

不平衡数据集下导管架平台结构损伤识别研究*

王维刚^{1,2}, 田 丰², 路敬祯¹

(1. 东北石油大学三亚海洋油气研究院 三亚, 572000)

(2. 东北石油大学机械科学与工程学院 大庆, 163318)

摘要 针对导管架平台结构损伤识别中数据不平衡问题,提出了一种基于一维条件生成对抗网络(one-dimensional conditional generative adversarial network,简称 1D-CGAN)与长短时记忆(long short-term memory,简称 LSTM)网络结合的导管架平台结构损伤识别方法。该方法将采集的一维损伤数据直接输入条件生成对抗网络(conditional generative adversarial network,简称 CGAN),通过在判别器和生成器中添加标签信息构建 1D-CGAN,利用标签信息控制其生成特定的新损伤样本,从而与完好状态样本组成平衡样本集。在此基础上,将划分的训练集输入到 LSTM 进行模型训练和损伤识别。实验结果表明:随着不平衡程度降低,所提出方法的识别准确率不断提高,当数据集达到平衡时,识别准确率能够达到 92.5%;与其他方法相比,所提出方法的识别准确率和精确率均有明显的提高。该方法提高了模型的分类性能,为不平衡数据下海洋结构损伤识别提供了一种新思路。

关键词 导管架平台;一维条件生成对抗网络;长短时记忆网络;不平衡问题;结构损伤识别

中图分类号 TH17;P751;TP181

引 言

海洋导管架平台作为海洋资源开发的基础性设施,是海上生产和生活的基地,世界上已建成的固定式海洋平台中大多数已进入平台结构设计使用寿命的后期,面临超期服役的问题。为保障这些平台的安全运行,急需发展延寿评估技术。结构健康监测^[1](structural health monitoring,简称 SHM)是一种评估结构完整性和剩余寿命的重要工具,结构损伤识别是 SHM 的基础科学问题。

深度学习^[2]具有强大的特征提取能力,在结构损伤识别领域应用广泛。该方法需要大量高质量数据进行训练,而获取平台损伤数据较为困难,损伤数据的匮乏将严重影响深度学习模型的识别精度。因此,解决导管架平台损伤识别过程中由于缺乏损伤数据导致的数据集不平衡问题非常必要。

目前,针对不平衡数据分类问题的方法主要有欠采样法、随机过采样法(random oversampling,简称 ROS)、合成少数过采样技术(synthetic minority oversampling technique,简称 SMOTE)和深度学习等方法。欠采样方法通过丢弃大样本中的一些数据,使不同类别的数据样本达到平衡。孟东霞等^[3]

提出了一种基于 k 近邻中心偏移因子对多数类样本欠采样的数据处理方法。Arefeen 等^[4]提出了基于神经网络的欠采样方法,但造成模型偏差较大。ROS 从少数类样本中反复抽取若干样本进行复制,并与多数类样本合并构成平衡数据集。该方法只是采取简单复制样本的策略来增加少数类样本,容易产生模型过拟合的问题。SMOTE 通过找到少数类样本数据的近邻进行合成,摆脱了对样本的简单重复。韩志艳等^[5]通过引入邻域的概念,提出了加权合成少数类过采样技术,在一定程度上提升了传统 SMOTE 方法的性能。严远亭等^[6]提出了基于构造性覆盖算法的过采样技术,并在 12 个典型数据集上取得了良好效果。Wei 等^[7]提出了一种改进的随机合成少数类过采样技术(improved random synthesis minority class oversampling technique,简称 IR-SMOTE),该方法虽然可以避免对少数类样本的单纯重复,但无法克服非平衡数据集的数据分布问题,容易产生分布边缘化。随着深度学习技术的发展,生成对抗网络(generative adversarial networks,简称 GAN)^[8]通过生成器与判别器的对抗,可以生成多样性更强的高质量样本。GAN 及其变体结构在海洋及其他土木结构领域得到了初步应

* 海南省院士创新平台科研资助项目(YSP TZ X 202301);海南省自然科学基金资助项目(623MS071)

收稿日期:2023-09-21;修回日期:2024-05-06

用。金禄渊^[8]提出非线性系统 GAN,将深度学习方法融入海洋平台损伤识别中。Dunphy 等^[9]利用 GAN 生成水泥结构表面损伤图片。在 GAN 变体结构中,CGAN 在结构损伤识别与故障诊断领域应用广泛。Rastin 等^[10]利用 CGAN 的判别器识别桥梁结构损伤位置。Luo 等^[11]利用 CGAN 生成齿轮故障状态下的频域信号,利用卷积神经网络完成对齿轮的故障诊断。在损伤识别领域,大多通过训练 GAN 的过程训练 GAN 的判别器,将 GAN 达到纳什均衡后的判别器作为分类器,或者通过 GAN 生成结构损伤表面的照片来扩充损伤样本集。但这 2 种方式都有缺陷,前者易导致 GAN 在训练过程中失去平衡,难以对判别器的训练效果做出评估;后者需要预先知道大致损伤区域,且要求检测人员能够接近待检测部位,不适用于导管架平台这类大型结构的损伤识别^[12]。

笔者在 CGAN 基础上,直接将采集的一维振动数据引入 CGAN 中,提出 1D-CGAN 和 LSTM 网络结合的海油导管架平台结构损伤识别方法。该方法不但能快速处理获取的数据,还能生成新的损伤样本,同时利用 LSTM 能更好地解决长振动数据序列在训练过程中的梯度消失和梯度爆炸问题。

1 基于 1D-CGAN 与 LSTM 的导管架平台结构损伤识别方法

1.1 生成对抗网络

GAN 主要包含生成器和判别器 2 个部分。图 1 为 GAN 结构示意图。生成器通过生成与真实样本 x 相似的样本 $G(z)$ 来欺骗判别器,直到判别器无法区分真实样本与生成样本。通常假设随机噪声 z 服从高斯分布。

在 GAN 训练过程中,生成器和判别器交替更新其模型参数,直到达到纳什均衡。判别器无法判断生成器的输入数据是真是假,对于给定数据预测

为真的概率稳定在 0.5。该过程可以表示为关于价值函数最大化和最小化的非合作博弈,其优化目标函数为

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

其中: $V(D, G)$ 为判别器和生成器之间的价值函数,反映了判别器正确判断数据真假的概率; E 代表数学期望; x 为真实数据; z 为随机噪声; $G(z)$ 为生成数据; $P_{\text{data}}(x)$ 为真实数据 x 的分布; $x \sim P_{\text{data}}$ 表示从真实数据中抽样; $D(x)$ 为判别器网络将真实样本判定为真的概率; $P_z(z)$ 为噪声向量 z 的先验分布; $z \sim P_z(z)$ 代表从 $P_z(z)$ 中获取抽样; $D(G(z))$ 为判别器网络将生成样本判定为真的概率。

1.2 一维条件生成对抗网络

1D-CGAN 在判别器和生成器中都添加了标签信息 y , 可以通过 y 控制 1D-CGAN 生成特定的损伤样本。优化目标函数可表示为

$$\min_G \max_D V(D, G) = E_{x \sim P_{\text{data}}(x)} [\log D(x|y)] + E_{z \sim P_z(z)} [\log (1 - D(G(z|y)))] \quad (2)$$

本研究中生成器和判别器分别采用与深度卷积生成对抗网络(deep convolutional generative adversarial network, 简称 DCGAN)类似的转置卷积和卷积结构。在判别器中,卷积降低数据样本的尺寸,增加样本维度,从而提取数据样本特征。转置卷积为卷积过程的逆运算,通过降低样本维度,增加样本尺寸实现新样本生成。首先,生成器可以根据输入的振动信号样本生成新样本,并将生成样本与原始样本一起输入到判别器中进行真伪判别;然后,以真实性标签误差和类别标签误差作为损失函数,交替迭代训练生成器和判别器,通过生成器和判别器不断对抗训练,最终生成满足质量要求的损伤样本。为了观察 1D-CGAN 是否收敛以及评估生成样本质量,笔者引进生成器和判别器损失得分。生成器损失得分公式为

$$L_{\text{Generator}} = -\text{mean}(\log(\hat{Y}_{\text{Generated}})) \quad (3)$$

其中: $\hat{Y}_{\text{Generated}}$ 为判别器将生成数据判定为真实数据的概率。

判别器损失得分公式为

$$L_{\text{Discriminator}} = -\text{mean}(\log(\hat{Y}_{\text{Real}})) - \text{mean}(\log(1 - \hat{Y}_{\text{Generated}})) \quad (4)$$

其中: $\hat{Y}_{\text{Real}}$ 为判别器将真实数据判定为生成数据的概率。

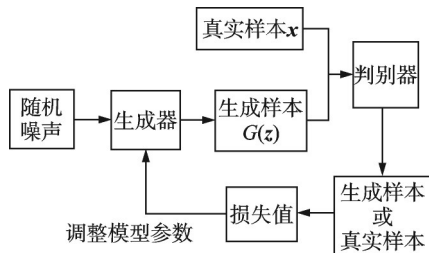


图1 GAN结构示意图

Fig.1 The schematic diagram of GAN structure

1.3 长短时记忆网络

LSTM网络是循环神经网络(recurrent neural network,简称RNN)的一个主要分支,在时间信息处理方面应用广泛。相比于RNN,LSTM网络引入了门单元来控制不同时刻的状态和输出,可以更好地解决长序列训练过程中的梯度消失和梯度爆炸问题。LSTM网络的门单元主要有遗忘门、输入门和输出门3种。

1.3.1 遗忘门

遗忘门的主要任务是通过sigmoid函数层接受上一个模块传递的输出并决定需要过滤掉哪些信息,输出公式为

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (5)$$

其中: σ 为sigmoid函数,表示为 $\text{sigmoid}(x) = 1/(1 + e^{-x})$; h_{t-1} 为前一个单元模块的输出; x_t 为当前输入; W_f 为权重; b_f 为偏置项。

1.3.2 输入门

输入门的主要任务是通过sigmoid函数层决定哪些信息需要更新以及通过正切函数层创建一个新的状态向量 $\tilde{C}_t$,生成候选记忆,并通过旧状态向量 C_{t-1} 乘以 f_t ,将一些不想保留的信息过滤掉,然后加上输入的 $i_t * \tilde{C}_t$ 形成新的状态向量 C_t ,输出公式为

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (6)$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (7)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (8)$$

其中: W_i 和 W_c 为权重; b_i 和 b_c 为偏置项。

1.3.3 输出门

输出门通过sigmoid函数层决定 C_t 的输出信息,随后与正切函数层处理后的 C_t 相乘,得到最后的输出公式为

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (9)$$

其中: W_o 为权重; b_o 为偏差项; h_t 为该单元模块最后输出结果, $h_t = o_t * \tanh(C_t)$; $*$ 为元素乘积; o_t 为经过sigmoid函数层后的输出。

1.4 导管架平台结构损伤识别框架

针对实际导管架平台结构损伤样本不足的问题,提出了基于1D-CGAN和LSTM网络的导管架平台结构损伤识别方法,其框架如图2所示,具体步骤如下:

1) 通过电动伺服装置对导管架平台施加激励,利用加速度传感器采集正常工况及不同损伤工况下的原始振动数据;

2) 利用采集的原始振动数据构建包含损伤和

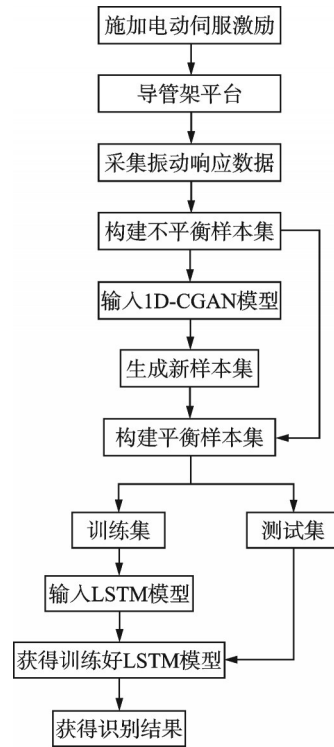


图2 基于1D-CGAN和LSTM的导管架平台结构损伤识别框架

Fig.2 Structural damage identification framework of offshore jacket platform based on 1D-CGAN and LSTM

正常样本的不平衡样本集;

3) 从不平衡样本集选一定数量的损伤样本输入1D-CGAN进行训练,生成新的样本集;

4) 通过原始不平衡样本和新生成的样本构建平衡样本集,并按一定比例分为训练集和测试集;

5) 采用训练集学习LSTM模型;

6) 利用训练好的LSTM模型对测试集进行分类,获得结构损伤识别结果。

2 实例研究

2.1 导管架平台简介

为验证所提方法的性能和有效性,将其用于导管架平台结构损伤识别。导管架平台损伤位置及测点布置如图3所示,主要由8根桩腿、若干斜撑与横撑以及上部组块组成。导管架平台模型底层长度为1 735 mm、宽度为1 390 mm、总高度为3 383 mm,总质量为285 kg,每根斜撑壁厚为3 mm,每根桩腿采用4个螺栓固定于底座。

2.2 损伤工况设置与振动数据获取

通过拆除不同位置处的斜撑来模拟不同损伤位置,更换不同壁厚的斜撑来模拟相同位置不同损伤

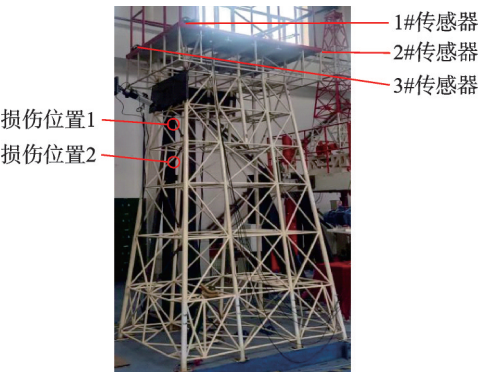


图3 导管架平台损伤位置及测点布置

Fig.3 Damage location and measuring point layout of jacket platform

程度,损伤位置分别处于第2~5层。导管架平台模型损伤工况如表1所示。

表1 导管架平台模型损伤工况

Tab.1 Damage condition of jacket platform model		
工况	位置	标签
1	完好	1
2	拆除位置1处斜撑	2
3	拆除位置2处斜撑	3
4	位置1更换2 mm 钢管	4
5	位置1更换1.5 mm 钢管	5

给如图4所示的电动伺服装置施加大小为1 kN、频率为1 Hz的正弦载荷,用于模拟平台受到的波浪载荷。根据有效独立法确定加速度传感器数量为3,布设点位置见图3,分别位于甲板的3个顶点处。传感器选择力平衡拾振器,型号为991B,并将其设置成加速度档,采集 x 和 y 方向的加速度信号。根据采样定理及平台结构模态参数,采样频率设为263 Hz。

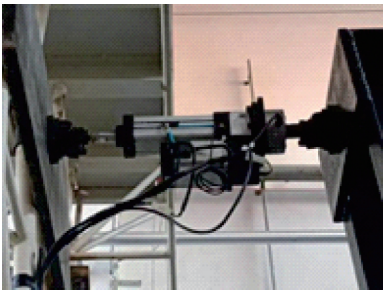


图4 电动伺服装置

Fig.4 Electric servo device

图5为工况1与工况2的时域信号。可见,通过时域信号无法区分2类信号。

2.3 1D-CGAN 结构超参数设置与生成样本质量评估

采用滑动窗口法对采集到的每个测点 y 方向振动信号进行采样,其中:信号长度为512,步长为256,训练集包含的正常工况样本数为3 000,每种损

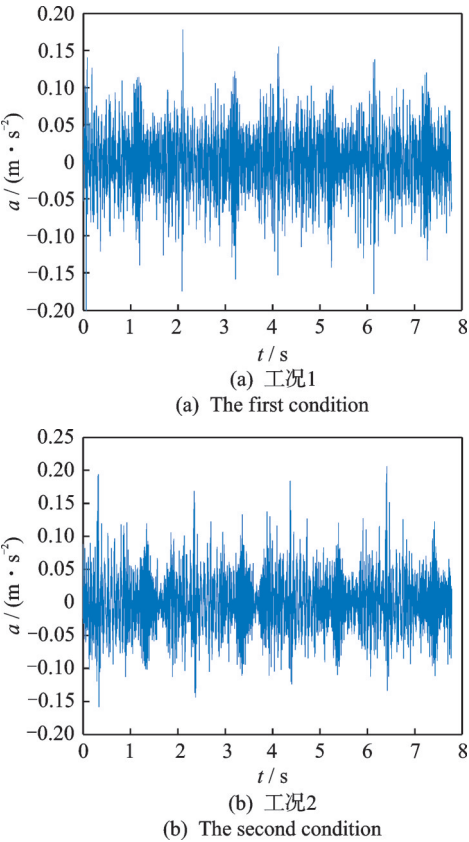


图5 工况1与工况2的时域信号

Fig.5 The time domain signals of the first and second condition

伤工况的样本数各为300,损伤样本全部用于训练1D-CGAN;测试集全部由原始样本组成,每种工况各1 000;所有样本的维数均为 512×3 。

模型的超参数对于GAN的稳定性以及生成样本的质量至关重要,笔者设计的1D-CGAN生成器部分主要由5个转置卷积层和4个Relu激活函数层组成;判别器部分主要由5个卷积层和4个Leaky Relu激活函数层组成。1D-CGAN的主要超参数如表2所示。为提高模型性能,在设计生成器和判别器时采用如下策略:①为了更有利于模型收敛,将判别器的学习率设为0.000 2,生成器学习率设为0.000 1;②为防止过拟合,在判别器前4个卷积层后面添加比率为0.3的丢弃层。

图6为生成器与判别器损失值得分曲线。可以看出,得分在训练初期出现大幅震荡,随后判别器损失值得分逐渐减小,生成器损失值得分逐渐增大,经过2 500次迭代后收敛至0.5附近。这说明网络已经学习了原始样本的特征分布,且生成的损伤样本特征分布能够接近原始样本。由于生成器和判别器之间的对抗学习,因此观察到这2个网络在训练过程中存在一定的波动。

表2 1D-CGAN的主要超参数

Tab.2 The main hyperparameters of 1D-CGAN

项目名称	网络层	卷积核大小	卷积核数量	步长	裁剪	填充	激活函数
生成器	转置卷积层1	5×1	512	0	0	—	ReLU
	转置卷积层2	11×1	256	3	(1,0)	—	ReLU
	转置卷积层3	13×1	128	4	(1,0)	—	ReLU
	转置卷积层4	12×1	64	4	(2,0)	—	ReLU
	转置卷积层5	5×1	1	3	(1,0)	—	—
判别器	卷积层1	8×1	512	4	—	0	Leaky ReLU
	卷积层2	9×1	256	4	—	(1,0)	Leaky ReLU
	卷积层3	10×1	128	3	—	(1,0)	Leaky ReLU
	卷积层4	9×1	64	3	—	0	Leaky ReLU
	卷积层5	8×1	1	1	—	0	—

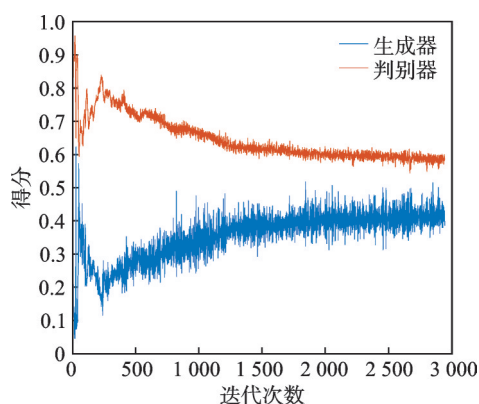


图6 生成器与判别器损失值得分曲线

Fig.6 The score curve of generator and discriminator

图7为工况2原始样本与生成样本时域波形。可以看出,生成样本与原始样本大体相似,但并不完全等同,这说明生成的样本避免了对原始样本的简单复制,可以增强模型的鲁棒性,并提高模型的泛化能力。

为进一步评估原始损伤样本与生成样本的相似性,t分布随机邻域嵌入(t-distributed stochastic neighbor embedding,简称t-SNE)聚类图如图8所示。图中,标签为6、7、8、9的样本分别代表类别标签为2、3、4、5样本的生成样本。可以看出,1D-CGAN生成的损伤样本和原始损伤样本的特征都能够聚合在一起,这表明生成样本的关键特征与原始样本具有相似的分布。因此,1D-CGAN能够

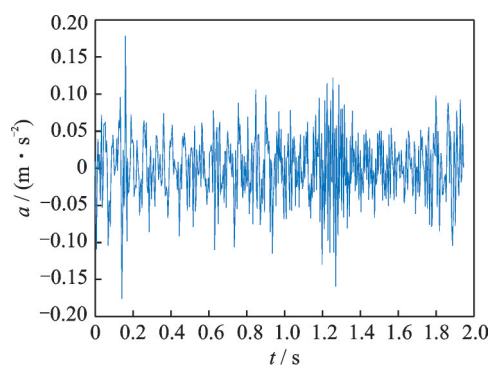
(b) 生成样本
(b) Generated sample

图7 工况2原始样本与生成样本时域波形

Fig.7 The time domain waveform of the original sample and the generated sample of the second working condition

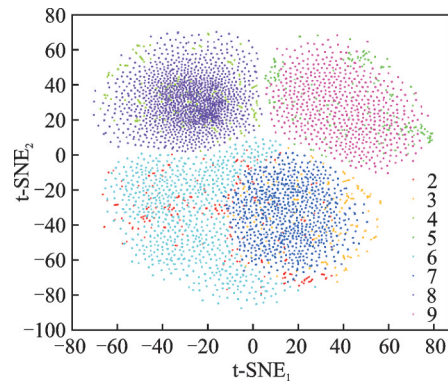


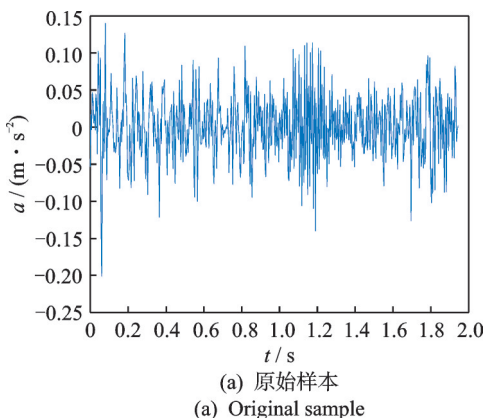
图8 t-SNE聚类图

Fig.8 T-SNE clustering diagram

生成与原始样本相似的样本。

2.4 结构损伤识别结果

为验证笔者所提方法的有效性,定义不平衡率为 $i_r = N_{\text{minority}} / N_{\text{majority}}$,其中: N_{minority} 为每一类损伤样本数量; N_{majority} 为正常样本数量。不平衡率的数值在0~1之间,越接近1,则不平衡程度越小。首先,设置原始数据集不平衡率为1:10,即每一类损伤样本与正常样本的数量比例为1:10,将原始训练集中

(a) 原始样本
(a) Original sample

所有损伤样本进行 1D-CGAN 训练,获得生成损伤样本;其次,向原始数据集添加生成的损伤样本,得到不平衡率分别为 1:4、1:2 和 1:1 的数据集;最后,使用不同不平衡率的训练集训练 LSTM 网络,并用测试集评估所提出方法的识别效果。不平衡率达到 1:1 时,平衡数据集下准确率变化曲线如图 9 所示。

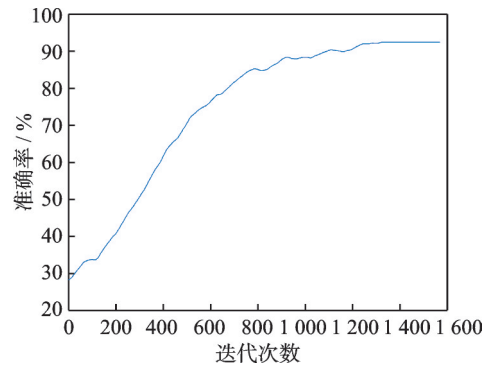


图 9 平衡数据集下准确率变化曲线
Fig.9 The change curve of accuracy under balanced dataset

基于 1D-CGAN 和 LSTM 的识别准确率如表 3 所示。可以看出:①当不平衡率为 1:10 时,所提出方法几乎无法学习到损伤样本的特征,识别准确率很低;②当不平衡率为 1:4 和 1:2 时,识别准确率都有了较大幅度的提升;③当达到平衡状态时,识别准确率达到 92.5%。可见,所提出方法对于损伤位置和损伤程度均具有良好的识别效果。

表 3 基于 1D-CGAN 和 LSTM 的识别准确率
Tab.3 The recognition accuracy based on 1D-CGAN and LSTM

不平衡率	识别准确率/%
1:10	32.4
1:4	42.9
1:2	73.0
1:1	92.5

3 方法对比

3.1 与其他不平衡方法对比

为体现笔者所提出方法的优越性,将基于 1D-CGAN 的损伤识别方法与 ROS、SMOTE、一维瓦瑟斯坦生成对抗网络(one-dimensional Wasserstein generative adversarial network,简称 1D-WGAN)和 1D-DCGAN 的损伤识别方法进行对比分析,分类模型均采用 LSTM,并采用准确率、精确率和召回率 3 个参数作为评价指标。基于不同不平衡方法的识别准确率如表 4 所示。

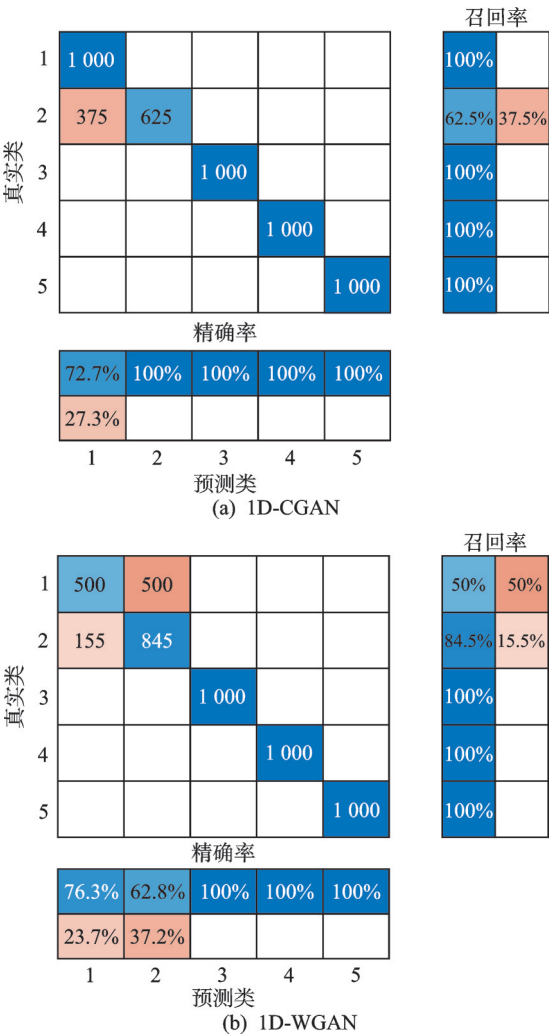
针对平衡数据集(不平衡率为 1:1),在不同不平衡

表 4 基于不同不平衡方法的识别准确率
Tab.4 Recognition accuracy based on different imbalance methods

不平衡方法	不平衡率			
	1:10	1:4	1:2	1:1
1D-CGAN	32.4	42.9	73.0	92.5
1D-WGAN	31.6	38.9	69.5	86.9
1D-DCGAN	24.9	31.0	60.0	81.6
SMOTE	30.0	38.4	68.4	82.0
ROS	28.6	37.3	67.3	78.5

方法下混淆矩阵如图 10 所示。混淆矩阵给出了每种工况的精确率和召回率,其中:混淆矩阵最右侧的列为每个工况的召回率;底部的行为每个工况的精确率。

可以看出,随着不平衡程度的降低,5 种方法的识别准确率和精确率都有所提升。从整体看,基于 1D-CGAN 方法的识别准确率和精确率高于其他方法。由于 1D-CGAN 采用对抗生成技术,生成样本质量高于 ROS 和 SMOTE,因此基于 1D-CGAN 的识别准确率高出其他方法。与 1D-CGAN 相比,由于 1D-WGAN 处理利普希茨限制条件时直接采用了“梯度裁剪”,容易出现梯度消失或梯度爆炸,从而



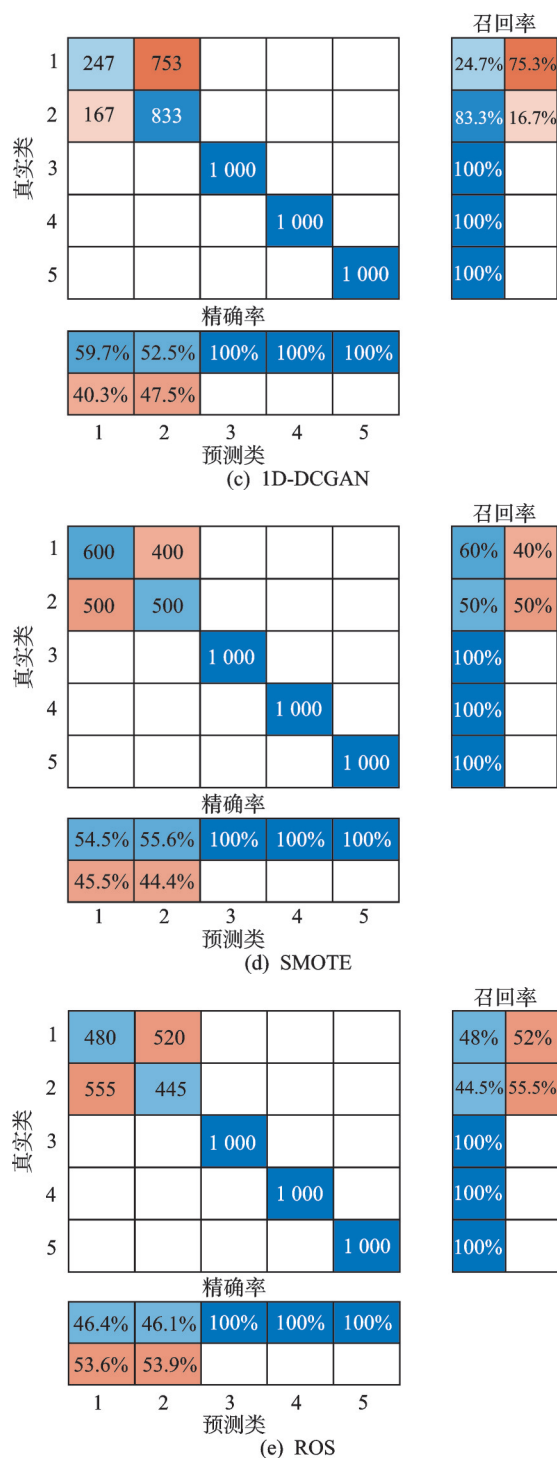


图10 不同不平衡方法下混淆矩阵

Fig.10 Confusion matrix under different unbalanced methods

降低了生成样本的质量,因此基于1D-CGAN的识别准确率高于1D-WGAN。

3.2 与其他识别网络对比

采用各不平衡率下的数据集分别训练一维信号分类模型LSTM、一维卷积神经网络(one-dimensional convolutional neural network,简称1D-CNN)和反向

传播(back propagation,简称BP)神经网络,并用测试集评估模型性能。针对平衡数据集(不平衡率为1:1),不同识别网络下混淆矩阵如图11所示,其给出了不同识别网络下每种工况的精确率和召回率。表5为不同识别网络的结构损伤识别准确率。

可以看出:随着不平衡率的降低,LSTM与1D-CNN的识别准确率均有所上升,而BP神经网络的准确率却几乎没有变化;基于LSTM的精确率和准确

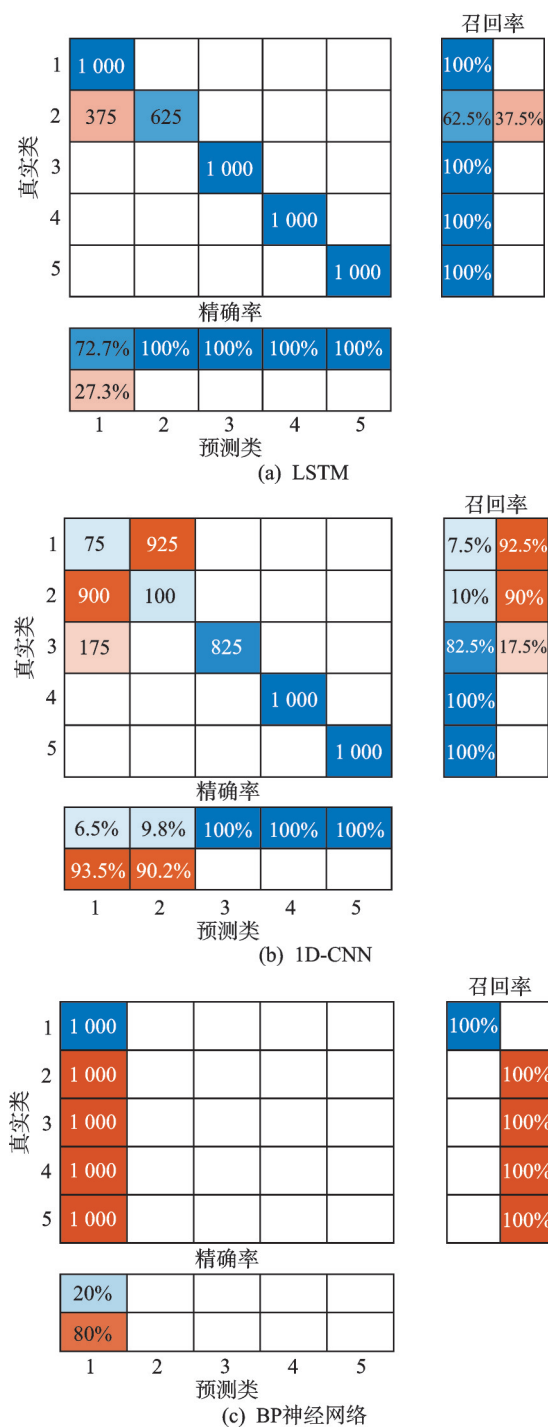


图11 不同识别网络下混淆矩阵

Fig.11 Confusion matrix under different recognition networks

率均高于其他2种方法。其原因是由于LSTM属于循环神经网络,具有记忆功能,在一维信号分类任务中的准确率高于1D-CNN,而BP神经网络属于非深度学习模型,在没有进行人工特征提取时其性能不佳。

表 5 不同识别网络的结构损伤识别准确率

识别网络	不平衡率			
	1:10	1:4	1:2	1:1
LSTM	32.4	42.9	73.0	92.5
1D-CNN	20.0	32.5	58.4	60.0
BP神经网络	20.0	21.0	20.0	20.0

4 结 论

1) 构建了基于1D-CGAN和LSTM网络的导管架平台结构损伤识别方法,实验结果验证了该方法能够较好地解决不平衡数据带来的损伤识别准确率低的问题。

2) 通过1D-CGAN可以生成损伤样本,与导管架平台结构完好状态样本构建平衡数据集,从而解决了初始不平衡数据问题,实验结果也验证了该方法优于其他方法。

3) 通过LSTM对构建的平衡数据集进行分类,实验结果验证了采用LSTM能够更好地捕捉序列数据中的长期依赖关系,在一维信号分类任务中具有更高的准确率。

4) 本研究方法在一定程度上提高了不平衡数据集下导管架平台损伤识别得准确率,但整体损伤识别的准确率有待进一步很高。

参 考 文 献

[1] 邵冬明. 海洋平台岸基监测和健康预测系统的研究[D]. 镇江: 江苏科技大学, 2020.

[2] 梁韬, 叶涛萍, 李守文, 等. 结合多通道MTF和CNN的框架结构损伤识别方法[J]. 振动、测试与诊断, 2024, 44(2): 217-224.

LIANG Tao, YE Taoping, LI Shouwen, et al. A framework structure damage identification method combining multi-channel MTF and CNN [J]. Journal of Vibration, Measurement & Diagnosis, 2024, 44 (2) : 217-224. (in Chinese)

[3] 孟东霞, 谢林燕. 基于k近邻中心偏移因子的欠采样方法[J]. 统计与决策, 2023, 39(12): 40-44.

MENG Dongxia, XIE Linyan. An undersampling method based on k-nearest neighbor center offset factor [J]. Statistics and Decision, 2023, 39(12): 40-44. (in Chinese)

[4] AREFEEN M A, NIMI S T, RAHMAN M S. Neural network-based under sampling techniques [J]. IEEE Transactions on Systems, Man and Cybernetics Systems, 2022, 52(2): 1111-1120.

[5] 韩志艳, 王健. 基于加权合成少数类过采样技术的故障诊断[J]. 计算机技术与发展, 2016, 26(9): 43-46.

HAN Zhiyan, WANG Jian. Fault diagnosis method based on weighted synthetic minority oversampling technique [J]. Computer Technology and Development, 2016, 26(9): 43-46. (in Chinese)

[6] 严远亭, 朱原玮, 吴增宝, 等. 构造性覆盖算法的SMOTE过采样方法[J]. 计算机科学与探索, 2020, 14(6): 975-984.

YAN Yuanting, ZHU Yuanwei, WU Zengbao, et al. Constructive covering algorithm-based SMOTE oversampling method [J]. Journal of Frontiers of Computer Science and Technology, 2020, 14 (6) : 975-984. (in Chinese)

[7] WEI G L, MU W M, SONG Y, et al. An improved and random synthetic minority oversampling technique for imbalanced data [J]. Knowledge-Based Systems, 2022, 248: 108839.

[8] 金禄渊. 基于机器学习的海洋平台在线损伤识别研究[D]. 大连: 大连理工大学, 2019.

[9] DUNPHY K, FEKRI M N, GROLINGER K, et al. Data augmentation for deep-learning-based multiclass structural damage detection using limited information [J]. Sensors, 2022, 22(16): 6193.

[10] RASTIN Z, AMIRI G G, DARVISHAN E. Generative adversarial network for damage identification in civil structures [J]. Shock and Vibration, 2021(1): 398-405.

[11] LUO J, HUANG J, LI H. A case study of conditional deep convolutional generative adversarial networks in machine fault diagnosis [J]. Journal of Intelligent Manufacturing, 2021, 32(2): 407-425.

[12] 蔡东成, 张健飞. 不平衡样本下基于CVAE和CNN的结构损伤识别方法[J]. 土木工程与管理学报, 2023, 40(2): 108-116, 129.

CAI Dongcheng, ZHANG Jianfei. Method of structural damage identification based on CVAE and CNN under imbalanced samples [J]. Journal of Civil Engineering and Management, 2023, 40(2): 108-116, 129. (in Chinese)



第一作者简介:王维刚,男,1976年3月生,博士,副教授。主要研究方向为海洋工程结构与装备结构健康监测、数字孪生、损伤识别和机器学习。曾发表《Sparsity discriminant preserving projection for machinery fault diagnosis》(《Measurement》2021, Vol.173, No.108488)等论文。

E-mail:wwg@nepu.edu.cn