文章编号:1671-6833(2024)04-0030-08

基于改进 WGAN-GP 和 ResNet 的车联网入侵检测方法

魏明军[1,2], 李 凤[1], 刘亚志[1,2], 李 辉[1]

(1. 华北理工大学 人工智能学院,河北 唐山 063210;2. 河北省工业智能感知重点实验室,河北 唐山 063210)

摘 要: 为保护车联网系统免受网络攻击的威胁,同时提高车联网入侵检测的准确率,针对车辆网络数据流量大且攻击类型不平衡的特点,提出了一种新的车联网入侵检测方法(AQVAE-RGSNet)。该方法通过一种对抗量化变分自编码器以对车辆网络数据进行不平衡处理,该编码器通过结合矢量量化变分自编码器与带梯度惩罚的生成对抗网络进行构建,以缓解数据集中异常攻击类型样本数量极度不平衡的问题,并使用 ResNet 网络与改进的分段残差神经网络对输入的样本数据进行联合学习并预测其攻击类型。实验结果表明:AQVAE-RGSNet 在车联网数据集 CICIDS2017 和 CAN-intrusion-dataset 上的 *F*1 得分分别达到了 0.998 6 和 0.999 7;在保证最佳训练效果的前提下,能够更有效地识别车辆网络之中的攻击威胁。

关键词:车联网; 入侵检测; 生成对抗网络; 残差神经网络; 特征融合

中图分类号: TP393; TN929.5 **文献标志码**:A **doi**:10.13705/j.issn.1671-6833.2024.04.008

近年来,机器学习(ML)和深度学习(DL)技术在网络安全和车辆系统中的应用引起了研究人员和汽车制造商的广泛关注。现代车辆(包括自动驾驶和联网车辆)越来越多地与外部世界相连,以便实现各种功能和服务。然而,互联性的提高也使车联网更容易受网络攻击威胁。网络攻击者既可以通过 OBD Ⅱ(on-board diagnostics Ⅱ)接口对车载网络发起内部攻击,也可以通过无线接口发送恶意流量报文对车载外部网络发起外部攻击。所以,利用 ML 算法和 DL 算法提高车辆网络入侵检测系统(intrusion detection system,IDS)的准确性和训练速度是非常必要的。

目前,ML 和 DL 模型已被广泛应用于车联网入侵检测任务中。例如,Mehedi 等[1]提出了基于深度迁移学习的 P-LeNet 车载网络入侵检测方法,在 Car-Hacking 数据集上获得了 0.978 的 *F*1 得分。Hossain 等[2]提出了一种用于车内入侵检测的基于一维卷积神经网络的 IDS,在许多时间序列数据分析问题中表现良好。Song 等[3]提出了一种基于深度卷积神经网络的 IDS 模型,该模型使用简化的 Inception-ResNet 来检测导航系统中的攻击。Yang 等[4]讨论了车内和车外网络的漏洞,提出了一种基于特征和异常的多层混合入侵检测系统,在 CAN-intrusion-dataset 和 CICIDS2017 数据集上 *F*1 得分分别达到了 0.963 和 0.800。Yang 等[5]利用卷积神经网络和超参数优化技术,提出了一种基于迁移学习和集成学习的智能入侵检测系统,在公共汽车安全基准数据集上的检测率超过 99.25%。Yang 等[6]提出了一种新的集成 IDS 框架 LCCDE,通过在 3 种高级 ML 算法(XGBoost、LightGBM 和 CatBoost)中确定针对每种类型的攻击的最佳 ML 模型来构建,并在 Car-Hacking 和 CICIDS2017 数据集上分别获得了 0.999 和 0.998 的 *F*1 得分。

虽然上述方法在车辆网络攻击检测任务中取得了显著成果,但性能仍有提升空间。为此,本文提出了一种新的车联网入侵检测模型(AQVAE-RGSNet)。本文的主要工作如下。

(1)提出了一种对抗量化变分自编码器以进行数据的不平衡处理,结合了两种数据生成方法的优势,有效缓解了车联网入侵检测中攻击类型不平衡

收稿日期:2023-12-30;**修订日期**:2024-02-05

基金项目:河北省高等学校科学技术研究项目(ZD2022102)

作者简介:魏明军(1969—),男,河北唐山人,华北理工大学教授,主要从事计算机视觉、入侵检测、机器学习、数据挖掘研究,E-mail:weimj@ncst.edu.cn。

引用本文:魏明军,李凤,刘亚志,等. 基于改进 WGAN-GP 和 ResNet 的车联网入侵检测方法[J]. 郑州大学学报(工学版),2024,45(4):30-37. (WEI M J, LI F, LIU Y Z, et al. An intrusion detection method for Internet of Vehicles based on improved WGAN-GP and ResNet[J]. Journal of Zhengzhou University (Engineering Science),2024,45(4):30-37.)

的问题。

(2)为了提高入侵检测的准确率和训练效率,本文提出了一种用来筛选车辆网络数据中可能的入侵攻击的 RGSNet 模型,由基于迁移学习的 ResNet 网络与改进的分段残差神经网络联合构建而成,能够进行更有效的深层特征提取。

1　本文方法

图 1 为本文所提出的 AQVAE-RGSNet 的整体结构,其中包含 5 个主要模块:数据输入、数据预处理、不平衡处理、数据转换和数据分类。模型训练结束后,在实际应用中,车辆网络数据经过数据预处理后直接进入数据转换模块,之后进入 RGSNet 进行入侵检测。

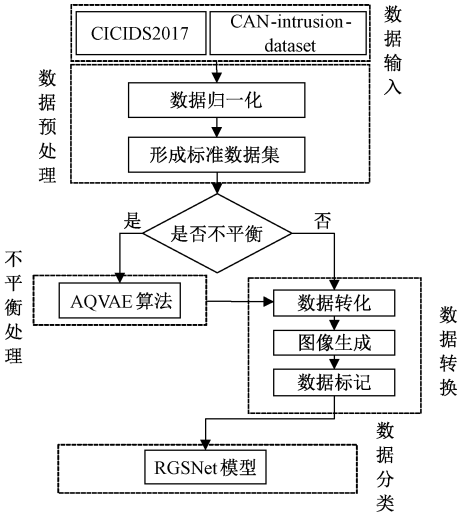


图 1　车联网入侵检测模型框架

Figure 1　Intrusion detection model framework for Internet of Vehicles

1.1　数据预处理

在特征提取前,首先需要对网络数据进行归一化处理。归一化技术中,最小-最大值归一化和分位数归一化是 2 种常用方法。由于最小-最大值归一化方法不能很好地处理异常值,并且可能导致多数数据样本值过小,因此本文使用分位数归一化。分位数归一化方法将特征分布转换为正态分布,并根据正态分布重新计算所有特征的特征值。其所得到的大部分特征值都接近于中值,这对于处理异常值是有效的。

1.2　不平衡处理

为了应对车辆网络中攻击类型不平衡的问题,本文提出对抗量化变分自编码器(adversarial quantized variational auto encoder,AQVAE)对数据集进行不平衡处理,AQVAE 算法是将 VQ-VAE-2 和 WGAN-GP 相结合的一种算法。

1.2.1　VQ-VAE-2

双层矢量量化变分自编码器(vector quantized variational auto encoder-2,VQ-VAE-2)是变分自动编码器的改进版本,它改善了变分自编码器生成图像模糊的问题,并且将连续数据离散化以生成所需要的特定类型数据。为此,本文选取其进行数据样本生成。

图 2 显示了 VQ-VAE-2 的整体结构。给定一个原始样本 $\boldsymbol{x}$,首先使用下层和上层的编码器(Encoder)进行特征提取和信息压缩,得到的特征为 $z_e(\boldsymbol{x}_{\text{top}})$。之后,算法通过矢量量化(VQ)的方式对 $z_e(\boldsymbol{x}_{\text{top}})$ 进行离散处理,以得到上层离散特征 $z(\boldsymbol{x}_{\text{top}})$。$z(\boldsymbol{x}_{\text{top}})$ 在经过解码器(Decoder)后与下层 Encoder 提取得到的特征进行级联。级联得到的特征 $z_e(\boldsymbol{x}_{\text{bottom}})$ 会被馈送到 VQ 中,以生成下层离散特征 $z(\boldsymbol{x}_{\text{bottom}})$。最后,离散特征 $z(\boldsymbol{x}_{\text{top}})$ 和 $z(\boldsymbol{x}_{\text{bottom}})$ 将被作为训练数据,通过 Decoder 训练生成新的数据样本。

1.2.2　WGAN-GP

WGAN-GP (Wasserstein generative adversarial network with gradient penalty,WGAN-GP)是 WGAN 的改进版本。WGAN-GP 能够在几乎没有超参数调优的情况下稳定地训练各种 GAN 架构。本文选取其作为判别器,以对生成器生成的数据继续判别和优化。

具体算法步骤如图 3 所示,WGAN-GP 由一个生成器 G 和一个判别器 D 组成。对于输入的高斯

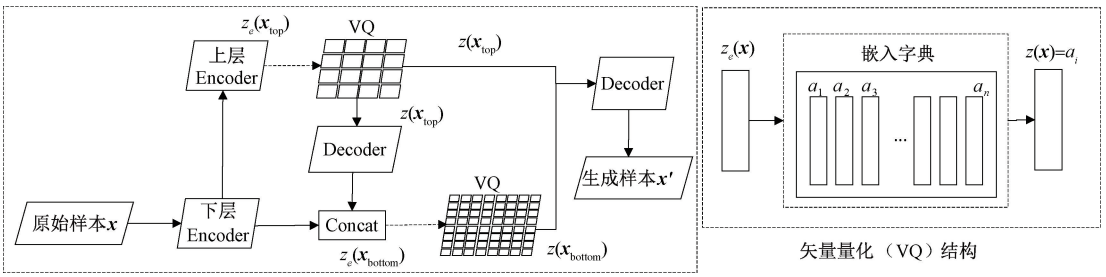


图 2　VQ-VAE-2 网络结构图

Figure 2　VQ-VAE-2 network structure diagram

随机变量,算法首先使用生成器进行虚拟样本生成。之后,生成的虚拟样本与真实样本将被同时馈送到判别器中,以判别样本类型并对生成器进行损失优化。

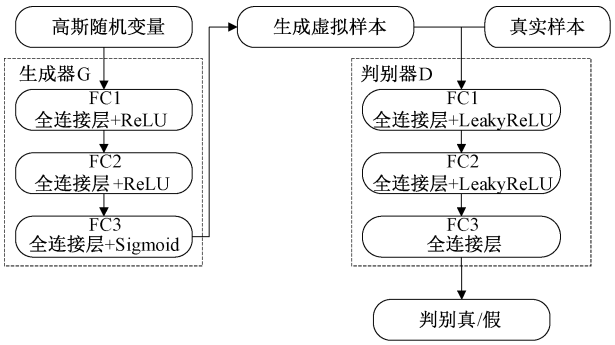


图3 WGAN-GP网络结构图

Figure 3 WGAN-GP network structure diagram

1.2.3 AQVAE算法

WGAN-GP中的梯度惩罚能够显著提升训练速度,解决了梯度二值化和梯度消失爆炸的问题,但在数据生成方面,其不能生成任意指定类型的数据。而使用VQ-VAE-2可将连续数据离散化以生成所需要类型的数据。因此,本文集成VQ-VAE-2和WGAN-GP的优势提出了AQVAE算法,能够以少数类型的样本数据为依据,通过过采样的方式生成同类型数据,以解决数据集中的样本不平衡问题。

AQVAE算法仍采用生成器-判别器架构。其中,生成器由VQ-VAE-2组成,判别器由WGAN-GP中的判别器D组成。所提出的AQVAE能够结合两者的优点,在保持较高训练效率的同时生成指定类型的数据,从而获得更有效的不平衡处理效果,其流程如图4所示。

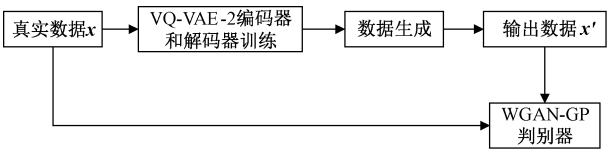


图4 AQVAE流程图

Figure 4 AQVAE flow chart

由表1可以看出CICIDS2017数据集中Botnets、Brute-Force和Web Attack这3类攻击类型的数量占比过少,异常攻击类型数据样本的极度不平衡会导致网络在分类时将注意力集中于样本数量较大的类型。为了提高少数类样本的准确率,同时确保多数类样本的训练速度和准确率,本文着重针对该数据集进行平衡处理。具体来说,根据车内网络数据经验,本文使用AQVAE算法分别生成20 000条Botnets、Brute-Force和Web Attack类型的攻击数据,不平衡处理后,数据集中此3类攻击的占比可达到3%以上,结果见表1。

表1 CICIDS2017数据集的数据样本分布

Table 1 Data sample distribution of CICIDS2017

类型	不平衡处理前		不平衡处理后	
	数据数量	数据占比/%	数据数量	数据占比/%
BENIGN	200 000	32.06	200 000	29.25
Botnets	1 938	0.31	21 938	3.21
DoS Attack	320 295	51.34	320 295	46.84
Sniffing	90 379	14.49	90 379	13.22
Brute-Force	9 117	1.46	29 117	4.26
Web Attack	2 133	0.34	22 133	3.24

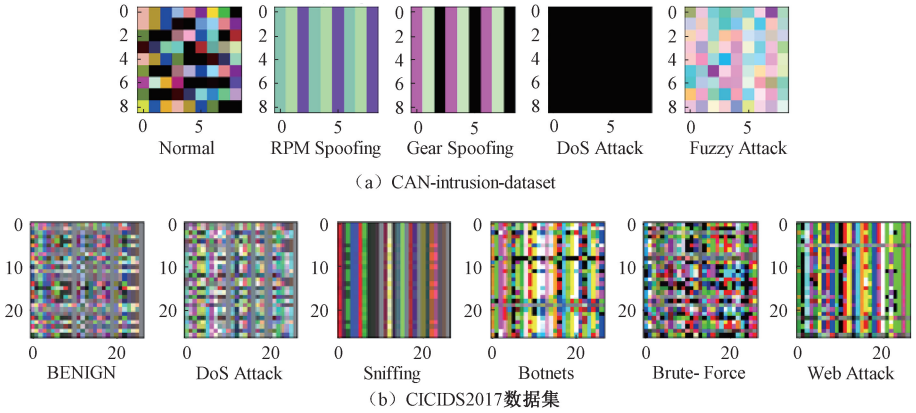
1.3 数据转换

由于CNN是为解决计算机视觉问题而设计的深度学习网络,而车辆网络流量数据通常以csv和txt文件的形式进行存储,为了有效挖掘CNN在入侵检测中的潜力,需要将网络流量数据转换为图像形式。在数据处理完成后,本文利用RGSNet模型来对特征进行提取。不同形式下的样本在CNN上的训练效果以及时间都会有所不同,其中,三维数据比一维数据训练的效果要更好[7]。为此,在使用提出的卷积网络RGSNet进行特征提取前,需要先将一维数据转换成为三维数据。转换时,首先根据时间戳和特征大小将数据样本转换成块(chunk):对于CAN-intrusion-dataset,由于其中的每条数据包含9个重要特征,因此可以将27条连续数据中的共243(27×9)个chunk转换成为大小为9×9×3的RGB图像;同理,本文将CICIDS2017数据集的78列特征属性扩充为81列后,将其换成大小为27×27×3的RGB图像集。接下来,需要根据数据块的攻击模式对转换后的图像进行标记:如果一个块/图像中的所有样本都是正常样本,则该图像被标记为“正常”;相反,如果一个数据块/图像包含攻击样本,则该图像将被标记为该数据块中出现最频繁的攻击类型。CAN-intrusion-dataset和CICIDS2017数据集中每种攻击类型的代表样本如图5所示。

经过上述数据转换操作后,生成代表了2个数据集的网络数据图像集。该图像集将作为数据分类模块RGSNet的输入,本文按7:3比例将其分为训练集和测试集,并把训练集中的30%用于验证。

1.4 数据分类

本文提出RGSNet对车辆网络数据集进行深层特征提取,以获得更精确的网络数据分类结果。RGSNet由基于迁移学习的残差神经网络ResNet18与改进的分段残差神经网络RGStage50并行组成,



（a）CAN-intrusion-dataset

（b）CICIDS2017数据集

图 5　2 个数据集中各类型的代表性样本图像

Figure 5　Representative sample images of each category in two datasets

将接收数据转换后的网络数据图像集作为输入,并对车辆网络数据集的入侵检测结果进行分类及输出。

1.4.1　ResNet18 网络

RGSNet 中所使用的 ResNet18 网络是利用迁移学习进行训练[8]。网络在训练时首先加载在 ImageNet 数据集上训练好的预训练权重,之后通过微调使其更好地拟合车联网数据集。

1.4.2　RGStage50 网络

分段残差神经网络(ResStage)是 Duta 等[9]提出的 iResNet 中的一个分支。在网络加深时,ResStage 更易于模型训练及优化。因此,本文选择其来进行更有效的深层特征提取。同时,为了提高网络的特征多样性和预测准确率并加快训练效率,本文将原始 ResStage50 网络中每个 ResBlock 的 3×3 卷积(Conv3×3)替换为紧凑卷积(CompConv3×3),其次将 ResBlock 中的最后一个 ReLU 激活函数替换为 Sigmoid,并加入一个额外的全局响应归一化层[10](global response normalization,GRN),从而提出了新的 RGStage 网络。

激活函数是深层神经网络中必不可少的一部分,能够使神经网络对非线性问题进行拟合。ResNet 中使用 ReLU 作为隐藏层的激活函数,但过多使用 ReLU 会导致网络中的负信号归零,这会阻碍主路径上的信息传播。同时,ResNet 的主路径中缺少标准化处理,这同样加大了深层网络中特征学习的难度。

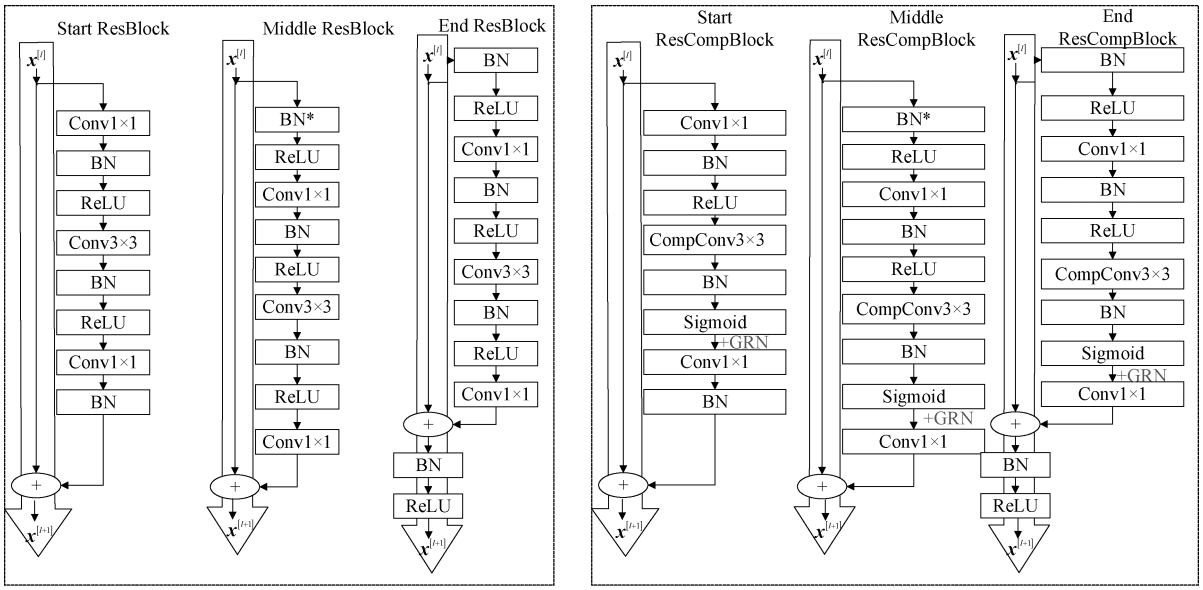
为了解决上述问题,分段残差神经网络中使用了一种分段 ResBlock 结构,整体的 ResStage 网络总共分为 3 个阶段,包括 1 个开始阶段、4 个主要阶段和 1 个结束阶段。如图 6(a) 所示,每个主要阶段均由 1 个 Start ResBlock、若干个 Middle ResBlock 和 1 个 End ResBlock 堆叠而成,4 个主要阶段中 Block 的堆叠数量分别为[3,4,6,3]。可以看出,ResStage 中激活函数的数量只与网络中主要阶段中主路径的数量相关,而与网络的大小和深度无关。因此,随着网络深度的增加,主路径中始终能够保持固定数量的激活函数,以便信息向前和向后传播。此外,每个主要阶段的 Start ResBlock 与 End ResBlock 最后都被添加了额外的批量归一化(BN)层,以使信号能够稳定进入下一阶段,这更加有利于网络对特征进行抽象和提取。

传统卷积的作用是学习一个特征变换,将 c_{in} 维度的输入转换为 c_{out} 维度的输出,其卷积核具有一定的过度参数化和冗余性问题。因此,为了加快模型的训练效率,本文将原始 ResStage50 网络中每个 ResBlock 的 3×3 卷积(Conv3×3)替换为了紧凑卷积(CompConv3×3)。紧凑卷积通过分治策略进行高效的特征学习,其输出的特征一半由卷积生成,一半由输入直接生成,这种方式能够以最小的参数量在卷积中最大程度地传递所学习到的信息。同时,紧凑卷积的核心单元可以以递归计算的方式执行,从而产生分而治之的策略,能够节省大量的计算量和参数量,并更高效地提取特征。

使用 Sigmoid 替换残差神经网络最后一层的 ReLU 激活函数可以提高网络的准确率[11]。本文同样将 ResBlock 中的最后一个 ReLU 激活函数替换成 Sigmoid。此外,为了增强网络中的特征竞争、提高模型性能,本文进一步在 ResBlock 中添加了 GRN 层。GRN 是一种归一化结构,能够通过增强通道间特征竞争的方式来减少深层网络中的特征崩溃问题。最终的 RGStage 网络如图 6(b)所示。

1.4.3　特征融合与分类

在使用 ResNet18 与所提出的 RGStage50 进行并行特征提取后,RGSNet 利用级联的方式对所提取到的特征进行融合,最后通过全连接层输出最终的



(a) iResNet-ResStage50 (b) RGStage50

图 6 RGStage 模型图

Figure 6 RGStage model diagram

入侵检测分类结果。

2 实验设计

2.1 实验所需数据集

CAN-intrusion-dataset[12]是在控制局域网(controller area network,CAN)攻击发起时通过车辆的 OBD Ⅱ端口记录 CAN 流量生成的,数据集中包含正常类型(Normal)数据和 4 种主要攻击类型(Fuzzy Attack、Gear Spoofing、DoS Attack 和 RPM Spoofing)数据。由于 CAN-intrusion-dataset 中的攻击类型数据样本相对平衡,因此不需要进行平衡处理。

在车辆网络外部安全数据集如 KDD-99、NSL-KDD 和 CICIDS2017[13]中,CICIDS2017 数据集包括更多的特征、实例和网络攻击类型。因此,本文选取 CICIDS2017 数据集作为车联网入侵检测的外部网络数据集。CICIDS2017 数据集共包含 7 大类型数据,其中由于 Infiltration 攻击类型的数量过少,因此本文不对其进行研究。表 1 统计了 CICIDS2017 数据集中除 Infiltration 攻击类型外的数据样本分布,其中 BENIGN 代表正常样本。

2.2 评估指标和参数设置

本文使用 4 个广泛流行的评价指标综合评估所提车联网入侵检测模型的性能,分别为准确率 AR、召回率 RR、精确率 PR 和 $F1$ 得分。其中,RR、PR 和 $F1$ 得分采用 macro-average 方法:在计算均值时,每个类型具有相同的权重,最后的结果是每个类型指标值的算术平均。计算公式分别为

$$\begin{cases} AR = \dfrac{TP + TN}{TP + FP + FN + TN}; \\ PR = \dfrac{TP}{TP + FP}; \\ RR = \dfrac{TP}{TP + FN}; \\ macro\_PR = \dfrac{1}{n}\sum_{i=1}^{n} PR; \\ macro\_RR = \dfrac{1}{n}\sum_{i=1}^{n} RR; \\ macro\_F1 = \dfrac{2macro\_PR \cdot macro\_RR}{macro\_PR + macro\_RR}。 \end{cases} \tag{1}$$

式中:TP 为真正例;FP 为假正例;TN 为真反例;FN 为假反例。

本文使用 PyTorch 框架实现所提车联网入侵检测系统,实验时所使用的计算机 CPU 型号为 Intel(R)Xeon(R)Platinum 8255C(2.50 GHz)。实验所使用的优化器为 Adam;dropout 率为 0.3;学习率为 0.001;$batch\_size$ 为 16。训练时,车内网络数据训练周期为 5,车外网络数据训练周期为 10。

3 实验结果与分析

3.1 整体对比实验

为了验证算法的整体效果,本文在 CAN-intrusion-dataset 和 CICIDS2017 数据集上将 AQVAE-RGSNet 与相关领域的方法进行了比较。评估结果如表 2、表 3 所示。

表 2　CAN-intrusion-dataset 上模型性能评估

Table 2　Performance evaluation of models on CAN-intrusion-dataset

方法	*AR*/%	*PR*/%	*RR*/%	*F*1
KNN[14]	97.40	96.30	94.70	0.934 0
SAIDuCANT[15]	87.21	88.66	98.24	0.920 0
P-LeNet[1]	98.10	98.14	98.04	0.978 3
LSTM-AE[16]	99.00	99.00	99.90	0.990 0
ID-CNN[2]	99.96	99.94	99.63	0.998 0
DCNN[3]	99.93	99.84	99.84	0.999 1
AQVAE-RGSNet	**99.99**	**99.95**	**99.99**	**0.999 7**

表 3　CICIDS2017 数据集上模型性能评估

Table 3　Performance evaluation of models on CICIDS2017 dataset

方法	*AR*/%	*PR*/%	*RR*/%	*F*1
KNN[17]	96.30	96.20	93.70	0.963 0
PCA-RF[18]	99.60	99.60	99.00	0.996 0
DBN[19]	98.95	95.82	95.81	0.958 1
LCCDE[6]	99.81	99.81	**99.91**	0.998 1
AQVAE-RGSNet	**99.84**	**99.88**	99.85	**0.998 6**

从表 2 可以看出,本文提出的方法表现出了最好的性能。用于对比的模型由于缺少数据转换,并且仅使用一维数据作为输入,准确率(87.21%~99.96%)和 *F*1 得分(0.920 0~0.999 1)较低。而本文提出的 AQVAE-RGSNet 的 *F*1 得分达到了 0.999 7,准确率和召回率指标也均达到了 99.99%。这主要得益于所使用的数据转换方法,转换后的图像集能够精准地区分出数据集中的正常数据和攻击类型数据。

如表 3 所示,本文提出的模型与文献[6, 17-19]中的先进方法进行了定量比较,这些方法在 CICIDS2017 数据集上取得了良好的性能,但由于缺乏数据的不平衡处理,准确率(96.30%~99.81%)和 *F*1 得分(0.958 1~0.998 1)相对较低。而本文提出的方法取得较好的结果,其中 *F*1 得分达到了最高的 0.998 6。这主要得益于所使用的不平衡处理方法,能够通过多样化的数据生成有效降低 CICIDS2017 数据集中数据不平衡所带来的影响。同时,使用 ResNet18 与 RGStage50 并行提取特征的方式增强了算法对于车辆网络数据的分析和分类能力,使其能够获得更精确的入侵检测结果。

3.2　消融研究

为了验证所提出的不平衡处理方法、数据转换方法和数据分类方法的有效性,本文分别对三者进行了消融实验。针对不平衡处理方法 AQVAE 的消融实验结果如表 4、表 5 所示。

表 4　CICIDS2017 数据集上的样本不平衡处理结果

Table 4　Result of sample imbalance processing on CICIDS2017 dataset

实验方法	*AR*/%	*PR*/%	*RR*/%	*F*1
w/o AQVAE	99.30	98.57	97.72	0.980 2
SMOTE-RGSNet	99.71	99.51	99.84	0.996 7
BorderlineSMOTE-RGSNet	99.49	99.17	99.63	0.994 0
AQVAE-RGSNet	**99.84**	**99.88**	**99.85**	**0.998 6**

表 5　CICIDS2017 数据集中各数据类型的不平衡处理结果

Table 5　Unbalanced processing results of each data category in CICIDS2017 dataset

类型	无不平衡处理			AQVAE-RGSNet		
	PR/%	*RR*/%	*F*1	*PR*/%	*RR*/%	*F*1
BENIGN	98.11	99.89	0.989 9	99.82	99.72	0.997 7
DoS Attack	99.92	98.76	0.993 4	99.81	99.89	0.998 5
Sniffing	99.97	100.00	0.999 8	100.00	100.00	1.000 0
Botnets	94.29	90.00	0.914 7	99.69	99.75	0.997 5
Brute-Force	99.14	100.00	0.995 7	99.90	99.71	0.998 0
Web Attack	99.99	97.65	0.987 9	100.00	100.00	1.000 0

从表 4 可以看出,使用 AQVAE 对 CICIDS2017 数据集进行不平衡处理后,提高了 CICIDS2017 数据集的整体检测精度。同时,本文使用了 2 种先进的不平衡处理方法 SMOTE[20] 和 BorderlineSMOTE[21] 进行对比验证。可以看出,本文提出的 AQVAE 的不平衡处理效果在 4 种评价指标上均优于其他算法,证明了其有效性。

从表 5 可以看出,CICIDS2017 数据集中不平衡攻击类型数据(Botnets、Brute-Force 和 Web Attack)的指标得到了明显的改善,同时其他攻击类型的指标并没有受到影响,这充分验证了所提出的 AQVAE 不平衡处理方法的有效性。

不平衡处理后,本文针对数据转换的方法进行了消融实验,分别以原本的网络数据(csv)、灰度图数据(一维)和本文数据转换后的三维图像数据作为分类网络的输入,以验证数据转换方法的有效性。以 CICIDS2017 数据集为例,实验结果如表 6 所示。本文的数据转换方法在 *AR*、*PR*、*RR*、*F*1 得分和训练时间上都表现出了最好的效果,这充分证明了其有效性。

在此之后,本文将 RGSNet 中的 RGStage50 替换为了其他的先进特征提取网络,分别为 ResNet50、Res2Next50[22]、ConvNext-tiny[23] 和 iResNet50[9],以验证数据分类方法中所提出的改进分段残差神经网络(RGStage)的有效性。实验结果如表 7 所示。

表6 数据转换模块的消融实验结果

Table 6 Results of ablation experiments on the data conversion module

数据格式	*AR*/%	*PR*/%	*RR*/%	*F*1	训练时间/min
csv	93.28	66.19	78.35	0.704 8	197.90
灰度图	99.59	99.70	98.86	0.992 6	304.00
本文三维图像	**99.84**	**99.88**	**99.85**	**0.998 6**	**10.85**

从表7可以看出,虽然ResNet50模型的训练时间较短,但其入侵检测的准确率最低;ConvNext-tiny与Res2Next50虽然能够获得较高的*AR*、*PR*、*RR*和*F*1得分,但是训练时间较长;iResNet50的准确率虽然相较于ResNet50有所提高,但是检测效果低于ConvNext-tiny和Res2Next50,而本文所提出的RGSNet方法能在保持高训练效率的同时获得最好的入侵检测结果,验证了其有效性。

表7 数据分类模块的消融实验结果

Table 7 Ablation experiment results of the data classification module

模型	CAN-intrusion-dataset(*epoch*=5)					CICIDS2017(*epoch*=10)				
	AR/%	*PR*/%	*RR*/%	*F*1	训练时间/min	*AR*/%	*PR*/%	*RR*/%	*F*1	训练时间/min
ResNet18+ResNet50	98.68	94.40	99.69	0.966 2	6.63	99.39	99.65	99.42	0.995 3	11.12
ResNet18+ConvNext-tiny	99.90	99.43	99.98	0.997 0	7.51	99.67	99.59	99.53	0.995 5	12.51
ResNet18+Res2Next50	99.98	99.90	99.99	0.999 4	8.57	99.49	99.81	99.45	0.996 3	14.17
ResNet18+iResNet50	99.68	98.26	99.92	0.990 6	6.82	99.80	99.73	99.72	0.997 2	11.63
AQVAE-RGSNet	**99.99**	**99.95**	**99.99**	**0.999 7**	**6.55**	**99.84**	**99.88**	**99.85**	**0.998 6**	**10.85**

4 结论

本文提出了一种用于车联网入侵检测的AQVAE-RGSNet方法,该方法使用所提出的RGSNet网络对输入的样本数据进行特征提取,能够有效降低入侵检测系统对于训练样本的数量依赖,并同时提高训练效率。此外,为了应对车辆网络中攻击类型不平衡的问题,本文提出了一种新的不平衡处理方法AQVAE,以缓解少数类型对模型检测性能所带来的负面影响。在车内和车外网络数据集上的实验结果表明,本文所提出的方法具有比以往方法更优秀的入侵检测性能。接下来,本文将继续在轻量化方面对模型进行改进,以使其能够适用于更多的车辆网络入侵检测场景。

参考文献:

[1] MEHEDI S T, ANWAR A, RAHMAN Z, et al. Deep transfer learning based intrusion detection system for electric vehicular networks[J]. Sensors (Basel, Switzerland), 2021, 21(14): 4736.

[2] HOSSAIN M D, INOUE H, OCHIAI H, et al. An effective in-vehicle CAN bus intrusion detection system using CNN deep learning approach[C]//2020 IEEE Global Communications Conference. Piscataway: IEEE, 2021: 1-6.

[3] SONG H M, WOO J, KIM H K. In-vehicle network intrusion detection using deep convolutional neural network[J]. Vehicular Communications, 2020, 21: 100198.

[4] YANG L, MOUBAYED A, SHAMI A. MTH-IDS: a multitiered hybrid intrusion detection system for Internet of Vehicles[J]. IEEE Internet of Things Journal, 2022, 9(1): 616-632.

[5] YANG L, SHAMI A. A transfer learning and optimized CNN based intrusion detection system for Internet of Vehicles[C]//IEEE International Conference on Communications. Piscataway: IEEE, 2022: 2774-2779.

[6] YANG L, SHAMI A, STEVENS G, et al. LCCDE: a decision-based ensemble framework for intrusion detection in the Internet of Vehicles[EB/OL]. (2022-09-01)[2023-12-19]. https://arxiv.org/abs/2208.03399.

[7] WU K H, CHEN Z G, LI W. A novel intrusion detection model for a massive network using convolutional neural networks[J]. IEEE Access, 2018, 6: 50850-50859.

[8] LI X H, HU Z Y, XU M F, et al. Transfer learning based intrusion detection scheme for Internet of Vehicles[J]. Information Sciences, 2021, 547: 119-135.

[9] DUTA I C, LIU L, ZHU F, et al. Improved residual networks for image and video recognition[EB/OL]. (2020-04-10)[2023-12-19]. https://arxiv.org/abs/2004.04989.

[10] WOO S, DEBNATH S, HU R H, et al. ConvNeXt V2: co-designing and scaling ConvNets with masked autoencoders[EB/OL]. (2023-01-02)[2023-12-19]. https://arxiv.org/abs/2301.00808.

[11] SANDLER M, HOWARD A, ZHU M L, et al. MobileNetV2: inverted residuals and linear bottlenecks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 4510-4520.

[12] SEO E, SONG H M, KIM H K. GIDS: GAN based in-

trusion detection system for in-vehicle network[C]//2018 16th Annual Conference on Privacy, Security and Trust (PST). Piscataway: IEEE, 2018: 1-6.

[13] ROSAY A, CARLIER F, LEROUX P. MLP4NIDS: an efficient MLP-based network intrusion detection for CIC-IDS2017 dataset[C]// International Conference on Machine Learning for Networking. Cham: Springer, 2020: 240-254.

[14] ALSHAMMARI A, ZOHDY M A, DEBNATH D, et al. Classification approach for intrusion detection in vehicle systems[J]. Wireless Engineering and Technology, 2018, 9(4): 79-94.

[15] OLUFOWOBI H, YOUNG C, ZAMBRENO J, et al. SAIDuCANT: specification-based automotive intrusion detection using controller area network (CAN) timing[J]. IEEE Transactions on Vehicular Technology, 2020, 69(2): 1484-1494.

[16] ASHRAF J, BAKHSHI A D, MOUSTAFA N, et al. Novel deep learning-enabled LSTM autoencoder architecture for discovering anomalous events from intelligent transportation systems[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(7): 4507-4518.

[17] SHARAFALDIN I, HABIBI LASHKARI A, GHORBANI A A. Toward generating a new intrusion detection dataset and intrusion traffic characterization[C]//Proceedings of the 4th International Conference on Information Systems Security and Privacy. Cham: Springer, 2018: 108-116.

[18] ABDULHAMMED R, FAEZIPOUR M, MUSAFER H, et al. Efficient network intrusion detection using PCA-based dimensionality reduction of features[C]//2019 International Symposium on Networks, Computers and Communications (ISNCC). Piscataway: IEEE, 2019: 1-6.

[19] ELMASRY W, AKBULUT A, ZAIM A H. Evolving deep learning architectures for network intrusion detection using a double PSO metaheuristic[J]. Computer Networks, 2020, 168: 107042.

[20] 张安琳, 张启坤, 黄道颖, 等. 基于 CNN 与 BiGRU 融合神经网络的入侵检测模型[J]. 郑州大学学报(工学版), 2022, 43(3): 37-43.
ZHANG A L, ZHANG Q K, HUANG D Y, et al. Intrusion detection model based on CNN and BiGRU fused neural network[J]. Journal of Zhengzhou University (Engineering Science), 2022, 43(3): 37-43.

[21] 吴正江, 杨天, 郑爱玲, 等. 融合拟单层覆盖粗集的集值数据平衡方法研究[J]. 计算机工程与应用, 2022, 58(19): 166-173.
WU Z J, YANG T, ZHENG A L, et al. Study on set-valued data balancing method by semi-monolayer covering rough set[J]. Computer Engineering and Applications, 2022, 58(19): 166-173.

[22] GAO S H, CHENG M M, ZHAO K, et al. Res2Net: a new multi-scale backbone architecture[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(2): 652-662.

[23] LIU Z, MAO H Z, WU C Y, et al. A ConvNet for the 2020s[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2022: 11966-11976.

An Intrusion Detection Method for Internet of Vehicles Based on Improved WGAN-GP and ResNet

WEI Mingjun[1,2], LI Feng[1], LIU Yazhi[1,2], LI Hui[1]

(1. College of Artificial Intelligence, North China University of Science and Technology, Tangshan 063210, China; 2. Hebei Provincial Key Laboratory of Industrial Intelligent Perception, Tangshan 063210, China)

Abstract: In order to protect the Internet of Vehicles system from the threat of network attacks and improve the accuracy of intrusion detection, a new intrusion detection method (AQVAE-RGSNet) was proposed for the characteristics of large data flow and unbalanced attack types in the vehicle network. Firstly, the adversarial quantized variational auto encoder was used to process the vehicle network data imbalance. And it was constructed by combining the vector quantized variational auto encoder-2 and the generative adversarial network with gradient penalty to alleviate the extremely unbalanced number of samples of abnormal attack types in the dataset. Afterwards, the ResNet and improved segmented residual neural network were used to learn the input sample data and predict its attack category. The experimental results indicated that AQVAE-RGSNet achieved $F1$ scores of 0.998 6 and 0.999 7 on the vehicle networking dataset CICDS2017 and CAN-intrusion-dataset, respectively. On the premise of ensuring the best training effect, it could identify attack threats more effectively in the vehicle network.

Keywords: Internet of Vehicles; intrusion detection; generate adversarial networks; residual neural network; feature fusion