

高小洋,何文祥,胡勇. 基于深度生成模型的河道砂建模方法[J]. 地质科技通报, 2023, 42(2): 94-104.

Gao Xiaoyang, He Wenxiang, Hu Yong. A modeling method of channel sand based on deep generation models[J]. Bulletin of Geological Science and Technology, 2023, 42(2): 94-104.

基于深度生成模型的河道砂建模方法

高小洋^{a,b}, 何文祥^{a,b}, 胡勇^{a,b}

(长江大学 a. 资源与环境学院; b. 油气地球化学与环境湖北省重点实验室, 武汉 430100)

摘要: 在传统的河道砂体建模方法中, 基于目标建模的方法难以条件化, 多点地质统计学难以再现连续河道砂体的形态, 导致建模成果难以直接应用于油田生产。深度生成模型通过深度学习, 可以生成足够精确的河道砂体模型, 能再现复杂的河道砂体形态, 很好地满足井点条件, 弥补了传统建模算法的不足。在建模过程中, 首先基于目标模拟方法与计算机匹配操作建立了 20 000 个河道砂体模型与对应的条件集, 并结合变分自编码(VAE)与生成对抗网络(GAN)的理论, 建立深度生成模型, 其中包括分类器、编码器、解码器与判别器。将条件数据与真实模型输入深度生成模型中得到对应的河道砂体模型, 通过大量的训练建立了可以生成满足井点条件的河道砂体的生成器, 最后将井点数据输入生成器中建立相应的河道砂体模型。研究表明, 深度生成模型建模算法与传统建模算法相比不仅展现出了连续、清晰的河道砂体, 并且在给定的条件下可以生成多个河道砂体模型。该建模方法克服了传统河道砂建模方法的不足, 为河道砂体储层建模提供了新的解决思路, 建立的河道模型可为油田开发提供参考。

关键词: 人工智能; 深度生成模型; 储层建模; 河道砂体

中图分类号: P628 **文章编号:** 2096-8523(2023)02-0094-11 **收稿日期:** 2021-09-01

doi: 10.19509/j.cnki.dzkg.2022.0208

开放科学(资源服务)标识码(OSID):



A modeling method of channel sand based on deep generation models

Gao Xiaoyang^{a,b}, He Wenxiang^{a,b}, Hu Yong^{a,b}

(a. College of Resources and Environment; b. Hubei Key Laboratory of Petroleum Geochemistry and Environment, Yangtze University, Wuhan 430100, China)

Abstract: In traditional channel sand body modeling methods, the method based on target modeling is difficult to condition, and multipoint geostatistics have difficulty reproducing the shape of continuous channel sand bodies, which makes the modeling results difficult to directly apply to oilfield production. Through depth learning, the depth generation model can generate a sufficiently accurate channel sand body model, reproduce the complex channel sand body shape, meet the well point conditions, and compensate for the shortcomings of the traditional modeling algorithm. In the process of modeling, 20 000 channel sand body models and corresponding condition sets are established based on the target simulation method and computer matching operation. Combined with the theory of variational autoencoders (VAE) and generative adversarial networks (GAN), a depth generation model is established, including a classifier, encoder, decoder and discriminator. The condition data and the real model are input into the depth generation model to obtain the corresponding channel sand body model. Through extensive training, a generator that can gen-

作者简介: 高小洋(1995—), 男, 现正攻读矿产普查与勘探专业博士学位, 主要从事地球化学和人工智能地质建模研究工作。
E-mail: 18734311896@163.com

通信作者: 胡勇(1980—), 男, 副教授, 主要从事油藏沉积学和地质建模研究工作。E-mail: 64421847@qq.com

erate channel sand bodies that meet the well point conditions is established. Finally, the well point data are input into the generator to establish the corresponding channel sand body model. The results show that compared with the traditional modeling algorithm, the depth generation model modeling algorithm not only shows a continuous and clear channel sand body but can also generate multiple channel sand body models under given conditions. This modeling method overcomes the shortcomings of traditional channel sand modeling methods and provides a new solution for channel sand reservoir modeling. The established channel model can provide a reference for the oilfield development stage.

Key words: artificial intelligence; deep generation model; reservoir modeling; channel sand body

河道砂体是陆相环境中常见的油气储集体,这类砂体是我国已发现油田主要的储层类型,也是我国油气最富集的储油层。在我国东部各油田中,河流相储层在各类碎屑岩储油层中占有重要比例。据统计,河流相储层原油地质储量占我国已开发油田动用地质储量的46.2%^[1];河流相储层地质储量采出程度为23%左右,大约仍有近3/4的地质储量滞留在储层内。因此对河道砂体开展精细研究,准确预测河道砂体的形态及分布,建立高精度的河道砂体模型,具有重要的意义^[2-9]。

在河道砂体的建模研究中,以基于目标的模拟方法(OBJM)和多点地质统计学建模(MPS)最常见。OBJM主要根据预先确定的目标体形态进行模拟^[10-12],但OBJM在模拟过程中使用多次迭代来匹配井数据,并且通常不能满足所有井数据。MPS采用训练图像作为主要输入参数,比OBJM更尊重井数据^[13-16],但对训练图像要求较高,如要求训练图像具有平稳性,实际情况却很少如此,导致模拟的河道出现间断的现象^[17-18]。

目前,人工智能已经应用到各行各业,并且取得了突破性的进展,在地质勘探研究领域也得到了很好的应用^[19-23]。深度学习是人工智能的一个重要的研究方向,通过建立深层神经网络来进行分析、学习和解释数据^[24]。其中图像的生成技术是应用相对较广的一项技术^[25-28],主要通过深度生成模型理论生成图像。目前主要有2个分支,分别为变分自编码器(variational auto-encoders,简称VAE)与生成对抗网络(generative adversarial networks,简称GAN)。其中VAE由Kingma等^[29]于2014年提出的基于变分贝叶斯(variational bayes,简称VB)推断的生成式网络结构,它以概率的方式描述对潜在空间的观察,在数据生成方面表现出了巨大的应用价值。而GAN是2014年由Goodfellow等^[30]提出,通过对抗训练的方式使生成网络产生的样本服从真实数据的分布。VAE和GAN被视为无监督式学习领域最具研究价值的方法之一。

随着深度生成模型理论的发展,许多学者将VAE与GAN理论应用于地质建模领域,例如:

Canchumuni等^[31]使用卷积变分自动编码器和多重数据同化的集合平滑器对河道砂体进行建模,生成了连续的河道砂体,但由于自编码器需要高斯分布等概率模型假设,导致生成的模型精度低且条件化难度大。Laloy等^[32]基于GAN理论提出了一种空间生成对抗网络的复杂地质介质反演方法,通过定义一个低维参数,使用马尔可夫链蒙特卡罗(MC-MC)方法进行有效的概率反演,最终可以快速生成二维和三维无条件模型;Zhang等^[33]将井数据作为条件,基于GAN理论建立复杂的地质模型,该方法能够捕获更真实的相结构,但由于GAN在训练时常常出现模型崩溃、过拟合、无法收敛等问题,训练难度较大,导致生成的模型质量较差,无法满足实际的河道砂体特征。

为了模拟陆相曲流河河道砂体展布,结合VAE与GAN理论,以模拟的井点数据为条件,通过建立的生成器生成河道砂体模型,解决传统建模算法的缺陷,生成的模型满足给定的井点条件,可以生成连续的河道砂体,满足油田的实际开发需求。

1 研究方法

在研究过程中,无论用VAE还是GAN建立目标河道砂体模型,总会出现生成的模型质量较差、无法生成多个满足井点条件的河道砂体模型,甚至所生成的砂体不连续等问题,导致建立的河道砂体模型精度较低,无法满足实际的地质特征。当用VAE建模时,生成的河道砂体模型较为模糊,但可以清晰地看到2条河道,河道的形态、方向与数据集中的河道砂体模型相同,并且可以生成多个不同的储层模型;当用GAN建模时总会生成相同的模型,没有多样性,但生成的模型清晰,对河道砂体的形态描述较为精确。本研究结合VAE与GAN理论进行砂体建模,以期得到合理的砂体储层模型。

1.1 变分自编码模拟

假设一个生成模型(图1)中包含隐变量,即有部分变量是不可观测的,其中隐向量 z 是一个低维空间 Z 中的随机向量,而 x 是一个高维空间 X 中的

$$V^*(\Phi, \theta | x) = -\frac{1}{2} \|x - \mu_G\|^2 + \lambda \text{KL}(N(\mu_I, \sigma_I), N(\mu_y, I)) \quad (14)$$

式中： μ_y 为井点条件在隐空间上的映射，即变分分布 $q(z|y; \varphi)$ (假设为高斯分布) 的均值，通过另一个推断网络 E_y 来计算， φ 为该网络的参数。

1.2 生成对抗网络模拟

为了使生成的河道砂体模型更加精确，引入 GAN 的思想，它是一种无监督学习方法，通过对抗训练的方式使生成网络产生的样本服从真实数据的分布。首先建立一个生成网络，它以随机向量(潜在空间中的随机点 z) 作为输入，并将其解码为一个模型 $G(z)$ 。然后建立一个判别网络，以河道砂体模型 x 作为输入，判断该模型是来自训练集还是由生成网络所创建，并输出 x 为真实模型的概率，若输出为 1 则代表判断它为真实模型，若为 0 则判断它为“赝品”(图 2)。

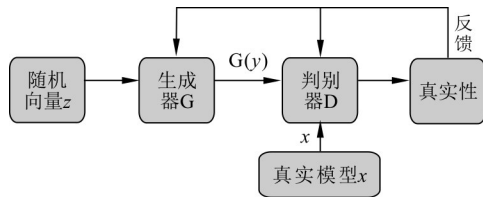


图 2 GAN 结构图

Fig. 2 GAN structural diagram

通过大量的迭代及优化，生成器 G 可以生成越来越真实的模型 $G(z)$ 使判别器 D 判别不出 $G(z)$ 是由生成器 G 生成的还是真实模型 x 。而判别器 D 可以不断地提高判别能力来区分 $G(z)$ 与真实模型 x 。通过这种对抗训练最终可以生成真实的河道砂体模型，在理想状态下 $D(G(z)) = 0.5$ 时，GAN 训练结束。

GAN 的判别器目标函数如下：

$$L_D = \max_w \{E_{x \sim P_{\text{data}}(x)} [\ln D(x)] + E_{z \sim P_z(z)} [\ln (1 - D(G(z)))]\} \quad (15)$$

式中： x 为真实模型，分布为 $P_{\text{data}}(x)$ ，是未知的。 $E_{x \sim P_{\text{data}}(\cdot)}$ 与 $E_{z \sim P_z(\cdot)}$ 为判别器分类概率的数学期望，其中 $E_{x \sim P_{\text{data}}(x)}$ 是直接取在真实模型数据集 x 中取得的真实样本，而 $E_{z \sim P_z(z)}$ 是从已知的输入分布 z 中取得的样本。 $D(x)$ 表示判别器 D 判断真实模型 x 是否真实的概率(该值不断地趋近于 1)。 $D(G(z))$ 是判别器 D 判断 G 生成的模型真实的概率。随着训练的进行， $D(x)$ 会逐渐增大， $D(G(z))$ 会逐渐变小，所以 L_D 会逐渐增大。

生成器 G 尝试去最小化： $L'_{GD} = -E_{z \sim P_z(z)} [\ln (D(G(z)))]$ 。但在实际训练时，真实模型和生成的

模型可能互不相关，经常出现梯度消失的现象^[34]。所以，本研究的生成器使用平均特征匹配目标函数来解决该问题：

$$L_{GD} = \frac{1}{2} \|E_{x \sim P_{\text{data}}(x)} f_D(x) - E_{z \sim P_z} f_D(G(z))\|^2 \quad (16)$$

该目标函数需要合成样本的特征中心去匹配真实样本的特征中心。 $f_D(x)$ 表示判别器中间层的特征，生成器 G 的目标是最小化该函数。

1.3 模型优化

由于 GAN 训练不稳定，总会出现模型崩溃的现象，而 VAE 训练后所得到的模型很模糊，难以达到预期效果。所以本研究在 GAN 的基础上引入 VAE 理论，训练过程如图 3。

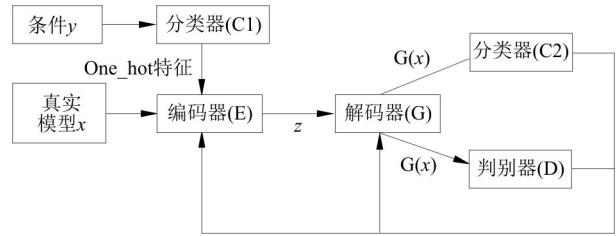


图 3 深度生成模型训练过程

Fig. 3 Training process of the deep generative model

所用的条件数据 y 是随机模拟的井点位置图，直接输入网络所生成的河道砂体模型效果较差，在实验过程中损失函数经常出现难以收敛的现象。所以，在训练之前建立了一个分类器 $C1$ 对条件数据进行分类，目的是将条件数据映射为对应的 one_hot 的形式(条件设定的具体方式请参考 3.1 节)，相当于将数据 y 进行压缩分类。分类器的目标函数为：

$$V_C = -\frac{1}{n} \sum_{n=1}^n L_n \times \log(P_{n \sim A'}) \quad (17)$$

式中： n 为样本个数； L_n 为第 n 个样本的 one_hot 特征； $P_{n \sim A'}$ 为第 n 个样本通过神经网络得到的概率 $P_{A'}$ (与 one_hot 维度相同)。

将真实模型 x 输入编码器 E 得到式(14)中的 μ_I, σ_I ，相当于变分分布 $q(z|x; \Phi)$ 对应的均值与方差，通过式(12)再参数化生成隐向量 z' ；将 one_hot 特征输入另一个编码器 E_y (本研究设定了 2 个编码器，在图 3 中总称为编码器 E)，输出一个 100 维向量，并且得到式(14)中的 μ_y 。将隐向量 z' 与该 100 维向量合并为新的隐向量 z ，输入解码器得到生成模型 $G(x)$ ，分别输入判别器 D 与分类器 $C2$ 中，其中判别器 D 使用 GAN 模型判别 $G(x)$ 的真实性，分类器 $C2$ 是提前用真实模型训练好的分类器(结构与 $C1$ 相同，目标函数为式(17))，通过反馈 $G(x)$ 属于对应类别的概率(V_C)来优化深度生成模型的参数。当判别器判

别不出 $G(x)$ 的真实性以及分类器 C2 可以正确地区分 $G(x)$ 所属的类别时,模型训练结束。

最终目标函数为:

$$L = V^*(\Phi, \theta | x) + L_{GD} + V_C + L_D \quad (18)$$

式中: V_C 为式(17)计算, V_C 中的参数 $P_{n \sim A'}$ 通过训练好的分类器 C2 获得。

将目标井点数据输入分类器 C1 得到对应的 one_hot 特征,然后通过编码器 E_y 得到一个 100 维的向量,在正态分布中随机抽取 n 个分位数,分别与该向量进行拼接得到隐向量 z ,输入解码器中得到多个满足井点条件的河道砂体模型。

2 河道砂体建模过程

2.1 数据准备及数据处理

深度生成模型的关键在于需要大量的训练数据,数据集越丰富,生成的模型越真实可靠。本研究设定的河道平均宽度为 500 m,河道的弯曲度约为 1.2,由此确定了河道的基本参数。采用 Petrel 软件基于目标的模拟方法随机生成了 2 000 个河道砂体的模型(图 4),以 2 条完整的河道为例,流向为东西向。通过对数据集的平移、缩放等数据集增强操作,将数据集扩大到 50 000 张。

通过随机模拟建立了 500 个井点图(图 5),其中黄色的点代表钻遇河道砂体的井,灰色的点代表钻遇河道间的井,黄色的河道砂数据是模型中主要的条件。用每一张井点图在建立的 50 000 个数据

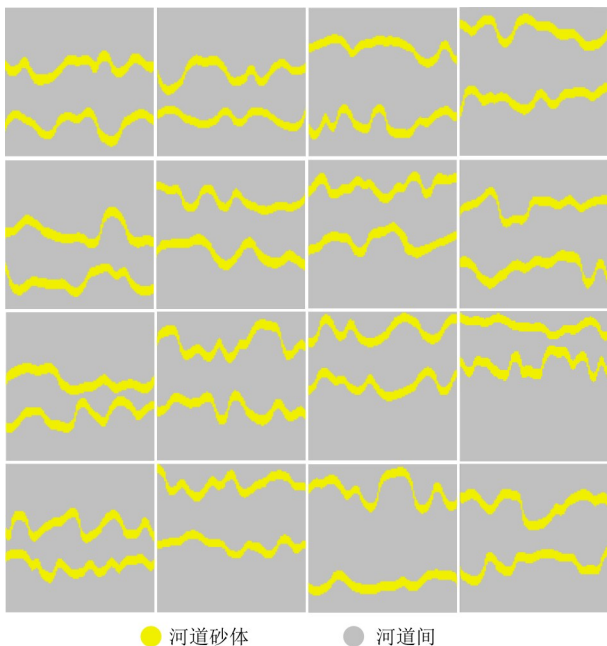


图 4 随机生成的河道砂体模型数据集(此处只显示了 16 个模型的数据)

Fig. 4 Randomly generated data set of channel sand body model

集中进行匹配,匹配时以图 5 中的点为中心设定一个矩形范围(设定为 10×10),该矩形必须与对应的河道砂体或河道间相匹配,即当井点是河道砂体时该矩形全部为黄色,当井点是河道间时全部为灰色,对之后的目标井点图(图 6)分类时增大正确分类概率。每一张井点图在 50 000 个数据集中挑选所有满足条件的砂体模型,即黄色的点全部落在河道砂体上,灰色的井点落在河道间上。通过计算机匹配以及人为操作,每一张井点图匹配了大约 20~50 张砂体模型,得到 20 000 个数据集。也就是说,将 20 000 个数据集分为了 500 个类别,每一个类别对应一个井点图。对每张井点图在原先设定的矩形范围内随机模拟井点位置,使每一个河道砂体模型可以对应不同的井位条件。

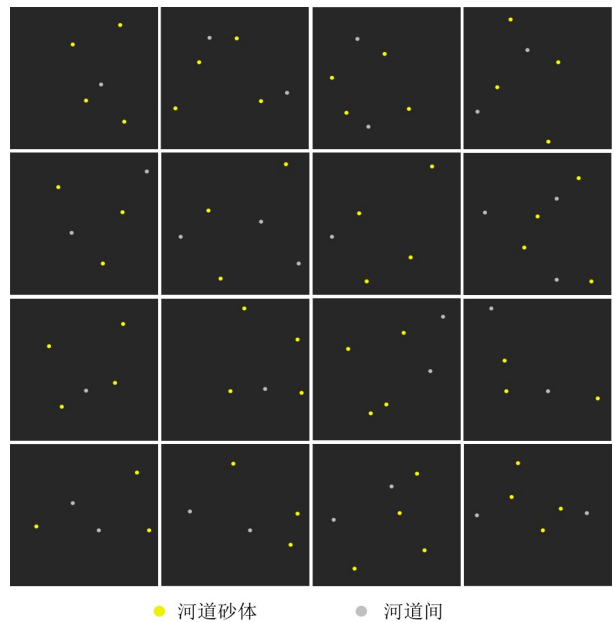


图 5 输入训练的条件集

Fig. 5 Condition set of input training

图 6 为验证过程中用的井点条件数据 y ,以 3 种随机井位进行验证,分别模拟了 10 口井,该数据不参与训练,同样也不在条件集中。本研究希望通过建立的深度生成模型模拟出来的河道能很好地满足井点条件,即模拟的河道能很好地覆盖井点数据是河道砂体的区域。

2.2 分类器的建立

通过实验发现,直接在深度生成模型中使用井点条件数据,会导致模型训练不稳定,难以收敛等问题,所以将井点条件数据转换为 one_hot 特征,即建立一个 500 阶的单位矩阵,每一个井点条件数据对应单位矩阵中的一个单位向量,为所有的河道模型加一个 one_hot 的标签,井点条件数据集与河道模型数据集将会以一个 one_hot 的标签一一对应,如图 7。

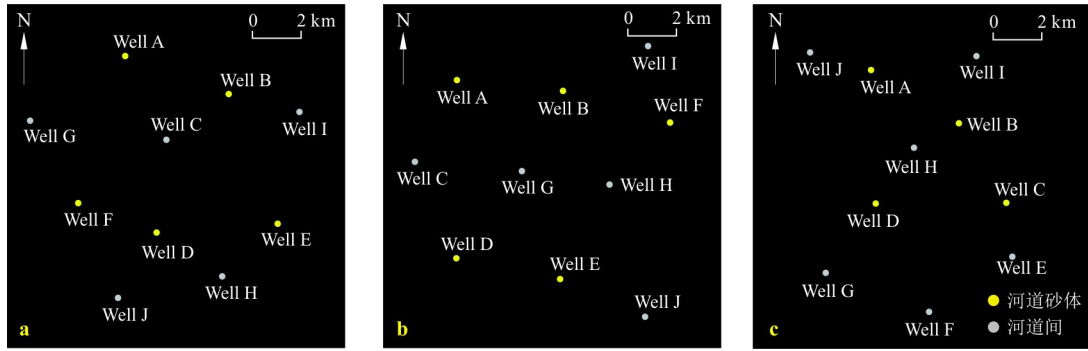


图 6 用于测试的井点图

Fig. 6 Well point diagram for testing

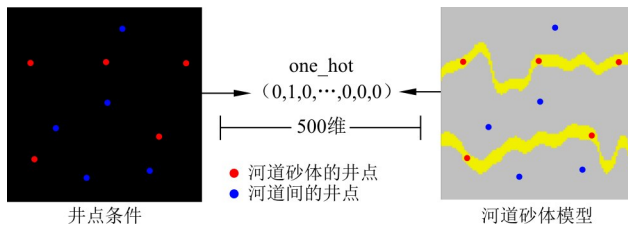


图 7 条件数据与河道模型通过 one_hot 特征进行连接

Fig. 7 Conditional data and channel model are connected through one_hot features

首先基于卷积神经网络建立分类器 C2 对河道模型数据集进行分类, 设定的输入模型的大小为 $256 \times 256 \times 3$, 设定了 6 个卷积层将河道砂体模型映射为 $4 \times 4 \times 1024$ 的数组 A, 其中卷积核的大小为 3×3 , 步长(strides)为 2。通过一个全连接层将数组 A 映射为一个 500 维的向量 A' 。通过 softmax 函数计算向量 A' 中每个元素的概率 $P_{A'}$ 。分类器 C1 主要对条件模型进行分类, 结构与分类器 C2 相同。

经前人研究, Adma 算法优化的效果较传统梯度下降算法要好^[35], 故采用 Adma 优化器来优化模型参数。

2.3 深度生成模型的建立与训练

对于编码器的建立, 建立了 2 个编码器, 其一是对河道砂体模型数据集 x 的编码器 E, 其二是对

于相应的 one_hot 的编码器 E_y 。编码器 E 采用卷积神经网络来建立, 与分类器的结构相同, 通过卷积运算将河道砂体模型映射为 $4 \times 4 \times 1024$ 的数组, 不同的是在卷积完成之后建立了 2 个全连接层分别输出 μ_I, σ_I , 将它们映射为一个 100 维的向量, 通过式 (12) 再参数化得到隐向量 z' 。编码器 E_y 是通过一个全连接层将 one_hot 特征直接映射为 100 维的向量, 再通过一个全连接层输出 μ_y , 将该 100 维的向量与 z' 拼接得到新的隐向量 z 。

解码器 G 是编码器 E 的逆过程, 不同的是在每一个卷积层中加入了反最大池化操作来增强生成器的能力。首先建立一个全连接层将隐向量 z 映射为 $4 \times 4 \times 1024$ 维的数组, 通过反卷积运算(反卷积层与编码器 E 的卷积层层数相同)将隐向量再映射为与原始模型大小一样的数组 ($256 \times 256 \times 3$) (图 8)。

判别器采用 GAN 的判别器。同样建立一个与编码器的卷积层层数相同的卷积网络将解码器 G 生成的模型 $G(z)$ 映射为 $4 \times 4 \times 1024$ 的数组, 然后通过一个全连接层映射为一个数字 t , 通过 sigmoid 函数将 t 映射到 0 到 1 的区间上, 用来计算损失函数。

在模型训练时, 首先需要训练分类器 C1 与分类器 C2, 通过交叉验证的方法进行训练(将数据集分为 7 份, 其中 6 份为训练数据, 1 份为测试数据 (500 个)), 当损失函数(式 (17))降低到一个基本恒

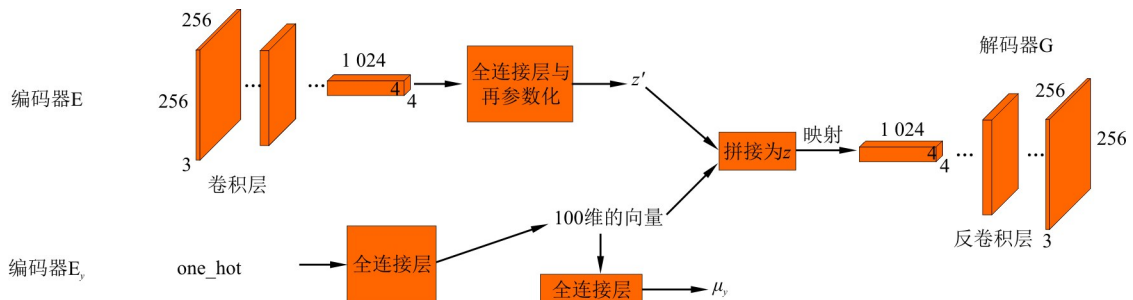


图 8 编码器 E 与解码器 G 的结构图

Fig. 8 Structural diagram of Encoder E and Decoder G

定的值时,即当继续训练时损失函数平均值的变化小于 0.01 时,并且训练精度与测试精度不会有所升高时,分类器训练完成。

将河道砂体模型数据集 x 与对应的 one_hot 特征分别输入到编码器 E 与编码器 E_y 中,通过再参数化与拼接操作,得到隐向量 z ,输入解码器生成河道砂体模型 $G(z)$ 。将原始河道砂体模型 x 与生成的河道砂体模型 $G(z)$ 输入判别器 D 判断 x 与 $G(z)$ 的真实概率 $D(x)$ 与 $D(G(z))$,将 $G(z)$ 输入分类器 C2 中判别 $G(z)$ 满足对应的 one_hot 条件的概率 V_{C2} 。通过不断地迭代优化模型参数。训练时通过优化 $V^*(\Phi, \theta | x), L_{GD}, L_y$ 和 V_C 来调节编码器 E 以及解码器 G 的参数,即 $L_G = V^*(\Phi, \theta | x) + L_{GD} + V_C$;通过优化 L_D 来调节判别器 D 的参数。在理想状态下当判别器判别不出 $G(z)$ 的真实性以及 L_G 降低到一个恒定的值时,深度生成模型训练完成。

将测试数据 y (图 6) 输入到判别器 C1 中输出 one_hot 特征,然后输入到编码器 E_y 中得到一个 100 维的向量,在正态分布中随机抽取 n 个 100 维的分位数,与该向量拼接得到隐向量 z ,输入到解码器中得到多个满足井点条件的河道砂体模型。

3 结果与讨论

结合 VAE 与 GAN 的理论,建立了一个深度生成网络,以大量训练集为基础,经过多层的卷积运算以及大量迭代,建立了一个可以生成较为精确的河道砂体储层模型的生成器。

其中分类器训练了 1 000 个 epochs,大约迭代了 29.6 万次得到比较精确的分类效果(图 9)。图 9-a 为分类器 C1 的损失函数变化曲线,训练大约 900 个 epochs 之后趋于稳定,最终收敛于 0.01 左右,通过统计分类精度达到了 98%,可以基本正确地将井点数据分类到对应的 one_hot 特征下。图 9-b 为分类器 C2 的损失函数变化曲线,目标函数在训练 700 个 epochs 之后趋于稳定,最终降到了 0.1 左右,通过统计分类精度达到 95% 左右,已经基本可以将河道砂体模型数据集进行正确的分类。

深度生成模型(包括编码器 E、解码器 G、判别器以及 V_C) 训练时,以 500 个数据为测试数据,20 000 个数据为训练数据,训练了 1 000 个迭代次数,大约迭代了 31.2 万次,损失函数变化如图 10 所示。

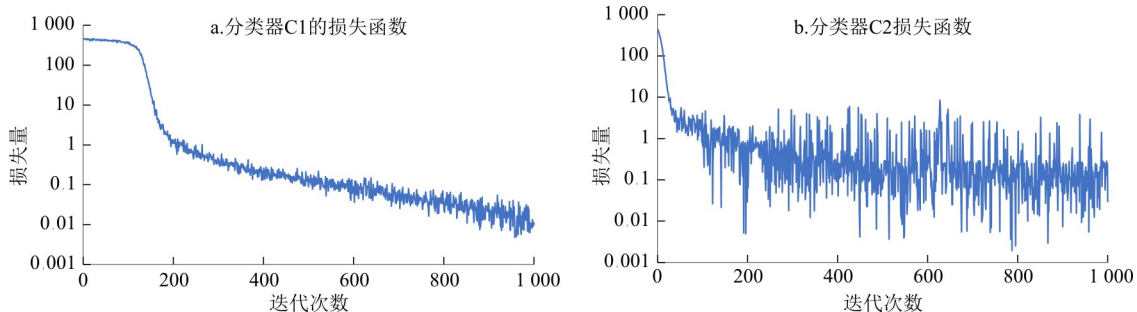


图 9 分类器 C1、C2 损失函数变化曲线

Fig. 9 Variation curve of loss function of the classifier C1, C2

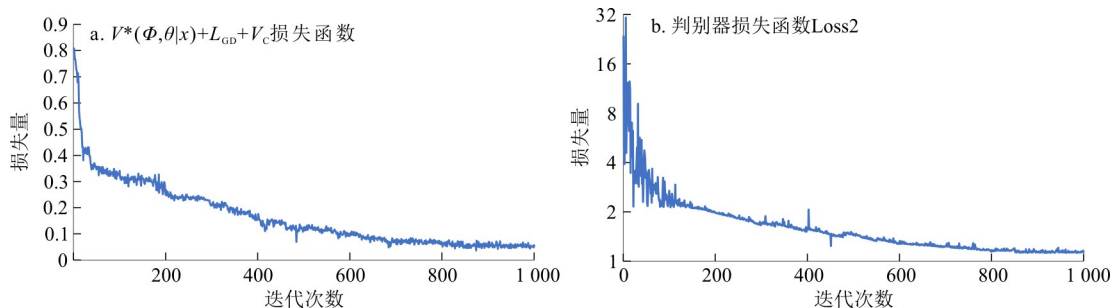


图 10 深度生成模型损失函数

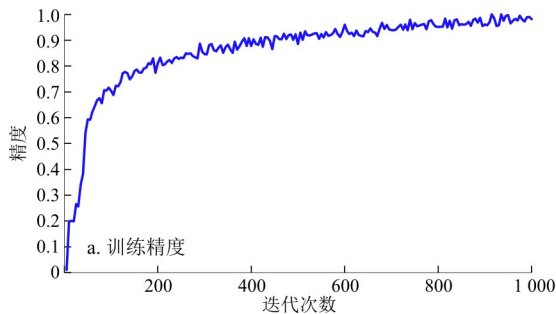
Fig. 10 Loss function of the deep generative model

将编码器 E 和解码器 G 的参数一起优化,所以 $Loss1 = V^*(\Phi, \theta | x) + L_{GD} + V_C$ 。由图 10 可知,当训练了 800 个迭代次数之后,Loss1 基本收敛,最终收敛于 0.05 左右。Loss2 训练 800 个迭代次数之

后,损失函数基本收敛,最终收敛于 1.2 左右。

为了验证模型生成的质量,统计了训练过程中生成的模型与条件数据的对应关系。将生成的模型与对应的井点条件做匹配,即在对应的井点位置生

成对应的沉积相则匹配成功,通过计算匹配成功的概率反映所建模型的精度。如图 11 所示,训练精度



在 800 个迭代次数之后收敛于 0.97 左右,测试精度训练 800 个迭代次数之后收敛于 0.95 左右,证明生

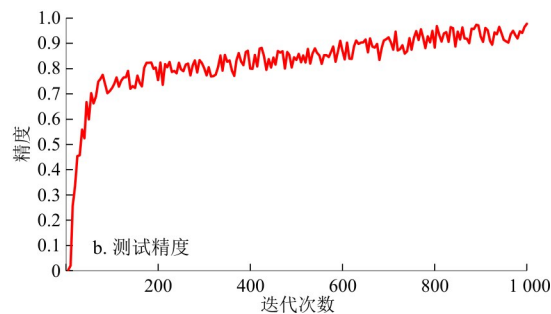


图 11 训练精度(a)和测试精度(b)变化曲线

Fig. 11 Change curve of training accuracy (a) and testing accuracy (b)

成的模型可以满足对应的井点条件。

模型训练完成后输入条件数据 y (图 6) 到分类器 C1 中得到 one_hot 特征,输入到编码器 E_y 中得到 100 维的向量,从正态分布中分别抽取 3 组维度为 100 的随机数组,与该向量拼接得到隐向量 z ,输入训练好的解码器中分别得到 3 个河道模型,如图 12 所示。

生成的所有河道砂体模型都满足给定的井点特征,在对应的井点处生成了对应的沉积相,得到了清晰的 2 条河道,并且没有断开的现象。通过不断地从正态分布中抽取不同的分位数就可以得到不同的满足井点条件的模型。

为了明确对生成的模型的影响因素以及对比传统建模算法的优势,展开以下 3 方面的讨论。

(1) 训练迭代次数的数量

在训练初期由于无法确定训练的次数,故在训练初期预先给定一个较大的值进行测试,本研究训练初期设定为 5 000 个迭代次数,根据损失函数、训

练精度与测试精度的变化确定深度生成模型是否训练完成,并判断模型是否出现过拟合(即训练精度较高,但测试精度较低)。通过调整网络参数并使用第 2.3 节所提出的优化方案解决过拟合问题。由图 11 可知,测试精度与预测精度较高,模型未出现过拟合现象。由图 10 可知,当训练 800 个迭代次数之后,损失函数基本收敛,模型训练完成,在训练至 1 000 个迭代次数时损失函数基本变化不大,故本节将迭代次数设定为 300,500,700,1 000 个迭代次数进行讨论。

在训练过程中,不同的迭代次数得到的效果不同,如图 13 所示。

训练的次数越多,生成的河道砂体模型越完整,细节越丰富。在训练 300 个迭代次数之后所生成的模型(图 13-a),砂体不连续,看不出河道的特征,仅

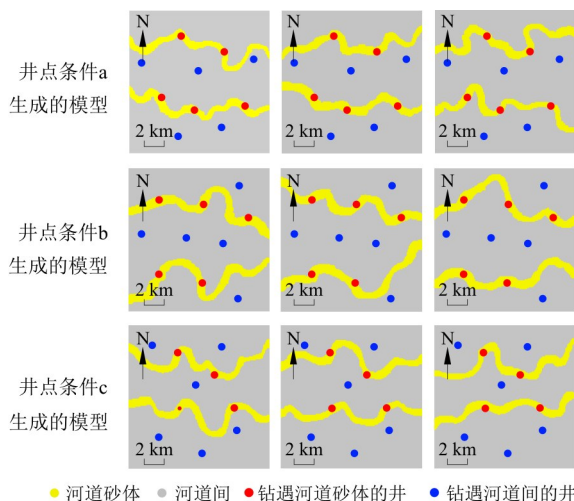
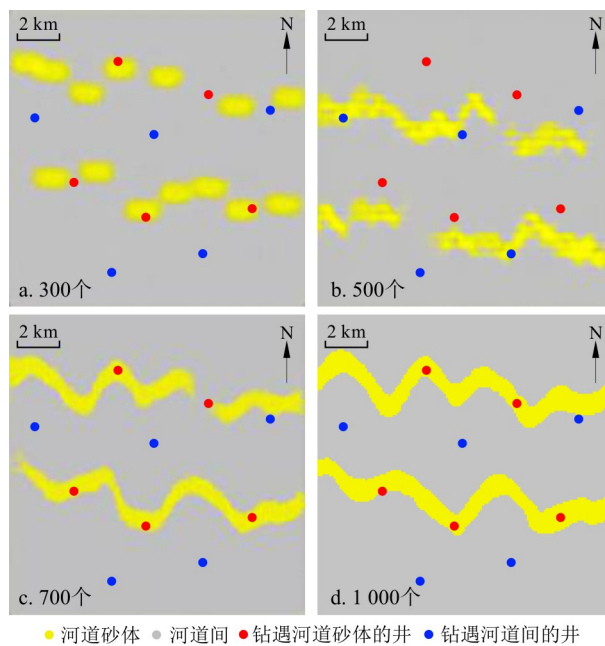


图 12 生成的模型

Fig. 12 Generated model



● 河道砂体 ● 河道间 ● 钻遇河道砂体的井 ● 钻遇河道间的井

图 13 训练不同的迭代次数生成的河道砂体模型

Fig. 13 Training different numbers of epochs to generate the channel sand body model

生成了一些零星的砂体;训练 500 个迭代次数后(图 13-b),砂体逐渐连续,但依然存在断开的现象,能看到河道的基本形态;训练 700 个迭代次数后(图 13-c),河道砂体整体变得连续,但是下部河道左端以及上部河道中部仍有部分缺失;训练 1 000 个迭代次数之后(图 13-d),砂体连续,河道没有缺失,2 条完整的河道被清晰地展现出来,生成了较好的河道砂体模型,并且完全满足井点条件。

深度生成模型需要对数据样本进行大量的矩阵运算,GPU(显卡)因为其处理矩阵计算的高效性被广泛运用到深度学习之中,在训练过程中采用英伟达的 GeForce RTX 2060 型号 GPU,该 GPU 有 2 170 个流处理器单元,训练的模型大小为 $256 \times 256 \times 3$ 。根据本研究讨论的参数,分类器训练速度较快,训练 20 000 个模型,训练 1 000 个迭代次数需耗时 1 h 左右;深度生成模型训练 20 000 个模型、训练 1 000 个迭代次数所花费的时间为 40 h。当模型训练完成后,生成一个模型只需要 0.07 s。

(2) VAE 与本研究的深度生成模型对比

使用 VAE 方法对数据集进行训练并得到了图 14 的效果。虽然生成 2 条清晰的河道,但是生成的

模型较模糊,边界不清晰,只能看出河道大概的走向。而用本研究所建立的深度生成模型得到的河道砂体模型(图 12),不仅生成 2 条清晰的河道,而且边界刻画清晰,与图 14 相比,对于河道砂体的刻画更加的精确,效果更好。

(3) 对比传统建模算法

本研究以基于目标模拟与多点地质统计 2 种传统建模方法,以图 6-a 的模拟井数据为例建立对应的河道模型,如图 15,16 所示。基于目标模拟所生

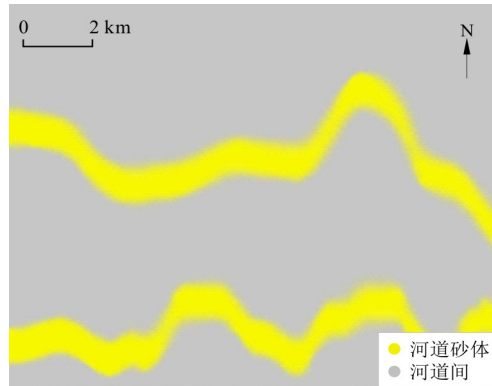


图 14 VAE 生成的河道砂体模型

Fig. 14 Channel sand body model generated by the VAE

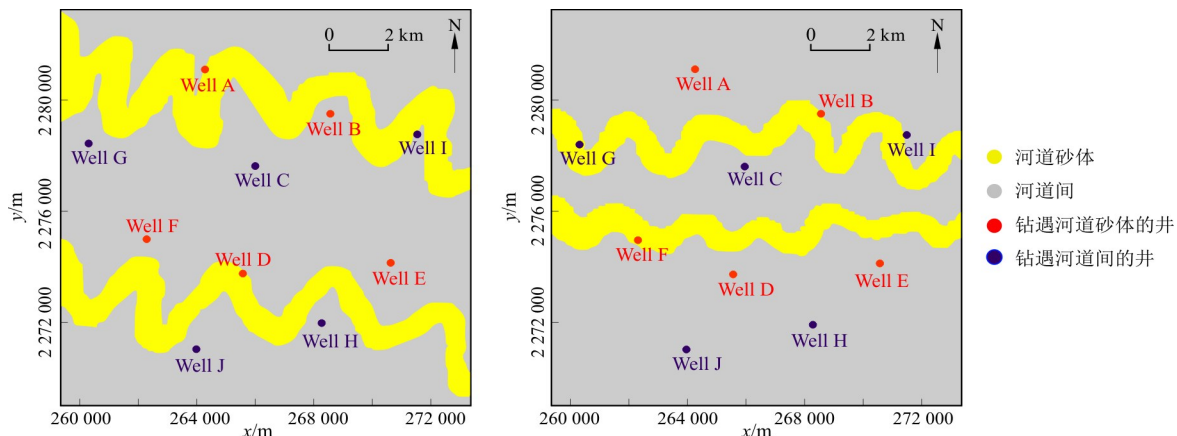


图 15 基于目标模拟生成的河道砂体模型

Fig. 15 Channel model generated based on target simulation

成的模型无法满足给定的井点条件(图 15),生成的河道不能全部覆盖条件为砂岩的井,不满足设定的条件;多点地质统计所生成的河道可以覆盖条件为砂岩的井点,但河道不连续,分布杂乱,不满足实际的沉积特征。

综上所述,传统建模算法生成的模型不满足给定的条件,生成的河道太过随机,采用的深度生成模型克服了传统建模算法的缺点,生成的河道不仅连续,而且满足给定的井点条件(图 12),生成的模型更精确。

4 结 论

(1)河道砂体为中国东部陆相油气田的主要储集砂体,随着大部分油气田进入开发后期,需要对河

道砂体储层精细刻画以明确储层砂体分布特征。储层建模是广泛应用的储层表征技术,很多油田都开展了河道砂体建模的研究,但传统的建模算法如基于目标的模拟和多点地质统计学模拟经过多年的发展,仍存在不足导致模拟结果不能直接应用于生产。变分自编码(VAE)与生成对抗网络(GAN)等深度生成模型的出现,为河道砂体建模提供了新的思路。

(2)通过传统的建模算法—基于目标的模拟建立了 2 000 个河道砂体模型,基于 Tensorflow 的深度学习框架建立了可以生成河道砂体储层模型的解码器以及针对井点条件的分类器与编码器。研究表明,训练的迭代次数越多,获得的模型特征数据库越

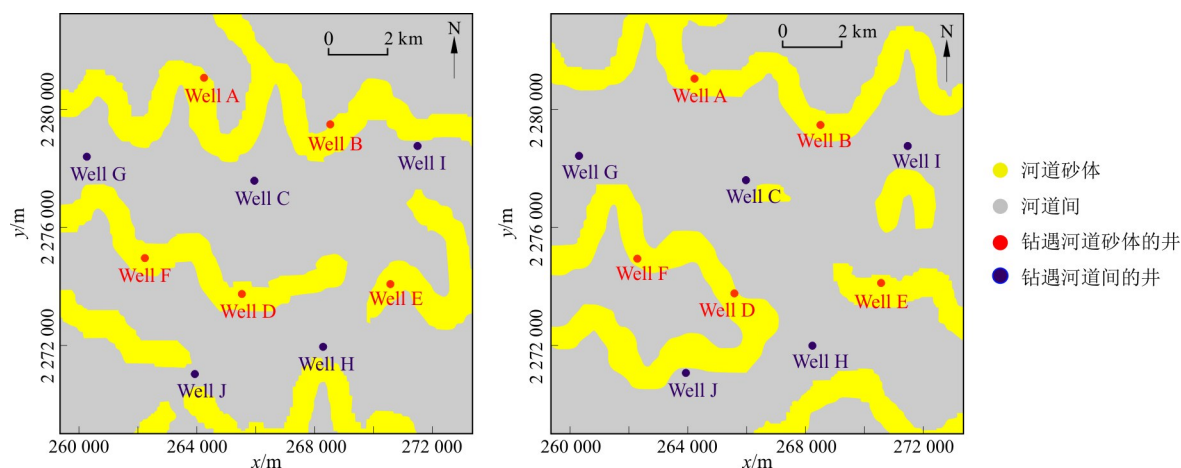


图 16 多点统计生成的河道砂体模型

Fig. 16 Channel sand body model generated by multipoint statistics

大,包含的河道细节越丰富,当训练了足够的迭代次数,并且损失函数、训练精度、预测精度都不会有太大的变化时,深度生成模型训练完成,生成的模型越能反映输入模型的形态特征。

(3)该算法克服了传统建模方法的缺点,与多点地质统计学相比,基于深度生成模型的建模方法在数据集建立方面,对砂体的空间位置和相互位置关系没有严格的要求,在对数据集的准备上相较于训练图像限制要更少,所生成的河道更连续;与基于目标模拟的建模算法相比,生成的河道满足设定的条件,生成的模型更具说服力。

(4)结合 VAE 与 GAN 的理论,以井点条件为限制生成对应的河道砂体模型,克服了其他建模方法条件化困难等问题。建立的模型很好地体现目标区的河道砂体分布特征,为储层建模的研究提供了另一个方向,更好地服务于油田生产。

参考文献:

- [1] 束青林. 河道砂储层油藏动态模型和剩余油预测[M]. 北京:石油工业出版社,2004.
Shu Q L. Reservoir performance model and remaining oil prediction of channel sand reservoir[M]. Beijing: Petroleum Industry Press,2004(in Chinese).
- [2] 李勤,张利锋,孙丽. 遗传模拟退火算法在储层属性建模中的应用[J]. 大庆石油地质与开发,2005(1):31-32,106.
Li Q, Zhang L F, Sun L. Influence of asphaltene deposition on oil seepage characteristics[J]. Petroleum Geology & Oilfield Development in Daqing, 2005(1): 31-32, 106 (in Chinese with English abstract).
- [3] 陈建阳,于兴河,张志杰,等. 储层地质建模在油藏描述中的应用[J]. 大庆石油地质与开发,2005(3):17-18,104.
Chen J Y, Yu X H, Zhang Z J, et al. Application of reservoir modeling in reservoir description of Baolige Oil Field[J]. Petroleum Geology & Oilfield Development in Daqing, 2005(3): 17-18, 104 (in Chinese with English abstract).
- [4] 徐伟,房磊,刘钧,等. 浅层高渗砂岩油藏分级相控地质建模

[J]. 地质科技情报,2017,36(1):197-201.

Xu W, Fang L, Liu J, et al. Hierarchic facies-constrained geological modeling of shallow high permeability sandy reservoir [J]. Geological Science and Technology Information, 2017, 36 (1): 197-201 (in Chinese with English abstract).

- [5] 付志国,石成方,赵翰卿,等. 喇萨杏油田河道砂岩厚油层夹层分布特征[J]. 大庆石油地质与开发,2007(4):55-58.

Fu Z G, Shi C F, Zhao H Q, et al. The distribution characteristics of interlayer in thick channel sand oil reservoir in Lasaxing Oilfield [J]. Petroleum Geology & Oilfield Development in Daqing, 2007(4): 55-58 (in Chinese with English abstract).

- [6] 舒志华,张立有,刘刚. 复合砂体中单一河道的识别方法[J]. 大庆石油地质与开发,2006(4):18-20,120.

Shu Z H, Zhang L Y, Liu G. Identification of single channel in compound sand body [J]. Petroleum Geology & Oilfield Development in Daqing, 2006(4): 18-20, 120 (in Chinese with English abstract).

- [7] 赵翰卿. 高分辨率层序地层对比与我国的小层对比[J]. 大庆石油地质与开发,2005(1):5-9,12-105.

Zhao H Q. High-resolution sequential stratigraphy correlation and Chinese subzone correlation [J]. Petroleum Geology & Oilfield Development in Daqing, 2005(1): 5-9, 12-105 (in Chinese with English abstract).

- [8] 何宇航,于开春. 分流平原相复合砂体单一河道识别及效果分析[J]. 大庆石油地质与开发,2005(2):17-19,104.

He Y H, Yu K C. Single channel identification and effect analysis of compound sand body in distributary plain facies [J]. Petroleum Geology & Oilfield Development in Daqing, 2005(2): 17-19, 104 (in Chinese with English abstract).

- [9] 沈华,尹微,徐佑平. 提高砂岩油藏储层预测精度的方法[J]. 大庆石油地质与开发,2005(3):24-27,104.

Shen H, Yin W, Xu Y P. Improving predicting accuracy of sandstone oil reservoirs [J]. Petroleum Geology & Oilfield Development in Daqing, 2005(3): 24-27, 104 (in Chinese with English abstract).

- [10] Haldorsen H H, Chang D M. Notes on stochastic shales from outcrop to simulation models[C]//Lake L W, Carol H B. Reservoir characterization. New York, USA: Academic Press,

- 1986;152-167.
- [11] Haldorsen H H, Damsleth E. Stochastic modelling[J]. J. Pet. Technol., 1990, 42(4): 404-412.
- [12] Holden L, Hauge R, Skare O, et al. Modeling of fluvial reservoirs with object models[J]. Math. Geol., 1998, 30: 473-496.
- [13] Strebelle S B, Journel A G. Reservoir modeling using multiple-point statistics[C]//Anon. SPE Paper 71324 Presented at the SPE Annual Technical Conference and Exhibition Held in New Orleans, Louisiana, 2001; 1-11.
- [14] Strebelle S B. Conditional simulation of complex geological structures using multiple-point statistics [J]. Math. Geol., 2002, 34: 1-21.
- [15] Mariethoz G, Renard P, Straubhaar J. The direct sampling method to perform multiple-point geostatistical simulations [J]. Water Resour. Res., 2010, 46: 11536
- [16] Rezaee H, Marcotte D, Tahmasebi P, et al. Multiple-point geostatistical simulation using enriched pattern databases[J]. Stoch. Environ. Res. Risk Assess., 2015, 29: 893-913.
- [17] 陈麒玉, 刘刚, 何珍文, 等. 面向地质大数据的结构-属性一体化三维地质建模技术现状与展望[J]. 地质科技通报, 2020, 39(4): 51-58.
Chen Q Y, Liu G, He Z W, et al. Current situation and prospect of structure-attribute integrated 3D geological modeling technology for geological big data[J]. Bulletin of Geological Science and Technology, 2020, 39(4): 51-58 (in Chinese with English abstract).
- [18] 吴胜和, 李文克. 多点地质统计学: 理论、应用与展望[J]. 古地理学报, 2005(1): 137-144.
Wu S H, Li W K. Multiple-point geostatistics: Theory, application and perspective[J]. Journal of Palaeogeography, 2005(1): 137-144 (in Chinese with English abstract).
- [19] 郝慧珍, 顾庆, 胡修棉. 基于机器学习的矿物智能识别方法研究进展与展望[J]. 地球科学, 2021, 46(9): 3091-3106.
Hao H Z, Gu Q, Hu X M. Research advances and prospective in mineral intelligent identification based on machine learning [J]. Earth Science, 2021, 46(9): 3091-3106 (in Chinese with English abstract).
- [20] 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 Y J, Zhou Z, Lin H X, et al. The mineral intelligence identification method based on deep learning algorithms[J]. Earth Science Frontiers, 2020, 27(5): 39-47 (in Chinese with English abstract).
- [21] 刘彦锋, 张文彪, 段太忠, 等. 深度学习油气藏地质建模研究进展[J]. 地质科技通报, 2021, 40(4): 235-241.
Liu Y F, Zhang W B, Duan T Z, et al. Progress of deep learning in oil and gas reservoir geological modeling[J]. Bulletin of Geological Science and Technology, 2021, 40(4): 235-241 (in Chinese with English abstract).
- [22] 马瑶, 赵江南. 机器学习方法在矿产资源定量预测应用研究进展[J]. 地质科技通报, 2021, 40(1): 132-141.
Ma Y, Zhao J N. Advances in the application of machine learning methods in mineral prospectivity mapping[J]. Bulletin of Geological Science and Technology, 2021, 40(1): 132-141 (in Chinese with English abstract).
- [23] 万里明, 吴均, 卢军凯, 等. 基于 Adam-神经网络的致密砂岩脆性评价方法: 以南堡凹陷高北边坡为例[J]. 地质科技通报, 2020, 39(2): 94-103.
Wan L M, Wu J, Lu J K, et al. Brittleness evaluation method of tight sandstone based on Adam-neural network: A case study of a block in Gaobei slope, Nanpu Sag[J]. Bulletin of Geological Science and Technology, 2020, 39(2): 94-103 (in Chinese with English abstract).
- [24] 陈仲为. 深度学习的发展以及应用[J]. 现代计算机, 2019(17): 46-50.
Chen Z W. Development and application of deep learning[J]. Modern Computer, 2019(17): 46-50 (in Chinese with English abstract).
- [25] 张福临. 深度学习在岩石薄片图像生成中的研究与应用[D]. 西安: 西安石油大学, 2021.
Zhang F L. Research and application of depth learning in rock slice image generation [D]. Xi'an: Xi'an Shiyou University, 2021 (in Chinese with English abstract).
- [26] 李兴保. 基于生成对抗网络的数字岩心重构研究[D]. 合肥: 合肥工业大学, 2020.
Li X B. Research on digital core reconstruction based on generative countermeasure network [D]. Hefei: Hefei University of Technology, 2020 (in Chinese with English abstract).
- [27] 张挺, 王先武, 杜奕, 等. 基于 DCGANs 的二维页岩图像重构方法[J]. 上海电力大学学报, 2021, 37(4): 402-406.
Zhang T, Wang X W, Du Y, et al. 2D shale image reconstruction based on DCGANs [J]. Journal of Shanghai University of Electric Power, 2021, 37(4): 402-406 (in Chinese with English abstract).
- [28] 陈龙. 基于生成对抗网络的多孔介质重构[D]. 西安: 长安大学, 2020.
Chen L. Reconstruction of porous media based on generative countermeasure network [D]. Xi'an: Chang'an University, 2020 (in Chinese with English abstract).
- [29] Diederik P. Kingma and max welling, auto-encoding variational bayes [J]. CoRR, 2013, abs/1312. 6114.
- [30] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets [C]//Anon. Proceedings of the 2014 Conference on Advances in Neural Information Processing Systems, [S. l.]: [s. n.], 2014: 2672-2680.
- [31] Canchumuni S W A, Emerick A A, Pacheco M A C. Towards a robust parameterization for conditioning facies models using deep variational autoencoders and ensemble smoother [J]. Computers and Geosciences, 2019, 128: 87-102.
- [32] Laloy E, Hérault R, Jacques D, et al. Training-image based geostatistical inversion using a spatial generative adversarial neural network [J]. Water Resources Research, 2018, 54(1): 381-406.
- [33] Zhang T F, Peter T K, Emilien D, et al. Generating geologically realistic 3D reservoir facies models using deep learning of sedimentary architecture with generative adversarial networks [J]. Petroleum Science, 2019, 16: 541-549.
- [34] Bao J M, Chen D, Wen F, et al. CVAE-GAN: Fine-grained image generation through asymmetric training [C]//Anon. Proceedings of the IEEE International Conference on Computer Vision, [S. l.]: [s. n.], 2017: 2745-2754.
- [35] Kingma D, Ba J. Adam: A method for stochastic optimization [C]//Anon. Proceedings of International Conference on Learning Representations, [S. l.]: [s. n.], 2015.