



DOI:10.12404/j.issn.1671-1815.2404568

引用格式:唐浩耀,崔鑫,张艺炜,等.基于数据增强与 CNN-Optuna-Attention 改进的网络流量分类方法[J].科学技术与工程,2025,25(15):6419-6430.

Tang Haoyao, Cui Xin, Zhang Yiwei, et al. Network traffic classification method improved based on data augmentation and CNN-Optuna-Attention[J]. Science Technology and Engineering, 2025, 25(15): 6419-6430.

# 基于数据增强与 CNN-Optuna-Attention 改进的网络流量分类方法

唐浩耀,崔鑫\*,张艺炜,赵庆慧

(山东理工大学计算机科学与技术学院,淄博 255049)

**摘要** 针对基于卷积神经网络的流量分类方法难以捕捉序列中不同部分的重要性、特征提取不足的问题,为提升网络流量分类精度,提出一种注意力机制与卷积神经网络相结合的流量分类方法,在卷积神经网络模型的基础上设计实现一种注意力机制层,通过接收全连接层的输出作为输入,计算输入特征的权重并乘以原始特征,实现对关键特征的加强,进而有助于提高模型对关键信息的捕捉能力。其次针对网络流量类别样本数不均导致模型过拟合于高比例类别,难以识别分类小比例类别的问题,提出了一种对数据集进行数据增强的方法。并且考虑到超参数组合优化的角度,提出一种基于贝叶斯优化的超参数搜索策略和五折交叉验证的方式对模型的超参数组合进行优化。通过上述方法研究确定模型的超参数组合。使用公开数据集进行上述实验与模型测试,结果表明:与其他方法相比,总体准确率、精确率以及  $F_1$  分数都有明显的提升,验证了本文所提方法具有更好的分类性能。

**关键词** 深度学习;流量分类;注意力机制;卷积神经网络;超参数优化;五折交叉验证

中图分类号 TP393; 文献标志码 A

## Network Traffic Classification Method Improved Based on Data Augmentation and CNN-Optuna-Attention

TANG Hao-yao, CUI Xin\*, ZHANG Yi-wei, ZHAO Qing-hui

(School of Computer Science and Technology, Shandong University of Technology, Zibo 255049, China)

**[Abstract]** In order to improve the accuracy of network traffic classification, a traffic classification method combining an attention mechanism and a convolutional neural network was proposed. An attention mechanism layer was designed and implemented on the basis of the convolutional neural network model, which received the output of the fully connected layer as input, calculated the weight of the input features, and multiplied it by the original features to strengthen the key features. This, in turn, helped to improve the model's ability to capture key information. Secondly, in order to solve the problem that the model was overfitting to the high-proportion category due to the unbalanced sample number of network traffic categories, and it was difficult to identify the small-proportion categories, a method to augment the dataset was proposed. Considering the perspective of hyperparameter combination optimization, a hyperparameter search strategy based on Bayesian optimization and five-fold cross-validation was proposed to optimize the hyperparameter combination of the model. The combination of hyperparameters of the model was determined by the above methods. The public dataset was used for the above experiments and model tests. The results show that compared with other methods, the overall accuracy, precision, and  $F_1$  score are significantly improved, which verifies that the proposed method has better classification performance.

**[Keywords]** deep learning; traffic classification; attention mechanisms; convolutional neural networks; hyperparameter optimization; five-fold cross-validation

据中国互联网络信息中心(China Internet Network Information Center, CNNIC)在北京发布的第 53 次报告<sup>[1]</sup>可知,截至 2023 年 12 月,中国网民总数突

破 10.92 亿人。移动互联网流量达 3 015 亿 GB。这一数据既体现了网络应用的日益丰富与普及,也揭示了网络资源使用紧张、分配不均衡和利用率低

收稿日期:2024-06-18 修订日期:2025-11-17

基金项目:山东理工大学科技博士项目(4041422007)

第一作者:唐浩耀(2000—),男,汉族,山东济南人,硕士研究生。研究方向:网络服务与信息安全。E-mail:1393124747@qq.com。

\* 通信作者:崔鑫(1972—),女,汉族,山东淄博人,博士,副教授。研究方向:下一代互联网技术、网络安全、网络大数据、无线传感网。E-mail:cx@sdut.edu.cn。

等问题<sup>[2]</sup>。

网络流量分类 (network traffic classification, NTC)<sup>[3]</sup> 作为现代网络管理与安全不可或缺的一环,对维持网络稳定和保障网络安全具有不可忽视的作用。NTC 的核心任务是将互联网流量按照预设的类别进行区分,如协议类型 (UDP、TCP 等)、应用类型 (Youtube、Facebook 等) 以及异常检测 (正常、异常)。它不仅可以被互联网服务提供商 (internet service provider, ISP) 用于服务质量 (quality of service, QoS)<sup>[4]</sup> 和故障排除,同时也广泛用于入侵检测系统,以保护用户设备免受潜在的网络攻击<sup>[5]</sup>。因此 NTC 一直是网络安全研究的重要领域<sup>[6-7]</sup>。

早期网络环境较简单,NTC 依赖基于端口与有效载荷的方法<sup>[8]</sup>,然而伴随着随机或动态端口号的使用,基于端口的技术准确率下降。Moore 等<sup>[9]</sup> 的研究揭示了一个重要事实,采用 IANA (internet assigned numbers authority) 列表的基于端口的流量分类技术存在局限性,其准确率不高于 70%; Madhukar 等<sup>[10]</sup> 的研究发现基于端口的技术无法识别他们所调查的流量比例高达 30% ~ 70%。

因此有效载荷的分类技术<sup>[8]</sup> 被研究者们开发并用于流量分类任务,这种新技术也称为深度包检测 (deep packet inspection, DPI), DPI 技术克服了端口变更带来的影响。然而新应用层出不穷,导致特征库中产生大量 DPI 技术无法识别的未知流量特征。Gringoli 等<sup>[11]</sup> 对 DPI 方法的性能进行了评估,实验结果显示,在 UNIBS 数据集上, L7-filter 等 DPI 工具的正确分类率为 67.73%,而在 POLITO 数据集上,其正确分类率则仅为 58.79%,这凸显了 DPI 技术面对未知流量特征时的问题。

为了解决 DPI 方法的问题,研究人员纷纷尝试将机器学习 (machine learning, ML) 算法应用于 NTC 领域<sup>[12-14]</sup>; 如决策树<sup>[15]</sup>、随机森林、支持向量机等。罗冬梅<sup>[12]</sup> 提出一种基于机器学习算法的流量分类方法,针对不同类别的在线流量样本流集合筛选出若干最近邻样本流,并计算每个样本流特征权重来确定各个特征与类别之间的相关性,并将相关性大的特征视为在线流量特征。根据得到的特征选取部分标识在线流量数据,以确定  $K$  中值聚类的起始中心来构造映射关系,从而获取未知的在线流量种类。实验结果表明该方法有很高的分类精度,且扩展性和适应能力较强。然而,随着互联网流量的激增和新型的网络流量不断涌现,仅仅依赖人工分析来提取网络流量特征已经无法满足需求。

为了克服传统方法和 ML 方法在分类任务上的局限性,深度学习 (deep learning, DL)<sup>[16]</sup> 应运而

生。DL 作为 ML 的前沿分支,其优势在于能够自动从大规模数据中提取特征<sup>[17]</sup>。这种端到端 NTC 的架构大幅减少了人工参与特征提取的繁琐过程,并且凭借其出色的特征识别能力能够揭示人类分析难以察觉的复杂特征,从而实现更一致和准确的分类效果<sup>[18]</sup>。宋继红等<sup>[19]</sup> 通过改进现有模型,引入 Inception 模块,利用不同尺寸和数量的卷积核来提取和重构特征,显著提升了流量分类的精度;李道全等<sup>[20]</sup> 提出一种基于一维卷积神经网络 (convolutional neural networks, CNN) 的模型,其在  $F_1$ 、召回率等多个指标上均优于经典 ML 方法;王勇等<sup>[21]</sup> 提出了一种将数据进行归一化处理转化为灰度图片作为 CNN 的输入,然后基于 LeNet-5 深度 CNN 构建了一个模型,该模型能够自主学习流量特征,从而实现对网络流量的高效分类的方法;孔镇等<sup>[22]</sup> 则设计了一种新的流量特征扩展方式,通过将详细的特征信息转化为图像格式,并采用 CNN 对代表流数据的图像进行细粒度分类;于帅等<sup>[23]</sup> 另辟蹊径,提出了一种基于深度特征融合的流量分类策略。通过对原始的统计特征进行融合处理,随后应用 CNN 对转换后的灰度图像进行分类;马继焯等<sup>[24]</sup> 通过数据归一化、数据增强和标签加噪等预处理步骤,结合 Res2Net 深度残差神经网络,构建了一种能够容忍噪声的深度神经网络模型,有效提升了流量分类的准确性和鲁棒性。何迎利等<sup>[25]</sup> 针对网络流量具有强烈的非线性和不确定性导致其预测精度不高的问题,在传统时序序列预测模型的基础上设计实现了一种局部上下文信息增强的注意力机制,这种机制能够从微观层面深入解析时间序列的内在结构,显著提升了预测模型对局部变化的感知能力,进而提升网络流量预测精度。

然而使用 CNN 进行流量分类时仍然会存在众多问题,如数据集样本数不足或类别比例不均衡、分类模型构建不合适、超参数选取欠佳、数据预处理不当、训练不充分或过度训练等因素都会影响分类任务的效果。针对上述问题提出改进方法,以期改善 NTC 任务的效果。

## 1 研究内容

针对上述使用 CNN 进行 NTC 任务存在的问题,在现有工作的基础上,提出一种新的基于数据增强与 CNN-Optuna-Attention 改进进行 NTC 的方法,该方法主要贡献如下。

(1) 针对单一 CNN 模型分类效果不佳等问题,本文提出一种在现有 CNN 模型<sup>[21]</sup> 的基础上加以改进的方法,将 CNN 与注意力机制 (attention mecha-

nism, AM)相结合,通过接收全连接层的输出作为输入,计算输入特征的权重并乘以原始特征,实现对关键特征的加强,有助于提高模型对关键信息的捕捉能力,从而提高网络流量分类的准确性。

(2) 针对网络流量数据的特征数无法准确转换为灰度图像的问题,提出一种对特征进行无关字段替换和均值填充的改进方式。根据流量特征构建矩阵时采取将每行数据除缺失值和布尔型值外的特征进行均值计算,并在末尾进行八次均值填充以适应 CNN 的输入特性的方式进行改进。

(3) 针对现有数据集中网络流量类别比例差距过大或不均衡可能导致模型过拟合于高比例类别,以及小比例类别数据难以识别分类的问题,通过对流量数据进行旋转、缩放、镜像<sup>[24]</sup>等操作从而对小比例类别流量数据进行数据增强。该方式可以使各类别数据比例差距缩小。

(4) 针对模型超参数组合的选择不佳导致分类准确率不高的问题,提出一种五折交叉验证划分数据集,并且基于贝叶斯优化进行超参数组合优化的方法<sup>[26]</sup>,该方法可以更充分地模型的未知超

参数组合的性能进行评估,从而比较不同模型参数下的评价结果,进行模型超参数优化,确定最终的模型。

本文方法流程图如图 1 所示。

## 2 数据集

### 2.1 网络流量数据集构建

实验采用两个真实数据集。Moore 数据集<sup>[9]</sup>是由专业用户采用抽样算法得到的 377 526 个网络数据样本,并根据应用类型分为 12 种类别。例如应用类型为 WWW 的具体数据格式如图 2 所示。Moore 流量数据集统计信息如表 1 所示。为了能够更好地验证本文所提方法的性能,同样使用了 ISCX non-

VPN 数据集<sup>[23]</sup>进行分类,使用其中 4 个网络应用类型,分别是视频通话 (video)、语音通话 (audio)、文字聊天 (chat)、IP 语音 (VoIP)。ISCX non-VPN 流量数据集统计信息如表 2 所示。

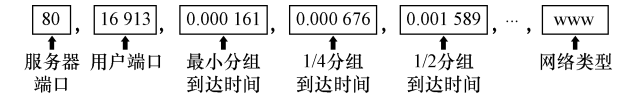


图 2 应用类型为 WWW 的数据格式

Fig. 2 Data format with application type WWW

### 2.2 数据集预处理

为保证 Moore 数据集特征的有效性,将数据集中每一条网络流样本的 249 个特征依次读取,替换特征中的无关特征字段,具体替换形式为

$$Y \rightarrow (1.0) \tag{1}$$

$$N \rightarrow (0.0) \tag{2}$$

$$? \rightarrow (0.0) \tag{3}$$

由于二维 CNN 的输入格式为数值矩阵,因此需要将预处理好的特征进行矩阵化处理,对于 Moore 数据集具有的 249 位特征而言,选择构建一个 16 × 16 的矩阵。由于特征维数少于矩阵元素个数,故而需要进行填充操作。不同于文献[21]采取 0 填充的方式,采取将每行数据除标签、缺失值和布尔型值外的特征进行均值计算,在矩阵的末位进行 8 次均值填充的方式,上述改进不仅可以满足 CNN 的输入特性。而且有利于减少信息损失、增加模型的鲁棒性。

将填充后归一化的矩阵中每个元素作为像素点映射到灰度图像中,Moore 数据集中每种应用类型对应的灰度图如图 3 所示。对于 ISCX non-VPN 数据集而言,采用文献[23]的图像处理方式,将原始特征通过特征融合扩展到 21 616 个特征,规整到 0 ~ 255,按照特征对应一个像素点映射到灰度图像中,ISCX non-VPN 每种应用类型对应的灰度图如图 4 所示。

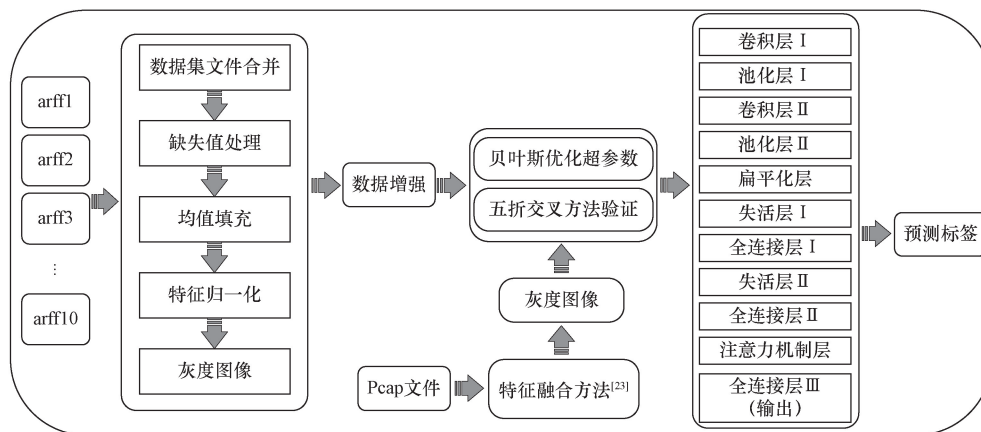


图 1 本文方法流程图

Fig. 1 Flowchart of the methodology in this paper

表1 Moore 流量数据统计信息

Table 1 Statistics of Moore traffic data

类别	数量	应用来源
WWW	328 092	Web browsers, web applications
MAIL	28 567	IMAP, POP, SMTP
FTP-CONTROL	3 054	FTP
FTP-PASV	2 688	FTP
ATTACK	1 793	Port scans, worms, viruses, sql injections
P2P	2 094	Napster, WKazaa, Gnutella, eDonkey, Bit Torrent
DATABASE	2 648	MySQL, dbase, Oracle
FTP-DATA	5 797	FTP
MULTIMEDIA	576	Windows Media Player, Real, iTunes
SERVICES	2 099	X11, DNS, IDENT, LDAP, NTP
INTERACTIVE	110	SSH, TELNET, VNC, Goto-MyPC
GAMES	8	Half-Life
总计	377 526	

表2 ISCX non-VPN 流量数据统计信息

Table 2 Statistics of ISCX non-VPN traffic data

类别	数量	应用来源
语音	500	AIM, ICQ, Facebook, Skype, Hangout, Google
聊天	500	Skype, Hangout, Facebook
视频	500	Skype, Hangout, Facebook
IP 语音	500	Voipbuster
总计	2 000	

### 3 数据增强

Moore 数据集作为经典网络流量数据集虽然被研究人员广泛使用,但仔细观察表 1 可知,Moore 数据集的 377 526 个网络流量数据样本中,INTERACTIVE 的类别样本数在整个数据集中只有 110 个,占比仅有 0.028%,GAMES 类别样本数则更少,仅有 8 个,占比 0.002%。

#### 3.1 ImageDataGenerator 生成器

为了防止 Moore 数据集中网络流量类别比例差距过大或不均衡可能导致模型过拟合于高比例类别,以及难以识别分类小比例类别数据的问题发生,

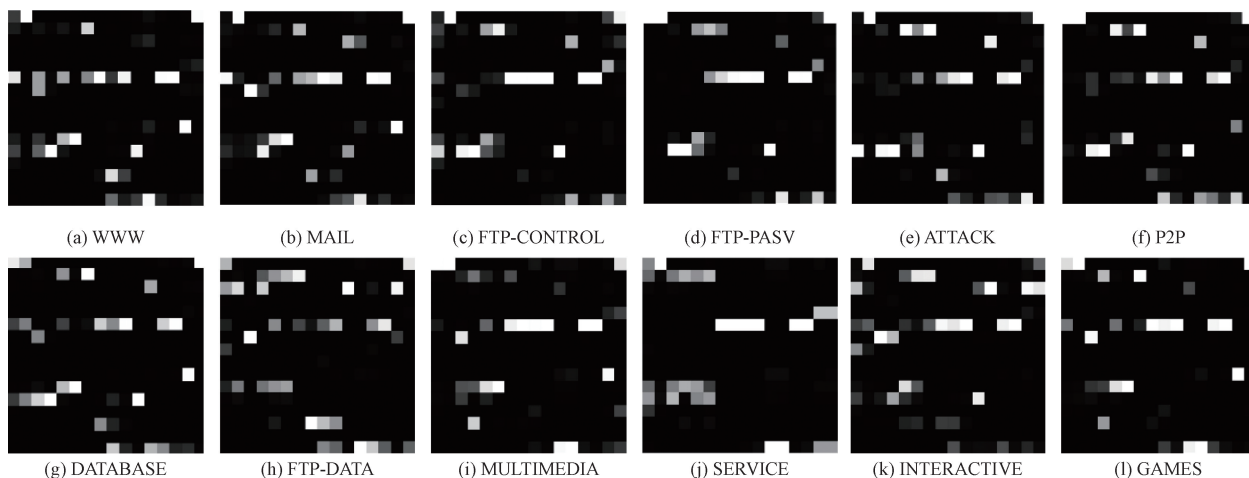


图3 Moore 流量数据对应类别的灰度图

Fig. 3 Grayscale graph of Moore traffic data corresponding to the category

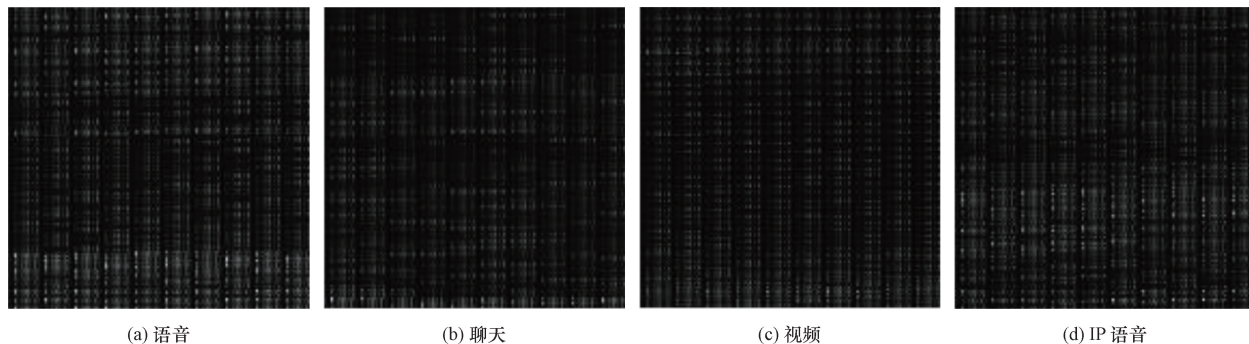


图4 ISCX non-VPN 流量数据对应类别的灰度图

Fig. 4 Grayscale graph of ISCX non-VPN traffic data corresponding to the category

选择使用 ImageDataGenerator 进行实验设计, ImageDataGenerator 作为 Keras 深度学习框架中的一个图像数据生成器,它在深度学习中常用于数据增强。ImageDataGenerator 通过随机变换、裁剪、缩放、旋转、平移等操作来增加图像数据的多样性,从而扩充数据量。这种方法可以帮助模型更好地泛化到不同的数据分布,提高模型的性能和鲁棒性。

### 3.2 参数设计

设计的 ImageDataGenerator 生成器的详细参数如表 3 所示。通过对数据集中小比例类别进行旋转、平移、翻转、剪切、缩放操作从而达到扩充数据集的效果。

表 3 ImageDataGenerator 参数设置

Table 3 Parameter settings for ImageDataGenerator

参数	设置值	具体描述
rotation_range	20	随机旋转的角度范围,以度为单位,设置为 $-20^{\circ} \sim +20^{\circ}$
width_shift_range	0.1	随机水平平移的宽度范围,设置为 10%
height_shift_range	0.1	随机垂直平移的高度范围,设置为 10%
horizontal_flip	True	是否随机水平翻转,设置为 True 表示会进行水平翻转
vertical_flip	True	是否随机垂直翻转,设置为 True 表示会进行垂直翻转
shear_range	0.2	随机剪切变换的角度范围,设置为 $-0.2 \sim +0.2$ rad
zoom_range	0.2	随机缩放的范围,设置为原始尺寸的 80% ~ 120%
fill_mode	"nearest"	当进行变换时,超出图像边界的像素填充方式,设置为"nearest"表示最近邻插值

### 3.3 实验结果

通过对 Moore 数据集进行分析,本文中使用的表 3 设计的 ImageDataGenerator 对 INTERACTIVE 和 GAMES 类别进行个性化数据增强,具体样本数量变化如表 4 所示,对 INTERACTIVE 和 GAMES 类别进行数据增强的图像如图 5 所示,该图展示了两种类别数据增强后的图像。

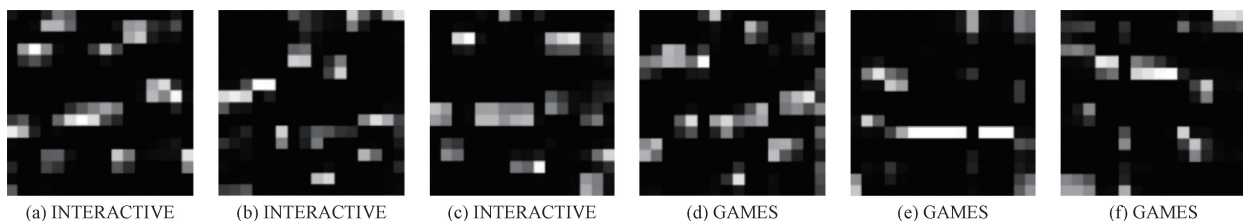


图 5 数据增强生成图像

Fig. 5 Data augmentation generates images

表 4 Moore 数据集的样本数量变化

Table 4 Changes in the sample size of the Moore dataset

类别	原样本数量	数据增强后样本数量
INTERACTIVE	110	3 410
GAMES	8	3 208
总计	377 526	384 026

## 4 超参数组合优化

针对超参数组合选择不当可能会导致分类模型出现欠拟合或过拟合的问题,提出了一种五折交叉验证与基于贝叶斯优化的 Optuna 相结合的超参数组合优化方式<sup>[26]</sup>。通过使用 Optuna 生成超参数组合,利用五折交叉验证方法进行模型训练从而优化超参数组合。

### 4.1 五折交叉验证方法

使用五折交叉验证对模型超参数组合进行优化,其核心思想如图 6 所示,将原始训练集划分为 5 份进行 5 次实验,选择其中 4 份数据用于训练,剩余 1 份用于评价效果,重复 5 次后保证每份数据都扮演过训练集和验证集的角色,验证集在训练过程中用于检验模型的训练情况,确定合适的超参数组合。最后得到模型的 5 次评价结果,取均值作为该超参数组合下的模型的最终评价结果,从而可以比较不同超参数下模型的评价结果,进行模型超参数优化。

### 4.2 基于贝叶斯优化的超参数优化框架 Optuna

相比于以往研究中超参数优化使用网格搜索算法遍历给定的超参数组合来优化模型,使用 Optuna 进行超参数组合的选择。Optuna 方法较网格搜索方法的优势如表 5 所示。

网格搜索方法需要研究者提前给出需要超参数遍历的空间,获取的超参数取值都是在给定的准确值内进行迭代选择的,对于提出的基于数据增强与 AM 改进的 CNN 分类模型来讲,其超参数组合是无法准确手动定义的。且需确定学习率、丢弃率、批次大小、迭代次数的复杂超参数组合,并不是简单、单一的超参数,使用网格搜索需要经过嵌套循环遍历,无疑增加了运行时间。因此选择 Optuna 方法进行超参数优化。

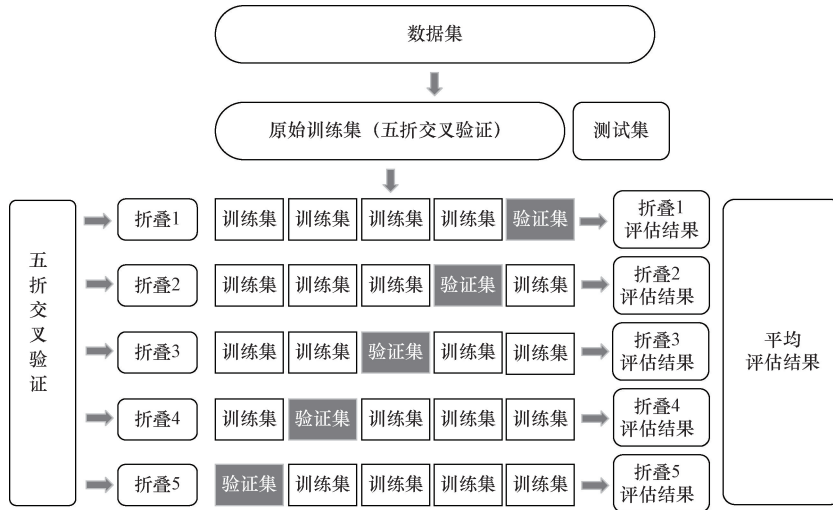


图6 五折交叉验证方式图

Fig. 6 Five-fold cross-validation graph

表5 Optuna 与网格搜索方法对比  
Table 5 Comparison between pptuna and grid search methods

参数	Optuna 方法	网格搜索方法
灵活性	高	低
超参数搜索范围	连续、离散参数等自动定义	手动定义参数组合
搜索方式	TPE、随即搜索等	穷举迭代搜索
收敛速度	快	慢
使用场景	复杂超参数空间	简单超参数空间
适用性	多样化模型优化	简单模型调优

### 4.3 超参数优化分析与结果

采用 Optuna 与五折交叉验证相结合的方式进入超参数组合优化,首先使用 Optuna 定义超参数组合的范围,在范围内限制 30 次超参数组合进行实验,并且 30 种不同的超参数组合都经过五折交叉验证来获取其平均准确率。为了防止模型过拟合,本文使用早停机制防止过度迭代,利用 Optuna 指定的学习率、丢弃率、批次大小的搜索空间如表 6 所示。

由实验所得 30 种超参数组合的准确率得分可知,采用学习率、丢弃率、迭代次数、批次大小设置值为

$$\text{learning\_rate} = 0.000484496235947 \quad (4)$$

$$\text{dropout\_rate} = 0.062456633270408 \quad (5)$$

$$\text{batch\_size} = 256 \quad (6)$$

防止分类模型的超参数选择不佳可能会影响分类准确率的问题发生。

表6 超参数的搜索空间

Table 6 Search space for hyperparameters

超参数	搜索空间
学习率	0.0001~0.1
丢弃率	0.0~0.3
批次大小	64/128/256
迭代次数	200(使用早停策略)

## 5 本文 CNN-Optuna-Attention 模型

### 5.1 CNN 模型维度选取

CNN 作为 DL 的重要领域之一,其广泛运用于计算机视觉以及自然语言处理方面。一维 CNN 适用于序列数据或语言数据,二维 CNN 擅长处理图像或音频频谱图等数据,三维 CNN 适用于视频或体积图像等数据。根据 Moore 和 ISCX non-VPN 的数据集可知,针对 Moore 和 ISCX non-VPN 数据集转换的二维图像数据,使用二维 CNN 进行分类是最合适的选择。因此本文模型在二维 CNN 的基础上进行研究。

然而,传统的 CNN 模型在处理复杂图像时,往往难以捕捉到图像中的关键信息。AM 允许模型在处理信息时,能够动态地聚焦于输入数据的某些部分,同时忽略其他不相关的信息。这种机制模拟了人类处理复杂信息时的注意力分配过程,有助于模型更好地捕捉关键特征,提高任务的准确性。本文通过引入 AM 来增强模型对关键特征的关注,提高分类性能。

### 5.2 CNN-Optuna-Attention 模型架构

本文中提出的模型采用序贯模型(sequential)结构,主要包括卷积层、池化层、全连接层以及自定义的 AM 层。CNN-Optuna-Attention 模型具体架构如图 7 所示。CNN-Optuna-Attention 模型网络参数结构如表 7 所示。

(1)卷积层。卷积层作为 CNN 的核心模块,用于提取输入图像的局部特征。本文模型使用两个卷积层,第一个卷积层有 8 个  $3 \times 3$  的卷积核,输入形状为  $16 \times 16 \times 1$ ,匹配二维图像的格式,第二个卷积层有 16 个  $3 \times 3$  的卷积核。每个卷积层后都使用

表 7 CNN-Optuna-Attention 模型网络参数结构  
Table 7 The network structure of the CNN-optuna-attention model

层类型	输入尺寸	卷积核/ 池化大小	卷积核/ 神经元数	步长	填充 方式	输出 尺寸	激活 函数
卷积层 1	(batch_size, 16, 16, 1)	(3, 3)	8	1	'same'	(batch_size, 16, 16, 8)	ReLU
池化层 1	(batch_size, 16, 16, 8)	(2, 2)	—	2	'same'	(batch_size, 8, 8, 8)	—
卷积层 2	(batch_size, 8, 8, 8)	(3, 3)	16	1	'same'	(batch_size, 8, 8, 16)	ReLU
池化层 2	(batch_size, 8, 8, 16)	(2, 2)	—	2	'same'	(batch_size, 4, 4, 16)	—
扁平化层	(batch_size, 4, 4, 16)	—	—	—	—	(batch_size, 256)	—
失活层 1	(batch_size, 256)	—	—	—	—	(batch_size, 256)	—
全连接层 1	(batch_size, 256)	—	256	—	—	(batch_size, 256)	ReLU
失活层 2	(batch_size, 256)	—	—	—	—	(batch_size, 256)	—
全连接层 2	(batch_size, 256)	—	128	—	—	(batch_size, 128)	ReLU
注意力机制层	(batch_size, 128)	—	—	—	—	(batch_size, 128)	—
全连接层 3 (输出层)	(batch_size, 128)	—	12/4	—	—	(batch_size, 12/4)	softmax

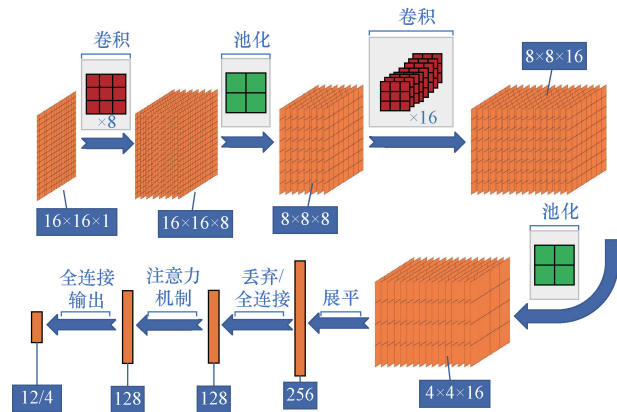


图 7 CNN-Optuna-Attention 模型架构图

Fig. 7 The architecture of the CNN-optuna-attention model

了 ReLU 激活函数, 以增强模型的非线性表达能力。特征图大小计算方式如下。

假设输入体积大小为  $H_1 \times W_1 \times D_1$  (高度  $\times$  宽度  $\times$  通道数), 卷积核数量为  $K$ , 卷积核大小为  $F \times F$ , 步长为  $S$ , 零填充大小为  $P$  可得

$$H_2 = \frac{H_1 - F + 2P}{S} + 1 \quad (7)$$

$$W_2 = \frac{W_1 - F + 2P}{S} + 1 \quad (8)$$

$$D_2 = K \quad (9)$$

式中:  $H_2$  为输出高度;  $W_2$  为输出宽度;  $D_2$  为输出通道数。

(2) 池化层。池化层用于对卷积层输出的特征图进行下采样, 以减少计算量并提取更重要的特征。本文模型使用两个最大池化层, 池化窗口大小为  $2 \times 2$ , 步长为 2。

(3) 全连接层。连接层用于将卷积层和池化层提取的特征整合起来并进行分类。每个全连接层后都使用了 ReLU 激活函数, 并添加了 Dropout 层以

防止过拟合。

(4) 自定义 AM 层。注意力层用于增强模型对关键特征的关注。该层应用于第二个全连接层之后, 接受全连接层的输出作为输入, 通过计算输入特征的权重并乘以原始特征, 实现对关键特征的加强。本文 AM 层具体实现细节如下。

设输入张量为  $X$ , 且

$$X \in \mathbf{R}^{N \times T \times D} \quad (10)$$

式(10)中:  $N$  为批次大小 (batch\_size);  $T$  为序列长度;  $D$  为特征维度。

① 权重和偏置初始化:

权重矩阵为  $W$ , 且

$$W \in \mathbf{R}^{D \times 1} \quad (11)$$

偏置向量为  $b$ , 且

$$b \in \mathbf{R}^{T \times 1} \quad (12)$$

② 计算能量:

对于输入张量  $X$  中的每一个元素  $x_{ntd}$  (其中  $n$  表示批次索引,  $t$  表示时间步索引,  $d$  表示特征索引), 其对应的能量  $e_{nt}$  计算公式为

$$e_{nt} \in \tanh\left(\sum_{d=1}^D x_{ntd} w_d + b_t\right) \quad (13)$$

式(13)中:  $w_d$  为权重矩阵  $W$  的第  $d$  个元素;  $b_t$  为偏置向量  $b$  的第  $t$  个元素。

③ 计算注意力权重:

注意力权重  $a_{nt}$  可以通过对能量  $e_{nt}$  应用 softmax 函数得到, 即

$$a_{nt} = \frac{\exp(e_{nt})}{\sum_{k=1}^T \exp(e_{nk})} \quad (14)$$

④ 应用注意力权重并求和:

最终的输出  $y_n$  是输入张量  $X$  与注意力权重  $a$  的加权和, 沿着时间步  $t$  进行求和, 即

$$y_n = \sum_{t=1}^T a_{nt} \left( \sum_{d=1}^D x_{ntd} \right) \quad (15)$$

式(15)中: $y_n$ 为批次中第 $n$ 个样本的输出。

(5)输出层。输出层用于输出最终的分类结果。本文模型使用 softmax 激活函数,输出层有 12/4 个神经元,对应于 Moore 和 ISCX non-VPN 数据集的流量类别数,将全连接层的输出转换为概率分布,从而进行多分类任务。

## 6 模型实验与结果分析

### 6.1 实验环境

实验设备方面,实验采用配备 CPU 为 i5-8300H,4 核 2.30 GHz,内存为 8 GB,GPU 为 NVIDIA GeForce GTX 1050 的计算机和 Windows 11 的操作系统。编程方面,选择 Python 作为编程语言并使用 PyCharm 作为开发工具。在数据处理方面,采用 Numpy 和 Pandas 库,深度学习框架选择 TensorFlow。

### 6.2 评估指标

为了评估本文所提方法的分类性能,本文中使用了准确率  $A$ 、精确率  $P$  以及  $F_1$  分数这 3 个指标进行性能评估。将预测结果用符号定义为真正例 (TP)、假正例 (FP)、假反例 (FN)、真反例 (TN),如表 8 所示。

表 8 计算评估指标的混淆矩阵

Table 8 Confusion matrix for calculating assessment metrics

真实值	预测值	
	Negative	Positive
Positive	False Negative (FN)	True Positive (TP)
Negative	True Negative (TN)	False Positive (FP)

(1)准确率  $A$ :准确率作为分类问题最基本的评估指标,其表示模型正确预测的样本数量与所有预测样本数量之间的比率,该指标用于评估模型的整体性能,准确率越高代表分类效果越好,其表达式为

$$A = \frac{TP + TN}{TP + TN + FP + FN} \quad (16)$$

(2)精确率  $P$ :精确率能够有效评估分类模型对每一个类别的分类好坏,其表示在所有被模型判定为正例的样本中,实际上的确是正例的样本所占的比例,其表达式为

$$P = \frac{TP}{TP + FP} \quad (17)$$

(3) $F_1$  分数: $F_1$  分数是精确率  $P$  和召回率  $R$  的调和平均数,其表示为

$$F_1 = \frac{2PR}{P + R} \quad (18)$$

$$R = \frac{TP}{TP + FN} \quad (19)$$

### 6.3 实验结果及分析

将本文设计的方法在 Moore 数据集上的实验结果与 BP 神经网络方法、文献[21]分类方法、使用默认超参数且没有引入 AM 机制的本文模型方法(简称对比方法 1),以及基于条件生成对抗网络<sup>[27]</sup>实现数据增强且引入本文 AM 机制的方法(简称对比方法 2)做对比。将本文方法在 ISCX non-VPN 数据集上的实验结果与文献[15,22-23]的方法作对比。通过表 9 给出的 Moore 数据集下 5 种方法的准确率对比可知,本文所提方法在准确率方面要优于其他对比方法。图 8 和图 9 分别给出了不同方法在 Moore 数据集上的精确率和  $F_1$  分数的对比柱状图表,经仔细分析可知本文所提方法的改进模型要较优于 BP 神经网络和文献[21]方法,并且通过本文所以方法与对比方法 1 相比,验证了本文所提方法在超参数优化以及五折交叉验证方法上的有效性,通过本文所提方法与对比方法 2 相比,验证了本文所提方法在数据增强方面的优势。表 10 给出了 ISCX non-VPN 数据集下 4 种方法的准确率对比。可以看出本文所提方法与模型在准确率方面要优于文献[15,22-23]。图 10 和图 11 分别给出了不同方法在 ISCX non-VPN 数据集上的精确率和  $F_1$  分数的对比柱状图表。由图表可知,本文所提模型与优化方法的精确率与  $F_1$  分数优于文献[5,22]的方法,同是采用相同特征融合方式处理的 ISCX non-VPN 数据集,与文献[23]相比在聊天类别方面略低,但在视频、语音、IP 语音均比其他方法效果更好。

表 9 Moore 数据集下采用不同方法的准确率

Table 9 The accuracy of different methods under the Moore dataset

方法	准确率
BP 神经网络	0.993 1
文献[21]的方法	0.993 0
对比方法 1	0.992 7
对比方法 2	0.992 5
本文方法	0.995 9

表 10 ISCX non-VPN 数据集下采用不同方法的准确率

Table 10 The accuracy of different methods under the ISCX non-VPN dataset

方法	准确率
文献[15]的方法	0.977 2
文献[22]的方法	0.985 8
文献[23]的方法	0.996 5
本文方法	0.997 5

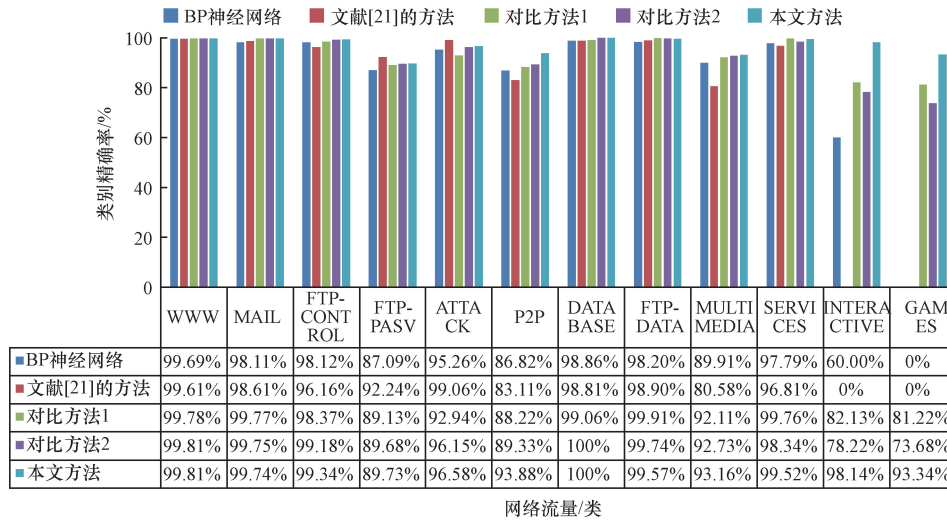


图8 Moore数据集下不同方法的类别精确率对比图表

Fig. 8 Comparison chart of the categorical precision of different methods in the Moore dataset

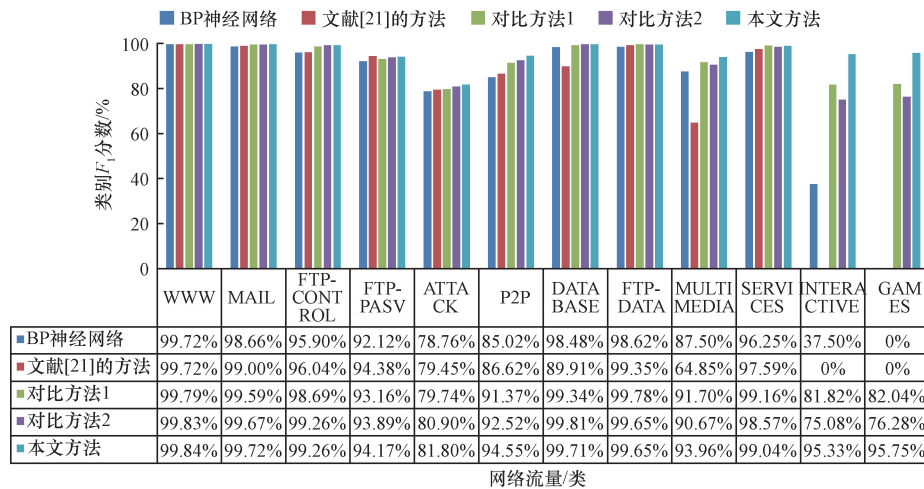


图9 Moore数据集下不同方法的类别  $F_1$  分数对比图表

Fig. 9 Comparison chart of  $F_1$  scores of different methods in the Moore dataset

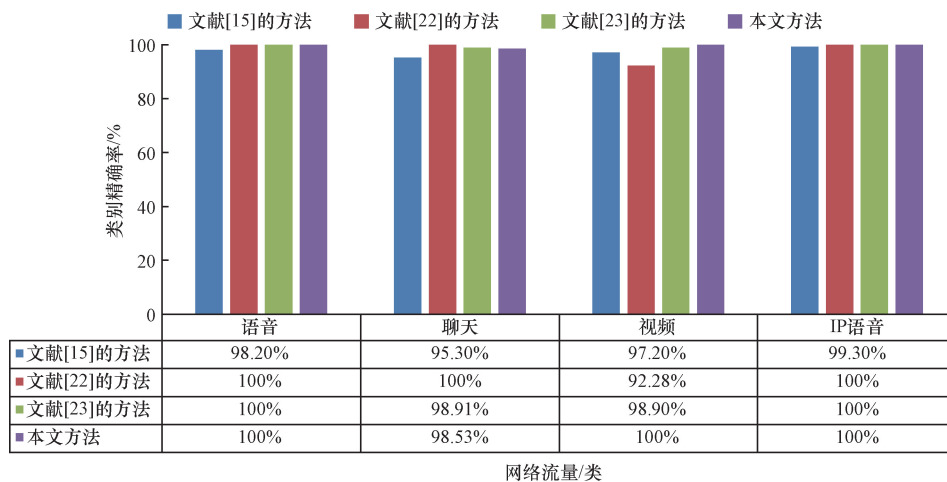


图10 ISCX non-VPN数据集下不同方法的类别精确率对比图表

Fig. 10 Comparison chart of the categorical precision of different methods in the ISCX non-VPN dataset

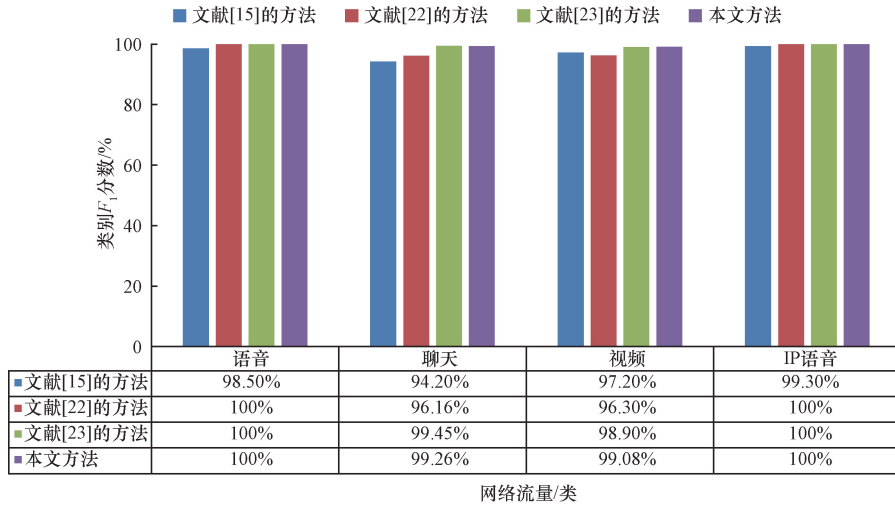


图 11 ISCX non-VPN 数据集下不同方法的类别  $F_1$  分数对比图表

Fig. 11 Comparison chart of  $F_1$  scores of different methods in the ISCX non-VPN dataset

为了更为直观地展示本文所提方法的分类性能,图 12 给出了本文所提方法在 Moore 与 ISCX non-VPN 数据集上的混淆矩阵。通过以上两个数据集的模型测试结果对比。进一步验证了本文中所使用的数据增强方法、超参数优化方法和引入 AM 改进

的 CNN 分类模型较其他方法的有效性。

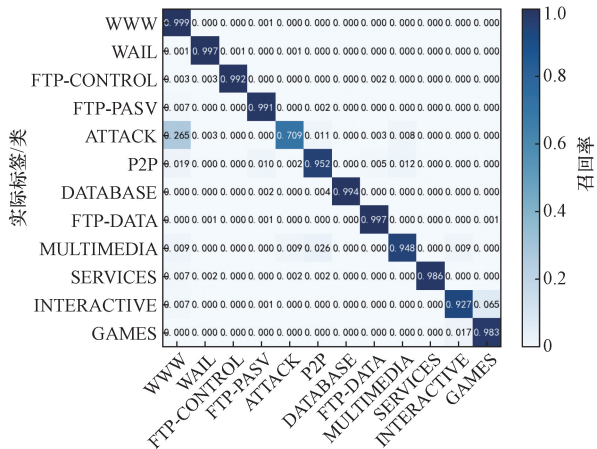
## 7 结论

针对当前使用 DL 进行 NTC 任务存在的问题,提出了一种基于数据增强技术和 CNN-Optuna-Attention 改进的网络流量分类方法,该方法首先对 Moore 与 ISCX non-VPN 数据集进行恰当的预处理,然后使用不同的方式转换为二维图像,通过分析 Moore 数据集的数据量可知其类别分布存在很大的问题,因此采取 ImageDataGenerator 对 Moore 数据集进行个性化的数据增强,扩充了 GAMES 和 INTERACTIVE 的类别样本。其次使用五折交叉验证和 Optuna 对分类模型的超参数组合进行优化,选择出适合分类模型的超参数,最后将上述两个数据集转换出的图像、超参数组合提供给本文经 AM 改进的 CNN 分类模型,实现高效准确的分类。

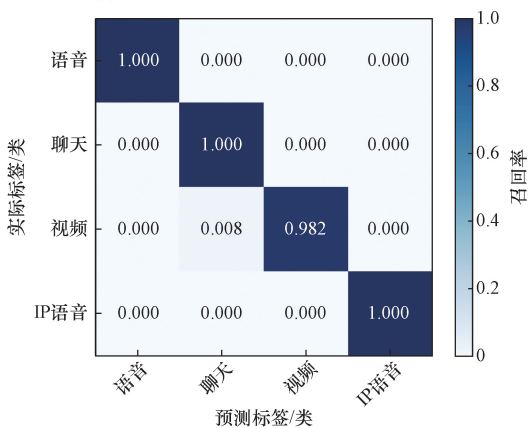
通过与其他文献的研究方法对比可知,本文所提方法在准确率、精确率、 $F_1$  分数上都存在着提高。在未来的工作中,将根据以下几个方面进行深入研究:①适当改变 CNN 分类模型的结构,比较网络流量分类的效果。②尝试利用其他数据增强方法对 Moore 数据集进行数据扩充,比较数据增强后生成图像的效果好坏。③在进行超参数优化时将 Optuna 框架中的 n\_trials 的取值适当的增大,增加超参数组合的种类,将更多模型中需要考虑的超参数加入到超参数优化的范围内,通过该方式获取分类模型的多种类型超参数组合的结果。

## 参 考 文 献

[1] 中国互联网信息中心(CNNIC). 第 53 次《中国互联网络发展状况统计报告》[EB/OL]. (2024-03-22) [2024-06-18].



(a) 本文方法在 Moore 数据集上的混淆矩阵



(b) 本文方法在 ISCX non-VPN 数据集上的混淆矩阵

图 12 本文方法的混淆矩阵

Fig. 12 The confusion matrix of the proposed method

- https://www.cnnic.net.cn/n4/2024/0321/c208-10962.html. China Internet Information Center (CNNIC). The 53rd statistical report on the internet development in China[EB/OL]. (2024-03-22)[2024-06-18]https://www.cnnic.net.cn/n4/2024/0321/c208-10962.html.
- [2] 郑春艳. 浅谈计算机网络发展现状[J]. 现代计算机, 2022, 28(21): 77-79.  
Zheng Chunyan. Introduction to the current status of computer network development[J]. Modern Computer, 2022, 28(21): 77-79.
- [3] Azab A, Khasawneh M, Alrabaee S, et al. Network traffic classification: techniques, datasets, and challenges[J]. Digital Communications and Networks, 2022, 10(3): 676-692.
- [4] Munarko A. Determinants of service quality: efficiency and workspace[J]. AKADEMIK: Jurnal Mahasiswa Humanis, 2022, 2(3): 136-144.
- [5] 任成刚. 大数据下的计算机网络安全技术分析[J]. 网络安全和信息化, 2023, (11): 148-150.  
Ren Chenggang. Analysis of computer network security technology under big data[J]. Network Security and Informatization, 2023, (11): 148-150.
- [6] Zhao J, Jing X, Yan Z, et al. Network traffic classification for data fusion: a survey[J]. Information Fusion, 2021, 72: 22-47.
- [7] 杨宇, 闫钰, 申芳, 等. 基于机器和深度学习的入侵检测综述[J]. 科学技术与工程, 2023, 23(18): 7607-7621.  
Yang Yu, Yan Yu, Shen Fang, et al. A review of intrusion detection based on machine and deep learning[J]. Science Technology and Engineering, 2023, 23(18): 7607-7621.
- [8] Shahraki A, Abbasi M, Taherkordi A, et al. Active learning for network traffic classification: a technical study[J]. IEEE Transactions on Cognitive Communications and Networking, 2021, 8(1): 422-439.
- [9] Moore A W, Zuev D. Internet traffic classification using Bayesian analysis techniques [C]//Proceedings of the 2005 ACM Sigmetrics International Conference on Measurement and Modeling of Computer Systems. New York: ACM, 2005: 50-60.
- [10] Madhukar A, Williamson C. A longitudinal study of P2P traffic classification[C]//14th IEEE International Symposium on Modeling, Analysis, and Simulation. Monterey: IEEE, 2006: 179-188.
- [11] Gringoli F, Salgarelli L, Dusi M, et al. GT: picking up the truth from the ground for internet traffic[J]. Computer Communication Review, 2009, 39: 12-18.
- [12] 罗冬梅. 网络协议流不平衡环境下基于机器学习算法的在线流量分类方法[J]. 科学技术与工程, 2017, 17(28): 103-107.  
Luo Dongmei. Online traffic classification method based on machine learning algorithm in network protocol flow imbalance environment [J]. Science Technology and Engineering, 2017, 17(28): 103-107.
- [13] 王瑞敏, 汤云凯, 郝高鑫, 等. 基于机器学习的网络流量分类研究[J]. 山西电子技术, 2024(2): 104-106, 119.  
Wang Ruimin, Tang Yunkai, Hao Gaoxin, et al. Research on network traffic classification based on machine learning[J]. Shanxi Electronic Technology, 2024(2): 104-106, 119.
- [14] 于治平, 刘彩霞, 刘树新, 等. 基于机器学习的网络流量分类综述[J]. 信息工程大学学报, 2023, 24(4): 447-453, 483.  
Yu Zhiping, Liu Caixia, Liu Shuxin, et al. A review of network traffic classification based on machine learning[J]. Journal of Information Engineering University, 2023, 24(4): 447-453, 483.
- [15] Perna G. Online classification of RTC traffic [C]//2021 IEEE 18th Annual Consumer Communications & Networking Conference (CCNC). Las Vegas: IEEE, 2021: 1-6.
- [16] 张超群, 易云恒, 周文娟, 等. 基于深度学习与数据增强技术的小样本岩石分类[J]. 科学技术与工程, 2022, 22(33): 14786-14794.  
Zhang Chaoqun, Yi Yunheng, Zhou Wenjuan, et al. Small-sample rock classification based on deep learning and data enhancement techniques[J]. Science Technology and Engineering, 2022, 22(33): 14786-14794.
- [17] Matsuo Y, Lecun Y, Sahani M, et al. Deep learning, reinforcement learning, and world models [J]. Neural Networks, 2022, 152: 267-275.
- [18] 顾玥, 李丹, 高凯辉. 基于机器学习和深度学习的网络流量分类研究[J]. 电信科学, 2021, 37(3): 105-113.  
Gu Yue, Li Dan, Gao Kaihui. Research on network traffic classification based on machine learning and deep learning[J]. Telecommunications Science, 2021, 37(3): 105-113.
- [19] 宋继红, 王桐桐. 基于特征融合卷积神经网络的加密流量分类方法[J]. 信息与电脑(理论版), 2021, 33(2): 49-51.  
Song Jihong, Wang Tongtong. An encrypted traffic classification method based on feature fusion convolutional neural network[J]. Information and Computer Science (Theoretical Edition), 2021, 33(2): 49-51.
- [20] 李道全, 王雪, 于波, 等. 基于一维卷积神经网络的网络流量分类方法[J]. 计算机工程与应用, 2020, 56(3): 94-99.  
Li Daoquan, Wang Xue, Yu Bo, et al. A network traffic classification method based on one-dimensional convolutional neural network[J]. Computer Engineering and Applications, 2020, 56(3): 94-99.
- [21] 王勇, 周慧怡, 俸皓, 等. 基于深度卷积神经网络的网络流量分类方法[J]. 通信学报, 2018, 39(1): 14-23.  
Wang Y, Zhou Huiyi, Feng Hao, et al. Network traffic classification method based on deep convolutional neural network[J]. Journal on Communications, 2018, 39(1): 14-23.
- [22] 孔镇, 董育宁. 一种基于深度学习的网络流量细粒度分类方法[J]. 南京邮电大学学报(自然科学版), 2021, 41(3): 100-108.  
Kong Zhen, Dong Yuning. A deep learning-based approach for fine-grained classification of network traffic[J]. Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition), 2021, 41(3): 100-108.
- [23] 于帅, 董育宁, 邱晓晖. 一种基于深度特征融合的网络流量分类方法[J]. 南京邮电大学学报(自然科学版), 2022, 42(3): 82-89.  
Yu Shuai, Dong Yuning, Qiu Xiaohui. A network traffic classification method based on deep feature fusion[J]. Journal of Nanjing

- University of Posts and Telecommunications (Natural Science Edition), 2022, 42(3): 82-89.
- [24] 马继焯, 朱国胜, 卫操, 等. 一种噪声容忍的网络流量分类方法[J]. 计算机科学, 2023, 50(S2): 775-781.
- Ma Jiye, Zhu Guosheng, Wei Cao, et al. A noise-tolerant network traffic classification method [J]. Computer Science, 2023, 50 (S2): 775-781.
- [25] 何迎利, 胡光宇, 张浩, 等. 基于局部信息增强注意力机制的网络流量预测[J]. 科学技术与工程, 2023, 23(30): 13014-13022.
- He Yingli, Hu Guangyu, Zhang Hao, et al. Network traffic prediction based on local information-enhanced attention mechanism [J]. Science Technology and Engineering, 2023, 23(30): 13014-13022.
- [26] 李艳辉, 王衍萌. 基于数据增强技术与 CNN-BiLSTM-Attention 的油田注水流量预测及效果[J]. 科学技术与工程, 2023, 23(32): 13896-13902.
- Li Yanhui, Wang Yanmeng. Oilfield water injection flow prediction and effect based on data enhancement technique with CNN-BiLSTM-Attention [J]. Science Technology and Engineering, 2023, 23(32): 13896-13902.
- [27] 金麒旻. 生成对抗网络的质量提升技术研究[D]. 北京: 北京邮电大学, 2023.
- Jin Qimin. Research on quality improvement technology of generative adversarial network [D]. Beijing: Beijing University of Posts and Telecommunications, 2023.