DOI:10.19651/j.cnki.emt.2519108

基于 SqueezeNet 轻量化网络的车载 CAN 总线入侵检测方法*

樊 炳[1] 曹 燚[2]
(1.南京信息工程大学计算机学院 南京 210044; 2.无锡学院网络空间安全学院 无锡 214105)

摘 要: 针对现有基于深度学习的 CAN 总线入侵检测方法普遍存在结构复杂、资源开销大和延迟较高的问题,本文提出一种基于改进 SqueezeNet 的轻量化 CAN 总线入侵检测模型。首先,将 CAN 报文数据转换为彩色图像,以增强其空间和通道特征的表达能力;其次,引入高效通道注意力(ECA)机制,加强对异常通信特征的细粒度建模;然后,对网络结构进行优化,采用深度可分离卷积和 Ghost 模块替代标准卷积,裁剪冗余层次以降低计算开销和参数量;最后,统一采用 Hardswish 激活函数,提升模型非线性表达能力与推理效率。在 Car-Hacking 公开数据集上的实验结果表明,所提方法达到 100%的检测准确率,模型大小仅为 0.35 MB,平均响应时间为 1.6 ms,具备高性能、低延迟及低资源占用的部署优势。

关键词: 控制器局域网络;车载网络;SqueezeNet;入侵检测;注意力机制;轻量化
中图分类号: TP393;TN915.08 **文献标识码**: A **国家标准学科分类代码**: 510.4010;520.2040

Vehicle CAN bus intrusion detection method based on SqueezeNet lightweight network

Fan Bing[1] Cao Yi[2]
(1. School of Computer Science, Nanjing University of Information Science and Technology, Nanjing 210044, China;
2. School of Cyber Science and Engineering, Wuxi University, Wuxi 214105, China)

Abstract: To address the common issues of complex architecture, high resource consumption, and significant latency in existing deep learning-based CAN bus intrusion detection methods, this paper proposes a lightweight CAN bus intrusion detection model based on an improved SqueezeNet architecture. First, CAN message data is converted into color images to enhance spatial and channel feature representation. Second, an efficient channel attention (ECA) mechanism is introduced to enable fine-grained modeling of anomalous communication patterns. Third, the network architecture is optimized by replacing standard convolutions with deep separable convolutions and Ghost modules, while pruning redundant layers to reduce computational overhead and parameter count. Finally, the Hardswish activation function is uniformly applied to enhance nonlinear expressiveness and inference efficiency. Experimental results on the Car-Hacking public dataset demonstrate that the proposed method achieves 100% detection accuracy with a model size of only 0.35 MB and an average response time of 1.6 ms, offering deployment advantages of high performance, low latency, and minimal resource consumption.

Keywords: controller area network;in-vehicle network;SqueezeNet;intrusion detection;attention mechanism;lightweight

0 引 言

为了实现自动驾驶、主动安全等智能功能,越来越多的电子元件被集成至车载系统中。车载局域网络(in-vehicle network,IVN)作为车内电子控制单元(electronic control unit,ECU)之间的主要通信方式,其安全性已成为保障整车功能稳定运行的关键环节[1-2]。当前主流的 IVN 通信协议为控制器局域网络(controller area network,CAN),其因

收稿日期:2025-06-16
*基金项目:无锡市"太湖之光"科技攻关(基础研究)项目(K20241046)、2023 年度江苏高校哲学社会科学研究一般项目(2023SJYB0919)、国家传感网工程技术研究中心开放课题基金(2024YJZXKFKT02)、无锡学院高层次人才科研启动专项经费(2022r043)资助

高效、实时与成本低廉等优势被广泛部署于动力系统、底盘控制系统和信息娱乐系统之中。值得注意的是，CAN 总线在设计之初并未考虑信息安全防护，其通信机制为开放式广播，既无认证机制，也不具备数据加密功能[3-4]。这一缺陷使得车辆在暴露于外部网络环境中后，极易遭受攻击。

近年来，车联网(internet of vehicles，IoV)概念的提出与实践不断深化，车辆与云端、基础设施及其他车辆之间的通信日益频繁，系统攻击面也随之显著扩大[5-6]。攻击者可通过物理接口或无线信道侵入 CAN 网络，实施拒绝服务、伪装或模糊测试等攻击，进而干扰车辆正常通信流程，可能引发刹车失灵、转向受控等严重交通安全事故[7-8]。

当前，基于深度学习的车载网络入侵检测方法已成为研究热点。深度学习方法的核心优势在于能够自动提取深层特征，从而在处理复杂和大规模数据时实现高效识别，因此具有广阔的应用前景，但同时也存在模型复杂度高与泛化能力不足等问题。已有研究在不同角度对检测方法进行了探索。在特征提取与分类模型融合方面，周志豪等[9]提出了一种基于 SMOTE-SDSAE-SVM 的检测方法，通过过采样缓解类别不平衡问题，并利用堆叠去噪自编码器提取特征，再结合支持向量机完成分类。该方法在不平衡数据场景下表现出较好的检测效果。Khan 等[10]提出了基于集成学习的 DivaCAN 方法，将深度神经网络、MLP、LightGBM 和随机森林等多种分类器融合，通过投票机制提升检测的稳定性，在多个公开数据集上的结果显示均优于单一模型。在时序建模方法方面，银鹰等[11]提出了一种基于 LSTM 的 CAN 总线入侵检测模型。该方法利用循环神经网络对报文间的时间依赖进行建模，能够识别 DoS、重放与模糊攻击。实验结果表明，该模型在多种攻击场景下均取得了较高的检测效果。许秀锋等[12]针对 LSTM 计算开销较大的问题，提出了一种基于 GRU 的检测方法。通过简化循环结构，该方法在减少参数量和训练时间的同时，仍保持了与 LSTM 接近的检测准确率，在资源受限的车载设备中更具应用潜力。Hossain 等[13]利用真实车辆采集的正常通信数据，并通过注入 DoS、模糊和欺骗等攻击流量构建实验数据集，提出了一种基于 LSTM 的入侵检测系统。该方法在实际采集数据上进行训练和验证，能够同时完成二分类与多分类任务，实验结果显示在不同攻击场景下均取得了较高的检测准确率。Taneia 等[14]提出了 ACL-IDS 模型，在 CNN-LSTM 架构的基础上引入注意力机制。该方法通过卷积层提取局部特征，结合 LSTM 建模时序特征，并利用注意力机制突出关键信息。实验结果表明，该模型在多类攻击检测任务中整体精度较高。在深层特征融合与跨场景适应方面，陈彦彬等[15]提出了一种基于双线性自注意力机制的 CAN 入侵检测方法。该模型先通过 DNN、CNN 和 LSTM 捕捉多层次的空间与时序特征，再引入 Transformer 与 FNet 结构构建双线性自注意力模块，以增强特征融合能力。与此同时，残差连接被用于缓解深层特征提取中的梯度消失问题，从而保证训练的稳定性。该方法在 Car-Hacking 数据集和物联网实验平台上均完成了验证，能够有效应对不同类型的攻击。李向荣等[16]提出了一种基于 BERT 与迁移学习的自适应入侵检测方法。该方法首先将 CAN 报文的标识符序列转化为上下文语义输入，由 BERT 模型进行双向特征建模，并利用掩码语言模型实现无监督检测。随后引入迁移学习策略，使模型能够在不同车型的数据上进行快速适配，减少从零开始训练的需求。相关实验覆盖 DoS、模糊与欺骗等多类攻击场景，验证了方法在跨车型检测任务中的适用性。Le 等[17]提出了一种基于自编码器与时间嵌入式 Transformer 的 CAN 总线入侵检测方法。该方法在报文级别通过自编码器提取数据帧特征，并利用时间戳嵌入的 Transformer 对序列进行建模，以适应 CAN 报文时间间隔不均匀的特点。在 Car-Hacking 与 ROAD 等数据集上的实验验证表明，该方法能够有效处理不同类型的入侵场景。

总体而言，现有基于深度学习的车载 CAN 总线入侵检测方法虽在特征提取与攻击识别方面取得了积极进展，但普遍存在结构复杂、资源开销大和延迟较高的问题，难以满足车载环境对轻量化与实时性的需求。为此，本文提出了一种基于改进 SqueezeNet 的轻量化 CAN 总线入侵检测模型。主要工作包括：

1)在主干网络的输入层及 Fire 模块结构中，引入深度可分离卷积和 Ghost 模块，通过卷积分解与冗余特征生成，替代传统标准卷积操作，在保持特征提取能力的基础上，有效减少模型参数量与计算开销。

2)在特征提取网络的高层输出阶段嵌入高效通道注意力机制(efficient channel attention，ECA)，通过局部一维卷积建模通道间依赖关系，引导模型更加关注关键通信特征维度，提升对 CAN 报文中异常模式的表征能力，增强检测精度与鲁棒性。

3)使用 Hardswish 激活函数代替传统 ReLU 激活方式，以提升模型的非线性建模能力与梯度传播效率，实现更稳定的训练过程与更快的推理响应，进一步提升模型在车载终端上的部署性能。

1　相关知识

1.1　CAN 总线原理

控制器局域网络由德国博世公司于 20 世纪 80 年代提出，旨在满足汽车内部电子控制单元之间对于高效、实时、可靠通信的需求[18]。

在通信机制方面，CAN 协议采用广播式传输，所有节点均能接收总线上传输的消息。每条报文通过标识符(identifier，ID)标明数据类别，同时确定优先级，ID 数值越小，优先级越高。为避免通信冲突，CAN 引入了非破坏性位仲裁机制，在总线竞争时，由 ID 优先级高的节点继续发送，其余节点自动退出，确保总线资源被高优先级信息

占用[19-21]。

标准 CAN 数据帧结构如图 1 所示。

Length	1 bit	12 bit	6 bit	0 to 8 bytes	16 bits	2 bits	7 bits	3 bits
Desc.	Start Of Frame	Arbitration Field	Control Field	Duta Field	CRC	ACK	End of Frame	Inter Frame Space

图 1　CAN 帧格式

Fig. 1　CAN frame format

Start of Frame(帧起始):表示数据帧的开始,用于同步通信。

Arbitration Field(仲裁字段):用于仲裁总线优先级,包含标识符和远程请求位。

Control Field(控制字段):指明数据长度并包含帧的控制信息。

Data Field(数据字段):存放实际传输的数据内容,最大为 8 字节。

CRC(循环冗余校验):用于检测帧在传输过程中是否发生错误。

ACK(应答字段):接收节点用来确认是否正确接收到该数据帧。

End of Frame(帧结束):标识数据帧传输的结束,保持总线稳定。

Inter Frame Space(帧间隔):帧与帧之间的间隔时间,便于节点准备下一帧。

1.2　CAN 总线攻击类型

由于 CAN 总线采用开放式广播通信机制,缺乏报文认证与加密措施,使其在面向外部通信接口时容易成为攻击目标[22]。攻击者一旦接入总线,即可监听或注入伪造报文,从而影响车辆的正常运行。根据攻击方式的不同,CAN 总线面临的主要威胁可归纳为以下 3 类:

1)伪造攻击(Spoofing)

攻击者主动构造格式合法的伪造报文,冒充车内某一节点发送控制或状态信息。如图 2 所示,攻击者可伪造高优先级标识符(ID)并在总线仲裁中获得主导权,从而抑制其它节点的正常发送并改变总线上的消息流。伪造方式可表现为持续占用(长期注入高优先级报文)、间歇注入(周期性或按需注入以保持隐蔽)或时序精确的短脉冲注入(针对特定控制事件触发),其后果可能包括状态信息错乱、控制逻辑被误导或执行单元发生异常动作,严重时可影响车辆安全功能的正常运行。

2)拒绝服务攻击(DoS)

如图 3 所示,攻击者通过持续注入频繁或高优先级的无效报文,占用总线带宽,导致其他节点无法完成正常通信,严重时可能导致整个车载网络陷入瘫痪状态。此类攻击通常利用 CAN 总线仲裁机制的特性,通过伪造高优先级 ID 报文实现持续占用,从而阻塞总线资源,造成低优先级节点丧失发送机会。该过程不仅影响数据传输的实时性,还可能引发关键安全控制报文延迟或丢失,进而危及车辆的运行安全。

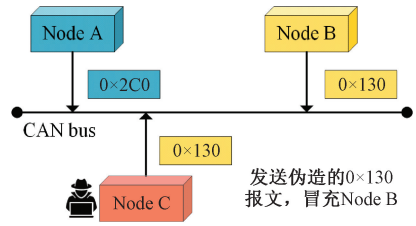


图 2　Spoofing 攻击

Fig. 2　Spoofing attacks

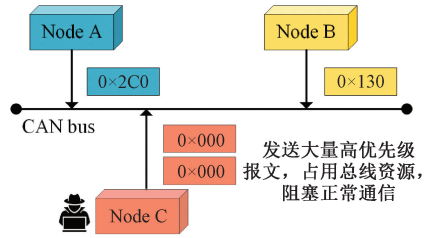


图 3　Dos 攻击

Fig. 3　Dos attacks

3)模糊测试攻击(Fuzzing)

模糊测试攻击是一种通过注入异常或非法格式报文来探索系统未知漏洞的手段。如图 4 所示,攻击者通过随机注入包含边界值、非法标识符或随机数据的报文,诱发车载系统中未覆盖的异常处理分支或故障状态,因而具有较强的不可预测性。该方法通过构造大量异构输入试图触发软件解析错误或未处理的异常分支。在车载环境中,持续或高频的模糊注入还可能引起 ECU 解析异常、缓冲区错误或占用总线带宽,进而导致关键控制帧的延迟或丢失并引发功能异常。

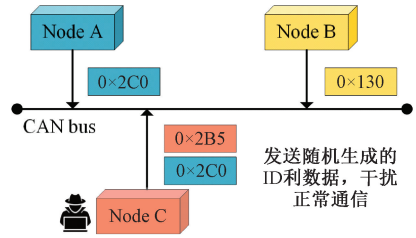


图 4　Fuzzy 攻击

Fig. 4　Fuzzy attacks

针对上述威胁,CAN 总线安全研究正逐步从被动防御转向主动检测。其中,基于通信特征提取的入侵检测系统(intrusion detection system,IDS)已成为提升整车通信安全性的核心防护手段之一,在保障通信过程的完整性与可靠性方面展现出较高应用价值[23]。

1.3　SqueezeNet 轻量级模型

SqueezeNet 是一种轻量级卷积神经网络架构,其整体结构如图 5 所示。该网络由输入层、卷积层与池化层、Fire 模块堆叠层以及最终的分类层组成。输入首先经过一个标

准卷积层完成初步特征提取，随后通过池化操作缩小空间维度，降低计算开销。在主干部分，多个 Fire 模块依次连接，在保持较低参数量的同时逐步增强特征表达能力，形成从浅层到深层的逐级抽象。为了进一步压缩模型规模，网络在后段采用 1×1 卷积取代大卷积核操作，并通过全局平均池化实现特征聚合，最后由 Softmax 层完成分类输出。

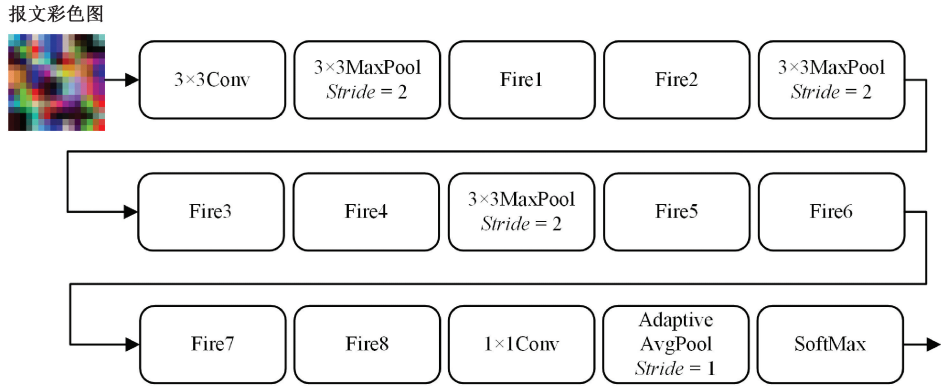


图 5　SqueezeNet 网络模型结构

Fig. 5　SqueezeNet network model structur

SqueezeNet 的主要结构为卷积层、Fire 模块、全局平均池化层和分类层，具体的网络结构如表 1 所示。

表 1　SqueezeNet 网络模型结构

Table 1　SqueezeNet network model structure

输入尺寸	操作	卷积核大小	步长
224×224×3	Conv2d	3×3	2
112×112×64	MaxPool2d	3×3	2
56×56×64	Fire	3×3,1×1	1
56×56×128	Fire	3×3,1×1	1
56×56×128	MaxPool2d	3×3	2
28×28×128	Fire	3×3,1×1	1
28×28×256	Fire	3×3,1×1	1
28×28×256	MaxPool2d	3×3	2
14×14×256	Fire	3×3,1×1	1
14×14×384	Fire	3×3,1×1	1
14×14×384	Fire	3×3,1×1	1
14×14×512	Fire	3×3,1×1	1
14×14×512	Conv2d	1×1	1
14×14×1 000	AdaptiveAvgPool2d	1×1	1

Fire 模块的结构包含 Squeeze 层与 Expand 层两部分。Squeeze 层通过 1×1 的卷积核对输入张量进行通道压缩，将输入通道数 N 减小到 M，从而降低通道数减少网络的计算复杂度。Expand 层包含 1×1 和 3×3 两种卷积核，分别将 Squeeze 层输出的 M 通道特征图扩展为 E_1 和 E_2 通道特征图。最后将两种卷积核卷积得到的特征图进行拼接，输出通道数为 E_1+E_2 的特征图。

其中 Fire 模块的具体结构如图 6 所示。

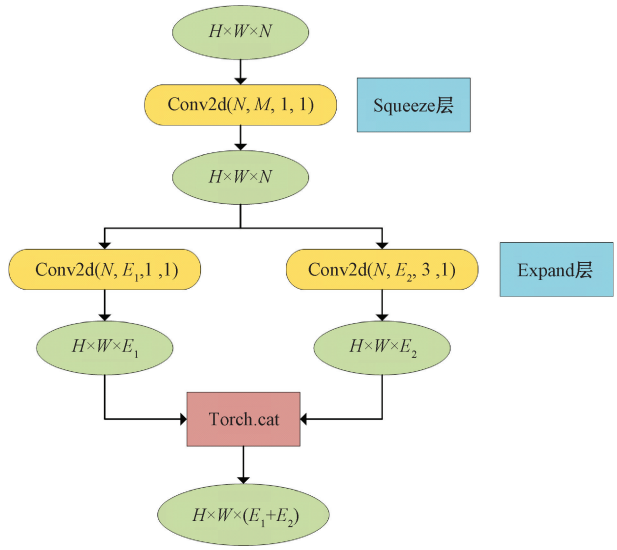


图 6　Fire 模块

Fig. 6　Fire module

2　车载 CAN 总线入侵检测模型的构建与改进

2.1　入侵检测流程

为提升车载 CAN 总线环境中入侵检测系统的响应效率与部署可行性，本文提出一种融合注意力机制与轻量级卷积神经网络的入侵检测方法，由数据处理模块与改进的轻量卷积神经网络模型 SqueezeNet 组成（如图 7 所示）。入侵检测流程如下：数据预处理模块首先对原始 CAN 报文进行清洗与筛选，随后完成特征提取与格式转换，生成可用于模型输入的图像化数据帧，划分训练集和测试集，输入模型进行训练，最后将模型进行入侵行为的检测。

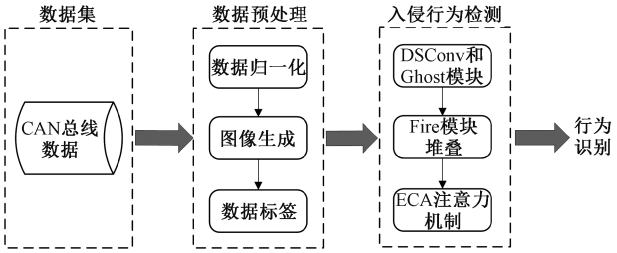


图 7　本文模型结构

Fig. 7　Model structure of this paper

2.2　数据预处理

考虑到卷积神经网络在图像识别任务中的卓越性能，SqueezeNet 是一种旨在显著减少模型参数数量的同时保持卷积神经网络准确率的轻量级网络结构，将原始网络通信报文数据进行图像化转换，利用 SqueezeNet 网络模型进行特征提取与行为分类。

将 CAN 总线报文数据转换过程首先进行数据清洗，对原始 CAN 总线数据中的缺失值进行填充，统一采用 0 进行补全；删除重复记录以避免样本冗余；将报文中表示标识符的字符串字段编码为数值型标签；并将包含字母的十六进制数转换为十进制，以满足后续模型处理的输入格式要求。

在数据清洗后，为适配图像模型的输入要求，需将网络流量数据归一化至 0～255 范围。本文采用分位数归一化方法，将特征分布转换为标准正态分布并重新计算特征值，使数据集中分布于中值附近，同时有效抑制异常值干扰，提升输入特征的鲁棒性和表征一致性。

在 Car-Hacking 数据集中，每条网络流量记录包含 9 个关键字段，包括 CAN 标识符(CAN ID)和 8 个数据字节(DATA[0]～DATA[7])。归一化处理后，将每连续 27 条报文重组为一个 27×9 的数据矩阵，并按列划分为 3 个部分，分别对应图像的 R、G、B 通道。各通道数据 reshape 为 9×9 矩阵后在通道维度拼接，生成 9×9×3 的彩色图像，随后插值放大为 224×224 用于模型输入。图像标签采用动态众数标注策略：若时间窗内全为正常流量则标记为 Normal；若存在攻击样本，则以该窗口内出现频次最高的攻击类型作为图像标签。

对于 Car-Hacking 数据集，由图 8 可以看出，不同类型的攻击样本在图像中的表现形式差异明显。Fuzzy 攻击通过随机生成非法但格式合法的 CAN 报文，其图像呈现出颜色高度离散、像素分布杂乱的随机图案；DoS 攻击通过持续发送空载荷或重复报文以阻塞 CAN 总线，生成的图像几乎无颜色信息，表现为近似纯黑的图像；Gear 和 RPM 欺骗攻击则通过伪造特定的 CAN ID 与数据字段，其图像表现为规则的颜色条纹或重复块状结构，具有明显的伪装模式。这些差异性的视觉表现有效增强了模型对不同攻击行为的区分能力。

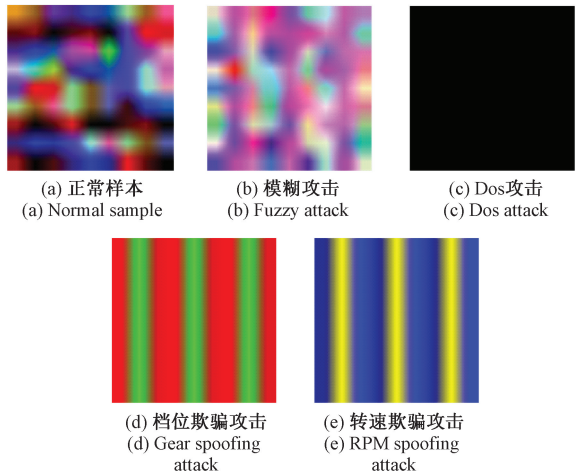


图 8　样本图片

Fig. 8　Sample image

2.3　改进的 SqueezeNet 轻量化模型

1) 改进基本结构

为提升 SqueezeNet 模型在嵌入式和移动设备中的计算效率和实时性，本文首先对其基本结构进行了系统优化。针对模型中标准卷积计算复杂度高、模型体积大的问题，本文引入了深度可分离卷积(depthwise separable convolution，DSConv)和 Ghost 模块(Ghost module)作为替代方案。

DSConv 原理如图 9 所示，通过将标准卷积分解为逐通道卷积和逐点卷积[24]，在保持空间特征提取能力的同时，显著降低了计算量和参数数量。

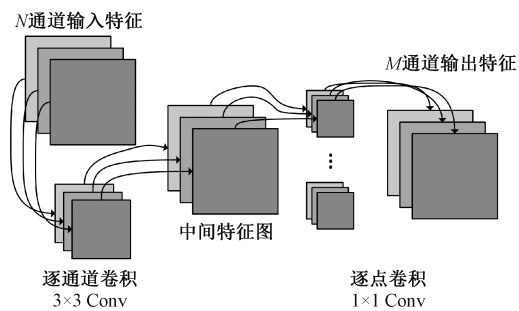


图 9　深度可分离卷积

Fig. 9　Depthwise separable convolution

在传统的二维卷积中，假设输入特征图大小为 $H\times W$，输入通道数为 N，输出通道数为 M，卷积核大小为 $K\times K$，则其参数量为：

$$Params_{standard}=N\times M\times K\times K \tag{1}$$

$$FLOPs_{standard}=H\times W\times N\times M\times K\times K \tag{2}$$

而采用深度可分离卷积后，卷积操作被分解为两步：第 1 步是对每个输入通道进行空间卷积(depthwise)，第 2 步是通过 1×1 卷积实现通道融合(pointwise)，其计算量为：

$$Params_{dsconv} = N \times K \times K + N \times M \tag{3}$$

$$FLOPs_{dsconv} = H \times W \times (N \times K \times K + N \times M) \tag{4}$$

由此可见，深度可分离卷积在保持特征提取能力的同时，能够显著减少参数总量与计算负担，尤其在通道数较大时效果尤为明显[25]。

为进一步降低模型计算复杂度与参数量，本文引入 Ghost 模块用于高效特征图生成。Ghost 模块的核心思想是先通过 1×1 卷积提取部分本征特征图，再利用一组线性变换操作生成冗余特征图，以替代标准卷积生成全部特征图的过程[26]。如图 10 所示，输入特征首先经过逐点卷积(PWConv)得到主特征图，随后通过一系列廉价操作(φ_1，φ_2，…)扩展为完整的输出特征图。该结构在保留特征表达能力的同时，显著降低了计算开销与模型体积。

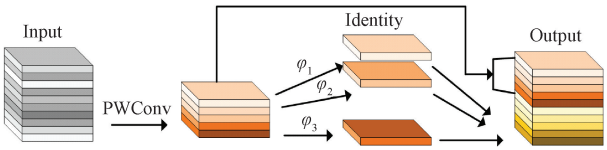


图 10　Ghost 模块

Fig. 10　Ghost module

2) 激活函数的改进

在深度神经网络中，激活函数是模型表达非线性特征、增强拟合能力的重要组成部分。传统 SqueezeNet 网络采用 ReLU(rectified linear unit)作为激活函数，其定义为：

$$f(x) = \begin{cases} x, & x \geqslant 0 \\ 0, & x < 0 \end{cases} = \max(0, x) \tag{5}$$

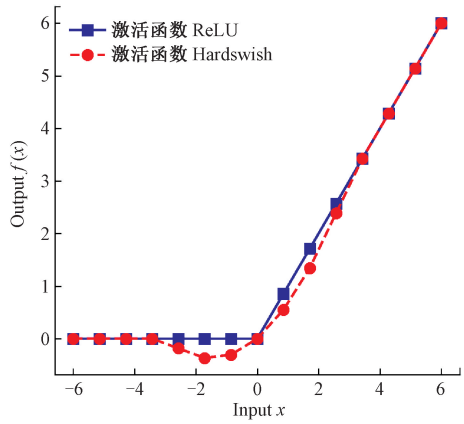
其中 x 表示输入变量。ReLU 激活函数在输入为负时，输出恒为零，容易导致部分神经元在训练中“失活”，即无法继续更新梯度，降低模型表达能力和训练效率。同时，ReLU 在输入接近零点时梯度不连续，可能造成训练不稳定、收敛速度减慢，影响模型的整体性能。

为了解决这些问题，本文采用 Hardswish 激活函数替代 ReLU。Hardswish 的定义为：

$$f(x) = \begin{cases} 0, & x \leqslant 3 \\ \dfrac{(x+3)^2}{6}, & -3 < x < 3 \\ x, & x \geqslant 3 \end{cases} \tag{6}$$

Hardswish 在整个输入区间内提供了平滑、连续的非线性映射，避免了 ReLU 在负区间神经元“失活”及梯度不连续的问题。它不仅提升了模型训练过程的稳定性和收敛速度，还保留了 ReLU 的计算效率和简洁实现，特别适合在嵌入式和边缘计算设备中部署。

图 11 展示了 Hardswish 和 ReLU 激活函数在不同输入区间的响应曲线对比。Hardswish 在负区间和接近零点处的过度相对平滑，显著减少了训练中的神经元“死区”现象，提升了模型的泛化能力和稳定性。



注：$f(x)$为激活函数因变量，x 为激活函数变量。

图 11　激活函数对比图

Fig. 11　Activation function comparison chart

3) 通道注意力机制 ECA 的引入

SqueezeNet 原始结构由于高度压缩参数，未设计用于建模通道间的相关性。在实际应用中，特定通道可能携带更具判别性的特征，因此通道注意力机制可用于增强模型对关键通道的响应能力。

为增强关键通道的判别能力，进一步提升网络在入侵检测任务中的通道选择能力，本文在 SqueezeNet 网络的最后一组 Fire 模块后引入了 ECA 机制。ECA 模块通过轻量的一维卷积实现通道间交互建模[27]，相较于传统的 squeeze-and-excitation(SE)机制，其无需引入全连接层，计算更为高效，参数更为紧凑，特别适用于嵌入式和车载等资源受限的实际部署环境。通过在最后阶段引入 ECA，模型在保持高效计算的同时，增强了对关键通道的选择与响应能力，从而提升了整体检测性能。ECA 模块的计算流程如下，如图 12 所示。

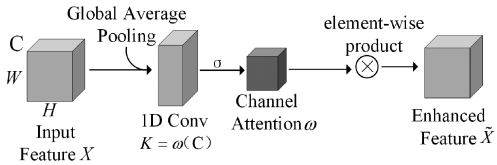


图 12　ECA 通道注意力机制结构

Fig. 12　Structure of the ECA channel attention mechanism

(1)通道特征压缩：对输入特征图 $X \in R^{C \times H \times W}$ 进行全局平均池化，提取每个通道的统计信息，形成通道描述向量 $z \in R^c$：

$$z_c = \frac{1}{H \cdot W} \sum_{i=1}^{H} \sum_{j=1}^{W} X_c(i, j) \tag{7}$$

(2)通道间依赖建模：对通道描述向量 z 应用一维卷积(卷积核大小为 K)，建模通道之间的局部交互关系，并通过 sigmoid 函数生成通道注意力权重 $\omega \in R^C$：

$$\omega = \sigma(Conv1D(z)) \tag{8}$$

(3)通道重加权增强：将原始特征图与通道注意力权

重进行逐通道加权融合,得到增强特征图。

$$\tilde{X}_c = \omega_c \cdot X_c \tag{9}$$

所有改进后的 SqueezeNet 网络结构如图 13 所示,其中虚线框部分为改进部位。

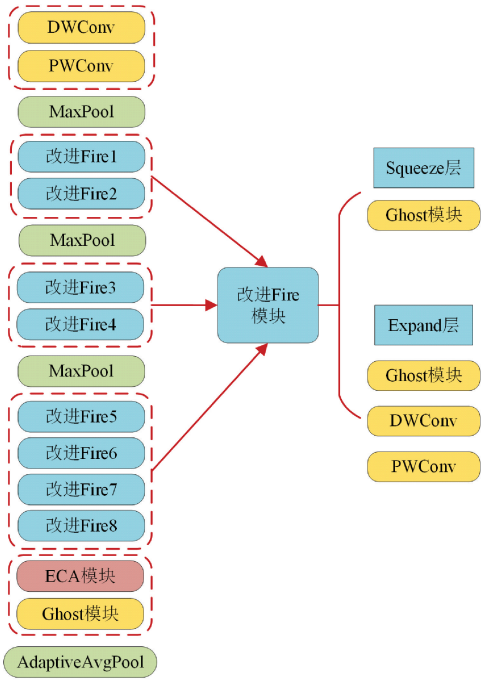


图 13　改进后 SqueezeNet 网络结构

Fig. 13　Improved SqueezeNet network structure

3　实验与分析

实验环境:操作系统为 Windows 11(64 位)CPU 为 Intel Core i9-14900HX@3.5 GHz;内存大小为 16 GB;GPU 为 NVIDIA GeForce RTX 4060 Laptop 8 GB;深度学习框架为 Pytorch。

3.1　数据集

数据集选取了韩国高丽大学 HCRL 实验室的 Car-Hacking 公开车载网络数据集[28]。Car-Hacking 数据集包含正常数据、拒绝服务攻击(DoS)、模糊攻击(Fuzzy)和伪造 CAN ID 的欺骗攻击(Spoofing)。具体攻击类型和数量分布如表 2 所示。

表 2　攻击数据组成

Table 2　Attack data composition

数据集	攻击类型	数量	Label
Car-Hacking	Normal	988 872	0
	RPM Spoofing attack	654 897	1
	Gear Spoofing attack	597 252	2
	DoS attack	587 521	3
	Fuzzy attack	491 847	4

3.2　超参数设置

网络训练过程中,批量大小为 32,训练轮数为 5,模型优化器为 Adam,初始学习率设定为 0.000 1,损失函数采用类别均衡焦点损失函数。

3.3　评估指标

车载数据中攻击与正常行为的严重类别不平衡导致单一准确率指标不足以全面反映模型性能,因此本文引入多种评价指标对检测结果进行评估。所选指标包括准确率(Accuracy)、精确率(Precision)、召回率(Recall)与 $F1$ 分数(F1-score),计算公式如式(10)~(13)所示,公式中相关符号的含义如表 3 所示。

$$\text{Accurary} = \frac{TP + TN}{TP + FP + TN + FN} \tag{10}$$

$$\text{Precison} = \frac{TP}{TP + PF} \tag{11}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{12}$$

$$F1\text{-score} = 2 \times \frac{Precison \times Recall}{Precison + Recall} \tag{13}$$

表 3　混淆矩阵

Table 3　Confusion matrix

类型	识别为攻击	识别为正常
原始为攻击流量	TP	FN
原始为正常流量	FP	TN

此外,本实验指标还包括推理时间(inference time)和模型大小(model size)。

3.4　实验及结果分析

1) 时序模型对比

为验证报文图像化表示相较于直接时序输入的效果,实验选取了一维卷积网络(one-dimensional convolutional neural network,1D-CNN)、LSTM、门控循环单元(gated recurrent unit,GRU)和时序卷积网络(temporal convolutional network,TCN)进行对比。这些模型均直接处理 CAN 报文序列,不经过图像化转换,用于评估不同输入方式对检测性能的影响。实验中,各模型在相同的数据预处理和划分策略下训练,输入为窗口长度 27 的 CAN 报文序列(特征维度为十进制 CAN ID 与 8 字节数据,共 9 维),其参数规模和训练配置保持一致,结果如表 4 所示。

表 4 给出了不同时序模型与所提方法在 Car-Hacking 数据集上的性能对比结果。整体来看,1D-CNN、LSTM、GRU 和 TCN 在各项指标上均超过 99%,说明时序建模方法在 CAN 总线入侵检测任务中能够取得较高的检测性能。然而,从精度(P/%)、召回率(R/%)和 F1-score 等指标来看,不同模型之间仍存在一定差异。例如,1D-CNN 与 LSTM 的召回率略低于 100%,部分攻击样本检测表现略低;而 GRU 与 TCN 的精度低于 99.2%,在正常与攻击

表 4　不同时序模型与所提方法在 Car-Hacking 数据集上的检测性能对比

Table 4　The detection performance comparison of different sequence models and the proposed method on the Car-Hacking dataset　%

模型	Acc	P	R	F1-score
1D-CNN	99.25	99.45	99.79	99.62
LSTM	99.32	99.55	99.76	99.66
GRU	99.12	99.13	99.98	99.56
TCN	99.11	99.11	100	99.55
本文模型	**100**	**100**	**100**	**100**

报文区分方面表现相对较弱。相比之下，所提方法在 4 个指标上均达到 100%，实现了更全面和稳定的检测效果，进一步验证了图像化表示结合改进卷积结构在车载入侵检测任务中的适用性。

2）卷积网络模型对比

本文在统一的数据预处理策略基础上，选取了两类具有代表性的模型进行对比：第 1 类为结构相对复杂的卷积神经网络，包括 ResNet-18 和 EfficientNet-B0；第 2 类为经典轻量化卷积网络，包括 ShuffleNetV2 和 MobileViT；上述模型将在数据集上完成训练与评估，结果如表 5 所示。

由表 5 实验结果可知，所提出的改进 SqueezeNet 模型在 Car-Hacking 数据集上表现良好。与 ResNet-18、EfficientNet、ShuffleNetV2 和 MobileViT 等对比模型相比，本文模型在检测准确率、精确率、召回率和 F1 分数方面均达到 100%，与 ResNet-18 持平，优于其他对比模型。

表 5　不同卷积网络模型与所提方法在 Car-Hacking 数据集上的检测性能对比

Table 5　Detection performance of different convolutional models and the proposed method on the Car-Hacking dataset

模型	Acc/%	P/%	R/%	F1-score/%	Inference Time/ms	Model size/MB
ResNet-18	100	100	100	100	2.5	42.72
EfficientNet	99.83	99.83	99.83	99.83	4.0	15.59
ShuffleNetV2	99.57	99.56	99.54	99.55	2.1	4.96
MobileViT	99.84	99.82	99.85	99.83	3.1	19.01
本文模型	**100**	**100**	**100**	**100**	**1.6**	**0.35**

从检测效率角度看，本文模型的推理时间为 1.6 ms，较 ResNet-18 的 2.5 ms 和 EfficientNet 的 4.0 ms 显著缩短，检测速度提升分别约 36% 和 60%，同时也优于 ShuffleNetV2 和 MobileViT 的 2.1 ms 和 3.1 ms。这主要得益于模型中采用了 Ghost 模块和深度可分离卷积架构，有效降低了计算复杂度和参数量。模型大小方面，本文模型仅为 0.35 MB，显著小于对比模型，减少了存储需求和计算负担，验证了其在嵌入式与资源受限环境中的应用价值。车载 CAN 总线异常检测对实时性有较高要求，通常要求每个数据包的检测时间不超过 10 ms，即入侵检测系统需在 10 ms 内完成判断与响应[29]。本文方法的平均检测时间低于该标准，能够满足车载 CAN 总线入侵检测的实时性需求。

3）其他方法对比

在对本文方法的可行性进行实验验证和进一步分析之后，为了能够更加全面地评估其相较于其他方法的效果，本文分别选取具有代表性的方法进行了对比实验，具体结果如表 6 所示。

通过表 6 可知，本文方法在 Car-Hacking 数据集上展现出优异的综合性能。在保持检测准确性的前提下，模型平均推理时间仅为 1.6 ms，模型体积控制在 0.35 MB，远低于其他对比方法，展现出显著的轻量化与高效率优势。模型能够快速响应 CAN 报文输入，有助于满足车载控制系统对实时性与资源占用的严格要求，适合在实际应用场景中部署。

表 6　对比实验结果

Table 6　Comparative experimental results

方法	Inference Time/ms	Model Size/MB
CANTransfer[30]	82.2	57.48
RecCNN[31]	5.5	2.02
SupCon-ResNet[32]	5.9	2.67
TIDL-IDS[33]	10.6	1.72
Our method	**1.6**	**0.35**

进一步观察各项指标，本文方法在推理效率与模型体积方面表现更优。与 CANTransfer[25] 相比，推理时间减少约 98%，模型体积缩小约 99%；与 RecCNN[26] 相比，推理时间降低约 71%，模型体积减少约 83%；与 SupCon-ResNet[27] 相比，推理时间减少约 73%，模型体积缩小约 86%；与 TIDL-IDS[28] 相比，推理时间减少约 85%，模型体积缩小约 80%。结果表明，本文方法在保持检测性能的同时，具有更高的运行效率与更小的模型规模，适用于对资源和响应速度要求严格的车载场景。

4）消融实验

为进一步分析改进的 SqueezeNet 模型中各改进点对模型性能的影响，本文设计并开展了消融实验，分别去除或保留深度可分离卷积（DSConv）、Ghost 模块、Hardswish 激活函数和 ECA 注意力机制，并在 Car-Hacking 数据集上

进行对比测试。不同模块组合下模型在准确率、测试时间和模型大小方面的结果如表 7 所示。

表 7　Car-Hacking 数据集消融实验结果

Table 7　Results of ablation experiments for the Car-Hacking dataset

DSConv	Ghost	Hardswish	ECA	Acc/%	Inference Time/ms	Model Size/MB
×	×	×	×	99.29	2.5	1.29
√	×	×	×	98.43	2.0	0.42
√	√	×	×	99.08	1.9	0.35
√	√	√	×	99.97	1.6	0.35
√	√	√	√	**100**	**1.6**	**0.35**

由表 7 可知,baseline 模型未引入任何模块,准确率为 99.29%,推理时间为 2.5 ms,模型大小为 1.29 MB。在此基础上,加入 DSConv 和 Ghost 模块后,准确率下降至 98.43%,推理时间缩短至 2.0 ms,模型大小显著减小至 0.42 MB。进一步引入 Hardswish 模块,准确率回升至 99.08%,测试时间降至 1.6 ms,模型大小减少至 0.35 MB。最后,在引入 ECA 模块后,准确率提升至 100%,推理时间保持不变,模型大小保持在 0.35 MB。结果表明,虽然 DSConv 和 Ghost 模块的引入使准确率有所下降,但有效减少了模型体积和推理时间,而 Hardswish 和 ECA 的加入在保持模型高效性的同时并提升了准确率。

5) 模型可解释性分析

为进一步探讨所提模型在分类过程中的判别依据,并增强模型的可解释性,采用 Grad-CAM(gradient-weighted class activation mapping)对训练完成的模型进行可视化分析。实验选取了正常报文及多种典型攻击样本,包括 DoS、Fuzzy、RPM 欺骗攻击与 Gear 欺骗攻击,将模型在分类决策过程中卷积层的激活信息映射到输入图像上,生成原图、热力图以及叠加图三联可视化结果,用于直观呈现模型在不同类别样本中的关注区域,如图 14 所示。

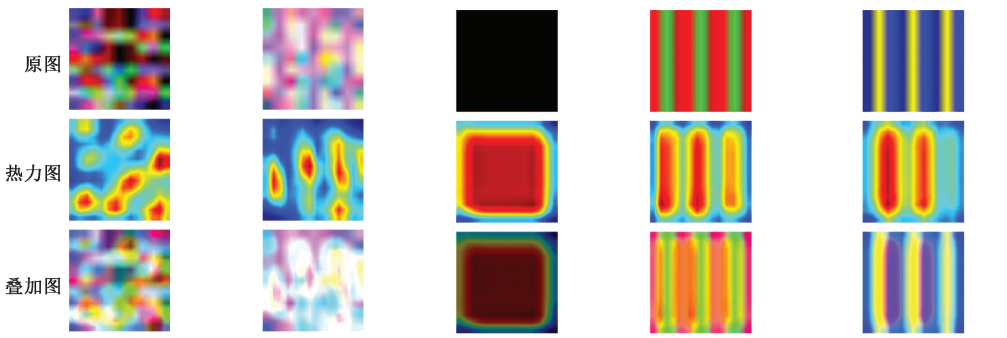


图 14　不同报文类型的 Grad-CAM 可视化结果

Fig. 14　Grad-CAM visualization results for different CAN message categories

从 Grad-CAM 可视化结果可以看出,模型在不同报文类型上的关注区域与各类攻击的特征分布具有较好的一致性。正常样本的激活区域呈现多点分散分布,且未形成连续或高度集中的热点,说明其通信数据不包含稳定的异常模式;模糊攻击的激活区域沿纵向呈弥散分布,反映出其扰动随机、缺乏稳定结构;DoS 攻击则导致模型在整幅图像主体区域产生大范围激活,体现了模型对全局信息缺失和填充值异常的敏感性;而在档位欺骗与转速欺骗攻击中,模型的激活区域集中于规则条带位置,恰好对应伪造报文在特定字段上的重复性注入。上述结果表明,图像化表示能够将不同攻击类型的特征显式化,模型通过聚焦于关键区域完成判别,从而提升了检测结果的可解释性与可靠性。

4　结　　论

针对车载 CAN 总线入侵检测中模型复杂度高、计算资源受限及实时响应要求严苛的问题,本文提出了一种基于改进 SqueezeNet 的轻量化检测方法。该方法通过引入深度可分离卷积与 Ghost 模块降低冗余计算,引入 ECA 通道注意力机制增强特征提取能力,并结合 Hardswish 激活函数提升非线性表达与训练稳定性。实验结果表明,该方法在保证高检测精度的同时显著减少了模型参数量与推理时间,展现出在车载资源受限环境中的应用潜力。本研究主要在传统 CAN 数据集上进行了验证,攻击类型有限,尚未涵盖更复杂的场景。在更高速的 CAN-FD 网络中,由于带宽和负载特性不同,模型的实时性和特征提取

效果仍需进一步考察。未来工作将进一步评估方法在更复杂攻击场景下的适应性，并探索与多种安全机制的融合，例如基于加密与认证的消息完整性保护、远程证明机制以及硬件安全模块支持。同时，将尝试构建能够跨车型、跨网络协议（如 CAN-FD 与车载以太网）的通用 IDS 框架，以增强系统的防御能力和方法的泛化价值。

参考文献

[1] 贾先锋，王鹏程，刘天宇. 基于智能网联汽车车载网络防护技术的研究[J]. 汽车实用技术，2022，47(1)：32-35.
JIA X F，WANG P CH，LIU T Y，Research on the on-board network protection technology of intelligent networked vehicles[J]. Automobile Applied Technology，2022，47(1)：32-35.

[2] 邓淼磊，阚雨培，孙川川，等. 基于深度学习的网络入侵检测系统综述[J]. 计算机应用，2025，45(2)：453-466.
DENG M L，KAN Y P，SUN CH CH，et al. Summary of network intrusion detection systems based on deep learning[J]. Journal of Computer Applications，2025，45(2)：453-466.

[3] 张海春，姜荣帅，王颉，等. 基于熵的车载 CAN 总线异常检测研究[J]. 汽车工程，2021，43(10)：1543-1548.
ZHANG H CH，JIANG R SH，WANG J，et al. Research on anomaly detection of in-vehicle CAN bus based on entropy[J]. Automotive Engineering，2021，43(10)：1543-1548.

[4] 陈博言，沈晴霓，张晓磊，等. 智能网联汽车的车载网络攻防技术研究进展[J]. 软件学报，2025，36(1)：341-370.
CHEN B Y，SHEN Q N，ZHANG X L，et al. Research progress on attacks and defenses technologies forin-vehicle network of intelligent connected vehicle[J]. Journal of Software，2025，36(1)：341 -370.

[5] 戴银飞，周秀贞，刘玉宝，等. 基于 CAN 总线数据的车载网络入侵检测系统[J]. 吉林大学学报（工学版），2025，55(3)：857-865.
DAI Y F，ZHOU X ZH，LIU Y B，et al. In-vehicle network intrusion detection system based on CAN bus data[J]. Journal of Jilin University（Engineering and Technology Edition），2025，55(3)：857-865.

[6] 况博裕，李雨泽，顾芳铭，等. 车联网安全研究综述：威胁、对策与未来展望[J]. 计算机研究与发展，2023，60(10)：2304-2321.
KUANG B Y，LI Y ZH，GU F M，et al. Review of internet of vehicle security research：Threats，countermeasures，andfuture prospects[J]. Journal of Computer Research and Development，2023，60(10)：2304-2321.

[7] 陈滢媛，董振江，董建阔，等. 车联网安全防护技术综述[J]. 电信科学，2023，39(3)：1-15.
CHEN F Y，DONG ZH J，DONG J K，et al. A survey of V2X security protection technologies[J]. Telecommunications Science，2023，39(3)：1-15.

[8] 关宇昕，冀浩杰，崔哲，等. 智能网联汽车车载 CAN 网络入侵检测方法综述[J]. 汽车工程，2023，45(6)：922-935.
GUAN Y X，JI H J，CUI ZH，et al. An overview of intrusion detection methods for in-vehicle CAN network ofintelligent networked vehicles[J]. Automotive Engineering，2023，45(6)：922-935.

[9] 周志豪，陈磊，伍翔，等. 基于 SMOTE-SDSAE-SVM 的车载 CAN 总线入侵检测算法[J]. 计算机科学，2022，49(S1)：562-570，801.
ZHOU ZH H，CHEN L，WU X，et al. SMOTE-SDSAE-SVM-Based vehicle CAN bus intrusion detection algorithm[J]. Computer Science，2022，49(S1)：562-570，801.

[10] KHAN M H，JAVED A R，IQBAI Z，et al. DivaCAN：Detecting in-vehicle intrusion attacks on a controller area network using ensemble learning[J]. Computers & Security，2024，139：103712.

[11] 银鹰，周志洪，姚立红. 基于 LSTM 的 CAN 入侵检测模型研究[J]. 信息网络安全，2022，22(12)：57-66.
YIN Y，ZHOU ZH H，YAO L H. Research on LSTM-based CAN intrusion detection model[J]. Netinfo Security，2022，22(12)：57-66.

[12] 许秀锋，蒲家坤，周爱国，等. 基于门控循环单元的车载控制器局域网络总线入侵检测方法[J]. 科学技术与工程，2021，21(16)：6786-6793.
XU X F，PU J K，ZHOU AI G，et al. Gated recurrent unit-based intrusion detection method for in-vehicle controller area network bus[J]. Science Technology and Engineering，2021，21(16)：6786-6793.

[13] HOSSAIN M D，INOUE H，OCHIAI H，et al. LSTM-based intrusion detection system for in-vehicle CAN bus communications[J]. IEEE Access，2020，8：185489-185502.

[14] TANEIA A，KUMAR G. Attention-cnn-lstm based intrusion detection system（ACL-IDS）for in-vehicle networks[J]. Soft Computing，2024，28(23)：13429-13441.

[15] 陈彦彬，刘桂雄. 双线性自注意力机制 CAN 总线入侵检测方法研究[J]. 电子测量技术，2025，48(2)：122-130.
CHEN Y B，LIU G X. Study on bilinear self-attention mechanism for CAN bus intrusion detection method[J].

Electronic Measurement Technology,2025,48(2):122-130.

[16] 李向荣,张运胜.融合 BERT 与迁移学习的车载 CAN 网络自适应入侵检测[J].汽车技术,2024(12):31-37.
LI X R, ZHANG Y SH. Integrated BERT and transfer learning for adaptive intrusion detection in vehicle CAN network[J]. Automobile Technology, 2024(12):31-37.

[17] LE T D, TRUONG H B Y, KIM D. Multi-classification in-vehicle intrusion detection system using packet-and sequence-level characteristics from time-embedded transformer with autoencoder[J]. Knowledge-Based Systems, 2024, 299: 112091.

[18] BOZDAL M, SAMIE M, JENNIONS I. A survey on can bus protocol: Attacks, challenges, and potential solutions[C]. 2018 International Conference on Computing, Electronics & Communications Engineering(ICCECE), 2018: 201-205.

[19] BOZDAL M, SAMIE M, ASLAM S, et al. Evaluation of can bus security challenges[J]. Sensors, 2020, 20(8): 2364.

[20] SHI D X, KOU L, HUO CH B, et al. A CAN bus security testbed framework for automotive cyber-physical systems[J]. Wireless Communications and Mobile Computing, 2022, 2022(1): 7176194.

[21] JAVED A R, UR REHMAN S, KHAN M U, et al. CANintelliIDS: Detecting in-vehicle intrusion attacks on a controller area network using CNN and attention-based GRU[J]. IEEE Transactions on Network Science and Engineering, 2021, 8(2): 1456-1466.

[22] LAMPE B, MENG W. Intrusion detection in the automotive domain: A comprehensive review[J]. IEEE Communications Surveys & Tutorials, 2023, 25(4): 2356-2426.

[23] MAN D P, ZENG F Y, LYU J G, et al. AI-based intrusion detection for intelligence internet of vehicles[J]. IEEE Consumer Electronics Magazine, 2021, 12(1): 109-116.

[24] 高杨,曹仰杰,段鹏松.神经网络模型轻量化方法综述[J].计算机科学,2024,51(S1):23-33.
GAO Y, CAO Y J, DUAN P S. Lightweighting methods for neural network models: A review[J]. Computer Science,2024,51(S1):23-33.

[25] 牛雅睿,武一,孙昆,等.基于轻量级卷积神经网络的手势识别检测[J].电子测量技术,2022,45(4):91-98.
NIU Y R, WU Y, SUN K, et al. Gesture recognition and detection based on lightweight convolutional neural network[J]. Electronic Measurement Technology, 2022,45(4):91-98.

[26] HAN K, WANG Y H, TIAN Q, et al. Ghostnet: More features from cheap operations[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 1580-1589.

[27] WANG Q L, WU B G, ZHU P F, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 11534-11542.

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

[29] 陈秀真,吴越,李建华.车载信息系统的安全测评体系及方法[J].信息安全学报,2017,2(2):15-23.
CHEN X ZH, WU Y, LI J H. System and approach of security testing and evaluationfor in-vehicle information systems[J]. Journal of Cyber Security, 2017,2(2):15-23.

[30] TARIQ S, LEE S, WOO S S. CANTransfer: Transfer learning based intrusion detection on a controller area network using convolutional LSTM network[C]. 35th Annual ACM Symposium on Applied Computing, 2020: 1048-1055.

[31] DESTA A K, OHIRA S, ARAI I, et al. Rec-CNN: In-vehicle networks intrusion detection using convolutional neural networks trained on recurrence plots[J]. Vehicular Communications, 2022, 35: 100470.

[32] HOANG T N, KIM D. Supervised contrastive ResNet and transfer learning for the in-vehicle intrusion detection system[J]. Expert Systems with Applications, 2024, 238: 122181.

[33] XIA ZH, HUANG L F, TAN J J, et al. TIDL-IDS: A time-series imaging and deep learning-based IDS for connected autonomous vehicles[C]. International Conference on Information Security, 2024: 269-285.

作者简介

樊炳,硕士研究生,主要研究方向为车联网安全。
E-mail:fanbing5082@163.com

曹燚(通信作者),博士,副教授,主要研究方向为计算机视觉、车联网安全。
E-mail:caoyi@cwxu.edu.cn