基于大语言模型的 HTTP/HTTPS 网络资产设备类型识别方法

陈倩怡[1]，苏马婧[1,2]，陈紫璇[1,2]，张永奇[1,2]，马　琰[1,2]
（1. 华北计算机系统工程研究所，北京　100083；
2. 中国信息安全研究院有限公司，北京　102200）

摘　要：针对传统基于静态指纹规则和判别式模型在复杂开放环境下泛化能力不足的现状，提出了一种基于大语言模型指令微调的 HTTP/HTTPS 网络资产设备类型识别方法。通过多源采集与多平台标签聚合的数据收集方案构造原始网络资产数据集，对该数据集进行关键特征优先保留的数据预处理，有效降低冗余噪声对模型输入的影响，然后通过融合 HTTP/HTTPS 响应体、响应头、SSL 证书、端口及协议等多源异构特征，构建统一的序列化表示；在此基础上，利用 LoRA 技术对 LLaMA-3-8B-Instruct 模型进行参数高效微调，引导模型学习网络资产特征与设备类型之间的语义关联关系。实验结果表明，在包含 38 万条真实网络资产的测试集中，该方法在样本高度不均衡和长尾设备场景下仍能保持稳定性能，Weighted F1-score 达到 0.959 1，相比未微调模型效果显著提升。同时，模型推理吞吐量提高 62.81%，验证了所提方法在大规模网络资产自动识别任务中的有效性与实用性。

关键词：网络资产识别；大语言模型；指令微调；多源异构特征
中图分类号：TP393　　**文献标志码**：A　　**DOI**：10.19358/j.issn.2097－1788.2026.05.001

中文引用格式：陈倩怡，苏马婧，陈紫璇，等. 基于大语言模型的 HTTP/HTTPS 网络资产设备类型识别方法[J]. 网络安全与数据治理，2026，45(5)：1－10.
英文引用格式：Chen Qianyi，Su Majing，Chen Zixuan，et al. Device type identification of HTTP/HTTPS network asset based on large language models[J]. Cyber Security and Data Governance，2026，45(5)：1－10.

Device type identification of HTTP/HTTPS network asset based on large language models

Chen Qianyi[1]，Su Majing[1,2]，Chen Zixuan[1,2]，Zhang Yongqi[1,2]，Ma Yan[1,2]
（1. National Computer System Engineering Research Institute of China，Beijing 100083，China；
2. China Information Security Research Institute Co.，Ltd.，Beijing 102200，China）

Abstract：To address the limited generalization ability of traditional network asset identification methods based on static fingerprint rules and discriminative models in complex and open environments，this paper proposes an HTTP/HTTPS network asset device type identification method based on instruction fine-tuning of a large language model. A multi-source data collection scheme with multi-platform label aggregation is designed to construct the original network asset dataset. A data preprocessing strategy that prioritizes key feature retention is applied to reduce redundant noise in model inputs. Multiple heterogeneous features，including HTTP/HTTPS response bodies，response headers，SSL certificates，ports，and protocols，are further integrated to construct a unified serialized representation. Based on this representation，the LoRA technique is employed to perform parameter-efficient fine-tuning on the LLaMA-3-8B-Instruct model，enabling the model to learn the semantic associations between network asset characteristics and device types. Experimental results on a test dataset containing 380 000 real-world network assets demonstrate that the proposed method maintains stable performance under highly imbalanced samples and long-tail device scenarios，achieving a Weighted F1-score of 0.959 1，which significantly outperforms the unfine-tuned base model. In addition，the model inference throughput is improved by 62.81%. These results verify the effectiveness and practicality of the proposed method for large-scale automated network asset device identification.

Key words：network asset identification；large language models；instruction tuning；multi-source heterogeneous features

0 引言

路由器、防火墙、网络摄像头、负载均衡设备等网络设备类资产广泛部署于公共网络和专用网络环境中，承担着网络转发、边界防护、数据采集与控制等关键功能，其安全性和运行状态直接关系到网络空间的整体安全态势。因此，对网络资产中设备类型及其厂商属性进行准确、自动化的识别，对网络空间测绘、安全风险评估以及漏洞关联分析等工作具有重要意义。目前，网络设备识别方法主要是基于静态指纹库规则匹配的方法，该方法在已知类型资产的识别中表现良好，但其依赖专家人工编写的正则表达式指纹库，维护成本高，泛化能力不足，在实际的开放网络环境中，准确识别多厂商、多类型的设备面临着严峻挑战。首先，不同厂商、甚至同一厂商的不同产品线，其 HTTP/HTTPS 响应头、HTML 页面结构及 SSL 证书均可能存在差异，导致指纹特征梳理工作量巨大；其次，特征的隐蔽与混淆日益普遍，现代网络设备倾向于使用最小化的登录页面，或通过 JavaScript 动态加载内容，导致传统的基于静态关键词匹配的方法难以提取有效识别信息；第三，新设备类型或型号出现速度快，基于人工指纹分析速度慢，难以快速发现新品种；此外，单一维度的特征往往不可靠，如端口可能被修改，Favicon 可能缺失，Banner 信息可能被管理员自定义掩盖等，导致基于规则特征的方法识别准确性低。

针对上述问题，本文提出了一种基于大语言模型指令微调[1-2]的网络资产设备类型识别方法，以实现 HTTP/HTTPS 协议网络资产中的设备类型智能化识别。利用大语言模型强大的语义理解与上下文推理能力，将设备识别任务重构为语义理解与生成任务。具体而言，本文主要工作如下：

（1）提出了一种多源采集与多平台聚合相结合的网络资产数据收集方案：通过将主动探测[3]与被动监听[4]相结合，并融合多种主流网络资产测绘平台的结果，构建了一个覆盖范围广、标签可靠性高的网络资产数据集。

（2）设计了一种面向大语言模型的多维异构特征融合与 Prompt[5]构建策略：针对网络资产数据在结构形式和语义层次上的高度异构性，本文提出将非结构化的 HTTP/HTTPS Body 与结构化的 SSL 证书、Header 键值对、端口及协议等特征进行统一序列化表达，并通过任务驱动的 Prompt 设计，将多源特征映射为大语言模型可理解的上下文输入形式。

（3）探索并验证了基于大语言模型的网络设备识别新范式。本文将 LLaMA[6]系列等通用大语言模型引入网络资产设备识别任务，在少量高质量指令样本的条件下，通过指令微调引导模型学习网络资产特征与设备类型之间的语义关联关系。实验结果表明，该方法不仅能够有效识别已知厂商与设备类型，在面对未知厂商或新型设备时也展现出良好的泛化能力，验证了大语言模型在网络资产识别场景中的有效性。

1 相关工作

随着网络规模的扩大以及网络设备类型的不断丰富，网络资产识别技术经历了由规则驱动向数据驱动的演进过程。现有研究根据其识别机制和建模方式的不同，主要可分为三类：基于静态指纹库的规则匹配方法、基于传统机器学习的识别方法，以及基于深度学习的识别方法。

1.1 基于静态指纹库的规则匹配方法

基于静态指纹库的规则匹配方法通过对网络资产进行探测，获取其在协议交互过程中暴露出的稳定特征，并将这些特征以规则的形式进行组织和存储，从而实现对资产类型的识别。它的核心流程分为三个阶段，分别是特征获取、字典构建和查表匹配。在特征获取阶段，通常采用主动探测与被动探测相结合的方式获取目标资产的特征信息。主动探测通过向目标主机发送特定协议请求，根据返回的响应来推断目标主机的状态、操作系统或服务类型。被动探测则是不发送任何数据包，通过监听网络流量，分析捕获到的数据包特征来识别资产。通过主动探测和被动探测相结合的方式，可以收集包括开放端口信息、协议类型、HTTP/HTTPS 响应头与响应体内容、TLS/SSL 证书字段以及站点图标哈希值等多种静态特征。在字典构建阶段，研究人员从目标资产的特征信息中提取具有高区分度的特征，将其映射为标准化的设备标签，从而构建起包含“特征—设备”映射关系的静态指纹库。在查表匹配阶段，识别引擎将实时采集的探测数据与静态指纹库中的规则进行匹配。若目标响应特征命中了库中的预定义规则，系统即依据规则逻辑判定该设备的厂商、型号及版本信息。

1.2 基于传统机器学习的识别方法

基于传统机器学习的识别方法通过从网络流量或协议交互数据中提取统计特征和语义特征，并对这些特征进行标准化处理，然后将处理后的数据作为训练集，用于对支持向量机[7]、随机森林[8]、朴素贝叶斯[9]或 K 近邻[10]等传统机器学习模型的训练，从而实现对设备类型进行分类。通过对已标注资产样本进行

监督学习，模型能够学习不同设备类型在特征空间中的分布差异，从而在未知资产上实现类型预测。

Li[11]等提出一种自动化设备指纹框架，对暴露在互联网的打印机、摄像头等物理设备，先爬取其内置网页，再利用网页结构、资源链接等特征无监督聚类生成指纹，替代手工维护关键词，原型系统对监控设备发现实验显示可扩展且准确。Miettinen[12]等提出了一个自动化 IoT 设备类型识别与安全隔离系统，系统利用设备首次联网的配网阶段流量，提取 23 维包级特征，如协议、端口、包长、目的 IP 计数等，使用随机森林算法进行分类识别。Sivanathan[13]等提出了一种基于网络流量特征的物联网设备分类框架，其核心是多阶段机器学习分类算法，结合网络流量的多种统计属性实现设备识别。

1.3 基于深度学习的识别方法

相比于传统机器学习方法，基于深度学习的识别方法展现出显著的优势。首先，深度学习具备强大的自动特征学习能力。深度神经网络通过多层非线性映射，能够在训练过程中从原始或仅需低阶预处理的数据中逐层学习具有判别性的中间表示，从而减少对人工设计特征函数和领域先验知识的依赖，实现了从原始输入到高层特征空间的端到端建模。除此之外，深度学习在非结构化序列数据建模方面具有天然优势，能够有效捕捉网络协议载荷及文本指纹中的长距离依赖关系与时序上下文信息，克服了传统方法丢失位置特征的缺陷。

Kotak[14]等通过将 IoT 设备的网络通信载荷转换为二维图像，并利用 CNN 进行分类，实现了基于通信行为的设备自动识别。该方法无需人工特征工程，能够有效区分不同 IoT 设备，并支持未知设备的检测。Dong[15]等提出了一种基于深度学习的 IoT 流量指纹识别方法，通过提取网络通信的时间序列特征并输入到 LSTM 模型中进行训练，使模型能够自动学习不同设备在包序列上的行为模式，并在包含 NAT/VPN 等现实网络干扰的环境中区分不同设备类型，实现了较高的设备分类准确率。

1.4 小结

尽管上述三种方法在各自的适用场景下取得了一定的成果，但它们在面对当前日益复杂的网络空间资产时，仍面临一个共同的问题：缺乏深层的语义理解与开放域的逻辑推理能力。无论是传统分类器还是深度神经网络，其本质上均为判别式模型。它们依赖于封闭集内的标注数据进行训练，仅仅学习了数据形态上的模式映射，而无法真正理解指纹文本背后的含义。当面对长尾资产、零样本新设备或描述极其模糊的指纹时，这些模型因缺乏通用的世界知识而往往判定失效。

2 数据收集与处理

2.1 数据收集与分析

本文构建了一套融合多源采集与多平台聚合的数据收集体系。如图 1 所示，该体系通过主动探测与被动监听相结合的方式，获取维度丰富、准确率高的网络资产数据集。

在主动探测阶段，以 Nmap 为核心引擎，对目标网络进行端口与服务深度扫描，获取资产的网络层与传输层基础信息，包括 IP 地址、开放端口、协议类型、端口关联服务及版本信息，以及操作系统类型与版本等。被动监听作为主动探测的有效补充，在网络边缘部署 Zeek 等流量分析工具，对实时网络通信数据进行捕获与解析，从而识别主动扫描难以发现的隐蔽资产，有效弥补主动探测在覆盖范围上的不足。在此基础上，本文进一步引入多平台数据聚合机制，以补充资产的业务层特征与厂商标签，即将主动探测获得的 IP 地址批量提交至 Shodan、Censys、DayDaymap 等网络资产测绘平台，获取其返回的设备类型、业务特征及厂商信息。针对不同平台标签结果可能存在不一致的问题，本文设计了一种基于置信度加权的投票融合策略：根据各平台历史准确率分配权重，对候选标签进行加权评分，并选取得分最高的标签作为最终结果，从而提高聚合标签的准确性和一致性，为后续模型训练提供可靠的数据支撑。

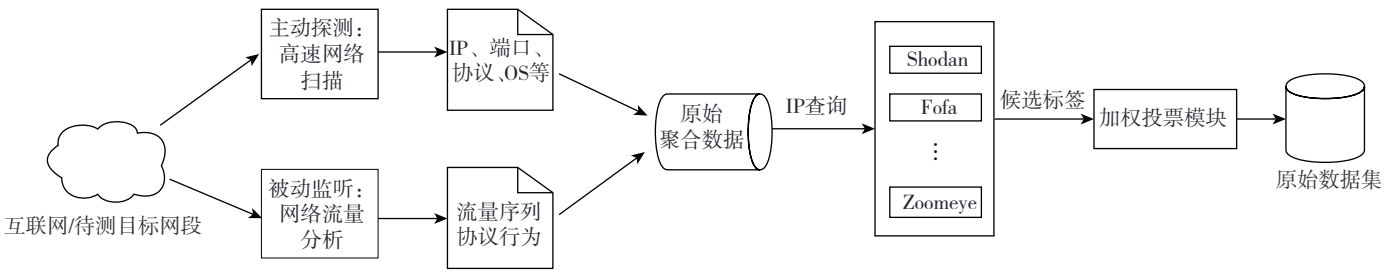


图 1 网络资产数据多源采集与标签融合处理流程

为了形式化描述标签冲突消解过程，假设针对某一资产 x，从 K 个测绘平台获取的候选标签集合为 $L_x = \{l_1, l_2, \cdots, l_K\}$。定义第 k 个平台的置信度权重为 w_k，其中，$\sum w_k = 1$，对于任意候选标签 T，其加权得分 Score(T) 计算如下：

$$\text{Score}(T) = \sum_{k=1}^{K} w_k \cdot \mathbb{I}(l_k = T) \tag{1}$$

其中，$\mathbb{I}(\cdot)$ 为指示函数，当平台 k 返回的标签 l_k 与 T 一致时取 1，否则取 0。

最终选定得分最高的标签 T^* 作为该资产的最终类型标注：

$$T^* = \underset{T}{\text{argmax}}\,\text{Score}(T) \tag{2}$$

通过上述数据收集方法收集的原始网络资产数据由多个层级的信息组成，包括网络层与传输层的基础通信特征、应用层的 HTTP/ HTTPS 交互内容以及加密通信场景下的 SSL 证书信息等。其中，部分字段如 HTTP/HTTPS 响应体（Body）具有较大的长度波动，而部分关键字段，如 HTTP/HTTPS 响应头（Header）和证书字段则具有较强的结构稳定性。具体字段组成如表 1 所示。

表 1　原始网络资产数据字段组成

数据维度	关键字段	示例
网络层	IP，Port	192. 168. 1. 1，80，443
传输层	Protocol	TCP
应用层	HTTP/HTTPS Header	Server，WWW-Authenticate
应用层	HTTP/HTTPS Body	HTML/JS 内容
安全传输层	SSL 证书	subject. CN，issuer
扩展标签层	厂商	ASUS，华为
扩展标签层	设备类型	路由交换设备，安全防护设备

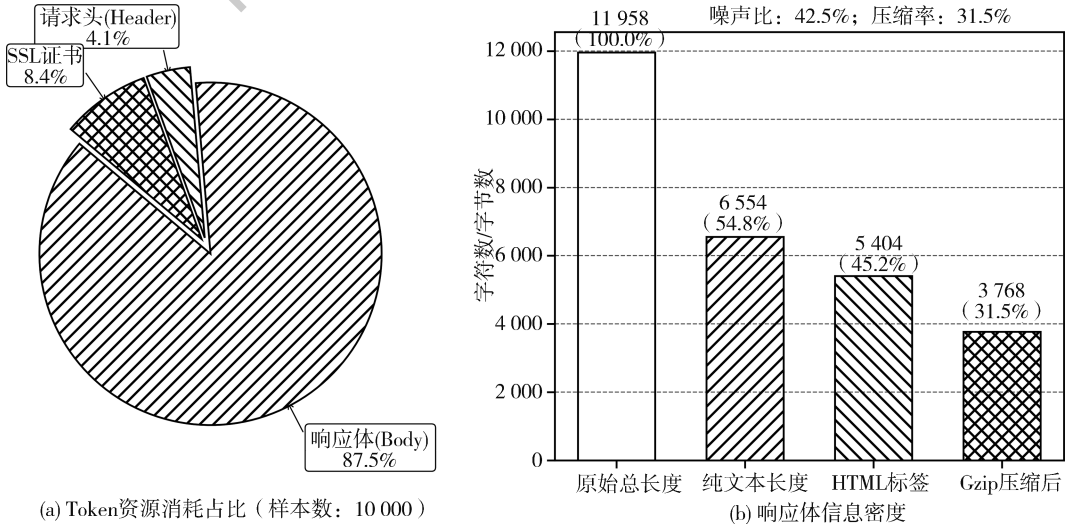
为了定量分析不同字段在模型输入中的资源占用及其有效信息密度，本文从原始数据集中随机抽取 10 000 条网络资产样本，对其中 Body、Header 及 SSL 证书字段的 token 消耗情况进行统计分析，并进一步对响应体字段内部的有效文本与结构噪声进行细粒度拆分，其统计结果如图 2 所示。

图 2（a）展示了在 10 000 条原始网络资产数据中，不同字段的平均 token 消耗占比情况。可以观察到，Body 在整体输入中占据绝对主导地位，其 token 消耗占比高达 87. 5%；相比之下，Header 和 SSL 证书字段的 token 占比分别仅为 4. 1% 和 8. 4%。这一结果表明，在直接使用原始数据作为模型输入的情况下，Body 极易消耗大量 token 预算，从而导致其他关键字段在输入长度受限时被截断甚至完全丢失，为后续的资产类型识别带来潜在风险。

为进一步探究 Body 字段内部的有效信息密度，本文对其字符构成进行了细粒度拆解，如图 2（b）所示。结果显示，原始 Body 中仅有 54. 8% 为纯文本信息，其余 45. 2% 均为 HTML 标签等结构化噪音。由于 HTML 标签主要服务于页面渲染，缺乏实质性的设备指纹特征，这部分内容构成了输入数据中的低价值噪声。此外，Gzip 压缩实验显示 Body 的平均压缩比达到了 31. 5%，即体积减少了 68. 5%，这一较高的压缩空间有力地证明了原始响应体中存在严重的语义冗余。

2. 2　数据处理

通过 2. 1 节对原始数据组成字段中 HTTP/HTTPS Body、Header 以及 SSL 证书的分析可以发现，Body 在原始数据中占据了主要的输入长度，但其内部包含大量结构性噪声和冗余信息，而 Header 和 SSL 证书字段



(a) Token资源消耗占比（样本数：10 000）　　(b) 响应体信息密度

图 2　原始网络资产数据的 Token 消耗分布与响应体信息密度分析

虽然 token 占比较低，却往往承载着与资产类型高度相关的关键特征。因此，若不加区分地对原始数据进行截断，极易导致模型丢失核心判别依据。针对这一问题，本文提出了一种关键特征优先保留的数据预处理策略，在输入长度受限的条件下优先保留高价值字段，并对低信息密度的冗余内容进行压缩与裁剪。

2.2.1 基于结构化清洗与语义去重的 HTTP/HTTPS 响应体处理机制

如前文分析所述，HTTP/HTTPS Body 通常包含完整的页面内容，在原始网络资产数据中占据了主要的信息规模。然而，Body 内部同时混杂了大量噪声内容，例如页面结构标签、通用脚本库以及样式定义等。这类内容不仅信息冗余度高，而且与设备类型判别的相关性较弱。若直接将 Body 输入模型，容易在有限的输入长度约束下挤占 Header、SSL 证书等高价值字段的上下文空间，同时引入无关特征干扰模型的判别过程。因此，针对 Body 的处理有必要从去除噪声与关键信息提取两个层面进行设计。基于上述考虑，本文构建了一套面向设备识别任务的多级结构化 Body 预处理流水线，其整体流程如图 3 所示。在该流水线的特征提取阶段，为了将非结构化的 Body 转化为高区分度的设备指纹，本文并非对所有 DOM 节点进行无差别保留，而是从页面元信息、资源依赖、交互结构以及关键语义线索等多个层面，对 Body 进行结构化特征抽取。

在页面元信息层面，提取 < title > 及部分具有设备指纹意义的 < meta > 标签，用于捕获页面生成框架、开发工具或厂商相关的隐含信息。在脚本与资源依赖层面，提取页面中引用的外部脚本与关键内联脚本，同时通过黑名单机制过滤常见的通用前端库以及统计分析类脚本。该设计的目的是保留与设备前端实现相关的专用脚本特征，避免模型受到高度普遍化和与设备类型无关的通用库干扰，从而降低噪声特征对模型学习的影响。在交互结构层面，重点提取页面中的表单结构，包括表单的提交方式、输入字段类型以及下拉框和文本域等元素。设备管理界面的表单结构直接反映了设备的功能模块和管理逻辑，是区分不同设备类型和厂商的重要结构性特征。在多媒体与页面布局层面，提取图片资源、框架结构及页面内部链接拓扑信息。图片路径、页面分栏方式以及功能页面之间的跳转关系，往往反映了设备管理界面的整体架构特征，具有跨版本、跨环境的稳定性。同时，通过过滤外部社交媒体等无关链接，进一步抑制噪声。在页面样式与图标层面，提取 CSS 样式表及 Favicon 等链接信息。这类资源路径在嵌入式设备中往往与厂商 UI 设计风格和固件结构密切相关，具有一定的指纹价值。最后，在关键文本信息层面，针对性地从页脚及正文中提取包含版权声明、厂商标识或设备型号信息的短文本片段。这类文本虽然在页面中所占比例较小，但通常直接暴露设备厂商或产品系列，是具有高判别力的显式指纹特征。

通过对上述特征的结构化提取，原本杂乱的 HTML 响应体被重构为一个语义清晰、维度丰富的特征字典。这不仅大幅降低了模型的输入长度压力，更确保

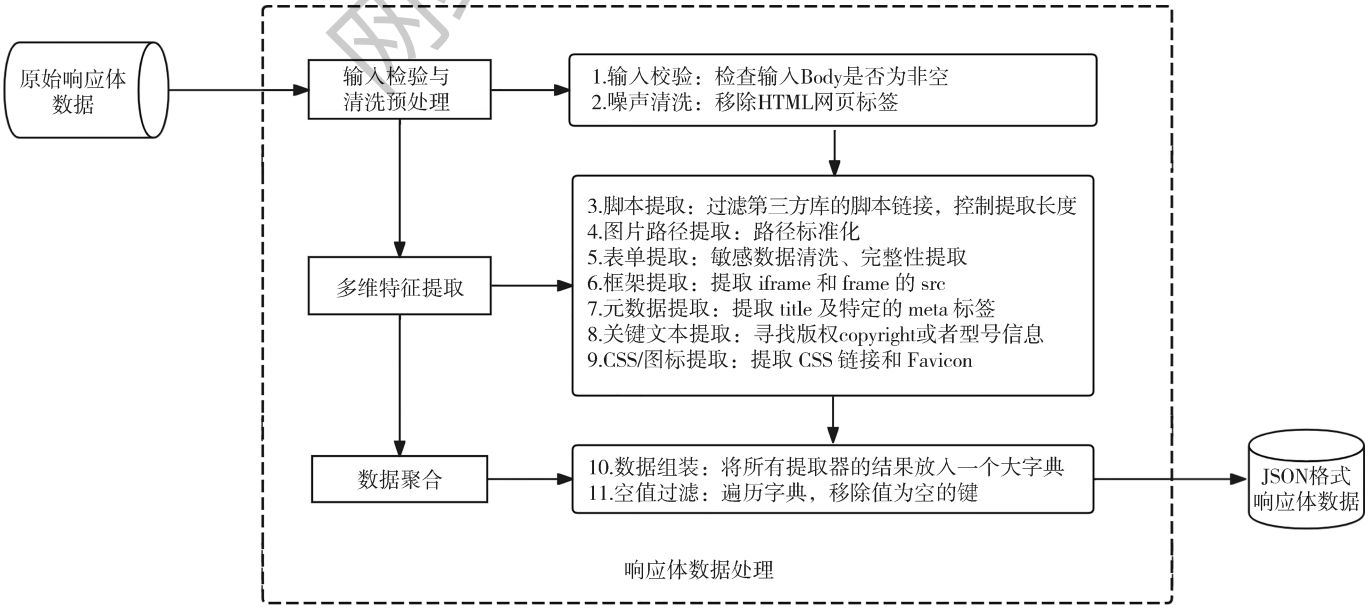


图 3　响应体数据结构化处理流程图

了模型能够聚焦于那些具有决定性意义的设备指纹信息。

2.2.2 基于黑名单的 HTTP/HTTPS 头部特征提取

HTTP/ HTTPS Header 是识别网络资产类型的关键依据，其中 Server、WWW-Authenticate 等字段往往直接暴露了设备的固件版本或厂商信息。然而，原始头部数据中包含大量动态生成的字段和传输层字段，这些字段在不同请求间高度变化，若直接输入模型，会导致模型关注点偏移，产生严重的过拟合。同时，因为网络资产种类繁多，可能会有各种自定义 Header，如 X-Juniper-Security，这个字段的值是 Juniper 品牌的安全防护产品。如果用白名单，可能会漏掉这些关键特征。因此，本模块采用黑名单过滤的策略，旨在最大化保留设备指纹的同时，彻底剔除时变噪声。具体需要过滤的字段如表 2 所示。

表 2　黑名单字段表

噪声类别	典型字段示例	剔除原因
时变型噪声	Date，Last-Modified，Expires，Age	随请求时间即时变化，与设备固有属性无关
传输层状态	Connection，Keep-Alive，Content-Length，Transfer-Encoding	反映当前 TCP/HTTP 会话状态，受网络环境影响大，非设备指纹特征
中间件痕迹	Via，X-Forwarded-For，CF-Ray，X-Real-IP	记录流量经过的代理、CDN 或负载均衡器信息，干扰对终端设备的识别
高熵动态值	ETag，X-Session-ID，X-CSRF-Token	包含随机生成的哈希或令牌，属于高维稀疏特征，极易引入无关噪声
安全策略头	X-Frame-Options，X-XSS-Protection，X-Content-Type-Options	通用的 Web 安全配置，区分度低，且在同类 Web 服务中高度重复

2.2.3 基于 X.509 解析的证书身份指纹提取

在 HTTPS 场景下，SSL/TLS 证书是网络资产的数字身份证，其中包含了高价值的身份认证信息。然而，原始的证书数据通常以 PEM 或 DER 格式存在，对于模型而言，这是一串高熵的随机字符，缺乏语义可读性。为此，本文设计了基于 X.509 标准的解析模块，将非结构化的证书数据解码为结构化的身份字典。其中重点关注 Subject 证书持有者和 Issuer 证书颁发者两个核心字段。Subject 字段直接揭示了设备的制造商或者组织，是判断资产归属的最强证据。Issuer 揭示了证书的信任链。在 IoT 场景中，大量设备使用厂商自签名的证书，此时 Issuer 与 Subject 相同，这本身就是一种强相关的设备指纹特征。对于缺失字段、默认占位值或明显无意义的证书信息，本文在预处理阶段进行统一过滤，以减少噪声特征对模型学习过程的干扰。经过上述处理后，SSL 证书特征被表示为由 Subject、Issuer 及 Country 字段组成的结构化文本片段，并与 HTTP/HTTPS Header 和 Body 特征一同构建统一的模型输入序列。

3　实验验证

3.1　数据集构造与实验环境设置

通过多级数据处理流程，原始网络资产数据已经转化为由 HTTP/HTTPS Header、Body 及 SSL 证书等多源特征组成的结构化表示。然而，以 LLaMA-3 为代表的生成式大语言模型本质上是基于序列 – 序列范式进行训练的，无法直接处理离散的键值对数据。基于此，本文在实验阶段引入指令微调范式，并将处理后的特征构造成 Alpaca 格式的数据集，用以支撑后续模型训练与评估。Alpaca 格式训练样本由 Instruction、Input 和 Output 三部分组成，分别对应设备识别任务描述、多源资产特征输入以及厂商和设备类型标签输出。

为了验证大语言模型在有限标注样本条件下对海量网络资产的识别与泛化能力，本实验构建了一个具有高度异构性的实验数据集与轻量化训练环境。在数据构建方面，笔者从真实的互联网测绘流量中选择了 600 余条高置信度样本作为指令微调的训练集，该集合覆盖了 Asus、Fortinet、Cisco 等主流厂商及路由器、防火墙、网络摄像头等核心设备类型，旨在作为高质量的种子数据引导模型学习设备指纹的语义映射关系。与此同时，为了严格评估模型在开放网络环境中的鲁棒性，实验构建了包含 38 万条样本的大规模测试集。为了确保评估环境与模型训练环境的一致性，同时也为了验证本研究提出的预处理策略在大规模真实数据上的有效性，在测试阶段对 38 万条测试样本执行了与训练集完全相同的预处理流程。所以，对于任何新捕获的未知资产流量，系统均先进行标准化清洗，再由模型进行判别。在模型与计算环境方面，本研究选用 LLaMA-3-8B-Instruct 作为基座模型，采用 LoRA 技术进行参数高效微调，通过冻结预训练参数并注入低秩矩阵，在单张具备 32 GB 显存的 GPU 平台上完成模型的领域适配训练，充分验证了该方法在有限算力资源下的可部署性与推理效能。

3.2　实验结果分析

本文采用混淆矩阵以及 Precision、Recall 和 F1-score 作为主要评价指标。预测为正的样本中，实际为正的样本数为 TP，实际为负的样本数为 FP；预测为负

的样本中，实际为正的样本数为 FN，实际为负的样本数为 TN。

混淆矩阵用于可视化展示模型在各类别间的判别细节，辅助分析模型在不同设备类型上的混淆模式与边界界定能力。

精确率（Precision）衡量模型预测为某一类别时的可信度。

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{3}$$

召回率（Recall）衡量模型对真实类别样本的覆盖能力。

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{4}$$

F1-score 是 Precision 与 Recall 的调和平均，用于综合评价模型性能。

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5}$$

由于测试集中包含大量厂商与设备类型，直接对全部类别绘制混淆矩阵将导致结果难以解读。所以，本文从测试集中厂商设备的样本占比以及设备类型的多样性这两个方面出发，选取了 15 个具有代表性的厂商设备用于混淆矩阵分析。图 4 与图 5 分别给出了未经微调的 LLaMA-3-8B 基座模型与经指令微调后的模型在测试集上的混淆矩阵结果。

由图 4 和图 5 可得，在针对前 14 个厂商设备的分类任务中，未经微调的 LLaMA-3-8B 基座模型与经指令微调后的模型识别表现整体接近，主对角线的值均在 99% 以上。这说明，一方面基座模型本身具备较强的通用特征提取能力，另一方面，标准化的数据清洗与结构化的输入构建过程降低了任务的解析难度，使得基座模型在处理常规样本时，能够达到与微调模型相近的判别水平。然而，在处理 Chipspoint_NVR 这一特定类别时，未经微调的基座模型表现出明显的分布外失效现象。具体而言，基座模型在 Chipspoint_NVR 对应的混淆矩阵行与列均为 0，表明该模型在推理过程

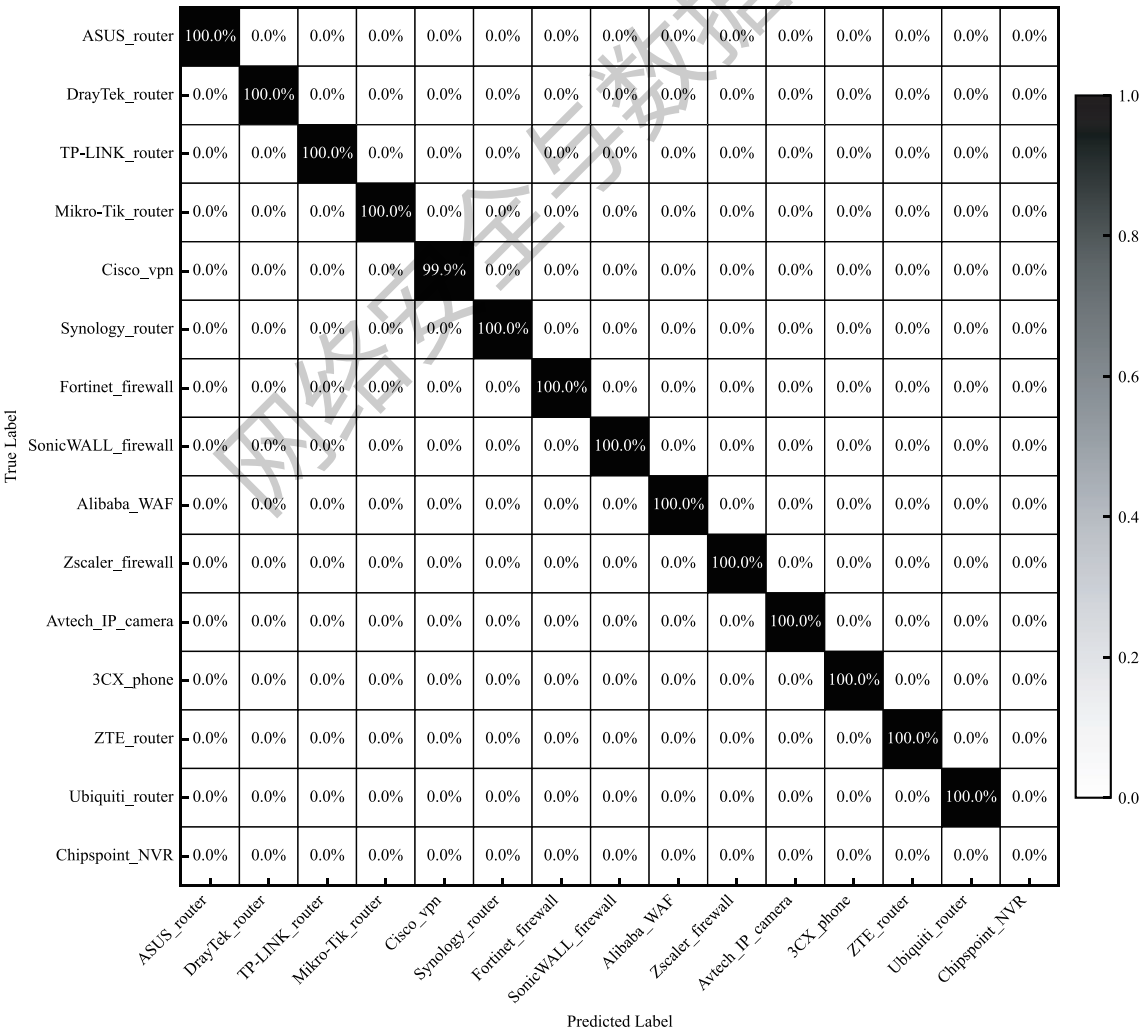


图 4　基座模型在 15 个厂商设备上的混淆矩阵

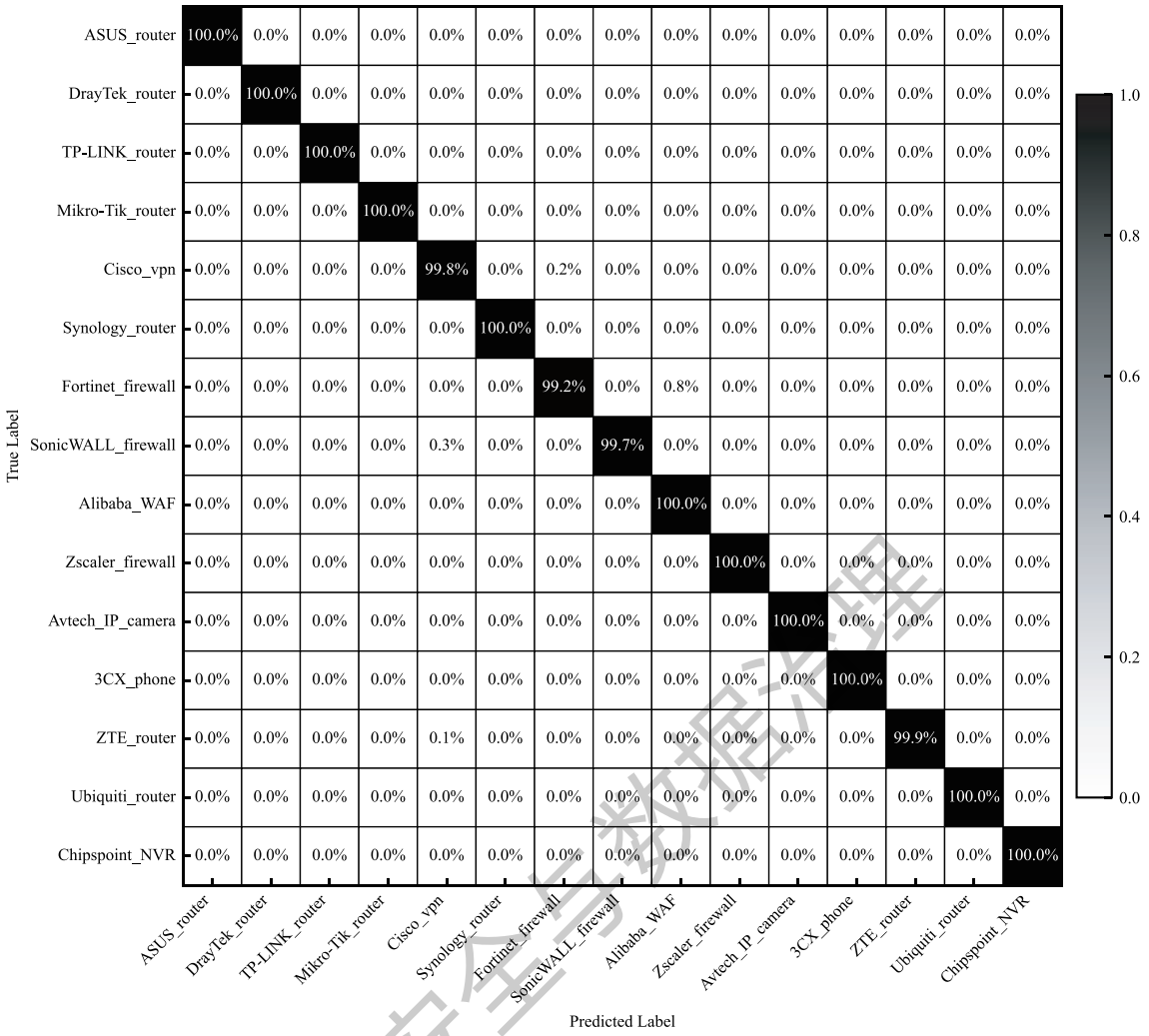


图 5　微调后的模型在 15 个厂商设备上的混淆矩阵

中未能将任何样本预测为该类别。这一结果说明，对于基座模型而言，该标签在当前任务设置下不仅难以区分，且在生成阶段具有极低的输出概率。

造成该现象的主要原因在于，Chipspoint_ NVR 属于典型的长尾厂商标签，其实体名称在通用预训练语料中出现频率较低。在缺乏领域对齐约束的情况下，基座模型更倾向于生成语义上更为通用、先验概率更高的厂商或设备类型词汇，而非输出具体且低频的实体名称。

相比之下，指令微调通过少量高质量标注样本，引导模型在给定网络资产特征条件下，将 Chipspoint_ NVR 这一特定标签纳入其条件生成分布中。微调后在该类别上达到 100% 的命中率，表明模型已成功建立起从多源输入特征到该厂商设备标签之间的稳定映射关系，完成了从通用语言建模能力向垂直领域设备识别能力的有效迁移。

表 3 和表 4 列出了未经微调的 LLaMA-3-8B 基座模型与经指令微调后的模型在这 15 个厂商设备上的 Precision、Recall、F1-score、样本数量结果。

表 3　基座模型在 15 个厂商设备上评估结果

厂商设备	精确率	召回率	F1 分数	样本数量
ASUS_ router	0. 998 4	0. 997 4	0. 997 9	43 084
DrayTek_ router	1. 000 0	0. 826 5	0. 905 0	30 926
TP_ LINK_ router	0. 978 7	0. 934 4	0. 956 0	14 185
MikroTik_ router	0. 999 2	1. 000 0	0. 999 6	3 767
Cisco_ vpn	0. 995 9	0. 755 5	0. 859 2	11 279
Synology_ router	0. 999 3	0. 984 3	0. 991 8	9 055
Fortinet_ firewall	0. 974 6	0. 408 1	0. 575 3	35 961
SonicWALL_ firewall	0. 999 7	0. 960 6	0. 979 8	10 692
Alibaba_ WAF	0. 999 6	0. 852 4	0. 920 2	16 285
Zscaler_ firewall	1. 000 0	0. 999 5	0. 999 8	4 247
Avtech_ IP_ camera	0. 941 4	0. 999 8	0. 969 7	5 353
3CX_ phone	1. 000 0	0. 987 8	0. 993 8	1 799
ZTE_ router	0. 997 9	0. 994 4	0. 996 2	1 440
Ubiquiti_ router	0. 950 4	0. 931 2	0. 940 7	1 686
Chipspoint_ NVR	0. 000 0	0. 000 0	0. 000 0	35 508

表 4　微调后的模型在 15 个厂商设备上评估结果

厂商设备	精确率	召回率	F1 分数	样本数量
ASUS_ router	0. 998 7	0. 997 6	0. 998 2	43 084
DrayTek_ router	0. 999 8	0. 965 8	0. 982 5	30 926
TP_ LINK_ router	0. 974 8	0. 935 3	0. 954 6	14 185
MikroTik_ router	0. 997 9	0. 999 2	0. 998 5	3 767
Cisco_ vpn	0. 943 2	0. 866 8	0. 903 4	11 279
Synology_ router	0. 998 9	0. 977 6	0. 988 1	9 055
Fortinet_ firewall	0. 947 7	0. 807 2	0. 871 8	35 961
SonicWALL_ firewall	0. 999 9	0. 972 8	0. 986 2	10 692
Alibaba_ WAF	0. 977 9	0. 994 4	0. 986 1	16 285
Zscaler_ firewall	1. 000 0	0. 994 8	0. 997 4	4 247
Avtech_ IP_ camera	0. 943 4	0. 953 1	0. 948 2	5 353
3CX_ phone	1. 000 0	0. 998 3	0. 999 2	1 799
ZTE_ router	0. 870 5	0. 994 4	0. 928 4	1 440
Ubiquiti_ router	0. 948 2	0. 934 2	0. 941 1	1 686
Chipspoint_ NVR	0. 998 2	0. 932 9	0. 964 5	35 508

由于测试集中不同厂商设备的样本数量分布高度不均衡，例如 ASUS_ router 样本量高达 43 084 条，而 ZTE_ router 仅有 1 440 条，在这种情况下，若直接采用 Macro F1-score 对各类别进行等权平均，小样本类别的性能波动将被不成比例地放大，从而过度影响整体评估结果，难以真实反映模型在大规模网络资产识别场景中的实际表现。因此，本文在总体性能评估中引入 Weighted F1-score 作为主要评价指标，以更加客观地衡量模型在样本分布不均衡条件下的综合识别能力。

Weighted F1-score 在各类别 F1-score 的基础上，引入类别样本数量作为权重进行加权求和，其定义如下：

$$\text{Weighted-F1} = \sum_{i=1}^{C} \frac{n_i}{N} \cdot \text{F1}_i \tag{6}$$

其中 C 表示类别总数，n_i 表示第 i 类的样本数量，$N = \sum_{i=1}^{C} n_i$ 为测试集样本总数，F1_i 为第 i 类对应的 F1-score。

经计算可得，未经微调的 LLaMA-3-8B 基座模型在测试集上的 Weighted F1-score 为 0. 743 0，而经指令微调后的模型 Weighted F1-score 提升至 0. 959 1，实现了 21. 61% 的绝对性能提升。该结果表明，在高度类别不均衡且包含大量噪声与长尾设备的真实网络资产数据环境中，通用大语言模型的零样本能力仍存在明显局限；而通过少量高质量指令样本进行领域对齐后，模型能够显著改善整体判别稳定性，在主流类别与长尾类别之间实现更加均衡的识别性能。

图 6 展示了未经微调的 LLaMA-3-8B 基座模型与经指令微调后的模型的推理吞吐量趋势对比，每推理 5 000 条时记录下时间，以推理前 20 个 5 000 条的时间做对比。

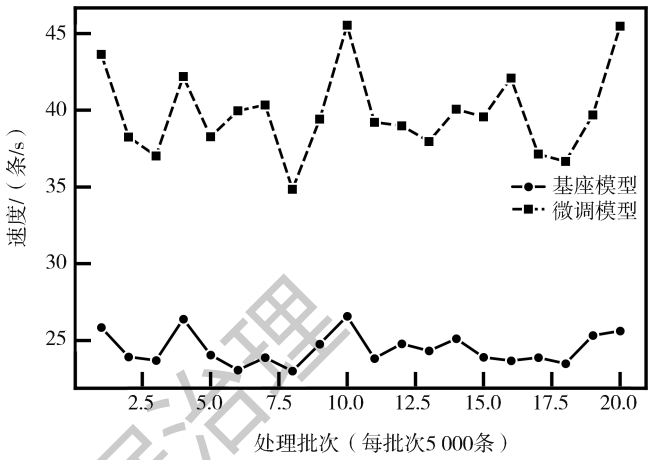


图 6　推理吞吐量趋势对比

经计算可得，未经微调的 Base LLaMA-3-8B 模型平均推理速度为 24. 46 条/s，而经指令微调后的模型平均推理速度提升至 39. 82 条/s，实现了 62. 81% 的推理速度提升。该结果表明，指令微调不仅未引入额外的推理开销，反而在实际批量推理场景中显著提升了模型的生成效率，为大规模网络资产自动识别任务提供了更高的系统吞吐能力。

综合准确性与推理效率两方面的实验结果可以看出，指令微调不仅有效提升了模型在复杂网络资产识别任务中的判别能力，还显著提高了其在大规模批量推理场景下的处理效率，为实际系统部署提供了有力支撑。

此外，本文还将微调后的 LLaMA-3-8B 与近年来先进的网络资产识备识别框架进行了对比，结果如表 5 所示，其中标“ –”的部分表示该指标在对应文献中未提及。对比后发现微调模型在平均识别准确率和单个设备平均识别时间上均优于其他框架。

4　结论

本文针对传统网络资产识别方法在复杂异构网络环境中依赖人工规则、泛化能力不足以及难以应对长尾设备识别的问题，提出了一种基于大语言模型指令微调的网络资产设备类型识别方法。通过构建多源采集与多平台聚合的数据收集体系，并结合关键特征优先保留的预处理策略，实现了对 HTTP/HTTPS 响应体、响应头及 SSL 证书等异构特征的高效融合与表达。

表 5　现有主流网络资产识别系统的效能对比

研究	分类识别算法/模型	选取特征	平均准确率/%	单个设备平均识别时间/ms
IoTSentinel[12]	随机森林 + 编辑距离 tiebreak 策略	协议、端口、IP	81.5	157.7
HomeMole[15]	随机森林	端口、包方向、协议、包间隔、方向	84.5	—
HomeMole[15]	双向 LSTM	协议、包间隔、包大小、方向、端口	92.1	—
IoTDevID[16]	决策树	包大小、载荷大小、IP、协议	83.3	—
微调模型	LLaMA-3-8B	SSL 证书、响应体、响应头、IP 等	93.8	25.1

在此基础上，将设备识别任务重构为指令驱动的生成式语义理解任务，并采用 LoRA 技术对通用大语言模型进行参数高效微调，使模型能够在少量高质量标注样本条件下学习网络资产特征与设备类型之间的稳定映射关系。

实验结果表明，与未微调的基座模型相比，指令微调后的模型在大规模真实网络资产数据集上显著提升了整体识别性能，尤其在长尾厂商和低频设备类型上表现出更强的鲁棒性与泛化能力。同时，该方法在提升识别准确率的同时并未引入额外推理开销，在批量推理场景下显著提高了系统吞吐量，具备良好的工程可部署性。

综上所述，本文验证了大语言模型在网络资产设备识别任务中的可行性与优势，为网络空间测绘与安全资产管理提供了一种新的技术范式。未来工作将进一步探索更精细的设备层级识别、跨协议特征融合以及持续学习机制，以提升模型在动态网络环境中的自适应能力。

参考文献

[1] WEI J, BOSMA M, ZHAO V Y, et al. Finetuned language models are zero-shot learners[C]//Proceedings of the 10th International Conference on Learning Representations, 2022.

[2] HU E J, SHEN Y, WALLIS P, et al. LoRA: low-rank adaptation of large language models[C]//Proceedings of the 10th International Conference on Learning Representations, 2022.

[3] 谢镕镕. 基于主动探测的物联网设备识别技术研究[D]. 北京：中国人民公安大学, 2023.

[4] 刘丹丹. 基于被动检测的物联网设备识别技术研究[D]. 北京：中国人民公安大学, 2023.

[5] ZHANG Z S, ZHANG A, LI M, et al. Automatic chain of thought prompting in large language models[C]//Proceedings of the 11th International Conference on Learning Representations, 2023: 1 - 18.

[6] DUBEY A, JAUHRI A, PANDEY A, et al. The Llama 3 Herd of Models [R]. arXiv: 2407.21783, 2024.

[7] CHANG C C, LIN C J. LIBSVM: a library for support vector machines[J]. ACM Transactions on Intelligent Systems and Technology, 2011, 2 (3): 18 - 22.

[8] LIAW A, WIENER M. Classification and regression by random forest[J]. R News, 2002, 2 (3): 18 - 22.

[9] BREIMAN L. Estimating the error rate of a prediction rule: a simple randomization method[J]. Machine Learning, 1992, 9 (1): 31 - 43.

[10] PARZEN E. A nonparametric estimation of the probability density runction and the mode[J]. Annals of Mathematical Statistics, 1962, 33 (3): 1065 - 1076.

[11] LI Q, FENG X, LI Z, et al. GUIDE: graphical user interface fingerprints physical devices[C]//2016 IEEE 24th International Conference on Network Protocols (ICNP). IEEE, 2016: 1 - 2.

[12] MIETTINEN M, MARCHAL S, HAFEEZI M, et al. IoT Sentinel: automated device-type identification for security enforcement in IoT[C]//2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS). 2017: 2177 - 2184.

[13] SIVANATHAN A, GHARAKHEILI H, LOI F, et al. Classifying IoT devices in smart environments using network traffic characteristics[J]. IEEE Transactions on Mobile Computing, 2019, 18 (8): 1745 - 1759.

[14] KOTAK J, ELOVICI Y. IoT device identification based on network communication analysis using deep learning [J]. Journal of Ambient Intelligence and Humanized Computing, 2023, 14 (7): 9113 - 9129.

[15] Dong Shuaike, Li Zhou, Tang Di, et al. Your smart home can't jeep a secret: towards automated fingerprinting of IoT traffic with neural networks [C]//Proceedings of the 15th ACM Asia Conference on Computer and Communications Security (ASIA CCS' 20), 2020: 47 - 59.

[16] KOSTAS K, JUST M, LONES M A. IoTDevID: a behavior-based device identification method for the IoT[J]. IEEE Internet of Things Journal, 2022, 9 (23): 23741 - 23749.

（收稿日期：2026 - 03 - 05）

作者简介：

陈倩怡（1999 - ），女，硕士研究生，主要研究方向：网络空间测绘、人工智能。

苏马婧（1985 - ），通信作者，女，博士，正高级工程师，主要研究方向：网络空间测绘、网络安全。E-mail：sumj@ncse. com. cn。

陈紫璇（1999 - ），女，硕士，助理工程师，主要研究方向：网络资产测绘。

版权声明

凡《网络安全与数据治理》录用的文章，如作者没有关于汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权等版权的特殊声明，即视作该文章署名作者同意将该文章的汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权授予本刊，本刊有权授权本刊合作数据库、合作媒体等合作伙伴使用。同时，本刊支付的稿酬已包含上述使用的费用，特此声明。

《网络安全与数据治理》编辑部