马剑波，左翔，丛小飞，等. 基于深度学习的水利工控网络流量异常检测方法[J]. 水利水电技术(中英文)，2025，56(4)：167-178. DOI：10.13928/j.cnki.wrahe.2025.04.014

MA Jianbo，ZUO Xiang，CONG Xiaofei，et al. Network traffic anomaly detection method for water conservancy industrial control systems based on deep learning[J]. Water Resources and Hydropower Engineering，2025，56(4)：167-178. DOI：10.13928/j.cnki.wrahe.2025.04.014

基于深度学习的水利工控网络流量异常检测方法

马剑波[1]，左 翔[2,3]，丛小飞[3]，叶瑞禄[3]，刘威风[3]

(1. 江苏省秦淮河水利工程管理处，江苏 南京 210022；2. 水资源高效利用与工程安全国家工程研究中心，江苏 南京 210098；3. 南京中禹智慧水利研究院有限公司，江苏 南京 210012)

摘 要：【目的】针对水利工控网络流量数据集不平衡、特征维数多和检测效率低等问题，提出一种结合改进条件生成对抗网络(ICGAN)、深度残差收缩网络(DRSN)、长短期记忆网络(LSTM)的流量异常检测方法。【方法】利用 ICGAN 构建了网络流量平衡数据集，利用 DRSN-LSTM 混合深度学习模型对网络异常流量数据进行检测，其中 DRSN 负责提取数据的空间特征，其残差连接可以解决网络退化与过拟合问题，压缩和激励网络可自动为每个特征图分配权重系数以提高检测效果，LSTM 负责提取数据的时间特征。【结果】以秦淮河武定门闸站为应用场景对该方法进行测试，结果表明采用 ICGAN 优化后的数据集训练的各类检测模型，其流量分类精度高于原始数据集。DRSN-LSTM 的网络流量异常检测的总体准确率达到了 98.76%，其中正常数据分类的 P、R 和 $F1$ 值，分别达到了 99.22%、99.69%和 99.46%，在评价指标上优于比较模型。【结论】融合 ICGAN、DRSN 和 LSTM 算法优势的水利工控网络流量异常检测方法，能够有效改善原始数据集中的类别不平衡性问题，提高对异常工控网络流量的检测能力，保障水利工程安全稳定运行。

关键词：水利工控；网络流量异常检测；深度学习；条件生成对抗网络；深度残差收缩网络；长短期记忆网络；评价指标

DOI：10.13928/j.cnki.wrahe.2025.04.014 **开放科学(资源服务)标志码(OSID)**：

中图分类号：TP301.6 **文献标志码**：A **文章编号**：1000-0860(2025)04-0167-12

Network traffic anomaly detection method for water conservancy industrial control systems based on deep learning

MA Jianbo[1], ZUO Xiang[2,3], CONG Xiaofei[3], YE Ruilu[3], LIU Weifeng[3]

(1. Jiangsu Qinhuai River Water Conservancy Engineering Management Office, Nanjing 210022, Jiangsu, China; 2. National Engineering Research Center of Water Resources Efficient Utilization and Engineering Safety, Nanjing 210098, Jiangsu, China; 3. Nanjing Zhongyu Intelligent Water Conservancy Research Institute Co., Ltd., Nanjing 210012, Jiangsu, China)

Abstract: [Objective] This study proposes a network traffic anomaly detection method that addresses the issues of data

收稿日期：2024-04-08；**修回日期**：2024-05-22；**录用日期**：2024-05-27；**网络出版日期**：2024-06-28

基金项目：国家重点研发计划(2023YFC3006500)；江苏省水利科技项目(2022052，2022064)

作者简介：马剑波(1973—)，男，高级工程师，学士，主要从事水利工程控制研究。E-mail：734051253@qq.com

通信作者：左 翔(1984—)，男，高级工程师，博士，主要从事水利信息化研究。E-mail：knightzuo@163.com

imbalance, high feature dimensionality, and low detection efficiency in water conservancy industrial control networks. The method integrates an improved Conditional Generative Adversarial Network (ICGAN), Deep Residual Shrinking Network (DRSN), and Long Short-Term Memory Network (LSTM). [Methods] ICGAN was used to construct a balanced network traffic dataset, and a DRSN-LSTM hybrid deep learning model was employed for anomaly detection in network traffic. DRSN was responsible for extracting spatial features, with residual connections addressing network degradation and overfitting issues. The compression and excitation network automatically assigned weight coefficients to each feature map to improve detection performance. Lastly, LSTM extracted temporal features from the data. [Results] The method was tested in the application scenario of the Qinhuai River Wudingmen Sluice Station. The result showed that models trained on the ICGAN-optimized dataset achieved higher traffic classification accuracy than those trained on the original dataset. Overall, DRSN-LSTM achieved an accuracy of 98.76% in detecting network traffic anomalies. P, R, and $F1$ values for normal data classification were 99.22%, 99.69%, and 99.46%, respectively, which outperformed the comparison models in terms of these evaluation indicators. [Conclusion] By integrating the advantages of ICGAN, DRSN, and LSTM algorithms, the anomaly detection method for water conservancy industrial network traffic effectively alleviates the type imbalance in the original dataset, improves the detection ability of abnormal industrial control network traffic, and ensures the safe and stable operation of water conservancy projects.

Keywords: water conservancy industrial control; network traffic anomaly detection; deep learning; conditional generative adversarial networks; deep residual shrinkage network; long short-term memory network; evaluation indicator

0 引 言

水利工控系统作为国家关键信息基础设施，其网络安全具有十分重要的意义，一旦遭到攻击与破坏导致丧失关键核心功能或者引起数据泄露，会对国家安全、国计民生、公共利益等造成巨大损失[1]。与互联网系统相比，水利工控系统是由传感器、现地控制系统和远程监控系统等多个环节构成的封闭系统，目前存在标准不统一、缺乏加密认证机制、设备与协议存在安全漏洞等问题，导致系统的网络安全问题非常复杂[2]。基于工控协议的网络流量作为水利生产运行数据传输和交互的载体，包含了大量正常与异常行为相关信息。若能及时发现并捕获异常网络流量并及时预警，可以有效保障水利工控系统的安全运行。传统的网络异常流量检测方法主要有基于数据挖掘[3]、数理统计[4]、信息论[5]的网络异常流量检测方法。随着攻击手段的不断变化，传统检测方法已经无法适应当前的网络环境，存在误报率高、检测率低等问题[6]。

近年来，随着人工智能的发展，研究人员尝试将机器学习应用于网络异常流量检测，主要包括浅层学习模型(如支持向量机、决策树、朴素贝叶斯等)和深度学习模型(如卷积神经网络、循环神经网络、长短期记忆网络等)，其中深度学习在网络安全领域发挥了重要作用，通过利用不同的深度学习模型降低误报率，检测异常网络流量并分类，对识别出的网络攻击行为进行预警[7-11]。在实际应用中，针对网络流量数据集不平衡的问题，通常采用重采样、特征选择、交叉验证等方法进行处理[12]；燕昺昊等[13]采用SMOTE模型，有效弥补了少数类随机过采样的缺陷，但是新生成的样本可能会出现在多数类的决策区域中，随机生成的样本结果是两种类别决策区域的重叠概率会增加，导致两种类别更难以区分[14]。YANG等[15]采用一种条件变分自动编码器用于生成网络攻击流量，提高训练样本多样性，但是存在生成数据模糊的问题。YU等[16-17]研究了一种改进的生成对抗网络，通过利用外部条件信息和添加距离损失函数，强化了模型的样本生成能力。面向网络流量高维的数据特征，浅层学习模型的学习能力表现不佳，深度学习具有多层的结构，可以自动学习数据的高级特征和复杂模式，有效提高异常流量的检测率。KUMAR等[18-19]建立了卷积神经网络模型(CNN)并应用于网络入侵检测，通过卷积操作提取特征的局部相关性，利用多层“卷积层-下采样层”对网络中正常行为和异常行为的特征进行刻画，与经典的反向传播网络(BP)和支持向量机(SVM)等相比，CNN有效提高了入侵检测识别的分类准确性；常晓燕等[20]提出叠加LSTM模型针对网络流量中时序相关数据进行异常检测，使用正常数据来训练模型，并取不同的时间步长进行预测；SHERAZ等[21]比较了结合不同分类器的自编码器(AE)、LSTM和CNN不同变体在基于异常检测的入侵检测系统的适用性，结果表明LSTM的准确性最高，其次是CNN和自编码器。单一深度学习模型在处理复杂数据时可能存在特征提取不足的问题，例如CNN能够提取数据的空间特征，但其学习序列数据相关特征的能力不强；LSTM在RNN的基

础上增加了遗忘门，可以有效解决长序列训练过程中的存在的梯度爆炸问题，但其只能读取单方向的序列数据，且误报率较高[17]。混合模型可以有效提高入侵检测性能，ALEESA 等[22]提出了一种循环长短时记忆神经网络模型，通过构建 RNN-LSTM 混合模型，提高了模型的网络入侵检测准确率。

现有的研究成果仍然存在如下诸多问题：优化不平衡数据集生成少数类数据时，无法控制生成数据的模式，模型难以训练；处理高维数据和使用复杂模型时，存在梯度消失、局部最优解等现象；没有根据特征重要性高效利用数据；缺乏对网络流量时间和空间特征的全面考虑等。针对上述问题，采用一种新型的 ICGAN 模型对水利工控网络流量数据集进行平衡性处理，为异常流量检测任务奠定训练基础；采用 DRSN 模型提取流量的局部空间特征，增强对数据特征的敏感度，缓解梯度消失等问题；将 DRSN 提取的特征输出传递到 LSTM，利用 LSTM 提取流量的时间特征，结合 Softmax 分类器，实现对网络流量的异常检测。该方法在检测的准确率、精确度、召回率和 $F1$ 分数指标上具有较好表现，可以为水利工控网络安全提供技术支撑。

1 水利工控网络流量的特点

由于水利工控系统高要求的稳定性、兼容性和扩展性等需求，需要采用开放、简单、易于调试和维护的网络通信协议。Modbus 协议作为最早应用于工业现场的开源工业以太网协议，并可以采用多种通信方式，因此在水利工控领域得到了广泛的应用[23]。Modbus TCP 协议是一种采用请求/应答方式的应用层消息协议，采用 Master/Slave 方式进行通讯[24]，其报文格式如图 1 所示。本文中主要对基于 Modbus TCP 协议的水利工控网络流量进行分析与研究。结合水利工控系统的应用场景，其网络流量具有实时性强、稳定性高、可靠性高和数据密集等特点。由于 Modbus 协议在被设计时缺乏访问控制、身份验证等安全设定，容易发生未经授权的操作。恶意入侵者只要通过有效的 Modbus 地址和合法的功能码，就可以进行 Modbus 会话，而且 Modbus 未经加密，地址和指令容易被破解[25]，因此针对网络流量的异常检测是保护水利工控网络免受未经授权的访问和攻击的重要手段。

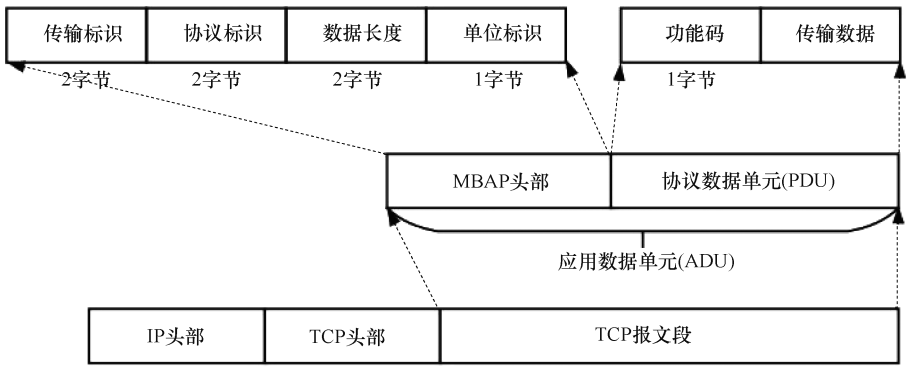

图 1 Modbus TCP 报文格式

Fig. 1 Modbus TCP message format

2 数据集平衡处理

由于水利工控网络在实际的运行过程中恶意攻击和入侵行为是小概率事件，导致正常与异常网络流量数据量的比例悬殊，当采用网络流量异常检测模型进行分类时，分类结果会偏向于多数类[26]，解决数据集不平衡问题对于提升异常检测方法的性能具有重要意义。条件生成对抗网络(CGAN)是 GAN 的改进型，它在传统的 GAN 基础上引入了条件信息，使得生成器可以根据条件信息生成特定类别的数据[27-29]。CGAN 的基本原理和 GAN 类似，包含一个生成器(G)和一个判别器(D)，通过对抗学习的方式进行训练。生成器接收随机噪声 z 和条件信息 c，负责生成逼真的数据 y；判别器接收真实数据 x 和生成数据 y，负责将 x 和 y 分别判别为真实数据和生成数据。

在实际应用中发现，当 G 将生成数据 y 输入 D 进行判别后，若判断为非真实数据，可能存在两个原因：(1)不符合真实流量数据的特征；(2)符合真实流量的特征，但不是符合条件信息 c 所指定的类别。因此当 G 要进行优化时无法选择正确的方向，容易出现 D 收敛、G 发散的现象。针对上述问题，研究提出了一种 ICGAN 模型，该模型采用双判别器模式(D1 和 D2)，其中 D1 负责对数据的真实性进行判别，而 D2 负责对生成数据的类别进行鉴定，通过双判别器模式，有利于 G 明确优化方向，其原理如图 2 所示。

判别器 D1 采用 3 层 MLP 网络结构，输入层神经元节点数与真实流量数据相同；隐藏层神经元节点数为 512，激活函数为 ReLU；输出层神经元节点数为 1，输出层的激活函数为 Sigmoid 函数，输出结果为[0，1]范围内的概率值，表示输入数据为真实数据的

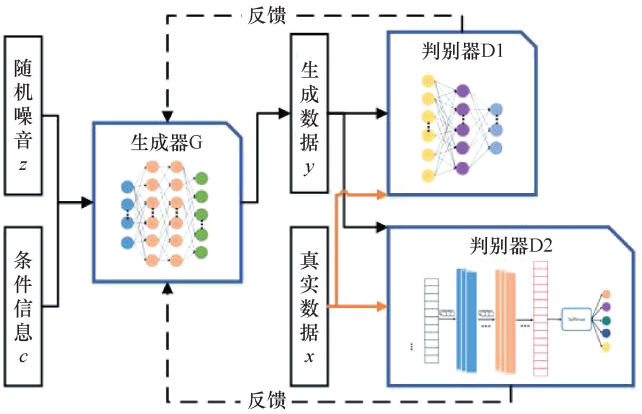

图 2　ICGAN 原理

Fig. 2　ICGAN principle

概率；损失函数($LOSS_{D1}$)使用二分类交叉熵损失函数，其公式为

$$LOSS_{D1} = -\gamma\ln(\bar{\gamma}) - (1-\gamma)\ln(1-\bar{\gamma}) \quad (1)$$

式中，γ 为真实值；$\bar{\gamma}$ 为预测值，且 $\bar{\gamma}$ 需要尽量靠近 γ，才能使 $LOSS_{D1}$ 最小化。

判别器 D2 本质是一个结构相对简单的网络流量异常检测模型，水利工控网络流量其一维数据的特性，与 1D-CNN 处理一维序列数据的优势相匹配，1D-CNN 能够保留流量数据原始特征中与异常相关的信息，最大化发挥学习特征的能力，并且训练速度快[30]，适用于当前流量数据快速生成的场景。设计的 1D-CNN 结构为 5 个一维卷积层、2 个一维池化层、1 个 Dropout 层、1 个平铺层、1 个全连接层。初始卷积核数目为 16，卷积核尺寸为 3×1，激活函数为 ReLU，Dropout 层丢弃神经元的概率值为 0.2，学习率为 0.001，采用 Adam 优化算法，softmax 分类器进行流量分类。D2 负责的是多分类任务，损失函数($LOSS_{D2}$)采用交叉熵损失函数，其公式为

$$LOSS_{D2} = -\sum_{i=1}^{k}\gamma_i\ln(\overline{\gamma_i}) \quad (2)$$

式中，k 为总类别数。

生成器 G 由 5 层的多层感知器(MLP)构建而成，输入层接收 z 和 c，其中 z 的长度设置为 80，c 的长度与流量类别数相同；隐藏层为 3 层，神经元节点数均设置为 1 024；输出层的数据长度与真实流量数据相同。损失函数($LOSS_G$)的为 $LOSS_{D1}$ 和 $LOSS_{D2}$ 的加权和，其公式为

$$LOSS_G = \omega \times LOSS_{D1} + (1-\omega) \times LOSS_{D2} \quad (3)$$

$$\omega = -\log(D1(y)) \quad (4)$$

式中，ω 为加权系数。

判别器 D1 的权重 ω 与概率值 $D1(y)$ 有关，当生成数据的 $D1(y)$ 值较大时，说明 D1 判别生成数据是真实数据的概率较大，因此需要加大 $LOSS_{D2}$ 所占比例，让生成器 G 向满足 D2 要求的方向优化；当 D1(y)值较小时，即生成数据还没有满足 D1 的要求，此时可以相对忽略生成数据是否符合指定类型的问题上，因此需要加大 $LOSS_{D1}$ 所占比例。

3　异常检测模型

针对水利工控网络流量的时空特征，以 DRSN 和 LSTM 为基础构建 DRSN-LSTM 混合模型，该模型同时具备卷积网络提取空间特征和循环神经网络提取时序特征的能力。

3.1　深度残差收缩网络

CNN 在处理高维数据和复杂模式时存在梯度消失、无法提取关键特征等问题。以 CNN 为基础的深度残差收缩网络(DRSN)由赵明航等[29]于 2020 年首次提出，DRSN 通过引入残差学习和收缩激励网络(SENet)来增强网络的性能，其原理如图 3 所示。残差学习利用恒等映射操作直接将浅层网络的特征跨层输入到深层网络，从而有效地避免了模型退化现象[30]，SENet(图 3 蓝框部分)自动学习各个特征通道的重要程度，用于各个特征通道的加权，从而提高从数据中提取重要特征的能力，将 SENet 得到的权重系数，输入软阈值模块，通过自动软阈值化，保留重要特征，丢弃冗余信息，提高了面向高维数据的特征提

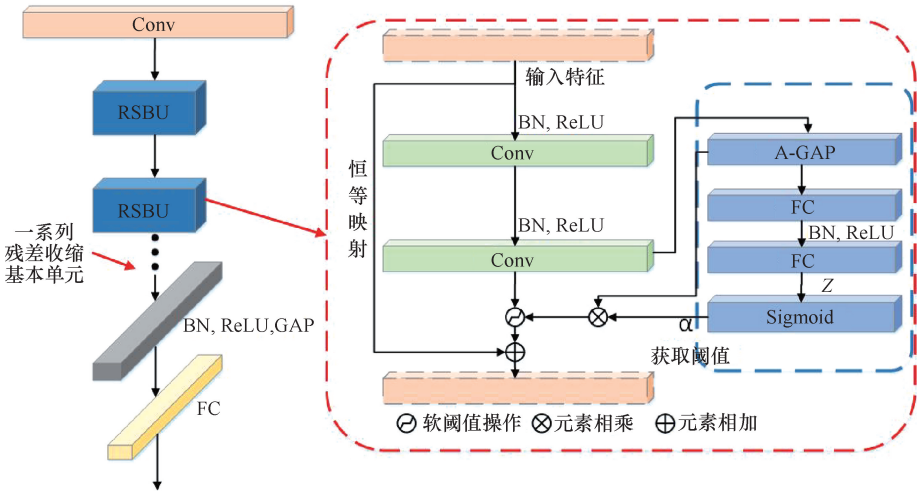

图 3　DRSN 原理

Fig. 3　DRSN principle

取能力，摆脱模型训练对先验知识和专家经验的依赖[31]。李天慧等[12]将 DRSN 与双向 LSTM 混合，构建了面向电力信息系统网络流量的异常检测模型，在性能优于 LSTM、门控循环单元单元网络（GRU）、CNN 等比较模型。

3.2 长短期记忆网络

长短期记忆网络（LSTM）是一种特殊的循环神经网（RNN），其原理如图 4 所示。RNN 具备循环连接结构，网络参数权重 w 的更新不仅依赖当前时刻 t 的数据 x_t 输入，还受到 t 时刻之前隐藏状态的影响。LSTM 采用了 RNN 基础结构，基于 t 时刻的输入 x_t 和 $t-1$ 时刻的隐藏状态 h_{t-1} 来计算 t 时刻的输出 y_t 和隐藏状态 h_t。与 RNN 不同的是，LSTM 在隐藏层中添加了门控结构（见图 4 红框部分）将短时间记忆和长时间记忆结合起来，在一定意义上克服了梯度消失或梯度爆炸的问题，能够学习工控网络流量数据周期性的规律，并且容易识别由多个数据包共同作用引起的攻击类型。

3.3 基于混合模型的流量异常检测

基于 DRSN-LSTM 的网络流量异常检测流程如图 5 所示。首先将采集的原始水利工控网络流量经过一系列预处理之后，形成平衡数据集并划分为训练集和测试集，将测试集输入到 DSRN 模块进行空间特征提取，为了进一步提取数据中包含的时间相关信息，经过 DRSN 模块处理的数据将被传递给 LSTM 模块进行长时间相关特征提取，最终提取完成的特征被送至

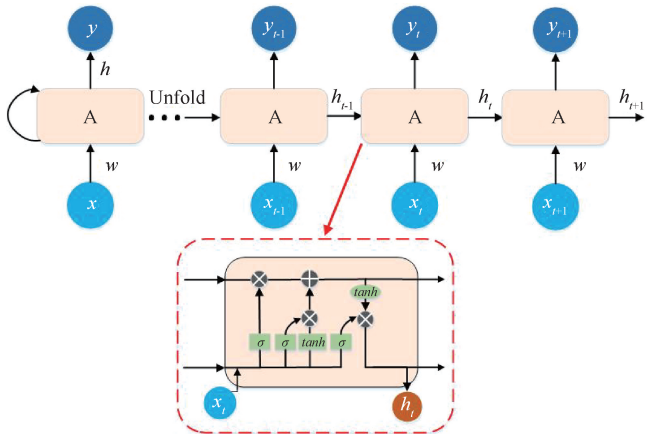

图 4 LSTM 原理

Fig. 4 LSTM principle

Softmax 分类器进行检测并获得分类结果，配合使用 Adam 优化器优化模型参数。根据研究结果，卷积核大小均为 3×3，步长为 2，LSTM 的学习率设置为 0.005，隐藏神经元数设置为 32，迭代次数设置为 100 次，批量大小设置为 64。

4 研究设计

4.1 数据来源

本文的研究场景为江苏省秦淮河武定门闸站，其水利工控系统的通信网络为以太网，利用网络流量分析平台（H3C SeerAnalyzer-NPA），对水利工控网内的网络流量进行抓包分析，数据集来自 2023 年江苏省水利厅下属水利工程管理单位网络攻防演练期间日常管理和网络攻击所产生的网络流量数据。

4.2 数据集构建

采集到的网络流量数据集经过清洗后共有 36 075 条数据，其中正常的数据有 26 334 条，非法遥控攻击占比 0.97%、字段组合攻击占比 9.49%、虚假数据注入攻击占比 1.94%、拒绝服务攻击数据占比 14.60%，异常数据的比例占 27%左右，非法遥控攻击和虚假数据注入攻击这两项数据均属于不平衡数据。经过数字化、归一化等预处理后，采用 ICGAN 模型进行平衡处理，其中正常流量需达到 50%的占比，保证异常检测模型能够学习到足够多的正常流量分布特征，降低误警率；

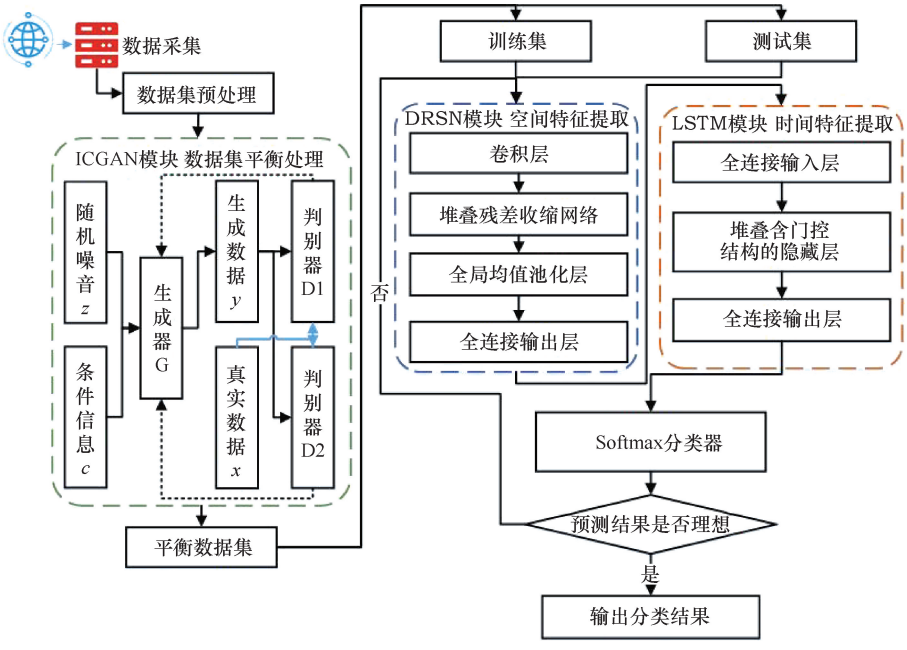

图 5 基于 DRSN-LSTM 的网络流量异常检测流程

Fig. 5 Network traffic anomaly detection process based on DRSN-LSTM

对于异常流量按照数量 3 750 进行分类，对数量超过 3 750 条的数据类别进行欠采样；对于数量不足 3 750 条的数据类别利用 ICGAN 模型进行过采样，并合并各类数据。在此基础上统计网络流量的属性特征，构建高维特征向量。最后抽取其中 70%的数据量作为训练数据集，剩下的 30%作为测试数据集。

4.3 验证方法

4.3.1 数据集质量评价方法

对采用 ICGAN 模型生成的数据集，利用余弦相似度(CS)对其质量进行评价，原理公式为

$$CS(\boldsymbol{A},\ \boldsymbol{B}) = \frac{\boldsymbol{A}\cdot\boldsymbol{B}}{\boldsymbol{A}\times\boldsymbol{B}} = \frac{\sum_{i=1}^{n}\boldsymbol{A}_i\times\boldsymbol{B}_i}{\sqrt{\sum_{i=1}^{n}\boldsymbol{A}_i^2}\times\sqrt{\sum_{i=1}^{n}\boldsymbol{B}_i^2}} \tag{5}$$

式中，$\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为两个向量；n 为向量的维度；$CS(\boldsymbol{A},\ \boldsymbol{B})$为两个向量的余弦相似度。

对于网络流量数据集，对于类别为 u 和 v 的数据集可以表示为

$$\boldsymbol{X}^u = [\boldsymbol{x}_1^u,\ \boldsymbol{x}_2^u,\ \boldsymbol{x}_3^u\cdots\boldsymbol{x}_{m^u}^u] \tag{6}$$

$$\boldsymbol{X}^v = [\boldsymbol{x}_1^v,\ \boldsymbol{x}_2^v,\ \boldsymbol{x}_3^v\cdots\boldsymbol{x}_{m^v}^v] \tag{7}$$

式中，m^u 和 m^v 分别为类别为 u 和 v 的数据集的数量；$\boldsymbol{x}$ 为数据向量，$\boldsymbol{x}=(x_1,\ x_2,\ x_3,\ \cdots x_n)$具有 n 维特征。

两类数据集之间的余弦相似度矩阵，可表示为

$$\boldsymbol{M}(CS(\boldsymbol{X}^u,\ \boldsymbol{X}^v)) = \begin{pmatrix} \mathrm{CS}(\boldsymbol{x}_1^u,\ \boldsymbol{x}_1^v), & \mathrm{CS}(\boldsymbol{x}_1^u,\ \boldsymbol{x}_2^v) & \cdots & \mathrm{CS}(\boldsymbol{x}_1^u,\ \cdots\boldsymbol{x}_{m^v}^u) \\ \mathrm{CS}(\boldsymbol{x}_2^u,\ \boldsymbol{x}_1^v), & \mathrm{CS}(\boldsymbol{x}_2^u,\ \boldsymbol{x}_2^v) & \cdots & \mathrm{CS}(\boldsymbol{x}_2^u,\ \cdots\boldsymbol{x}_{m^v}^u) \\ \cdots & & & \cdots \\ \mathrm{CS}(\boldsymbol{x}_{m^u}^u,\ \boldsymbol{x}_1^v), & \mathrm{CS}(\boldsymbol{x}_{m^u}^u,\ \boldsymbol{x}_2^v) & \cdots & \mathrm{CS}(\boldsymbol{x}_{m^u}^u,\ \boldsymbol{x}_{m^v}^u) \end{pmatrix} \tag{8}$$

矩阵 $\boldsymbol{M}$ 的大小为 $m^u\times m^v$，其元素值为 $\boldsymbol{X}^u$ 和 $\boldsymbol{X}^v$ 任意两个数据向量之间的余弦相似度值，通过对该矩阵求平均值，可以得到两类数据集之间的平均余弦相似度值(ACS)。

4.3.2 分类精度评价方法

基于深度学习的水利工控网络流量异常检测方法旨在精确识别网络中的异常流量数据并进行分类，提升网络流量检测的准确率，检测精度评价指标主要包括：准确率(ACC)、精准率(P)、召回率(R)和 $F1$ 值[26]，各指标定义为

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

$$P = \frac{TP}{TP + FP} \tag{10}$$

$$R = \frac{TP}{TP + FN} \tag{11}$$

$$F1 = \frac{2\times P\times R}{P + R} \tag{12}$$

式中，TP 为分类正确的正常数据量；TN 为分类正确的异常数据量；FP 为分类错误的正常数据量；FN 为分类错误的异常数据量。

在多分类问题中，ACC 为整体的准确率，而 P、R 和 $F1$ 则是针对每一类的分类效果的评价指标。

5 结果与分析

5.1 ICGAN 训练过程

生成器和判别器目标函数的收敛速度是衡量 CGAN 及其变体模型运行稳定性和训练难易程度的指标之一，较快的收敛速度和较小的收敛范围表明模型能够较快完成训练，性能更佳。CGAN 和 ICGAN 模型对抗训练过程中目标函数的变化情况如图 6 所示，两种模型刚开始训练时目标函数值波动较大，经训练后减小且趋于平缓，生成器初始目标值较大，判别器初始目标函数值较小，在训练过程中分别出现下降和上涨，而后逐渐平缓的趋势，体现了生成器和判别器训练过程中的对抗特征。从图 6(a)可以看出随着训练次数的增加，CGAN 目标函数的波动范围很大，表明生成器和判别器对抗激烈，模型难以训练。通过对 CGAN 进行改进，采用双判别器模式后，利用目标函数[式(3)]可以更好地引导生成器的训练，如图 6(b)所示，随着训练次数的增加，ICGAN 的生成器和判别器的目标函数能够较快达到收敛的状态，在迭代次数为 100～200 之间时，偶尔有较大的波动幅度，在迭代次数到达 250 之后，目标函数曲线呈现出较为稳定的状态，仅在很小的范围内波动，表明 ICGAN 更易训练，模型运行更加稳定。

5.2 数据集质量评价

对秦淮河武定门闸站的水利工控网络流量数据集进行平衡处理后，正常流量数据为 15 000(占 50%)，四类异常数据均为 3750 条(各占 12.25%)，其中非法遥控攻击、字段组合攻击和虚假数据注入攻击三类数据采用 ICGAN 模型进行了过采样，提升了在总体数据集中所占的比例。为验证 ICGAN 模型生成的数据质量，对占 30%的测试数据分别采用余弦相似度和分类测试表现进行评价。

(1)余弦相似度评价。采用余弦相似度评价生

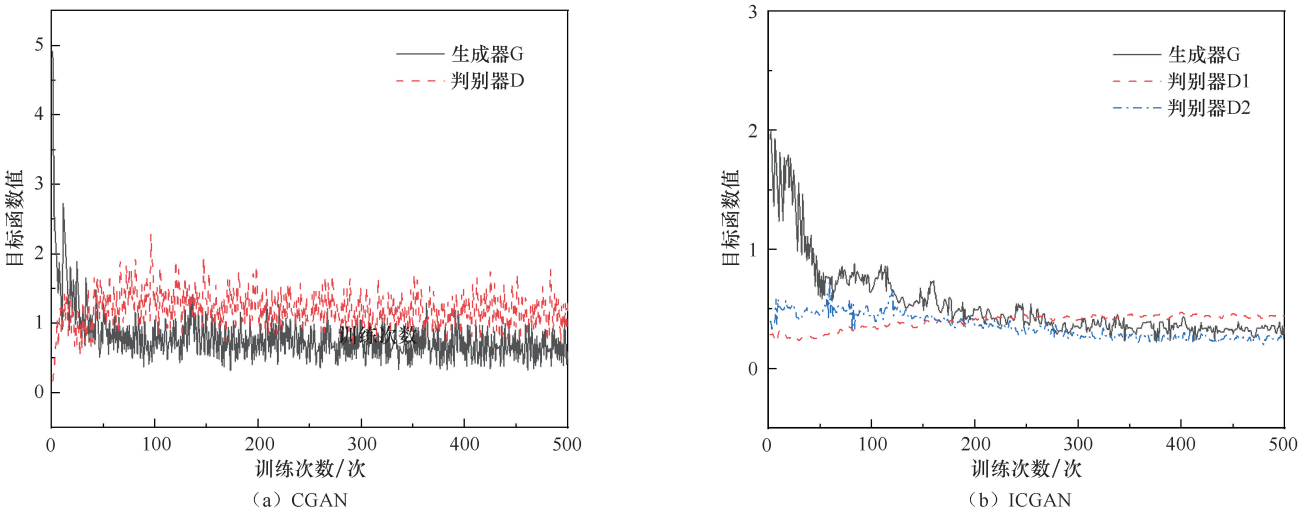

图 6　CGAN 和 ICGAN 的目标函数曲线

Fig. 6　Objective function curves of CGAN and ICGAN

成数据与真实数据的接近程度，原始数据集中各个类别流量数据内部和相互之间的平均余弦相似度值（*ACS*）如图 7 所示。可以看出对于同类别数据的内部之间相似度较高（对角线部分），明显区别于不同类别的数据相互之间的相似度。

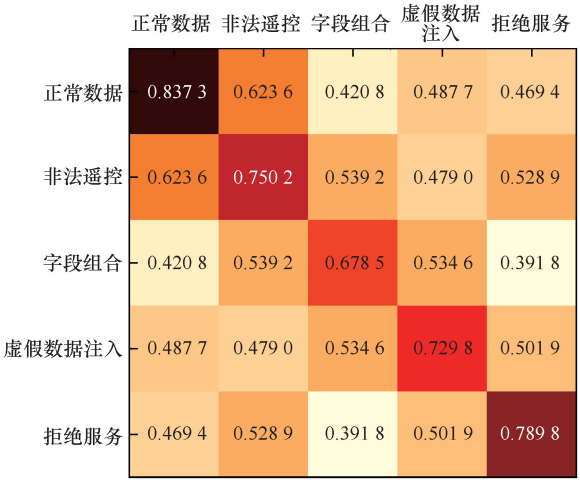

图 7　原始数据集中不同类别数据之间的 *ACS* 值

Fig. 7　*ACS* values of different data types in the original dataset

原始与生成数据集中相同类别之间的 *ACS* 值如表 1 所列。可以看到原始和生成数据集中同类别数据的 *ACS* 值比其他类别数据更加接近，采用 ICGAN 模型生成的数据能够较好地逼近真实数据。

（2）分类测试表现评价。为了更加直观地评价生成数据的质量，采用 DRSN-LSTM 模型分别在原始和平衡数据集上进行训练，DRSN-LSTM 在不同测试集上的表现如表 2 所列。利用原始数据集对模型进行训练后，测试的总体准确率（*ACC*）为 94. 27%，对于正常数据的检测精准率（*P*）和召回率（*R*）分别达到了 96. 99%和 97. 93%，说明检测模型能够较好地分辨正常流量和异常流量；但是对于非法遥控攻击和虚假数据注入攻击，*P* 值分别为 58. 33%和 79. 68%，*R* 值分别为 61. 82%和 63. 95%，说明检测模型对于异常流量的分类精度较低，主要原因在于是原始数据集是不平衡数据集，其中正常流量占整个数据集的 70%以上，异常流量中的非法遥控攻击和虚假数据注入攻击分别仅占 0. 97%和 1. 94%，导致在实际训练时，检

表 1　网络流量的数据类型

Table 1　Data types of network traffic

流量类别	原始数据集中同类别数据之间的 *ACS* 值	原始数据集和生成数据集中同类别数据之间的 *ACS* 值
非法遥控攻击	0. 750 2	0. 747 6
字段组合攻击	0. 678 5	0. 669 2
虚假数据注入攻击	0. 729 8	0. 728 5

表 2　DRSN-LSTM 在原始和平衡数据集上的分类测试表现

Table 2　Classification test performance of DRSN-LSTM on original and balanced datasets

数据集类型	数据类别	*ACC*/%	*P*/%	*R*/%	F_1/%
原始数据集	正常数据	94. 27	96. 99	97. 93	97. 46
	非法遥控攻击		58. 33	61. 82	60. 03
	字段组合攻击		87. 13	83. 64	85. 35
	虚假数据注入攻击		79. 68	63. 95	70. 95
	拒绝服务攻击		89. 10	89. 09	89. 09
平衡数据集	正常数据	98. 76	99. 22	99. 69	99. 46
	非法遥控攻击		98. 39	97. 69	98. 04
	字段组合攻击		97. 70	98. 04	97. 87
	虚假数据注入攻击		99. 18	97. 16	98. 16
	拒绝服务攻击		97. 88	98. 40	98. 14

测模型主要学习的是正常流量的特征，并且在根据损失函数更新模型参数时，异常流量的影响权重较小。

采用 ICGAN 模型对原始数据集进行平衡处理后，降低了正常流量在数据集中的占比，并提高了各类异常流量的数据量，检测模型能够有效的学习各类流量的数据特征，因此在 *ACC*、*P*、*R* 和 F_1 值上有更好的表现，说明经过平衡化处理的数据集能够有效提高检测模型分类精度。

为了比较 ICGAN 与传统数据过采样方法的性能差异，选取了随机过采样、SMOTE 和改进的边界 SMOTE(Border-line SMOTE)模型作为对比。利用不同的模型分别对原始数据集进行平衡化处理，然后利用处理后的数据集对 DRSN-LSTM 进行训练，测试结果如表 3 所列。结果表明，相比于原始数据集，使用不同的过采样方法构建的平衡数据集，所训练的 DRSN-LSTM 模型在分类精度的评价指标上有明显提升，而采用 ICGAN 模型的数据集，其评价指标均明显优于其他方法。

表 3　DRSN-LSTM 在不同平衡数据集上的分类测试表现

Table 3　Classification test performance of DRSN-LSTM on different balanced dataset

生成数据集的方式	*ACC*/%	*P*/%	*R*/%	F_1/%
随机过采样	95.58	95.43	95.29	95.36
SMOTE	96.38	96.12	96.32	96.22
Border-line SMOTE	96.91	96.54	96.77	96.65
ICGAN	98.76	98.48	98.19	98.33

注：表中 *P*、*R* 和 *F*1 值为多分类评价指标的平均值。

5.3　检测模型性能比较

将常见的网络流量异常检测模型，如随机森林(RF)、支持向量机(SVM)、1 D-CNN、LSTM、CNN-LSTM 和本文的 DRSN-LSTM 模型分别在平衡数据集上进行训练，各模型的 *ACC*、*P*、*R* 和 F_1 值检测精度评价指标如图 8 所示，采用混淆矩阵展示了各模型的分类检测结果(见图 9)，DRSN-LSTM 的网络流量异常检测的准确率达到了 98.76%，其中正常数据分类的 *P*、*R* 和 F_1 值，分别达到了 99.22%、99.69% 和 99.46%，各分类的平均 *P*、*R* 和 F_1 值分别为 98.48%、98.19% 和 98.33%，评价指标方面均优于其他模型；从混淆矩阵的分类结果来看 DRSN-LSTM 的误报率和漏报率方面也低于其他模型，验证了本文方法对水利工控网络流量进行异常检测的有效性。

从图 8 可以看出传统 1 D-CNN、LSTM 和 CNN-LSTM 也具有较好的检测结果，如果为了进一步提高对数据高维特征的利用率并改善检测效果，还需要对

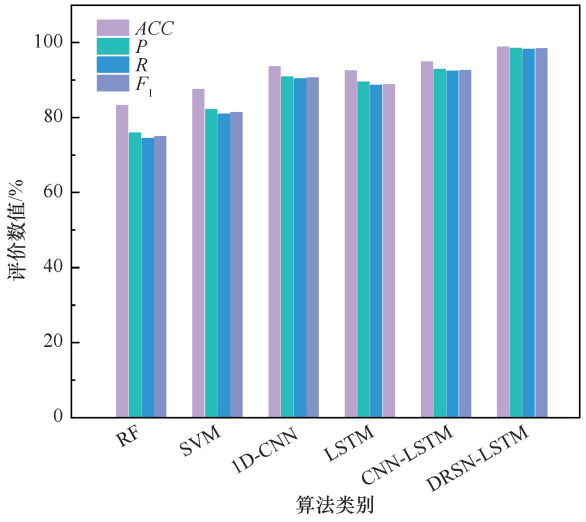

图 8　不同模型的检测精度评价指标值

Fig. 8　Evaluation indicator values of detection accuracy of different models

数据进行降维和特征选择处理。本文方法中 DRSN 模块作为 CNN 的变体，引入了软阈值操作，可以对不同的特征图选择不同阈值，从而能够自适应保留重要的特征信息，相比于改进前的 CNN-LSTM 模型，其总体准确率提高了 3.91%。

各模型训练和测试所耗费时间如表 4 所列。其中 1 D-CNN、LSTM、CNN-LSTM 和 DRSN-LSTM 训练时间是在训练集上收敛的平均值，测试时间是在测试集上测试的平均值。其中 LSTM 训练耗时最长，对于混合模型来说，时序特征提取模块对整体运行速度影响较大。CNN-LSTM 模型的训练耗时明显高于 DRSN-LSTM，原因在于 DRSN 模块利用收缩激励网络(SENet)能够有效提升高维数据特征的提取能力，加快模型训练的收敛速度。测试时间的结果表明，尽管在测试效率上 DRSN-LSTM 没有表现出优势，但就分类准确率而言，其效果高于传统的 CNN-LSTM 分类模型。

表 4　不同模型的训练与测试时间　s

Table 4　Training and testing time of different models　s

时　间	RF	SVM	1D-CNN	LSTM	CNN-LSTM	DRSN-LSTM
训练时间	64.32	95.26	323.62	2 048.71	1 320.54	505.23
测试时间	1.13	5.45	0.84	98.33	2.04	2.32

5.4　方法通用性研究

为了验证本文方法的通用性，采用另一个常见的工控网络流量数据集，该数据集来源于 2014 年密西西比州立大学实验室的天然气场景下的测试数据[32]，

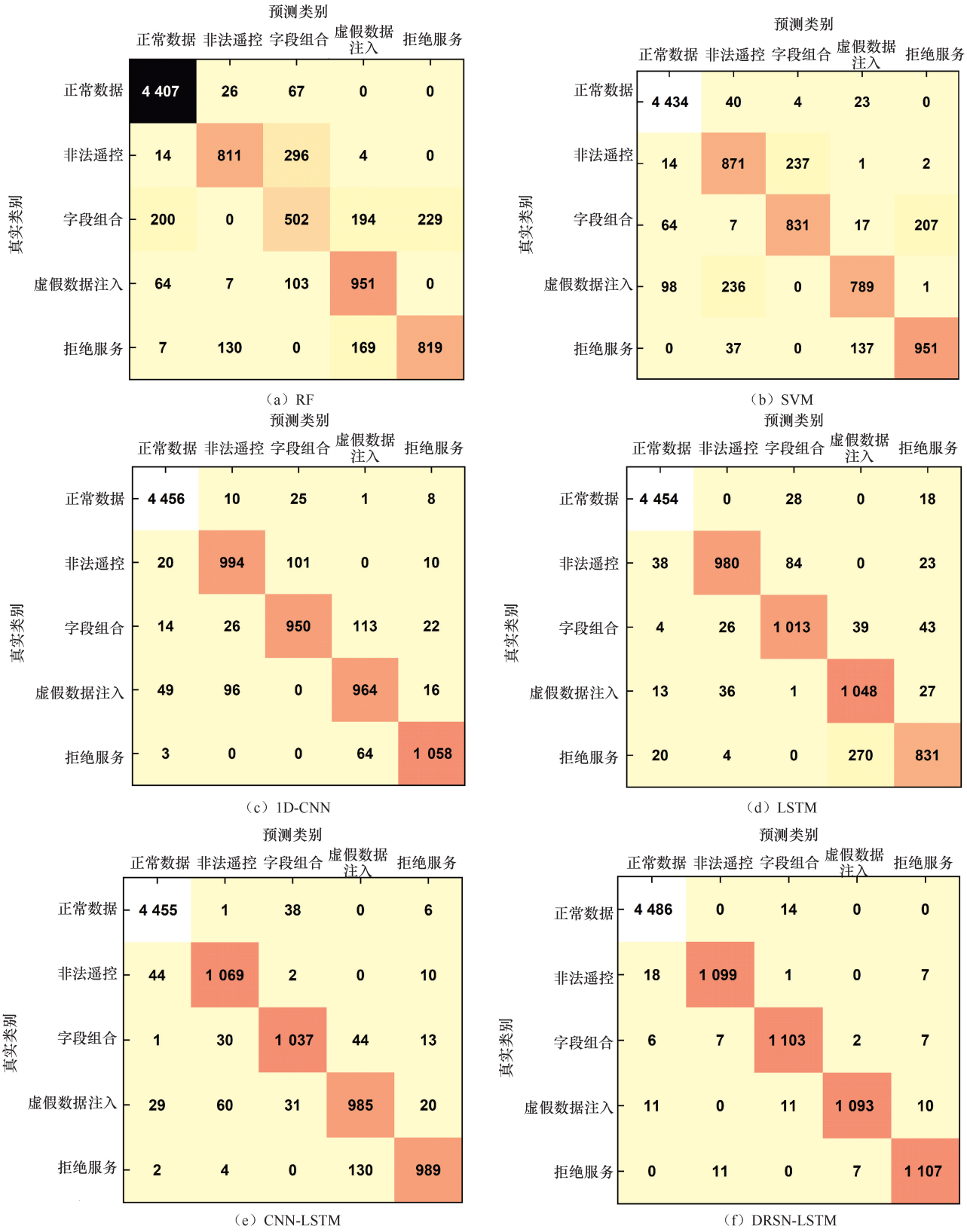

图 9　不同模型的分类检测结果混淆矩阵

Fig. 9　Confusion matrix of classification detection results of different models

是真实天然气管道和储水罐系统等场景中的工控网络流量数据。数据集共有 97 019 条样本，包括来自正常操作产生的 61 156 条真实流量数据和针对 SCADA 系统进行攻击的 35 863 条异常流量数据，其中异常

流量数据主要包含：侦察注入攻击、恶意响应注入攻击、恶意命令注入攻击和拒绝服务攻击，按照水利工控网络流量数据集中各类样本数量和比例，抽取训练集和测试集。

DRSN-LSTM 模型在储水系统网络流量数据集上的 *ACC*、*P*、*R* 和 *F*1 值检测精度评价指标，以及训练和测试所耗费时间如表 5 所列，该模型在非水利工控流量数据集上同样具有较好的性能，总体准确率达到 97.84%，五个类别的精准率分别为 98.74%、94.90%、98.62%、92.12%和 97.78%，同时各个类别的 *R* 和 *F*1 指标也都达到 90%以上，测试集的检测时间为 2.15 s，验证表明 DRSN-LSTM 对本文涉及的两个数据集的检测性能都较好，检测方法具有较好的通用性。

5.5 结果讨论

本文提出的水利工控网络流量异常检测方法改善了基于生成对抗网络的数据生成模型难以训练，数据生成模式难以控制；流量检测模型训练过程存在梯度消失和梯度爆炸等现象；数据时空特征利用不足，高维数据特征难以提取等方面的问题，可以在不影响检测效率的前提下，提高模型检测的准确率。利用 ICGAN 来解决水利工控网络流量数据集不平衡的问题，与随机过采样、SMOTE 和 Border-line SMOTE 传统数据过采样方法相比，ICGAN 生成的样本兼顾真实性和类别属性，且对抗训练过程稳定，能够有效提高少数类异常流量的检测效果。采用平衡后的数据集验证 DRSN-LSTM 混合模型的性能，与 RF、SVM、1D-CNN、LSTM 和 CNN-LSTM 等模型相比，在整体性能指标上本文所提模型优于其他比较模型。利用真实天然气管道和储水罐系统等场景中的工控网络流量数据集进一步验证 DRSN-LSTM 模型的性能，在检测准确性、训练时间和检测效率上仍保持优异的性能，表明本方法具有较好的通用性。

在同样类似的案例研究中，李茹[2]基于深度信息网络(DBN)对水利泵站工控网络流量数据进行检测，并与 PCA 降维的 DNN 模型、自编码降维的 DNN 模型进行比较，研究表明重要特征的提取对于提高检测模型的性能具有重要意义，但是由于没有对数据集进行平衡处理，仅采用有限的数据集进行训练，训练样本中的异常流量的数据并不充足，分类准确率为 96.52%，还有进一步提升的空间。李金娜[33]利用与本文同样的天然气场景下的工控流量数据集，验证其 SVM-RF-GSKNN 混合模型的性能，由于该研究采用的是浅层学习模型，模型复杂度低，特征提取能力相对有限，在同等训练数据的前提下，训练时间为 30 s 左右，但是漏报率较高，达到 5.74%。与以上研究相比，本研究通过数据集平衡处理，提高了少数类异常流量的样本数，模型经过训练后避免了检测结果偏向于多数类；同时具有较好的时空特征提取能力，结合 SENet 网络能够自动提取重要特征，进一步提高了模型的检测性能。由于本研究采用的数据集是通过自动化脚本模拟得到，相较于真实的水利工控网络流量数据还是存在一定的差别，后续可以使用真实工控网络流量数据进行验证。

在后续工作中，可以从以下几个方面开展进一步的研究：(1)将该方法应用于其他行业或领域的业务网络，根据场景特征进一步优化模型，扩大应用范围；(2)DRSN 中的软阈值函数由于渐进性不佳，在设置阈值时易将区间范围内的数值强制置“0”或“1”，从而增大特征阈值的偏差，后继可以研究可变软阈值函数对其进行替换，进一步提高模型性能；(3)深度学习类模型的训练效果比较依赖前期的特征构造，本文特征构造主要采用计数、百分比和均值统计类指标，后继可以进一步探索其他指标的应用效果。

6 结 论

提高网络流量异常检测的能力对于保障水利工控网络安全具有重要意义。针对实际采集的网络流量数据集不平衡的问题，提出了一种 ICGAN 的数据过采样模型，利用该模型对原始数据集进行了平衡处理，经验证平衡后的数据集对于开展网络流量检测的多分

表 5 DRSN-LSTM 在储水系统网络流量数据集上的分类测试表现

Table 5 Classification test performance of DRSN-LSTM on network flow dataset of water storage system

数据集类型	数据类别	*ACC*/%	*P*/%	*R*/%	*F*1/%	训练时间/s	测试时间/s
平衡数据集	正常数据	97.84	98.74	99.32	99.03	569.73	2.15
	侦察注入攻击		94.90	95.01	94.96		
	恶意响应注入攻击		98.62	98.27	98.44		
	恶意命令注入攻击		92.12	91.57	91.85		
	拒绝服务攻击		97.78	98.89	98.33		

类任务效果更好，分类精度更高；针对网络流量的时空特征，提出基于 DRSN 和 LSTM 的混合深度学习模型 DRSN-LSTM，DRSN 负责提取局部空间特征，其残差网络能够有效缓解网络退化现象，收缩激励网络可以有效提取重要特征信息；LSTM 负责提取长序列时间特征，通过两种不同方法将网络流量的时空信息进行提取并学习，从而实现对网络流量的异常检测。通过两种工控流量数据集的验证，证明本文的方法在准确率(*ACC*)，精准率(*P*)，召回率(*R*)、*F*1 值等网络流量分类评价指标上均优于传统的随机森林(RF)、支持向量机(SVM)、1 D-CNN、LSTM 和 CNN-LSTM 模型。

参考文献(References)：

[1] 谢遵党. 水利水电工程数字设计工厂建设构想与实践[J]. 水利水电技术(中英文)，2023，54(2)：60-72.
XIE Zundang. Concept and practice of digitalized design factory for water conservancy and hydropower engineering[J]. Water Resources and Hydropower Engineering，2023，54(2)：60-72.

[2] 李茹. 基于深度神经网络的水利泵站工控安全研究[D]. 镇江：江苏科技大学，2019.
LI R. Research on Industrial Control Safety of Water Pump Station Based on Deep Neural Network [D]. Zhenjiang：Jiangsu University of Science and Technology，2019.

[3] YASAMI Y，MOZAFFARI S P. A novel unsupervised classification approach for network anomaly detection by k-Means clustering and ID3 decision tree learning methods[J]. Journal of supercomputing，2010，53(1)：231-245.

[4] YEUNG D Y，DING Y. Host-based intrusion detection using dynamic and static behavioral models [J]. Pattern Recognition，2003，36(1)：229-243.

[5] 刘奕，李建华，张一瑶，等. 基于特征属性信息熵的网络异常流量检测方法[J]. 信息网络安全，2021(2)：78-86.
LIU Y，LI J H，ZHANG Y D，et al. Abnormal network traffic detection method based on feature attribute information entropy [J]. Information Network Security，2021(2)：78-86.

[6] 陈戈. 基于机器学习的网络异常流量检测方法研究[D]. 沈阳：沈阳理工大学，2022.
CHEN G. Research on Abnormal Network Traffic Detection Method Based on Machine Learning[D]. Shenyang：Shenyang Polytechnic University，2022.

[7] SARKER I H. Deep cybersecurity：A comprehensive overview from neural network and deep learning perspective [J]. SN Computer Science，2021，2(3)：154.

[8] MAHMOUD A，SHAHRAKI A，AMIR T. Deep learning for network traffic monitoring and analysis (NTMA)：A survey[J]. Computer Communications，2021，170(15)：19-41.

[9] SHONE N，TRAB N N，PHAI D V，et al. A deep learning approach to network intrusion detection[J]. IEEE Trans. Emerging Topics in Comput. Intellig.，2018，2(1)：41-50.

[10] 张昊，张小雨，张振友，等. 基于深度学习的入侵检测模型综述[J]. 计算机工程与应用，2022，58(6)：17-28.
ZHANG H，ZHANG X Y，ZHANG Z Y，et al. Overview of intrusion detection models based on deep learning[J]. Computer Engineering and Applications，2022，58(6)：17-28.

[11] 武建. 人工智能技术在水利行业中的应用实践与展望[J]. 水利发展研究，2024，24(8)：44-49.
WU Jian. Application practice and prospect of artificial intelligence in water sector[J]. Water Resources Development Research，2024，24(8)：44-49.

[12] 李天慧，谢云澄，车荣花，等. 基于 DRSN-BiLSTM 的电力信息网络入侵检测模型[J]. 电力信息与通信技术，2023，21(9)：30-37.
LI T H，XIE Y C，CHE R H，et al. Intrusion detection model of power information network based on DRSN-BiLSTM [J]. Electric Power Information and Communication Technology，2019，21(9)：30-37.

[13] 燕昊昊，韩国栋，黄雅静，等. 非平衡网络流量识别方法[J]. 计算机应用，2018，38(1)：20-25.
YAN B H，HAN G D，HUANG Y J，et al. Non-equilibrium network traffic identification method[J]. Journal of Computer Applications，2018，38(1)：20-25.

[14] PACHECO F，EXPÓSITO E，GINESTE M. A framework to classify heterogeneous internet traffic with machine learning and deep learning techniques for satellite communications [J]. Computer Networks，2020，173：107213.

[15] YANG Y，ZHENG K，WU C，et al. Improving the classification effectiveness of intrusion detection by using improved conditional variational auto encoder and deep neural network[J]. Sensors，2019，19：2528.

[16] YU Y，TANG B，LINR，et al. CWGAN：Conditional wasserstein generative adversarial nets for fault data generation[C]// IEEE 2019 IEEE International Conference on Robotics and Biomimetics (ROBIO). Dali：IEEE. 2019. DOI：10.1109/ROBIO49542.2019.8961501.

[17] MA Z，LI J，SONG Y，et al. Network intrusion detection method based on FCWGAN and BiLSTM[J]. Computational Intelligence and Neuroscience，2022. DOl：org/10.1155/2022/6591140.

[18] KUMAR R，ZHANG X，KHAN R U，et al. Malicious code detection based on image processing using deep learning [C] //ICCAI. Proceedings of the 2018 International Conference on Computing and Artificial Intelligence. New York，U.S.ICCAI，2018. DOI：10.1145/3194452.3194459.

[19] 贾凡，孔令智. 基于卷积神经网络的入侵检测算法[J]. 北京理工大学学报，2017，37(12)：1271-1275.
JIA F，KONG L Z. Intrusion detection algorithm based on convolutional neural networks [J]. Journal of Beijing Institute of Technology，2017，37(12)：1271-1275.

[20] 常晓燕. 基于深度神经网络的多维时间序列异常检测方法研究[D]. 上海：东华大学，2022.
CHANG X Y. Research on Anomaly Detection Method of Multi-Dimensional Time Series Based on Deep Neural Network [D].

Shanghai：Donghua University，2022.

[21] SHERAZ N，YASIR S，SHEHZAD K，et al. Enhanced network anomaly detection based on deep neural networks[J]. IEEE Access，2018，6：48231-48246.

[22] ALEESA AM，YOUNIS M，MOHAMMED A A，et al. Deep-intrusion detection system with enhanced UNSW-NB15 dataset based on deep learning techniques[J]. Journal of Engineering Science and Technology，2021，16(1)：711-727.

[23] 戴甦，马媛，吴亚男. 基于事件驱动的数字孪生太浦河“四预”业务应用[J]. 水利发展研究，2024，24(9)：28-32.
DAI Su，MA Yuan，WU Yanan. Business application of “forecasting，early warning，rehearsal and pre-planning” in digital twin Taipu River[J]. Water Resources Development Research，2024，24(9)：28-32.

[24] 罗旋，李永忠. Modbus TCP 安全协议的研究与设计[J]. 数据采集与处理，2019，34(6)：1110-1117.
LUO X，LI Y Z. Research and design of Modbus TCP security protocol[J]. Data Acquisition and Processing，2019，34(6)：1110-1117.

[25] 朱智燊. 基于 Modbus-TCP 协议的工控数据安全技术研究[D]. 广州：广东工业大学，2020.
ZHU Z S. Research on Industrial Control Data Security Technology Based on Modbus-TCP Protocol[D]. Guangzhou：Guangdong University of Technology，2020.

[26] 赵忠斌. 基于机器学习的网络异常流量检测系统研究[D]. 北京：中国人民公安大学，2023.
ZHAO Z B. Research on Abnormal Network Traffic Detection System Based on Machine Learning[D]. Beijing：People's Public Security University of China，2023.

[27] MIRZA M，OSINDERO S. Conditional generative adversarial nets[J]. Computer Science，2014：2672-2680.

[28] 王文博，刘绚，林海，等. 基于深度学习的电力工控流量应用层报文异常检测[J]. 电力系统自动化，2023，47(11)：69-76.
WANG W B，LIU X，LIN H，et al. Application layer message anomaly detection of electric power industrial control traffic based on deep learning[J]. Automation of Electric Power Systems，2023，47(11)：69-76.

[29] ZHAO M H，ZHONG S S，FU X Y，et al. Deep residual shrinkage networks for fault diagnosis[J]. 2020，16(7)：4681-4690.

[30] HE K，ZHANG X，REN S，et al. Deep residual learning for image recognition[J]. Institute of Electrical and Electronics Engineers，2016：770-778.

[31] 童轶之. 结合残差收缩和长短期记忆网络的轴承故障诊断[D]. 杭州：浙江理工大学，2022.
TONG Y Z. Bearing Fault Diagnosis Via Combining Deep Residual Shrinkage and Long Short-Term Memory Network[D]. Hangzhou：Zhejiang Sci-Tech University，2022.

[32] MORRIS T，GAO W. Industrial control system traffic data sets for intrusion detection research[C]//ICCIP. Critical Infrastructure Protection Ⅷ. Berlin，Germany：ICCIP，2014：65-76.

[33] 李金娜. 基于机器学习的工控流量入侵检测技术研究[D]. 长沙：湖南大学，2022.
LI J N. Research on Intrusion Detection Technology of Industrial Control Traffic Based on Machine Learning[D]. Changsha：Hunan University，2022.

（责任编辑　王　璐）