

doi:10.16112/j.cnki.53-1223/n.2026.01.202505290001

引用格式: 郇战, 王文韬, 王澄, 等. 基于特征筛选与数据增强的图卷积神经网络在 TSN 网络配置检测中的应用[J]. 昆明理工大学学报(自然科学版), 2026, 51(1): 137-145.

Citation: HUAN Zhan, WANG Wentao, WANG Cheng, et al. Application of Graph Convolutional Network Based on Feature Selection and Data Augmentation in TSN Network Configuration Detection[J]. Journal of Kunming University of Science and Technology (Natural Science), 2026, 51(1): 137-145.

## 基于特征筛选与数据增强的图卷积神经网络 在 TSN 网络配置检测中的应用

郇战<sup>1</sup>, 王文韬<sup>1</sup>, 王澄<sup>2\*</sup>, 王毅<sup>2</sup>, 陈瑛<sup>1</sup>, 胡芬<sup>1</sup>

(1. 常州大学 王净微电子学院/集成电路产业学院, 江苏 常州 213164;

2. 常州大学 计算机科学与人工智能学院, 江苏 常州 213164)

**摘要:** 为了提升时间敏感网络(Time Sensitive Networking, TSN)网络配置检测的准确率, 特别是在数据不平衡条件下的分类性能, 提出一种基于特征筛选和条件表格生成对抗网络(Conditional Tabular Generative Adversarial Network, CTGAN)数据增强的图卷积神经网络(Graph Convolutional Network, GCN) TSN 网络配置检测模型. 首先通过计算互信息量(Mutual Information, MI)筛选得到强相关特征, 在此基础上使用 CTGAN 针对原始数据集不平衡问题进行数据增强, 最后构建 GCN 网络模型得到网络配置的分类结果. 计算机仿真表明, 使用 MI-CTGAN-GCN 模型进行网络配置的可行性预测可以提高对不平衡数据集的分类能力, 与现有检测算法相比, 模型分类准确率更高, 达到了 96.28%, 验证了该方法的可行性与优越性.

**关键词:** 时间敏感网络(TSN); 特征筛选; 互信息量; 生成对抗网络; 图卷积神经网络

中图分类号: TP183; TP393.11 文献标识码: A 文章编号: 1007-855X(2026)01-0137-09

### Application of Graph Convolutional Network Based on Feature Selection and Data Augmentation in TSN Network Configuration Detection

HUAN Zhan<sup>1</sup>, WANG Wentao<sup>1</sup>, WANG Cheng<sup>2\*</sup>, WANG Yi<sup>2</sup>, CHEN Ying<sup>1</sup>, HU Fen<sup>1</sup>

(1. Wangzheng School of Microelectronics/School of Integrated Circuits Industry, Changzhou University, Changzhou

213164, Jiangsu, China; 2. School of Computer Science and Artificial Intelligence,

Changzhou University, Changzhou 213164, Jiangsu, China)

**Abstract:** To improve the accuracy of Time-Sensitive Network (TSN) network configuration detection, especially under imbalanced data conditions, this paper proposes a Graph Convolutional Network (GCN) model based on feature selection and Conditional Tabular Generative Adversarial Network (CTGAN) data enhancement for TSN network configuration detection. First, strongly relevant features are selected by calculating Mutual Information (MI). Then, CTGAN is applied to address the issue of imbalanced data in the original dataset. Finally, a GCN model is constructed to obtain the classification results of network configurations. Computer simulations demon-

收稿日期: 2025-05-29. 基金项目: 国家自然科学基金青年项目(62201093); 江苏省研究生科研与实践创新计划项目(SJCX23\_1595).

作者简介: 郇战(1969-), 男, 硕士, 教授. 主要研究方向: 物联网和智能控制技术. E-mail: hzh@cczu.edu.cn

通信作者: 王澄(1991-), 男, 博士, 讲师. 主要研究方向: 工业物联网. E-mail: chengwang@cczu.edu.cn

strate that the MI - CTGAN - GCN model proposed in this paper enhances the classification capability for imbalanced datasets and achieves a higher accuracy rate compared to existing detection algorithms, with a classification accuracy of 96.28%, validating the feasibility and superiority of the proposed method.

**Keywords:** time - sensitive networking (TSN); feature selection; mutual information; generative adversarial network; graph convolutional network

## 0 引言

工业物联网 (IIoT) 等新一代信息通信技术的飞速发展,为制造业带来了巨大的机遇与挑战<sup>[1]</sup>,推动了新兴经济体中制造企业的转型.在这一背景下,实时通信能力变得至关重要<sup>[2]</sup>.传统以太网虽然具有良好的带宽和通用性<sup>[3]</sup>,但在某些服务质量 (QoS) 要求较高的领域,如工业自动化、汽车电子等<sup>[4]</sup>,其在实时性和确定性方面的局限性使其难以满足工业场景中严苛的通信要求<sup>[5]</sup>.为了应对这一问题,TSN 技术应运而生.TSN 是一系列协议的总称,它定义并实现了各种 QoS 策略的机制和协议,如使用优先级、流量整形、时间触发和帧抢占<sup>[6]</sup>.通过各类协议,TSN 旨在为传统以太网引入实时性和确定性<sup>[7]</sup>,确保网络中高优先级的流量能够按时传输.

IEEE 802.1 工作组为提高 TSN 配置的灵活性和便捷性,提出了一系列协议对网络配置和管理模型进行优化<sup>[8]</sup>.随着网络规模的扩大和复杂性的增加,在工业网络上部署大量传感器也使得 TSN 网络配置成为一项具有挑战性的任务<sup>[9]</sup>.TSN 网络配置的准确性会直接影响到整个系统的性能和稳定性.现有的 TSN 网络配置检测算法和模型仍有不足,在准确率上仍需要提高,且未考虑数据集不平衡性对分类结果的影响.

汽车平台的电子电气 (E/E) 架构需要设计空间探索 (Design Space Exploration, DSE) 算法来完成复杂嵌入式系统的设计和配置.DSE 算法所创建的大量候选解决方案使得性能评估成为一种计算密集型的工作,基于机器学习的方法在 TSN 网络配置检测方面是高效的,与传统可调度性分析相比快了 70 倍<sup>[10]</sup>.然而现有基于机器学习的方法并未在模型输入中充分融入网络拓扑的图结构信息,导致模型难以捕捉 TSN 网络配置在图结构上的关键特征,从而限制了分类准确率的提升.

本文提出一种新的模型构建方法,旨在优化 TSN 的网络配置检测.该方法首先采用 MI 指标对特征进行筛选,筛选出关联度较高的特征;然后使用 CTGAN<sup>[11]</sup> 算法生成少数类样本,从而解决数据集中的类别不平衡问题;接着通过 GCN 模型检测 TSN 网络配置,构建 MI - CTGAN - GCN 模型;最后本文将该模型与其他数据增强模型进行横向比较.

## 1 相关工作

### 1.1 网络配置

针对如何高效地进行 TSN 网络配置检测的问题,文献[10]采用了深度学习算法,使用图神经网络 (Graph Neural Network, GNN) 模型进行分类检测,该模型使用一个包含图信息以及组成部分之间相关关系的全局向量作为神经网络的输入,在应对训练集拓扑结构和流量模式变化方面表现出了较好的鲁棒性.文献[12]通过优化训练集的构建和模型架构,对基础的 GNN 模型进行了改进,改进后的模型对网络配置检测的分类准确率提高至 90.5%,但仍未达到理想水平.文献[13]采用特征优选和麻雀搜索算法优化支持向量机模型,对网络配置进行分类检测,分类准确率有所提高,但是未考虑不同拓扑结构对模型的影响.文献[14]分别使用硬投票集成与粒子群优化 (PSO) 算法优化后的 BPNN 模型对两种拓扑结构的网络配置进行分类检测,准确率进一步提高,但是未解决数据集不平衡性对准确率的影响.

### 1.2 可调度性分析

TSN 中使用调度策略的目的是在流量出队列进行链路传输时对其加以限制,从而提升数据传输的确定性.不同的策略在满足传输确定性需求 (如时间约束) 上有不同的权衡.本文使用以下 3 种调度策略<sup>[14]</sup>:

- 1) 先进先出调度 (FIFO): 只有一个流量类别,所有流的优先级都是相同的.
- 2) 手动分类的优先级调度 (Manual): 此方法被视作基于流的相对关键性进行尝试的基线解决方案.

流量会被分为 3 种类别,分别是音频流、视频流和命令与控制流. 3 种类别的流量具有不同的优先级,它们之间的关系如下:命令与控制流 > 音频流 > 视频流.

3) 具有 8 个优先级的简明优先级(Concise Priority, CP)调度(CP8): CP 是 RTaW - Pegase 中优先级分配算法的名称,可以将相同类型的流量分配给不同的流量类别,该算法实现了最优优先级分配.

上述调度解决方案属于使用静态优先级调度的 QoS 策略,它们可以通过仅提供 8 个优先级的基本 TSN 网络设备来实现. 这 3 种调度方案与标准的静态优先级调度分析是兼容的,即计算每个流的最坏情况遍历时间(Worst Case Traversal Time, WCTT),以此作为判别每个配置的可行性标签的依据. 本文针对 3 种调度方法分别计算每个流的 WCTT. 当且仅当每个流的 WCTT 小于系统对流的时间约束的时候,网络配置为可行,标签设置为 1; 否则,认为配置不可行,标签设置为 0.

## 2 基于 MI - CTGAN - GCN 模型的解决方案

本文从网络拓扑中收集含有流量以及链路各类信息的数据集,通过计算各类特征与标签之间的 MI 值对数据集进行特征筛选,寻找相关性较高的特征类别. 针对原始数据集不平衡的问题,使用 CTGAN 对不同调度方法中的少数类样本进行数据增强,生成与原始数据集中流量与链路等特征参数分布相近的样本,进一步提高模型的分能力,将增强后的数据集输入到 GCN 网络中进行节点特征与边特征的卷积聚合,构建 MI - CTGAN - GCN 模型.

### 2.1 特征筛选

网络中的流量由 3 种类型组成,分别是音频流、视频流、控制流,其特征如表 1 所示. 流的特征及其比例以及网络拓扑由一家汽车 OEM 提供. 实验过程中,流的数量会发生变化,但是流的比例作为流生成过程的固定参数. 本文考虑的是截止时间约束,即通过可调度性分析计算出的每个流的 WCTT 必须小于流的截止时间.

本文考虑的拓扑如图 1 所示<sup>[15]</sup>,该网络由 5 个交换机和 14 个节点,即电子控制单元(Electric Control Unit, ECU)组成. 除了 ECU12 到 switch3 的链路为 1 Gbps 外,所有链路的数据传输速率均为 100 Mbps. 原始数据集从汽车电子系统中采集得到. 在 ECU 互相传输数据的过程中,会产生待检测配置,配置由不同类型的流量流组成.

表 1 三种流量类型的特性<sup>[10]</sup>

Tab. 1 Characteristics of three traffic types<sup>[10]</sup>

流量类型	特性
音频流	①帧为 128 或 256 字节 ②周期为 1.25 ms ③截止时间为 5 或 10 ms ④比例:7/46
视频流	①ADAS + 视觉流 ②每 33 ms 有 30 个 1 500 字节的帧(30 FPS 视觉摄像头) ③每 33 ms 有 15 个 1 000 字节的帧(30 FPS ADAS 摄像头) ④截止时间为 10 ms(ADAS)或 30 ms(视觉) ⑤比例:7/46
命令与控制流	①帧为 53 到 300 字节 ②周期为 5 到 80 ms ③截止时间等于周期 ④比例:32/46

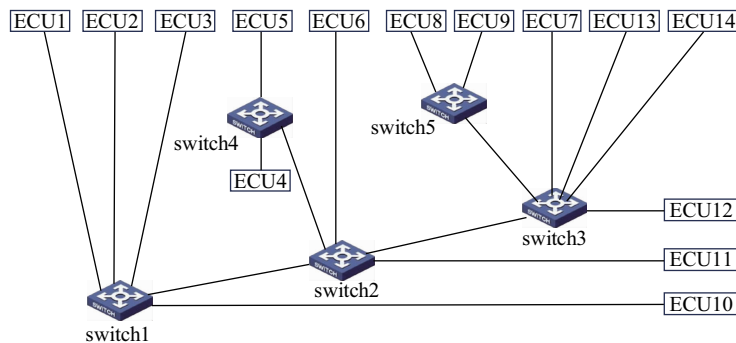


图 1 实验中使用的拓扑

Fig. 1 Topology used in the experiment

特征筛选可以从原始数据集中筛选出与标签强相关的特征,使得模型获得更强的分类能力. MI 值的

计算公式如式(1)所示:

$$I(X, Y) = \sum_{y \in Y} \sum_{x \in X} p(x, y) \log \left( \frac{p(x, y)}{p(x)p(y)} \right) \quad (1)$$

式中:  $p(x)$  和  $p(y)$  分别为变量  $X$  和  $Y$  的边缘概率分布函数,  $p(x, y)$  为  $X$  和  $Y$  的联合概率分布函数.

首先计算原始数据集中所有变量与标签之间的 MI 值, 然后筛选出一组最具有相关性的特征. 这些特征是在综合考虑与 FIFO、CP8 和 Manual 3 种调度方法可行性标签的相关性后得出的, 旨在挑选出与所有标签的相关性都尽可能大的特征. 所筛选特征的 MI 值如表 2 所示. 7 个特征分别表示网络拓扑中的总流量数、关键流数、音频流数、视频流数、链路最大负载, 执行 WCTT 所需的时间以及根据各条链路负载所计算得到的链路 Gini 系数.

由表 2 可知, 不同特征与不同调度方法可行性标签的 MI 值大小有较为明显的差距, 这表明它们之间的关联程度是不一样的.

总体来看, 所选特征与 CP8 标签的 MI 值最大, 表明二者之间的相关性程度最高. 在 Manual 标签中, 特征的 MI 值也都基本大于 0.2, 说明这些特征都与标签有较强的相关性. 而在 FIFO 标签中, 虽然整体相关性不如 CP8 和 Manual, 但是所选特征与 FIFO 相关性差异最小, 而且 MI 值也基本大于 0.1, 说明这些特征也与标签存在一定的关联性. 上述结果表明, 使用 MI 值选取此 7 种特征用于后续网络配置检测模型进行分类训练是有依据的.

## 2.2 MI-CTGAN-GCN 模型

利用 MI-CTGAN-GCN 模型进行网络配置检测的过程如图 2 所示.

MI-CTGAN-GCN 算法模型具体实现步骤如下:

(1) 对网络配置原始数据集进行数据预处理, 比较每个流的 WCTT 与系统对流的时间约束, 对配置可行性进行标签化处理, 同时基于 MI 值筛选出最具相关性的 7 个特征.

(2) 将原始配置数据集按照 0.85:0.15 的比例划分为测试集与训练集. 通过原始数据集中配置可行与不可行标签的比例计算每种调度方法中可行配置的百分比, 即  $f$  值, 将训练数据集中的少数类配置样本作为进行 CTGAN 数据增强的对象.

CTGAN 是一种基于生成对抗网络 (Generative Adversarial Networks, GAN) 的数据增强方法. GAN 是一种深度学习模型<sup>[16]</sup>, 可以在复杂分布上进行无监督学习. GAN 有两个模块, 分别为生成器 G (generator) 和判别器 D (discriminator), 二者相互博弈学习从而产生新的数据. 其本质是让生成器对随机噪声的映射尽可能地拟合真实数据<sup>[17]</sup>. CTGAN 在标准 GAN 的结构基础上针对表格数据的特点引入了分模态归一化, 从而解决非高斯和多模式分布的连续列数据归一化问题, 并设计了条件向量来指导 G 产生数据. 具体过程如图 3 所示.

表 2 TSN 网络特征与标签之间的 MI 值

Tab. 2 MI between TSN network features and labels

特征	FIFO	CP8	Manual
totalIndividualFlows	0.151 8	0.344 7	0.282 4
critical	0.127 3	0.295 2	0.247 1
audio	0.095 6	0.218 6	0.187 7
video	0.187 1	0.442 2	0.359 0
maxLoad	0.205 3	0.533 9	0.370 5
timeToRunOneWCTT	0.159 1	0.308 9	0.286 7
Gini Coefficient	0.153 3	0.252 7	0.268 7

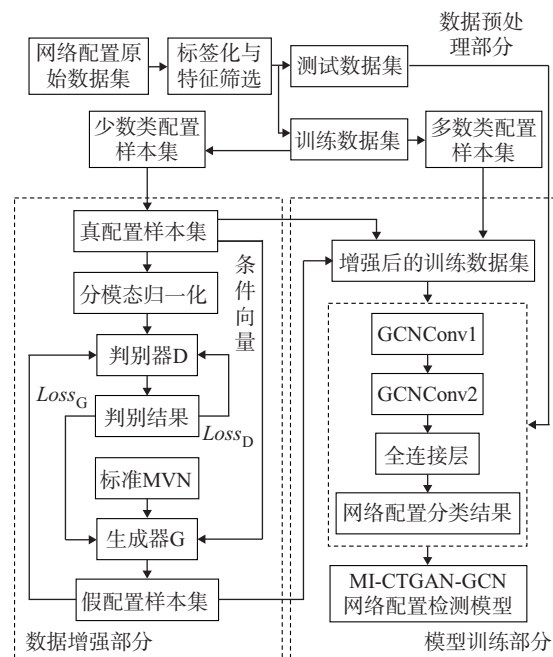


图 2 MI-CTGAN-GCN 模型流程图

Fig. 2 MI-CTGAN-GCN model flowchart

(3)在数据归一化步骤中,使用独热向量对训练集中离散列进行归一化,并对具有复杂分布的连续列进行分模态归一化,步骤如下:

1)对于数据集中每个连续列  $C_i$  和其中的每一个值  $c_{i,j}$ ,使用变分高斯混合模型来拟合高斯混合并估计模式数  $m_i$ .学习到的高斯混合可以表示为:

$$P_{C_i}(c_{i,j}) = \sum_{k=1}^{m_i} \mu_k N(c_{i,j}; \eta_k, \varphi_k) \quad (2)$$

式中: $\eta_k$  为模式的数学期望, $\mu_k$  为模式的权值, $\varphi_k$  为模式的标准差.

2)对  $C_i$  中的每个  $c_{i,j}$  计算其来自每种模式的概率. 概率密度:

$$\rho_k = \mu_k N(c_{i,j}; \eta_k, \varphi_k) \quad (3)$$

3)从给定的概率密度中采样一个模式,并使用采样的模式对值进行归一化. 使用独热向量来表示  $c_{i,j}$  所采样的模式,并使用标量  $\alpha_{i,j}$  表示模式内的值.  $\alpha_{i,j}$  的计算公式:

$$\alpha_{i,j} = \frac{c_{i,j} - \eta_k}{4\varphi_k} \quad (4)$$

(4)引入条件向量  $cond$  来指示条件  $D_{i^*} = k^*$  以指导生成器 G 产生样本. 将数据集中的离散列  $D_1, \dots, D_{N_d}$  转换成独热向量  $d_1, \dots, d_{N_d}$ , 其中  $d_i = [d_i^{(k)}], k = 1, \dots, |D_i|$ . 设第  $i$  个独热向量  $d_i$  的第  $i$  个掩模向量  $m_i = [m_i^{(k)}], k = 1, \dots, |D_i|$ , 条件可以由掩模向量表示为式(5):

$$m_i^{(k)} = \begin{cases} 1, & \text{if } i = i^* \text{ and } k = k^* \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

而后条件向量  $cond$  可用掩模向量表示为: $cond = m_1 \oplus \dots \oplus m_{N_d}$ .  $cond$  用于辅助 G 产生样本. G 由两个全连接隐藏层和一个输出层组成并使用  $Relu$  激活函数. 标准 MVN 变量会输入 G 中并映射产生一组独热向量  $\{\hat{d}_1, \dots, \hat{d}_{N_d}\}$ . 在训练过程中,通过在所有批处理实例中添加  $m_{i^*}$  和  $\hat{d}_{i^*}$  之间的交叉熵以惩罚损失,使得 G 将给定的  $m_{i^*}$  复制到生成的  $\hat{d}_{i^*}$  上. 经过训练后的 G 能够学习生成满足特定条件的样本,从而产生虚假样本,作为判别器 D 的一个输入.

(5)将原始数据集中的少数类样本设置为真实样本,参与到 G 与 D 之间的博弈学习中,作为 D 的另一个输入.

(6)判别器 D 由两个全连接隐藏层和输出层组成,使用  $leaky\ relu$  作为激活函数,同时在隐藏层后添加  $dropout$  减少模型的过拟合. 虚假样本与真实样本送往 D 进行判别,并输出对样本的判别结果.

(7)通过判别结果计算两个损失函数  $Loss_G$  和  $Loss_D$ , 并反向传播交替迭代更新 G 和 D 的网络模型参数,二者的构建如式(6)和式(7)所示:

$$Loss_G = -\frac{1}{m} \sum_{i=1}^m [D(x'^{(i)})] + H \quad (6)$$

$$Loss_D = \frac{1}{m} \sum_{i=1}^m [D(x'^{(i)}) - D(x^{(i)})] \quad (7)$$

式中: $m$  为批次中的样本数量; $x$  为真实样本; $x'$  为虚假样本; $D$  为判别器的输出; $H$  为交叉熵分数,始终为正. 判别器 D 的网络模型在每次迭代更新中提高分辨虚假样本与真实样本的能力. 生成器 G 在每次迭代中提高生成的虚假样本与真实样本之间的相似度,以混淆判别器.

(8)使用迭代训练完成后的 G 生成与真实样本高度近似的数据增强样本,这些样本与原始数据集中的训练集一同组成数据增强后的训练集.

(9)构建如图 4 所示的 GCN 网络模型. GCN 的核心思想是聚合邻居节点的信息来更新每个节点的特

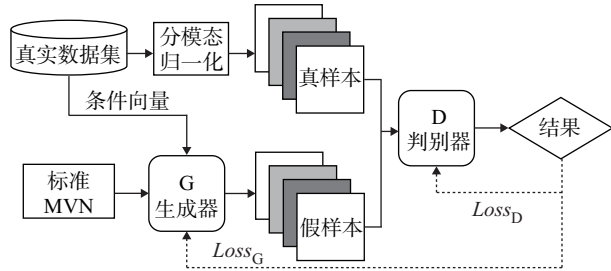


图 3 CTGAN 进行数据增强的过程

Fig. 3 Process of data augmentation using CTGAN

征<sup>[18]</sup>. 由各条链路负载计算得出 Gini 系数并通过特征筛选得到共 7 项特征数据作为节点特征, 网络拓扑中的链路负载数据作为边特征. GCN 由两层图卷积层构成, 并添加批归一化层 (Batch Normalization, BN) 减小过拟合, 使用 ReLU 作为激活函数以提供更有效率的梯度下降和反向传播<sup>[19]</sup>. 卷积层对各项网络配置数据的卷积逐步提取出配置可行性的信息. ReLU 激活函数的表达式如式 (8) 所示:

$$f = \max(0, x) \quad (8)$$

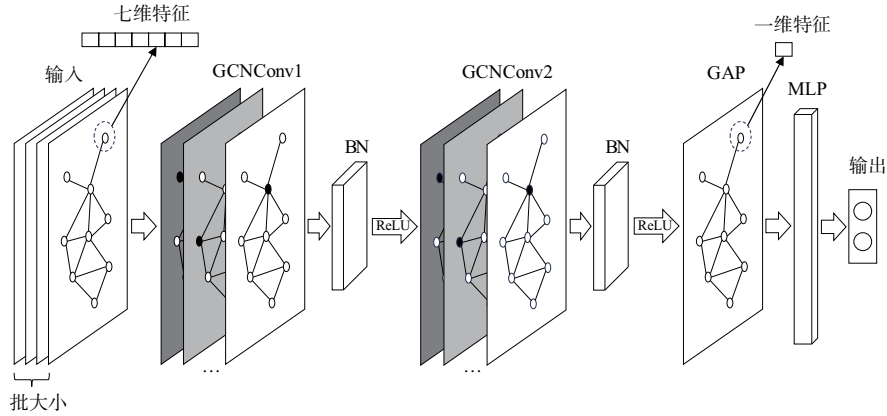


图 4 实验所构建的 GCN 网络模型

Fig. 4 GCN network model constructed in the experiment

(10) 对图进行全局平均池化 (Global Average Pooling, GAP), 将所有节点特征进行平均. 在此过程中, 拓扑中每个节点的七维特征会转化为一维特征. 对于图  $g_i$ , 公式如式 (9):

$$r_i = \frac{1}{N_i} \sum_{n=1}^{N_i} x_n \quad (9)$$

式中:  $r_i$  为  $g_i$  的全局特征向量,  $N_i$  为  $g_i$  中节点的数量,  $x_n$  为  $g_i$  中第  $n$  个节点的特征向量. 将池化后的特征通过 MLP 映射到输出维度, 得到分类结果. 交叉熵损失函数通过反向传播更新 GCN 分类模型的参数, 交叉熵损失函数公式如式 (10) 所示:

$$Loss = - \sum_{i=1}^N y_i \log(\hat{y}_i) \quad (10)$$

式中:  $N$  为类别数, 由于此任务为二分类任务, 因此  $N=2$ ;  $y_i$  为真实标签;  $\hat{y}_i$  为模型的预测概率.

(11) 使用经过数据增强后的训练集进行训练得到的 GCN 网络模型对原始数据集中的测试样本集进行网络配置可行性分类, 输出最终的分类结果.

### 3 仿真实验

#### 3.1 环境配置

本文实验将原始数据集划分为训练集和测试集, 比例为 0.85:0.15. 由于数据集的不平衡性, 模型训练前需要先对训练集进行数据增强, 测试集不变. 表 3 展示了不同调度方法的  $f$  值,  $f$  值为每种调度方法中可行配置的百分比. 可以看出 FIFO 和 CP8 的数据明显不平衡, 其中 FIFO 的不可行配置与可行配置的比例大于 10:1.

CTGAN 中的生成器和判别器都由两个全连接隐藏层和一个输出层组成. 全连接隐藏层维度为 256, Adam 优化器的学习率为 0.0002, 进行 600 次迭代. GCN 模型的第一个图卷积层输入特征维度为 7, 输出特征维度为 32, 第二个图卷积层输入特征维度为 32; 输出特征维度为 64, 批次大小为 32, 学习率设置为 0.001, 进行 350 次迭代. 原始数据集中总共有 8 964 个样本, 每个样本有 264 个特征. 本文使用 CTGAN 根据 3 种调度方式的  $f$  值分别为

表 3 不同调度方法的  $f$  值

Tab. 3  $f$  of different scheduling methods

调度方法	$f/\%$
FIFO	8.54
CP8	22.08
Manual	51.38

FIFO、CP8 和 Manual 调度方法生成 3 642、248 和 2 466 个样本,并将它们添加到训练集中对模型进行训练.

### 3.2 仿真结果

图 5 比较了 3 种调度方法在不使用数据增强算法,以及使用 3 种不同数据增强方法—Borderline – SMOTE(BS)、标准 GAN 和 CTGAN 的 GCN 模型的准确率情况. 过往研究<sup>[10]</sup>使用 GNN 模型对 CP 调度方法在相同拓扑结构上的检测准确率最高为 90.5%. 可以看出,本文构建的 GCN 模型相较于以往算法模型的准确率有较大的提高,通过数据增强后的准确率有更大的提升. 不经过数据增强的 GCN 模型对 CP8 调度方法的准确率达到 91.41%,在使用 CTGAN 数据增强后准确率得到进一步提升,达到了 92.56%. 3 种调度算法的分类准确率都表明数据增强后的模型分类准确率有较大提高,CTGAN 算法相比较 Borderline – SMOTE 算法和标准 GAN 算法有更明显的优势,其中 CTGAN – GCN 模型对 FIFO 调度方法的准确率已经达到 96.28%. 此结果表明,由 CTGAN 数据增强算法得到的训练集可以强化 GCN 模型对不平衡数据集的分类能力.

图 6 展示了 CTGAN – GCN 算法模型在 30 次实验中的准确率的分布情况,可以看出,模型在 3 种调度方式下的准确率都较为稳定,上下浮动在 0.5% 左右. 该结果数据表明所建模型是一个较为稳定的模型.

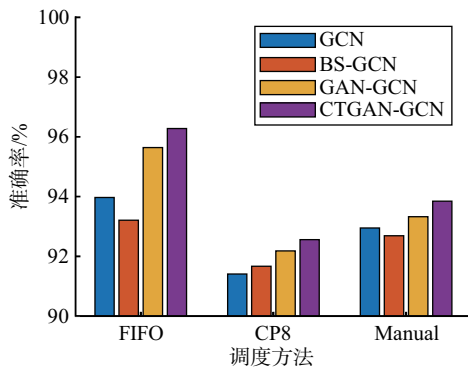


图 5 四种模型的 TSN 网络配置准确率对比图  
Fig. 5 Comparison of TSN network configuration accuracy for four models

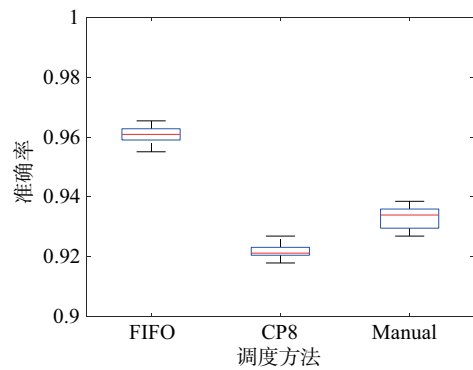


图 6 CTGAN – GCN 算法模型箱线图  
Fig. 6 Boxplot of the CTGAN – GCN algorithm model

图 7 和图 8 展示了采用 3 种数据增强方法后的真正率 (TPR) 以及真负率 (TNR) 的对比图. 经过数据增强后的模型对 3 种调度方法的真正率和真负率都有所提高. 数据集中 FIFO 和 Manual 调度方法中的正例样本 (可行配置样本) 以及 CP8 调度方法中的负例样本 (不可行配置样本) 都为少数类样本. 由仿真实验结果对比图中 FIFO 和 Manual 调度方法的 TNR 以及 CP8 的 TPR 可以看出,CTGAN 数据增强算法能够最显著地增强模型对少数类样本的分类能力,在提高模型分类检测 TPR 和 TNR 的效果上最为出色.

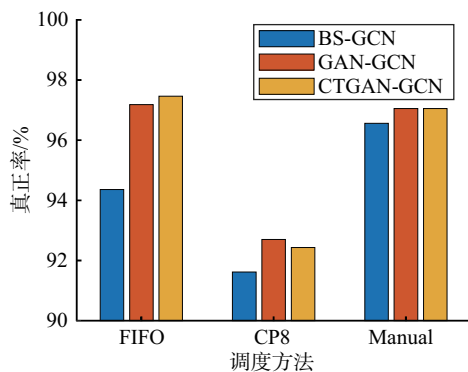


图 7 三种数据增强模型的 TPR 对比图  
Fig. 7 Comparison of TPR for three data augmentation models

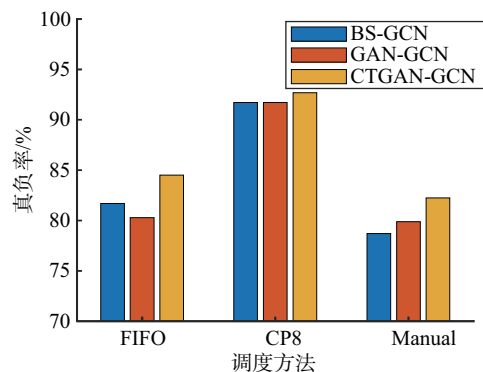


图 8 三种数据增强模型的 TNR 对比图  
Fig. 8 Comparison of TNR for three data augmentation models

表 4 展示了 3 种调度方法与对应 4 种模型在精确率 (precision)、召回率 (recall)、 $F1$  值和 Kappa 系数 ( $K$  值) 方面的统计量。

表 4 三种调度方法对应四种模型的精确率、召回率、 $F1$  值和 Kappa 系数

Tab. 4 Precision, recall,  $F1$  score, and Kappa coefficient of four models corresponding to three scheduling methods

调度方法	模型	精确率/%	召回率/%	$F1$ /%	$K$ /%
FIFO	GCN	93.90	99.86	96.79	49.02
	BS - GCN	98.09	94.36	96.19	64.96
	GAN - GCN	98.01	97.18	97.59	74.62
	CTGAN - GCN	98.43	97.49	97.94	78.48
CP8	GCN	90.61	91.35	90.98	82.78
	BS - GCN	90.89	91.62	91.25	83.30
	GAN - GCN	90.98	92.70	91.83	84.33
	CTGAN - GCN	91.93	92.43	92.18	85.08
Manual	GCN	93.44	97.87	95.60	77.86
	BS - GCN	94.25	96.56	95.39	77.75
	GAN - GCN	94.58	97.05	95.80	79.65
	CTGAN - GCN	95.18	97.05	96.11	81.41

由表 4 可以看出,在进行数据增强后,模型的性能均有所提高,在 4 类统计量上都有不同程度的增量。其中 CTGAN - GCN 算法模型提升效果最为显著,其精确率、 $F1$  值和  $K$  值的提高最为明显,在召回率上偶有下降是因为模型牺牲了一部分对多数类样本的分类能力,用以改进对少数类样本的分类能力。该结果表明,CTGAN - GCN 算法有助于缓解数据集的类别不平衡问题,从而帮助模型捕捉到更多少数类样本的特征,减少对多数类样本的过拟合,同时降低模型预测的随机性。

## 4 结论

针对以往深度学习算法在 TSN 网络配置可行性检测问题上准确率不佳的问题,本文引入 MI 值,从原始数据集中筛选得到强相关性的特征;其次使用 CTGAN 数据增强算法生成少数类样本解决数据集不平衡问题,最后构建 GCN 模型进行 TSN 网络配置检测。本文对比了 4 种模型并比较了 3 种数据增强算法,仿真结果表明:MI - CTGAN - GCN 时间敏感网络配置检测模型在增强原始数据集中的少数类样本,提升数据集的平衡性后,实现了更高的分类准确率,同时该模型通过图卷积神经网络中充分融合了 TSN 网络配置的图结构信息,有效验证了该模型在提升检测性能方面的有效性和创新性。

## 参考文献:

- [1] QIU F J, KUMAR A, HU J, et al. A review on integrating IoT, IIoT, and industry 4.0: A pathway to smart manufacturing and digital transformation[J]. IET Information Security, 2025, 2025: 9275962.
- [2] KIANGALA K S, WANG Z H, KIANGALA K S, et al. An effective communication prototype for time - critical IIoT manufacturing factories using zero - loss redundancy protocols, time - sensitive networking, and edge - computing in an industry 4.0 environment[J]. Processes, 2021, 9(11): 1 - 25.
- [3] CIRAOLO M. Automotive Ethernet: Construction and performance analysis of a 10BASE - T1S network[D]. Torino: Politecnico di Torino, 2024.
- [4] TABASSUM N, REDDYY C R K. Review on QoS and security challenges associated with the Internet of vehicles in cloud computing[J]. Measurement: Sensors, 2023, 27: 100562.
- [5] FIORI T. Exploring avionic connectivity in modern space systems: Design and experimental evaluation of ethernet - based communication architectures with FTT - Ethernet and Time Sensitive Networking[D]. Rome: Sapienza University of Rome, 2025.
- [6] ARESTOVA A, JENS HIELSCHER K S, GERMAN R. Simulative evaluation of the TSN mechanisms time - aware shaper and frame preemption and their suitability for industrial use cases[C]//2021 IFIP Networking Conference (IFIP Networking),

- 21 – 24 June 2021, Espoo and Helsinki, Finland. IEEE, 2021 : 1 – 6.
- [ 7 ] VLK M, HANZÁLEK Z, BREJCHOVÁ K, et al. Enhancing schedulability and throughput of time – triggered traffic in IEEE 802.1Qbv time – sensitive networks[ J ]. IEEE Transactions on Communications, 2020, 68( 11 ): 7023 – 7038.
- [ 8 ] ERGENÇ D, BRÛLHART C, NEUMANN J, et al. On the security of IEEE 802.1 time – sensitive networking[ C ]//2021 IEEE International Conference on Communications Workshops ( ICC Workshops ), 14 – 23 June 2021, Montreal, QC, Canada. IEEE, 2021 : 1 – 6.
- [ 9 ] FEDULLO T, MORATO A, TRAMARIN F, et al. A comprehensive review on time sensitive networks with a special focus on its applicability to industrial smart and distributed measurement systems[ J ]. Sensors, 2022, 22( 4 ): 1638.
- [ 10 ] MAI T L, NAVET N. Deep learning to predict the feasibility of priority – based Ethernet network configurations[ J ]. ACM Transactions on Cyber – Physical Systems, 2021, 5( 4 ): 1 – 26.
- [ 11 ] XU L. Synthesizing tabular data using conditional GAN[ D ]. Cambridge: Massachusetts Institute of Technology, 2020.
- [ 12 ] LONG MAI T, NAVET N. Improvements to deep – learning – based feasibility prediction of switched Ethernet network configurations[ C ]//29th International Conference on Real – Time Networks and Systems, NANTES France. ACM, 2021 : 89 – 99.
- [ 13 ] 唐铖杰,王澄,郇战,等.基于 MIC – SSA – SVM 的时间敏感网络配置可行性检测[ J ].昆明理工大学学报(自然科学版),2024,49( 1 ): 73 – 82.
- TANG C J, WANG C, HUAN Z, et al. Feasibility detection for time – sensitive network configurations based on MIC – SSA – SVM[ J ]. Journal of Kunming University of Science and Technology ( Natural Science ), 2024, 49( 1 ): 73 – 82.
- [ 14 ] WANG C, CHEN L, TANG C J, et al. Enhanced time – sensitive networking configuration detection using optimized BPNN with feature selection for industry 4.0[ J ]. Cluster Computing, 2024, 27( 7 ): 9795 – 9810.
- [ 15 ] MAI T L. Machine learning in the design space exploration of TSN networks[ D ]. Luxembourg: University of Luxembourg, 2022.
- [ 16 ] AGGARWAL A, MITTAL M, BATTINENI G. Generative adversarial network: An overview of theory and applications[ J ]. International Journal of Information Management Data Insights, 2021, 1( 1 ): 100004.
- [ 17 ] CHAKRABORTY T, REDDY K S U, NAIK S M, et al. Ten years of generative adversarial nets ( GANs ): A survey of the state – of – the – art[ J ]. Machine Learning: Science and Technology, 2024, 5( 1 ): 011001.
- [ 18 ] BHATTI U A, TANG H, WU G L, et al. Deep learning with graph convolutional networks: An overview and latest applications in computational intelligence[ J ]. International Journal of Intelligent Systems, 2023, 2023 : 8342104.
- [ 19 ] HAMMAD M M. Deep learning activation functions: Fixed – shape, parametric, adaptive, stochastic, miscellaneous, non-standard, ensemble[ EB/OL ]. Arxiv, 2024 : 2407. 11090. <https://arxiv.org/abs/2407.11090>.

( 编 辑:朱银周)