

基于深度学习的车载CAN网络入侵检测模型

陶 艺^{1,2}, 刘亚丽^{1,2}, 刘景昊^{1,2}

(1. 江苏师范大学 计算机科学与技术学院, 江苏 徐州 221116;

2. 广西密码学与信息安全重点实验室(桂林电子科技大学), 广西 桂林 541004)

摘 要: 深度学习技术因其高效性被广泛应用于车载控制器局域网络入侵攻击检测。车载CAN网络通过引入深度学习技术训练入侵检测模型,以增强车辆的安全防护能力。然而,现有车载CAN网络入侵检测模型存在无法识别特定类型攻击、模型准确率低和模型泛化能力差等问题。为了解决上述问题,提出一种基于深度学习的车载CAN网络入侵检测模型CNN-KAN,实现了准确的入侵攻击检测。CNN-KAN利用Kolmogorov-Arnold定理实现对车载CAN网络报文消息的具体识别,利用迁移学习提高CNN-KAN的泛化能力。实验表明,CNN-KAN在数据集B-CAN和Survival上的准确率分别为94.50%和100%,同时具有较快的收敛速度,适用于车载CAN网络场景下的入侵攻击检测。

关键词: 深度学习;控制器局域网络;入侵检测;CNN;KAN;迁移学习

DOI:10.11907/rjdk.251171

扫描二维码阅读全文:



中图分类号:TP309

文献标识码:A

文章编号:1672-7800(2025)012-0001-07

Deep Learning-Based Intrusion Detection Model for Automobile CAN Networks

TAO Yi^{1,2}, LIU Yali^{1,2}, LIU Jinghao^{1,2}

(1. College of Computer Science and Technology, Jiangsu Normal University, Xuzhou 221116, China;

2. Guangxi Key Laboratory of Cryptography and Information Security (Guilin University of Electronic Technology), Guilin 541004, China)

Abstract: Deep learning technology is widely used in automobile controller area network (CAN) intrusion attack detection due to its high efficiency. Automobile controller area network uses deep learning techniques to train intrusion detection model, which enhances vehicle security protection capability. However, existing automobile CAN network intrusion detection models have many problems such as inability to identify specific types of attacks, low model accuracy, and poor model generalization ability. This paper proposes an automobile CAN network intrusion detection model based on deep learning, named CNN-KAN, which solves the above problems and achieves accurate attack detection. CNN-KAN utilizes the Kolmogorov-Arnold theorem to achieve the specific recognition of automobile CAN network messages, and uses transfer learning to improve the generalization ability of CNN-KAN. Experiments show that CNN-KAN achieves accuracies of 94.50% and 100% on the B-CAN and Survival datasets, respectively, while exhibiting fast convergence, making it well suited for intrusion detection in in-vehicle CAN network scenarios.

Key Words: deep learning; controller area network; intrusion detection; CNN; KAN; transfer learning

0 引言

随着5G技术的普及,车联网已逐渐融入现代交通体

系的各个层面^[1]。智能车辆与网络的高度互联,使得网络通信成为现代交通系统中不可或缺的信息传输方式。在车联网架构中,控制器局域网络(Controller Area Network, CAN)总线协议作为核心通信标准,已在智能车辆通信领

收稿日期:2025-03-13

基金项目: 国家自然科学基金项目(61702237);徐州市科技计划项目(KC22052);广西密码学与信息安全重点实验室(桂林电子科技大学)研究课题(GCIS202114);福建省网络安全与密码技术重点实验室(福建师范大学)开放课题(NSCL-KF2021-04);江苏师范大学研究生科研与实践创新计划项目(2024XKT2632, 2024XKT2589);河南省网络密码技术重点实验室研究课题(LNCT2021-A07);教育部产学研合作协同育人项目(202101374001)

作者简介: 陶艺(1999-),男,CCF会员,江苏师范大学计算机科学与技术学院硕士研究生,研究方向为信息安全、车联网安全、密码学;刘亚丽(1981-),女,博士,江苏师范大学计算机科学与技术学院教授,研究方向为物联网安全、隐私保护和认证技术、应用密码;刘景昊(2000-),男,江苏师范大学计算机科学与技术学院硕士研究生,研究方向为信息安全、区块链与密码学。本文通讯作者:刘亚丽。

域占据了主导地位。随着车辆智能化程度的提升,车载电子控制单元(Electronic Control Unit, ECU)的数量也呈显著增长的态势,导致车载CAN网络系统的功能日趋复杂^[2]。然而,智能化带来的不仅仅是便捷与高效,也威胁到车辆的信息安全。车载CAN网络正面临着越来越多的网络攻击,攻击者通过入侵车载CAN网络,可能造成严重的交通安全事故,直接危及驾驶员和乘客的生命安全^[3]。例如2015年,网络攻击者远程控制车辆运行,导致车辆失控引发交通事故^[4]。近年来,类似的汽车入侵事件层出不穷。据2019–2024年间发生的汽车安全事件统计数据显示,无论是物理攻击还是远程攻击,涉及车辆安全的入侵事件数量均呈逐年增长趋势^[5]。因此,开发可靠的车载CAN网络入侵检测模型具有重要的理论价值,对保障交通安全具有深远的现实意义。

传统入侵检测模型通过预先在数据库中存储已有的攻击类型,当需要检测时再进行模式匹配,从而识别入侵攻击^[6–7]。然而,该方法需要预先收集大量攻击类型数据并存储在数据库中,大大加重了车辆的存储负担。此外,为了确保入侵检测模型的检测能力,需要定期更新数据库,使得车辆不得不频繁收集和分析传输的报文,显著增加了车辆的工作负担。深度学习是一种提取数据特征并依此构造预测模型的技术,在检测攻击方面具有很高的准确率,可以减轻车辆工作负担^[8–12]。然而,基于深度学习的入侵检测模型,其性能在很大程度上依赖于数据集的质量。数据集的多样性与完整性直接决定了模型最终的检测效果。因此,基于深度学习的入侵检测模型,应同时考虑模型检测准确率和泛化能力。

近年来,车载CAN网络的安全问题受到学术界的广泛关注。Song等^[8]提出了一种基于简化卷积神经网络(Convolutional Neural Networks, CNN)的检测方法,通过将CAN ID序列以二进制形式进行处理并用于模型训练。然而,该方法在检测模糊攻击消息时准确率较低。Sun等^[9]设计了一种融合卷积神经网络与长短期记忆网络(Long Short-Term Memory, LSTM)的异常检测框架,以应对复杂的攻击场景。但该框架仅针对二分类问题进行设计,无法进一步识别具体的入侵攻击类型,限制了其在实际应用中的价值。Wu等^[10]提出了一种基于残差块的深度学习架构,结合卷积神经网络和循环神经网络学习时空特征,并在两个通用网络入侵数据集上取得了显著的检测效果。然而,该架构存在过拟合的风险,实验评估未使用车载CAN网络数据集,不适用于车载CAN网络场景下的攻击检测。Hoang等^[11]提出了一种结合监督对比学习和ResNet架构的模型SupCon。SupCon在检测四种攻击类型时,总体误报率降低了80%。然而,SupCon在不同车型CAN总线数据上的泛化能力仍需进一步研究。Aljabri等^[12]设计了一个适用于资源受限的车载CAN网络深度学习模型,通过自适应方法提高了检测系统对新型攻击的响应能力。然而,该模型面

对完全未知或复杂的攻击模式,仍存在检测盲区。

综上所述,现有入侵检测模型存在以下问题:①模型无法识别特定类型攻击;②模型攻击检测准确率较低;③模型泛化能力较差。

针对上述问题,本文提出了一种基于深度学习的车载CAN网络入侵检测模型CNN-KAN。主要贡献如下:

(1)特定类型攻击识别。提出一种可靠多分类深度学习模型CNN-KAN,根据Kolmogorov–Arnold定理,实现模型对车载CAN网络攻击类型的具体识别。

(2)攻击检测准确率高。CNN-KAN在Car-Hacking数据集上进行训练,并将Car-Hacking数据集转换为图像。实验结果表明,CNN-KAN测试准确率可达97.56%,能够有效区分各种入侵攻击模式。

(3)模型泛化能力强。在Car-Hacking数据集训练后的CNN-KAN,通过迁移学习在B-CAN数据集和Survival数据集进行微调。实验结果表明,CNN-KAN测试准确率分别高达94.50%和100%,测试结果优于其他微调后的模型。

1 准备工作

1.1 CAN总线协议

CAN总线协议是车载网络中标准信息传输协议,部署在总线上的节点可以相互发送消息^[13]。通信消息具有固定传输格式,最大传输8字节数据。CAN总线协议以广播形式发送报文,位于总线上的其他节点可同时收到广播报文。图1为CAN标准数据帧格式,数据帧由帧起始、仲裁场、控制场、数据场、CRC场、应答场和帧结束7个字段组成。

帧起始	仲裁场	控制场	数据场	CRC场	应答场	帧结束
1	12	6	0–64	16	2	7

Fig. 1 CAN standard data frame format

图1 CAN标准数据帧格式

其中,帧起始字段长度为1 bit,标志一帧数据的开始,用于所有节点的同步;仲裁场字段长度为12 bit,包含标识符ID和远程传输请求位(Remote Transmission Request, RTR),决定总线上的传输优先级,实现总线访问控制;控制场字段长度为6 bit,包含标识符扩展位(Identifier Extension, IDE)、保留位(r0)和数据长度码(Data Length Code, DLC),为接收端提供数据场的长度信息,以便正确解析数据;数据场字段长度为64 bit,承载传输报文的数据内容;CRC场字段长度为16 bit,包含循环冗余校验码(Cyclic Redundancy Check, CRC)和CRC分隔符(CRC Delimiter),用以提高数据传输的可靠性;应答场字段长度为2 bit,包含应答槽(ACK Slot)和应答分隔符(ACK Delimiter),用以确保帧的接收情况;帧结束字段长度为7 bit,标志帧的结束。

1.2 KAN 网络

Kolmogorov-Arnold 网络 (Kolmogorov-Arnold Network, KAN), 相较于传统神经网络具有更高的准确率^[14]。Kolmogorov-Arnold 表示定理指出, 任何多元连续函数都可以表示为一元函数与加法运算的组合^[15]。基于 Kolmogorov-Arnold 表示定理的 KAN 定义如公式(1)所示:

$$f(x_1, \dots, x_n) = \sum_{q=1}^{2n+1} \Phi_q \left(\sum_{p=1}^n \phi_{q,p}(x_p) \right) \quad (1)$$

其中, $\phi_{q,p}$ 是映射每个输入变量 (x_p) 的单变量函数, 包含 $\phi_{q,p}: [0, 1] \rightarrow R$ 和 $\phi_q: R \rightarrow R$ 。由于所有待学习的函数均为单变量函数, 因此可将每个一维函数的参数化为 B 样条曲线, 并具有局部 B 样条基函数的可学习系数^[16]。KAN 层定义如公式(2)所示:

$$\Phi = \{\phi_{q,p}\}, \quad p = 1, 2, \dots, n_{in}, \quad q = 1, 2, \dots, n_{out} \quad (2)$$

其中, $\phi_{q,p}$ 是可学习参数的参数化函数。

1.3 迁移学习

迁移学习 (Transfer Learning, TL) 是一种支持跨领域迁移知识的机器学习方法^[17]。相较于深度学习模型, 迁移学习是将一个数据集上训练的神经网络模型的权重迁移到另一个数据集的过程。考虑到 CNN 模型底层提取的特征模式适用于其他任务, 而顶层提取的特征则专门针对特定数据集, 因此, 迁移学习被广泛应用于图像处理等领域。为提高迁移学习性能, 在深度学习模型的迁移学习过程中进行微调, 冻结预训练模型的大多数底层, 仅有部分顶层参与微调, 以便在新数据集上重新训练模型。微调将通过更新预训练模型中的部分高阶特征, 使其能更好地拟合目

标数据集。

2 入侵检测模型

2.1 数据集介绍

CNN-KAN 的训练采用韩国黑客与对策研究实验室发布的 3 个流行数据集: Car-Hacking 数据集^[18]、B-CAN 数据集^[19]和 Survival 数据集^[20]。其中, CNN-KAN 使用 Car-Hacking 数据集作为源数据集训练模型, 将 B-CAN 数据集和 Survival 数据集作为迁移学习目标数据集。源数据集和目标数据集的标签空间各自包含不同信息, 其中, 源数据集的标签包括正常、拒绝服务、模糊、转速和齿轮欺骗信息; 目标数据集包括正常、拒绝服务、模糊和故障信息。表 1 列出了 Car-Hacking 数据集、B-CAN 数据集和 Survival 数据集的消息数量。

Table 1 Number of messages in Car Hacking, B-CAN and Survival
表 1 Car Hacking、B-CAN 和 Survival 的消息数量

消息类型	Car-Hacking	B-CAN	Survival
DoS 攻击	458 976	100 000	33 141
Fuzzy 攻击	491 847	100 000	39 812
Gear 攻击	597 252	—	—
RPM 攻击	654 897	—	—
Malfunction 攻击	—	—	7 401
正常消息	928 136	100 000	100 000

2.2 入侵检测流程

入侵检测流程图如图 2 所示, 主要分为数据预处理、构建模型、训练模型和测试模型 4 个部分。

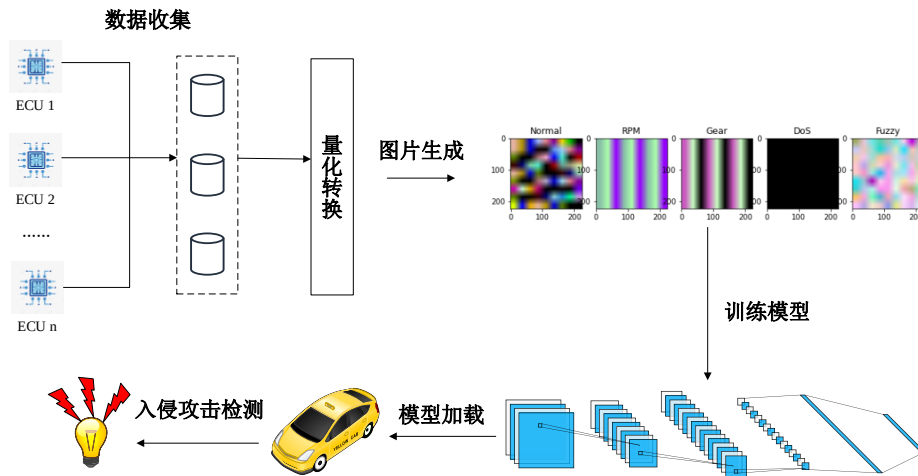


Fig. 2 Intrusion detection flow chart

图 2 入侵检测流程图

2.2.1 数据预处理

CNN-KAN 在获取消息数据后, 需要先进行预处理, 以便为模型生成合适的输入。由于 CNN 模型在图像数据集上的预测效果较好, 而 Car-Hacking 数据集和 B-CAN 数据集为表格型数据, 因此需要将表格数据转换为图像数据形式^[21]。

图像的像素值区间为 0~255, 表格数据也应归一化为 0~255 的尺度。在 CNN-KAN 中使用分位数归一化方法, 其可将数据特征的分布调整为正态分布, 并基于此重新计算所有特征值^[22]。该方法使大多数特征值接近中位数, 从而有效降低了异常值的影响。归一化处理后, 数据样本会依据时间戳和特征规模进行分块转换。

对于Car-Hacking数据集和B-CAN数据集,每条数据拥有9个特征(CAN ID和DATA[0]-DATA[7]),首先将连续的27个样本中的9个特征转换为 $9 \times 9 \times 3$ 的图像,随后将其尺寸统一调整为 $224 \times 224 \times 3$,以适应深度学习模型的输入要求。每个转换后的图像都是一个具有红、绿、蓝通道的彩色图像。

2.2.2 模型构建

CNN-KAN结构图如图3所示。首先,卷积层采用 3×3 卷积核,池化层使用 2×2 窗口;然后,使用ReLU激活函数;最后,KAN层连接所有特征以生成输出的管道。

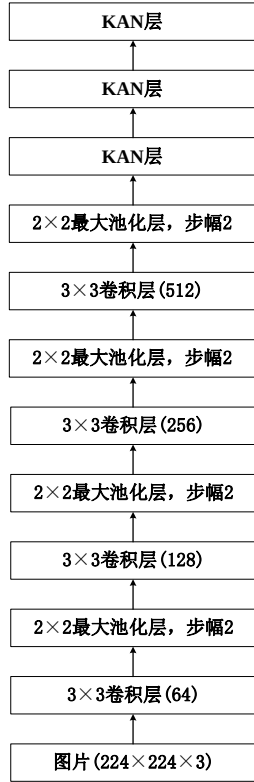


Fig. 3 CNN-KAN structure diagram

图3 CNN-KAN结构图

传统CNN模型通常使用固定形式的激活函数,但在面对复杂数据分布或特征变化时存在表达能力受限的问题。CNN-KAN在卷积层之后引入KAN层,对每个特征通道的激活操作进行基于B样条核函数的可学习激活变换,显著增强了非线性建模能力。KAN层结构通过引入基础激活分支和B样条激活分支来改进特征表示。基础激活分支保留了传统激活函数的线性映射能力,而B样条激活分支则以B样条函数的形式捕捉复杂的输入响应模式。二者融合能够在保证模型稳定性的同时,有效提升特征的代表能力。每个输入通道的激活函数均由多个B样条基函数组成,其系数可学习,并随训练动态调整。KAN结构示意图如图4所示。

每个KAN层包含基础激活项与B样条基函数项的线性组合,具体如公式(3)所示:

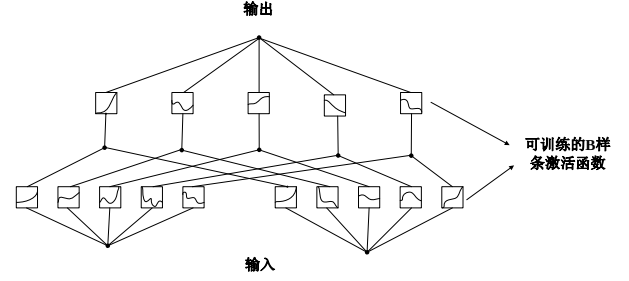


Fig. 4 KAN structure diagram

图4 KAN结构图

$$x^{(l+1)} = \underbrace{W_b^{(l)} \cdot \sigma(x^{(l)})}_{\text{基础激活项}} + \underbrace{W_s^{(l)} \cdot B^{(l)}(x^{(l)})}_{\text{B样条基函数项}} \quad (3)$$

其中, $\sigma(\cdot)$ 为SiLu基函数, W_b 为可学习权重, $B^{(l)}(x)$ 为B样条基张量, W_s 为样条权重。B样条函数通过De Boor算法递推计算,具体如公式(4)~(5)所示:

$$B_{i,0}(x) = \begin{cases} 1 & t_i \leq x < t_{i+1} \\ 0 & \text{其他} \end{cases} \quad (4)$$

$$B_{i,k}(x) = \frac{x - t_i}{t_{i+k} - t_i} B_{i,k-1}(x) + \frac{t_{i+k+1} - x}{t_{i+k+1} - t_{i+1}} B_{i+1,k-1}(x) \quad (5)$$

其中, k 为样条阶数, t_i 为动态调整网格点。KAN根据输入分布自适应调整网格点以提升模型的灵活性。首先,对输入数据按通道排序,提取分位数点 x_{sort} ,用于生成自适应网格 t_{adapt} ;其次,加权融合均匀网格 t_{unif} 与自适应网格 t_{adapt} ,更新B样条函数参数,具体如公式(6)所示:

$$t_{\text{new}} = \epsilon \cdot t_{\text{unif}} + (1 - \epsilon) \cdot t_{\text{adapt}}, \quad \epsilon \in [0, 1] \quad (6)$$

最后,深度KAN定义为层序列 $[n_0, n_1, \dots, n_l]$,其中, n_l 为第 l 层神经元数。KAN的输出定义如公式(7)所示:

$$\text{KAN}(\mathbf{x}) = (\Phi_{L-1} \circ \Phi_{L-2} \circ \dots \circ \Phi_1 \circ \Phi_0)\mathbf{x} \quad (7)$$

每层函数矩阵 Φ_l 由基础激活项与B样条基函数项联合实现。

算法1介绍了CNN-KAN的伪代码。其中,输入图片数据集 $\{X_i, y_i\}$,周期数量为 N ,初始学习率为 η_0 ,提前停止训练的周期数为 P ,输出具有最小验证损失的模型。该算法首先将图像数据转换为张量并进行归一化处理,随后,通过小批量训练方式进行数据加载。在训练过程中,采用AdamW优化器进行参数更新,并使用指数衰减策略调整学习率,以提高模型的稳定性和收敛速度。

在每个训练周期中,算法会计算交叉熵损失,并同时评估模型的验证损失和准确率。如果在设定的轮数内交叉熵损失没有减少,则系统会保存当前最佳模型,并可能触发提前停止机制,从而避免过拟合并提升训练效率。

算法1 CNN-KAN

输入: 图片数据集 $\{X_i, y_i\}$, 其中, $0 \leq i \leq N$, N 是样本总数, $X_i \in \mathbb{R}^{224 \times 224 \times 3}$, $y_i \in [0, C]$, C 为数据集标签类别。 N 是周期数量, η_0 是初始学习率, P 是提前停止训练的周期数。

输出: 具有最少验证损失的模型。

1. 将图像转换为张量,像素值归一化为 $[-1, 1]$
2. 以指定批量大小初始化数据加载器
3. FOR 周期从1到 NDO

4. AdamW 参数更新: $\theta_{i+1} \leftarrow (1 - \lambda)\theta_i - \alpha \mathbf{M}_i \nabla f_i(\theta_i)$
5. 学习率衰减调度器指数衰减学习率: $\eta_i = \eta_0 \cdot \gamma^i$
6. 更新模型交叉熵损失: $L(x, c) = -\log \left(\frac{\exp(x[c])}{\sum_j \exp(x[j])} \right)$
7. 验证模型损失和准确率
8. IF 交叉熵损失在 P 内没有减少 THEN
9. 保存当前最优的模型
10. END
11. END

3 实验和分析

本文提出的模型 CNN-KAN 在配备 Intel(R) Xeon(R) Gold 5418Y * 12 核处理器和 24 GB 显存的 RTX 3090 GPU 的云服务器上训练,并在配备 AMD Ryzen 5 4500U 64 位 CPU 的笔记本电脑上测试,开发环境为 Python3.11,深度学习框架为 PyTorch2.1.0 和 Tensorflow2.16.2。

3.1 评估指标

为了评估经过训练的 CNN-KAN 模型性能,对数据集输入特征进行预测,其预测结果通常归为以下四类^[11]:①真阳性(TP):预测和实际都为攻击型报文;②真阴性(TN):预测和实际都为正常型报文;③假阳性(FP):预测为攻击型报文,实际为正常型报文;④假阴性(FN):预测为正常型报文,实际为攻击型报文。

为应对数据集中攻击与正常报文数量不均衡的问题,本文选用准确率(Accuracy)、精确率(Precision)、召回率(Recall)和 F1 四个指标测试模型性能。具体测试指标如公式(8)所示:

$$\left\{ \begin{aligned} Accuracy &= \frac{TP + TN}{TP + TN + FP + FN} \\ Precision &= \frac{TP}{TP + FP} \\ Recall &= \frac{TP}{TP + FN} \\ F_1 &= 2 \times \frac{Precision \times Recall}{Precision + Recall} \end{aligned} \right. \quad (8)$$

3.2 实验分析

本文所提出的模型 CNN-KAN 在 Car-Hacking 数据集上进行训练,并通过迁移学习在 B-CAN 数据集和 Survival 数据集上进行了微调。在微调过程中,仅训练 KAN 模块的最后一层参数,其余参数均被冻结,以保留原有特征提取结构并专注于分类头的微调。CNN-KAN 实验参数设置如表 2 所示。

本文选取深度学习领域的经典模型 Inception^[23]、ResNet50^[24]、InceptionResNet^[25]、Xception^[26] 和 VGG19^[27] 作为对照,与所提出的 CNN-KAN 进行对比。Car-Hacking 数据集上的实验结果如表 3 所示。

在 Car-Hacking 数据集上,各模型均展现出良好的入

Table 2 Experimental parameter settings
表 2 实验参数设置

实验参数	参数值
初始学习率	0.001
指数衰减率	0.8
最大周期	20
早停周期	3
批量大小	128
B 样条网格大小	5
B 样条阶数	3
B 样条定义域	$[-1, 1]$

Table 3 Experimental results on the Car-Hacking dataset
表 3 Car-Hacking 数据集的实验结果

模型	Accuracy	Prediction	Recall	F1
Xception	0.968 5	0.970 2	0.968 5	0.968 7
VGG19	0.963 0	0.966 5	0.963 0	0.963 4
Inception	0.966 2	0.968 2	0.966 2	0.966 4
ResNet50	0.963 8	0.967 3	0.963 8	0.964 2
InceptionResNet	0.968 9	0.967 9	0.968 9	0.969 1
CNN-KAN	0.975 6	0.977 9	0.975 6	0.977 7

侵检测性能。其中,CNN-KAN 模型在所有评估指标上均取得最佳表现,准确率、精确率、召回率和 F1 分别达到 0.975 6、0.977 9、0.975 6 和 0.977 7,性能优于对比模型。此外,通过迁移学习在 B-CAN 数据集上的实验结果如表 4 所示;在 Survival 数据集上的实验结果如表 5 所示。

Table 4 Experimental results on the B-CAN dataset
表 4 B-CAN 数据集的实验结果

模型	Accuracy	Prediction	Recall	F1
Inception	0.658 1	0.489 9	0.658 1	0.547 6
ResNet50	0.535 7	0.463 8	0.535 7	0.448 6
InceptionResNet	0.894 2	0.909 0	0.894 2	0.888 4
Xception	0.925 3	0.936 3	0.925 3	0.924 5
VGG19	0.666 7	0.500 0	0.666 7	0.555 6
CNN-KAN	0.945 0	0.941 7	0.958 1	0.927 8

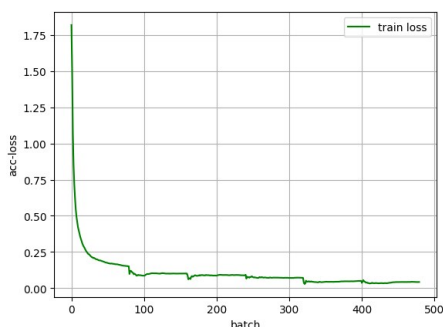
Table 5 Experimental results on the Survival dataset
表 5 Survival 数据集的实验结果

模型	Accuracy	Prediction	Recall	F1
Inception	0.949 8	0.953 5	0.949 8	0.947 6
ResNet50	0.755 1	0.630 3	0.755 1	0.674 1
InceptionResNet	0.986 5	0.987 2	0.986 5	0.985 7
Xception	0.983 7	0.983 6	0.983 7	0.982 9
VGG19	0.945 6	0.901 1	0.945 6	0.921 5
CNN-KAN	1.000 0	1.000 0	1.000 0	1.000 0

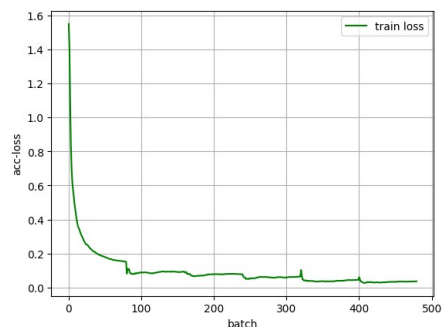
在迁移学习的微调阶段,不同模型在 B-CAN 和 Survival 数据集上的表现存在明显差异。CNN-KAN 模型在 B-CAN 和 Survival 两个数据集上均取得了最优表现。在 B-CAN 数据集中,CNN-KAN 的准确率、精确率、召回率和

F1 值分别达到了 0.945 0、0.941 7、0.958 1 和 0.927 8。而在 Survival 数据集中, CNN-KAN 的准确率、精确率、召回率和 F1 值均达到 1.000 0, 充分展示了其强大的拟合能力和泛化性能。尽管其他对比模型在部分指标上表现出色, 但它们在两个数据集上的整体表现仍未能超越 CNN-KAN 模型。

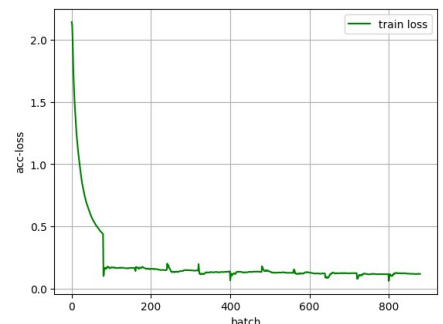
图 5 展示了不同模型的训练损失下降曲线。可以看到, CNN-KAN 模型在训练初期收敛更快, 且曲线较为平稳, 体现出训练过程中的高稳定性。尽管 CNN-KAN 是一个全参数训练模型, 但它在整体训练轮数和结构复杂性上表现优异。此外, KAN 层的可学习激活结构仅作用于网络后端, 计算路径简洁, 在实际部署中具有良好的效率与可行性。



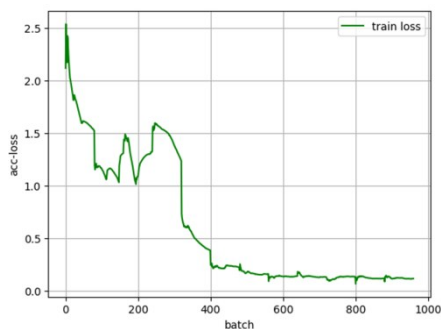
(a) The training loss descent curve of the Inception model
(a) Inception 模型的训练损失下降曲线



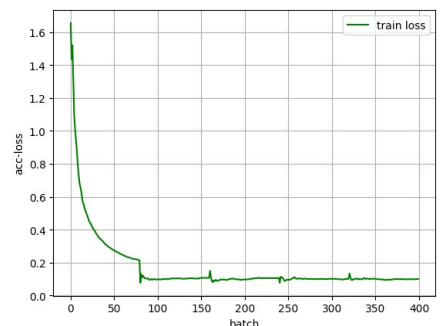
(b) The training loss descent curve of the InceptionResNet model
(b) InceptionResNet 模型的训练损失下降曲线



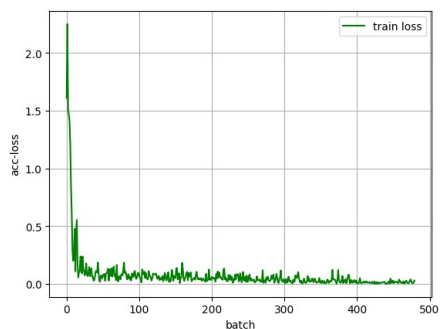
(c) The training loss descent curve of the ResNet50 model
(c) ResNet50 模型的训练损失下降曲线



(d) The training loss descent curve of the VGG19 model
(d) VGG19 模型的训练损失下降曲线



(e) The training loss descent curve of the Xception model
(e) Xception 模型的训练损失下降曲线



(f) The training loss descent curve of the CNN-KAN model
(f) CNN-KAN 模型的训练损失下降曲线

Fig. 5 Convergence comparison of different models

图 5 不同模型收敛性对比

综上所述, 本文提出的 CNN-KAN 模型在多个经典 CNN 架构的基础上引入 KAN 层, 显著增强了模型对复杂特征的表达能力和非线性建模能力, 在两个迁移学习数据集中均展现出优越的分类性能和泛化能力, 验证了其作为深度学习模型的有效性。

4 结语

为了实现准确的入侵攻击检测, 本文提出一种基于深度学习的车载 CAN 网络入侵检测模型 CNN-KAN。针对现有车载 CAN 网络入侵检测模型存在无法识别特定类型攻击、模型准确率低和模型泛化能力差等问题, CNN-KAN

采取了以下具体措施:①利用Kolmogorov-Arnold定理对车载CAN网络中报文消息进行具体识别,提高了入侵检测模型的检测准确率;②利用迁移学习微调CNN-KAN,提高了模型泛化能力。实验结果表明,CNN-KAN在数据集B-CAN和Survival上的准确率分别为94.50%和100%,适用于车载CAN网络场景下的入侵攻击检测。

尽管本文提出的CNN-KAN模型在多个数据集上表现出良好的性能,并验证了其在收敛速度、分类准确率和泛化能力方面的优势,但计算效率仍有待提升。此外,当前模型采用全参数训练,在训练时间上相较于迁移学习模型仍存在劣势。未来可以尝试结合预训练模型,仅将KAN层作为可插拔组件增强特征,以降低训练代价。

参考文献:

- [1] WU C X. Current status and development of automotive network communication bus[J]. *Automobile Applied Technology*, 2019(2): 203-204.
吴晨曦. 汽车网络通信总线现状及发展[J]. *汽车实用技术*, 2019(2): 203-204.
- [2] YU L. Research on the application of the key technology of vehicle network[J]. *Automobile Applied Technology*, 2018(22): 161-163.
喻露. 现阶段车载网络关键技术的应用探究[J]. *汽车实用技术*, 2018(22): 161-163.
- [3] XU H, WU D, LU J W, et al. An intrusion detection algorithm and its hardware acceleration for CAN in vehicles[J]. *Journal of Computer Research and Development*, 2023, 60(12): 2783-2796.
许鹤, 吴迪, 卢继武, 等. 一种车载控制器局域网络入侵检测算法及硬件加速[J]. *计算机研究与发展*, 2023, 60(12): 2783-2796.
- [4] MILLER C, VALASEK C. Remote exploitation of an unaltered passenger vehicle[J]. *Black Hat USA*, 2015 (S 91): 1-91.
- [5] UPSTREAM SECURITY. 2024 Global automotive cybersecurity report[R]. *Global Automotive Cybersecurity Report*, 2024.
- [6] RAJAPAKSHA S, KALUTARAGE H, ALKADRI M O, et al. Ai-based intrusion detection systems for in-vehicle networks: a survey[J]. *ACM Computing Surveys*, 2023, 55(11): 1-40.
- [7] YU Y, ZENG X, XUE X, et al. LSTM-based intrusion detection system for VANETs: a time series classification approach to false message detection[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2022, 23(12): 23906-23918.
- [8] SONG H M, WOO J, KIM H K. In-vehicle network intrusion detection using deep convolutional neural network[J]. *Vehicular Communications*, 2020, 21: 100198.
- [9] SUN H, CHEN M, WENG J, et al. Anomaly detection for in-vehicle network using CNN-LSTM with attention mechanism[J]. *IEEE Transactions on Vehicular Technology*, 2021, 70(10): 10880-10893.
- [10] WU P, GUO H, MOUSTAFA N. Pelican: a deep residual network for network intrusion detection[C]//*International Conference on Dependable Systems and Networks Workshops*, 2020: 55-62.
- [11] 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.
- [12] ALJABRI W, HAMID M A, MOSLI R. Lightweight and adaptive data-driven intrusion detection system for autonomous vehicles[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2025, 26(2): 2282-2292.
- [13] GE L, ZHOU W H, XU H. Research on the application of CAN communication network in automobiles[J]. *Automobile Technology*, 2000(11): 1-4.
葛林, 周文华, 徐航. CAN通信网络在汽车中的应用研究[J]. *汽车技术*, 2000(11): 1-4.
- [14] LIU Z, WANG Y, VAIDYA S, et al. Kan: kolmogorov-arnold networks[DB/OL]. <https://arxiv.org/abs/2404.19756>.
- [15] KOLMOGOROV A N. On the representation of continuous functions of several variables by superpositions of continuous functions of a smaller number of variables[M]. Providence: American Mathematical Society, 1961.
- [16] GORDON W J, RIESENFELD R F. B-spline curves and surfaces[M]. New York: Academic Press, 1974.
- [17] LEONARDO M M, CARVALHO T J, REZENDE E, et al. Deep feature-based classifiers for fruit fly identification (diptera: tephritidae)[C]//*Conference on Graphics, Patterns and Images*, 2018: 41-47.
- [18] ALSAADI F W, AL-ADHAILEH M H. Cyber attack detection for self-driving vehicle networks using deep autoencoder algorithms[J]. *Sensors*, 2023, 23(8): 4086-4112.
- [19] LAMPE B, MENG W. Can-train-and-test: a new CAN intrusion detection dataset[C]//*98th Vehicular Technology Conference*, 2023: 1-7.
- [20] HAN M L, KWAK B I, KIM H K. Anomaly intrusion detection method for vehicular networks based on survival analysis[J]. *Vehicular Communications*, 2018, 14: 52-63.
- [21] WEN L, LI X, GAO L. A new reinforcement learning based learning rate scheduler for convolutional neural network in fault classification[J]. *IEEE Transactions on Industrial Electronics*, 2020, 68(12): 12890-12900.
- [22] HUSSAIN F, ABBAS S G, HUSNAIN M, et al. IoT DoS and DDoS attack detection using ResNet[C]//*International Multitopic Conference*, 2020: 1-6.
- [23] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision[C]//*Conference on Computer Vision and Pattern Recognition*, 2016: 2818-2826.
- [24] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//*Conference on Computer Vision and Pattern Recognition*, 2016: 770-778.
- [25] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception-v4, inception-resnet and the impact of residual connections on learning[C]//*Conference on Artificial Intelligence*, 2017: 4278-4284.
- [26] CHOLLET F. Xception: deep learning with depthwise separable convolutions[C]//*Conference on Computer Vision and Pattern Recognition*, 2017: 1251-1258.
- [27] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks[C]//*Conference on Computer Vision and Pattern Recognition*, 2016: 1646-1654.

(责任编辑:陈维捷)