



基于并行深度神经网络的物联网边缘侧隐私数据检测

周少鹏^{1,3}, 王滨^{2,3*}, 李超豪^{1,3}, 周政演¹, 张峰^{3,4}, 吴春明¹

1. 浙江大学计算机科学与技术学院, 杭州 310027

2. 西安电子科技大学杭州研究院, 杭州 311231

3. 浙江全省智能物联网与数据安全重点实验室, 杭州 310000

4. 西安电子科技大学网络与信息安全学院, 西安 710071

* 通信作者. E-mail: wangbin02@xidian.edu.cn

收稿日期: 2025-05-30; 修回日期: 2025-07-23; 接受日期: 2025-08-27; 网络出版日期: 2025-10-17

国家自然科学基金 (批准号: 62472335) 和浙江省科技厅“尖兵领雁”项目 (批准号: 2024C01165, 2024C01066) 资助

摘要 针对物联网隐私检测中边缘设备资源受限、新型隐私泄露泛化检测能力不足及精度与效率难以平衡三大挑战, 本文提出基于并行深度神经网络的物联网边缘侧隐私数据检测系统 (PDNN-PDE). 该系统首先通过轻量级多模态数据编码器, 高效处理图像、时序传感器信号及日志文本等异构数据; 再依托并行深度神经网络隐私检测引擎, 在边缘侧设备实现多模态隐私数据的分布式实时检测; 同时基于云-边协同的动态增量学习算法, 周期性更新模型以增强对新型隐私泄露模式的泛化检测能力. 实验结果表明, PDNN-PDE 检测精度较单模态轻量模型平均提升 3.6%; 融合模态在检测精度仅降低 0.7% 的情况下, 内存占用较现有专用模型减少 76%, 推理延迟减少 21.3%; 在新型泄露模式自适应方面, 经 70 轮增量学习后模型准确率恢复至 89%, 较静态模型提升约 44%. 上述结果证明 PDNN-PDE 可在资源受限物联网边缘环境中实现高效鲁棒且自适应的隐私数据检测.

关键词 物联网边缘计算, 隐私数据检测, 并行深度学习, 多模态数据, 增量学习

1 引言

物联网 (IoT) 技术正以前所未有的速度蓬勃发展, 深刻改变着人类与物理世界的互动方式^[1]. 预计到 2030 年, 全球物联网设备数量将从 2023 年的 159 亿台激增至 321 亿台以上^[2], 致使数据生成范式发生根本性转变. 预计到 2026 年, 全球物联网设备产生的数据总量将达到 100 ZB^[3], 涵盖文本、图像、时序等多种模态数据. 然而这些数据往往包含大量敏感隐私信息, 如个人信息、商业机密、关键基础设施信息等. 物联网设备与数据量的爆发式增长, 显著加剧了隐私泄露风险: 海量异构数据令人工审查失效且传统查控机制超载, 设备离散部署造成攻击面扩散和集中监控困难, 边缘资源受限阻碍复

引用格式: 周少鹏, 王滨, 李超豪, 等. 基于并行深度神经网络的物联网边缘侧隐私数据检测. 中国科学: 信息科学, 2025, 55: 2870–2885, doi: 10.1360/SSI-2025-0253
Zhou S P, Wang B, Li C H, et al. PDNN-PDE: privacy data detection for the edge IoT based on P-DNN. Sci Sin Inform, 2025, 55: 2870–2885, doi: 10.1360/SSI-2025-0253

杂模型部署,业务数据连续性则对检测实效性提出了更高要求.相较于云端依赖高带宽传输、集中式处理、复杂模型机制的范式,边缘侧隐私数据泄露检测需实现强实时响应、数据边缘化处理、碎片化资源适配及不稳定网络环境适应.这使其超越基础合规需求,成为物联网可持续发展与国家关键基础设施安全的战略基石.

物联网边缘侧新型隐私数据泄露是指,在物联网边缘侧因设备安全性不足、数据传输过程缺乏保护机制、边缘节点存在漏洞和数据处理不当等原因,导致隐私数据未经授权被泄露、截获或篡改.与传统的云端数据泄露不同,边缘侧泄露发生在接近数据源的地方,面临更复杂的安全挑战和多层攻击面.当前隐私数据检测的方法大致可分为规则匹配、机器学习、深度学习、轻量级神经网络以及更先进的多模态模型.尽管这些方法在不同程度上表现出一定的有效性,但在物联网边缘设备资源受限的环境下难以有效运行,对新型隐私泄露模式的泛化检测能力不足,并且难以平衡检测精度与运行效率.规则匹配技术原理简单、部署成本低^[4],但在动态物联网环境中,面对新型传输协议和异构数据,其静态规则集难以实现语义深度解析,导致漏报率与误报率显著升高,且规则库更新滞后.机器学习(如支持向量机、决策树)、深度学习(如卷积神经网络、长短期记忆网络)能够自动提取复杂模式特征,在多模态数据处理及检测中表现出良好的泛化性能;然而其对高质量标注数据存在强依赖,且模型训练和推理开销在边缘节点上易导致资源过载^[5].深度学习(如卷积神经网络、长短期记忆网络)能够捕捉更高维度的时空特征,在复杂隐私泄露场景中表现出更优的检测精度;但其复杂网络结构同样带来显著的计算和存储开销,在资源受限的物联网设备上难以直接部署^[6].轻量级深度学习模型如 MobileNet V2^[7] 和 ShuffleNet V2^[8] 通过模型压缩、深度可分离卷积等技术大幅降低参数数量和计算复杂度,但往往伴随检测性能的下降;此外,现有轻量级架构在多模态融合和对新型隐私泄露模式的动态适应能力方面仍有不足.更先进的多模态模型则尝试融合网络流量、设备上下文、传感器信号等多源异构信息,以提升对复杂隐私泄露行为的识别能力;但此类模型在边缘侧的部署仍面临更高的资源开销和协同调度挑战^[9].

为解决上述物联网边缘侧隐私数据检测中的关键挑战,本文提出一种基于并行深度神经网络的物联网边缘侧隐私数据检测框架(PDNN-PDE),首先通过轻量级多模态数据编码器,高效处理图像、时序传感器信号及日志文本等异构数据,然后依托并行深度神经网络隐私检测引擎,在边缘侧设备实现多模态隐私数据的分布式实时检测.同时基于云-边^[10]协同的动态增量学习算法,周期性更新模型以增强对新型隐私泄露模式的泛化检测能力.

本文提出的 PDNN-PDE 系统在文本隐私数据集、Privacy Alert 图像数据集^[11] 和 MultiIoT^[12] 时序数据集上进行了全面实验评估,并在工业 PC (personal computer), Jetson Nano 和 Raspberry Pi 4 等多种异构边缘计算平台上验证了其性能.实验结果表明,PDNN-PDE 检测精度较单模态轻量模型平均提升 3.6%;融合模态在检测精度仅降低 0.7% 的情况下,内存占用较现有专用模型减少 76%,推理延迟减少 21.3%;在新型泄露模式自适应方面,经 70 轮增量学习后模型准确率恢复至 89%,较静态模型提升约 44%.实验结果证明,PDNN-PDE 可在资源受限物联网边缘环境中实现高效鲁棒且自适应的隐私数据检测.

本文的主要贡献包括以下 3 点.

(1) 发现并深入分析了物联网边缘侧隐私数据检测中资源受限与检测性能效率之间的核心矛盾,并揭示了现有方法在应对新型隐私泄露模式和边缘设备约束时的固有局限.

(2) 设计并提出了基于并行深度神经网络的物联网边缘侧隐私数据检测系统(PDNN-PDE),包括轻量级多模态数据编码器、并行深度神经网络(PDNN)隐私检测引擎以及云-边协同动态增量学习模型更新机制,实现物联网边缘侧多模数据高效并行处理、动态隐私泄露检测与边缘算力适配.

(3) 实现了 PDNN-PDE 原型系统并在多模态数据集和异构边缘计算设备上进行广泛实验.结果表明 PDNN-PDE 检测精度较单模态轻量模型平均提升 3.6%;融合模态在检测精度仅降低 0.7% 的情况下,内存占用较现有专用模型减少 76%,推理延迟减少 21.3%;在新型泄露模式自适应方面,经 70 轮

增量学习后模型准确率恢复至 89%, 较静态模型提升约 44%. 充分证明了 PDNN-PDE 在资源受限物联网边缘环境中实现高效、鲁棒和自适应隐私数据检测的显著优势.

2 相关工作

2.1 物联网隐私数据特点及风险

物联网环境中的隐私数据类型多样, 主要包括用户隐私数据、位置隐私数据、模型隐私数据和设备隐私数据等. 用户隐私数据涵盖个人身份信息与行为习惯数据^[13], 前者作为用户身份识别的核心标识, 包括姓名、身份证号码、电话号码等关键信息, 其泄露可能导致身份盗用、诈骗等安全风险; 后者通过智能设备对用户行为的持续监测生成, 如智能家居记录的作息规律、智能穿戴设备采集的运动偏好等^[14], 这类数据反映的用户生活模式若被非法获取, 可能引发精准广告轰炸甚至行为操纵等问题. 位置隐私数据^[15] 包含通过智能手机、智能车辆等设备实时采集的位置坐标及移动轨迹, 其泄露可能导致用户行踪被追踪, 在共享出行等场景中可能直接威胁人身安全. 模型隐私数据涉及机器学习模型在训练与应用过程中的敏感信息, 包括原始训练数据 (如身份信息、健康数据)、模型参数、算法结构及中间训练结果等^[16], 其泄露可能引发成员推断、模型反演等攻击. 设备隐私数据则包含设备状态信息 (如设备运行参数) 与配置信息 (如设备密钥), 这些数据的泄露可能导致设备被恶意控制, 在工业物联网场景中甚至可能引发生产事故与经济损失.

物联网异构设备的广泛部署导致物联网数据模态、格式及结构的高度复杂化, 显著增加了隐私数据检测的技术难度. 同时物联网数据具有动态性特征, 表现为数据随时间与用户行为持续更新, 要求隐私数据检测技术具备实时适应性, 能够对增量数据进行精准检测. 此外, 物联网数据具有海量且分布式特征, 数据在多设备、多环节的存储与传输, 这种去中心化的架构显著扩大了潜在隐私数据泄露面, 进一步提升了物联网隐私数据检测方案的实施难度.

2.2 现有隐私数据识别方法

早期隐私数据识别通常依赖于规则匹配技术, 通过使用预定义的规则、关键词和正则表达式来识别隐私数据. 这类方法原理简单, 部署成本较低, 主要应用于隐私模式相对静态和明确的场景, 例如在结构化日志中检测特定身份号码. 然而在动态的物联网环境中, 规则匹配技术在语义深度解析方面存在固有限制, 尤其是在面对新型传输协议和异构数据时, 其静态规则集缺乏灵活性, 无法有效适应不断变化的物联网数据特征, 从而导致了漏报率和误报率显著上升. 此外, 规则库的更新往往滞后于新型隐私泄露模式的出现, 使得其无法应对新型隐私威胁. 同样, 传统机器学习方法, 如 K 近邻 (KNN) 和支持向量机 (SVM) 等分类算法通过数据训练自动学习隐私数据与非隐私数据的特征边界, 在复杂数据模式识别中展现出更强的泛化能力. 然而其高度依赖人工特征工程且捕获多模态数据中复杂非线性关系的能力不足, 难以处理物联网环境中多模异构、大规模的动态数据流.

经典深度学习模型的出现, 包括卷积神经网络 (CNN)、循环神经网络 (RNN)、长短期记忆 (LSTM) 网络和门控循环单元 (GRU), 标志着数据处理和模式识别的重大进步. 这些模型已被广泛应用于图像识别、自然语言处理和时间序列分析等任务, 并展现出优于传统机器学习方法的性能. 更专业的深度学习模型, 如 Faster R-CNN^[17] 和 YOLO^[18] 等目标检测模型, 通过准确识别图像中的敏感区域, 在视觉隐私检测方面取得了显著成功. 同样, Bi-LSTM-CRF^[19], RoBERTa^[20] 和 DistilBERT^[21] 等先进的自然语言处理模型, 通过捕获日志数据中复杂的语义关系, 显著提升了基于文本的隐私检测能力. 对于时间序列数据, Informer 等模型在长序列预测和异常检测方面表现出卓越能力^[22]. 然而深度学习模型因其庞大的架构和高计算/高内存占用等特性, 无法直接部署在资源受限的物联网边缘设备上, 尤其是在考虑多模态数据流和实时处理需求时. 同样, 多模态融合技术可以整合来自图像、文本和时序传感器等不同来源的信息, 从而实现更全面的理解. 例如, ViLBERT^[23] 利用协同注意力层并行处理文

本和视觉特征, 捕获细微的模态间关系. 然而, 这些复杂的多模态融合模型虽然功能强大, 但通常伴随着巨大的计算和内存开销, 同样无法在资源受限的边缘设备上直接部署.

2.3 面向边缘设备的轻量级神经网络

为了应对边缘设备的计算和内存限制, 大量研究致力于开发轻量级神经网络. MobileNet V2 和 ShuffleNet V2 等模型采用高效架构 (如深度可分离卷积) 来减少参数数量和计算成本, 使其适用于边缘设备的图像处理. 类似地, Tiny-LSTM^[24] 和 Tiny-TCN^[25] 已在资源受限环境中用于高效的时间序列处理. 这些模型优先考虑效率, 通常比其大型对应模型实现更低的单样本推理延迟和内存占用. 然而其规模的显著压缩可能导致检测精度下降. 除了架构设计, 模型压缩技术被广泛用于优化深度学习模型以适应边缘部署. 关键策略包括量化 (降低权重和激活的数值精度, 例如转换为 INT8)、剪枝 (移除冗余连接或神经元) 和知识蒸馏 (训练一个较小的“学生”模型以模仿一个较大的“教师”模型). 尽管这些技术有效地减小了模型大小、内存使用和推理时间, 但它们常常面临压缩比和性能下降之间的权衡. 在保持最先进精度的同时实现“无损”压缩, 特别是对于复杂的多模态任务, 仍然是一个重大挑战. 此外, 许多轻量级模型通常是单模态任务设计的, 缺乏有效且鲁棒地融合多模态异构物联网数据流的固有机制.

与 MobileNet V2 和 Tiny-LSTM 等仅针对单模态设计的轻量模型不同, PDNN-PDE 采用并行多模态架构, 为图像、文本、时序数据分别构建专用轻量分支, 实现异构数据并行处理. 其轻量级的编码器针对物联网数据特性定制, 完成去噪、关键词过滤与 ROI (region of interest) 提取, 输出由模态专用分支精准检测, 兼顾效率与精度. 同时 PDNN-PDE 引入云-边协同动态增量学习机制, 可在线识别新型隐私模式并轻量更新模型, 而传统轻量模型多为静态部署, 无法自适应演化威胁. 这些设计使 PDNN-PDE 在保持低资源开销的同时, 具备更强的实用性、灵活性与长期鲁棒性.

3 基于并行深度神经网络的物联网边缘侧隐私数据检测系统设计

本文提出了一种基于并行深度神经网络的物联网边缘侧隐私数据检测系统, 旨在解决边缘设备资源受限、新型隐私泄露模式泛化检测能力不足以及检测精度与效率难以平衡的三大挑战. 系统主要包括以下 3 个模块. (1) 轻量级多模态数据编码器: 负责对多模态物联网异构数据, 包括图像、时序传感器信号、日志文本等进行数据清洗、归一化及跨模态特征融合, 为检测模型提供高质量输入. (2) PDNN 隐私检测引擎: 基于并行深度神经网络实现多特征空间的分布式检测, 在边缘侧设备完成隐私数据实时检测. (3) 云-边协同动态增量学习模型更新模块: 基于云-边协同的强化学习, 通过设备间的策略交互与云端参数聚合, 动态优化检测模型对新型隐私泄露模式的泛化能力, 并平衡检测精度与开销. 系统整体架构如图 1 所示.

3.1 轻量级多模态数据编码器

轻量级多模态数据编码器旨在实现异构模态数据的轻量化统一处理. 针对物联网边缘环境下面临的噪声污染、格式异构及维度爆炸等关键挑战, 采用模态差异化策略, 适配不同数据模态的独特噪声模式与隐私特性, 为后续隐私泄露检测引擎提供低维、标准化的特征向量. 该编码器包含数据清洗模块、特征归一化模块和维度精简模块, 具体设计如图 2 所示.

(1) 文本数据处理. 轻量级多模态数据编码器物联网设备日志通常为非结构化文本 (如 “2023-05-20 14:30:23 [INFO] upload private data to server: 192.168.1.10”), 含大量冗余信息 (如时间戳、日志级别), 基于词袋模型的原始维度一般为 1000~2000 维. 针对文本数据, 定义原始输入文本数据为 T , 首先进行关键词清洗与标准化, 过滤停用词 (如 “[INFO]” “to”). 定义隐私敏感关键词集合 $\mathcal{K} = \{k_1, k_2, \dots, k_M\}$ (其中 $M = 200$ 为关键词总数), 保留隐私敏感关键词 (如 “upload” “private data” “server” “GPS” “token”),

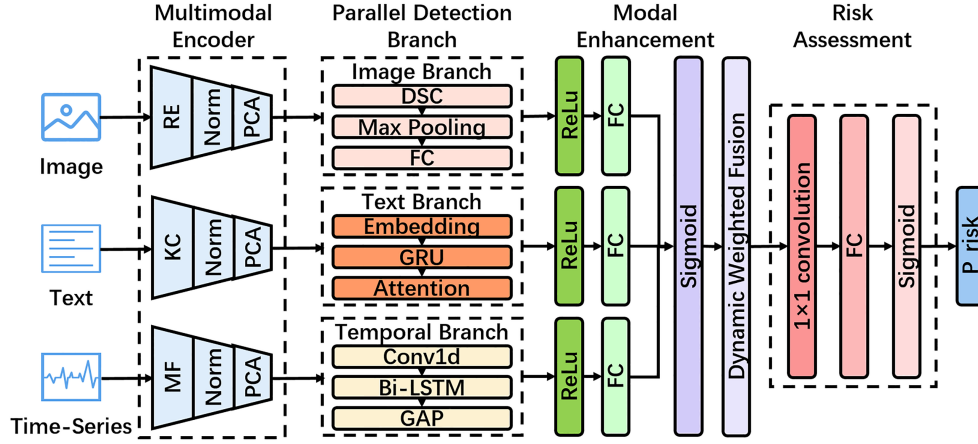


图 1 (网络版彩图) 基于并行深度神经网络的物联网边缘侧隐私数据检测系统架构.

Figure 1 (Color online) Architecture of the privacy data detection framework for the edge IoT based on PDNN-PDE.

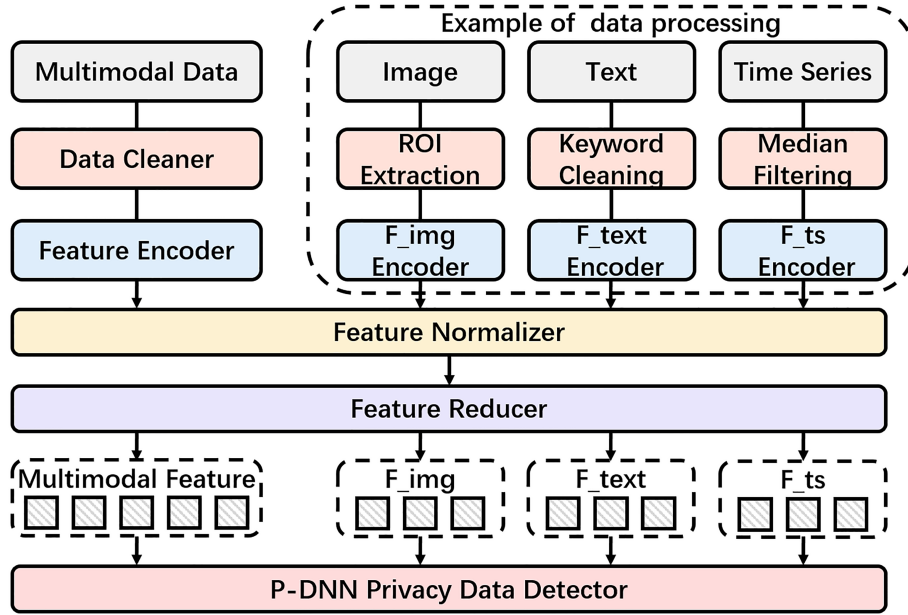


图 2 (网络版彩图) 轻量级多模态数据编码器架构.

Figure 2 (Color online) Illustration of the lightweight multimodal encoder.

统一为小写并去除特殊符号(如“,”“.”). 然后构建关键词向量 F_{text} , 将清洗后的日志转换为二进制向量(存在关键词为 1, 否则为 0) 作为输出, 向量维度为 $M = 200$, 对应每个关键词的存在性标记,

$$F_{\text{text}}[i] = \begin{cases} 1, & \text{若 } k_i \in T, \\ 0, & \text{其他.} \end{cases} \quad (1)$$

(2) 图像数据处理. 物联网图像数据(如摄像头帧为 32×32 灰度图(1024 维)), 含大量背景噪声(如设备外壳、静态环境), 隐私敏感区域(如车牌、人脸) 占比较小. 针对物联网图像数据, 以摄像头帧为例, 定义原始灰度图像 $I \in \mathbb{R}^{32 \times 32}$, 首先对每个像素点 (i, j) , 取其 3×3 邻域内的所有像素值:

$$\mathcal{N}(i, j) = \{(i', j') \mid i' \in [i-1, i+1], j' \in [j-1, j+1]\}. \quad (2)$$

计算中值作为输出值, 去除噪声并保留边缘特征:

$$\tilde{I}(i, j) = \text{median} \{I(i', j') \mid (i', j') \in \mathcal{N}(i, j)\}. \quad (3)$$

然后基于边缘检测算法定位图像中的高梯度区域 (边缘密集区), 计算梯度幅值 $G(i, j)$ 和方向 $\theta(i, j)$ 分别为

$$G(i, j) = \sqrt{G_x(i, j)^2 + G_y(i, j)^2}, \quad \theta(i, j) = \arctan \left(\frac{G_y(i, j)}{G_x(i, j)} \right), \quad (4)$$

其中 G_x, G_y 为 Sobel 算子在 x 和 y 方向的梯度响应. 然后进行非极大值抑制和双阈值滞后处理, 生成二值边缘图 $E(i, j)$. 提取边缘图中所有连通区域的最小外接矩形 $I_{\text{ROI}} = \tilde{I}(y_1 : y_2, x_1 : x_2)$, 坐标为 (x_1, y_1, x_2, y_2) , 其中 $x_1 < x_2, y_1 < y_2$ 提取最小外接矩形作为 ROI, 裁剪为 24×24 像素的子图, 聚焦隐私敏感区域, 若 ROI 尺寸不足 24×24 , 则通过镜像填充至目标尺寸. 将 24×24 的 ROI 子图展平为 576 维向量, 通过主成分分析 (PCA) 降维至 50 维, 保留 95% 的方差, 去除背景冗余. 最后对降维后的 50 维特征向量进行标准化, 确保各维度均值为 0, 标准差为 1, 提升训练稳定性, 具体表示为

$$\begin{aligned} X &= \text{flatten}(I_{\text{ROI}}), \\ X_{\text{centered}} &= X - \mu, \quad \mu = \frac{1}{N} \sum_{k=1}^N X_k, \\ \Sigma &= \frac{1}{N} X_{\text{centered}}^T X_{\text{centered}}, \\ F_{\text{image}} &= X_{\text{centered}} \cdot U_k, \quad U_k = [u_1, u_2, \dots, u_{50}], \\ F_{\text{image}}^{\text{norm}} &= \frac{F_{\text{image}} - \mu_{\text{feat}}}{\sigma_{\text{feat}} + \epsilon}, \quad \epsilon = 10^{-8}. \end{aligned} \quad (5)$$

(3) 时序数据处理. 物联网传感器时序数据 (如温湿度、加速度计读数) 具有时间依赖性, 含高频噪声与偶发异常值, 原始维度通常为 100~500 维 (采样频率 10 Hz×50 s), 针对传感器时序数据, 定义原始时序序列 $X = \{x_1, x_2, \dots, x_N\}$, 长度为 N . 首先采用动态中值滤波, 窗口大小为 $k = \lfloor s \cdot \sigma \rfloor$, 其中 σ 为序列标准差, $s = 2$ 为噪声敏感度系数. 对每个时间点 i , 取窗口 $[i - k, i + k]$ 内的中值作为输出: $\tilde{x}_i = \text{median}(x_{i-k}, x_{i-k+1}, \dots, x_{i+k})$ 边界点采用镜像填充处理 (如 $i - k < 1$ 时, 窗口为 $[1, 2k + 1]$). 然后计算序列的四分位距 (interquartile range, IQR) 范围 $\text{IQR} = Q_3 - Q_1$, 其中 $Q_1 = \text{quantile}(X, 0.25)$, $Q_3 = \text{quantile}(X, 0.75)$. 定义异常边界为 $[Q_1 - 1.5\text{IQR}, Q_3 + 1.5\text{IQR}]$. 超出边界的异常值用相邻有效值的均值替换:

$$\hat{x}_i = \begin{cases} \text{LowerBound}, & \tilde{x}_i < Q_1 - 1.5 \cdot \text{IQR}, \\ \tilde{x}_i, & Q_1 - 1.5 \cdot \text{IQR} \leq \tilde{x}_i \leq Q_3 + 1.5 \cdot \text{IQR}, \\ \text{UpperBound}, & \tilde{x}_i > Q_3 + 1.5 \cdot \text{IQR}. \end{cases} \quad (6)$$

将时序序列划分为 5 个等长片段, 对每段独立进行线性归一化, 避免全局极值对局部特征的影响. 序列长度为 N , 每段长度 $L = \lceil N/5 \rceil$, 第 m 段为 $X^{(m)} = \{x_{(m-1)L+1}, \dots, x_{mL}\}$ ($m = 1, 2, 3, 4, 5$), 则分段归一化可具体表示为

$$x_{\text{norm}}^{(m)} = \frac{x^{(m)} - \min(X^{(m)})}{\max(X^{(m)}) - \min(X^{(m)}) + 10^{-8}}, \quad (7)$$

其中 $\min(X^{(m)})$ 和 $\max(X^{(m)})$ 为第 m 段的极值. 从去噪后的序列中提取 24 维统计特征, 包括均值 $\mu^{(m)}$ 、方差 $\sigma^{(m)2}$ 、最大值 $\max^{(m)}$ 、最小值 $\min^{(m)}$ 、峰度 $\text{kurtosis}^{(m)}$ 、偏度 $\text{skewness}^{(m)}$ 和相邻段均值差 $\Delta\mu^{(m)}$, 最终特征向量表示为

$$F_{\text{time}} = [\mu^{(1)}, \sigma^{(1)}, \max^{(1)}, \min^{(1)}, \text{kurtosis}^{(1)}, \text{skewness}^{(1)}, \dots, \Delta\mu^{(4)}]. \quad (8)$$

维度精简模块通过“过滤式筛选 + PCA 降维”组合策略压缩特征空间. 首先利用方差阈值法和互信息法去除无效特征, 再通过 PCA 保留累计方差贡献度 $\geq 95\%$ 的主成分. 针对边缘设备算力差异, MCU (microprogrammed control unit) 级设备直接使用云端预训练的 PCA 矩阵, Edge GPU 设备通过随机奇异值分解 (single value decomposition, SVD) 加速降维, 确保维度压缩率. 各模态经专用处理后, 输出统一维度的特征向量, 通过“时间戳对齐”与“设备 ID 绑定”实现多模态数据同步, 最终输入 PDNN 隐私检测引擎的模态专用分支.

3.2 PDNN 隐私检测引擎

PDNN 隐私检测引擎聚焦于多模态数据并行处理、动态隐私风险判别与边缘算力适配, 实现在资源受限的边缘设备上高效、鲁棒的隐私泄露检测. 检测引擎总体结构分为 3 层. (1) 模态专用分支层: 针对时序、文本、图像 3 类异构数据, 设计差异化的特征提取网络, 分别捕捉传感器信号的时间依赖、日志文本的语义关联及摄像头帧的视觉轮廓. (2) 特征增强层: 通过模态内残差增强保留各模态核心信息, 结合模态间门控融合, 动态分配跨模态权重, 实现信息互补. (3) 融合决策层: 从增强特征中提取细粒度风险模式, 输出隐私泄露概率, 并通过动态阈值判定风险等级. PDNN-PDE 可形式化表示为多模态特征提取与融合检测的级联过程, 具体如下所示:

$$\text{PDNN-PDE}(X) = \text{Fusion}(\text{Branch}_{ts}(X_{ts}), \text{Branch}_{txt}(X_{txt}), \text{Branch}_{img}(X_{img})), \quad (9)$$

其中, 输入 $X = \{X_{ts}, X_{txt}, X_{img}\}$ 分别表示预处理后的时序统计特征、文本关键词向量和图像降维特征; $\text{Branch}_{ts}, \text{Branch}_{txt}, \text{Branch}_{img}$ 为模态专用分支, 分别提取时序、文本、图像的专属特征; $\text{Fusion}(\cdot)$ 为融合决策层, 通过跨模态特征增强与多粒度风险判断输出隐私泄露概率 $P(\text{risk}) \in [0, 1]$.

各模态分支针对数据特性设计差异化网络结构, 提取模态专属特征. 时序数据的核心挑战是捕捉短期波动与长期依赖, 其分支结构定义为 $\text{Branch}_{ts}(X_{ts}) = \text{GAP}(\text{BiLSTM}(\text{Conv1D}(X_{ts})))$, 具体构建步骤如下.

- 1D 卷积层: 通过局部感受野 ($\text{kernel} = 3$) 捕捉短期时序模式, 输出特征 H_{conv} , 其中 σ 为 ReLU 激活函数, $*$ 表示卷积操作:

$$H_{\text{conv}} = \sigma(W_{\text{conv}} * X_{ts} + b_{\text{conv}}), \quad W_{\text{conv}} \in \mathbb{R}^{3 \times 24 \times 32}. \quad (10)$$

- 双向 LSTM 层: 通过正向与反向序列处理提取长期依赖, 输出序列 H_{lstm} , 其中 T 为序列长度, $\overleftarrow{H}_{\text{lstm}}$ (反向) 与 $\overrightarrow{H}_{\text{lstm}}$ (正向) 通过 LSTM 单元的状态转移计算:

$$H_{\text{lstm}} = [\overleftarrow{H}_{\text{lstm}}; \overrightarrow{H}_{\text{lstm}}], \quad \overleftarrow{H}_{\text{lstm}}, \overrightarrow{H}_{\text{lstm}} \in \mathbb{R}^{T \times 16}. \quad (11)$$

- 全局平均池化 (GAP): 将序列压缩为固定长度特征 $F_{ts} \in \mathbb{R}^{32}$:

$$F_{ts} = \frac{1}{T} \sum_{t=1}^T H_{\text{lstm}}(t). \quad (12)$$

文本数据的核心目标是聚焦隐私敏感关键词的语义关联, 其分支结构定义为 $\text{Branch}_{txt}(X_{txt}) = \text{Attention}(\text{GRU}(\text{Embedding}(X_{txt})))$, 具体构建步骤如下.

- 嵌入层: 将稀疏二进制关键词向量 (200 维) 转换为密集语义表示 $H_{\text{embed}} \in \mathbb{R}^{T \times 32}$:

$$H_{\text{embed}} = X_{txt} \cdot W_{\text{embed}}, \quad W_{\text{embed}} \in \mathbb{R}^{200 \times 32}. \quad (13)$$

- GRU 层: 通过门控机制捕捉文本上下文关联, 输出隐状态序列 H_{gru} . GRU 单元的状态转移方

程如式 (14) 所示, 其中 z_t (更新门) 和 r_t (重置门) 控制信息传递, h_t 为当前隐状态,

$$\begin{aligned} z_t &= \sigma(W_z \cdot h_{t-1} + U_z \cdot x_t), \\ r_t &= \sigma(W_r \cdot h_{t-1} + U_r \cdot x_t), \\ \tilde{h}_t &= \tanh(W \cdot (r_t \odot h_{t-1}) + U \cdot x_t), \\ h_t &= (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t. \end{aligned} \quad (14)$$

• 注意力层: 通过可学习权重聚焦敏感关键词, 输出上下文向量 $c \in \mathbb{R}^{32}$, 其中 $W_a \in \mathbb{R}^{32 \times 32}$ 为注意力权重矩阵, α_t 为第 t 个词的注意力系数,

$$\alpha_t = \frac{\exp(h_t^T W_a s_{t-1})}{\sum_{k=1}^T \exp(h_k^T W_a s_{t-1})}, \quad c = \sum_{t=1}^T \alpha_t h_t. \quad (15)$$

图像数据的核心需求是在低算力下保留关键视觉特征. 其分支结构定义为 $\text{Branch}_{img}(X_{img}) = \text{FC}(\text{MaxPool}(\text{DWConv}(X_{img})))$, 具体构建步骤如下.

• 深度可分离卷积 (DWConv): 将标准卷积分解为逐通道卷积 (depthwise) 和逐点卷积 (pointwise), 进一步降低参数量. 输出特征 H_{dw} 计算为

$$H_{dw} = \text{ReLU}(\text{DepthwiseConv}(X_{img}) * W_{pw} + b_{pw}), \quad W_{pw} \in \mathbb{R}^{1 \times 1 \times 16 \times 32}. \quad (16)$$

- 最大池化层: 通过 2×2 窗口降低空间分辨率, 保留边缘特征 (如设备 ID 轮廓), 输出 H_{pool} .
- 全连接层: 将池化特征展平后映射为 32 维特征 F_{img} :

$$F_{img} = \text{ReLU}(W_{fc} \cdot \text{flatten}(H_{pool}) + b_{fc}), \quad W_{fc} \in \mathbb{R}^{72 \times 32}. \quad (17)$$

特征增强与融合决策层接入各模态分支输出特征, 通过双路径增强模块提取跨模态关联, 最终由融合决策层输出风险概率.

特征增强模块包含模态内增强 (intra-modal) 与模态间增强 (inter-modal). 模态内增强采用残差块缓解梯度消失, 保留模态核心信息. 增强特征 F_{intra} 计算如式 (18) 所示, 其中 $F \in \{F_{ts}, F_{txt}, F_{img}\}$, $W_1, W_2 \in \mathbb{R}^{32 \times 32}$ 为全连接层权重,

$$F_{intra} = F + \text{ReLU}(W_2 \cdot \text{ReLU}(W_1 \cdot F + b_1) + b_2). \quad (18)$$

模态间增强通过门控融合机制动态分配模态权重. 门控权重 $\alpha \in \mathbb{R}^3$ 和跨模态特征 F_{inter} 计算如公式 (19), 其中 σ 为 Sigmoid 函数, $W_\alpha \in \mathbb{R}^{3 \times 96}$ 为门控权重矩阵,

$$\alpha = \sigma(W_\alpha \cdot [F_{ts}^{intra}, F_{txt}^{intra}, F_{img}^{intra}] + b_\alpha), \quad F_{inter} = \sum_{i=1}^3 \alpha_i \cdot F_i^{intra}. \quad (19)$$

融合决策层通过多粒度特征提取与风险评分输出最终概率 $P(\text{risk})$. 首先进行多粒度特征提取, 通过 1×1 卷积 ($32 \rightarrow 16$ 维) 和全连接层 ($16 \rightarrow 8$ 维) 提取粗粒度与细粒度特征, 然后进行风险评分与决策, 通过 Sigmoid 激活输出概率, 并基于动态阈值 θ 判定结果:

$$P(\text{risk}) = \sigma(W_{out} \cdot H_{fine} + b_{out}), \quad \text{prediction} = \begin{cases} \text{leakage}, & P(\text{risk}) \geq \theta, \\ \text{non-leakage}, & P(\text{risk}) < \theta. \end{cases} \quad (20)$$

3.3 云-边协同动态增量学习模型更新

针对物联网边缘设备在隐私检测中面临的动态场景适配与资源受限问题,本文设计了基于边缘智能筛选、云端知识融合与轻量化模型更新的云-边协同动态增量学习模型更新机制,确保模型在保留旧知识的同时高效适应新隐私模式,其核心算法流程如下.

(1) 初始化配置与边缘设备部署. 系统启动时,在旧任务 $\mathcal{T}_{\text{old}}$ 上预训练的初始模型 $\mathcal{M}_0$ 被部署至边缘设备集合 $\mathcal{D} = \{d_1, d_2, \dots, d_m\}$, 各设备初始化样本缓冲区 $B_i = \emptyset$ ($i = 1, 2, \dots, m$). 边缘筛选器 $\mathcal{F}$ 定义为仅保留置信度极端的样本: 对于样本 x_t , 若模型预测置信度 $p_t = \text{softmax}(\hat{y}_t)_{\max}$ 高于 $\theta_{\text{high}} = 0.95$ (高确定性已知模式) 或低于 $\theta_{\text{low}} = 0.1$ (低置信潜在新威胁), 则将其以 $(x_t, \hat{y}_t)$ 形式 (无标签样本使用模型预测标签) 暂存于本地缓冲区, 为后续筛选上传作准备.

(2) 边缘设备的样本筛选与批量上传. 在新样本流 $\mathcal{S} = \{s_t\}_{t=1}^T$ 的实时处理中, 每个边缘设备 d_i 执行以下操作: 首先通过本地模型 $\mathcal{M}_0$ 推理生成预测 $\hat{y}_t$, 经温度缩放校准后计算置信度 p_t ; 若满足筛选条件 ($p_t > \theta_{\text{high}}$ 或 $p_t < \theta_{\text{low}}$), 则将样本存入缓冲区 B_i . 当缓冲区容量达到上传阈值 $K = 100$ 时, 设备通过 AES-256 加密将数据批量传输至云端, 并清空缓冲区. 该策略有效过滤了无效样本, 使云端接收的样本中高价值数据占比提升, 显著降低了边缘-云端的通信负载.

(3) 云端数据处理与模型微调. 云端接收所有设备上传的样本集 $\mathcal{B} = \bigcup_{i=1}^m B_i$ 后, 首先过滤低置信度且标签与旧模型预测冲突的样本, 保留一致性高的训练数据. 随后初始化微调模型 $\mathcal{M}_{\text{ft}} \leftarrow \mathcal{M}_0$, 冻结特征提取层以保留旧任务的底层特征表示, 仅开放分类器层进行训练. 微调过程采用联合损失函数, 具体如公式 (21) 所示. 其中, 二元交叉熵 BCE (binary cross entropy) 损失拟合新样本标签, 知识蒸馏损失 KL 通过约束新旧模型的特征分布, 避免“灾难性遗忘”, 在高效融入新样本信息的同时保护了旧任务知识,

$$\mathcal{L} = 0.7 \cdot \text{BCE}(\hat{y}, y) + 0.3 \cdot \text{KL}(\mathcal{M}_0(x) \parallel \mathcal{M}_{\text{ft}}(x)). \quad (21)$$

(4) 轻量化模型压缩与边缘部署优化. 微调完成后, 云端通过动态量化技术将模型压缩为 INT8 格式 ($\mathcal{M}_{\text{quant}}$), 通过公式 $W_{\text{int8}} = \text{round}((W_{\text{FP32}} - Z)/S)$ (Z 为零点偏移, S 为缩放因子) 将模型体积进一步缩小. 进一步计算增量参数 $\Delta = \mathcal{M}_{\text{quant}} - \mathcal{M}_0$, 仅传输参数差异, 有效降低带宽需求. 边缘设备接收 Δ 后, 通过 $\mathcal{M}_i \leftarrow \mathcal{M}_0 + \Delta$ 更新本地模型, 并验证旧任务精度: 若衰减超过 2%, 则触发回滚机制, 确保模型在新旧任务上的稳定性.

4 实验评估

4.1 实验设置

本研究选取各类代表性方法进行对比, 包括传统规则匹配 (基于关键词和正则表达式的匹配系统 [26])、传统机器学习方法 (KNN, SVM)、经典深度学习模型 (CNN, RNN, LSTM, GRU) [27]、目标检测模型 (YOLO, Faster R-CNN)、轻量级神经网络 (MobileNet V2, ShuffleNet V2, Tiny-LSTM, Tiny-TCN)、自然语言处理模型 (Bi-LSTM-CRF, RoBERTa, DistilBERT)、时序预测模型 (Informer) 和多模态融合模型 (ViLBERT, MMFFHS).

实验硬件涵盖边缘与云端场景. 云端使用高性能服务器: 配备 4 块 NVIDIA RTX 3090 GPU, CPU 为 Intel Xeon E5-2698 v4 (20 核 40 线程, 基频 2.3 GHz, 睿频 3.6 GHz), 内存 128 GB DDR4 ECC RAM, 存储为 2TB NVMe SSD. 边缘侧测试设备包括: 工业 PC, 搭载 Intel Core i7-9700 (8 核 8 线程, 基频 3.0 GHz, 睿频 4.7 GHz), 16GB DDR4 RAM, 256GB SATA SSD; Raspberry Pi 4, 采用 Broadcom BCM2711 四核 ARM Cortex-A72 处理器 (主频 1.5 GHz), 4GB LPDDR4 内存, 32GB Class 10 MicroSD 卡; Jetson Nano, 集成 NVIDIA Maxwell 架构 GPU (128 个 CUDA 核心), 四核 ARM Cortex-A57 (主

频 1.43 GHz), 4GB LPDDR4 内存, 16GB eMMC 存储. 所有设备均运行 Ubuntu 20.04 系统, CUDA 版本 11.6, PyTorch 1.12.

数据集方面, 本研究通过部署轻量级监控代理, 真实采集了 73 家不同厂商 (覆盖智能终端、工业传感器、网关设备等) 的设备运行日志与业务网络流量数据, 共获取 85481 条原始样本. 原始数据长度差异显著, 最长记录为 32693 字符, 最短仅为 73 字符, 随后对数据进行预处理 (去噪, 如移除时间戳; 标准化, 如统一编码), 并严格依据《中华人民共和国网络安全法》《中华人民共和国个人信息保护法》及行业规范, 将每条记录标注为“隐私”或“非隐私”类别; 标注中识别并归类出 8 类典型违规场景. (a) 设备密钥等敏感信息泄露; (b) 弱口令风险; (c) 敏感关键词; (d) 证件护照数据; (e) 私有网络链路及内部邮箱; (f) 人权歧视言论; (g) 涉政内容; (h) 其他违规 (隐私泄露样本占比约 40%), 形成涵盖用户指令、设备异常、访问日志、数据上传等边缘场景的文本隐私数据集. 同时引入图像隐私检测数据集 PrivacyAlert^[11], 包含 6800 张在 2011~2021 年间发布的网络图像, 图像标注了“私人”或“公共”类别, 用以模拟物联网隐私图像数据. 同时使用 IoT 多模态基准数据集 MultiIoT 的时序数据部分, 涵盖 12 种传感器模态、8 个任务, 总样本量超过 115 万, 用于模拟真实物联网环境下的连续传感数据. 所有模态数据均经过归一化、去噪及增强处理, 以保证评估的公平可靠.

模型配置方面, PDNN-PDE 包括三路模态专用分支 (时序: $24 \rightarrow 32 \rightarrow 32$; 文本: $200 \rightarrow 64 \rightarrow 32$; 图像: $1 \rightarrow 8 \rightarrow 32$ 全连接网络), 融合层采用 4 头 32 维多头自注意力, 分类器为 $32 \rightarrow 64 \rightarrow 1$ 两层全连接网络, Dropout = 0.3, 优化器为 AdamW (学习率 1×10^{-4} , 权重衰减 1×10^{-4}), 训练轮次 80 轮, 云端批量大小 128, 边缘批量大小 32, 损失函数为正样本权重 2.0 的二元交叉熵.

实验评估指标包括检测性能指标和系统资源指标: 检测性能评估指标包括准确率 (accuracy)、误报率 (false positive rate)、漏报率 (false negative rate) 和 F1 分数等, 用于衡量检测模型的整体检测质量. 资源开销评估指标包括单样本推理延迟 (single sample inference latency, SSIL)、内存占用 (memory footprint, MF) 和模型大小 (model size, MS) 等, 用于评估模型在不同设备上的实用性和部署成本. 并行机制效率评估指标包括并行处理带来的吞吐量提升和资源利用率. 模型更新能力评估指标包括新模式检测率、旧模式保持率 (避免灾难性遗忘)、模型收敛速度和更新频率.

4.2 与当前隐私数据识别方法检测能力对比

本研究在所有数据集上分别训练并测试上述对比方法与所提 PDNN-PDE 模型的隐私数据检测能力. 所有方法均使用相同预处理和训练集/测试集划分, 并采用交叉验证以保证评估结果的可靠性. 对比实验关注各方法在单一模态和多模态下的检测精度.

实验结果如表 1 所示, PDNN-PDE 在所有评估模态的隐私数据检测中均展现出具有一定竞争力的检测性能. 在文本模态中, PDNN-PDE 的 F1 分数为 92.5%, 略高于预训练语言模型, 如 RoBERTa (91.9%) 和 DistilBERT (90.8%), 同时保持了更低的误报率 (FPR 4.0%) 和漏报率 (FNR 5.0%). 这一优势表明, PDNN-PDE 的轻量级编码器和并行架构结合自研的物联网日志数据集特性, 能够提供更精准高效的识别能力. 对于图像 (PrivacyAlert)、时序 (MultiIoT) 和融合 (跨模态) 模态, PDNN-PDE 的性能 (F1 分数分别为 88.1%, 89.2% 和 89.8%) 表现出较好的竞争力, 略低于或非常接近各自领域较为先进的方法 (Faster R-CNN, Informer, ViLBERT). 在所有模态中, PDNN-PDE 检测性能较单模态轻量模型平均提升 3.6% (文本 0.6%、图像 2.5%、时序 4.7%), 且接近专用模型性能 (差值 $\leq 0.8\%$). 同时 PDNN-PDE 在所有模态中均表现出较低的 FPR 和 FNR 值, 证明 PDNN-PDE 能够较好地减少误报以及更精准地捕捉实际隐私泄露问题.

4.3 物联网边缘侧设备计算开销与并行机制效率分析

为了量化 PDNN-PDE 在不同边缘计算环境下的部署成本, 本研究设计了资源开销评估实验. 实验选择了 3 种典型硬件平台: 计算能力强大的工业 PC、具备中等计算能力的 GPU 加速平台 Jetson

表 1 PDNN-PDE 与各类方法在不同模态下的检测性能对比.

Table 1 Comparison of detection performance of PDNN-PDE and various methods under different modalities.

Modality	Method	Accuracy (%)	FPR (%)	FNR (%)	F1-score (%)
Image	KNN	71.8	21.2	26.5	69.1
	SVM	75.3	19.5	23.8	74.2
	CNN	83.7	11.5	13.2	82.9
	YOLO	86.9	9.8	11.1	86.4
	Faster R-CNN	89.1	8.9	10.1	88.7
	MobileNet V2	86.2	10.4	12.0	85.6
	ShuffleNet V2	85.5	10.9	12.6	84.8
	PDNN-PDE	88.5	9.3	10.5	88.1
Text	Rule match	63.5	7.9	37.1	58.7
	KNN	77.2	16.3	21.9	75.5
	SVM	81.1	13.8	16.5	80.4
	Bi-LSTM-CRF	89.5	5.8	7.9	88.9
	RoBERTa	92.4	4.5	5.5	91.9
	DistilBERT	91.3	5.1	6.3	90.8
	PDNN-PDE	92.9	4.0	5.0	92.5
Time-series	SVM	73.9	19.8	24.5	71.9
	RNN	83.2	11.3	15.2	82.3
	LSTM	87.1	7.6	10.8	86.6
	GRU	86.7	8.1	11.3	86.2
	Tiny-LSTM	85.4	8.8	12.9	84.5
	Tiny-TCN	85.9	8.5	12.4	85.0
	Informer	90.5	5.4	7.6	90.0
	PDNN-PDE	89.7	6.0	8.0	89.2
Fusion	KNN	74.3	19.1	26.1	71.5
	SVM	79.5	15.8	20.9	77.8
	ViLBERT	91.0	5.0	6.5	90.5
	MMFFHS	90.0	5.7	7.5	89.5
	PDNN-PDE	90.3	5.5	7.0	89.8

Nano 以及资源极端受限的 Raspberry Pi 4. 在每个平台上, 本研究针对图像、文本、时序和融合 4 种模态的隐私数据检测任务, 分别测量了 PDNN-PDE 与当前主流轻量级模型与多模态模型 (如 MobileNet V2, ShuffleNet V2, DistilBERT, Tiny-LSTM, Tiny-TCN, ViLBERT 和 MMFFHS) 的单样本推理延迟 (SSIL)、模型存储 (MS) 和模型推理内存占用 (MF).

实验结果如表 2 所示, PDNN-PDE 在资源开销方面表现较好. 在所有测试模态和设备上, PDNN-PDE 均实现了最低的单样本推理延迟 (SSIL), 相较于对比模型, PDNN-PDE 的推理速度更快, 如在工业 PC 上处理图像模态时, PDNN-PDE 的 SSIL 仅为 14.8 ms, 优于 MobileNet V2 的 18.5 ms 和 ShuffleNet V2 的 16.2 ms. 同时, PDNN-PDE 的模型推理内存占用 (MF) 也显著低于其他对比模型, 这表明 PDNN-PDE 在执行推理时对内存的需求更小, 这对于内存受限的边缘设备至关重要. 虽然 PDNN-PDE 的模型存储 (MS) 在某些情况下可能略高于某些极度压缩的模型, 但其在推理延迟和内存占用上的优化证明了其轻量化设计和高效运行的特点.

为了深入探究 PDNN-PDE 的并行处理机制对其性能的提升作用, 本研究进一步开展了并行机制

表 2 PDNN-PDE 与轻量级模型在各模态不同边缘设备上的资源开销对比.

Table 2 Comparison of resource overhead between PDNN-PDE and lightweight models on edge devices with different modalities.

Modality	Method	Industrial PC			Jetson Nano			Raspberry Pi 4		
		SSIL (ms)	MS (MB)	MF (MB)	SSIL (ms)	MS (MB)	MF (MB)	SSIL (ms)	MS (MB)	MF (MB)
Image	MobileNet V2	18.5	13.1	15.2	22.5	13.1	18.0	115.3	13.1	25.5
	ShuffleNet V2	16.2	11.2	12.8	19.5	11.2	15.5	98.2	11.2	22.0
	PDNN-PDE	14.8	12.5	11.5	16.5	12.5	13.5	85.1	12.5	20.0
Text	DistilBERT	28.1	250.3	60.5	35.2	250.3	70.3	180.5	250.3	95.2
	PDNN-PDE	22.5	80.2	35.1	28.1	80.2	40.1	140.3	80.2	50.1
Time-series	Tiny-LSTM	12.1	8.1	10.5	18.1	8.1	12.0	75.1	8.1	18.0
	Tiny-TCN	11.5	7.6	9.8	17.1	7.6	11.0	70.2	7.6	16.5
	PDNN-PDE	9.0	7.8	8.0	13.1	7.8	9.5	55.1	7.8	14.0
Fusion	ViLBERT	120.1	350.3	800.5	40.1	350.3	900.5	700.1	350.3	1200.5
	MMFFHS	105.2	280.1	650.3	35.1	280.1	750.3	600.2	280.1	900.3
	PDNN-PDE	45.1	100.2	150.1	20.1	100.2	180.1	250.1	100.2	250.1

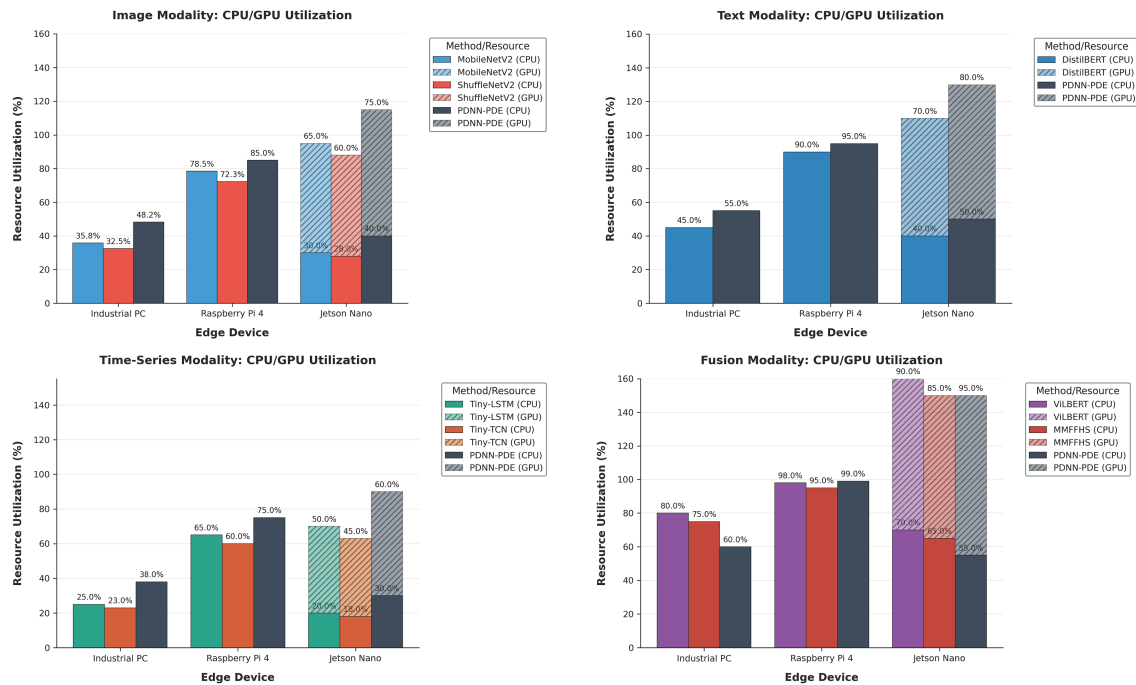


图 3 (网络版彩图) PDNN-PDE 资源利用率对比评估.

Figure 3 (Color online) Comparative evaluation of PDNN-PDE resource utilization.

效率对比实验. 此实验同样在工业 PC, Jetson Nano 和 Raspberry Pi 4 上进行, 并覆盖上述 4 种数据模态. 本研究重点关注两个指标: 吞吐量 (throughput) 和 CPU/GPU 利用率 (resource utilization). 吞吐量直接反映了模型在单位时间内处理样本的能力, 是衡量并行效率的关键指标; 而 CPU/GPU 利用率则揭示了模型对硬件资源的实际利用程度. 通过分析 PDNN-PDE 在高负载下的吞吐量表现及其对多核 CPU 和 Jetson Nano GPU 的利用情况, 评估 PDNN-PDE 所采用的并行深度学习架构的计算资源协调有效性.

PDNN-PDE 并行机制效率对比实验结果分别如图 3 和 4 所示. 在吞吐量方面, PDNN-PDE 在所

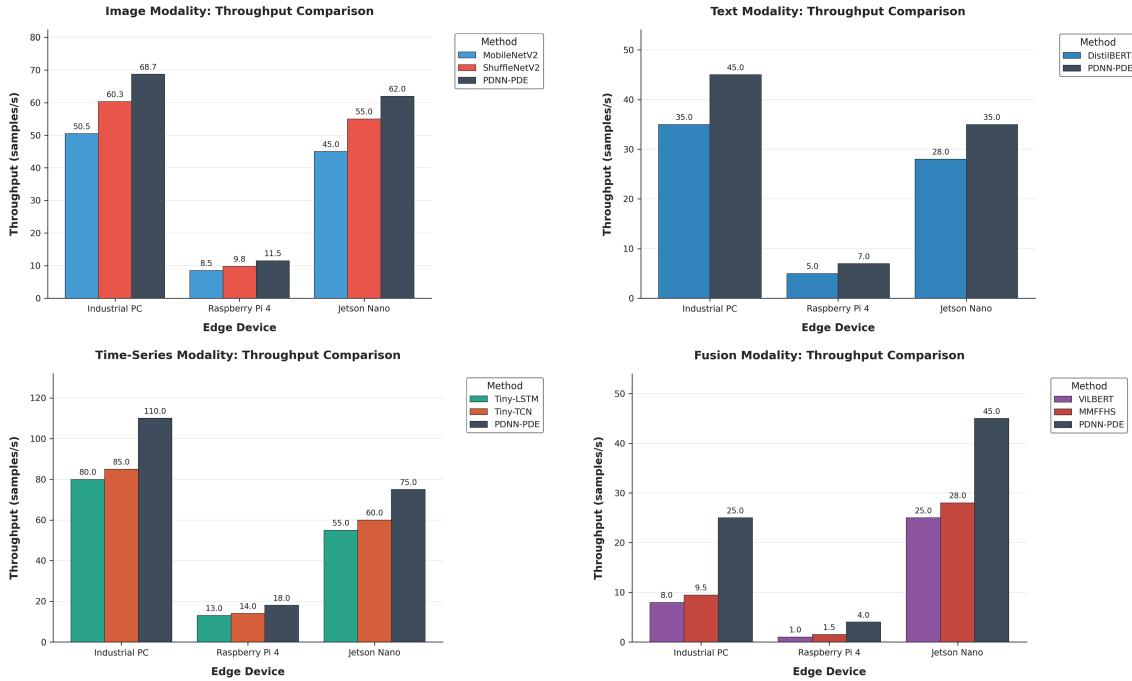


图 4 (网络版彩图) PDNN-PDE 吞吐量对比评估.

Figure 4 (Color online) Comparison and evaluation of PDNN-PDE throughput.

有模态和 3 种边缘设备上都实现了最高的吞吐量, 这证明了其并行架构能够有效提升数据处理能力. 例如, 在工业 PC 处理时序数据时, PDNN-PDE 的吞吐量高达 110.0 samples/s, 远超 Tiny-LSTM 的 80.0 samples/s 和 Tiny-TCN 的 85.0 samples/s.

在资源利用率方面, PDNN-PDE 在工业 PC 和 Raspberry Pi 4 上展现出对 CPU 资源的高效利用, 而在具备 GPU 的 Jetson Nano 上, PDNN-PDE 更是能充分利用 GPU 的并行计算能力, 实现了 CPU 和 GPU 的协同高效工作. 例如, 在融合模态下, PDNN-PDE 在 Jetson Nano 上的 GPU 利用率达到了 95.0%. 这表明 PDNN-PDE 的并行设计不仅提升了处理速度, 还能够有效发挥边缘设备的异构计算潜力, 确保在严苛的物联网环境中也能进行高效、实时的隐私数据检测.

4.4 云-边协同动态增量学习模型更新能力评估

为模拟新型隐私泄露模式出现后的场景, 本研究在原训练模型基础上引入新的隐私样本 (模拟概念漂移), 观察增量学习的准确率恢复情况. 具体做法是, 先在原始数据上训练 PDNN-PDE 模型; 然后在测试阶段动态加入新的隐私数据, 并使用云-边协同的增量学习模块进行模型微调, 每个迭代轮次后评估检测准确率. 实验结果如图 5 所示, 初始模型在新模式样本上的检测准确率骤降 (由约 85% 降至 45%), 但随着增量训练的进行, 准确率迅速恢复. 在约 70 轮增量训练后, 检测准确率重回原水平, 达到约 89%. 增量更新前后的精度对比说明所提动态学习机制能有效帮助模型适应新型隐私数据模式, 恢复检测性能.

5 结论与展望

本研究旨在解决物联网边缘侧设备资源受限、新型隐私泄露模式泛化检测能力不足以及检测精度与效率难以平衡的三大挑战. 为此, 本文提出了一种基于并行深度神经网络的物联网边缘侧隐私数据检测系统 (PDNN-PDE), 该系统包含轻量级多模态数据编码器、PDNN 隐私检测引擎和云-边协同动态增量学习模型更新模块. 实验结果表明, PDNN-PDE 在文本、图像和时间序列等多种模态上均表现

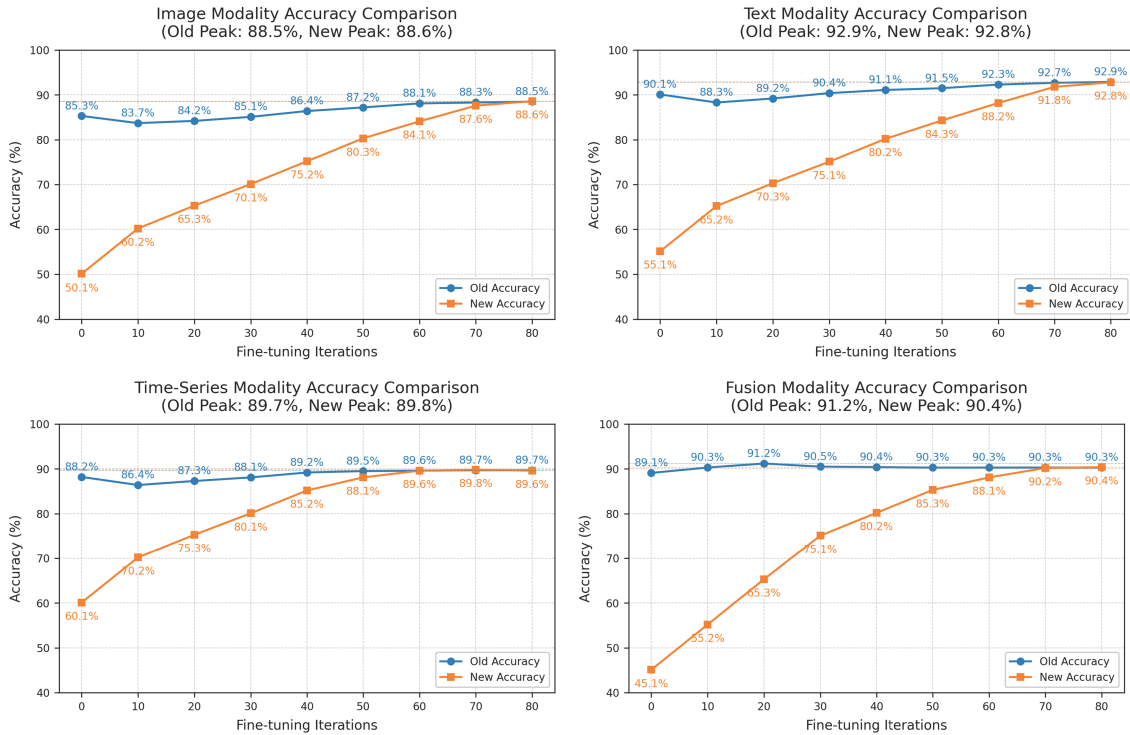


图 5 (网络版彩图) PDNN-PDE 动态增量学习模型检测精度更新评估。

Figure 5 (Color online) Update evaluation of detection accuracy of PDNN-PDE dynamic incremental learning model.

出具有竞争力的检测性能, F1 分数优于传统方法和经典深度学习架构, 同时保持了较低的误报率和漏报率. 此外, PDNN-PDE 在边缘设备上的资源开销显著降低, 实现了较低的单样本推理延迟和内存占用, 同时通过并行机制显著提升了吞吐量和 CPU/GPU 利用率. 这些结果有力地验证了 PDNN-PDE 能够在资源受限的物联网边缘环境中实现高效、鲁棒隐私数据检测, 并通过云-边协同动态增量学习有效提升了模型对新型隐私泄露模式的泛化检测能力. 未来工作将进一步探索联邦学习与差分隐私等技术, 以增强模型在多设备协作下的隐私保护能力和泛化性能, 从而应对更复杂多变的物联网隐私泄露挑战.

参考文献

- 1 Ji W, Yang Z M, Wang Z, et al. Visual end-edge-cloud fusion architecture: key technologies of future super metropolitan clusters. *Sci Sin Inform*, 2024, 54: 2518–2532 [纪雯, 杨哲铭, 王智, 等. 视觉端边云融合架构: 面向超级智慧城市群演进的关键技术. *中国科学: 信息科学*, 2024, 54: 2518–2532]
- 2 Sheikh A M, Islam M R, Habaebi M H, et al. A survey on edge computing (EC) security challenges: classification, threats, and mitigation strategies. *Future Internet*, 2025, 17: 175
- 3 Albreem M A, Sheikh A M, Bashir M J K, et al. Towards green internet of things (IoT) for a sustainable future in Gulf Cooperation Council countries: current practices, challenges and future prospective. *Wireless Netw*, 2023, 29: 539–567
- 4 Hashmat F, Abbas S G, Hina S, et al. An automated context-aware IoT vulnerability assessment rule-set generator. *Comput Commun*, 2022, 186: 133–152
- 5 Musa A, Kakudi H A, Hassan M, et al. Lightweight deep learning models for edge devices—a survey. *Int J Comput Inform Syst Ind Manag Appl*, 2025, 17: 18
- 6 Wang X, Tang Z, Guo J, et al. Empowering edge intelligence: a comprehensive survey on on-device AI models. *ACM Comput Surv*, 2025, 57: 1–39
- 7 Sandler M, Howard A, Zhu M, et al. Mobilenetv2: inverted residuals and linear bottlenecks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018. 4510–4520

- 8 Ma N, Zhang X, Zheng H T, et al. Shufflenet v2: practical guidelines for efficient CNN architecture design. In: Proceedings of the European Conference on Computer Vision (ECCV), 2018. 116–131
- 9 Dellys H N, Mokeddem H, Sliman L. On the integration of social context for enhanced fake news detection using multimodal fusion attention mechanism. AI, 2025, 6: 78
- 10 Li Y, Zhang J X, Yao Z X, et al. Neighborhood-aware distributed intelligent computing offloading and resource allocation for edge computing. Sci Sin Inform, 2024, 54: 413–429 [李云, 张剑鑫, 姚枝秀, 等. 邻域感知的分布式智能边缘计算卸载和资源分配算法. 中国科学: 信息科学, 2024, 54: 413–429]
- 11 Zhao C, Mangat J, Koujalgi S, et al. PrivacyAlert: a dataset for image privacy prediction. In: Proceedings of the International AAAI Conference on Web and Social Media, 2022. 1352–1361
- 12 Mo S, Morency L P, Salakhutdinov R, et al. MultiIoT: benchmarking machine learning for the internet of things. 2023. ArXiv:2311.06217
- 13 Roman R, Zhou J, Lopez J. On the features and challenges of security and privacy in distributed internet of things. Comput Networks, 2013, 57: 2266–2279
- 14 Weber R H. Internet of things—new security and privacy challenges. Comput Law Secur Rev, 2010, 26: 23–30
- 15 Gao Y, Liu S C, Guo B, et al. Lightweight sensing-computing-decision collaboration enhancement for multi-mobile terminals. Sci Sin Inform, 2024, 54: 2136–2156 [高元, 刘思聪, 郭斌, 等. 多移动终端轻量化感 – 算 – 策协同增强方法. 中国科学: 信息科学, 2024, 54: 2136–2156]
- 16 Aouedi O, Piamrat K. F-BIDS: federated-blending based intrusion detection system. Pervasive Mob Comput, 2023, 89: 101750
- 17 Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell, 2016, 39: 1137–1149
- 18 Zhang Z, Lu X, Cao G, et al. ViT-YOLO: transformer-based YOLO for object detection. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021. 2799–2808
- 19 Alzaidi R, Caragea C, Giles C L. Bi-LSTM-CRF sequence labeling for keyphrase extraction from scholarly documents. In: Proceedings of the World Wide Web Conference, 2019. 2551–2557
- 20 Tan K L, Lee C P, Anbananthen K S M, et al. RoBERTa-LSTM: a hybrid model for sentiment analysis with transformer and recurrent neural network. IEEE Access, 2022, 10: 21517–21525
- 21 Zhang J, Huang K, Lang D, et al. Research on multifeature fusion false review detection based on DynDistilBERT-BiLSTM-CNN. IEEE Internet Things J, 2024, 11: 30040–30053
- 22 Zhou H, Zhang S, Peng J, et al. Informer: beyond efficient transformer for long sequence time-series forecasting. AAAI, 2021, 35: 11106–11115
- 23 Dellys H N, Mokeddem H, Sliman L. On the integration of social context for enhanced fake news detection using multimodal fusion attention mechanism. AI, 2025, 6: 78
- 24 Chang Y, Thompson R, Hixenbaugh C. Tiny long-short term memory model for resource-constrained prediction of battery cycle life. In: International Conference on Machine Learning, Optimization, and Data Science. Cham: Springer, 2024. 131–144
- 25 Boner S, Vogt C, Magno M. Tiny TCN model for gesture recognition with multi-point low power TOF-sensors. In: Proceedings of the IEEE 4th International Conference on Artificial Intelligence Circuits and Systems (AICAS), 2022. 356–359
- 26 Babun L, Celik Z B, McDaniel P, et al. Real-time analysis of privacy-(un)aware IoT applications. Proc Privacy Enhanc Tech, 2021, 2021: 145–166
- 27 Walling S, Lodh S. An extensive review of machine learning and deep learning techniques on network intrusion detection for IoT. T Emerg Telecommun T, 2025, 36: e70064

PDNN-PDE: privacy data detection for the edge IoT based on P-DNN

Shaopeng ZHOU^{1,3}, Bin WANG^{2,3*}, Chaohao LI^{1,3}, Zhengyan ZHOU¹,
Feng ZHANG^{3,4} & Chunming WU¹

1. *College of Computer Science and Technology, Zhejiang University, Hangzhou 310027, China*

2. *Hangzhou Institute of Technology, Xidian University, Hangzhou 311231, China*

3. *Zhejiang Key Laboratory of Artificial Intelligence of Things (AIoT) Network and Data Security, Hangzhou 310000, China*

4. *School of Cyber Engineering, Xidian University, Xi'an 710071, China*

* Corresponding author. E-mail: wangbin02@xidian.edu.cn

Abstract Addressing the three major challenges in IoT privacy detection—limited resources on edge devices, insufficient generalization capabilities for detecting new types of privacy leaks, and the difficulty in balancing accuracy and efficiency—this paper proposes a parallel deep neural network-based IoT edge-side privacy data detection system (PDNN-PDE). The system first uses a lightweight multimodal data encoder to efficiently process heterogeneous data such as images, time-series sensor signals, and log texts. Then, it relies on a parallel deep neural network privacy detection engine to achieve distributed real-time detection of multimodal privacy data on edge devices. At the same time, based on a cloud-edge collaborative dynamic incremental learning algorithm, the model is periodically updated to enhance its generalization detection capabilities for unknown privacy leakage patterns. Experimental results show that PDNN-PDE achieves an average accuracy improvement of 3.6% compared to single-modal lightweight models; when fusing modalities, the detection accuracy decreases by only 0.7%, while memory usage is reduced by 76% and inference latency is reduced by 21.3% compared to existing specialized models. In terms of adaptability to new leakage patterns, after 70 rounds of incremental learning, the model's accuracy recovered to 89%, representing an approximately 44% improvement over static models. These results demonstrate that PDNN-PDE can achieve efficient, robust, and adaptive privacy data detection in resource-constrained IoT edge environments.

Keywords IoT edge computing, privacy data detection, parallel deep learning, multimodal data, incremental learning