

DOI: 10.3901/JME.2021.24.268

基于深度学习的无人驾驶汽车导航传感器 异常诊断方法*

官文峰^{1,2} 王元哲² 陈辉¹ WANG Danwei²

(1. 武汉理工大学高性能舰船技术教育部重点实验室 武汉 430063;

2. 南洋理工大学电机与电子工程学院 新加坡 639798 新加坡)

摘要: 近年来,无人驾驶汽车(Unmanned autonomous vehicle, UAV)作为未来智能交通系统(Intelligent transportation system, ITS)的重点发展方向已成为业内研究的热点。作为一种新兴智能交通工具,无人汽车的行驶完全依赖于导航传感器提供的精确位置和路径数据。GPS传感器因计算机黑客的恶意网络攻击或物理故障而造成导航位置数据异常或篡改给无人车的安全行驶带来了巨大的威胁和挑战。针对无人汽车GPS传感器易受攻击的现实问题,提出一种基于深度学习的无人车GPS传感器异常检测新方法。该方法通过改进传统一维卷积神经网络(1D Convolutional neural network, 1D-CNN)的拓扑结构,设计出1DGAP-CNN算法框架,使其满足快速实时性的诊断要求。首先,原始的多传感器位置数据直接输入到提出的算法中进行数据融合和预处理;其次,提出的算法自动的进行特征提取、降维减参和模式辨识;最后,模型直接输出诊断结果,整个诊断过程由模型自主完成。结果表明,提出的方法相比于主流的智能诊断算法具有更高的诊断准确率和更快的检测速度。

关键词: 无人驾驶汽车;深度学习;异常诊断;导航传感器;网络攻击

中图分类号: TH165; TP277; TP306

Anomaly Diagnosis for Navigation Sensors of Unmanned Autonomous Vehicles Based on Deep Learning

GONG Wenfeng^{1,2} WANG Yuanzhe² CHEN Hui¹ WANG Danwei²

(1. Key Laboratory of High Performance Ship Technology of Ministry of Education in China,

Wuhan University of Technology, Wuhan 430063;

2. School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, Singapore)

Abstract: As a key technology of future intelligent transportation systems(ITS), unmanned autonomous vehicles(UAVs) have become a research hotspot in recent years. As an emerging intelligent transportation tool, unmanned vehicles rely on the precise location observations provided by navigation sensors. Once compromised, navigation sensors may generate abnormal observations deviating away from the ground truth, which as a result will cause severe consequences or even fatal accidents. To enhance the security of UAVs, a new deep learning based anomaly diagnosis method is proposed in this paper for the detection and identification of sensor anomalies in UAVs. The proposed method improves the topological structure of the traditional 1D Convolutional neural network (1D-CNN) and designs a 1DGAP-CNN algorithm framework to achieve a real-time rapid diagnosis. First, the original pose measurements from multiple sensors are directly input into the proposed algorithm for data fusion and preprocessing. Secondly, the proposed algorithm automatically performs feature extraction, dimension transformation, parameter reduction, and anomaly identification. Finally, the diagnosis results are automatically generated. Evaluation results show that the proposed method has higher

* 国家重点研发计划(2019YFE0104600)、工信部“绿色智能内河船舶创新专项”、国家自然科学基金(51579200, U1709215)、广西自然科学基金(2020GXNSFBA159058)、中央高校基本科研业务费专项武汉理工大学优秀博士学位论文培育(2019-YB-023)和中国国家留学基金委博士联合培养(CSC201906950020)资助项目。20210521 收到初稿, 20211025 收到修改稿

diagnostic accuracy and faster detection speed than the state-of-the-art intelligent diagnostic algorithms.

Key words: unmanned autonomous vehicles; deep learning; anomaly diagnosis; navigation sensor; cyber attacks

0 前言

近十年来,随着“互联网+”、人工智能、“5G”通信、工业大数据、物联网和智能车联网等高新技术的快速发展和普及应用,在全球范围内的交通运输系统中正发生着由“交通 3.0”向“智能交通 4.0”转变的巨大变革^[1-2]。无人驾驶汽车(Unmanned autonomous vehicle, UAV)作为一种新兴交通工具已成为智能交通运输系统(Intelligent transportation system, ITS)的重要组成部分^[3-5]。UAV 集成了诸如计算机视觉、智能控制、语音识别、目标检测和人机交互等一系列先进技术,可以自动完成诸如自动驾驶、自主蔽障、自主路径规划和自动泊车等一系列“人工操作”,不仅有效释放了驾驶员的操控时间,同时还有效避免和减少了诸如疲劳驾驶、酒驾、交通拥堵和违章等交通事故的发生^[6-7]。然而,UAV 作为一种高新技术产物在给人们带来了极大便利的同时,其安全稳定性也面临着严峻挑战^[8-9]。

自 UAV 问世以来,国内外无人汽车开发商诸如谷歌、特斯拉、百度和 Uber 等公司在路测阶段时频频遭遇各种事故发生^[10-11]。如 2015 年 7 月,一辆吉普汽车被黑客入侵仪表盘电脑,导致车辆失控并撞入深沟^[10]。2016 年 9 月,网络黑客从 12 英里以外的位置远程入侵特斯拉 Model S 无人驾驶汽车,干扰无人车的刹车、门锁和仪表盘等电子系统,给乘客和车辆的安全带来极大的危险^[11]。分析其原因主要有两个方面:其一是硬件和软件技术问题,如导航传感器损坏、设备老化和信号延迟等,这类问题随着科技水平的发展已逐渐得到解决和完善^[9];另一方面则是来自于恐怖分子或黑客的恶意网络攻击,干扰正常传感器或系统的真实数据,使无人车执行错误指令,从而造成令人防不胜防的危害,目前已成为影响无人汽车大规模上市前必须解决的首要问题^[8]。

通常,无人车的网络攻击主要分为内部攻击和外部攻击两个方面^[9],外部攻击主要入侵无人车以外的设备,例如对路基测量单元、交通信号、道路标识、通信设备和其他车辆数据等基础设施的攻击和干扰;内部攻击主要是入侵无人车本体的系统和传感器,典型的包括车载控制系统、导航传感器、车载诊断系统以及其他车载传感器等^[12]。在众多的

攻击类型中,导航传感器的攻击是最直接、最常见和最具危害性的攻击类型之一,它可以在人们毫无明显察觉的情况下突然或缓慢的改变汽车的行驶车道、行驶速度和方向,该攻击极具隐蔽性,给人们的心理产生巨大的恐惧感,是当前业内迫于解决的关键技术难题之一^[9]。

目前,业内学者和科研院所主要聚焦于无人车功能的实现和性能提升的研究,直接针对无人车导航传感器网络攻击检测的研究还较少,比较相关的有 LIU 等^[8]提出的一种基于扩展卡尔曼滤波器(Kalman filter, KF)模型的无人车网络攻击安全姿态评估方法。MO 等^[13]建立了一种基于 KF 残差的 χ^2 检测器用于传感器攻击检测。WANG 等^[9]提出了一种基于支持向量机与 KF 相结合的无人车传感器异常检测方法。LI 等^[14]提出了一种基于贝叶斯博弈的网络攻击防御方法用于列车控制系统的攻击检测。RASHEED 等^[3]提出了一种基于深度强化学习的方法用于自动车辆的系统安全维护。ULLAH 等^[15]提出了一种基于深度学习的方法用于物联网的网络安全预测。以上方法虽然在不同的攻击检测问题上得到应用,然而仍存在一些不足:首先,基于模型的异常检测方法需要建立复杂的数学模型,并且检测效果很大程度上取决于模型精度,然而在实际中建立高精度的数学模型是十分困难的^[16]。其次,现行基于机器学习的方法仍然需要借助 KF 对原始数据计算残差后再输入到 SVM 等机器学习模型中进行异常检测,此类方法的诊断性能容易受到 KF 的残差精度的影响。目前基于深度强化学习的网络攻击防御策略研究仍处于初步探索阶段,其攻击检测效果仍需要进一步提升。

针对以上问题的不足,本文将人工智能最前沿的深度学习技术引入到无人车导航传感器异常检测领域,提出了一种基于一维卷积神经网络(1D Convolutional neural network, 1D-CNN)的无人车导航传感器异常智能诊断方法和系统框架。该模型由输入层、1D 卷积层、1D 池化层、1D 全局均值池化层(1D Global average pooling, 1D-GAP)和 Softmax 分类器组成。本文的主要贡献包含以下三个方面:首先,给出了一种直接采用无人车多传感器原始位置数据进行网络攻击检测的新方案;其次,将深度学习技术引入无人车网络攻击检测领域,通过其强大的特征提取能力解决缓变微小攻击难以被检测的

难题;第三,改进了现有一维卷积神经网络算法,设计了 1D-GAP-CNN 算法结构,使其更加符合无人车网络攻击检测的快速实时性要求。结果表明,提出的方法可以有效诊断无人车 GPS 传感器在不同的网络攻击下的异常状况,相比于主流的智能诊断算法具有更高的诊断准确率和更快的异常检测速度。

1 问题描述

无人汽车在行驶过程中需要依赖大量的传感器实时的读取和接收来自于路基单元、交通信号、雷达基站和其他车辆的路况数据^[8],为了实现高可靠性的不间断导航,通常在车体上同时安装多种传感器以获得足够冗余量的位置信息^[12]。图 1 所示为作者团队开发的无人车试验平台,安装有全球定位系统(Global position system, GPS)、激光探测与测量传感器(Light detection and ranging, LIDAR)、摄像头、无线电探测与测距传统器(Radio detection and ranging, RADAR)和超声传感器等。

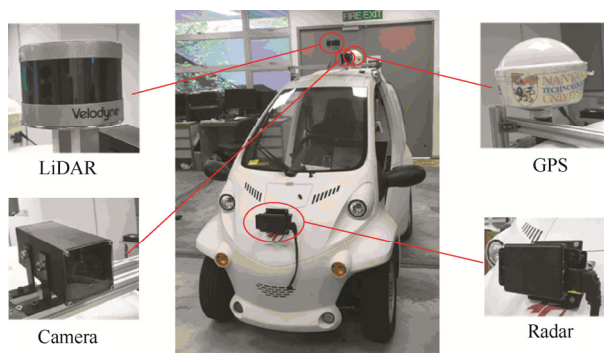


图 1 无人车的传感器布置

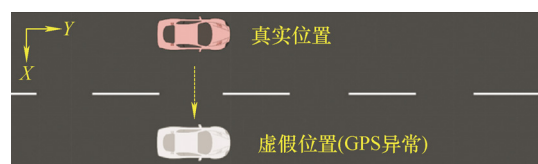
在上述传感器中, GPS 是最重要的导航单元之一,可以实时的引导车辆行驶,是目前网络黑客重点攻击的对象^[8]。一旦 GPS 发生异常,无人车中央控制系统会立即中止使用 GPS 数据,并启动应急措施,否则错误的位置数据将会导致无人车偏离正确的行驶轨迹,从而给车辆和乘客带来巨大的危险。据调查研究表明, GPS 因物理失效或恶意网络攻击而导致的异常主要有以下五种类型^[8-9]。

(1) 定点异常:即无人车实际上处于移动状态时, GPS 却始终显示一个固定的位置坐标,此攻击效果也类似于 GPS 物理损坏。

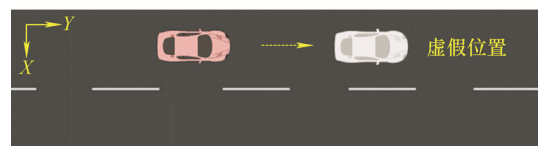
(2) X 向偏移异常:即 GPS 测量值与无人车真实位置在左右 X 方向偏移一定距离,如图 2a 所示。

(3) Y 向偏移异常:即 GPS 测量值与无人车真实位置在前后 Y 方向偏移一定距离,如图 2b 所示。

(4) X 向缓变异常:即 GPS 测量值相对无人车真实位置在左右 X 方向以一定倍率缓慢漂移,该攻击由于残差增量较小且达不到检测阈值,极具隐蔽性。



(a) X 向左右偏移攻击



(b) Y 向前后偏移攻击

图 2 无人车 GPS 网络攻击示意图

(5) Y 向缓变异常:即 GPS 测量值相对无人车真实位置在前后 Y 方向以一定倍率缓慢偏移,该攻击也极具隐蔽性。

以上五种异常中, X 向缓变和 Y 向缓变是最难辨识的攻击类型,由于两者发生时, GPS 的虚假位置量是逐渐增加的,此时的异常特征极不明显,很难被有效检测,也是黑客重点设计的攻击类型之一。在本文中,着重于针对以上五种因人为攻击或物理损坏而引起的 GPS 异常状态进行诊断和分析。

2 诊断方法

2.1 卷积神经网络

CNN 是当前人工智能领域最具代表性的深度学习算法之一^[16],它采用深层网络结构模拟动物视觉系统的工作机理,可以直接从原始数据中自动的提取关键特征,具备强大的特征提取能力^[17-18]。CNN 主要采用卷积层和池化层两种操作算子进行特征处理^[19],首先,卷积层通过构建多个滤波器(卷积核)用于提取特征,类似于视觉系统的感受野,每个滤波器一次仅关注一个局部区域的数据^[18];其次,由池化层再对卷积层生成的特征图进行尺寸压缩和强化特征差异^[19]。如图 3 所示, CNN 通过设计多层堆叠的卷积层、激活层和池化层对目标逐层提取特征^[20],最终提取出隐藏于数据中的关键代表特征。根据卷积核的操作类型, CNN 通常分为一维卷积网络(1D-CNN)和二维卷积网络(2D-CNN), 1D-CNN 适用于处理一维时间序列数据,主要用于语音识别任务^[18], 2D-CNN 擅长处理二维图像或三维视频流数据,主要用于计算机视觉领域^[21]。

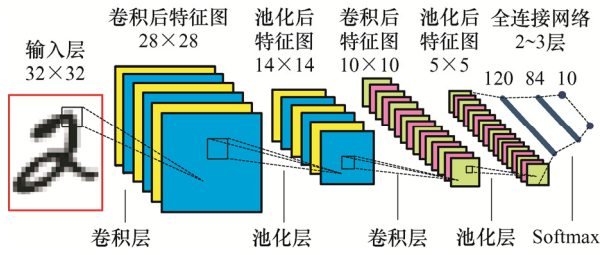


图3 卷积神经网络基本结构

不同于计算机视觉和语音识别领域中的模式分类任务，无人车网络攻击异常诊断任务既要求诊断模型的准确率高，又要求实时检测速度快。在本文中，由于无人驾驶汽车的GPS和LIDAR导航传感器采集的数据均为多通道一维时间序列数据，因此，1D-CNN更加适合处理无人车导航传感器异常诊断问题。然而，经大量研究发现，传统的1D-CNN模型中通常使用了一个Flatten层和2~3层的全连接层(Fully connection, FC)部分作为输出层，而FC层带入了过多的训练参数量，约占CNN总参数量的80%~90%^[22]，巨大的模型参数量导致模型训练时间和诊断等待时间较长，不利于无人车网络攻击的快速实时性检测。因此，主流的基于语音识别的1D-CNN框架尚不能直接套用在本研究问题上。

2.2 提出的1DGAP-CNN诊断算法

针对无人车网络攻击诊断的现实问题，提出了一种1DGAP-CNN的深度学习新算法用于GPS导航传感器的异常智能诊断，其算法结构如图4所示，主要由输入数据融合层、一维卷积网络层、降维减参层和输出层四个部分组成。首先，原始的无人车GPS和LIDAR导航传感器多通道测量数据被直接输入到输入层自动完成数据融合和数据处理。其次，一维卷积网络层内的多个1D卷积层、Relu激活层和1D池化层依次对原始数据进行特征提取。然后，1D-GAP层自动的完成降维和参数量压缩，并且自适应的关联一维卷积网络层与输出层之间的维度变换；最后，诊断结果由Softmax分类器输出。

提出的诊断方法主要基于以下两点考虑而设计：首先，考虑到无人车的GPS和LIDAR导航传感器采集的数据均为多通道一维时间序列数据，因此，本文引入了类似于语音识别任务的一维卷积神经网络算法^[23]用于处理无人车多传感器异常检测问题。其次，考虑到无人车传感器异常诊断的实时性要求，提出的方法对现行的1D-CNN算法进行了改进，通过设计了一个降维减参层，采用一维全局均值池化技术代替传统1D-CNN中的Flatten层和2~3层的全连接层部分，有效减少模型参数量达80%~90%，从而显著的减少算法训练时间和诊断

等待时间，提升批量样本的诊断速度，达到实时性要求。

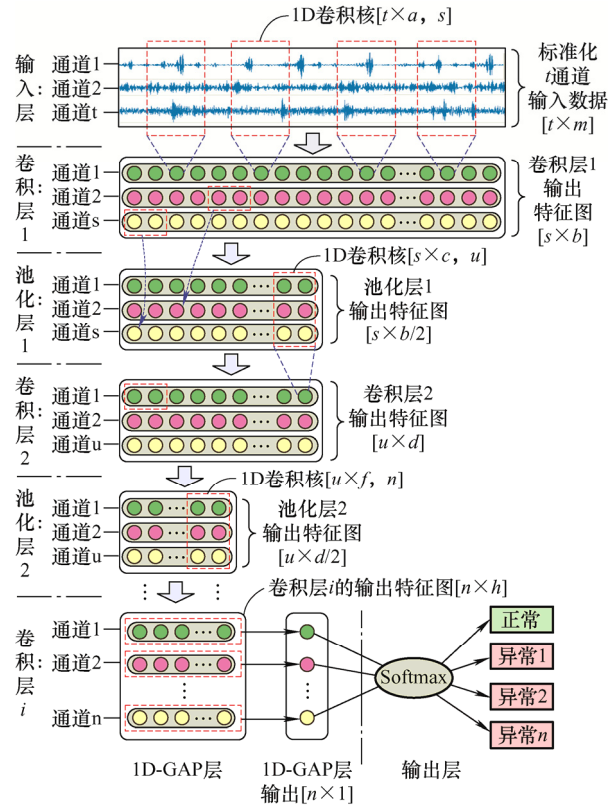


图4 提出的1DGAP-CNN深度学习算法结构图

图5为提出的无人车导航传感器异常诊断系统框架图，主要包含三个模块：底层为无人车导航传感器的数据采集模块，顶层为诊断结果实时输出和可视化模块，中间层为提出的1DGAP-CNN智能诊断

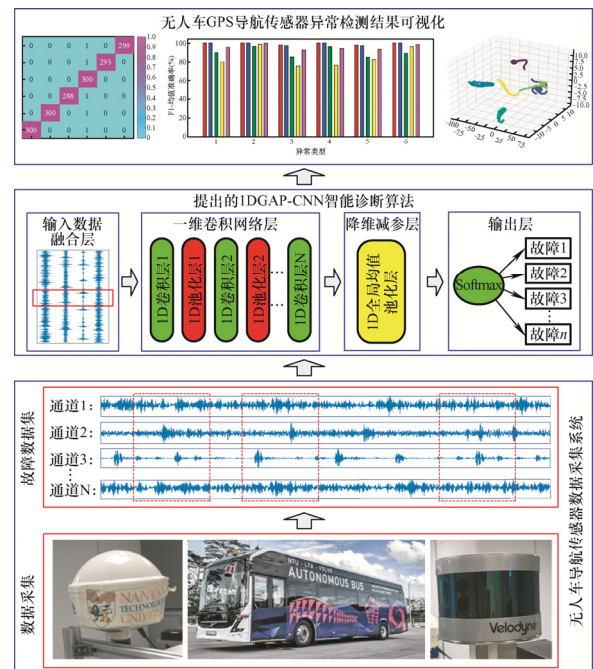


图5 提出的无人车导航传感器网络攻击异常诊断框架图

断算法模块。该方法无需采用任何的数据处理方法对原始数据进行人工处理, 1DGAP-CNN 算法可自动对 GPS 和 LIDAR 的传感器数据进行多通道数据融合、特征提取和异常诊断结果输出, 端到端的算法结构具有优越的自适应性和灵活通用性。

在提出的算法中, 训练过程主要包含前向传播和误差反向传播两个过程。

2.2.1 前向传播计算

2.2.1.1 输入数据融合层

输入层用于接收无人车导航传感器的原始测量数据, 并对这些数据进行数据融合、数据截断和标准化处理等, 将原始数据处理成 1DGAP-CNN 模型可训练的样本格式。

(1) 多传感器数据融合。在输入层中, 首先对无人车的 GPS 和 LIDAR 等多个传感器的测量数据进行数据融合。无人车上安装的 GPS 输出定位信息, LIDAR 经 SLAM 可得到定位数据, 假设有 n 种异常类型, 每种异常同时具有 k 个监测传感器, 每个传感器具有 m 个数据通道, 每个通道采集了 l 个数据点, 从而可以构建一个 $[n, [k, m], l]$ 的多维矩阵原始数据集, 如图 6 所示。

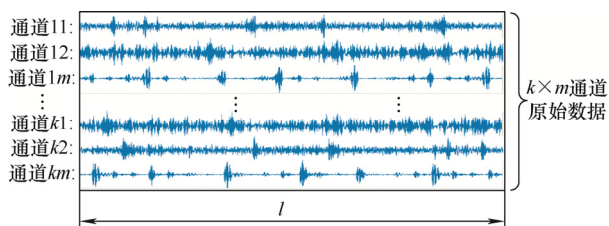


图 6 无人车多导航传感器数据融合

(2) 数据截断和样本生成。在图 6 中的每个传感器通道获得的数据均是一个长时间序列数据, 该数据若一次性输入 CNN 模型将会导致内存溢出而无法训练^[24]。为了生成诊断模型可训练的样本, 本文根据 GPS 和 LIDAR 的采样频率对原始的长时间序列数据进行数据截断^[25], 如图 7 所示。在图 7 中, 假设采样周期是 400 个点, 原始样本长度为 2000 个数据点, 则等分截断后可获得 5 个长度为 400 个点的新训练样本。

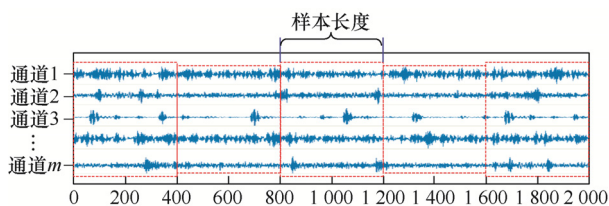


图 7 数据等分截断的示意图

2.2.1.2 一维卷积网络层

一维卷积网络层主要由多个 1D 卷积层、Relu 激活层和 1D 池化层依次堆叠^[23], 组成深层网络结构对原始训练样本数据进行特征提取。

(1) 一维卷积层。一维卷积层包含有多个一维卷积核, 每一个卷积核为一个权重矩阵, 1D 卷积核按照一定的步长滑动遍历整个输入特征图, 从而提取原始数据中的特征要素^[18], 不同的卷积核提取不同的特征, 通过设置多个卷积核即可提取多个不同的特征^[24], 一维卷积的数学表达为^[20]

$$C_{l' \times w'}^{(k)} = W_{l \times w}^{(k)} \otimes X_{l \times w}^{(k)} + B^{(k)} \quad (1)$$

式中, \otimes 表示卷积运算; k 表示第 k 组一维卷积核; $W_{l \times w}^{(k)}$ 为第 k 组一维卷积核的权重; l 为卷积核的长度, w 为卷积核的宽度(训练样本通道数); $X_{l \times w}^{(k)}$ 表示输入数据; v 代表每个训练样本的长度; $B^{(k)}$ 为第 k 组卷积核对应的偏置; $C_{l' \times w'}^{(k)}$ 为卷积操作后的输出; l' 为输出特征图的长度; w' 输出特征图的宽度。

(2) 激活层。CNN 中常用的激活函数有修正线性单元(Rectified linear, Relu)、Sigmoid 和 Tanh 函数^[16]。考虑到收敛速度和改善梯度消失问题, 本文选用 Relu 作为激活函数, 其数学表达为^[24]

$$A_{l' \times w'}^{(k)} = f(C_{l' \times w'}^{(k)}) = \max\{0, C_{l' \times w'}^{(k)}\} \quad (2)$$

(3) 一维池化层。池化的作用是在保持原有主要特征不丢失的前提下尽量压缩图像尺寸和精减数据维度, 同时增加不同特征之间的差别, 具有特征增强的功能^[19]。在 CNN 中, 池化层一般设置在卷积层之后, 常用的池化操作有最大池化和均值池化两种^[16]。在 CNN 中应用较多的是最大池化, 其表达式为

$$P_{l'' \times w'}^{(k)} = \max_{(j-1)S \leq i \leq jS} \{A_{l' \times w'}^{(k)}(t_{z, w'})\} \quad (3)$$

式中, z 代表一维池化核的长度; w' 代表一维池化核的宽度, 该值始终与被池化操作的特征图的宽度相同; $t_{z, w'}$ 为池化核内对应的元素值; S 为池化核的步长, 一般为 2; $P_{l'' \times w'}^{(k)}$ 为经过一维池化操作后的输出特征图; l'' 代表经一维池化操作后输出特征图的长度。

2.2.1.3 一维全局均值池化层

1D-GAP 层是一种具有维度自适应的新池化结构, 用于替代传统 1D-CNN 算法中的 Flatten 层和 2~3 层的全连接网络层, 同时自适应的匹配前一层输出与后一层输入之间的维度变换关系。1D-GAP 与普通 1D 池化相类似, 其独特之处在于 1D-GAP 的

池化核的尺寸与被池化的特征图纵向尺寸完全相同,如图8所示。1D-GAP操作的数学表达可写为

$$GAP_{1D}^s = \frac{1}{h} \sum_{i=1}^h X_i^s \quad (4)$$

式中, GAP_{1D}^s 表示经全局均值池化处理后的输出结果; s 表示第 s 个通道; X_i^s 表示第 s 个通道内第 i 个特征元素值。在图8中, 假设有 n 种分类状态, 即 Softmax 分类器的输出维度为 $[n \times 1]$, 一维卷积网络层的最后一组 1D 卷积层的输出特征图尺寸为 $[n \times h]$, h 代表特征图长度, n 代表宽度(n 个状态类型), 则 1D-GAP 层通过构建 n 个 $[1 \times h]$ 的 1D-GAP 池化核对卷积层输出 $[n \times h]$ 执行一维全局均值池化操作, 每一个池化核内的所有数值求均值后均可得到一个特征值, 从而 1D-GAP 层的输出为 $[n \times 1]$, 该值然后输入到 Softmax 分类器进行故障分类。在本算法中, 1D-GAP 层可减少 80%~90% 的模型训练参数量, 从而可有效提升算法的诊断速度和训练效率。

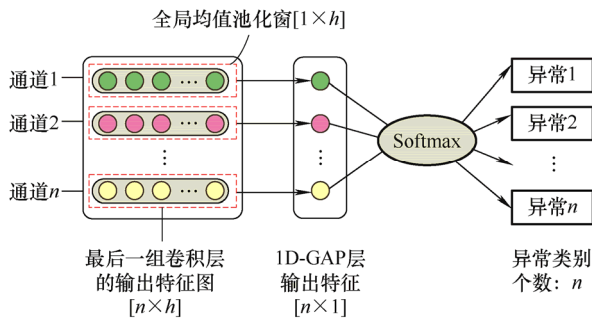


图8 一维全局均值池化变换操作示意图

2.2.1.4 Softmax 输出层

经 1D-GAP 层计算输出的结果仍然是一个量值不统一且不符合概率分布的 n 维数组 $[X]_{1 \times n}$, 这将无法统一度量当前样本所属故障类别的概率大小, 并且不利于模型参数训练和误差反向传播^[19]。为此, 在 1D-GAP 层之后设置了一个 Softmax 分类器, 对 1D-GAP 的输出值进行归一化操作^[24], 使 1D-GAP 的输出符合概率分布, 其数学表达式为^[20]

$$Y_i = \text{Softmax}(X_i) = \frac{\exp(X_i^{(1)}), \exp(X_i^{(2)}), \dots, \exp(X_i^{(n)})}{\sum_{l=1}^n \exp(X_i^{(l)})} \quad (5)$$

式中, l 表示 X_i 的第 l 个元素索引号; $e^{X_i^l}$ 表示把 X_i 的第 l 个元素值转变为 0 至 1 之间; Y_i 为 X_i 经过 Softmax 函数归一化后的输出结果。

2.2.2 误差反向传播阶段

未经训练的模型通过前向传播计算得到的结果

与真实标签具有较大的差距, 为了使模型预测值无限接近样本标签, 模型需要通过不断的训练和寻优。提出的算法采用梯度下降误差反向传播算法(Back propagation, BP)进行模型训练和参数权重更新, 并且采用交叉熵损失函数计算 Softmax 层输出结果与真实标签之间的误差值, 其数学表达式^[23]为

$$J(w) = -\frac{1}{m} \left[\sum_{i=1}^m \sum_{j=1}^n I\{\bar{y}_i = j\} \lg \frac{\exp(x_i^T \cdot w_j)}{\sum_{l=1}^n \exp(x_i^T \cdot w_l)} \right] \quad (6)$$

式中, i 表示第 i 个训练样本, j 表示属于第 j 个类别(共有 n 个类别); $I\{\cdot\}$ 为逻辑指示函数, 当大括号内的值为真时, $I=1$, 否则 $I=0$; $\bar{y}_{(i)}$ 表示第 i 个样本的真实标签; $J(w)$ 为交叉熵损失函数。对式(6)求一阶偏导数, 即可以逐层更新 CNN 的权重参数 w 和偏置 b ^[24]。

$$w' = w - \eta \frac{\partial J}{\partial w} \quad (7)$$

$$b' = b - \eta \frac{\partial J}{\partial b} \quad (8)$$

式中, w' 和 b' 分别为更新后的权重和偏置; w 和 b 为当前的权重和偏置; $\eta \in (0, 1)$ 为学习率。训练的过程即是不断的调整 w 和 b , 以最小化 $J(w)$ 。

3 算法验证与评估

3.1 异常数据集

为验证提出方法的有效性和可行性, 本文对第1节所述的无人车 GPS 传感器常见的 5 种异常或网络攻击进行快速诊断。此处选用了 GPS 和 LIDAR 两个导航传感器的位置数据作为验证数据, 两者均包含有 x 、 y 和姿态角 θ 三个通道的数据, 其中 GPS 的数据加入了 5 种攻击, LIDAR 作为参考信号未加入攻击。两个传感器的采样频率均为 10 Hz, 测试时间为 1 000 s, 因此每类异常的每个通道均可获得一个包含有 10 000 个点的一维时间序列数据段。考虑到不同异常严重程度和网络攻击的不确定性, 为获取相对完备的异常状态训练数据集, 本文设置了多种异常等级, 详细说明如下。

(1) 正常状态和定点攻击: 无等级。

(2) X 向偏移攻击: 分别设置了 ± 1 m、 ± 3 m、 ± 5 m、 ± 7 m 和 ± 10 m 五个等级。

(3) Y 向偏移攻击: 分别设置了 ± 5 m、 ± 10 m、 ± 20 m、 ± 30 m 和 ± 50 m 五个等级。

(4) X 向缓变漂移攻击: 分别设置了 ± 0.2 m、 ± 0.4 m、 ± 0.6 m、 ± 0.8 m 和 ± 1.0 m 五种斜率等级。

(5) Y 向缓变漂移攻击: 分别设置了 ± 0.2 m、 ± 0.4 m、 ± 0.6 m、 ± 0.8 m 和 ± 1.0 m 五个斜率的攻击等级。以上每个异常等级的采样时间均为 100 s。

按照文章第 2.2.1 节所述的方法, 每种异常状态的原始数据在输入层中首先进行数据融合, 两个传感器获得 6 个通道数据, 从而构建一个 $[10\ 000, 6]$ 的二维特征图, 在 5 种异常状态和 1 个正常状态下最终可得到一个 $[6, 10\ 000, 6]$ 的原始数据集, 第一个 6 代表 6 种状态类型。其次, 对所有原始数据做标准化^[19]处理后进行样本切割, 此处无人车的 GPS 和 LIDAR 的采样频率均为 10 Hz, 因此, 在数据截断中将每个训练样本的长度设置为 10, 将原始数据进行等分截断, 从而每种状态类型可获得 1 000 个训练样本, 如表 1 所示。最后, 在每类异常的所有样本中随机取 70% 作为训练集, 30% 作为测试集, 在训练集中随机的选取 20% 进行交叉验证, 训练集用于模型训练, 测试集用于检测算法的诊断准确率, 最终建立的异常数据集如表 1 所示。

表 1 无人车 GPS 传感器异常数据集

标签	异常类型	原始数据	样本长度	总样本	训练集	测试集
1	正常状态	$[10\ 000, 6]$	$[10, 6]$	1 000	700	300
2	定点攻击	$[10\ 000, 6]$	$[10, 6]$	1 000	700	300
3	X -向偏移	$[10\ 000, 6]$	$[10, 6]$	1 000	700	300
4	Y -向偏移	$[10\ 000, 6]$	$[10, 6]$	1 000	700	300
5	X -向缓变	$[10\ 000, 6]$	$[10, 6]$	1 000	700	300
6	Y -向缓变	$[10\ 000, 6]$	$[10, 6]$	1 000	700	300

3.2 模型超参数选择与训练过程

根据图 5 所示的诊断算法框架, 本文对网络层数、卷积核大小、卷积核数量、激活函数、池化核大小和池化层数量等进行反复调参, 最终选取了如表 2 所示的模型超参数。该模型共包含 12 个网络层, 其中有 4 个 1D 卷积层、1 个 1D 池化层、3 个 Relu 激活层、1 个 1D-GAP 层和 Softmax 输出层, 所有的卷积核和池化核的补零方式都设置为 Padding="Same", 采用 Adam 自适应学习率优化器和 mini-batch 训练法^[24], 每批 64 个样本, 训练 200 轮。为了减少因深层网络而易产生的过拟合现象, 本模型引入了深度学习训练技巧, 在每一网络层都引入了批量正则化处理, 同时设置了两个 Dropout 层^[19], 随机的将 30% 的神经元置零。图 9 为未使用训练技巧的模型收敛效果, 图 10 为加入训练技巧后的模型收敛效果, 对比可见, 建立的模型明显减少了过拟合现象并具有更好的收敛效果。

表 2 建立的诊断模型超参数明细表

序号	网络层	超参数	输出形状	参数量
0	输入层	原始数据 $[10\ 000, 6]$	$[batch, 10, 6]$	0
1	1D 卷积层 1	核长: $[3, 6, 128]$, 步长: $[1, 0]$	$[batch, 10, 128]$	2 432
2	激活层 1	ReLU 激活函数	$[batch, 10, 128]$	0
3	Dropout 层 1	Dropout (0.3)	$[batch, 10, 128]$	0
4	1D 卷积层 2	核长: $[3, 128, 64]$, 步长: $[1, 0]$	$[batch, 10, 64]$	24 640
5	激活层 2	ReLU 激活函数	$[batch, 10, 64]$	0
6	1D 池化层	池化核: $[2, 1, 64]$, 步长: $[2, 0]$	$[batch, 5, 64]$	0
7	1D 卷积层 3	核长: $[3, 64, 32]$, 步长: $[1, 0]$	$[batch, 5, 32]$	6 176
8	激活层 3	ReLU 激活函数	$[batch, 5, 32]$	0
9	Dropout 层 2	Dropout (0.3)	$[batch, 5, 32]$	0
10	1D 卷积层 4	核长: $[3, 32, 6]$, 步长: $[1, 0]$	$[batch, 5, 6]$	582
11	1D-GAP 层	池化核: $[5, 1, 6]$, 步长: $[5, 0]$	$[batch, 6]$	0
12	Softmax 层	Softmax 分类函数	$[batch, 6]$	0

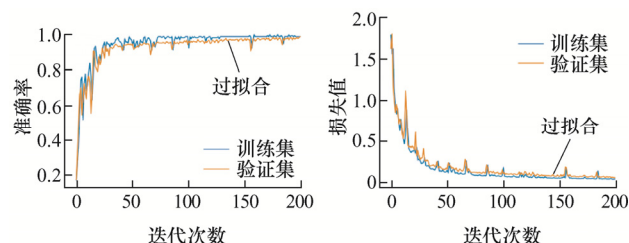


图 9 未加入训练技巧的模型准确率和误差收敛曲线

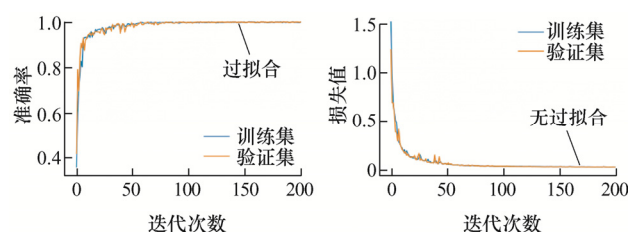


图 10 加入训练技巧后的模型准确率和误差收敛曲线

3.3 诊断结果分析

通过对表 1 所示的数据集进行训练, 最终得到的诊断结果如表 3 所示。在表 3 中, 本文采用了查准率(Precision ratio, PR, 又称精确率)、灵敏度(Recall ratio, RR, 又称召回率)以及 F1 均值^[26]对模型结果进行评估。在传统的 1DCNN 算法中, 模型末端通过采用了一个 2~3 层的全连接结构^[23], 为了进一步验证提出的算法的改进效果, 本文将表 2 中的 GAP 层更换为一个具有 3 层隐含层的全连接网络, 超参数为 256-128-6, 采用同样的数据集对传统 1DCNN-FC 算法进行训练, 两种模型最终得到的诊断结果如表 3 所示, 表 4 给出了两个诊断模型的训练参数的数量。

表 3 两种算法的诊断结果对比

模型	改进 1DGAP-CNN 算法			传统 1DCNN-全连接算法			
训练时间/s	162.83			201.45			
测试时间/s	0.202			0.358			
评估指标	查准率 (%)	召回率 (%)	F1 均值 (%)	查准率 (%)	召回率 (%)	F1 均值 (%)	样本数
正常状态	100.0	100.0	100.0	100.0	100.0	100.0	300
定点异常	100.0	100.0	100.0	99.67	100.0	99.83	300
X 向偏移	98.31	97.00	97.65	97.96	96.00	96.97	300
Y 向偏移	100.0	99.67	99.83	100.0	99.67	99.83	300
X 向缓变	96.72	98.33	97.52	96.08	98.00	97.03	300
Y 向缓变	100.0	100.0	100.0	100.0	100.0	100.0	300
平均值	99.17	99.17	99.17	98.95	98.94	98.94	1 800

表 4 两种诊断算法的参数量对比

网络层	改进的 1DGAP-CNN	传统的 1DCNN-全连接
1D 卷积层 1	2 432	2 432
1D 卷积层 2	24 640	24 640
1D 卷积层 3	6 176	6 176
1D 卷积层 4	582	无
1D-GAP 层	0	无
Flatten 层	无	0
全连接层 1	无	41 216
全连接层 2	无	32 896
全连接层 3	无	774
模型总参数量	33 830	108 134

通过对比表 3 和表 4 的结果可以看出。

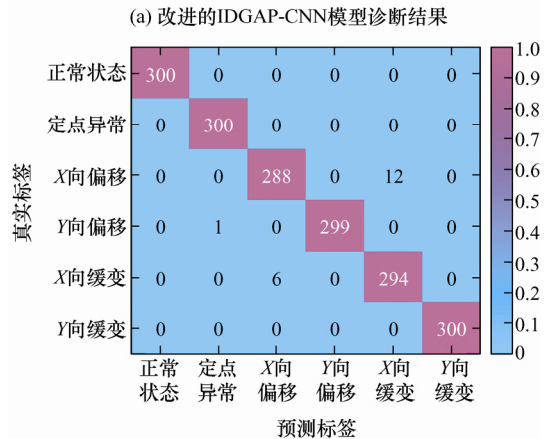
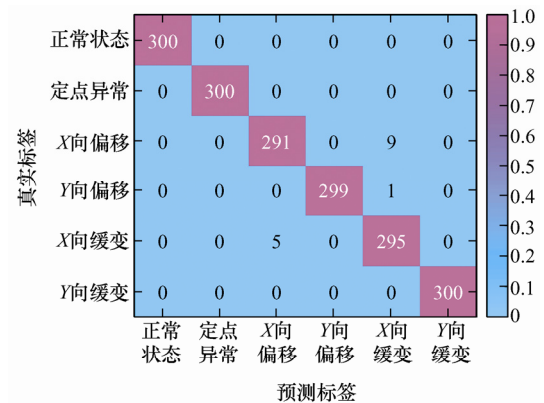
(1) 在诊断准确率方面, 传统的 1DCNN-FC 算法的准确率为 98.94%, 而提出的算法的诊断准确率已提升至 99.17%, 且对正常状态、定点攻击和 Y 向缓变攻击的辨识度达 100%, 对 X 向偏移、Y 向偏移和 X 向缓变攻击的辨识度分别为 97.65%、99.83% 和 97.52%, 相比传统方法具有更高的诊断准确率。

(2) 在模型参数量方面, 传统的 1DCNN-FC 算法总参数量为 108 134 个, 3 层的全连接网络参数量为 74 886 个; 而采用一维全局均值池化改进后的 1DGAP-CNN 算法的模型参数量仅为 33 830, 相比传统 1DCNN-FC 减少参数量 74 304 个, 约占总参数量的 70%, 有效减少了模型参数量。

(3) 在诊断时间方面, 传统的方法的训练时间和测试时间分别为 201.45 s 和 0.358 s, 而改进算法的训练时间和测试时间分别为 162.83 s 和 0.202 s, 尤其是测试时间上, 改进的 1DGAP-CNN 算法有效减少诊断时间约 43%, 给无人车网络攻击检测提供了更多的报警排障时间, 这对网络攻击的实时诊断具有重要现实意义。

为了进一步量化两种算法对不同异常样本误判的具体情况, 本文中引入了多分类错误量化矩阵^[16]对表 3 中的诊断结果进一步量化, 表 3 对应的错误矩阵如图 11 所示。在图 11 中, 横轴代表算法预测

的结果, 纵轴为样本真实的标签, 主对角线上的数值为诊断正确的样本数量, 非对角线上的数值代表误判的数量, 误判样本所在的位置的横纵坐标即显示了误判的信息。从图 11a 中可以看出, 改进的算法在 1 800 个样本中, 仅有 15 个样本被误判, 其中有 9 个 X 向偏移被误判为 X 向缓变, 5 个 X 向缓变误判为 X 向偏移, 分析原因主要是 X 向偏移和 X 向缓变在初始位置时的样本具有相似性, 虽然造成误判, 但是也说明诊断模型已察觉该样本的异常, 只是具体异常类型发生了混淆, 然而对正常和异常状态的辨识度已达 100%。从图 11b 中可以看出, 传统的 1DCNN-FC 算法有 19 个样本被误判, 效果略低于改进的方法。由此可见, 改进的算法相比传统 1DCNN-FC 算法具有更好的诊断性能。



(b) 传统的 1DCNN-FC 模型诊断结果

图 11 两种算法诊断结果的多分类错误矩阵

3.4 与其他算法对比验证

为进一步验证改进的 1DGAP-CNN 算法相比传统智能诊断方法的有效性和优越性, 本文与现行的支持向量机(Support vector machine, SVM)、K 近邻(K-nearest neighbor, KNN)和深层 BP 神经网络(Deep back-propagation neural network, DNN)三种主流算法进行对比验证^[24], 三种算法仍然采用表 1 所示的无人车数据进行训练和测试, 最终得到的诊断结果如表 5 所示。以上三种模型的超参数分别是: SVM

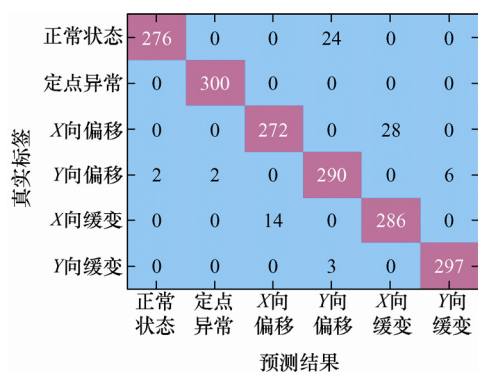
采用了高斯径向基核函数、惩罚系数 $C=10.0$ 、松弛变量 $\xi=0.01$ ；KNN 采用 Minkowski 距离， k 值取 10，叶子结点为 30；DNN 采用了 5 层网络，其中隐含层有 4 层，结点个数分别为 512-256-128-64-6(输出层)，采用 Tanh 激活函数和 Softmax 分类器，采用 Adam 优化器，加入正则化项($\lambda=0.001$)，交叉熵损失函数，迭代次数为 200 次。

对比表 5 可以看出，在准确率方面：SVM、KNN、DNN 和传统 1D-CNN 的准确率分别为 90.28%、84.74%、95.61%和 98.94%，而本文提出的方法准确率高达 99.17%。在训练时间方面：提出的方法训练用时为 162.83 s，处于 DNN 和传统 1DCNN 算法中间水平，由于 SVM 和 KNN 属于浅层机器学习

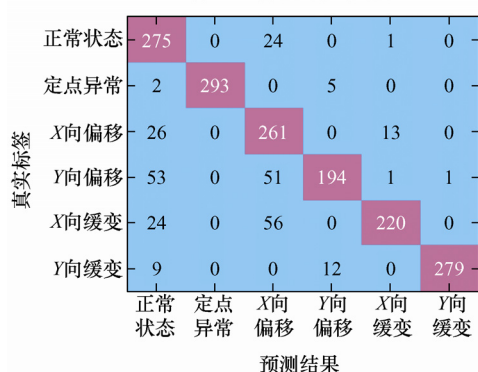
算法，其训练机理与深度学习算法不同，因此模型训练时间不具可比性；在测试时间方面，提出的算法用时最短，仅为 0.202 s，虽然 DNN 用时 0.247 s 与提出的算法比较接近，但是 DNN 的诊断准确率仅为 95.61%，其误诊率过大。图 12 给出了三种对比算法诊断结果的错误矩阵，从图 12 可以看出，SVM、KNN 和 DNN 三种算法的误判样本数量分别是 79、176 和 278 个，并且三种算法均出现了对正常样本的误判。因此，综合对比以上 5 种算法结果可以看出，提出的 1DGAP-CNN 算法相比 SVM、KNN、DNN 和传统 1DCNN 算法具有更加优越的诊断性能和更快的检测速度，更加适用于无人车导航传感器网络攻击异常的快速智能诊断。

表 5 五种诊断算法的结果对比

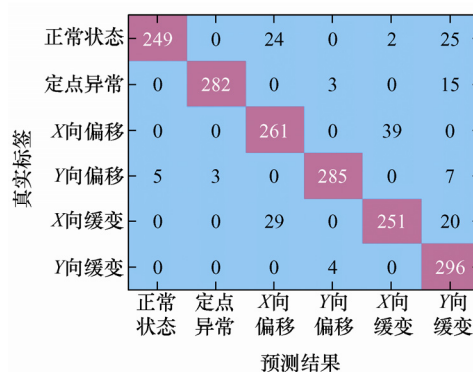
模型	改进1DGAP-CNN算法			传统1DCNN-全连接算法			SVM算法			KNN算法			DNN算法			样本数
训练时间/s	162.83			201.45			5.463			—			99.74			
测试时间/s	0.202			0.358			0.936			0.405			0.247			
评估指标	查准率 (%)	召回率 (%)	F1 均值 (%)	查准率 (%)	召回率 (%)	F1 均值 (%)	查准率 (%)	召回率 (%)	F1 均值 (%)	查准率 (%)	召回率 (%)	F1 均值 (%)	查准率 (%)	召回率 (%)	F1 均值 (%)	
正常状态	100.0	100.0	100.0	100.0	100.0	100.0	98.03	83.00	89.89	70.69	91.67	79.83	99.28	92.00	95.50	300
定点异常	100.0	100.0	100.0	99.67	100.0	99.83	98.95	94.00	96.41	100.0	97.67	98.82	99.34	100.0	99.67	300
X 向偏移	98.31	97.00	97.65	97.96	96.00	96.97	83.12	87.00	85.02	66.58	87.00	75.43	95.10	90.67	92.83	300
Y 向偏移	100.0	99.67	99.83	100.0	99.67	99.83	97.60	95.00	96.28	91.94	64.67	75.93	91.48	96.67	94.00	300
X 向缓变	96.72	98.33	97.52	96.08	98.00	97.03	85.96	83.67	84.80	93.62	73.33	82.24	91.08	95.33	93.16	300
Y 向缓变	100.0	100.0	100.0	100.0	100.0	100.0	81.54	98.67	89.29	99.64	93.00	96.21	98.02	99.00	98.51	300
平均值	99.17	99.17	99.17	98.95	98.94	98.94	90.87	90.22	90.28	87.08	84.56	84.74	95.72	95.61	95.61	1 800



(a) SVM 算法的诊断结果



(b) KNN 算法的诊断结果



(c) DNN 算法的诊断结果

图 12 三种对比算法诊断结果的多分类错误矩阵

4 结论

无人驾驶汽车作为一种新兴的智能交通运输工具，在给人们的出行带来巨大交通便利的同时也面临着诸多安全挑战。在无人驾驶环境下，网络黑客作为一种无形的“杀手”给无人车的安全运行带来了巨大威胁。本文针对无人车导航传感器易受黑客

攻击的现实问题, 提出了一种基于改进 1D-GAP-CNN 的深度学习新算法, 用于检测无人驾驶汽车的 GPS 导航传感器因物理故障或网络黑客恶意攻击而产生的多种异常问题。

本文首先将深度学习中最前沿的一维卷积神经网络算法引入到无人车传感器网络攻击诊断领域; 其次, 针对无人车网络攻击诊断的实时性要求和现行 1D-CNN 算法参数量过多的不足, 通过设计了一个一维全局均值池化层来代替现行 1D-CNN 中的 Flatten 层和 2~3 层的全连接网络部分。结果表明, 改进的方法可以有效减少模型参数量近 70%, 提升检测速度约 40%, 且诊断准确率达到 99% 以上。提出的方法无需对导航传感器原始数据做任何的人工干预和特征工程操作, 1D-GAP-CNN 算法可自动完成多传感器数据融合、特征提取、维度变换、降维减参和故障辨识等操作, 诊断结果实时输出。

通过与主流的 SVM、KNN、DNN 和传统的 1D-CNN 智能算法对比验证, 结果表明, 提出的方法具有更快的检测速度和更高的诊断准确率, 更加适用于无人车网络攻击的实时快速诊断。研究结果可为后续的无人车在线诊断与网络反攻击等提供新思路, 具有重要的参考意义和借鉴价值。

参 考 文 献

- [1] NING Z, ZHANG K, WANG X, et al. Joint computing and caching in 5G-envisioned Internet of vehicles: A deep reinforcement learning-based traffic control system[J]. IEEE Transactions on Intelligent Transportation Systems, 2020.1-12.
- [2] BAGLOEE S A, TAVANA M, ASADI M, et al. Autonomous vehicles: challenges, opportunities, and future implications for transportation policies[J]. Journal of Modern Transportation, 2016, 24(4): 284-303.
- [3] RASHEED I, HU F, ZHANG L. Deep reinforcement learning approach for autonomous vehicle systems for maintaining security and safety using LSTM-GAN[J]. Vehicular Communications, 2020, 26: 100266.
- [4] DETHE S N, SHEVATKAR V S, BIJWE R P. Google driverless car[J]. International Journal of Scientific Research in Science, Engineering and Technology, 2011, 2(2): 133-137.
- [5] FERDOWSI A, CHALLITA U, SAAD W, et al. Robust deep reinforcement learning for security and safety in autonomous vehicle systems [C]// 21st International Conference on Intelligent Transportation Systems (ITSC). IEEE, 2018, 2018: 1-8.
- [6] SHI X, WONG Y D, CHAI C, et al. An automated machine learning (AutoML) method of risk prediction for decision-making of autonomous vehicles[J]. IEEE Transactions on Intelligent Transportation Systems, 2020: 1-10.
- [7] 赵治国, 周良杰, 朱强. 无人驾驶车辆路径跟踪控制预瞄距离自适应优化[J]. 机械工程学报, 2018, 54(24): 166-173.
ZHAO Zhiguo, ZHOU Liangjie, ZHU Qiang. Preview distance adaptive optimization for the path tracking control of unmanned vehicle[J]. Journal of Mechanical Engineering, 2018, 54(24): 166-173.
- [8] LIU Q, MO Y, MO X, et al. Secure pose estimation for autonomous vehicles under cyber attacks [C]// 2019 IEEE Intelligent Vehicles Symposium (IV). IEEE, 2019: 1583-1588.
- [9] WANG Y, MASOUD N, KHOJANDI A. Real-Time sensor anomaly detection and recovery in connected automated vehicle sensors[J]. IEEE Transactions on Intelligent Transportation Systems, 2020: 1-11.
- [10] CURTIS S. Hacker remotely crashes Jeep from 10 miles away[J]. The Telegraph, 2015.
- [11] SOLON O. Team of hackers take remote control of Tesla Model S from 12 miles away[J]. The Guardian, 2016.
- [12] FRANCO V, WANG Y, KHOJANDI A, et al. Real-time sensor anomaly detection and identification in automated vehicles[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 21(3): 1264-1276.
- [13] MO Y, SINOPOLI B. On the performance degradation of cyber-physical systems under stealthy integrity attacks[J]. IEEE Transactions on Automatic Control. 2016, 61(9): 2618-2624.
- [14] LI Y, ZHU L. A Bayesian game based defense scheme for CBTC systems under man-in-the-middle attacks [C]// 2019 IEEE Intelligent Transportation Systems Conference, IEEE, 2019: 2172-2176.
- [15] ULLAH F, NAEEM H, JABBAR S, et al. Cyber security threats detection in internet of things using deep learning approach[J]. IEEE Access. 2019, 7(99): 124379-124389.
- [16] GONG W, CHEN H, ZHANG Z, et al. A data-driven-based fault diagnosis approach for electrical power DC-DC inverter by using modified convolutional neural network with global average pooling and 2-D feature image[J]. IEEE Access, 2020, 8: 73677-73697.
- [17] LECUN Y, BOSER B, DENKER J S, et al.

- Backpropagation applied to handwritten zip code recognition[J]. *Neural Computation*, 1989, 1(4): 541-551.
- [18] GOODFELLOW I, BENGIO Y, COURVILLE A. Deep learning [M]. The MIT Press, 2016.
- [19] 宫文峰, 陈辉, 张美玲, 等. 基于深度学习的电机轴承微小故障智能诊断方法[J]. *仪器仪表学报*, 2020, 41(01): 195-205.
- GONG Wenfeng, CHEN Hui, ZHANG Meiling, et al. Intelligent diagnosis method for incipient fault of motor bearing based on deep learning[J]. *Chinese Journal of Scientific Instrument*, 2020, 41(01): 195-205.
- [20] XIA M, LI T, XU L, et al. Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks[J]. *IEEE/ASME Transactions on Mechatronics*, 2018, 23(1): 101-110.
- [21] 姜洪权, 贺帅, 高建民, 等. 一种改进卷积神经网络模型的焊缝缺陷识别方法[J]. *机械工程学报*, 2020, 56(8): 235-242.
- JIANG Hongquan, HE Shuai, GAO Jianmin, et al. An improved convolutional neural network for weld defect recognition[J]. *Journal of Mechanical Engineering*, 2020, 56(8): 235-242.
- [22] LIN M, CHEN Q, YAN S C. Network in network [C]// *International Conference on Learning Representations*, 2014: 1-10.
- [23] 曲建岭, 余路, 袁涛, 等. 基于一维卷积神经网络的滚动轴承自适应故障诊断算法[J]. *仪器仪表学报*, 2018, 39(7): 134-143.
- QU Jianling, YU Lu, YUAN Tao, et al. Adaptive fault diagnosis algorithm for rolling bearings based on one-dimensional convolutional neural network[J]. *Chinese Journal of Scientific Instrument*, 2018, 39(7): 134-143.
- [24] 宫文峰, 陈辉, WANG Danwei, 等. 基于改进 CNN-GAP-SVM 的船舶电力变换器快速故障诊断方法[J/OL]. *计算机集成制造系统*, 2020: 1-18.
- GONG Wenfeng, Chen Hui, Danwei WANG, et al. Fast fault diagnosis method of marine electrical converter based on improved CNN-GAP-SVM algorithm[J]. *Computer Integrated Manufacturing Systems*, 2020: 1-18.
- [25] WEN L, LI X, GAO L, et al. A new convolutional ceural network-based data-driven fault diagnosis method[J]. *IEEE Transactions on Industrial Electronics*, 2018, 65(7): 5990-5998.
- [26] SHAO H, JIANG H, ZHAO H, et al. A novel deep autoencoder feature learning method for rotating machinery fault diagnosis[J]. *Mechanical Systems and Signal Processing*, 2017, 95: 187-204.
-
- 作者简介: 宫文峰, 男, 1987 年出生, 讲师/工程师, 武汉理工大学与新加坡南洋理工大学联合培养博士研究生。主要研究方向为复杂装备故障诊断与寿命预测、深度学习与人工智能、机械振动与 CAE 技术等。
E-mail: wfgongcn@163.com
- 王元哲, 男, 1990 年出生, 博士, 新加坡南洋理工大学研究员。主要研究方向为机器人学, 控制工程, 网络安全。
E-mail: wang0951@e.ntu.edu.sg
- 陈辉, 男, 1962 年出生, 博士, 博士研究生导师, 国家二级教授。主要研究方向为船舶电力推进系统和智能装备控制。
E-mail: hchen@whut.edu.cn
- WANG Danwei, 男, 1957 年出生, 博士, 博士研究生导师, 新加坡南洋理工大学终身教授、新加坡工程院院士, IEEE Fellow。主要研究方向为机器人学, 控制工程, 故障诊断。
E-mail: edwwang@ntu.edu.sg