型:引入自注意力机制显著改善了模型的训练表现,生成器和判别器的损失函数波动明显减小,表明自注意力机制能够提高模型在训练过程中的稳定性,并有效减少损失函数的波动幅度。

(3) Self-GP-DCGAN 模型:在引入自注意力机制的基础上,进一步采用 WGAN-GP 替代传统损失函数,显著提升了模型的稳定性。在这种设置下,生成器和判别器的损失函数几乎没有大幅波动,展现出极高的稳定性。

综上分析结果可以看出,引入自注意力机制和 WGAN-GP 损失函数的 Self-GP-DCGAN 模型在训练过程中表现出高度稳定性,大幅提高了模型在实际应用中的可靠性与有效性。

3.3 基于聚类分析的多样性评价

岩矿图像样本在颜色、形状和纹理等方面具有丰富的多样性。研究显示,模型生成的岩矿图像(如图 10 所示)与原始图像具有较高的相似性,同时在视觉上展现出了明显的个体差异。同时在纹理方面相较于原始图像表现出较大的差异性,这表明本文所构建的模型不仅能够生成多样化的岩矿样本,还

能够合成自然条件下出现概率较低的岩矿样本，进一步拓展了数据集的范围。

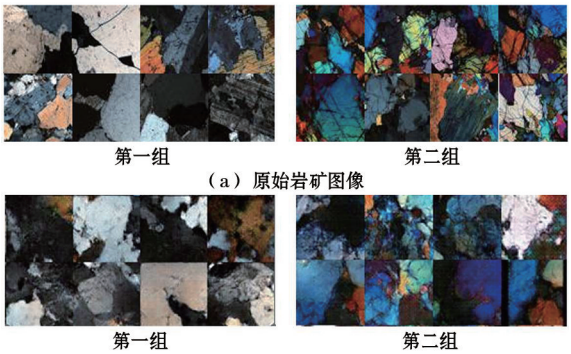


图 10　原始岩矿图像与生成岩矿图像对比

Fig.10　Comparison between original and generated rock and mineral images

为了更加科学地评估模型生成的岩矿图像的多样性，本研究引入聚类分析方法。该方法作为岩矿图像智能增量模型的创新性评价手段，可以有效量化和可视化岩矿图像的内在差异性。本次聚类分析采用了 DBSCAN(Density-Based Spatial Clustering of Applications with Noise)聚类算法，该算法在处理噪声和异常值方面具有较高的鲁棒性，尤其适用于多样性评价任务。

在试验过程中，选取了 64 组原始岩矿图像与生成岩矿图像进行 DBSCAN 聚类分析，结果如图 11 所示。

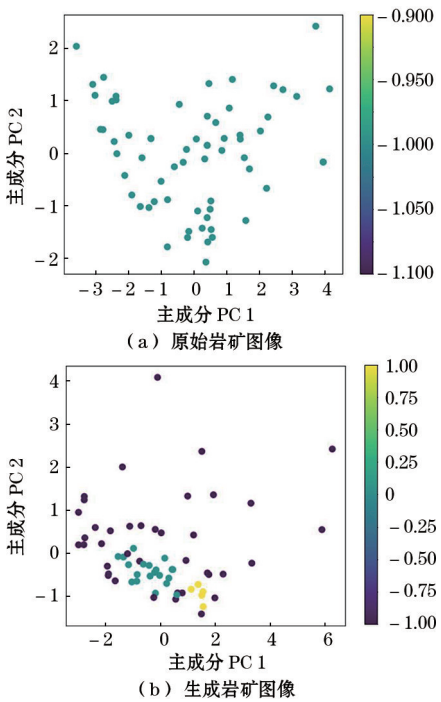


图 11　聚类分析对比

Fig.11　Comparison of cluster analysis

由图 11 可知，生成岩矿图像组在聚类过程中生成的簇数显著多于原始岩矿图像组。结果表明，生成岩矿图像样本在多样性方面具有显著优势，其复杂性和多样性均优于原始样本。

3.4　基于问卷调查的评价分析

3.4.1　问卷设计

图像生成指标是评判图像质量的重要依据，但仅凭个人的肉眼观察进行主观评价难以保证评价结果的广泛性和客观性。因此，本研究采用调查问卷的形式，对试验中生成的石英闪长岩岩矿图像从清晰度、色彩质量和细节表现等方面进行主观评价，以检验其结果的可行性。

本次问卷采用“三真一假”的形式设计，即在每组问题中混合三张真实岩矿图像与一张生成岩矿图像。问卷问题具体分为两部分：第 1，2，5，7，9 题主要考察生成岩矿图像的清晰度与细节表现；第 3，4，6，8，10 题主要针对生成岩矿图像的色彩质量。答题者需根据观察判断图像中哪一张为生成图像，问卷设计如图 12 所示。

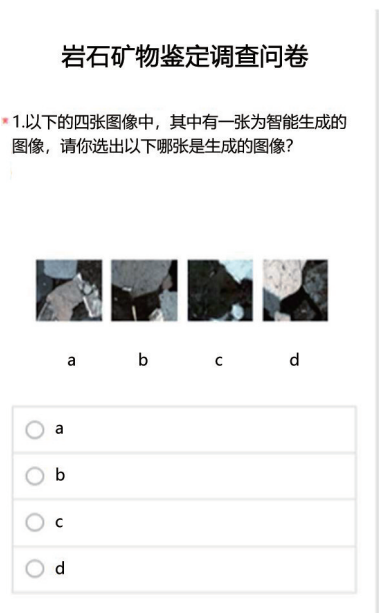


图 12　岩矿图像真假鉴别问卷

Fig.12　Questionnaire for authenticity identification of rock and mineral images

3.4.2　问卷发放与收集

问卷通过电子邮件和社交媒体平台发放给矿业工程领域的专业人士和学者，包括地质学家、矿物学家及在校学生。参与者群体的多样性得到了充分保证，从而提高了评价结果的代表性与可靠性。本次

问卷共包含 10 道题目,共收集有效答卷 213 份。

3.4.3 样本选择

问卷中使用的图像样本从生成的岩矿图像和真实岩矿图像中随机选取。为减少评价偏差,问卷采用双盲测试方法,即参与者和问卷发放者均无法知晓哪些图像为生成样本,哪些为真实样本。

3.4.4 结果分析

对收集的问卷数据进行严格的统计分析,以确保结果的客观性。评估的核心指标为问卷正确率,其中正确率越低表明模型生成的岩矿图像与真实图像的差异更小。通过该评价方法,可以获得对模型生成岩矿图像质量的主观评价结果,为模型性能的评估和优化方向确定提供依据。问卷调查结果见表 3。

表 3 问卷调查结果

Table 3 Survey questionnaire results

题目序号	正确数	正确率/%
1	75	35.21
2	64	30.05
3	49	23.00
4	60	28.17
5	78	36.62
6	51	23.94
7	70	32.86
8	46	21.60
9	64	30.05
10	43	20.19

由表 3 可知,模型的整体表现存在一定波动,在部分题目中的表现较为突出,而在其他题目中仍有提升空间。正确率最高的题目为第 5 题,达到 36.62%;其次为第 1 题和第 7 题,正确率分别为 35.21%和 32.86%。数据表明,模型在图像清晰度和细节捕捉方面具有一定的优化空间。

相比之下,第 3,4,6,8,10 题的正确率均低于 30%,表明在这些题目中,参与者难以区分生成的岩矿图像与真实岩矿图像,这一现象从侧面反映出模型生成图像的色彩质量接近真实图像。综上所述,本文构建基于改进的 DCGAN 的岩矿图像智能增量模型能够生成质量较高且具有高度真实性的岩矿图像。

4 结论

(1) 本文针对岩石矿物图像样本稀缺的问题,提出了一种基于改进深度卷积生成对抗网络(DCGAN)的岩矿图像智能增量方法。通过引入自注意力机制,提升了模型生成图像的情绪度和真实性;采用带有梯度惩罚机制的 WGAN-GP,增强了模型稳定性与收敛效率;扩展判别器和生成器的网络结构,提升了模型的整体表现力。同时,结合多种数据增强技术扩充样本多样性与数量,最终生成石英闪长岩样本 3 880 张,辉橄榄岩样本 2 208 张,有效缓解了样本稀缺问题。

(2) 试验结果显示,模型生成的岩矿图像在峰值信噪比(PSNR)、Fréchet 初始距离(FID)和结构相似性指数(SSIM)等指标上表现优异,其中 FID 指标较 GAN 模型和 DCGAN 模型分别降低了 14.5%和 5.8%。各模型损失函数曲线分析表明,引入自注意力机制和 WGAN-GP 损失函数的 Self-GP-DCGAN 模型在训练过程中表现出高度稳定性,大幅提高了模型在实际应用中的可靠性与有效性。

(3) 聚类分析结果表明,模型生成的岩矿图像在复杂性和多样性均优于原始岩矿图像,验证了改进模型在生成高质量岩矿图像方面的优势;问卷调查结果表明,模型生成图像的色彩质量与真实图像难以区分,题目最低正确率仅为 20.19%,进一步证明了模型能够生成高质量和高真实性的岩矿图像。

(4) 本研究有效缓解了岩石矿物图像样本稀缺问题,为岩石矿物的自动化识别与分类提供了数据支持。未来将优化生成器结构以生成更高分辨率图像,探索多模态数据融合技术,并推动成果在实时岩矿识别系统中的应用,以提高野外勘探效率,验证其在其他地质材料识别中的通用性。

参考文献(References):

[1] 郭艳军,周哲,林贺洵,等.基于深度学习的智能矿物识别方法研究[J].地学前缘,2020,27(5):39-47.
GUO Yanjun, ZHON Zhe, LIN Hexun, et al. The mineral intelligence identification method based on deep learning algorithms[J]. Earth Science Frontiers,2020,27(5):39-47.

[2] 朱君,张俊坤,陈尧,等.基于深度学习的显微镜下岩石薄片识别[J].现代信息科技,2022,6(24):66-68.
ZHU Jun, ZHANG Junkun, CHEN Yao, et al. Rock slice recognition under microscope based on deep learning[J]. Modern Information Technology,2022,6(24):66-68.

[3] 马泽栋,马雷,李科,等.基于岩石图像深度学习的多尺度岩性识别[J].地质科技通报,2022,41(6):316-322.
MA Zedong, MA Lei, LI Ke, et al. Multi-scale lithology recognition based on deep learning of rock images[J]. Bulletin of Geological Science and Technology,2022,41(6):316-322.

[4] NOTESCO G, OGEN Y, BEN-DOR E. Mineral classification of Makhtesh Ramon in Israel using hyperspectral longwave infrared (LWIR) remote-sensing data[J]. Remote Sensing,2015,7(9):12282-12296.

[5] 白林,魏昕,刘禹,等.基于VGG模型的岩石薄片图像识别[J].地质通报,2019,38(12):2053-2058.
BAI Lin, WEI Xin, LIU Yu, et al. Rock thin section image recognition and classification based on VGG mode[J]. Geological Bulletin of China,2019,38(12):2053-2058.

[6] 任伟,张盛,乔计花,等.基于深度学习的岩石矿物智能识别[J].地质论评,2021,67(增刊1):281-282.
REN Wei, ZHANG Sheng, QIAO Jihua, et al. The rock and mineral intelligence identification method based on deep learning[J]. Geological Review,2021,67(Suppl.1):281-282.

[7] LIU Chengzhao, LI Mingchao, ZHANG Ye, et al. An enhanced rock mineral recognition method integrating a deep learning model and clustering algorithm[J]. Minerals,2019,9(9):516.

[8] 徐圣嘉,苏程,朱孔阳,等.基于深度学习的岩石薄片矿物自动识别方法[J].浙江大学学报(理学版),2022,49(6):743-752.
XU Shengjia, SU Cheng, ZHU Kongyang, et al. Automatic identification of mineral in petrographic thin sections based on images using a deep learning method[J]. Journal of Zhejiang University (Science Edition),2022,49(6):743-752.

[9] 张利军,鲁文豪,张建东,等.基于深度学习的镜下岩石、矿物薄片识别[J].地学前缘,2024,31(3):498-510.
ZHANG Lijun, LU Wenhao, ZHANG Jiandong, et al. Rock and mineral thin section identification based on deep learning[J]Earth Science Frontiers,2024,31(3):498-510.

[10] 刘汉湖,杨武年,杨容浩.高光谱遥感岩矿识别方法对比研究[J].地质与勘探,2013,49(2):359-366.
LIU Hanhu, YANG Wunian, YANG Ronghao. A comparative study on the mineral identification methods using hyperspectral remote sensing data[J]. Geology and Exploration,2013,49(2):359-366.

[11] 杨锁荣,杨洪朝,申富饶,等.面向深度学习的图像数据增强综述[J].软件学报,2025,36(3):1390-1412.
YANG Suorong, YANG Hongchao, SHEN Furao, et al. Image data augmentation for deep learning: a survey[J]. Journal of Software,2025,36(3):1390-1412.

[12] SHORTEN C, KHOSHGOFTAAR T M. A survey on image data augmentation for deep learning[J]. Journal of Big Data,2019,6(1):1-48.

[13] WU Baokun, JI Xiaohui, HE Mingyue, et al. Mineral identification based on multi-label image classification[J]. Minerals,2022,12(11):1338.

[14] WANG Wenjian, WANG Huaying, YANG Shaokai, et al. Resolution enhancement in microscopic imaging based on generative adversarial network with unpaired data[J]. Optics Communications,2022,503:127454.

[15] OKADA N, MAEKAWA Y, OWADA N, et al. Automated identification of mineral types and grain size using hyperspectral imaging and deep learning for mineral processing[J]. Minerals,2020,10(9):809.

[16] GOODFELLOW I, POUGET-ABADIE J, MIEZA M, et al. Generative adversarial networks[J]. Communications of the ACM,2020,63(11):139-144.

[17] CHENG Jieren, YANG Yue, TANG Xiangyan, et al. Generative adversarial networks: a literature review[J]. KSII Transactions on Internet and Information Systems,2020,14(12):4625-4647.

[18] 陈波,袁志龙,陈龙,等.基于生成对抗网络的车轮踏面缺陷图像生成算法研究[J].冶金自动化,2022,46(5):56-64.
CHEN Bo, YUAN Zhilong, CHEN Long, et al. Research on image generation algorithm of wheel tread defects based on generative adversarial networks[J]. Metallurgical Industry Automation,2022,46(5):56-64.

[19] SARP S, KUZHU M, WILSON E, et al. WG2AN: synthetic wound image generation using generative adversarial network[J]. The Journal of Engineering,2021,2021(5):286-294.

[20] MENDES J, PEREIRA T, SILVA F, et al. Lung CT image synthesis using GANs[J]. Expert Systems with Applications,2023,215:119350.

[21] ONO H, SUZUKI S. Data augmentation for gross motor-activity recognition using DCGAN[C]//2020 IEEE/SICE International Symposium on System Integration (SII). New York: IEEE,2020:440-443.

[22] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks[J]. arXiv, 2015:1511.06434.

[23] HUSSAIN S, ANEES A, DAS A, et al. High-content image generation for drug discovery using generative adversarial networks[J]. Neural Networks,2020,132:353-363.

[24] 王琳,季晓慧,杨眉,等.基于数据增强和集成学习的矿物图像识别[J].地学前缘,2024,31(4):87-94.
WANG Lin, JI Xiaohui, YANG Mei, et al. Mineral identification based on data augmentation and ensemble learning[J]. Earth Science Frontiers,2024,31(4):87-94.

[25] ZHAO Gaochang, CAI Zhao, WANG Xin, et al. GAN data augmentation methods in rock classification[J]. Applied Sciences,2023,13(9):5316.

[26] ZHANG Lijun, DUAN Lixiang, HONG Xiaocui, et al. Imbalanced data enhancement method based on improved DCGAN and its application[J]. Journal of Intelligent & Fuzzy Systems,2021,41(2):3485-3498.

[27] ZHANG Han, GOODFELLOW I, METAXAS D N, et al. Self-attention generative adversarial networks[C]// International Conference on Machine Learning. New York: PMLR,2019:7354-7363.

[28] 魏玮,何飞鹏,陈海永,等.一种基于自注意力机制的自动人脸替换方法[J].计算机应用与软件,2024,41(1):198-203+239.

WEI Wei, HE Feipeng, CHEN Haiyong, et al. An automatic face replacement method based on self-attention mechanism[J]. Computer Applications and Software,2024,41(1):198-203+239.

[29] GULRAJANI I, AHMED F, ARJOVSKY M, et al. Improved training of wasserstein GANs[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems, New York: Curran Associates Inc., 2017:5769-5779.

[30] MA Tingtian, YE Wenhua, LENG Sheng, et al. Solid waste surface feature enhancement method based on gamma correction and wavelet transform[J]. Signal, Image and Video Processing,2021,15(7):1627-1634.

[31] WANG Zhou, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity[J].IEEE Transactions on Image Processing,2004, 13(4):600-612.