

M 现代计算机

XIANDAI JISUANJI

第30卷第20期（总第812期）

半月刊（1984年创刊）

2024年10月25日出版

主管单位 中山大学
主办单位 广州中山大学出版社有限公司
出版单位 广东现代计算机杂志社有限公司
发行 广东省报刊发行局（全国公开发行）
印刷 广州市友盛彩印有限公司
社长 黄少伟
主编 石玉珍
编委 邹岚萍 熊锡源 李文 石玉珍 梁嘉璐
地址 广州市海珠区新港西路135号
中山大学内（510275）
电话 020-84112089（编辑部）
网址 www.moderncomputer.cn
电子邮箱 tougao@moderncomputer.cn

ISSN 1007-1423
CN 44-1415/TP

邮发代码：46-121
定价：30.00元



邮局订刊二维码



现代计算机
官方网站二维码

ISSN 1007-1423



9 771007 142246

20

M 现代计算机

第30卷 第20期（总第812期）

2024年10月

2024年10月 第30卷

第20期（总第812期）

M

ISSN 1007-1423
CN 44-1415/TP

MODERN COMPUTER

现代计算机



中山大学出版社 主办

中国期刊数据库CNKI全文收录期刊
中国学术期刊（光盘版）收录期刊
中文科技期刊数据库全文收录期刊
中国核心期刊（遴选）数据库收录期刊
中国学术期刊综合评价数据库收录期刊

- ◆ 研究与开发：计算机发展和软、硬件开发的理论研究
- ◆ 图形图像：重点为与图形图像相关的理论及实践研究
- ◆ 开发案例：基于某方面的计算机开发案例研究与分析
- ◆ 实践与经验：计算机应用的实例及心得

版权声明

1. 本刊版权属于杂志社所有，其他报刊或网站如需转载，须经本刊同意，注明转载自本刊并付作者稿酬。
2. 本刊来稿恕不退还，请自留底稿。请勿一稿多投。来稿文责自负，严禁抄袭。对侵犯他人版权或其他权利的稿件，本刊概不承担连带责任。
3. 对所投稿件，本刊编辑有权根据刊物的需要进行删改或调整。
4. 凡是刊登在本刊的稿件，即表示作者同意稿件在《现代计算机》网站、中国期刊数据库CNKI、中国学术期刊（光盘版）、中文科技期刊数据库、中国科技期刊（遴选）数据库、中国学术期刊综合评价数据库等媒体发布。

目次

研究与开发

- 基于改进YOLOv8的轻量化小尺度行人和非机动车目标检测算法 帅勃宇, 张雅丽 (1)
- 植物染色质可及性深度学习预测研究综述 吴昭瑞, 王舒宇, 赵玉茗, 于洋, 魏金寒 (8)
- 基于改进YOLOv8的飞鸟检测算法 陈倩, 卢扬, 邵飞翔, 李师艳 (14)
- 轻量化改进YOLOv5s无人机遥感图像识别 杨磊, 田立勤, 吴骏一, 彭凯远 (20)
- 基于Faster R-CNN的低照度人脸检测方法研究 岑锐强, 冯广 (25)
- 基于深度学习的蚕病检测算法设计 曾沛杰, 周卫, 陈金良 (31)
- 基于图神经网络的药物相互作用预测方法研究综述
..... 郭文博, 龙伟, 蒋林华, 叶文标, 胡灵犀, 彭司华 (36)
- 基于机器视觉的圆形精密工件尺寸检测 于佳昊, 刘嘉承, 王连锴, 杨展, 宋德 (41)

图形图像

- SOIQA: 一种图像超分领域的实用型图像质量评价方法 法静怡, 严广宇 (45)

实践与经验

- DDoS攻击的发展与检测技术研究 马立鑫, 薛占双, 刘海燕 (52)
- 基于RFC优化算法的报文数据快速模式匹配 王瑞钦, 谭皇 (57)
- 通过集成域对抗训练和最大平均差异的故障诊断 张宇, 张公政 (63)
- 基于在线评测的主从式数据库信息快速同步方法研究 王位明 (68)
- 高并发Web应用系统性能优化研究 欧阳习彪, 徐宝林 (72)

开发案例

- 基于MLP的电动汽车充电桩共享推荐模型 王伟, 游凤芹, 徐晖 (76)
- “拾忆长征”红色旅游小程序设计与开发 沈楨, 李丽娟, 徐妍 (82)
- 情感陪伴智能桌面萌宠系统的研究与设计 李明亮, 谢桂芳, 刘灵丽, 李昱强 (87)
- 基于云计算的医药信息集成共享平台设计与应用 范颖颖 (92)
- 基于B/S架构的在线实验系统的设计与实现
..... 魏武华, 黄业坤, 于豪, 李森鑫, 郝茵茵, 梁静, 刘少华 (97)
- 手机端自习室预约系统的设计与实现 王菲, 张凤涛, 热依丽罕·图尔荪, 朱芳瑶 (101)
- 智能疫苗接种系统的设计 张琳婧, 李庆玲 (106)
- 基于自动跟随的智能购物车设计与研究 李文豪 (110)
- 基于动作识别的智能家居控制系统研究 方亮, 江权威, 黎思超, 刘伟 (114)
- 基于深度学习的中式菜品检测识别系统 李娟 (117)

Modern Computer

(Vol. 30, No. 20; Oct. 25, 2024)

CONTENTS

Research and Development

A lightweight small-scale pedestrian and non-motorised vehicle target detection algorithm based on improved YOLOv8	(1)
A review of deep learning prediction of chromatin accessibility in plants	(8)
Flying bird detection algorithm based on an improved YOLOv8	(14)
Application to UAV remote sensing image recognition with a lightweight improved model of YOLOv5s	(20)
Research on low-light face detection method based on Faster R-CNN	(25)
Design of silkworm disease detection algorithm based on deep learning	(31)
A review of prediction methods of drug interaction based on graph neural network	(36)
Dimension measurement of circular precision components based on machine vision	(41)

Graphic and Image

Stack objective image quality assessment: A practical evaluation method in super resolution task	(45)
--	------

Practice and Experience

A review of the development and detection technologies of DDoS attacks	(52)
Fast packet pattern-match based on optimized RFC algorithm	(57)
Fault diagnosis through integrated domain adversarial training and maximum average difference	(63)
Research on the rapid synchronization method of master-slave database information based on online evaluation	(68)
Research on performance optimization of high-concurrency web application system	(72)

Development Solution

An MLP-based smart recommendation model for electric vehicle charging pile sharing	(76)
Design and development of a red tourism mini program for “recalling the Long March”	(82)
A study and design of emotionally supportive intelligent desktop pet system	(87)
Design and application of medical information integration and sharing platform based on cloud computing	(92)
Design and implementation of online experiment system based on B/S architecture	(97)
Design and implementation of self-study room reservation system on mobile phone	(101)
Design of intelligent vaccination system	(106)
Design and research of smart shopping cart based on auto-following	(110)
Research on smart home control system based on motion recognition	(114)
Chinese food detection recognition system based on deep learning	(117)

研究与开发

文章编号: 1007-1423(2024)20-0001-07

DOI: 10.3969/j.issn.1007-1423.2024.20.001

基于改进YOLOv8的轻量化小尺度行人和非机动车目标检测算法

帅勃宇, 张雅丽*

(中国人民公安大学信息安全学院, 北京 100038)

摘要: 为解决交通路口行人和非机动车治理问题, 针对全景路口监控下行人和非机动车目标小等问题, 提出了一种基于改进YOLOv8的轻量化小尺度行人和非机动车目标检测算法——ACM-YOLO。首先, 提出轻量化的高效AFPN特征融合网络, 替换原始网络中的PAFPN算法, 提高小目标的识别效果; 其次, 基于PConv提出了CWPCConv通道权重部分卷积, 并进一步提出了CWPC2f, 有效降低了模型参数量和计算量; 最后, 使用MPDIoU函数优化网络的边界框损失。结果表明, 相较于YOLOv8m, mAP50在自建数据集、VisDrone2019数据集和CityPerson数据集上分别提高了2.2%、8.1%和3%, 参数量下降了17%, GFLOPs下降2。

关键词: 小目标检测; YOLOv8; 行人和非机动车检测; CWPCConv

0 引言

行人和非机动车几乎覆盖了城市中所有道路区域, 因其数量大、体积小等特点给城市交通管理带来诸多不便。随着交管部门对非机动车监管力度日益增强, 基于道路监控摄像头的非机动车智能在线检测对道路安全和社会发展具有重要意义^[1]。目前对于行人和非机动车的检测多是依赖近距离摄像头拍摄画面, 视野范围有限, 不能覆盖整个路口, 无法解决全局性问题, 给行人和非机动车检测带来了巨大困扰。

为解决上述问题, 本文结合小目标检测的重难点, 在实际交通路口场景下, 面向小尺度行人和非机动车目标, 提出一种基于改进YOLOv8的轻量化小尺度行人和非机动车检测算法——ACM-YOLO。

1 相关工作

目前, 行人和非机动车检测主要有Two-stage和One-stage两类算法。Two-stage网络将定位和分类任务分开, 先提取物体在图像中的位置, 然后再对确定区域进行分类识别。Ren等^[2]

提出了Faster-RCNN, 使用RPN网络生成较好的建议框, 设置不同的anchor, 引入兴趣池化固定长度输出。Xu等^[3]针对2D边界框的不足, 提出了Beta-RCNN, 引入beta表示, 主要解决拥挤和遮挡环境下的行人检测问题。Sun等^[4]在Faster-RCNN的基础上, 提出一种去RPN模块、复杂后处理和NMS模块的模型, 在检测中取得了不错的成绩。

One-stage网络将目标检测看作回归问题, 直接得到类别概率和位置信息。YOLO系列自问世以来不断迭代, 不同改进方法被相继提出, 叶佳林等^[5]、杨紫辉等^[6]分别在不同的YOLO网络中进行改进, 提出了更适合非机动车检测的网络。Guo等^[7]、Montenegro等^[8]通过添加不同的注意力机制, IoU损失等方法实现对算法的改进。Liu等^[9]、Li等^[10]通过引入部分新结构、新模块, 从而使网络更好地适用于行人检测。

2 ACM-YOLO

YOLOv8是ultralytics团队于2023年1月开源的anchor-free目标检测算法框架, 其在目标检

收稿日期: 2024-04-08 修稿日期: 2024-05-04

基金项目: 中国人民公安大学安全防范工程双一流创新研究专项(2023SYL08)

作者简介: 帅勃宇(2000—), 男, 辽宁辽阳人, 硕士研究生, CCF学生会会员, 研究方向为目标检测; *通信作者: 张雅丽(1977—), 女, 山西大同人, 硕士, 副教授, 研究方向为安全防范技术, E-mail: zhangyl_mail@163.com

测、实例分割等领域具有较好的效果。

ACM-YOLO从轻量化和精度两个层次，从多尺度融合、特征提取和损失函数三个角度对YOLOv8结构进行了优化，其整体结构如图1所示。

2.1 轻量化的高效AFPN多尺度特征融合网络

在原始YOLOv8模型中采用了PAFPN结构，虽然对不同层的特征进行了融合，但还存在着一定改进空间，一是不同层的语义信息使用率较低，原始特征语义信息未充分表达；二是该结构没有充分考虑到非相邻层间的语义信息差距，特征信息传递时可能造成语义信息丢失或退化。对此，本文提出了改进的轻量化高效AFPN^[11]网络结构。

一是从关注深层特征出发，拓展小目标检测层，强化小目标特征表达，在不同层融合时使用ASFF^[12]为不同层的融合分配不同的权重；二是从特征提取能力出发，使用具有更多梯度流分支和更少参数的C2f增强多尺度特征图融合后的特征表达；三是从轻量化的角度出发，在主干网络特征图传入AFPN前控制传入通道数，有效解决多尺度融合带来的参数量和计算量剧增问题。轻量化的高效AFPN特征融合网络如图2所示。

图2中B2~B5表示不同的主干网络层，E、L、M、S表示融合过程中不同尺度的层，缩小因子即是控制输入AFPN通道数的卷积核大小为1的普通卷积。以图2中M尺度特征图为例，该过程如式(1)~(6)所示：

$$M_1^{in} = Conv(B3) \tag{1}$$

$$M_1^{out} = C2f(ASFF(M_1^{in}, D_s(E_1^{in}))) \tag{2}$$

$$M_2^{in} = Conv(M_1^{out}) \tag{3}$$

$$M_2^{out} = C2f(ASFF(M_2^{in}, D_s(E_2^{in}), U_s(L_1^{in}))) \tag{4}$$

$$M_3^{in} = Conv(M_2^{out}) \tag{5}$$

$$M_3^{out} = C2f(ASFF(M_3^{in}, D_s(E_3^{in}), U_s(L_2^{in}), U_s(S_1^{in}))) \tag{6}$$

式中：*Conv*为普通标准卷积；*ASFF*为自适应空间权重多尺度特征融合；*C2f*为原始YOLOv8模型中对应模块；*D_s*为下采样；*U_s*为上采样。

2.2 CWPCConv

轻量化的同时为提高模型的特征提取能力，本文基于PCConv^[13]提出了CWPCConv(Chanel Weights Partial Conv)——一种具有通道权重的部分卷积。

PCConv通过部分操作代替全部，从有效降低计算量，实际计算量如式(7)所示：

$$h \times w \times k^2 \times c_p^2 \tag{7}$$

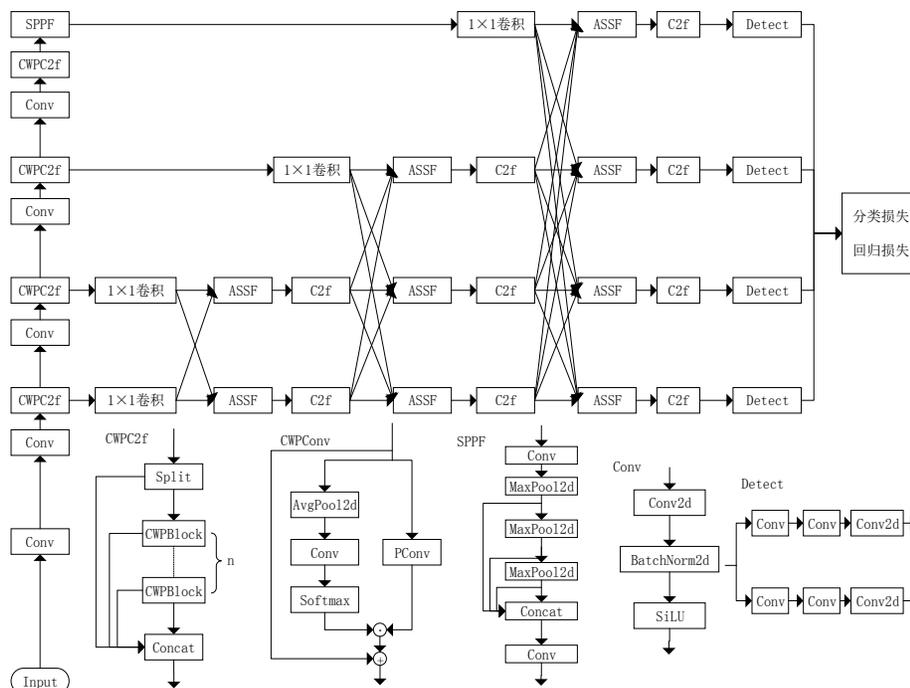


图1 ACM-YOLO网络结构图

式中： h 、 w 为特征图的高和宽； k 为卷积核大小； c_p 表示进行卷积操作的部分通道。但使用PCConv时部分重要特征信息无法充分表达，对此本文提出了CWPCConv，该卷积由两个分支组成，在第一个分支中，使用PCConv对输入特征图进行部分卷积；在第二个分支中，首先使用平均池化提取全局信息，并经过一个 1×1 的卷积进行特征表达，然后沿着通道维度进行权重分配，得到具有通道权重的特征图，最后将两个分支

特征图相乘得到具有通道权重的部分卷积，同时引入残差思想，对整个过程进行残差，其结构图如图3所示。

图3中AvgPool为平均池化，Conv为普通标准卷积，Softmax为加权操作。

在CWPCConv的基础上，本文进一步提出了CWPC2f，使用CWPCConv替换C2f中的Bottleneck模块，即为CWPBlock。CWPC2f如图4所示。

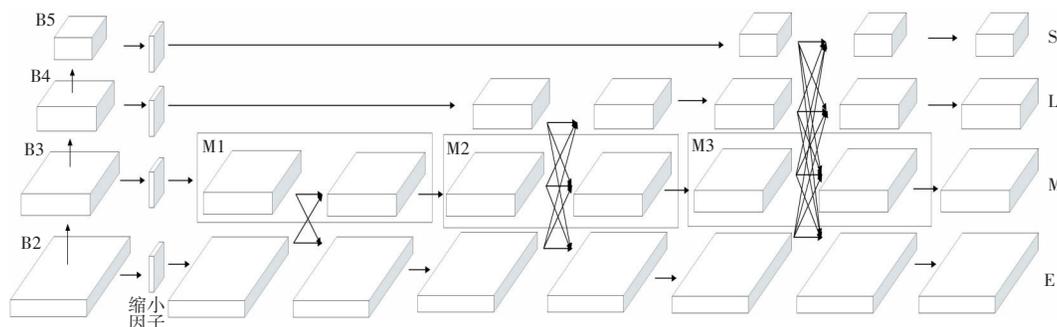


图2 轻量化的高效AFPN多尺度特征融合网络

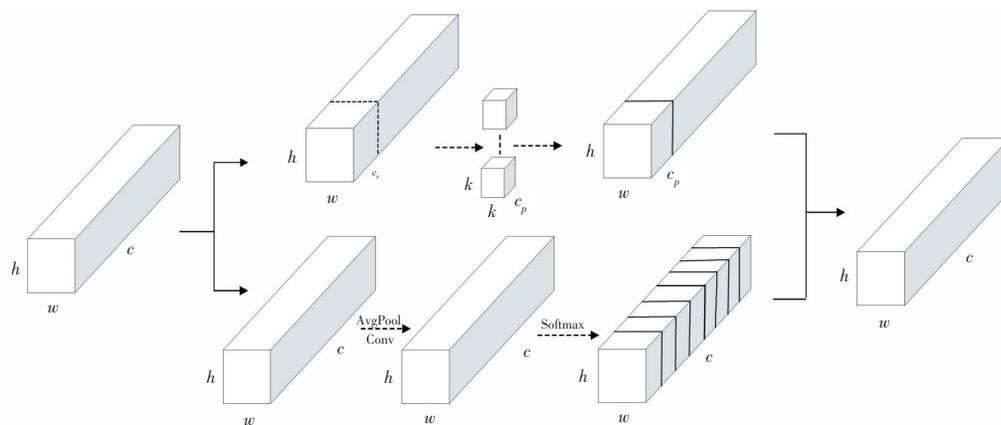


图3 CWPCConv结构图

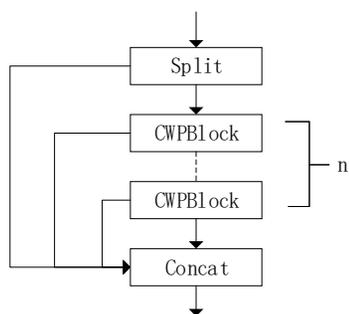


图4 CWPC2f结构图

2.3 MPDIoU loss

原始YOLOv8回归分支损失中利用交叉熵函数的思想来优化标签 y 附近左右两个位置的概率，使网络分布聚焦到标签值附近，其表达式如式(8)所示：

$$f_{DFL}(S_i, S_{i+1}) = -((y_{i+1} - y_i) \log(S_i) + (y - y_i) \log(S_{i+1})) \quad (8)$$

式中： S_i 是网络的Sigmoid输出； y 是标签值； y_i 和 y_{i+1} 是 y 相邻两侧的标签值，满足 $y_i < y < y_{i+1}$ ，

在推理时满足 $y^* = \sum_{i=0}^k (P(i) \cdot i)$, $P(i)$ 可以经过 Sigmoid, 即 S_i 实现。

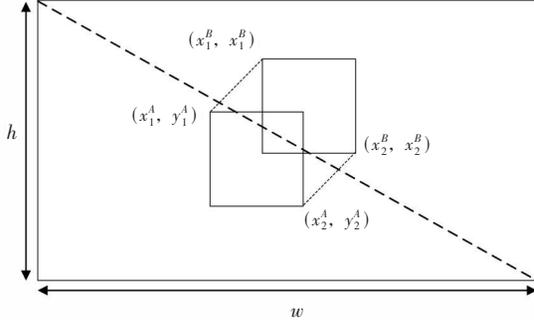


图 5 MPDIoU 结构图

本文使用 MPDIoU^[14] (如图 5 所示) 作为边界框损失回归, 其是一种基于最小点距离的回归损失, 一方面针对回归框具有相同宽高比时的损失优化, 另一方面针对目标遮挡, 即回归框重叠问题, 其表达式如式(9)所示:

$$MPDIoU = \frac{A \cap B}{A \cup B} - \frac{(x_1^B - x_1^A)^2 + (y_1^B - y_1^A)^2}{w^2 + h^2} - \frac{(x_2^B - x_2^A)^2 + (y_2^B - y_2^A)^2}{w^2 + h^2} \quad (9)$$

式中: A, B 表示任 $A, B \subseteq S \subseteq \mathbf{R}^n$; $(x_1^A, y_1^A), (x_2^A, y_2^A)$ 为回归框 A 的左上角点和右下角点坐标, 同理 $(x_1^B, y_1^B), (x_2^B, y_2^B)$ 为回归框 B 的左上角点和右下角点坐标, w, h 分别为输入图像的宽和高。其作为损失函数表达式如式(10)所示:

$$f_{MPDL} = 1 - MPDIoU \quad (10)$$

3 实验验证与结果分析

3.1 实验数据集

真实交通路况下, 行人和非机动车出现区域基本一致, 实际工作中时常将两者进行合并管理, 为更加贴合实际工作, 本文针对俯瞰远视距监控摄像头画面存在的视距远、目标小等问题, 提出了基于真实监控画面的行人和非机动车数据集 USS Pedestrian and NMV (Under surveillance small-scale pedestrian and non-motorized vehicle datasets), 该数据集截取至 213 分 37 秒的真实监控视频, 共 1357 张图像, 随机划分为 1102 张训练集图像和 255 张验证集图像, 平均每幅图像 57 个

实例, 共约 77349 个实例, 平均实例大小约为整幅图像的 0.116%。部分数据集如图 6 所示。



图 6 部分数据集

3.2 实验环境与设置

本实验使用计算机操作系统为 Ubuntu20.04, 12 vCPU Intel® Xeon® Silver 4214R CPU @ 2.40 GHz 处理器, NVIDIA GeForce RTX 3080 Ti 显卡, 显存 12 GB。以 YOLOv8 的 8.0.114 版本作为原始模型, 编译环境为 Python3.8.10。

未特殊说明实验学习率设置为 0.01, 权重衰减为 0.0005, 迭代次数为 200, 批次处理大小为 4。

3.3 实验评估指标

实验以查准率(P)、召回率(R)、平均精度值 mAP (mean Average Precision)、浮点运算量 (GFLOPs)、模型参数量 (Params) 和模型大小作为评价指标。其中 GFLOPs 代表每秒浮点运算次数, 用来评估模型的计算效率; mAP 又称为全类平均精度, 由各类别的平均正确率 (AP) 进行加权计算得到, 该值越高说明模型的性能越好。

3.4 消融实验

实验以 YOLOv8m 模型为基线, 定量分析不同改进方式下算法优越性。

3.4.1 改进 AFPN 的消融实验

(1) AFPN 算法检测能力的消融实验

本小节将分别对 AFPN 算法拓展小目标检测层的效果和使用 C2f 替换 BasicBlock 的效果进行消融实验。批次处理大小设置为 2, 实验结果见表 1。

从表 1 可以看出, 在引入 AFPN 算法后, 模型充分考虑非相邻层间的语义差距, 在更小的

行人目标上效果略有提升；拓展小目标检测层后，深层小目标特征信息充分表达，mAP50值提高了2.7%，在更小的行人类别上mAP50值提升了4.7%，但参数量提高了近8倍，进一步使用C2f替换BasicBlock后参数量下降至未使用C2f时的57%，同时精度提高了2.7%，实现了参数量和检测效果的巨大提升。

表1 AFPN算法检测能力的消融实验

	YOLOv8m	YOLOv8m+ AFPN (B345)	YOLOv8m+ AFPN (B2345)	YOLOv8m+ AFPN (B2345+ C2f)
P	0.472	0.465	0.519	0.566
R	0.406	0.404	0.392	0.412
mAP50	Non-motor vehicle	0.505	0.492	0.513
	Pedestrian	0.210	0.219	0.248
	All	0.357	0.356	0.380
模型大小/MB	52	396.1	419.4	182.8
参数量/百万	25.9	197.7	209.2	91
GFLOPs	78.9	472.3	778.9	288.2

(2) AFPN算法轻量化的消融实验

本小节将对缩小因子效果进行消融实验，以上一小节实验中表现最好的改进AFPN算法作为基准，设置缩小因子(factor)分别为1、2、4、8、16，实验结果见表2。

表2 AFPN算法轻量化的消融实验

	P	R	mAP50	模型大小/MB	参数量/百万	GFLOPs
factor=1	0.531	0.428	0.404	182.8	91	288.2
factor=2	0.527	0.421	0.390	105.8	52.5	169.4
factor=4	0.513	0.411	0.385	86.1	42.7	138.9
factor=8	0.503	0.398	0.379	80.9	40.1	130.8
factor=16	0.497	0.397	0.370	79.2	39.4	128.6

从表2可以看出，不进行输入通道控制时，模型的mAP50值最高，但参数量和计算量巨大。随着缩小因子的增大，mAP50值有所下降，参数量减少，但在缩小因子逐渐变大的过程中，mAP50与参数量下降幅度逐渐失衡。综合实验结果，当缩小因子为4时模型准确度和复杂度组合最优，因此本文选取缩小因子为4时的AFPN算法作为后续实验对象。

3.4.2 CWPCConv的消融实验

本小节将对不同卷积比例下CWPCConv特征提取能力、参数量和计算量进行消融实验，以缩小因子为4时的改进AFPN算法为基准，div表示卷积比例，当div为2时表示部分卷积比例为1/2，实验结果见表3。

表3 CWPCConv部分卷积比例的消融实验

	P	R	mAP50	模型大小/MB	参数量/百万	GFLOPs
div=2	0.363	0.297	0.264	48	23.7	83.3
div=4	0.491	0.401	0.371	43.6	21.5	76.9
div=8	0.496	0.388	0.365	42.5	21	75.3

从表3可以看出，当部分卷积比例减小时，模型的参数量和计算量有所减小，但参数量和计算量下降幅度明显小于准确度，综合实验结果，卷积比例为1/4时模型参数量、计算量和准确度结果组合最优，因此本文后续实验中均设置部分卷积比例为1/4。

3.4.3 ACM-YOLO的消融实验

本小节实验将验证模型进行不同改进时的效果，下述表格中将不同改进模块用字母代替，具体为A：AFPN(B2345+C2f)；B：CWPC2f；C：MPDIoU，实验结果见表4。

表4 ACM-YOLO的消融实验

	A	B	C	P	R	mAP50	模型大小 /MB	参数量 /百万	GFLOPs
YOLOv8m+				0.472	0.406	0.357	52	25.9	78.9
YOLOv8m+	√			0.513	0.411	0.385	86.1	42.7	138.9
YOLOv8m+	√	√		0.491	0.401	0.371	43.6	21.5	76.9
YOLOv8m+	√	√	√	0.499	0.408	0.379	43.6	21.5	76.9

从表4可以看出，对比YOLOv8m模型，使用改进的AFPN算法后，mAP50值提升了2.8%，但参数量增加了65%，计算量增加了76%；进一步使用CWPC2f模块后，mAP50值下降了1.4%，但参数量仅为改进前的51%，计算量下降至55%，相较于YOLOv8m模型参数量仍下降了17%，GFLOPs下降了2.4；进一步使用MPDIoU损失函数后，mAP50值提升了0.8%。综上所述，本文提出的ACM-YOLO相较于YOLOv8m，mAP50值提高2.2%，参数量下降了17%，计算量下降了2，验证了ACM-YOLO算法在小尺度行人和非机动车检测上的优越性。

3.5 对比实验

本小节实验选取实际应用中更主流的YOLO系列算法作为对比实验对象，具体为YOLOv3、YOLOv5m、YOLOv7、YOLOv8m。实验结果见表5。

表5 对比实验

	P	R	mAP50	模型大小 /MB	参数量 /百万	GFLOPs
YOLOv3	0.49	0.401	0.358	123.4	61.5	154.6
YOLOv5m	0.485	0.402	0.358	42.1	21.2	49
YOLOv7	0.465	0.423	0.342	74.8	36.5	103.2
YOLOv8m	0.472	0.406	0.357	52	25.9	79.3
Ours	0.495	0.421	0.379	43.6	21.5	76.9

从表5可以看出，本文提出的ACM-YOLO算法的精确度和参数量均达到了最优，相较于表现最好的YOLOv5m模型，在几乎不增加参数量的同时mAP50值提高了2.1%。因此，本文提出的ACM-YOLO算法在小尺度行人和非机动车检测上，在更轻量化的基础上有着更好的检测精度，能更好地适用于实际生活中对小尺度行人和非机动车的检测任务。

3.6 ACM-YOLO算法泛化能力验证

Visdrone2019^[15]数据集由天津大学机器学习和数据挖掘实验室的AISKEYEY团队制作，是目前小目标检测领域主流的数据集之一，共有包括行人和多种非机动车在内的10个类别。CityPerson数据集是由Zhang等^[16]于2016年在Cityscape数据集上提出的针对行人标注的数据集，每幅图像中行人的平均数量为7，并存在部分遮挡情况。本小节将使用上述两个数据集进行泛化能力验证实验。实验结果见表6、表7。

表6 ACM-YOLO在VisDrone2019数据集上的泛化能力实验

	P	R	mAP50	模型大小 /MB	参数量 /百万	GFLOPs
YOLOv8m	0.653	0.373	0.514	52	25.9	78.9
Ours	0.697	0.392	0.544	43.6	21.5	76.9

表7 ACM-YOLO在CityPerson数据集上的泛化能力实验

	P	R	mAP50	模型大小 /MB	参数量 /百万	GFLOPs
YOLOv8m	0.465	0.363	0.342	52	25.9	78.9
Ours	0.511	0.414	0.423	43.6	21.5	76.9

从表6、表7可以看出，ACM-YOLO的表现两个数据集上均好于YOLOv8m，mAP50值分别提升了8.1%和3%。因此，本文提出的ACM-YOLO算法在小尺度行人和非机动车检测上具有良好的泛化能力。

3.7 实验结果可视化

图7中(a)、(b)、(c)分别对应VisDrone2019数据集中的原始图像、YOLOv8m和ACM-YOLO的检测效果图，从结果可以看出ACM-YOLO对于小尺度行人和非机动车目标检测效果明显优于YOLOv8m。第一幅图片中ACM-YOLO正确检出所有YOLOv8m的漏检目标；第三幅图片中左下角处有一名骑电动车的人，YOLOv8m检测结果为自行车，ACM-YOLO检测结果为电动车和人。

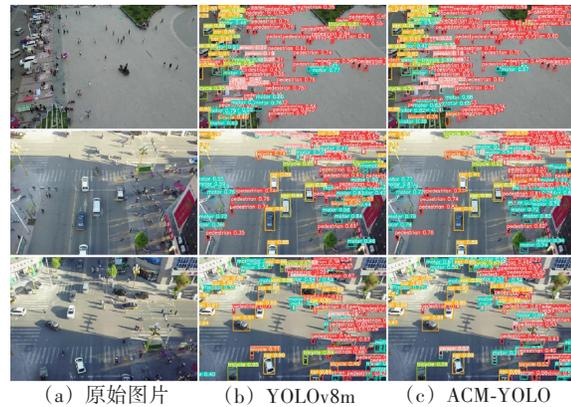


图7 不同算法检测效果对比图

4 结语

针对交通治理现状，本文立足于真实监控视野下小尺度行人和非机动车目标检测问题，提出了ACM-YOLO算法，分别针对小目标识别问题，提出了轻量化的高效AFPN特征融合网络，兼顾轻量化和特征表达能力；针对轻量化问题，提出了CWPCConv，对部分卷积进行通道权重分配，有效地解决了部分卷积去除冗余计算时的不确定性，又进一步提出了CWPC2f，实现了参数减少而精度提高的效果；针对目标遮挡问题，引入MPDIoU。经实验证明，ACM-YOLO算法在参数量、计算量和精度上均取得了良好的表现，有效解决了小尺度行人和非机动车目标检测难的问题，对现实交通中的行人和非机动车治理具有极大的现实意义。

参考文献:

- [1] 潘晓东, 马小翔, 赵晓翠. 信号交叉口非机动车骑行特性及安全性实验研究[J]. 交通科学与工程, 2010, 26(4): 60-64, 69.
- [2] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [3] XU Z, LI B, YUAN Y, et al. Beta R-CNN: looking into pedestrian detection from another perspective [C]//Proceedings of the Conference on Neural Information Processing Systems, 2020, 33: 19953-19963.
- [4] SUN P, ZHANG R, JIANG Y, et al. Sparse R-CNN: end-to-end object detection with learnable proposals [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 14454-14463.
- [5] 叶佳林, 苏子毅, 马浩炎, 等. 改进YOLOv3的非机动车检测与识别方法[J]. 计算机工程与应用, 2021, 57(1): 194-199.
- [6] 杨紫辉, 江磊, 任洪娟. 基于改进YOLOv3网络的非机动车检测[J]. 智能计算机与应用, 2021, 11(8): 87-91.
- [7] GUO W, SHEN N, ZHANG T. Overlapped pedestrian detection based on YOLOv5 in crowded scenes [C]//Proceedings of the 2022 3rd International Conference on Computer Vision, Image and Deep Learning & International Conference on Computer Engineering and Applications (CVIDL & ICCEA). IEEE, 2022: 412-416.
- [8] MONTENEGRO B, FLORES-CALERO M. Pedestrian detection at daytime and nighttime conditions based on YOLO-v5 [J]. Ingenius. Revista de Ciencia y Tecnología, 2022(27): 85-95.
- [9] LIU R, LI S, DUAN Y, et al. YOLO-DSRF: an improved small-scale pedestrian detection algorithm based on YOLOv4 [C]//Proceedings of the 2023 IEEE 5th International Conference on Civil Aviation Safety and Information Technology, 2022: 307-313.
- [10] LI Y, LI S, DU H, et al. YOLO-ACN: focusing on small target and occluded object detection [J]. IEEE Access, 2020, 8: 227288-227303.
- [11] YANG G, LEI J, ZHU Z, et al. AFPN: asymptotic feature pyramid network for object detection [EB/OL]. arXiv:2306.15988, 2023.
- [12] LIU S, HUANG D, WANG Y. Learning spatial fusion for single-shot object detection [EB/OL]. arXiv: 1911.09516, 2019.
- [13] CHEN J, KAO S, HE H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 12021-12031.
- [14] SILIANG M, YONG X. MPDIoU: a loss for efficient and accurate bounding box regression [EB/OL]. arXiv:2307.07662, 2023.
- [15] DU D, ZHU P, WEN L, et al. VisDrone-DET2019: the vision meets drone object detection in image challenge results [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, 2019.
- [16] ZHANG S, BENENSON R, SCHIELE B. Citypersons: a diverse dataset for pedestrian detection [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 3213-3221.

A lightweight small-scale pedestrian and non-motorised vehicle target detection algorithm based on improved YOLOv8

Shuai Boyu, Zhang Yali*

(School of Information Network Security, People's Public Security University of China, Beijing 100038, China)

Abstract: To solve the problem of pedestrian and non-motorized vehicle governance at traffic intersections, a small and lightweight small-scale pedestrian and non-motorized vehicle detection algorithm based on improved YOLOv8, ACM-YOLO is proposed, for the small size of pedestrian and non-motorized vehicle targets at panoramic intersection monitoring. Firstly, a lightweight and efficient AFPN feature fusion network is proposed to replace the PAFPN algorithm in the original network to improve the recognition effect of small targets; secondly, CWPCConv channel weight partial convolution is proposed based on PConv, and CWPC2f is further proposed to effectively reduce the number of model parameters and calculation; finally, the MPDIoU function is used to optimize the boundary box loss of the network. The results show that compared with YOLOv8m, the mAP50 is increased by 2.2%, 8.1% and 3% on the self-built dataset, VisDrone2019 dataset and CityPerson dataset respectively, the number of parameters is reduced by 17% and the GFLOPs is reduced by 2.

Keywords: small target detection; YOLOv8; pedestrians and non-motor vehicles detection; CWPCConv

文章编号: 1007-1423(2024)20-0008-06

DOI: 10.3969/j.issn.1007-1423.2024.20.002

植物染色质可及性深度学习预测研究综述

吴昭瑞, 王舒宇, 赵玉茗*, 于洋, 魏金寒

(东北林业大学计算机与控制工程学院, 哈尔滨 150000)

摘要: 高通量测序技术的发展推动了染色质可及性检测方法的进步, 但受限于湿实验的高成本、复杂性, 以及植物细胞特性, 植物染色质可及性研究相对滞后。研究人员尝试将植物染色质可及性与生物信息学方法相结合, 来探讨生物信息学方法在染色质可及性数据和预测分析的优势和局限性。

关键词: 染色质可及性; 深度学习; OCR 分类预测

0 引言

染色质可及性是核大分子与染色质 DNA 接触的程度, 具有高度的动态性, 因此检测的过程产生大量数据。染色质开放性与基因表达调控密切相关, 因为它决定了顺式调控元件是否能够招募转录因子形成蛋白质-DNA 复合物, 从而影响下游基因的转录。由于染色质可及性影响转录因子与基因顺式调控元件的结合^[1], 开放的染色质使 DNA 能被调控元件接近, 从而激活基因表达。全基因组范围内的染色质开放性研究对于理解不同植物中关键分子过程有重要意义^[2]。生物实验的 OCR (open chromatin region) 检测方法需要耗费高成本的材料, 临床样本在很大程度上难以进行表观遗传学分析。然而, 小规模理解复杂疾病的性质审查染色质可及性分析的进展带来了广泛的临床影响, 比如 TCGA 项目和 ENCODE 项目以及一些研究人员证实了染色质开放状态的变化对癌症的产生和演变具有重要影响^[3], 研究人员通过 OCR 测序技术来构建乳腺癌细胞表观调控网络, 识别预测顺式调控因子对细胞形态的改变, 由此来减少乳腺癌恶性转移和发展。除此之外, 通过映射 OCR

可以预测 DNA 相互作用^[4]。通过以上结论可以看出, OCR 检测是表观遗传学中一项重要的研究工作。

植物界 OCR 的研究对育种至关重要, 现有的 OCR 预测算法基本基于人类或者模式动物数据, 植物 OCR 研究领域缺少相关数据的整合和预测, 仍处于探索阶段, 而植物细胞特性是研究困难的主要原因。OCR 测序技术可分为三类: 酶切法、物理断裂法、甲基化修饰法^[5]。这些测序技术都伴随着细胞需求量大、实验环境复杂等问题, 因此, 将深度学习算法应用于全基因组测序技术已然成为一个好的趋势。本文主要综述了深度学习在植物 OCR 预测的运用。

1 染色质开放区域实验数据集的建立

在研究过程中研究人员构建了针对染色质可及性不同方面研究的数据集, 数据集的质量对于实验的成功起着决定性的作用, 不同的数据集在相同的算法中会出现不同的结果, 表 1 汇总了部分近年来研究人员和学者在实验过程中构建的数据集以及相关信息的分析。

对于表 1 中的数据集, 研究人员多为构建 ATAC-seq 文库, 构建实验流程如图 1 所示。

收稿日期: 2024-04-13 修稿日期: 2024-06-20

基金项目: 黑龙江省大学生创新创业计划项目 (S202310225208)

作者简介: 吴昭瑞 (2003—), 女, 湖南双峰人, 本科生, 研究方向为软件工程; 王舒宇 (2002—), 男, 本科生, 研究方向为软件工程; *通信作者: 赵玉茗 (1978—), 女, 山东莱州人, 博士, 教授, 研究方向为生物信息学、数据挖掘, E-mail: zym@nefu.edu.cn; 于洋 (2003—), 男, 山东人, 本科生, 研究方向为软件工程; 魏金寒 (2003—), 女, 黑龙江双鸭山人, 本科生, 研究方向为 Web 开发

表 1 常见植物染色质可及性数据集汇总

数据集物种	数据来源	数据集构建结果
大豆	大豆OCR光照生物节律实验(贾光红)	RNA-seq、ATAC-seq文库
禾本科植物：水稻、结缕草、小麦	邓兴旺实验室何航课题组提供，北京林业大学草坪所提供，中国科学院遗传与发育研究所李义文课题组提供	水稻、结缕草、小麦的 ATAC-seq 数据
模式植物：水稻、拟南芥、苜蓿、番茄	RAP-DB database、PlantDHS、NCBI、TAIR database、M.truncatula genome database、Phytozome database	正负样本比例1:1的训练集
37个物种的19个植物器官	NCBI GEO/SRA 数据资源平台、已发表文章中收集染色质可及性样本	PlantCADB Web 数据库平台
水稻、二穗短柄草、小米、高粱、玉米旗叶	实验田培养的植物	多个物种 ATAC-seq 文库

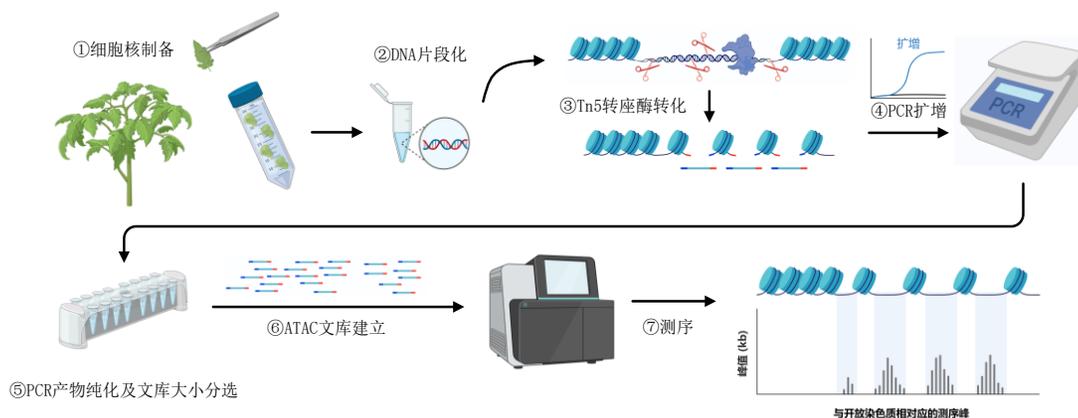


图 1 ATAC-seq文库构建实验流程

2 深度学习预测模型的分析

植物 OCR 的研究存在滞后性，实验研究使用深度学习预测潜在的 OCR 是有生物学意义的，一个完整的潜在 OCRs 全基因组图谱有助于研究核小体位置的变化，有助于发现基因组调控元件和基因调控机制^[6]。在 ATAC-seq 实验中仅能获得当前状态的染色质可及性，而通过深度学习算法预测可以得到潜在 OCR，Shen 等^[6]针对模式植物开发了染色质可及性预测系统，CharPlant 是第一个能够在整个基因组中识别潜在 OCR 的预测工具，并且包含预测植物染色质开放区域的神经网络模型，该系统适用于具有遥远进化关系、广泛的植物物种^[7]。

深度学习在基因组学应用越来越多，同时在生物学预测任务中表现出优越的性能，下面介绍常用于生物信息学问题的深度学习算法。

2.1 卷积神经网络

卷积神经网络(convolutional neural networks, CNN)是一种前馈神经网络，在研究领域应用十分广泛，研究人员多用卷积神经网络进行预测或与个人成果进行比较。

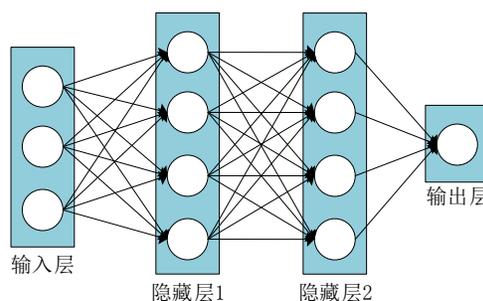


图 2 卷积神经网络原理图

DeepChrome^[8]是 CNN 的一个经典预测方法，能够对基因表达任务做出高精度的预测，在 2.0

版本^[9]中通过改进架构实现了跨细胞预测,在预测基因表达研究中,科研人员基于CNN和其他模型融合混合框架取得好的效果。DeepBind^[10]、DeepSEA^[11]和Basset^[12]等工具使用CNN通过学习序列特征,分别预测蛋白质结合位点、非编码区以及DNA序列的功能活性^[2]。Zhou等^[13]开发的DeepSEA用于非编码变量效应预测,预测染色质特征的准确性很高,可以准确地预测单个SNPs对DeepSEA TF结合分类器对TF结合的影响。Hu等^[14]提出了PlantDeepSEA,是一个基于DeepSEA模型的植物NCV优先级排序和CRE识别的在线Web服务器。

2.2 循环神经网络

循环神经网络(recurrent neural network, RNN)对序列特性数据进行预测方面十分强大。双向RNN能够结合上下文信息进行预测,在预测染色质可及性这样的动态性质过程中,融入时间序列能够提高模型的预测性能。

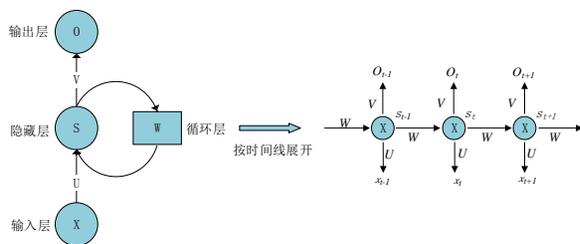


图3 循环神经网络原理图

Boza^[15]及其团队设计的深度双向循环神经网络DeepNano针对MinION测序技术进行了优化,显著降低了长读取时的测序错误率。这一创新有助于提升MinION在临床应用中的性能,尤其在监测传染病爆发和表征癌症结构变异等方面具有重要意义。相较于传统的隐马尔可夫模型(HMM)方法,DeepNano提供了更准确、计算效率更高的替代方案,使基于DeepNano调用的MinION测序能够更快速、准确地分析基因组信息。

2.3 长短时记忆神经网络

长短时记忆神经网络(long short-term memory, LSTM)是一种改进的循环神经网络(RNN)架构,由于RNN无法处理长序列数据的长程依赖关系,

过长的数据会使模型训练时发生梯度的爆炸或消失, Hochreiter等^[16]提出的LSTM模型解决了这一缺陷,可以对存在短期或长期依赖的数据进行建模。

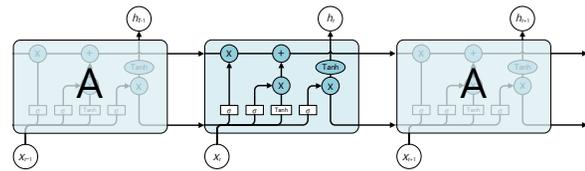


图4 长短时记忆神经网络原理图

Sekhon等^[17]开发了一种新的基于注意力的深度学习架构DeepDiff,使用了多个LSTM模块预测基因在两种不同细胞条件下的对数倍变化,该模块可以表示染色质标记不同位置的信号之间的相互作用,编码输入信号的空间结构。DeepDiff提供了一种深入理解HMs如何调控基因表达的新视角,通过利用深度学习从数据中提取丰富表征的能力,可以更深入地探索表观遗传因素对基因调控的原理。

2.4 卷积递归神经网络

卷积递归神经网络(convolutional recurrent neural networks, CRNN)是DCNN和RNN的结合,在CRNN模型中融合了二者的优点:卷积神经网络用于提取输入数据的特征表示,循环神经网络用于对这些特征进行序列建模,从而在文字识别等序列数据任务中表现良好。研究人员将CTC(connectionist temporal classification)解码应用于CRNN解决了不定长序列的对齐问题。与CNN模型相比,CRNN具有更少的参数:在CRNN中,所有层都具有权重共享连接,不需要完全连接的层^[18]。

Quang等^[19]提出的CRNN神经网络框架DanQ在预测DNA序列特性和功能方面表现优异,明显优于gkm-SVM和DeepSEA等其他方法。在ROC AUC和PR AUC指标上,DanQ分别在94.1%和97.6%的目标上优于DeepSEA。DanQ能够更准确地预测遗传变异导致的表观遗传变化,这对于鉴定因果变异和预测基因组编辑的表型结果至关重要。Symeonidi等^[20]提出的基于递归卷积神经网络的模型ReChrome用于预测基因表达。该模型通过加深卷积网络的深度以及允许感受野

呈指数级增长，显著提升了预测性能，还引入了全局池化模块，大幅减少了模型参数数量，降低了过拟合的风险，使模型能够灵活适应不同长度的输入数据。ReChrome 相比于 DeepChrome 和 AttentiveChrome 等模型在基因表达预测任务中展现出了更为卓越的性能，同时稳定性和泛化能力也显著优于其他模型，为基因组学研究提供了更为准确和可靠的工具，有望推动该领域的研究进展。

2.5 残差网络(ResNet)

在深度卷积网络的研究中提出了退化问题的概念，退化问题是指随着网络深度的增加，精度达到饱和，然后迅速退化。这种退化不是因为过拟合引起的，而是因为模型中添加过多的层数会导致更高的训练误差，针对退化问题 He 等^[21]提出了残差网络学习框架，在深层网络上能够降低训练误差。

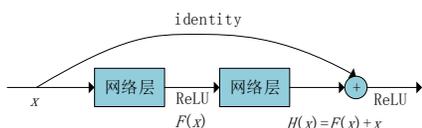


图5 残差网络原理图

Nair 等^[22]提出了一种多模态残差网络架构 ChromDragoNN，揭示了在 123 种不同的细胞环境中染色质动力学的顺式和反式调控因子，该架构能够增强跨细胞类型的共享和上下文特定染色质可访问位点的全基因组预测，这对于破译转录调控的动力学至关重要。

随着测序技术的突飞猛进和生物数据的爆炸式增长，利用计算机技术对生物数据进行高效处理，挖掘其中蕴含的生物学知识已成为科研领域的重点。深度学习模型以其强大的自动特征学习能力，准确捕捉数据中的复杂模式，被广泛应用于表观遗传数据的分析中，以预测基因表达情况及其差异性，这些模型不仅有助于我们深入揭示表观遗传因素在基因调控中的复杂机制，还为染色质可及性的精准预测提供了潜在的可靠工具。通过不断优化这些模型，我们有望更准确地解析生物数据，进一步推动生物学领域的研究发展。

表2 应用于生物信息研究中的深度学习模型

预测模型	预测方法	模型特点
卷积神经网络 (CNN)	DeepChrome ^[8]	可以自动学习组蛋白修饰标记之间的组合相互作用
循环神经网络 (RNN)	DeepNano ^[15]	用于序列标注的人工神经网络
长短时记忆神经网络 (LSTM)	AttentiveChrome ^[23]	能够对相关标记进行差异化关注
多个 LSTM 模块的多层次结构	DeepDiff ^[17]	对相关的修饰进行差异处理，为每个修饰定位重要的基因组位置
残差网络	ChromDragoNN ^[22]	减少深度，更容易优化
卷积递归神经网络 (CRNN)	ReChrome ^[20]	特征提取、序列建模和转录集成到一个统一的框架中

3 未来发展方向

3.1 集成策略

集成策略是通过整合不同方式得到的特征信息，增强数据质量的效果，提升预测的性能。集成策略需要先选择基础预测模型，再将多个模型组合。黄子墨^[24]提出了集成学习模型 EnDGE 通过五种核心组蛋白修饰特征数据来预测不同细胞类型对差异表达基因，其预测任务是一个二分类任务，预测基因是否为表达基因，通过融合特征利用多层感知机分类模型预测，结果显示 EnDGE 性能超过了单个分类器和 ReChrome。Cao 等^[25]提出 lncLocator 的最终预测模型是一个集成预测器，用于预测 lncRNA 亚细胞定位，其结果也是优于单个分类器。Zhang 等^[26]使用每个特征集训练基本模型，与合并所有特征的特征集相比维度更低，并且构建了基于 LightGBM 的双层集成模型策略，这种双层集成模型减少了特征的数量、提高了训练的效率，一个单层集成模型来表示每个特征组的预测贡献，然后这些单层集成模型的预测分数使用不同的权重被集成为最终的两层集成模型^[27]。徐斯文^[28]发现一些基于测序数据的无偏性设计的算法，在存在序列特异性的情况下会产生大量的假阳性。集成策略是解决以上问题的一种有效方案。

4 结语

表观遗传修饰对基因的表达起着重要的调控作用, 研究人员发现染色质可及性与基因的表达相关, 通过生物信息学实验, 无需复杂的实验流程, 仅需已测序且注释准确的染色质状态信息。深度学习因可以自动学习特征表示, 得到了更广泛的应用。集成学习则用于提高模型的泛化能力和预测性能, 提取更多的数据特征。当前染色质可及性的研究在很多方面存在研究深度不足的问题, 在机理方面和控制方面研究不深入, 在植物染色质可及性方面仍缺少模型的更新。将 ATAC-seq 与各种组学构建相结合, 是未来需要继续探索的方向。

参考文献:

- [1] 张雅娜, 牟苑菁, 李苇祥, 等. 基于单细胞 ATAC 测序数据的乳腺癌上皮细胞转录因子研究[J]. 生命科学, 2024, 28(2): 171-178.
- [2] 丁可. 植物染色质可及性分析预测研究及相关系统构建[D]. 哈尔滨: 东北林业大学, 2024.
- [3] 陈柳彬. 基于 cfDNA-seq 数据的染色质开放区域识别算法研究[D]. 西安: 西安交通大学, 2021.
- [4] 吕笑言. 单细胞染色质可及性测序数据在线分析平台设计与实现[D]. 哈尔滨: 哈尔滨工业大学, 2022.
- [5] 吴杰, 全建平, 叶勇, 等. 染色质转座酶可及性测序研究进展[J]. 遗传, 2020, 42(4): 333-346.
- [6] SHEN Y, CHEN L L, GAO J X. CharPlant: a de novo open chromatin region prediction tool for plant genomes [J]. *Genomics, Proteomics & Bioinformatics*, 2021, 19(5): 860-871.
- [7] 沈胤. 基于深度学习的染色质开放区域及相互作用预测[D]. 武汉: 华中农业大学, 2023.
- [8] SINGH R, LANCHANTIN J, ROBINS G, et al. DeepChrome: deep-learning for predicting gene expression from histone modifications [J]. *Bioinformatics*, 2016, 32(17): i639-i648.
- [9] KADAVATH S, PARADIS S, YEUNG J. DeepChrome 2.0: investigating and improving architectures, visualizations, & experiments [EB/OL]. arXiv: 2209.11923, 2022.
- [10] ALIPANAHI B, DELONG A, WEIRAUCH M T, et al. Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning [J]. *Nature Biotechnology*, 2015, 33: 831-838.
- [11] ZHANG S, ZHOU J, HU H, et al. A deep learning framework for modeling structural features of RNA-binding protein targets [J]. *Nucleic Acids Research*, 2016, 44: e32.
- [12] KELLEY D R, SNOEK J, RINN J L. Basset: learning the regulatory code of the accessible genome with deep convolutional neural networks [J]. *Genome Research*, 2016, 26: 990-999.
- [13] ZHOU J, TROYANSKAYA O G. Predicting effects of noncoding variants with deep learning-based sequence model [J]. *Nature Methods*, 2015, 12(10): 931-934.
- [14] HU Z, ZHUO T, LIU Y M, et al. PlantDeepSEA, a deep learning-based web service to predict the regulatory effects of genomic variants in plants [J]. *Nucleic Acids Research*, 2021, 49(W1): W523-W5291.
- [15] BOZA V, BREJOVA B, VINAR T. DeepNano: deep recurrent neural networks for base calling in MinION nanopore reads [J]. *PLoS ONE*, 2017, 12(6): e0178751.
- [16] HOCHREITER S, SCHMIDHUBER J. Long short-term memory [J]. *Neural Computation*, 1997, 9(8): 1735-1780.
- [17] SEKHON A, SINGH R, Qi Y. DeepDiff: DEEP-learning for predicting DIFFerential gene expression from histone modifications [J]. *Bioinformatics*, 2018, 34(17): i891-i900.
- [18] SHI B, BAI X, YAO C. An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(11): 2298-2304.
- [19] QUANG D, XIE X H. DanQ: a hybrid convolutional and recurrent deep neural network for quantifying the function of DNA sequences [J]. *Nucleic Acids Research*, 2016(11): e107.
- [20] SYMEONIDI A, NICOLAOU A, JOHANNES F, et al. Recursive convolutional neural networks for epigenomics [C] // Proceedings of the 2020 25th International Conference on International Conference on Pattern Recognition, Milan, Italy, 2021: 2567-2574.
- [21] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C] // Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 2016: 770-778.

- [22] NAIR S, KIM D S, PERRICONE J, et al. Integrating regulatory DNA sequence and gene expression to predict genome-wide chromatin accessibility across cellular contexts [J]. *Bioinformatics*, 2019, 35 (14) : i108-i116.
- [23] SINGH R, LANCHANTIN J, SEKHON A, et al. Attend and predict: understanding gene regulation by selective attention on chromatin [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017:6788-6798.
- [24] 黄子墨. 基于表观遗传数据的差异表达基因预测方法研究[D]. 济南:山东大学,2024.
- [25] CAO Z, PAN X, YANG Y, et al. The lncLocator: a subcellular localization predictor for long non-coding RNAs based on a stacked ensemble classifier[J]. *Bioinformatics*, 2018, 34(13):2185-2194.
- [26] ZHANG Y J, YU S, XIE R P, et al. PeNGaRoo, a combined gradient boosting and ensemble learning framework for predicting non-classical secreted proteins[J]. *Bioinformatics*, 2020, 36(3):704-712.
- [27] WANG J W, LI J H, YANG B J, et al. Bastion3: a two-layer ensemble predictor of type III secreted effectors[J]. *Bioinformatics*, 2018, 35(12):2017-2028.
- [28] 徐斯文. 基于染色质可及性测序数据的转录因子结合位点分析方法研究[D]. 哈尔滨:哈尔滨工程大学,2022.

A review of deep learning prediction of chromatin accessibility in plants

Wu Zhaorui, Wang Shuyu, Zhao Yuming*, Yu Yang, Wei Jinhan

(College of Computer and Control Engineering, Northeast Forestry University, Harbin 150000, China)

Abstract: The development of high-throughput sequencing technology has promoted the progress of chromatin open region detection methods. However, the high cost and complexity of wet lab experimentation, along with the characteristics of plant cells, have led to the relative lag of plant chromatin accessibility research. Researchers have tried to combine plant chromatin accessibility and bioinformatics methods. This paper discusses the application and advantages of bioinformatic methods in chromatin accessibility data and predictive analysis.

Keywords: chromatin accessibility; deep learning; OCR classification prediction

文章编号: 1007-1423(2024)20-0014-06

DOI: 10.3969/j.issn.1007-1423.2024.20.003

基于改进 YOLOv8 的飞鸟检测算法

陈倩, 卢扬*, 邵飞翔, 李师艳

(扬州市职业大学信息工程学院, 扬州 225009)

摘要: 旨在提升飞鸟检测任务的准确性, 针对传统算法在复杂背景和多变姿态飞鸟检测中的不足, 提出基于深度学习的改进 YOLOv8 算法。该算法通过融合 GOLD-YOLO 中的 GD 机制和 ASF-YOLO 中的尺度序列特征融合 (SSFF) 模块, 优化网络结构, 改进特征提取和融合方式, 增强多尺度特征融合能力, 提升小尺寸目标的特征表示。此外, 还收集了一个涵盖不同鸟类近照、飞翔鸟类和鸟群图片的飞鸟数据集, 用于算法的训练和测试。实验结果表明, 改进后的 YOLOv8 算法在飞鸟检测任务中性能得以提升, 尤其在处理小目标和复杂背景时表现更佳。

关键词: 目标检测; YOLO; 特征提取

0 引言

随着计算机视觉技术的快速发展, 目标检测作为其中的重要分支, 已经广泛应用于各个领域, 包括安全监控、自动驾驶、生态保护等。其中, 飞鸟检测作为目标检测的一个重要应用领域, 其准确性、实时性以及稳定性对于诸多场景而言都具有不可忽视的意义。然而, 由于飞鸟目标的多样性、快速移动性以及复杂背景干扰等特性, 使得飞鸟检测算法面临着诸多挑战^[1]。

传统的飞鸟检测算法往往依赖于复杂的特征提取和分类器设计, 不仅计算量大、实时性差, 而且在面对复杂背景和多变姿态的飞鸟时, 其检测效果往往不尽如人意^[2]。近年来, 深度学习技术的兴起为飞鸟检测提供了新的解决思路。其中, YOLO(You Only Look Once)系列算法以其高效的实时性能和准确的检测结果而备受关注。然而, 对于飞鸟检测这一特定任务, YOLOv8 算法仍然存在一些局限性, 如对小目标的检测精度不高、对复杂背景的鲁棒性不强等问题。

因此, 本文旨在通过改进 YOLOv8 算法, 提升其在飞鸟检测任务中的性能。具体而言, 我们将针对飞鸟目标的特性, 优化网络结构, 改进特征提取和融合方式, 提高算法对小目标和复杂背景的处理能力。主要贡献如下:

(1) 融合了 GOLD-YOLO 中的 GD 机制和 ASF-YOLO 中的尺度序列特征融合(SSFF)模块, 来增强多尺度特征融合能力, 提升小尺寸目标的特征表示。

(2) 收集了飞鸟数据集, 数据集中不仅仅包括不同鸟类的近照, 还包含飞翔的鸟类和鸟群的图片。

1 相关研究

经过多年的发展, YOLO 系列模型已在实时物体检测领域占据重要地位。YOLOv1-v3^[3]构建了最初的 YOLOs, 确立了由主干网络-颈部模块-检测头三部分组成的单阶段检测结构。YOLOv4 优化了之前使用的主干结构, 并提出了一系列改进方法, 如 Mish 激活函数、PANet 和数据增

收稿日期: 2024-05-07 修稿日期: 2024-06-17

基金项目: 2023 年度扬州市职业大学校级科研项目(RX230002164); 2023 年江苏省职业院校学生创新创业培育计划项目(G-2023-2013)

作者简介: 陈倩(1997—), 女, 江苏扬州人, 硕士, 助教, 研究方向为人工智能, 物联网; *通信作者: 卢扬(1977—), 女, 江苏扬州人, 硕士, 副教授, 研究方向为计算机技术, E-mail: 942535135@qq.com; 邵飞翔(2004—), 男, 江苏徐州人, 学生, 研究方向为软件技术; 李师艳(2003—), 女, 江苏常州人, 学生, 研究方向为网络技术

强方法。YOLOv5继承了YOLOv4的方案，改进了数据扩增策略，并提供了更多的模型变体。YOLOv6首次将重新参数化方法引入YOLO系列模型，提出了EfficientRep主干网和Rep-PAN Neck。YOLOv7^[4]着重分析了梯度路径对模型性能的影响，提出了E-ELAN结构，在不破坏原有梯度路径的情况下增强了模型能力。YOLOv8吸收了以往YOLO模型的优点，并对其进行了整合，以实现当前YOLO系列的SOTA。

在目前的大多数目标检测应用中，YOLO系列一直是最为主流的目标检测方法，而YOLOv8则是YOLO系列中最新且稳定的SOTA算法，因此选择该网络作为本文的主干网络模型。

2 YOLOv8算法改进策略

2.1 YOLOv8模型

YOLOv8网络包括主干网络Backbone、颈部模块Neck和检测头Head三个组成部分。Backbone主要负责对图像的特征提取，然后在主干中选择三个有效的特征层输入到YOLOv8的Neck中。Neck的主要任务是对提取的特征图进行特征融合处理。Neck采用PANet^[5]结构，该结构由FPN和PAN两部分组成。FPN首先从卷积神经网络中提取特征图，构建特征金字塔，然后自顶向下进行上采样和特征图的融合。PAN是一种用于增强FPN性能的网络结构，它引入了路径聚合的概念，通过自底向上的结构，从不同层级的特征图中提取信息并将其融合，

充分保留了空间特征信息。

在本文中，模型融入华为2023最新GOLD-YOLO^[6]中的Gather-and-Distribute(GD)与ASF-YOLO^[7]中的Attentional Scale Sequence Fusion(ASF)方法改进YOLOv8的颈部模块Neck。

2.2 融合Gather-and-Distribute(GD)

GOLD-YOLO提供了一种先进的GD机制，通过卷积和自注意力操作来实现，它增强了多尺度特征融合能力，并在所有模型尺度上实现了延迟和准确性之间的理想平衡。

GD机制包括两个分支：浅层聚集与分布分支(low-GD)和深层聚集与分布分支(high-GD)，分别通过基于卷积的区块和基于注意力的区块提取和融合特征信息。除此之外，GD分成3个模块：特征对齐模块(feature alignment module, FAM)、信息融合模块(information fusion module, IFM)、信息注入模块(information injection module, Inject)。GD机制如图1所示，颈部的输入包括骨干提取的特征图B2、B3、B4、B5。

low-GD主要用于融合模型浅层的特征信息，包含FAM模块和IFM模块。FAM将B2、B3、B4、B5统一到B4的 $[h/4, w/4]$ 的尺寸，然后在通道上合并。

$$F_{align} = Low_FAM([B2, B3, B4, B5]) \quad (1)$$

将FAM的输出先经过多层RepBlock提取信息，然后再按通道分成两个部分，见下式，得到inj_p3和inj_p4，作为后续Inject模块的输入。

$$F_{fuse} = RepBlock(F_{align}) \quad (2)$$

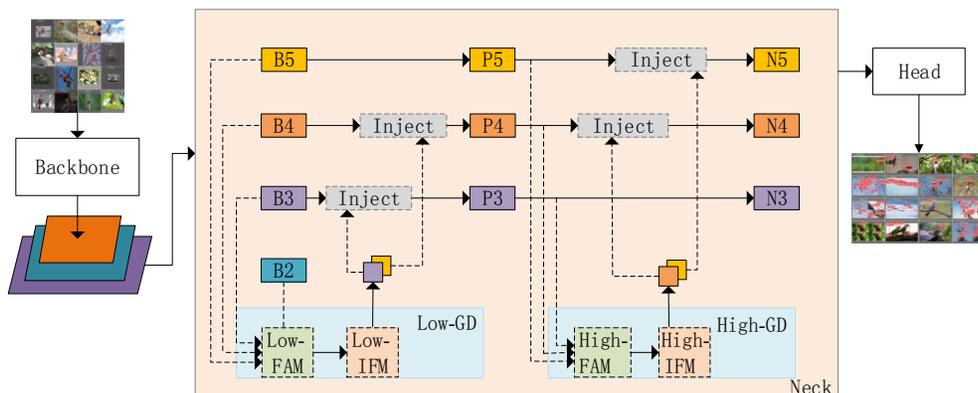


图1 GD机制示意图

$$F_{inj_P3}, F_{inj_P4} = Split(F_{fuse}) \quad (3)$$

为了更有效地将全局信息注入不同层次，GD采用注意力操作来融合信息。输入本地信息(指当前层级的特征)和全局注入信息(由 IFM 生成)，分别记为 F_{local} 和 F_{inj} 。我们使用两种不同的 Convs 与 F_{inj} 进行计算，得出 F_{global_embed} 和 F_{act} 。而 F_{local_embed} 是 Conv 与 F_{local} 计算得到的。然后通过注意力计算融合特征得到 F_{out} 。

在 low-GD 阶段， F_{local} 等于 B_i ，计算公式如下：

$$F_{global_act_Pi} = resize\left(Sigmoid\left(Conv_{act}\left(F_{inj_Pi}\right)\right)\right) \quad (4)$$

$$F_{global_embed_Pi} = resize\left(Conv_{global_embed_Pi}\left(F_{inj_Pi}\right)\right) \quad (5)$$

$$F_{att_fuse_Pi} = Conv_{local_embed_Pi}(B_i) * F_{ing_act_Pi} + F_{global_embed_Pi} \quad (6)$$

$$P_i = RepBlock(F_{att_fuse_Pi}) \quad (7)$$

high-GD 融合 low-GD 生成的特征 $\{P3, P4, P5\}$ ，如图 2 所示。

计算公式如下：

$$F_{align} = High_FAM([P3, P4, P5]) \quad (8)$$

由于 High-FAM 输出的特征尺寸变小了，为了充分融合全局信息，low-GD 中的 $RepBlock$ 替换成了 Transformer 模块，这里的 Transformer 是基于卷积实现的。

$$F_{fuse} = Transformer(F_{align}) \quad (9)$$

$$F_{inj_N3}, F_{inj_N4} = Split(Conv1 \times 1(F_{fuse})) \quad (10)$$

high-GD 的 Inject 模块与 low-GD 的此模块相同。在 high-GD 阶段， F_{local} 等于 P_i ，计算公式如下：

$$F_{global_act_Ni} = resize\left(Sigmoid\left(Conv_{act}\left(F_{inj_Ni}\right)\right)\right) \quad (11)$$

$$F_{global_embed_Ni} = resize\left(Conv_{global_embed_Ni}\left(F_{inj_Ni}\right)\right)$$

$$(12)$$

$$F_{att_fuse_Ni} = Conv_{local_embed_Ni}(P_i) * F_{ing_act_Ni} + F_{global_embed_Ni} \quad (13)$$

$$N_i = RepBlock(F_{att_fuse_Ni}) \quad (14)$$

2.3 融合Attentional Scale Sequence Fusion(ASF)

在 GOLD-YOLO 框架的基础上，再采用尺度序列特征融合(SSFF)模块增强网络的多尺度信息提取能力，并采用三重特征编码器(TPE)模块融合不同尺度的特征图以增加详细信息。

SSFF 可以更好地将深层特征图的高维信息与浅层特征图的详细信息相结合，其中在图像下采样过程中图像的大小发生变化，但尺度和特征保持不变。具体公示如下：

$$F_{\sigma}(w, h) = G_{\sigma}(w, h) \times f(w, h) \quad (15)$$

$$G_{\sigma}(w, h) = \frac{e^{-\frac{w^2 + h^2}{2\sigma^2}}}{2\pi\sigma^2} \quad (16)$$

式中： $f(w, h)$ 表示宽度为 w 、高度为 h 的二维输入图像。 $F_{\sigma}(w, h)$ 是由一系列卷积使用二维高斯滤波器 $G_{\sigma}(w, h)$ 进行平滑处理生成的。 σ 是用于卷积的二维高斯滤波器的标准差缩放参数。

这些生成的图像具有相同的分辨率和不同的尺度。因此，不同大小的特征图可以看作是尺度空间，不同分辨率的有效特征图可以调整到相同的分辨率进行拼接。

在对大尺寸特征图进行处理后，其通道数调整为 1C，然后使用最大池化+平均池化的混合结构进行下采样，这有助于保留高分辨率特征。对于小尺寸特征图，使用卷积模块来调整通道数，然后使用最近邻插值方法进行上采样。这有助于保持低分辨率图像的局部特征的丰富性，并防止小目标特征信息的丢失。最后，将三个大小相同的大、中、小尺寸的特征图进行一次卷积，然后在通道维度上进行拼接，公示

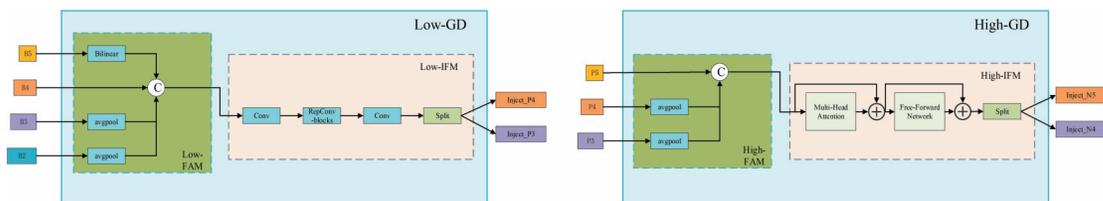


图 2 low-GD 与 high-GD 示意图

如下所示。

$$F_{TFE} = \text{Concat}(F_l, F_m, F_s) \quad (17)$$

式中： F_{TFE} 表示TFE模块输出的特征图。 F_l 、 F_m 、 F_s 分别表示大、中和小尺寸的特征图。 F_{TFE} 由 F_l 、 F_m 、 F_s 的拼接得到， F_{TFE} 与 F_m 具有相同的分辨率，且通道数是 F_m 的三倍。

对应的.yaml文件代码如下：

```
backbone:
# [from, repeats, module, args]
- [-1, 1, Conv, [64, 3, 2]] # 0-P1/2
- [-1, 1, Conv, [128, 3, 2]] # 1-P2/4
- [-1, 3, C2f, [128, True]]
- [-1, 1, Conv, [256, 3, 2]] # 3-P3/8
- [-1, 6, C2f, [256, True]]
- [-1, 1, Conv, [512, 3, 2]] # 5-P4/16
- [-1, 6, C2f, [512, True]]
- [-1, 1, Conv, [1024, 3, 2]] # 7-P5/32
- [-1, 3, C2f, [1024, True]]
- [-1, 1, SPPF, [1024, 5]] # 9
head:
- [[2, 4, 6, 9], 1, SimFusion_4in, []] # 10
- [-1, 1, IFM, [[64, 32]]] # 11
- [9, 1, Conv, [512, 1, 1]] # 12
- [[4, 6, -1], 1, SimFusion_3in, [512]] # 13
- [[-1, 11], 1, InjectionMultiSum_Auto_pool,
[512, [64, 32], 0]] # 14
- [-1, 3, C2f, [512]] # 15
- [6, 1, Conv, [256, 1, 1]] # 16
- [[2, 4, -1], 1, SimFusion_3in, [256]] # 17
- [[-1, 11], 1, InjectionMultiSum_Auto_pool,
[256, [64, 32], 1]] # 18
- [-1, 3, C2f, [256]] # 19
- [[19, 15, 9], 1, PyramidPoolAgg, [352, 2]] # 20
- [-1, 1, TopBasicLayer, [352, [64, 128]]] # 21
- [[19, 16], 1, AdvPoolFusion, []] # 22
- [[-1, 21], 1, InjectionMultiSum_Auto_pool,
[256, [64, 128], 0]] # 23
- [-1, 3, C2f, [256]] # 24
- [[-1, 12], 1, AdvPoolFusion, []] # 25
- [[-1, 21], 1, InjectionMultiSum_Auto_pool,
[512, [64, 128], 1]] # 26
- [-1, 3, C2f, [1024]] # 27
- [[4, 6, 8], 1, ScalSeq, [256]] # 28 args[inchane]
- [[19, -1], 1, Add, []] # 29
# - [[19, -1], 1, asf_attention_model, []] # 29
- [[29, 24, 27], 1, Detect, [nc]] # 28
```

3 实验设计与结果分析

3.1 数据集

本文选用的数据集是爬虫爬取的飞鸟图片和用labelimg标注的标签文件，其中共有4850张图片和对应标签。为了保证算法在实际场景中的应用，数据集不仅仅包括不同鸟类的近照，还包含飞翔的鸟类和鸟群的图片。

3.2 实验环境与参数配置

本文的实验环境为：操作系统：Windows11系统，CPU：13th Gen Intel® Core™ i7-13700H，GPU：NVIDIA GeForce RTX 4060。编程语言采用Python3.8，深度学习框架采用PyTorch，在Anaconda中搭建实验环境，在PyCharm平台上进行实验。

3.3 评价指标

本实验使用精确率Precision、召回率Recall、mAP作为评价模型检测性能的评价指标^[8]。Precision是指在所有模型预测为正类别的样本中，实际是正类别的比例。Precision衡量了模型在正类别预测中的准确性。Precision的计算公式：

$$Precision = \frac{TP}{TP + FP} \quad (18)$$

式中： TP 表示模型正确预测为正类别的样本数量， FP 表示模型错误预测为正类别的样本数量。

Recall是指在所有实际正类别的样本中，模型正确预测为正类别的比例。Recall衡量模型是否能够有效地找到所有正类别样本。Recall的计算公式：

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

其中： FN 表示模型错误预测为负类别的样本数量。

mAP是目标检测任务中常用的评估指标，它综合了模型在不同类别上的精确率-召回率曲线(Precision-Recall Curve)并计算了平均值。mAP衡量了模型在多类别上的检测性能，用于评估目标检测算法的整体效果。

3.4 实验结果与分析

为了证明本文方法的有效性，本文以未改进的YOLOv8为基线模型，加上GOLD-YOLO、

ASF-YOLO 和本文的融合算法相对比进行消融实验。在测试集上的实验结果见表1。引入GD和ASF模块后，Precision、Recall和mAP比原YOLOv8依次提升了1.6、1.3和0.6个百分点。

表1 消融实验结果

YOLOv8	GD	ASF	Precision	Recall	mAP
√			0.883	0.709	0.805
√	√		0.868	0.715	0.789
√		√	0.889	0.696	0.807
√	√	√	0.899	0.722	0.811

为了更加全面地展现算法的训练过程，对本文方法在150轮训练过程中的Precision、Recall、mAP_{0.5}(IoU阈值为0.5)、mAP_{0.5:0.95}(IoU阈值从0.5到0.95)变化情况进行分析。横坐标是epoch轮数，纵坐标分别表示上述的四个指标。从图3可以看出，150epoch时模型已经收敛。

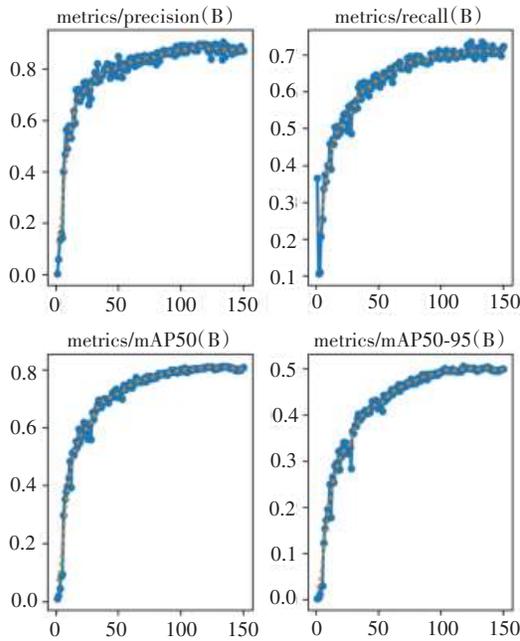
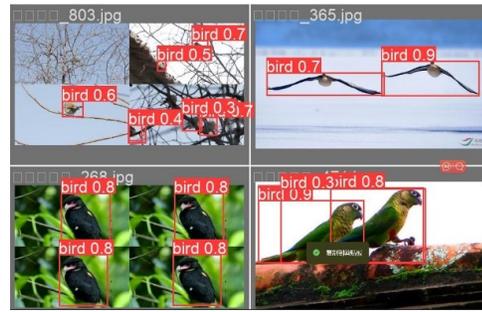
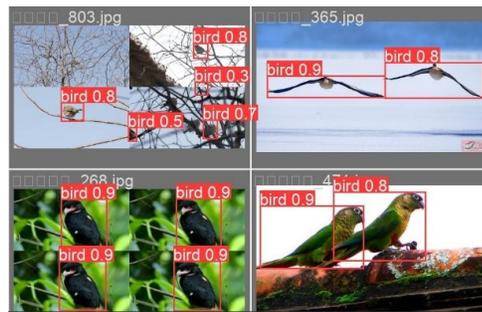


图3 训练结果

为了检验本文算法的实际检测效果，实验使用YOLOv8和本文融合改进的算法进行对比检测。不同鸟类的不同形态的检测对比结果如图4所示。两个模型都能正确检测出鸟类，但改进后的模型检测结果具有更高的置信度。



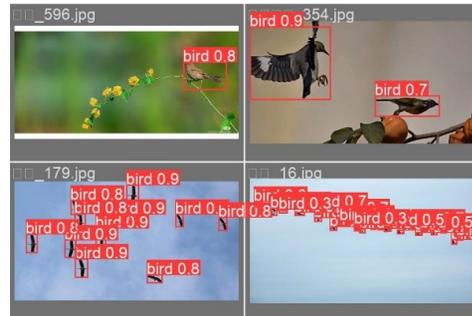
(a) 改进前



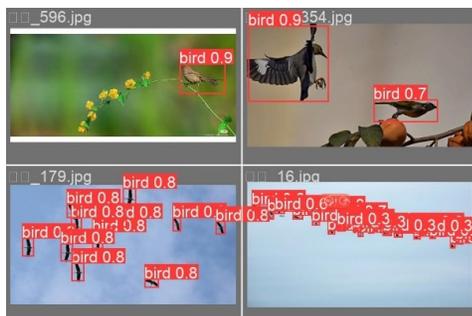
(b) 改进后

图4 不同鸟类形态的检测效果图

群鸟的检测对比结果如图5所示。原模型未出现明显漏检的情况，改进前后的模型能全部检测到所有小目标，改进后模型检测到的置信度要高于原模型。



(a) 改进前



(b) 改进后

图5 群鸟的检测效果图

综上，本文改进后的算法相比于原模型有更好的检测精度，对于不同类型、不同形态的鸟类及群鸟的检测效果都更加精确。在不同的评价指标Precision、Recall、mAP上，改进后模型的表现都要更佳。

4 结语

为了解决复杂背景下多变姿态的飞鸟检测精度低的问题，本文提出了对YOLOv8算法的改进方案，旨在提升算法在飞鸟检测任务中的性能。在算法改进方面，本文融合了GOLD-YOLO中的GD机制和ASF-YOLO中的尺度序列特征融合(SSFF)模块，来增强多尺度特征融合能力，提升小尺寸目标的特征表示。此外，为了支持飞鸟检测任务的研究，本文还收集了一个飞鸟数据集。该数据集不仅包含了不同鸟类的近照，还涵盖了飞翔的鸟类和鸟群的图片，为算法的训练和测试提供了丰富的数据资源。通过实验验证，本文改进的YOLOv8算法在飞鸟检测任务中取得了性能提升。

参考文献：

- [1] 陈天华,朱家焯,印杰. 基于注意力机制的鸟类识别算法[J]. 计算机应用,2024,44(4):1114-1120.
- [2] 王禹涵,张凯兵,蒋睿涛,等. 自然场景下飞鸟目标

检测与实时跟踪[J]. 湖北工程学院学报,2023,43(6):53-58.

- [3] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [4] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,2023:7464-7475.
- [5] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation [C] //Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition,2018:8759-8768.
- [6] WANG C, HE W, NIE Y, et al. Gold-YOLO: efficient object detector via gather-and-distribute mechanism [C] //Proceedings of the Conference on Neural Information Processing Systems,2024,36.
- [7] Kang M, Ting C M, Ting F F, et al. ASF-YOLO: A novel YOLO model with attentional scale sequence fusion for cell instance segmentation [J]. Image and Vision Computing,2024,147:105057.
- [8] 罗磊,谢竹逵. 基于改进YOLOv8的交通标志检测算法[J]. 机电工程技术,2024,53(3):205-210.

Flying bird detection algorithm based on an improved YOLOv8

Chen Qian, Lu Yang*, Shao FeiXiang Li Shiyan

(College of Information Engineering, Yangzhou Vocational University, Yangzhou 225009, China)

Abstract: In this paper, we aim to improve the accuracy of the flying bird detection task, and propose an improved YOLOv8 algorithm based on deep learning to address the shortcomings of traditional algorithms in the detection of flying birds with complex backgrounds and multi-variable attitudes. By fusing the GD mechanism in GOLD-YOLO and the Scale Sequence Feature Fusion (SSFF) module in ASF-YOLO, the network structure is optimized, the feature extraction and fusion methods are improved, the multi-scale feature fusion capability is enhanced, and the feature representation of small-size targets is improved. In addition, a flying bird dataset containing close-up photos of different birds, flying birds and flock images is collected to support the training and testing of the algorithm. The experimental results show that the improved YOLOv8 algorithm improves the performance in the flying bird detection task, especially when dealing with small targets and complex backgrounds.

Keywords: target detection; YOLO; feature extraction

文章编号: 1007-1423(2024)20-0020-06

DOI: 10.3969/j.issn.1007-1423.2024.20.004

轻量化改进 YOLOv5s 无人机遥感图像识别

杨磊¹, 田立勤^{1*}, 吴骏一², 彭凯远¹

(1. 华北科技学院计算机学院, 廊坊 065201; 2. 青海师范大学计算机学院, 西宁 810016)

摘要:近年来,“无人机+”消防、“无人机+”巡检、“无人机+”应急等领域广泛应用无人机,对无人机遥感图像进行识别是在各领域实现“无人机+”的关键。但是,无人机遥感图像复杂的信息,大量的小目标给目标检测带来了巨大的挑战,随着深度学习时代的到来,在目标检测性能提升的同时,随之而来的是模型参数的大量增加。基于此,以一种轻量化的 YOLOv5s 改进模型应用到无人机遥感图像识别当中,在提高目标检测性能的同时尽可能减少模型的参数量。实验结果表明,改进后的模型较原始的 YOLOv5s 模型 mAP₅₀ 和 mAP_{50:95} 分别提高了 7.9 和 4.3 个百分点,模型参数量只增加 0.14 M。

关键词: YOLOv5s; 轻量化; 无人机; 图像识别; 目标检测

0 引言

无人机遥感图像相较于地面拍摄的图像存在着如目标尺度不均衡、空间分布不均衡、样本数量不均衡、类别语义不均衡和优化目标不均衡五大不均衡问题^[1]。无人机遥感图像中复杂的背景、大量的小目标、云雾、光照、遮挡、正样本只覆盖整张图片中的一小部分等问题都会给无人机遥感图像的识别带来巨大的挑战。

随着深度学习时代的到来,很大程度上提升了目标检测的性能,基于深度学习的目标检测分为单阶段目标检测和双阶段目标检测^[2]。双阶段目标检测如 Faster-RCNN^[3],其工作原理是先用卷积神经网络(CNN)来提取特征,然后在图像上生成一些候选区域,在候选区域里面进行分类和回归,从而提高目标检测的性能。卷积神经网络是一种多层前馈神经网络,它可以直接对二维图像进行处理,能够有效地从大量样本中学习相应的特征,避免了复杂的特征提取过程^[4]。单阶段目标检测如 YOLO^[5],其工作原理是在进行目标检测时,特征提取网络

先将输入图像划分为 $S \times S$ 个网格单元,然后对每一个网格进行边框预测,最后去除冗余边框和阈值比较低的边框。使用 YOLO 进行目标检测不需要在图像上生成候选区域,只需要进行一次前馈网络,因此具有更快的检测速度。

由于无人机航拍的角度、拍照环境中光照强度的影响,导致无人机遥感图像中目标检测识别率不高^[6]。基于此,本文对单阶段目标检测中的 YOLOv5s 模型进行改进,在颈部网络(Neck)中替换原始模型中的 Upsample 上采样层为 CARAFE 上采样,在原始模型的损失函数中引入 OTA(Optimal Transport Assignment)损失函数,提出以一种轻量化 YOLOv5s 改进应用到无人机遥感图像的识别当中。改进后的模型在 mAP₅₀ 和 mAP_{50:95} 均有提升,在模型参数量(Params)上也只有轻微的涨幅。

1 YOLOv5s 网络结构

YOLOv5s 的网络结构由输入端(Input)、骨干网络(Backbone)、颈部网络(Neck)和头部网络(Head)四部分组成^[7],网络结构如图 1 所示。

收稿日期: 2024-04-02 修稿日期: 2024-09-24

作者简介: 杨磊(2000—),男,河南驻马店人,硕士研究生,研究方向为图像融合、目标检测; *通信作者: 田立勤(1970—),男,陕西定边人,教授,研究方向为网络安全评价与用户行为认证, E-mail: yl488248869@163.com; 吴骏一(1996—),男,河北承德人,博士研究生,研究方向为应急信息化与智能化、算法优化; 彭凯远(2000—),男,山西晋城人,硕士研究生,研究方向为计算机视觉

精确率和召回率的公式如下：

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

式中： TP 表示正样本中被预测为正； FP 表示负样本中被预测为正； FN 表示正样本中被预测为负。均值平均精度是指所有类别的平均精度求和后与所有类别的比值。均值平均精度的公式如下：

$$mAP = \frac{\sum_{i=1}^k AP_i}{k} \quad (5)$$

4.2 对比实验

为验证改进后模型的目标检测性能，同时选取当下主流的目标检测模型，如YOLOv3、YOLOv5s和YOLOv8s来进行对比实验。不同算法之间的对比实验结果见表2。

表2 对比实验结果

Methods	P/%	R/%	mAP ₅₀ /%	mAP _{50, 95} /%	Params/M	GFLOPs/G
YOLOv3	26.9	17.9	15.0	6.6	8.68	12.9
YOLOv5s	43.9	33.8	33.0	18.1	7.03	15.8
YOLOv8s	51.4	39.7	40.6	24.2	11.13	28.5
Ours	52.0	39.8	40.9	22.4	7.17	16.1

可以看出改进后的模型较原始模型相比在IoU阈值为0.5的均值平均精度和IoU阈值为0.5到0.95的均值平均精度分别提高7.9和4.3个百分点，同时模型参数量和GFLOPs分别只增加0.14 M和0.3 G。改进后的模型较YOLOv8s模型相比，IoU阈值为0.5的均值平均精度上提高了0.3个百分点，IoU阈值为0.5到0.95的均值平均精度提高0.2个百分点，同时在模型参数量和GFLOPs均比YOLOv8s模型要低。

4.3 消融实验

为了验证YOLOv5s算法在相同参数和环境配置下引入OTA损失函数和CARAFE上采样模块对目标检测性能的影响，对其进行消融实验，消融实验结果见表3。

表3 消融实验结果

Methods	P/%	R/%	mAP ₅₀ /%	mAP _{50, 95} /%	Params/M	GFLOPs/G
YOLOv5s	43.9	33.8	33.0	18.1	7.03	15.8
+CARAFE	44.2	34.2	33.1	18.1	7.17	16.1
+OTA	51.0	39.9	40.6	22.3	7.03	15.8
Ours	52.0	39.8	40.9	22.4	7.17	16.1

表中第一行是原始的YOLOv5s算法，将其作为消融实验的基准线。第二行是在原始YOLOv5s模型的基础上替换其Upsample上采样为CARAFE上采样，替换后的模型在IoU阈值为0.5的均值平均精度上提高了0.1个百分点，模型参数量上增加了0.14 M。第三行是在原始YOLOv5s模型的基础上引入OTA损失函数，替换后的模型在IoU阈值为0.5的均值平均精度上提高了7.5个百分点，IoU阈值为0.5到0.95的均值平均精度提高4.2个百分点，模型参数量上并未增加。最后一行是在原始YOLOv5s模型的基础上引入OTA损失函数并替换其Upsample上采样为CARAFE上采样，改进后的模型在IoU阈值为0.5的均值平均精度和IoU阈值为0.5到0.95的均值平均精度分别提高7.9和4.3个百分点，并且未增加模型参数量。

4.3 检测结果

为了验证改进后的模型较原始模型的目标检测性能，特别选取同一场景下不同视角的无人机遥感图片做目标检测可视化结果分析，如图5所示。

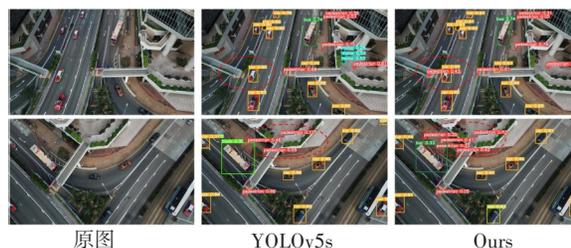


图5 可视化结果对比

由图5可以看出，原始YOLOv5s模型对小目标感知能力较弱，从而导致漏检和误检的情况发生。改进后的模型可以自动进行目标检测

标签分配并且具有更大的感受野,增强了模型对目标特征信息的提取能力,从而降低了对小目标的漏检率和误检率,提高了模型的目标检测性能。

5 结语

针对无人机遥感图像复杂的信息、小目标数量大、目标检测难度大和检测模型参数量大的问题,提出一种轻量化的YOLOv5s改进模型应用到无人机遥感图像识别中。通过在原始YOLOv5s模型中引入OTA损失函数和替换Upsample上采样为CARAFE上采样,可以在提升目标检测性能的同时不会增加太多模型参数量。实验结果表明,改进后的YOLOv5s模型,在VisDrone2019数据集上 mAP_{50} 可达40.9%,相较于原始YOLOv5s模型提高7.9个百分点,Params为7.17 M,相较于原始YOLOv5s模型只增加0.14 M。与其他算法相比,改进后的YOLOv5s模型在 mAP_{50} 和 $mAP_{50:95}$ 均有提高,增强了模型对目标特征信息的提取能力,从而降低了对小目标的漏检率和误检率,提高了模型的目标检测性能。但是模型对于无人机图像尺度变化导致的类别语义不平衡问题的解决能力有待提高,这将是下一阶段研究的重点。

参考文献:

- [1] 冷佳旭,周应华,叶永明,等. 无人机视角下的目标检测研究进展[J]. 中国图象图形学报,2023,28(9):2563-2586.
- [2] 郭伟,申磊,曲海成,等. 自适应权重金字塔和分支强相关的SAR图像舰船检测[J]. 中国图象图形学报,2022,27(10):3127-3138.
- [3] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [4] 桂方燧,武文星,任维康. 基于PSO-CNN神经网络的车牌识别系统[J]. 华北科技学院学报,2021,18(5):100-106.
- [5] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016:779-788.
- [6] 梁永春,田立勤,陈楠,等. 基于航拍图像的人员实时搜寻机制与分析[J]. 华北科技学院学报,2020,17(2):98-105.
- [7] 张光华,李聪发,李钢硬,等. 基于改进YOLOv7-tiny的无人机航拍图像小目标检测算法[J]. 工程科学与技术:1-14.[2024-09-24]. <https://doi.org/10.15961/j.jsuese.202300593>.
- [8] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C] // Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 936-944.
- [9] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation [C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018:8759-8768.
- [10] GE Z, LIU S, LI Z, et al. OTA: optimal transport assignment for object detection [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021:303-312.
- [11] KNIGHT P A. The Sinkhorn-Knopp algorithm: convergence and applications [J]. SIAM Journal on Matrix Analysis and Applications, 2008, 30(1): 261-275.
- [12] WANG J, CHEN K, XU R, et al. Carafe: content-aware reassembly of features [C] // Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019:3007-3016.
- [13] 吕恩辉,王雪松,程玉虎. 基于反卷积特征提取的深度卷积神经网络学习[J]. 控制与决策, 2018, 33(3):447-454.
- [14] DU D, ZHU P, WEN L, et al. VisDrone-DET2019: the vision meets drone object detection in image challenge results [C] // Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, Seoul, Korea (South), 2019:213-226.
- [15] 袁柱柱. 基于深度学习的行人检测与计数[D]. 南京:南京邮电大学,2021.

(下转第30页)

文章编号: 1007-1423(2024)20-0025-06

DOI: 10.3969/j.issn.1007-1423.2024.20.005

基于 Faster R-CNN 的低照度人脸检测方法研究

岑锐强^{1*}, 冯 广²

(1. 广东工业大学计算机学院, 广州 510006; 2. 广东工业大学自动化学学院, 广州 510006)

摘要: 随着深度学习的发展, 人脸检测算法在理想的环境下检测的精准率和检测速度已经达到了相当优秀的水平。但现实应用场景中很难达到理想中的检测环境, 低光照、图片模糊产生干扰, 导致人脸检测的精准度下降。为了解决这一问题, 提出了一种结合 MSRCR 算法的改良 Faster-RCNN 的人脸检测方法。首先针对低照度和图片模糊问题提出了使用 MSRCR 算法对图片进行自动白平衡(AWB)。最后使用 Soft NMS 对 Faster-RCNN 的候选框算选进行优化, 解决人脸重叠检测目标丢失的问题。实验结果表明, 低光照或像模糊的情况下结合 MSRCR 算法的改良 Faster R-CNN 对比传统的人脸检测有着更高的检测精准度。

关键词: 人脸检测; 深度学习; 图像增强; 低照度

0 引言

人脸目标检测是计算机视觉的一个重要研究方向。在现实生活中, 人脸检测广泛应用于实时监控、支付验证和目标监控等领域。目前, 许多现有的人脸检测算法在理想条件下可以达到较高的精度, 但这对检测到的图像有一些要求。在真实的检测场景中, 检测到的图像会存在光照过低、人脸尺度不一致、图像模糊等问题。起初, Viola 等^[1]提出 VJ 检测器, VJ 检测器的核心在于运用类 Haar 特征^[2], 旨在精准描述图像中的局部灰度变化。类 Haar 特征的基本形式丰富多样, 比如边缘特征、线性特征、中心特征以及对角线特征等, 这些特征能够全面捕捉图像中不同方向的灰度变化信息。但这种传统的机器学习方法需要设计有效的目标特征, 影响了整体效果。随着近年来深度学习的发展, 深度卷积网络在图像切割、分类和目标检测方面取得了许多研究成果。与传统检测方法相比, 深度学习具有良好的自适应能力, 提高了模型的通用性。因此, 基于深度学习的人脸检测方法成为一个重要的研究课题。

目前, 关于目标检测的研究很多。RCNN 算法是常用的检测模型之一。Girshick 等^[3]提出的 R-CNN 在目标检测方面具有较高的准确性。在此基础上, Chen 等^[4]提出了结合区域建议网络和 RCNN 的 STN(Srvised Transformer Network) 检测算法, 可以实现场景中的大姿态人脸检测。通过在原始网络中添加 RPN 区域选择网络来改进 RCNN, 提高了模型检测的准确性。Chen^[5]提出了引入特征金字塔的方法, 并使用网络不同深度的特征图进行检测, 使模型可以在多个尺度上检测人脸。尽管上述改进通过增加网络的区域选择或筛选无用的候选框来提高网络的准确性, 但它们实际上并没有解决自然环境对检测的影响。

1 实验工作

本文采用了 Faster R-CNN 目标检测模型进行低照度的人脸检测研究, Faster R-CNN 可以被视由两部分网络组成: Fast R-CNN 和 RPN (Region Proposal Network), 加入区域生成网络 RPN 是 Faster R-CNN 对比于其他检测模型具有的优点。虽然 Faster R-CNN 依然使用的是滑动

收稿日期: 2024-02-29 修稿日期: 2024-07-10

作者简介: *通信作者: 岑锐强(1999—), 男, 广西贵港人, 硕士研究生, 研究方向为计算机视觉人脸目标检测, E-mail: 1953027864@qq.com; 冯广(1973—), 男, 广东云浮人, 博士, 教授级高级实验师, 研究方向为网络控制、机器学习、大数据

窗口在特征图上依次遍历生成候选区域，但是 Faster R-CNN 将特征抽取、边框回归及分类都整合在一个网络中，有效提高检测精度和检测效率，在检测速度方面提升尤为明显。

Faster R-CNN 采用了双网络共同训练的模式，该模式不仅可以加快网络模型的检测时间而且可以提高模型整体的检测精度。与其他检测模型相比，同样使用 Wider Face 数据集进行训练对比其他的人脸检测网络，Faster R-CNN 拥有最高的检测速度和精度。

本文提出了一种基于 Faster R-CNN 检测网络和 MSRCR^[6] 图像增强算法的改进方法。该方法使用 MSRCR 算法增强初始图像，MSRCR 使用颜色恢复因子调整图像以避免失真。为了避免网络特征图中包含的语义信息太深而无法正确表达人脸，实验中使用 ResNet-101 代替 VGG-16。在此基础上，研究进一步使用 ROI align 代替原始候选框生成，并使用软非最大值抑制 (SOFT-NMS)^[7] 代替原始网络的 NMS 来提高 RPN^[8] 网络的筛选精度。实验表明，改进后的 Faster R-CNN 在相同环境下具有更高的检测精度。

2 检测模型改进

2.1 基于 MSRCR 图像增强算法

在目标检测网络中，要是直接将质量较差的人脸图像转化为高维的特征图进行人脸目标检测，得到的效果并不理想，经常出现人脸于环境模糊色调交叉从而无法被检测出来。针对上述存在的问题，本文采用对图片进行光照补强提高图像辨识度的方法，再进行随后的操作。常用的图像增强算法中 RetiNex 算法^[9] 属于效果比较优秀的一种，但是 RetiNex 算法处理过的人脸图像会出现颜色失真以及对比度低的问题，为此本文采用了带色彩恢复的多尺度视网膜增强算法 MSRCR。MSRCR 算法是在 RetiNex 算法上进行优化，其理论依据与 RetiNex 算法相同，认为图片中物体的亮度取决于环境的照亮和物体表面对光照的反射，其数学表达式为

$$I(x, y) = L(x, y) * R(x, y) \quad (1)$$

式中： $I(x, y)$ 代表被观察到或图像记录的信号； $L(x, y)$ 代表环境中光照射的分量； $R(x, y)$ 代表物体的反射分量。RetiNex 算法的核心思想就是

抛开环境中光照的分量，保留物体本身的折射分量，从而得到物体的原本外貌。所以算法对上式两边同时取对数，关系式如下：

$$\lg(I(x, y)) = \lg(L(x, y)) - \lg(R(x, y)) \quad (2)$$

针对于我们已知的一副图像数据 $I(x, y)$ ，计算得到的 $R(x, y)$ 被认为是 $R(x, y)$ 增强后得到的效果，主要问题就是如何得到 $L(x, y)$ 环境的光照分量。Rahman^[10] 等提出了使用高斯卷积可以从已知图像 $S(x, y)$ 估算出亮度分量 $L(x, y)$ ，其数学公式如下：

$$L(x, y) = R(x, y) * G(x, y) \quad (3)$$

$$G(x, y) = k * e^{-\frac{(x^2 + y^2)}{\sigma^2}} \quad (4)$$

然后将 $\text{Log}(R(x, y))$ 的输出值进行量化操作，约束在 $[0, 255]$ 的像素值之内，其量化方法就是计算出 $\text{Log}(R(x, y))$ 的最大值和最小值，然后对每一个 $\text{Log}(R(x, y))$ 进行线性量化，其数学公式如下：

$$R(x, y) = \frac{[\lg R(x, y) - \lg \min(x, y)]}{[\lg \max(x, y) - \lg \min(x, y)]} \quad (5)$$

为了进一步优化图片的呈现效果，我们在单尺度 RetiNex 算法的基础上，提出了多尺度视网膜增强算法 MSR (Multi-Scale RetiNex)。MSR 算法的核心理念在于，它首先对图像的 RGB 三个通道分别进行高斯模糊处理，随后再将各个通道的处理结果进行加权求和。通过这一改进方法，我们得以在多个尺度上提取图像特征，从而更加精准地增强图像质量。具体的算法表达式为

$$R_{\text{MSR}}(x, y) = \sum_{i=1}^n W_i * R_i \quad (6)$$

式中： W 是图像第 i 层的加权系数，所有权重之和为 1，通常会使用平均权重。该算法对灰度图像的处理具有很好的效果，但在针对彩色图像进行增强时可能会造成图像的失真，不能呈现图像中原本的色彩效果。

MSRCR 算法是在 MSR 的基础之上加入了色彩恢复因子 C 来调节由于图像局部区域对比度增强而导致颜色失真的缺陷，对 MSR 的结果做了色彩平衡、归一化、增益和偏差线性加权。MSRCR 算法在计算得到 RGB 各个通道的 $\text{Log}(R(x, y))$ 之后，再分别计算出各个通道的均值 Mean 和均方

差 Var 。然后设置动态参数 $Dynamic$ ，计算每个通道的 min 和 max 值，数学公式如下：

$$\lg[R(x, y)]_{\max} = Mean + Dynamic * Var \quad (7)$$

$$\lg[R(x, y)]_{\min} = Mean - Dynamic * Var \quad (8)$$

最后用式(4)进行线性映射，映射同时需要注意添加一个溢出判断：

$$R(x, y) = \begin{cases} 255, & R(x, y) > 255 \\ R(x, y), & 255 > R(x, y) > 0 \\ 0, & R(x, y) < 0 \end{cases} \quad (9)$$

在基于RGB的三通道的MSRCR图像增强算法，我们将通道数 $scale$ 固定为3，在此基础上通过实验对比不同动态参数对图像优化效果的影响。实验针对MSRCR算法使用不同的动态参数 $Dynamic$ 进行图像增强。图1给出了对比实验结果，动态参数取值较小，夜色街景图像的增强效果越好。当动态参数取值为3时，为本实验应用场景的最佳动态参数取值，此时MSRCR算法既能针对夜景图像进行最大程度的清晰度增强，又避免引入其他噪声对图像的影响。



图1 MSRRR图像增强算法不同参数效果对比图

2.2 RPN网络优化

在Faster R-CNN的RPN网络筛选候选框的过程中，图像经过区域生成网络RPNHead会生成一定数量的候选框集合 B ，以及每个候选框对应的分数 S 。每次选中候选框集合 B 中分数最高的候选框 M 将其移入集合 D 中，将 M 与集合 B 中所有的候选框做重叠判断，如果重叠率，即 IoU 大于给定的额度，就将其分数赋值为0。

NMS算法中分数重置函数如下：

$$S_i = \begin{cases} S_i, & IOU(M, b) < N \\ 0, & IOU(M, b) \geq N \end{cases} \quad (10)$$

式中： IoU 表示两个候选框之间的重叠率； N 是设定的阈值。通过非极大值抑制算法可以将初始的候选框数量筛除大部分重叠的候选框，但该方法依旧存在一个缺点。如图2所示，如果检测图像中两个或多个人脸出现重叠现象，一些正确框选人脸的候选框分数将会被重置为0，最终导致人脸无法被检查出来，降低了算法的平均检测率(mAP)。



图2 NMS算法过滤候选框

针对NMS存在的问题，本文采用了另一种处理候选框重叠的方法Soft-NMS算法。与NMS相似，Soft-NMS算法同样需要将当前集合得分最高的候选框与其他候选框进行 IoU 计算，并且判断二者的 IoU 是否超过阈值。而对比于NMS，Soft-NMS算法没有直接将超过重叠率阈值得分低的候选框分数重置为0，而是根据重叠率对重叠的候选框的得分进行线性衰减。Soft-NMS的分数衰减函数如下：

$$S_i = \begin{cases} S_i, & IOU(M, b) < N \\ S_i - IOU(M, b), & IOU(M, b) \geq N \end{cases} \quad (11)$$

Soft-NMS相比传统的NMS方法可以保留更多的检测目标(见图3)，传统的NMS方法会直接删除重叠的边界框，导致一些可能是真实目标的边界框被错误地删除。而Soft-NMS通过降低重叠边界框的得分，而不是直接删除它们，从而保留了更多的目标。这样可以提高目标检测的召回率，准确地检测到所有目标。而且Soft-NMS通过降低重叠边界框的得分，使得分较高的边

界框在非极大值抑制过程中仍然保留一定的权重，从而提高了目标的定位精度。Soft-NMS通过降低重叠边界框的得分，使得分较高的边界框在非极大值抑制过程中仍然保留一定的权重，从而提高了目标的定位精度。



图3 Soft-NMS算法过滤候选框

3 实验数据和结论

3.1 实验数据集

本次研究使用VOC2012作为模型的预训练数据集以及DarkFace_Train_2021中共计6000张街道夜景图片作为实验的数据集，每张图片的尺寸1080×720，其中每张街道夜景的图像都是背景复杂、对比度强、清晰度高。在DarkFace_Train_2021数据集中，实验随机选取1000张街道夜景图片作为检测模型的测试集，其余的5000张图片作为训练集，并从训练集中随机选500张夜景图片作为本次实验的验证集。

3.2 实验结果和分析

从表1可以看出，相较于直接使用DarkFace数据集、直接训练Faster R-CNN检测模型，预先使用MSRCR图像增强算法对图片进行增强处理，对模型检测的精准度有显著的提升。相对于使用ResNet101而言，使用VGG16作为特征提取网络的网络规模以及网络的深度较少，得到的图像特征网络包含的信息更少，使得整体模型的检测精准度低于ResNet101。直接使用Brightness算法对图片亮度进行25%的提高时，相对比直接使用源数据集进行训练，两个不同提取网络的模型都有精准度的提高，但提高数

值并不明显，因为Brightness算法仅提高了图像的曝光度，没有像MSRCR算法整体提高图像的清晰度。

在使用MSRCR算法进行图像增强后，使用不同特征提取网络的两个Faster R-CNN目标检测模型都有大幅度的效果提高，其中提升效果最优的是使用了ResNet101作为图像特征网络提取网络的检测模型。这是由于ResNet101^[11]相较于VGG16具有较多的网络层次，拥有更多的网络参数，需要更多的图像数据，而MSRCR图像增强原理是在一定程度上增加原本图像中的数据量。另一方面，虽然网络层次越深，网络就可以提取到更加丰富以及抽象的信息，但是随着网络层次的愈加增多，梯度消失或者梯度爆炸的问题也会随之出现。ResNet101引入了残差结构解决了退化问题，从而使得网络可以提取更加抽象的信息。因此使用ResNet101相对比VGG16具有更高的精准度。

此外，实验将原Faster R-CNN模型预先使用VOC2012数据集进行预训练，再进行迁移学习，用DarkFace进行进一步的优化训练。迁移学习的本质就是将先前学习到的知识迁移到新的任务上，从而加快新模型的训练速度并提高性能，所以我们可以冻住模型中部分网络层候选更加精细的调整训练。经过多次实验研究，统计结果表明，将学习率调整为 10^{-4} 进行迁移学习，模型的mAP提升最大，使用VGG16作为特征提取网络的模型mAP为0.831，使用ResNet的模型提升效果最大，mAP为0.925。

表1 实验模型图像检测的mAP结果

Enhancement	学习率	VGG16	ResNet101
False	0.001	0.708	0.791
Brightness	0.001	0.764	0.803
MSRCR	0.001	0.824	0.913
MSRCR	0.0001	0.831	0.925

图4和图5是Faster R-CNN网络的最终检测结果，可以看出，使用MSRCR算法的Faster R-CNN检测模型对于街道夜景的人脸目标检测结果有大幅度的提高。与传统的机器学习相比，研究中使用的改进Faster R-CNN目标检测模型

通过使用MSRCR对检测图片进行增强, 从而使模型能够从网络转化的图像特征中获取低耦合、高区分度的候选目标特征, 而且避免了候选目标筛选算法检测目标丢失问题的发生, 整体模型的检测精准度已经得到了很大程度的提升。



图4 夜景人脸检测效果对比图

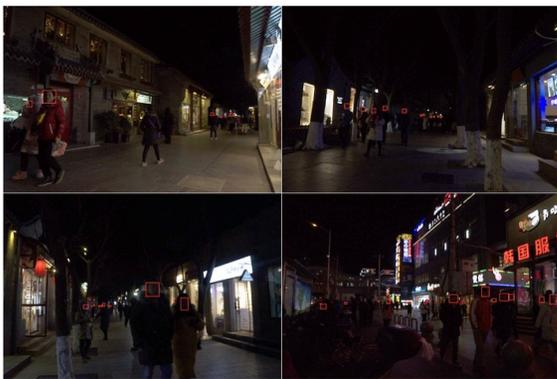


图5 街道人脸检测效果图

4 结语

本文提出了一种基于MSRCR算法的图像增强的改进Faster R-CNN低照度人脸检测模型, 并对该模型进行了分析和测试。根据实验结果, 我们认为MSRCR算法对低光照下的人脸目标检测有很好的优化效果。同时改进了模型原始的RPN网络, 通过引入更好的候选目标框算法, 减少了模型在检测过程中的重叠目标无法全部检测的现象。与其他模型相比, 本实验改进的Faster R-CNN在低照度、多检测目标的应用场景具有更高的检测精度以及更快的检测效率。

参考文献:

[1] VIOLA P, JONES M. Rapid object detection using a boosted cascade of simple features [C] // Proceedings of the 2001 IEEE Computer Society Conference on

Computer Vision and Pattern Recognition (CVPR 2001). IEEE, 2001, 1: I-I.

- [2] 陈锻生, 陈齐松, 刘政凯. 基于类灰度图的类Haar特征构建及其应用[J]. 郑州大学学报(理学版), 2007(1): 33-39.
- [3] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.
- [4] CHEN D, HUA G, WEN F, et al. Supervised transformer network for efficient face detection [C] // Proceedings of the 14th European Conference Computer Vision, Amsterdam, The Netherlands, 2016: 122-138.
- [5] CHEN K, CAO Y, LOY C C, et al. Feature pyramid grids [EB/OL]. arXiv: 2004.03580, 2020.
- [6] JOBSON D J, RAHMAN Z, WOODDELL G A. A multiscale retinex for bridging the gap between color images and the human observation of scenes [J]. IEEE Transactions on Image processing, 1997, 6 (7) : 965-976.
- [7] BODLA N, SINGH B, CHELLAPPA R, et al. Soft-NMS-improving object detection with one line of code [C] // Proceedings of the IEEE International Conference on Computer Vision, 2017: 5561-5569.
- [8] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39 (6) : 1137-1149.
- [9] 赵宏宇, 肖创柏, 禹晶, 等. 基于WLS的Retinex单幅夜间彩色图像增强方法[J]. 北京工业大学学报, 2014, 40(3): 404-410.
- [10] RAHMAN Z U, JOBSON D J, WOODDELL G A. Retinex processing for automatic image enhancement [C] // Human Vision and Electronic Imaging VII. International Society for Optics and Photonics, 2004: 100-110.
- [11] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.

Research on low-light face detection method based on Faster R-CNN

Cen Ruiqiang^{1*}, Feng Guang²

(1. School of Computers, Guangdong University of Technology, Guangzhou 510006, China;
2. School of Automation, Guangdong University of Technology, Guangzhou 510006, China)

Abstract: With the development of deep learning, face detection algorithms have reached an excellent level of accuracy and speed in ideal environments. However, in real-world scenarios, it is difficult to achieve the ideal detection environment, and factors such as low lighting and blurred images can cause interference, leading to a decrease in the accuracy of face detection. To address this issue, a face detection method based on an improved Faster-RCNN combined with the MSRCR algorithm is proposed. Firstly, the MSRCR algorithm is used to perform automatic white balance (AWB) on the images to address the issues of low lighting and blurring. Finally, Soft NMS is employed to optimize the candidate box selection in Faster-RCNN, solving the problem of missing overlapping detection targets. Experimental results show that the improved Faster R-CNN combined with the MSRCR algorithm has higher detection accuracy compared to traditional face detection methods in low-light or blurred conditions.

Keywords: face detection; deep learning; image enhancement; low illumination

(上接第24页)

Application to UAV remote sensing image recognition with a lightweight improved model of YOLOv5s

Yang Lei¹, Tian Liqin^{1*}, Wu Junyi², Peng Kaiyuan¹

(1. School of Computer, North China Institute of Science and Technology, Langfang 065201, China;
2. School of Computer, Qinghai Normal University, Xining 810016, China)

Abstract: In recent years, UAV have been widely used in various fields such as “UAV +” firefighting, “UAV +” inspection, “UAV +” emergency response, etc. The recognition of remote sensing images of UAVs is the key to realize “UAV+” in various fields. However, the complex information of UAV remote sensing images and a large number of small targets bring great challenges to target detection. With the arrival of the deep learning era, while the performance of target detection is improved, the number of model parameters increases substantially. Based on this, in this paper, a lightweight YOLOv5s improved model is applied to UAV remote sensing image recognition, which improves the target detection performance while reducing the number of model parameters as much as possible. The experimental results show that the improved model improves mAP_{50} and $mAP_{50:95}$ by 7.9 and 4.3 percentage point, respectively, compared with the original YOLOv5s model, and the number of model parameters increases by only 0.14 M.

Keywords: YOLOv5s; lightweight; UAV; image recognition; target detection

文章编号: 1007-1423(2024)20-0031-05

DOI: 10.3969/j.issn.1007-1423.2024.20.006

基于深度学习的蚕病检测算法设计

曾沛杰, 周卫*, 陈金良

(广西民族大学人工智能学院, 南宁 530006)

摘要: YOLOv8应用于蚕病识别任务时, 蚕病症的形态和外观会因为处于不同的阶段发生变化, 导致模型容易受到背景信息的干扰, 为了增强模型捕获不同阶段蚕病的特征的能力, 提出了一种名为DCNv2-Block的新特征提取模块。实验结果表明, 相比YOLOv8原模型, 提出的改进方法的mAP0.5达到96.5%, 提高了1.1个百分点, 能更有效捕捉蚕病的病症, 提高了蚕病害识别的准确性, 有助于及时发现病害并采取控制措施。

关键词: YOLOv8; 蚕病检测; 可变形卷积; 深度学习

0 引言

蚕在我国是一种重要的经济昆虫, 它的历史可以追溯到新石器时代。蚕丝可以用来生产丝绸, 是中国古代一项重要出口商品, 也是古代经济发展的重要支柱之一。一直以来, 蚕的养殖是中国农业的重要组成部分。随着我国西部大开发战略的实施和“东桑西移”工程的推动, 广西壮族自治区的蚕桑产业迎来了新一轮快速发展期, 广西成为全国最大的蚕茧生产基地和茧丝加工基地^[1]。近年来, 一些地区涌现出一批蚕桑新村和种养大户, 这些新兴力量为当地经济发展注入了新的活力, 也为助农增收提供了重要途径, 这在一定程度上改善了农民的生活水平, 展现了蚕桑产业对于农村经济的积极贡献。广西蚕桑生产区域覆盖71个县, 蚕桑从业人员约370万人, 蚕户年收入约1.84万元^[2]。然而, 蚕桑产业的发展也面临着一些挑战。原因之一是有部分养蚕人员缺乏经验, 不善于应用新的技术和方法, 这种情况可能导致养蚕过程中出现诸如蚕病等疾病的暴发, 严重影响产业的稳定和发展。因此, 蚕病的监控尤为关键。蚕病是蚕桑产业中常见的病害, 如白

僵病、脓病等, 如果不及时发现和控制, 将对产量和质量造成严重损失。特别是在养蚕人员缺乏经验的情况下, 对蚕病的监测和防治显得尤为重要。通过增强蚕病检测的效果, 及时发现病害的迹象, 并及时采取有效的防治措施, 可以最大限度地降低蚕病对产业的影响, 保障农民的收入。

近年来, 深度学习的迅速发展推动了基于计算机视觉的目标检测技术在农业工程领域的广泛应用, 尤其是对农作物病虫害的检测^[3-5]。目前已经有很多研究人员将深度学习算法应用到农作物病虫害检测中。王铭慧等^[6]针对番茄病虫害目标小而密的特点, 提出了基于Swin Transformer的YOLOX目标检测网络, 以精确定位图像中的病虫害目标。同时, 采用了基于经典卷积神经网络构建的旋转不变Fisher判别CNN分类网络, 使在原模型基础上病虫害分类的准确率提升6.1%。白雪松等^[7]提出一种基于Res2Net50改进模型的农作物病虫害检测算法, 优化了对病虫害症状相似导致类间差异小、难以区分的问题。罗云响^[8]为了解决样本不平衡和部分蚕病特征相似的问题, 在YOLOv5的主干

收稿日期: 2024-04-07 修稿日期: 2024-05-30

作者简介: 曾沛杰(1999—), 男, 广东清远人, 在读硕士研究生, 研究方向为计算机视觉; *通信作者: 周卫(1980—), 男, 广西钦州人, 博士, 副教授, 从事领域为大型软件系统、计算机视觉、自然语言处理等, E-mail: zhouwei@gxmzu.edu.cn; 陈金良(2001—), 男, 安徽六安人, 硕士研究生, 研究方向为行人重识别

部分的C3结构后面增加了SPPCSPC模块,并在检测头部分引入了Shuffle Attention注意力机制。使用了经过图像背景分割处理的病蚕图像混合原始数据集来训练改进的YOLOv5模型,与原始模型实验相比其mAP指标共提升了3.3%。丁继尧^[9]提出了一种基于深度哈希的蚕病图像检索算法,通过融合层和切片层分别增强了模型对蚕病图像局部细节信息的获取和表达能力。采用了配对损失来提高模型对不同类别蚕病图像的识别性能。改进的AlexNet-MSP模型的平均识别率达到了85.8%,对比算法至少提高了6.6%。

1 材料与方法

1.1 数据集构建

实验数据集包含1458张拍摄的蚕照片,包括健康、白僵病、脓病三种状态,数据集拍摄地点位于广西壮族自治区,在不同光照背景、多个角度下拍摄。为了更加贴合自然复杂环境,数据集中有一部分数据是采用专业相机进行拍摄,能清晰捕捉到多角度下的蚕体和清晰的背景。此外,为了更符合农户的需求和使用习惯,另一部分数据集是来自不同农户使用不同设备拍摄的。数据集中包含了两种比较常见的蚕病,白僵病通常由真菌感染引起,感染后蚕体皮肤表面会逐渐浮现暗褐色病斑,蚕病死后蚕身被白色分生孢子包裹,并处于僵硬状态。蚕患脓病后,蚕体表面呈现出油浸斑,体节肿大,呈现出竹节状。具体特征如图1所示。数据按7:3划分为训练集与验证集。数据集分布见表1。



健康 白僵病 脓病

图1 蚕的三种状态图像

表1 蚕病健康状况图像分布 单位:张

	健康	白僵病	脓病
专业相机拍摄	486	390	243
农户拍摄	0	96	243
总数	486	486	486

1.2 YOLOv8模型结构

YOLOv8是YOLO(You Only Look Once)系列目标检测算法的第八个版本。YOLO是一种OneStage快速的目标检测算法,由backbone、neck与head三部分组成,根据不同尺度提供了N/S/M/L/X五种尺寸,对应不同参数数量的变体,以满足不同场景需求,YOLOv8相比于之前的YOLO系列版本,在检测速度与精度上都有提升,能够更快更准确地检测出目标,因此使用YOLOv8作为基础网络进行进一步优化,提高蚕病检测的准确率。

1.3 YOLOv8模型结构改进

在YOLOv8中引入了可变形卷积DCNv2(Deformable ConvNets v2)^[10],网络结构如图2所示。通过学习偏移和加权,提高了模型对于变形目标特征的提取能力。在蚕病识别任务中,由于蚕的形态和外观可能会因为病害在不同的发病阶段发生变化,如脓病的病症,在四、五龄盛食期前后发病的蚕环节间膜高起、形状竹节,五龄盛食期到催熟期发病,全身肿胀、体色乳白。引入DCNv2结构能够增强模型捕获不同阶段蚕病的特征的能力。

1.4 DCNv2

可变形卷积DCNv2是一种在Deformable Convolutional Networks^[11]改进的卷积操作,为了解决可变形卷积的感受野区域大于感兴趣区域,从而会对目标区域进行错误识别的问题,DCNv2通过引入调制机制来调整卷积核的形状使感兴趣的区域更加集中,以准确地提取输入特征。如输入特征图为 $W \times H \times C$,首先是偏移量生成:通过卷积操作生成卷积核在输入特征图上沿着 x 和 y 方向的采样点偏移量,在这个过程中生成的Offsets field是通过卷积生成的,其中 W 、 H 是不变的,由于沿着 x 、 y 方向进行采样,所以特征图通道数 $2C$ 。接着通过双线性插值法计算出Offsets field偏移位置的特征值,在这个过程中梯度是通过双线性查找进行反向传播的。最后,利用这些采样点进行卷积操作。DCNv2结构如图3所示。

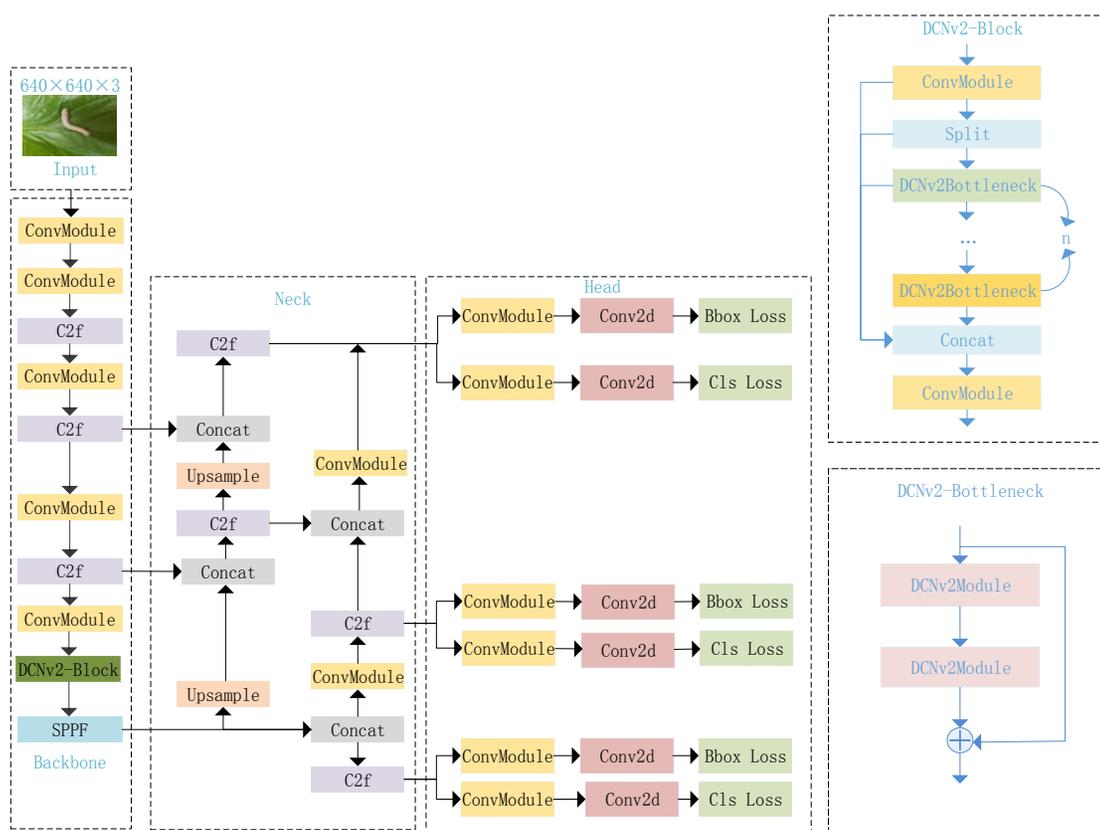


图2 YOLOv8-DCNv2网络结构图

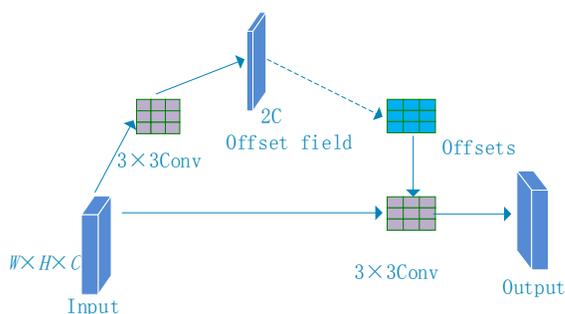


图3 DCNv2网络结构

2 模型预训练与测试

2.1 实验环境

本实验所用操作系统为64位 Ubuntu 18.04.6 LTS, CPU 型号为 Intel® Xeon® Gold 6136 CPU @ 3.00GHz, GPU 为 16G NVIDIA Tesla T4, 运行环境为 Torch1.11.0+cu113 和 Python 3.9.17 以及 CUDA 11.4。提出的研究是基于 YOLOv8 作为基准模型, 并在模型的基础上进行改进。

2.2 模型训练

模型训练时, 将所有训练集图像的分辨率统一调整为 640*640 像素。优化器使用 SGD 优化器。将批量训练(batch-size)设置为 8, 训练时初始学习率为 0.01, 权值的衰减速率(weight_decay)设为 0.0005, 动量(momentum)设为 0.937。模型的总 epoch 数设为 100, 并且设置了运行超过 50epochs 没有提升就提前终止模型训练, 自动保存模型识别精度最高的权重。将 Mosaic 设置为 1, 开启数据增强和尺度变换, 丰富图片资源, 增强病害识别能力。

2.3 评价指标

本文评价模型性能的指标采用均值平均精度 (Mean Average Precision, mAP)。均值平均精度与模型的精确率 (Precision, P) 和召回率 (Recall, R) 有关。

精确率 (Precision) 表示预测为正例的样本中, 真正的正例所占的比例, 计算公式如式(1)所示:

$$P = \frac{TP}{TP + FP} \times 100\% \quad (1)$$

召回率(Recall)表示真正的正例中, 被预测为正例的样本所占的比例, 计算公式如式(2)所示:

$$R = \frac{TP}{TP + FN} \times 100\% \quad (2)$$

AP对不同召回率点上的精度进行平均(即PR曲线下的面积), 针对一个类别而言的, 计算公式如式(3)所示:

$$AP = \int_0^1 P(R)dR \times 100\% \quad (3)$$

mAP是精确率的平均, 把所有类别的AP都计算出来后, 再对它们求平均值, 即可得到mAP, 计算公式如式(4)所示:

$$AP = \frac{1}{C} \sum_{i=1}^C AP_i \times 100\% \quad (4)$$

上述公式中, TP(True positives)为正确检测为某一类蚕病的数量; FP(False positives)表示被错误划分为某一类蚕病的数量; FN(False negatives)图像中被漏检的蚕病目标的数量。

3 结果与分析

3.1 效果对比

实验对比效果见表2。可以看到, 使用提出的DCNv2-Block的改进方法, 在蚕病识别中可以提供更好的特征学习和适应性。由于蚕病在不同的发病阶段的病态具有多样化的形态和尺度, 导致模型很容易受到背景信息的干扰。通过学习卷积核的形变增强了模型在感受野内自适应地捕捉到目标的特征的能力, 可以更好地捕捉

到目标周围的局部特征和区分蚕病病征与背景之间的微小差异, 从而能够更准确地捕捉到发生病变的蚕。

表2 添加DCNv2-Block前后效果对比

Model	脓病	白僵病
YOLOv8		
YOLOv-DCNv2		

3.2 消融实验

经过多次实验得出使用本文提出的DCNv2-Block替换YOLOv8基线模型Backbone其中一个c2f模块的效果最好, 通过消融实验验证了所提改进方法的可行性, 实验结果见表3。由结果可以看出, 将DCNv2-Block替换YOLOv8模型Backbone中的第四个c2f模块得到的效果最好, 虽然召回率相比于原模型下降了0.3个百分点, 但是模型的精确度提升了2.9个百分点, mAP_0.5提升了1.1个百分点, mAP_0.5 : 0.95提升了0.7个百分点。

表3 YOLOv8-DCNv2消融实验(%)

使用DCNv2-Block取代backbone中第一个c2f模块	使用DCNv2-Block取代backbone中第二个c2f模块	使用DCNv2-Block取代backbone中第三个c2f模块	使用DCNv2-Block取代backbone中第四个c2f模块	精确度 Precision P	召回率 Recall R	mAP_0.5	mAP_0.5 : 0.95
—	—	—	—	92.6	91.6	95.4	85.9
√	—	—	—	92.6	90	95	84.5
—	√	—	—	90.8	92.8	95.9	85.7
—	—	√	—	92.2	91.9	95.6	85.6
—	—	—	√	95.1	91.3	96.5	86.6

注: “—”表示没有用此模块; “√”表示使用此模块。

3.3 对比实验

为验证本研究提出模型的有效性，选择YOLOv5、YOLOv8、YOLOv9及本文提出的YOLOv8-DCNv2共四种网络模型，在相同的训练环境下进行对比实验。对比结果见表4。从结果可以看出，对于蚕病检测的效果，提出的YOLOv8-DCNv2模型比YOLOv5、YOLOv9和YOLOv8的平均精确率分别提高1.6个百分点、3.2个百分点和1.1个百分点。

表4 各模型对比结果(%)

模型 Model	精确度 Precision P	召回率 Recall R	mAP_0.5	mAP50:95
YOLOv5	93.6	92.9	94.9	80.5
YOLOv8	92.6	91.6	95.4	85.9
YOLOv9	88.3	87	92.3	82.1
YOLOv8-DCNv2	95.1	91.3	96.5	86.6

4 结语

在研究中，成功地将可变形卷积应用于改进YOLOv8模型，提高了蚕病识别的准确性和鲁棒性。通过引入可变形卷积，有效增强了模型对于变形目标特征的提取能力。实验结果表明，所提方法的准确性得到提升，并且降低了运算次数，为蚕病监测和预防提供了一种新的有效途径。未来的工作将继续探索更深层次的模型优化和数据增强技术，以进一步提升蚕病识别的性能，有助于农户及时采取措施进行治疗和预防，从而提高蚕宝宝的健康状况和产量。

参考文献:

- [1] 黄深惠,蒋满贵,黄旭华,等. 广西近几年蚕病的发生与防控[J]. 广西蚕业,2021,58(4):53-58.
- [2] 吕南,张小勤,凌亚山. 广西蚕桑产业存在的风险以及风险管理研究[J]. 中国蚕业,2023,44(4):39-44
- [3] 牛潘婷,张宝林,潘丽杰,等. 基于深度学习的农作物病虫害研究进展[J]. 内蒙古师范大学学报(自然科学汉文版),2024,53(1):93-102.
- [4] 苏令涛,李瑞泽,张功磊,等. 基于深度学习的农作物病虫害识别研究[J]. 数学建模及其应用,2022,11(4):1-12.
- [5] 王明,张倩. 我国基于深度学习的图像识别技术在农作物病虫害识别中的研究进展[J]. 中国蔬菜,2023(3):22-28.
- [6] 王铭慧,张怀清,樊江川,等. 基于深度学习网络实现番茄病虫害检测与识别[J]. 中国农业大学学报,2023,28(11):165-181.
- [7] 白雪松,吴建平,景文超,等. 基于改进残差网络的农作物病虫害检测研究[J]. 计算机技术与发展,2023,33(5):145-151.
- [8] 罗云昀. 基于深度学习的桑蚕病虫害检测技术研究[D]. 南宁:广西民族大学,2023.
- [9] 丁继尧. 基于远程监控的蚕病识别系统研究与实现[D]. 武汉:武汉理工大学,2021.
- [10] ZHU X, HU H, LIN S, et al. Deformable convnets v2: more deformable, better results[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Los Angeles CA, United States, 2019:9308-9316.
- [11] DAI J, QI H, XIONG Y, et al. Deformable convolutional networks[C]//Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 2017:764-773.

Design of silkworm disease detection algorithm based on deep learning

Zeng Peijie, Zhou Wei*, Chen Jinliang

(College of Artificial Intelligence, Guangxi University for Nationalities, Nanning 530006, China)

Abstract: In the context of silkworm disease recognition, the YOLOv8 model is susceptible to interference from background information due to changes in the morphology and appearance of silkworm disease at different stages. To address this issue, a new feature extraction module called DCNv2-Block was proposed to enhance the model's ability to capture characteristics of silkworm disease at different stages. Experimental results demonstrated that compared with the original YOLOv8 model, the proposed improved method achieved the mAP0.5 of 96.5%, representing 1.1 percentage point increase. This enhancement enables better capture of symptoms of silkworm diseases, improves accuracy in identifying silkworm diseases, and facilitates timely detection and control measures.

Keywords: YOLOv8; detection of silkworm disease; variable convolution; deep learning

文章编号: 1007-1423(2024)20-0036-05

DOI: 10.3969/j.issn.1007-1423.2024.20.007

基于图神经网络的药物相互作用预测方法研究综述

郭文博*, 龙伟, 蒋林华, 叶文标, 胡灵犀, 彭司华

(湖州师范学院信息工程学院数智农业研究院, 湖州 313000)

摘要: 对近年来基于图神经网络的 DDI 预测方法的研究进展进行了总结, 包括 DDI 预测的定义和分类、基于图神经网络模型的 DDI 预测方法的步骤、常用的数据集和评价指标以及未来的研究方向和应用前景等。对当前存在的问题, 如模型可解释性进行了探, 并对未来的研究方向例如多模态信息融合等进行了展望。

关键词: 药物相互作用; 图神经网络; 深度学习; 药物治疗

0 引言

药物相互作用(drug-drug interaction, DDI)是指当一个药物与另一个药物在体内接触后, 产生的一种影响或效应。药物相互作用对于药物治疗效果和安全性都具有重要影响, 因此药物相互作用的研究和预测也就成为了药物研究领域中的一个重要研究方向^[1]。

药物相互作用是药物治疗效果和安全性的重要因素, 药物的疗效和安全性都与其在体内的代谢和药动学有关, 而药物相互作用会影响药物的代谢和药动学过程, 从而导致不良反应的发生、药物疗效的增强或减弱等情况^[2]。因此, 药物相互作用的研究和预测可以为药物的安全性评价、剂量调整、疗效优化等提供重要依据。其次, 传统的药物相互作用预测方法存在一些问题, 其通常基于分子结构和生物活性等特征进行预测, 但这些特征往往无法准确反映药物相互作用的复杂性和多样性, 导致预测结果的不准确和误差较大。因此, 寻求一种更准确和有效的药物相互作用预测方法显得非常必要^[3]。随着人工智能和机器学习等技术的

发展, 基于机器学习的药物相互作用预测方法正在逐渐成为研究的热点^[4]。特别是近年来, 图神经网络等深度学习技术的应用^[5], 为药物相互作用预测提供了新的思路和方法, 其预测准确性和稳定性也得到了大幅提升^[6-7]。

药物相互作用预测研究的主要目的是提高临床用药的安全性和有效性, 优化药物的剂量和用药方案, 减少药物不良反应和药物误用的发生, 为药物研究和开发提供重要参考。药物相互作用可能会导致药效变化或不良反应, 因此了解药物之间的相互作用可以指导医生在临床用药时的决策, 从而避免不必要的风险和损害^[8]。药物相互作用预测也可以在药物研发初期发现药物间的相互作用, 避免研发失败和浪费资源, 同时也可以发现新的药物组合, 提高新药的研发效率和成功率。药物组合疗法在临床治疗中有着广泛的应用, 药物相互作用预测可以帮助医生和研究人员优化药物组合, 提高治疗效果和安全性。药物相互作用可能会导致治疗失败或不良反应, 从而增加医疗成本, 药物相互作用预测可以避免这种情况的发生, 节省医疗成本。

收稿日期: 2024-05-16 修稿日期: 2024-07-24

作者简介: *通信作者: 郭文博(2000—), 男, 河南商丘人, 在读硕士研究生, 研究方向为生物医学, E-mail: gw214308@163.com; 龙伟(1978—), 男, 四川成都人, 博士, 副教授, 研究方向为图像处理、数字化农业信息处理; 蒋林华(1977—), 男, 湖南永州人, 博士, 教授, 博士生导师, 研究方向为生物医学、电子光学; 叶文标(2000—), 男, 安徽阜阳人, 在读硕士研究生, 研究方向为生物医学、深度学习; 胡灵犀(1991—), 女, 四川广安人, 博士, 研究方向为光子准粒子和结构物质的相互作用、图像处理; 彭司华(1962—), 男, 湖南省湘西人, 博士, 副研究员, 主要从事生物信息学、基因组学研究

1 传统的药物相互作用预测方法

药物相互作用预测是药物研发过程中的重要环节，能够帮助科学家了解药物的安全性、疗效和药理作用。传统的药物相互作用预测方法主要基于实验数据、统计学和机器学习等技术，通过分析药物化学特征、分子结构和药理信息等来推断药物之间的相互作用^[9]。本节将概述传统的药物相互作用预测方法，包括药物结构相似性、药物化学信息以及基于统计学和机器学习等方面的方法。

1.1 基于药物相似性的方法

药物结构相似性方法是一种常见的传统药物相互作用预测方法^[10]。该方法基于假设：具有相似结构的药物往往具有相似的药理作用和相互作用。药物结构相似性可以通过计算分子之间的相似性指标来衡量，例如分子指纹、分子描述符和化学键等^[11]。然后，通过比较药物之间的结构相似性，推测它们是否具有相似的相互作用。药物结构相似性方法的优点是简单易用，但也存在局限性，例如无法捕捉复杂的结构-活性关系和药物多样性。

1.2 基于统计学的方法

基于统计学的方法主要是通过分析药物化学结构和属性之间的关系来预测药物相互作用^[8]。例如QSAR^[12] (Quantitative Structure-Activity Relationship, 定量构效关系)方法：通过建立药物的化学结构和活性之间的数学关系来预测药物相互作用。该方法需要大量的实验数据作为支撑，并且容易受到化学结构描述符的选择和建模方法的影响。还有一种方法是通过分析药物相互作用的生物学机制和分子交互作用来预测药物相互作用。但是这种方法需要对药物和其作用机制有较为深入的了解，因此在实际应用中存在一定的局限性。

1.3 基于机器学习的方法

基于机器学习的方法主要是通过训练模型来预测药物相互作用。基于特征的方法是将药物的分子结构和属性转换为数值型特征，然后使用传统的机器学习算法来训练模型。这种方法需要精心选择特征，且容易受到特征描述符

的选择和编码方式的影响。例如Naive Bayes(朴素贝叶斯)：是一种基于贝叶斯定理的分类算法，常用于文本分类^[13]。在DDI预测中，可以通过朴素贝叶斯算法将药物特征转化为概率分布，从而预测药物相互作用。Random Forest(随机森林)：是一种决策树集成方法，通过对多个决策树的结果进行平均或投票来进行分类或回归^[14]。在DDI预测中，可以通过随机森林算法综合考虑药物分子结构、药物代谢途径等多个因素，从而预测药物相互作用^[15]。Support Vector Machine(支持向量机)：是一种常用的分类算法，通过将数据映射到高维空间中，找到能够区分不同类别的超平面来进行分类^[16]。在DDI预测中，可以通过支持向量机算法对药物分子结构、药物代谢途径等因素进行分类，从而预测药物相互作用。

2 基于图神经网络的药物相互作用预测方法

2.1 图神经网络(GNN)

传统的药物相互作用预测方法通常只考虑药物分子的局部结构和生物活性参数，而基于图神经网络的药物相互作用预测方法可以充分考虑药物分子的全局结构，从而更准确地预测药物相互作用。图神经网络通过对药物分子的图结构进行学习，能够充分挖掘药物分子之间的相互作用关系，从而在药物相互作用预测方面具有更高的准确性^[17]。

基于图神经网络的药物相互作用预测方法主要涉及以下两个方面：①药物表示学习，药物可以被表示成一个图结构，其中节点表示原子或功能基团，边表示它们之间的化学键或空间距离等关系。图神经网络通过将药物的分子结构转换成图的形式，然后通过神经网络对图结构进行学习，将药物的分子结构转化为低维空间的向量表示，从而实现药物结构的表征和学习；②药物相互作用预测，在获得药物的低维向量表示后，图神经网络通过将不同药物的向量表示进行组合，预测不同药物之间的相互作用概率。这个过程通常涉及到相似性度量、分类和回归等多种模型和方法。

2.2 图神经网络模型的构建

预测药物相互作用的图神经网络模型通常包括以下几个步骤：数据集的构建、图表示的构建、图特征的提取、使用深度学习模型、模型的训练与优化以及模型的测试与评估。

数据集的构建是药物相互作用预测的关键步骤之一。它涉及收集药物分子结构及其相互作用的相关信息，如分子结构、蛋白质靶点、药效等。通常，药物相互作用预测需要包含大量的药物结构信息和相应的相互作用标签。在构建数据集时，需要考虑以下几个方面数据来源，药物相互作用预测的数据可以来自于已知药物相互作用数据库，如 DrugBank、STITCH、BindingDB 等。数据清洗在药物相互作用预测中，需要清洗掉无用的数据，如重复数据、缺失数据等。数据划分，数据集通常需要划分为训练集、验证集和测试集三部分，其中训练集用于模型的训练，验证集用于模型的性能调优和选择，测试集用于模型的性能评估。

图表示的构建，将药物分子结构、蛋白质靶点等转换为图形结构表示，每个节点表示分子中的原子或靶点，边表示它们之间的关系^[18]。图特征的提取，从构建的图表示中提取特征，可以使用传统的图特征提取方法，如子图、路径等，也可以使用基于深度学习的图嵌入方法。

使用深度学习模型，如图注意力网络 (graph attention network, GAN)、卷积神经网络 (convolutional neural network, CNN)、循环神经网络 (recurrent neural network, RNN) 等构建药物相互作用预测模型^[19]。模型的训练与优化，使用训练集对模型进行训练，并使用验证集对模型进行优化，调整模型的超参数，使其达到最佳的性能。模型的测试与评估，使用测试集对模型进行测试，评估模型的预测性能，常用的评估指标包括准确率、召回率、F1 值等^[20-21]。

2.3 基于图神经网络的方法

图神经网络是一种基于图结构数据的深度学习模型，可以有效地处理非欧几里得结构数据，包括分子结构和化合物网络等复杂的结构数据。近年来，图神经网络在药物相互作用预测中得到了广泛的应用，并取得了一系列优秀

的预测效果。下面对图神经网络在药物相互作用预测中的一些方法进行详细介绍。

DDI-GCN^[22]设计了一种基于图卷积神经网络的药物相互作用预测方法，能够对多个药物之间的相互作用进行预测。DDI-GCN 使用图卷积神经网络模型来对药物分子进行编码，并使用注意力机制来对编码进行交互和融合，从而预测不同药物之间的相互作用。DDI-GNN^[23]是一种基于图神经网络的药物-药物相互作用预测方法，能够考虑药物相似性、药物分类和相互作用网络结构等因素。DDI-GAT^[24]提出了一种基于图注意力网络的药物-药物相互作用预测方法，能够考虑药物相似性、药物分类和相互作用网络结构等因素。DDI-GAT 使用图注意力网络模型来对药物相互作用网络进行编码，并使用多层感知机来进行分类，从而预测它们之间的相互作用。此方法还使用一种新的药物表示方法，将药物的结构和药物相似性信息结合起来。MPNN-DDI^[25]使用分子图卷积神经网络模型来编码药物分子的结构信息，并使用多层感知机来进行分类，从而预测药物之间的相互作用。此方法还引入了一种新的药物表示方法，能够考虑药物分子的局部和全局结构信息。TDM-GCN^[26]使用了图卷积网络和时间动态模型来学习药物和疾病的特征表示，该方法还采用了一种新的动态图生成方法，能够更好地捕捉药物-疾病相互作用的时空动态信息，并提高预测性能。GAT-VSE^[27]实现了一种基于图注意力网络和视觉语义嵌入的药物-药物相互作用预测方法，它使用了图注意力网络和视觉语义嵌入模型来学习药物的特征表示。该方法还采用了一种新的特征重构方法和特征对齐方法，能够更好地捕捉药物之间的相互作用信息，并提高预测性能。

3 未来挑战与展望

由于模型的复杂性，很难理解模型如何进行决策，从而难以解释预测结果的合理性。基于图神经网络的药物相互作用预测方法需要进一步解释其预测结果的可解释性^[28]。未来，基于图神经网络的药物相互作用预测方法可能会从以下几个方面进行发展：

多模态信息融合：除了药物分子的结构信息外，药物相互作用预测还可以利用其他多模态信息，例如药物基因组学、药物代谢组学、药物蛋白质组学等信息。因此需要研究如何将多种信息进行融合以提高药物相互作用预测的准确性。

模型可解释性的提高：研究更加可解释的图神经网络模型，以提高模型的可解释性，设计更加可解释和可视化的模型。

基于图神经网络的药物相互作用预测方法在未来有着广阔的应用前景。目前的基于图神经网络的药物相互作用预测方法往往只考虑药物本身的结构信息^[29]，而忽略了其他重要的信息，例如药物的剂量、使用方式、患者基因型等。这些信息可能会对药物相互作用产生重要影响，因此将这些信息纳入预测模型中有望进一步提高预测准确性。

4 结语

DDI的发生严重影响了患者的治疗，并成为患者安全和药物管理的重大问题。如果深度学习能够有效地预测DDI，那么DDI造成的危害将被大大降低，因此迫切需要开发性能更好的深度学习方法。本文总结了现有的基于机器学习和图神经网络的DDI预测方法，并且介绍了图神经网络方法模型的构建流程。最后讨论了现有方法的挑战，并提出了在未来有趣的研究方向，在DDI预测方向上还存在巨大的发展前景。

参考文献：

- [1] 罗熹,曾智颖,王建新,等. 结合语义和依存关系的药物相互作用关系抽取[J]. 湖南大学学报(自然科学版),2022,49(6):90-100.
- [2] 薛斌,李益洲,李梦龙. 基于化学信息学方法预测药物副作用的研究进展[J]. 计算机与应用化学,2019,36(5):487-490.
- [3] 刘璐,陈笑艳. 国产小分子化学创新药临床药物相互作用研究现状[J]. 中国临床药理学与治疗学,2021,26(8):863-875.
- [4] 邵振杰. 基于深度学习方法的药物与药物的相互作用类别预测[D]. 上海:华东师范大学,2022.
- [5] 饶晓洁,张通,孟献兵,等. 基于多层注意力和消息传递网络的药物相互作用预测方法[J]. 自动化学报,2023,49(12):2507-2519.
- [6] 叶慧敏. 基于图神经网络的药物与药物相互作用预测方法研究[D]. 南京:南京邮电大学,2022.
- [7] 任忠豪. 基于图结构表示学习的药物与药物相互作用预测及应用研究[D]. 西安:西京学院,2022.
- [8] 黄紫阳. 基于生物医学知识图谱和药物分子结构的药物-药物相互作用预测[D]. 武汉:华中农业大学,2022.
- [9] 郝鑫坤. 基于知识图谱的药物相互作用预测方法研究[D]. 南宁:广西大学,2022.
- [10] HAN K, CAO P, WANG Y, et al. A review of approaches for predicting drug-drug interactions based on machine learning [J]. *Frontiers in Pharmacology*, 2022, 12:3966.
- [11] 张竣,袁锐,陈世龙,等. 基于核方法的协同药物组合预测方法[J]. 中国科学:生命科学,2023,53(11):1663-1672.
- [12] 欧阳佳明. 基于多模态表示学习的药物相互作用预测研究[D]. 上海:上海大学,2022.
- [13] 汪颖,李享云. 蛋白质-RNA相互作用预测中的几类分类器[J]. 渤海大学学报(自然科学版),2015,36(4):305-309,326.
- [14] 王振娟,韩乐,李雪婵,等. 机器学习在药物筛选中的研究进展[J]. 中国医院药学杂志,2023,43(22):2584-2588.
- [15] 代绍兴. 利用结构生物信息学方法促进药物研发[D]. 合肥:中国科学技术大学,2013.
- [16] 杨岩,肖佳妹,周晋,等. 支持向量机法及其在中药研究中的应用[J]. 中草药,2020,51(8):2258-2266.
- [17] WANG Y L, WANG F, SHI X X, et al. Cloud 3D-QSAR: a web tool for the development of quantitative structure-activity relationship models in drug discovery [J]. *Briefings in Bioinformatics*, 2021, 22(4):bbaa276.
- [18] SU X, HU L, YOU Z, et al. Attention-based knowledge graph representation learning for predicting drug-drug interactions [J]. *Briefings in bioinformatics*, 2022, 23(3):bbac140.
- [19] ZHANG C, LU Y, ZANG T. CNN-DDI: a learning-based method for predicting drug-drug interactions using convolution neural networks [J]. *BMC Bioinformatics*, 2022, 23(1):1-12.
- [20] VO T H, NGUYEN N T K, KHA Q H, et al. On the road to explainable AI in drug-drug interactions prediction: a systematic review [J]. *Computational and Structural Biotechnology Journal*, 2022, 20:2112-2123.

- [21] HAN K, CAO P, WANG Y, et al. A review of approaches for predicting drug-drug interactions based on machine learning[J]. *Frontiers in Pharmacology*, 2022, 12: 3966.
- [22] ZHONG Y, ZHENG H, CHEN X, et al. DDI-GCN: drug-drug interaction prediction via explainable graph convolutional networks[J]. *Artificial Intelligence in Medicine*, 2023, 144: 102640.
- [23] AL_RABEAH M H, LAKIZADEH A. GNN-DDI: a new data integration framework for predicting drug-drug interaction events based on graph neural networks[J]. *Scientific Reports*, 2022, 22: 55-69.
- [24] TANVIR F, SAIFUDDIN K M, AKBAS E. DDI prediction via heterogeneous graph attention networks[EB/OL]. arXiv:2207.05672, 2022.
- [25] CHE M, YAO K, CHE C, et al. Knowledge-graph-based drug repositioning against COVID-19 by graph convolutional network with attention mechanism[J]. *Future Internet*, 2021, 13(1): 13.
- [26] KARLOV D S, SOSNIN S, FEDOROV M V, et al. GraphDelta: MPNN scoring function for the affinity prediction of protein-ligand complexes [J]. *ACS Omega*, 2020, 5(10): 5150-5159.
- [27] TUZIMSKI T, PETRUCZYNIK A. Review of chromatographic methods coupled with modern detection techniques applied in the therapeutic drugs monitoring (TDM) [J]. *Molecules*, 2020, 25 (17) : 4026.
- [28] QIU Y, ZHANG Y, DENG Y, et al. A comprehensive review of computational methods for drug-drug interaction detection [J]. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2021, 19(4): 1968-1985.
- [29] WANG N N, ZHU B, LI X L, et al. Comprehensive review of drug-drug interaction prediction based on machine learning: current status, challenges, and opportunities [J]. *Journal of Chemical Information and Modeling*, 2023, 64(1): 96-109.

A review of prediction methods of drug interaction based on graph neural network

Guo Wenbo*, Long Wei, Jiang Linhua, Ye Wenbiao, Hu Lingxi, Peng Sihua

(Digital and Intelligent Agriculture Research Institute, School of Information Engineering, Huzhou University, Huzhou 313000, China)

Abstract: This paper summarizes the research progress of DDI prediction methods based on graph neural network in recent years, including the definition and classification of DDI prediction, the steps of DDI prediction method based on graph neural network model, commonly used data sets and evaluation indicators, as well as the future research direction and application prospect. The existing problems, such as model interpretability, are discussed, and the future research directions such as multi-modal information fusion are prospected.

Keywords: drug interaction; graph neural network; deep learning; drug therapy

文章编号: 1007-1423(2024)20-0041-04

DOI: 10.3969/j.issn.1007-1423.2024.20.008

基于机器视觉的圆形精密工件尺寸检测

于佳昊, 刘嘉承, 王连锴, 杨展, 宋德*

(长春理工大学物理学院, 长春 130013)

摘要: 针对圆形精密工件尺寸测量困难、检测成本较高等问题提出一种基于区域灰度模型的快速、精确的检测算法。对基于机器视觉的圆形精密尺寸检测方法进行了研究。首先, 通过形态学处理完成图像的预处理; 然后, 在区域灰度模型算法中引入OTSU自适应阈值, 提高检测精度; 之后, 通过Canny算子定义的圆环完成缺陷处理; 最后, 采用最小二乘法拟合圆形直径长度。该方法运算速度快、检测精度高, 最大测量误差小于0.002 mm, 平均测量误差精度小于0.0007 mm。满足工件检测精度要求。

关键词: 精密检测; 机器视觉; 区域灰度模型

0 引言

对于工件尺寸检测, 边缘检测的精确度是重要的评判标准。边缘检测的准确性可以直接影响轮廓提取。近些年, 亚像素边缘检测技术得到飞速发展, 逐渐成为提高检测精度的主流做法^[1]。

传统的亚像素边缘检测算法分为矩法、拟合法和插值法三类。Kisworo等^[2]提出边缘模型和局部能量函数相结合方法。该方法为一维和二维边缘类型的图像提供了良好结果, 但使用存在一定局限性。Bin等^[3]根据低阶径向数和矩旋转不变性的特征, 提出了基于正交傅里叶-梅林矩法的检测方式, 检测精度较低。Pap等^[4]提出一种结合高阶ENO插值法和差分高斯拉普拉斯来准确确定边缘点的方法。刘国栋等^[5]提出基于灰度面积插值的亚像素定位方法, 检测便捷但定位精度稍有欠缺。张宗华等^[6]提出高斯拟合和随机采样一致性算法相结合的方式, 有效减少了图像拍摄角度对精度的影响。于祥等^[7]提出一种

使用新型的logistic边缘检测模型和阶跃阈值自适应提取方法的Zernike矩亚像素边缘检测算法, 有效提高边缘识别精度。Trujillo-Pino等^[8]通过图像的局部灰度特征进行边缘检测, 由局部灰度特征推导出边缘采集模型。还有学者^[9-10]基于此方法进行改进, 但仍存在检测精度不够等问题。

针对圆型工件尺寸检测精度较低的问题, 本文首先通过形态学处理有效去除工件图像中存在的非边缘点, 减少其对尺寸检测的影响; 其次, 改进区域灰度模型, 在传统区域灰度模型的基础上引入OTSU自适应阈值, 较大地提高了亚像素边缘检测的精度和适用性; 之后, 基于Canny算子确定圆环型亚像素边缘点坐标的检测范围完成缺陷处理, 有效解决工件表面划痕影响检测的问题; 最后, 采用最小二乘法拟合圆形直径尺寸, 更加精确地表达了圆形图像信息, 便于检测工件尺寸。实际检测表明, 本文算法同时满足快速、精确的图形检测需求。

收稿日期: 2024-04-17 修稿日期: 2024-05-24

基金项目: 国家自然科学基金资助项目(U2031113)

作者简介: 于佳昊(1998—), 男, 黑龙江齐齐哈尔人, 在读硕士研究生, 研究方向为图像处理; 刘嘉承(1999—), 男, 江西赣州人, 在读硕士研究生, 研究方向为图像处理; 王连锴(1982—), 男, 吉林通化人, 博士, 副教授, 研究方向为半导体光电子器件; 杨展(2001—), 男, 山东菏泽人, 在读本科, 研究方向为图像处理; *通信作者: 宋德(1981—), 男, 吉林白山人, 博士, 副教授, 从事领域为微光成像与系统、图像处理等, E-mail: songde614@cust.edu.cn

1 尺寸检测算法流程

尺寸检测算法流程如图1所示。

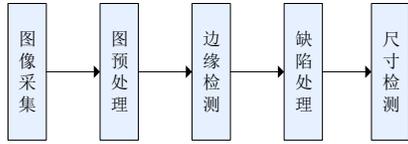


图1 算法框图

1.1 图像预处理

图像在拍摄过程中受环境等因素的影响，图片中工件表面极易沾染污渍，极大影响检测精度。因此，本文形态学处理中开运算，高效便捷地去除杂质。如图2所示，图2(a)和图2(b)分别为原图和形态学处理后的图像。原图中工件表面存在的少许杂质信息，经过形态学处理后被有效去除。

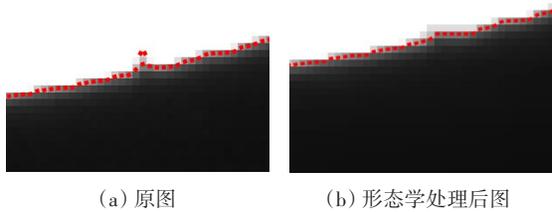


图2 形态学处理对比图

1.2 边缘检测

区域灰度模型是一种非连续的图像采集模型，以已采集到的边缘为中心建立像素窗口，通过区域面积及像素总强度得出边缘参数。以斜率小于1的曲线方程 $y = a + bx + cx^2$ 为例，边缘检测采用 5×3 的像素结构，因此每一列像素值均为5个像素值的和，如图3所示。

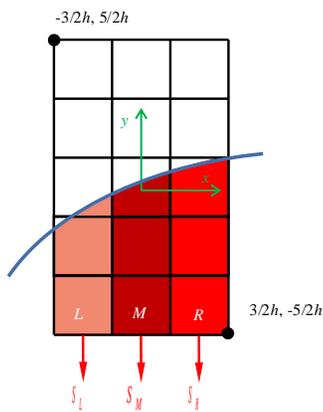


图3 曲线 5×3 结构图

每列像素值和如式(1)~(3)所示：

$$S_L = 5B + \frac{A - B}{h^2} L \quad (1)$$

$$S_M = 5B + \frac{A - B}{h^2} M \quad (2)$$

$$S_R = 5B + \frac{A - B}{h^2} R \quad (3)$$

式中： L 、 M 、 R 表示每一列中边缘线以下的面积； A 和 B 表示边缘上下两侧的强度值； h 表示像素宽度。最终求得曲线方程中 a 、 b 和 c 的表达式如式(4)~(6)所示：

$$c = \frac{S_L + S_R - 2S_M}{2(A - B)} \quad (4)$$

$$b = \frac{S_R - S_L}{2(A - B)} \quad (5)$$

$$a = \frac{2S_M - 5(A + B)}{2(A - B)} h - \frac{1}{12} c \quad (6)$$

区域灰度模型算法以相邻两个像素灰度差值为判定图像边缘的条件，原算法中选定一个固定灰度差作为阈值，人工选定适用性较小，同时存在精度检测不稳定、不准确等问题。本文引入OTSU自适应阈值法，该方法假定图像由前景和背景两大类像素组成，通过计算求取这两部分之间的像素值的类间方差最大化，进而推导出最优阈值。表达式如式(7)所示：

$$\sigma^2 = \omega_1 * \omega_2 * (\mu_1 - \mu_2)^2 \quad (7)$$

式中： σ^2 为所求类间方差； ω_1 和 ω_2 为图像中的像素点被分为两类像素的概率； μ_1 和 μ_2 为两类像素的均值。最终图像边缘检测结果如图4所示，图4(a)为预处理后的原图，其工件内部存在一些非边缘点被错误地识别成边缘点，经OTSU自适应阈值改进后，结果如图4(b)所示，这部分非边缘点被有效去除。

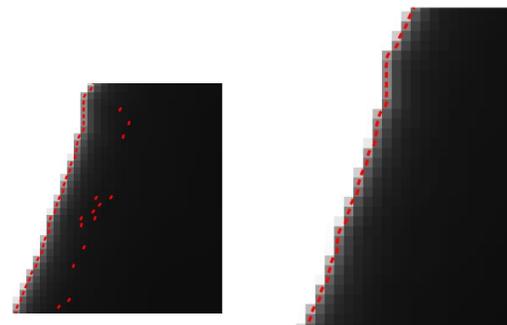


图4 区域灰度模型检测对比图

1.3 缺陷处理

圆形工件表面并不完全平整，存在凹痕，甚至工件内部圆孔在切削时呈倾角切削，这部分图像与邻近像素点存在较大的灰度差，因此在亚像素边缘检测时被识别成边缘像素点，严重影响尺寸检测精度。像素级边缘检测中Canny算子完成的边缘检测能够有效解决这一问题^[11]。但像素级边缘检测精度较低。本文提出一种基于Canny算子的圆形工件缺陷检测方法。

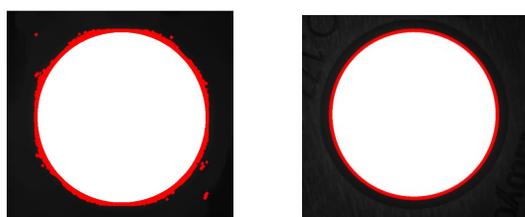
完成亚像素边缘检测后，采用Canny算子对图像实现像素级边缘检测，获取圆形图像像素级边缘点坐标；然后，采用像素级边缘点坐标的最值寻找圆形图像像素级中心点坐标和圆形半径，并通过半径定义一个像素级圆环；最后依据半径公式完成圆环内亚像素级边缘点坐标的搜寻。像素级半径 r 表示为

$$r^2 = \frac{x_{\max} - x_{\min}}{2} \quad (8)$$

式中： x_{\max} 和 x_{\min} 表示像素级边缘点最大和最小的横坐标。亚像素边缘点坐标与像素级中心点坐标形成的半径 r_n 满足以下约束条件：

$$r - 1 \leq r_n \leq r + 1 \quad (9)$$

此约束条件形成了亚像素边缘点的检测圆环，防止存在漏检的情况，阻止误将圆形图像的亚像素级边缘点去除的情况发生。检测结果如图5所示。图5(a)为采用矩形感兴趣区域完成缺陷处理的原图，圆形工件四个角的干扰像素点无法去除，本文改进后的缺陷处理结果如图5(b)所示，这部分缺陷点被有效去除。



(a) 矩形感兴趣区
缺陷处理原图

(b) 改进圆环型
缺陷处理原图

图5 圆形缺陷处理对比图

1.4 尺寸检测

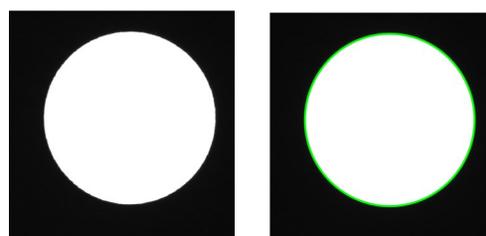
曲线边缘存在细小毛刺无法直接去除，影响尺寸检测结果，因此采用拟合的方式完成尺

寸检测，提高精准度。本文采用最小二乘法进行改进拟合，拟合公式如式(10)所示：

$$Q_{(a,b,c)} = \sum \sigma_i^2 = \sum [X_i^2 + Y_i^2 + aX_i + bY_i + c]^2 \quad (10)$$

式中： a 、 b 、 c 表示曲线方程的参数； X_i 和 Y_i 表示样本点的横纵坐标； σ_i 表示样本点与圆边缘点距离的平方和半径平方的差。

通过拟合后的曲线方程计算图形边缘点和圆中心点的距离求取圆形直径长度，结果如图6所示，图6(a)为圆形原图，图6(b)为圆形拟合图，其较好地拟合圆形边缘。



(a) 圆形原图

(b) 圆形拟合图

图6 圆形拟合边缘图

2 结果与结论

2.1 测量结果

测试相机选用海康威视600万像素工业面阵工业相机型号：MV-CA060-11GM，传感器型号：Sony IMX178，像元尺寸：2.4 μm*2.4 μm，分辨率：3072*2048。

拍摄镜头选用DTCM110-36，0.2倍双侧远心镜头。

实验测试设备选用处理器：Intel® Core™ i7-10750H CPU@ 2.60 GHz，内存：16.00 GB (3200 MHz)，显卡：NVIDIA GeForce GTX 1650 Ti。

本文针对圆形工件进行尺寸检测。工件标准尺寸直径为12 mm。在测量中对该工件进行多角度旋转拍摄，共进行10次测量，得到最终测量尺寸误差，见表1。圆形直径平均误差达到0.0007 mm，最大误差小于0.002 mm，相较于原区域灰度模型检测算法检测精度有较大提高。有部分测量数据变化较大，该类数据可能受光线变化等环境因素影响，对图像质量产生了影响。

表 1 圆形工件尺寸检测结果 单位:mm

检测算法	原算法	误差	本文算法	误差
1	11.9989	0.0011	11.9991	0.0009
2	11.9986	0.0014	11.9988	0.0012
3	11.9989	0.0011	11.9990	0.0010
4	11.9988	0.0012	11.9987	0.0013
5	12.0016	0.0016	11.9996	0.0004
6	12.0026	0.0026	11.9995	0.0005
7	11.9993	0.0007	11.9996	0.0004
8	12.0012	0.0012	11.9991	0.0009
9	12.0017	0.0017	11.9995	0.0005
10	12.0014	0.0014	11.9996	0.0004
均值	12.0003	0.0012	11.9992	0.0007

2.2 结论

为更精确地检测圆形工件尺寸,提出基于改进区域灰度模型亚边缘检测,基于Canny算子完成的改进缺陷处理算法,并通过最小二乘法完成拟合,仿真实验表明:通过形态学完成图像的预处理,并将OTSU自适应阈值与区域灰度模型检测相结合改进亚像素边缘检测,提升边缘检测精度;运用Canny算子确定像素级圆环,并依据该圆环搜寻圆形图像亚像素边缘点完成缺陷处理,再采用最小二乘法拟合边缘点,能够有效地提高尺寸检测精度。

参考文献:

[1] HUANG C, JIN W, XU Q, et al. Sub-pixel edge detection algorithm based on Canny-Zernike moment method[J]. Journal of Circuits Systems and Computers,

2020,29(15):2050238.

- [2] KISWORO M, VENKATESH S, WEST G. Modeling edges at subpixel accuracy using the local energy approach[C]//Proceedings of the IEEE Transactions on Pattern Analysis and Machine Intelligence. Los Alamitos:IEEE Computer Society, 1994:405-410.
- [3] BIN T J, AO L, CUI J W, et al. Subpixel edge location based on orthogonal Fourier-Mellin moments[J]. Image & Vision Computing, 2008, 26(4):563-569.
- [4] PAP L, ZOU J J. Sub-pixel edge detection for photogrammetry using laplace difference of Gaussian and 4th order ENO interpolation [C] //Proceedings of the 2010 IEEE International Conference on Image Processing. Piscataway: IEEE, 2010:2841-2844.
- [5] 刘国栋, 刘炳国, 陈凤东, 等. 亚像素定位算法精度评价方法的研究[J]. 光学学报, 2009(12): 3446-3451.
- [6] 张宗华, 王森, 王宇莹, 等. 基于编码环带径向直线拟合的圆环编码标志点中心提取方法[J]. 激光与光电子学进展, 2020, 57(7):212-220.
- [7] 于祥, 富容国, 黄明柱, 等. 基于改进 Zernike 矩的远场激光高精度中心测量[J]. 应用光学, 2023, 44(6):1244-1249.
- [8] TRUJILLO-PINO A, KRISIAN K, ALEMAN-FLORES M, et al. Accurate subpixel edge location based on partial area effect[J]. Image & Vision Computing, 2013, 31(1):72-90.
- [9] 张智凡, 于凤芹. 改进局部灰度特征的工件亚像素级边缘检测[J]. 软件导刊, 2019, 18(6):158-162.
- [10] 宿晓玮. 基于像元分解的标志点亚像素定位研究[D]. 徐州:中国矿业大学, 2014.
- [11] 申嘉锡, 齐华, 王晨. Canny算子对图像边缘检测的一种改进[J]. 现代计算机, 2022, 28(3):46-49.

Dimension measurement of circular precision components based on machine vision

Yu Jiahao, Liu Jiacheng, Wang Liankai, Yang Zhan, Song De*

(School of Physics, Changchun University of Science and Technology, Changchun 130013, China)

Abstract: A fast and accurate detection algorithm based on regional grayscale model is proposed to address the difficulties in measuring the dimensions and the high cost of detection for circular precision components. Research on Machine Vision-Based Circular Precision Dimension Measurement Methods. Firstly, preprocessing of the image is accomplished through morphological operations; Then, in the Partial-area-effect model algorithm, OTSU adaptive thresholding is introduced to improve detection accuracy; Afterwards, defect handling is accomplished through a circular ring defined by the Canny operator; Finally, the diameter length of the circular shape is fitted using the least squares method. The method boasts high computational speed and accuracy, with a maximum measurement error of less than 0.002 mm and an average measurement error precision of less than 0.0007 mm, meeting the precision requirements for workpiece inspection.

Keywords: precision testing; machine vision; partial-area-effect model

图形图像

文章编号: 1007-1423(2024)20-0045-07

DOI: 10.3969/j.issn.1007-1423.2024.20.009

SOIQA: 一种图像超分领域的实用型图像质量评价方法

法静怡, 严广宇*

(沐曦集成电路(上海)有限公司, 上海 200131)

摘要: 主观图像质量评价需要大量人员参与, 其结果 MOS 可信性高, 但需花费较多人力物力, 客观图像质量评价一般采用数学建模的方式, 计算过程简便, 但往往和主观评价的一致性不高。SOIQA 是一种图像超分领域的实用型图像质量评价方法, 基于已有方法, 结合深度学习, 以 UNet 为基础模型, 预测可靠的评价结果, 用于指导超分算法的迭代。实验证明, SOIQA 在数据集上的准确率比一般客观评价方式高出 12%, 和主观评价 MOS 有较高的一致性。

关键词: 图像质量评价; 深度学习; Unet

0 引言

图像质量评价通过对图像特征的分析, 衡量图像的失真程度, 对算法分析、系统性能评估等方面具有重要意义。MOS (Mean Opinion Score) 是一种常用的主观质量评价方法, 通过大量评分者对图像的清晰度、饱和度等方面进行评价, 得到人眼对图像的视觉观感优劣判断。这种评价方式耗费大量的人力, 需要设计缜密的对比实验, 因此长久以来, 研究者们致力于探索各种客观质量评价方法, 用来替代或补充 MOS 评价。

可靠的图像质量评价方法不仅能够满足研发人员对相关系统性能进行快速大批量的测试需求, 对特定领域的算法迭代也具有指导性意义。PSNR 和 SSIM 由于建模原理明确且计算简单, 是评价图像超分算法的主要手段, 但随着超分技术的发展, 不同算法在这两项客观指标上的效果差距越来越小, 在某些特定领域或场景下的主观视觉效果却有明显差异^[1]。

目前主流的方法是抽取图像的高级语义特征, 并基于不同的模型评估失真图像特征和参考图像特征的距离或相似度, 以此构建评价方式^[2-5]。除了持续探索不同的建模方式, 是否

够基于已有的、较为成熟的评价方法, 构建出一套综合性的图像质量评价方式, 其评估结果更加接近主观评价? SOIQA (Stack Objective Image Quality Assessment), 是一种基于 PSNR、SSIM、VMAF 和 NIQE 的评估结果, 并结合 UNet 模型^[6]的图像质量评价方法, 得到最终可靠的评价结果, 用来指导图像超分算法的迭代。SOIQA 在图像质量评价数据集^[3, 7-8]上的评价结果和 MOS 的一致性达到 76.73%, 而 PSNR、SSIM、VMAF 和 NIQE 的一致性分别是 60.38%、61.20%、64.47% 和 49.65%。

1 图像质量评价的相关工作

人眼对于脸部、树枝等尺度较小但细节丰富的纹理变化十分敏感, 但对天空、毛发、树林等尺度较大或纹理杂乱区域的色调轻微变化、过平滑(细节损失)等变化不敏感^[9]。PSNR 和 SSIM 作为一直沿用的图像超分评价方式, 由于其背后数学模型较为简单, 评价结果也相对片面。VMAF 和 NIQE 作为近年来较新的评价方法, 更接近主观评价结果, 但可靠性仍然受到色调变化、对比度变化的影响。本文第一节将进一步介绍这几种算法的优缺点。

收稿日期: 2024-04-10 修稿日期: 2024-06-19

基金项目: 上海市浦江人才计划(22PJD091)

作者简介: 法静怡(1994—), 女, 江苏镇江人, 硕士, 研究方向为视觉图像算法; *通信作者: 严广宇(1996—), 男, 湖北孝感人, 硕士, 研究方向为视觉图像算法, E-mail: guangyu.yan@metax-tech.com

1.1 PSNR

PSNR 通过比较失真图像和参考图像像素差来计算。由于其物理意义明确、计算简便，常被用于衡量图像处理算法的表现。PSNR 越高，经过处理的图像丢失的信息越少。PSNR 的核心部分是 MSE(均方误差)，常作为图像超分任务的损失函数使用。MSE 是像素级别的损失计算，没有考虑图像的纹理和整体观感，某些场景下的评价可能会失去意义^[3,10]。如图 1 所示，这组图中三种不同的噪音带来的主观视觉感受差异很大，但 MSE 的结果极为接近。

1.2 SSIM

SSIM 从亮度、对比度和结构三个方面衡量两张图像的相似性^[1]，和图像均值以及方差相关，计算结果在 0 到 1 之间，值越高，两张图像相似度越高。相比 PSNR，SSIM 在某些场景下可靠性更高，例如图 1 中能够较好地地区分不同的噪音。但是这种能力依然受限于它的建模方式，对于某些程度的噪音，或者对于非结构性的噪音，例如色调转变或明暗度转变，SSIM 无法给出合理的评价^[11-12]，具体参考图 2。

1.3 VMAF

VMAF 是 Netflix 在 2016 年提出的一种基于机器学习的视频质量评价方法，核心有三个模块：VIF(visual quality fidelity, 视觉信息保真度)^[13]、DLM(detail loss measure, 细节损失指标)^[14]以及 TI(temporal information, 时域运动指标/平均相关位置像素差)。VMAF 的值在 0 到 100 之间，分数越高代表画质越好。VMAF 对多种场景、多帧特征的视频质量评价具有良好的鲁棒性，是目前互联网视频最主流的客观视频评价指标^[15]。然而，如图 3 所示，对于同一张图像，在细节和信息量不变的情况下，仅通过调节对比度就能大幅提高 VMAF 值，因此某些场景下仍旧不能很好地客观反映图像质量的高低。

1.4 NIQE

NIQE 是一种无参考的图像质量评价方法，通过神经网络得到一组清晰图像的特征，并以相同的方式计算待评估图像的特征，比较两者距离，给出待评估图像的质量评价^[3]，分值越低，图像质量越好。但 NIQE 很难给出一个明确的阈值来界定高质量的图像应取得多少分，并且可靠性也受到图像对比度改变的干扰，如图 4



注：图片来自电影《千与千寻》。

图 1 不同观感的图像，MSE 的值极为接近



注：图片来自电影《千与千寻》。

图 2 不同观感的图像，SSIM 值极为接近



图3 改变图像对比度使得VMAF值提升



图4 NIQE评估结果和主观感受不一致

所示。

2 算法实现

由于在图像超分领域, PSNR和SSIM是最为常见的两种评价指标, 能在像素级别和整体结构层面对修复后的图像和无损图像进行比较, 而VMAF和NIQE的结果更加符合主观评价, SOIQA是一种基于神经网络的图像质量评价方法, 通过综合考虑几种已有的图像质量评价方法进行综合计算, 其评价结果更加符合视觉主观评价。

建立IQA (Image Quality Assessment) 评价体系, 对于数据集的要求是: 对于一张参考图像, 需要一组对应不同程度失真的有损图像, 以及对应各自的MOS得分。IQA已有的数据集包括LIVE2、TID2013、CSIQ等。此外, 针对图像超分领域, 也有专门的数据集QADS, 数据集中包含了超分图像及其对应的MOS评价, 更详细的介绍请参考第三节。

2.1 数据处理

假设对于一张无损图像叠加一组种类和强度各不相同的噪音, 例如高斯模糊、白噪音等, 得到一组和无损图像对应的不同损失程度的有损图像, 这N张有损图像同时各自对应一个MOS值。对这组有损图像分别计算PSNR、SSIM、VMAF和NIQE, 可以得到:

$$ass_i = [p\ snr_i, ssim_i, vmaf_i, niqe_i, mos_i] \quad (1)$$

如上所示, 一条数据中包含了对某张有损

图像的四个客观评价指标以及主观评价MOS值, MOS值需要被统一缩放到 $[0, 1]$ 区间内。在这组数据中, 任意选取一条数据作为基础数据, 除此以外其他 $N-1$ 条数据分别和基础数据相减,

$$dist_j = [dp\ snr_j - dp\ snr_i, ssim_j - ssim_i, vmaf_j - vmaf_i, niqe_j - niqe_i, mos_j - mos_i] \quad (2)$$

由于NIQE的值越低表示图片质量越高, 因此相减后得到差值需要取反, 和其他三个客观指标保持一致性。对于图像超分来说, 不同的算法, 如果效果接近, 那么PSNR的差值一般在 ± 0.3 dB左右, SSIM差值在 ± 0.1 左右, VMAF差值在 ± 25 左右, NIQE差值在 ± 2 左右。此外, 如果MOS差值超过 ± 0.3 , 则很容易区分图像质量高低。因此剔除所有差值不符合上述要求的数据, 最后使用min-max方法让所有差值归一化, 得到:

$$ratio_i = [rp\ snr_j, rssim_j, rvmaf_j, rniqe_j, rmos_j] \quad (3)$$

这样一条数据称作一个实例, 代表两张图像之间客观评价得分的距离和主观评价得分的距离。一个好的客观评价模型, 在比较两张图像质量时, 得到的结果需要与主观评价有一致性。

2.2 提出方法

一个合理的思路是分别给客观距离赋予不同的权重, 使其相加后的结果等于主观距离,

表1 不同图像质量评价方法比较

通道数	数据集名称(实例个数)								平均准确率/%	MAE
	LIVE2(128)		CSIQ(163)		TID2013(613)		QADS(503)			
	正确实例个数	准确率/%	正确实例个数	准确率/%	正确实例个数	准确率/%	正确实例个数	准确率/%		
PSNR	86	67.19	81	49.69	346	56.44	343	68.19	60.38	—
SSIM	94	73.44	72	44.17	347	56.61	355	70.58	61.20	—
VMAF	74	57.81	116	71.17	394	64.27	329	65.41	64.67	—
NIQE	56	43.75	74	45.40	337	54.98	274	54.47	49.65	—
线性回归	78	60.93	104	63.80	436	71.12	360	71.57	66.86	0.1307
SOIQA	109	85.16	117	71.78	443	72.27	391	77.73	76.73	0.1279

基于此想法，线性回归的结果展示在表1中，表中统计的是对于所有测试实例，预测结果和主观结果相符的实例个数，在一个实例中，如果某个评价方式给出的评价，和MOS给出的评价具有一致性，那么就被统计为正确实例。准确率即为正确实例占总实例个数的比例。从实验结果可知，将四种客观评价结果做线性回归得到的结果，比起单独某个客观评价更加接近主观评价。由此得到启发，可以设计更合理的模型SOIQA，综合考虑多种客观评价指标，预测得到和主观评价具有更高一致性的评价结果。

2.2.1 UNet模型

UNet模型^[6]在2015年被首次提出，其典型的结构是通过一条short-cut将不同阶段的特征进行融合，如图5所示。基于UNet模型，SOIQA将四种客观距离作为输入，经过网络预测后得到输出结果。模型的左侧将四个距离通过网络映射到高维空间，右侧则拟合系数，最终将不同维度特征分别和该维度下的系数相乘，映射回一维空间，整个过程即用神经网络模拟一个关于距离的高维函数。

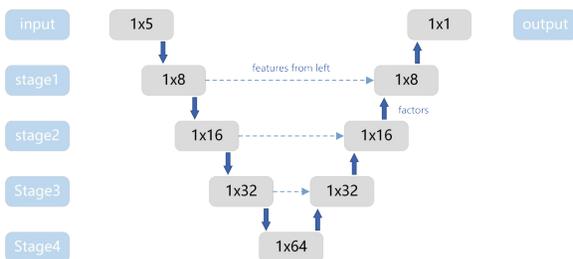


图5 SOIQA基于UNet模型将输入映射到高维空间

2.2.2 符号提示

比起直接拟合最终结果，如果能给模型一些提示，例如在某个实例中，对比图像比基础图像的质量更好还是更差，那么模型将更容易收敛到具体好多少或者差多少。这个提示被称为符号提示，它能一定程度地改善损失函数MAE(mean-absolute-error)所造成的收敛困扰。由于MAE计算的是平均绝对误差，并没有考虑正负性，因此如果在输入中额外增加一维关于正负性的提示，有助于提高模型准确性。得到符号提示的方法有很多，这里采取了GBDT(Gradient Boosting Decision Tree)^[16]算法，将符号问题当作分类问题，输入为四个距离，输出为0或1，表示对比图像比基础图像质量更好或更差。

3 实验

根据第二节内容，将QADS、LIVE2、TID2013和CSIQ处理成6804条训练实例以及1407条测试实例。优化器选择SGD，初始学习率设置为0.01，损失函数设置为MAE，训练170个epoch。

3.1 数据集介绍

LIVE图像质量评价数据集是首个高分辨率的主观图像质量评价数据集^[1]，收集了人眼对不同程度、不同噪音叠加的图片的主观MOS打分。2006年LIVE数据集^[1]经过重新校正后发布了LIVE2，并给出相应DMOS(difference MOS)。CSIQ数据集由Oklahoma State University制作提供^[7]，失真图像在四台校正过的LCD显示屏上供实验人员打分。TID2013在TID2008的基础上

扩展了噪音种类^[8]。在图像超分领域, QADS^[17]收集了21种超分算法, 并制造出980张超分后的图像, 填补了图像超分领域质量评价数据集的空白。以上这些数据集, 每张有损图像对应一张相应的无损图像, 并以MOS或DMOS作为主观评价, 制作图像质量评价数据集的过程十分繁琐, 并且需要大量的人力进行评估。

3.2 对比经典图像质量评价方法

将SOIQA模型预测结果和PSNR、SSIM、VMAF以及NIQE的评估结果作对比, 在测试实例中统计和主观评价一致的实例个数, 即客观评价对两张图片给出分数, 需要和主观评价的分数表现出一致性。统计结果见表1。根据实验结果, SOIQA模型取得平均76.73%的准确率, 而其他指标准准确率分别为60.38%(PSNR)、61.20%(SSIM)、64.67%(VMAF)以及49.65%(NIQE)。除此以外, 为了评估图像超分算法在影视剧上的效果, 实验选择了两种视频超分算法, 分别是VRT^[18]和BasicVSR++^[19], 以及一种图像超分算法SwinIR-GAN^[20]。测试图像来源于影视剧片段进行抽帧, 经过4倍降采样后得到低分辨率图像, 通过不同的超分辨算法, 还原到原始图像尺寸。VRT和BasicVSR++是视频超分算法中PSNR、SSIM指标较高的算法, 可以较好地还原真实图像, 但噪音也可能被随之放大, 而GAN类的算法则在填补细节上具有优势, 整体观感较好, 但过于平滑或凭空产生的细节和

原图差异较大, 从而导致PSNR等指标得分不高。如图6及表2所示, SwinIR-GAN算法效果在整体观感上最优, 但如果细节和原图对比相差过大, 也会影响最终评价。

表2 关于图6的图像质量评估结果

	BasicVSR++	VRT	SwinIR
PSNR	34.51	34.75	32.37
SSIM	0.934	0.936	0.935
VMAF	70.26	71.03	76.23
NIQE	13.86	14.16	11.50
SOIQA	60.98	63.33	64.27

3.3 消融实验

为了研究模型结构和符号提示对最终结果的提升, 实验设置了三组不同模型结构在不同通道数下的训练, 第一种非线性模型只设计了几层全连接层以及激活函数, 将四种评价距离作为输入, 预测最终得分; 第二种模型为基于UNet的模型, 输入为四种评价距离, 输出为最终得分。从实验结果可知, 当模型从简单的非线性模型换为UNet以后, 模型在不同数据集上平均准确率由54.83%提升到74.61%。第三种模型是在UNet模型基础上, 增加了一维符号提示作为输入, 模型准确率进一步提升到76.73%。消融实验证明了UNet比普通非线性模型能够更好地拟合关于距离的函数, 并且符号提示能够帮助模型进一步提升准确率。具体结果见表3。



(a) 无损原图

(b) BasicVSR++超分修复

(c) VRT超分修复

(d) SwinIR超分

注: 图片来自电影《少年派的奇幻漂流》。

图6 不同超分算法(4倍超分)细节对比

表3 消融实验

通道数	数据集名称(实例个数)								平均准确率/%	MAE	
	LIVE2(128)		CSIQ(163)		TID2013(613)		QADS(503)				
	正确实例个数	准确率/%	正确实例个数	准确率/%	正确实例个数	准确率/%	正确实例个数	准确率/%			
非线性模型	32	64	50.0	113	69.33	307	50.08	251	49.90	54.83	0.1312
	64	64	50.0	83	50.92	309	50.41	251	49.90	50.31	0.1505
	96	64	50.0	83	50.92	307	50.08	252	50.10	50.28	0.1424
SOIQA- 无符号提示	32	104	81.25	116	71.17	426	69.49	385	76.54	74.61	0.1127
	64	102	79.69	111	68.10	425	69.33	384	76.34	73.37	0.1251
	96	98	76.56	116	71.17	422	68.84	383	76.14	73.18	0.1127
SOIQA- 有符号提示	32	105	82.03	119	73.01	441	71.94	391	77.73	76.18	0.1181
	64	109	85.16	117	71.78	443	72.27	391	77.73	76.73	0.1279
	96	108	84.38	120	73.62	441	71.94	387	76.94	76.72	0.1122

4 结语

SOIQA 是一种基于 PSNR、SSIM、VMAF 和 NIQE 的评估结果, 结合深度学习方法的图像质量评估方法, 可以得到更加符合主观评价的评估结果, 用于比较不同图像超分算法的效果, 并以此结果指导算法的迭代。实验结果显示, SOIQA 的评价结果比单独使用 PSNR、SSIM、VMAF 以及 NIQE, 都更接近主观评价。这种综合评价方式不仅仅在图像质量评价领域, 在其他领域依然适用。

参考文献:

- [1] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.
- [2] RASSOOL R. VMAF reproducibility: validating a perceptual practical video quality metric [C] // Proceedings of the 2017 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB). IEEE, 2017: 1-2.
- [3] SAAD M A, BOVIK A C. Blind quality assessment of videos using a model of natural scene statistics and motion coherency [C] // Proceedings of the 2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR). IEEE, 2012: 332-336.
- [4] GOODALL T, BOVIK A C. No-reference task performance prediction on distorted LWIR images [C] // Proceedings of the 2014 Southwest Symposium on Image Analysis and Interpretation. IEEE, 2014: 89-92.
- [5] ZHANG R, ISOLA P, EFROS A A, et al. The unreasonable effectiveness of deep features as a perceptual metric [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 586-595.
- [6] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation [C] // Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer International Publishing, 2015.
- [7] LARSON E C, CHANDLER D M. Most apparent distortion: full-reference image quality assessment and the role of strategy[J]. Journal of Electronic Imaging, 2010, 19(1): 011006-011006-21.
- [8] PONOMARENKO N, JIN L, IEREMEIEV O, et al. Image database TID2013: peculiarities, results and perspectives-science direct[J]. Signal Processing: Image Communication, 2015, 30: 57-77.
- [9] LIANG J, ZENG H, ZHANG L. Details or artifacts: a locally discriminative learning approach to realistic image super-resolution [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 5657-5666.
- [10] NDAJAH P, KIKUCHI H, YUKAWA M, et al. SSIM image quality metric for denoised images [C] // Proc. 3rd WSEAS Int. Conf. on Visualization, Im-

- aging and Simulation, 2010:53-58.
- [11] BAO D Q. Image blur metric[EB/OL], 2024. <https://www.mathworks.com/matlabcentral/fileexchange/24676-image-blur-metric>.
- [12] ZORAN K, PECE M. Experimental comparison of PSNR and SSIM metrics for video quality estimation [C] //Proceedings of the International Conference on ICT Innovations. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009:357-366.
- [13] SHEIKH H R, BOVIK A C. Image information and visual quality [J]. IEEE Transactions on Image Processing, 2006, 15(2): 430-444.
- [14] LI S, ZHANG F, MA L, et al. Image quality assessment by separately evaluating detail losses and additive impairments[J]. IEEE Transactions on Multimedia, 2011, 13(5): 935-949.
- [15] RASSOOL R. VMAF reproducibility: validating a perceptual practical video quality metric [C] //Proceedings of the IEEE International Symposium on Broadband Multimedia Systems & Broadcasting. IEEE, 2017.
- [16] FRIEDMAN J H. Greedy function approximation: a gradient boosting machine [J]. Annals of Statistics, 2001, 29: 1189-1232.
- [17] ZHOU F, YAO R, LIU B, et al. Visual quality assessment for super-resolved images: database and method [J]. IEEE Transactions on Image Processing, 2019, 28(7): 3528-3541.
- [18] LIANG J, CAO J, FAN Y, et al. VRT: a video restoration transformer [J]. IEEE Transactions on Image Processing, 2022: 2171-2182.
- [19] CHAN K C K, ZHOU S, XU X, et al. BasicVSR++: improving video super-resolution with enhanced propagation and alignment [C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 5972-5981.
- [20] LIANG J, CAO J, SUN G, et al. SwinIR: image restoration using swin transformer [C] //Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 1833-1844.

Stack objective image quality assessment: A practical evaluation method in super resolution task

Fa Jingyi, Yan Guangyu*

(MetaX Integrated Circuits(Shanghai) Co., Ltd., Shanghai 200131, China)

Abstract: Subjective assessments provide reliable evaluations that are established through significant amounts of manpower and material resources. Objective assessments based on mathematical models can be easily implemented, but they fail to capture the nuances of human judgement and therefore lack consistency with subjective assessments. Stack Objective Image Quality Assessment (SOIQA) is a practical objective assessment, which combines existing classic image quality assessments based on deep learning to predict image quality close to subjective evaluation. SOIQA is applied to the task of image super-resolution (SR) and utilized PSNR, SSIM, VMAF, and NIQE assessments in SOIQA, as they are commonly used in SR tasks. To stack these assessments effectively, a model is designed encouraged by UNet model. Experimental results show that SOIQA can achieve better image quality evaluation performance with accuracy of 12% higher than the existing objective assessments in MOS consistency.

Keywords: image quality assessment; deep learning; Unet

实践与经验

文章编号: 1007-1423(2024)20-0052-06

DOI: 10.3969/j.issn.1007-1423.2024.20.010

DDoS 攻击的发展与检测技术研究

马立鑫, 薛占双, 刘海燕*

(北华航天工业学院计算机学院, 廊坊 065000)

摘要: 在网络安全领域, 分布式拒绝服务攻击(DDoS)因其对信息系统的严重威胁而备受关注。研究了 DDoS 攻击的演进历程, 从 20 世纪 90 年代的洪水攻击到使用僵尸网络、物联网(IoT)设备和云服务的复杂攻击模式, 展示了攻击方法和技术的多样化。还对 DDoS 攻击进行了详细的分类, 包括基于攻击手段和流量的分类方法, 以及基于流量分析、行为分析、机器学习和深度学习的检测技术, 指出了这些技术的优点和局限性。此外, 还介绍了当下较为成熟的 DDoS 检测工具与系统, 如 Suricata Snort, Bro(Zeek)以及防护服务 Cloudflare, 并讨论了 DDoS 攻击未来的发展趋势。

关键词: 网络安全; 分布式拒绝服务; 攻击检测; 深度学习

0 引言

随着计算机技术迅猛发展, 在推动全球信息化的同时也给网络安全带来了严峻的挑战。在诸多安全威胁中分布式拒绝服务攻击(distributed denial of service, DDoS)的危害尤为突出, 攻击者利用分布在不同地理位置的计算机或设备发送大量恶意流量至目标系统, 旨在消耗其资源、带宽或处理能力, 从而导致合法用户无法正常访问该系统的网络服务^[1-3]。

近年来, DDoS 攻击呈现愈演愈烈的趋势。攻击规模和手段都呈现出增多和多样化的态势。DDoS 攻击通常利用僵尸网络(botnet)来实施, 即由大量被恶意软件感染控制的计算机或设备组成的网络。这些受感染的设备在攻击者的操控下, 协同发起大规模的恶意流量攻击, 对目标系统造成压力和干扰^[4-5]。DDoS 攻击不仅导致正常用户无法访问网站或使用网络服务, 在攻击期间, 攻击者还可能进行数据窃取、篡改或破坏, 造成敏感数据泄露, 给企业带来严重的经济损失等危害^[6]。例如 2016 年 9 月, Mirai 僵

尸网络攻击物理设备致使美国多个城市的互联网瘫痪。2023 年 4 月西部数据遭网络攻击, 造成云服务瘫痪; 同年 6 月厦门一公司系统被攻击, 超百万条用户信息泄露并被黑客滥用。2024 年 2 月, 韩国某竞技联赛遭 DDoS 攻击, 一场系列赛打了 7 个小时, 导致了比赛的暂停与推迟^[7-9]。

深入研究 DDoS 攻击的发展和检测技术对于提升网络安全、减少经济损失、推动技术创新、增强网络防御能力具有重要意义。本文研究 DDoS 攻击的发展和检测技术, 首先介绍 DDoS 的发展和分类, 其次介绍其检测技术和工具。

1 DDoS 攻击的发展

根据 DDoS 的攻击手段可以将其分为以下五个阶段: 早期阶段、僵尸网络阶段、复杂化阶段、利用 IoT 设备阶段和云服务下的 DDoS^[10-14]。

1.1 早期阶段

在早期阶段, DDoS 攻击的手段主要采用简单且直接的洪水攻击方式, 如 SYN Flood、UDP Flood、ICMP Flood 等^[15]。该阶段主要以网络层

收稿日期: 2024-04-30 修稿日期: 2024-05-31

基金项目: 河北省教育厅项目(CXY2023013)

作者简介: 马立鑫(2000—), 男, 河北衡水人, 在读硕士研究生, 研究方向为计算机技术; 薛占双(1999—), 男, 河北唐山人, 在读硕士研究生, 研究方向为计算机技术; *通信作者: 刘海燕(1981—), 女, 河北廊坊人, 硕士研究生, 副教授, 研究方向为软件工程, E-mail: 37206446@qq.com

为攻击对象,通过发送大量伪造的请求或数据包使目标系统资源耗尽。

1.2 僵尸网络阶段

随着PC数量的增长,DDoS攻击进入僵尸网络阶段。该时期攻击类型依然包括SYN Flood、UDP Flood等基础类型,但攻击者利用大规模感染的僵尸网络发动攻击,分散攻击流量来源,增加攻击规模和威力,难以单点防御^[16]。

1.3 复杂化阶段

网络技术的不断进步,攻击工具和技术也在不断演进,更为复杂的反射放大型DDoS攻击出现,该阶段引入更多复杂的攻击手法,如HTTP Flood、DNS Amplification、Slowloris等,主要针对应用层协议或服务器的特定漏洞进行攻击^[17]。其攻击更具针对性和破坏力,能绕过传统的防护机制,对目标系统造成更大影响。

1.4 利用物联网设备阶段

随着物联网(Internet of Things, IoT)设备在智能家居、智慧医疗、生态环境监测、工业生产等领域取得成功,攻击者开始利用存在安全漏洞的IoT设备构建庞大的僵尸网络,增加攻击规模和威力。据绿盟科技发布的《2020 DDoS攻击态势报告》,在攻击者用来发起DDoS攻击的所有资源中,31%是物联网设备^[18]。

1.5 云服务下的DDoS

随着技术的进步以及市场的需求,云服务逐渐兴起,改变了传统的IT消费模式,使得计算资源像水和电一样,成为可按需获取的公共服务。正因云服务具有资源集中化、易于匿名和可扩展性等特性以及攻击工具的普及,使得攻击者产生了攻击云服务的动机。攻击者攻击云服务的手段包括传统DDoS攻击的平移,还包括云平台特有的新型攻击,如云滴冻结攻击、流表泛洪攻击、功率攻击等^[19-20]。

2 DDoS攻击的分类

DDoS攻击可以根据不同的分类方式进行区分,常见的分类方式基于攻击手段、基于攻击目标、基于攻击流量等。基于攻击手段可以将DDoS攻击分为三类:网络层攻击、传输层攻击

和应用层攻击。在这些攻击中又可以分为固定带宽和混合型攻击两种。

(1)网络层攻击:针对网络层协议进行攻击,通过发送大量伪造的数据包使目标系统资源耗尽。例如:ICMP Flood攻击就是一种经典的网络层攻击,攻击利用ICMP协议向目标系统发送大量的ICMP Echo请求,造成目标系统的网络带宽和处理能力耗尽,导致网络服务不可用或延迟增加^[21]。

(2)传输层攻击:主要针对传输层协议进行攻击,通常包括TCP连接洪泛和UDP洪泛等。例如:UDP Flood攻击会向目标系统发送大量UDP数据包,消耗目标系统的带宽和处理能力。

(3)应用层攻击:针对应用层协议进行攻击,利用协议或服务的特定漏洞对目标系统发动攻击。例如:HTTP Flood攻击会发送大量HTTP请求给目标网站,使其服务器超载无法正常提供服务^[22]。

基于攻击流量的特点可以分为固定带宽攻击和混合型攻击。固定带宽攻击中攻击者通过大规模洪泛式攻击消耗目标系统的带宽,导致合法流量无法正常传输。而对于混合型攻击则结合多种攻击手法同时对目标系统发动攻击,使得防御更加困难^[23-24]。

3 DDoS攻击检测技术

为了有效检测DDoS攻击,研究人员提出了多种检测技术,包括基于流量分析、行为分析、机器学习和深度学习的方法。流量分析关注网络流量特征,行为分析依赖于正常行为模式的建模,而机器学习和深度学习则通过训练模型来识别攻击流量。

基于流量分析的DDoS攻击检测技术以其实现简便性和对大规模流量波动的敏感性而具有优势,尤其对传统洪泛式攻击有效。然而,它在检测低带宽和复杂智能攻击方面存在局限性,可能导致误报或漏报,且对应用层攻击如HTTP Flood的识别能力不足。因此,该技术更适合于快速识别显著流量异常的场景^[25-27]。

基于行为分析的DDoS攻击检测技术通过监测系统行为特征,如连接数和请求频率,有效识别异常行为,尤其在检测应用层攻击方面表

现出色。然而，其对正常用户行为的敏感性可能导致较高的误报率，且在面对多变的攻击手法时适应性有限。因此，它更适用于需要基于特定行为模式进行检测的环境，对某些已知的应用层攻击具有较好的检测效果^[28-29]。

基于机器学习的DDoS攻击检测技术通过分析大量样本数据，能够有效识别复杂的攻击模式，并具备良好的自适应性，尤其对新型和智能攻击有较好的检测效果。然而，这种方法依赖于充足的训练数据，并且需要定期更新模型以应对新兴攻击。在实时性要求高的环境中，还需考虑算法的计算成本和可能的延迟。因此，它适合于需要持续学习和适应新攻击模式的场景，以提升检测的准确性和适应性^[30-32]。

基于深度学习的DDoS攻击检测技术具有卓越的复杂数据模式处理能力和高维数据处理能力。它能够自动提取关键特征，无需人工干预，并且具备强大的泛化能力，即使面对未知攻击类型也能进行有效检测。然而，这种方法依赖于大量的标记训练数据，并且模型训练的计算成本相对较高。因此，它最适合于需要处理大规模、高维度数据集，并且能够自动识别复杂攻击模式的场景^[33-35]。

4 DDoS攻击检测工具与系统

目前已经存在了一些较为成熟的DDoS攻击检测工具和系统，他们通常用于监视网络流量，识别异常模式，并采取措施来缓解或阻止攻击。例如：思科的Snort、开放信息安全基金会的Suricata和Bro等。

Snort是一个开源的网络入侵检测系统，它能够实时监控网络流量，并根据一系列预定义的规则来检测潜在的恶意活动。它可以用来检测DDoS攻击和其他网络入侵行为。Snort是一款轻量级、开源的网络入侵检测系统，适用于中小型组织或个人用户^[36-38]。

类似于Snort，Suricata也是一个开源的网络入侵检测系统，它支持多线程处理和高性能特性，可以用于检测和阻止DDoS攻击。适用于中大型组织或网络服务提供商，可以处理大量的网络流量^[39]。

Bro(现称为Zeek)是另一个流行的开源网络

安全监控系统，它能够实时分析网络流量，识别DDoS攻击并生成相应的日志和警报。适用于中大型组织或安全团队，可以用于深度分析网络流量^[40-41]。

Cloudflare是一家提供云安全服务的公司，其DDoS防御服务可以自动检测和缓解各种规模的攻击，并确保网站的可用性。适用于各种规模的网站和在线服务提供商，能够提供快速、自动化的DDoS防御服务，帮助客户保持其网络服务的可用性，并减少攻击造成的影响^[42-43]。

5 未来DDoS发展趋势

随着计算机技术的不断发展，在未来DDoS攻击可能会采用更加隐蔽和复杂的手段，以规避传统的防御措施。主要的攻击方式可能包含以下几个方面。

(1) 短时高强度攻击：攻击者可能会发起短时但高强度的DDoS攻击，以利用系统检测和缓解攻击所需的时间窗口，在防御系统响应之前造成尽可能大的破坏。

(2) 多向量攻击：攻击者可能会同时使用多个攻击向量，例如同时利用TCP、UDP和应用层攻击，以增加攻击的复杂性和防御的难度。

(3) 供应链攻击：攻击者可能会通过供应链中的薄弱环节发起攻击，例如通过感染第三方服务提供商的系统来间接影响目标组织。

(4) 自动化和人工智能：随着技术的发展，攻击者可能会利用自动化工具和人工智能来发起更复杂、更难以预测的DDoS攻击。

DDoS攻击检测技术的发展趋势是深度学习技术，深度学习模型可以更好地识别和区分正常流量和恶意攻击流量，它能够自动提取关键特征，无需人工干预，并且具备强大的泛化能力，即使面对未知攻击类型也能进行有效检测，因此深度学习技术在DDoS攻击检测中的应用将会更越来越广泛。

6 结语

DDoS攻击随着技术进步呈现多样化和复杂化趋势，对网络安全构成严重威胁。检测技术的发展，包括流量分析、行为分析、机器学习和深度学习，为应对DDoS攻击提供了多种手

段。现有工具和系统如Snort、Suricata等在监测和缓解攻击方面发挥重要作用。面对未来攻击手段的演进,研究应注重深度学习的应用,提高检测系统的实时性和自动化水平,并促进跨学科合作与数据共享,以增强全球网络安全防御能力。

参考文献:

- [1] 陈飞,毕小红,王晶晶,等. DDoS攻击防御技术发展综述[J]. 网络与信息安全学报, 2017, 3(10): 16-24.
- [2] 王印玺,黄华雪. DDoS攻击的发展与防御综述[J]. 现代计算机, 2021, 27(2): 51-56.
- [3] 靳燕. DDoS攻击特性分析与检测模型构建[J]. 网络空间安全, 2023, 14(1): 103-106.
- [4] ZEIDANLOO H R, MANAF A B A. Botnet detection by monitoring similar communication patterns [J]. International Journal of Computerence & Information Security, 2010, 7(3): 88-96.
- [5] 何娅莹,覃仁超,舒月,等. BRNet:基于特征复用的僵尸网络检测模型[J]. 计算机技术与发展, 2023, 33(4): 108-113.
- [6] FreeBuf网络安全门户. 2023年DDoS攻击现状及趋势报告 [EB/OL]. (2023-09-10) [2024-04-10]. <https://www.freebuf.com/articles/paper/377582.html>.
- [7] FreeBuf网络安全门户. 深度分析|Mirai新样本:新的伪装方式使其更难被识别 [EB/OL]. (2024-02-08) [2024-04-10]. <https://www.freebuf.com/articles/web/391832.html>.
- [8] 网络安全动态周刊. 2023年网络攻击事件盘点 [EB/OL]. (2023-12-29) [2023-12-29]. <https://zhuanlan.zhihu.com/p/675207396>.
- [9] 腾讯网. LCK官方遭受DDOS,解说声称感到无力 [EB/OL]. (2024-02-28) [2024-04-10]. <https://new.qq.com/rain/a/20240228A09LZ700>.
- [10] KIM Y. ATTENTION: ATtackEr traceback using MAC layer abNormality detecTION [C]//Advances in Information Security and Assurance: Third International Conference and Workshops, ISA 2009, Seoul, Korea, June 25-27, 2009. Proceedings 3. Springer Berlin Heidelberg, 2009: 239-249.
- [11] 陈亮,李志华. 面向物联网僵尸网络多阶段攻击的异常流量检测方法 [J/OL]. 计算机科学, 1-13 [2024-03-30]. <http://kns.cnki.net/kcms/detail/50.1075.TP.20231120.1028.012.html>.
- [12] 叶沛鑫. 面向物联网的DDoS攻击防御研究[D]. 重庆:重庆邮电大学, 2022.
- [13] 何娅莹. 基于神经网络的僵尸网络检测算法研究 [D]. 绵阳:西南科技大学, 2023.
- [14] 岳猛,王怀远,吴志军,等. 云计算中DDoS攻防技术研究综述 [J]. 计算机学报, 2020, 43(12): 2315-2336.
- [15] 柯一川. 浅谈TCP/IP协议与DoS/DDoS攻击方式 [J]. 网络安全技术与应用, 2024(2): 9-11.
- [16] 杨智乐. 僵尸网络检测研究[D]. 西安:西安电子科技大学, 2022.
- [17] 郝帅,白翼铭,李致成,等. 反射放大型DDoS攻击的预防策略研究 [J]. 信息技术与网络安全, 2021, 40(2): 7-13, 23.
- [18] 绿盟科技. 2020 DDoS攻击态势报告 [EB/OL]. (2021-01-21) [2024-04-10]. https://bestpractice.nsfocus.com.cn/html/2021/92_0121/148.html.
- [19] BOIN C, GUILLAUME X, GRIMAUD G, et al. One year of DDoS attacks against a cloud provider: an overview [C]//Proceedings of the 2022 4th International Conference on Advances in Computer Technology, Information Science and Communications (CTISC). IEEE, 2022: 1-5.
- [20] KANSAL V, DAVE M. Proactive DDoS attack mitigation in cloud-fog environment using moving target defense [EB/OL]. arXiv: 2012.01964, 2020.
- [21] BURGESS J. Modern DDoS attacks and defences: survey [EB/OL]. arXiv: 2211.15404, 2022.
- [22] ELSAYED M S, LE-KHAC N A, DEV S, et al. Ddosnet: a deep-learning model for detecting network attacks [C]//Proceedings of the 2020 IEEE 21st International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM). IEEE, 2020: 391-396.
- [23] KOPP D, DIETZEL C, HOHLFELD O. DDoS never dies? an IXP perspective on DDoS amplification attacks [C]//Proceedings of the International Conference on Passive and Active Network Measurement. Cham: Springer International Publishing, 2021: 284-301.
- [24] NAWROCKI M, KRISTOFF J, HIESGEN R, et al. SoK: a data-driven view on methods to detect reflective amplification DDoS attacks using honeypots

- [C] //Proceedings of the 2023 IEEE 8th European Symposium on Security and Privacy (EuroS&P). IEEE,2023:576-591.
- [25] GUPTA B B, JOSHI R C, MISRA M. An efficient analytical solution to thwart DDoS attacks in public domain [C] //Proceedings of the International Conference on Advances in Computing, Communication and Control,2009:503-509.
- [26] DANG Q V. Outlier detection on network flow analysis[EB/OL]. arXiv:1808.02024,2018.
- [27] WEI Y, JANG-JACCARD J, SINGH A, et al. Classification and explanation of distributed denial-of-service (DDoS) attack detection using machine learning and shapley additive explanation (SHAP) methods [EB/OL]. arXiv:2306.17190,2023.
- [28] NAVAZ A S S, SANGEETHA V, PRABHADEVI C. Entropy based anomaly detection system to prevent DDoS attacks in cloud [J]. International Journal of Computer Applications,2013,62(15):42-47.
- [29] CHEN J, TANG X, CHENG J, et al. DDoS attack detection method based on network abnormal behaviour in big data environment [J]. International Journal of Computational Science and Engineering, 2020, 23(1):22-30.
- [30] GUO H, FAN X, CAO A, et al. Peek inside the closed world: evaluating autoencoder-based detection of DDoS to cloud [EB/OL]. arXiv: 1912.05590, 2019.
- [31] HAMARSHE A, ASHQAR H I, HAMARSHEH M. Detection of DDoS attacks in software defined networking using machine learning models [C] //Proceedings of the International Conference on Advances in Computing Research. Cham: Springer Nature Switzerland,2023:640-651.
- [32] GOLDUZIAN A. Predict and prevent DDOS attacks using machine learning and statistical algorithms [EB/OL]. arXiv:2308.15674,2023.
- [33] DORIGUZZI-CORIN R, MILLAR S, SCOTT-HAYWARD S, et al. LUCID: a practical, lightweight deep learning solution for DDoS attack detection [J]. IEEE Transactions on Network and Service Management,2020,17(2):876-889.
- [34] HUSSAIN F, ABBAS S G, HUSNAIN M, et al. IoT DoS and DDoS attack detection using ResNet [C] //Proceedings of the 2020 IEEE 23rd International Multitopic Conference (INMIC). IEEE,2020:1-6.
- [35] ALFATEMI A, RAHOUTI M, AMIN R, et al. Advancing DDoS attack detection: a synergistic approach using deep residual neural networks and synthetic oversampling [EB/OL]. arXiv: 2401.03116, 2024.
- [36] FICKE E, SCHWEITZER K M, BATEMAN R M, et al. Analyzing root causes of intrusion detection false-negatives: Methodology and case study [C] //Proceedings of the MILCOM 2019-2019 IEEE Military Communications Conference (MILCOM). IEEE,2019:1-6.
- [37] ISHAQ K, JAVED H A. Implementing snort intrusion prevention system (IPS) for network forensic analysis [EB/OL]. arXiv:2308.13589,2023.
- [38] TAMAYO J, LÓPEZ L I B, CARAGUAY Á L V. Detection of distributed denial of service attacks carried out by botnets in software-defined networks [EB/OL]. arXiv:2401.09358,2024.
- [39] ALSAKRAN F, BENDIAB G, SHIAELES S, et al. Intrusion detection systems for smart home IoT devices: experimental comparison study [C] //Proceedings of the International Symposium on Security in Computing and Communication. Singapore: Springer Singapore,2019:87-98.
- [40] LIU Z. DeepTLS: comprehensive and high-performance feature extraction for encrypted traffic [EB/OL]. arXiv:2208.03862,2022.
- [41] ERGENÇ D, SCHENDERLEIN R, FISCHER M. Tsnzeek: an open-source intrusion detection system for IEEE 802.1 time-sensitive networking [C] //Proceedings of the 2023 IFIP Networking Conference (IFIP Networking). IEEE,2023:1-6.
- [42] NISENOFF A, SHARMA R, FEAMSTER N. Understanding user awareness and behaviors concerning encrypted DNS settings [EB/OL]. arXiv: 2208.04991,2022.
- [43] ZIRNGIBL J, SATTLER P, CARLE G. A first look at SVCB and HTTPS DNS resource records in the wild [C] //Proceedings of the 2023 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW). IEEE,2023:470-474.

文章编号: 1007-1423(2024)20-0057-06

DOI: 10.3969/j.issn.1007-1423.2024.20.011

基于 RFC 优化算法的报文数据快速模式匹配

王瑞钦¹, 谭 皇^{2*}

(1. 北京交通大学经济管理学院, 北京 100044; 2. 合肥泰能路通科技有限公司, 合肥 230000)

摘要: 为提高报文数据模式匹配速度, 对 RFC 算法进行优化。该优化算法通过选择优良的哈希算法对规则 CBM 比特位图进行去重匹配, 从而降低预处理过程的时间复杂度; 在“缩减”的最后阶段, 将等价类规则 CBM 比特位图改为索引数组, 使得数据平面的报文数据匹配查找时间复杂度由 $O(n^2)$ 降低到 $O(1)$, 进一步减少了对空间资源的消耗。实验结果表明, 该 RFC 优化算法有效降低了时间和空间复杂度, 达到 100 Gbps 的处理性能水平, 可以应用到各种流量攻击防护场景。

关键词: RFC 算法; CBM 位图; 哈希去重; 索引数组; 模式匹配

0 引言

随着 Internet 的快速发展, 以及网络链路速度的不断提升, 作为网络节点的流量攻击防护设备必须提高报文数据模式匹配的速度以避免影响正常的网络服务。快速报文数据模式匹配已成为流量攻击防护的一项关键技术。

报文数据模式匹配是指将报文的若干字段数据与预先配置的若干条规则进行比对的过程, 通常, 模式匹配规则包含源 IP 地址、目的 IP 地址、源端口、目的端口及协议类型等五元组固定字段, 以及非固定长度的若干个七层数据字段。因此, 报文数据快速模式匹配可以转化为报文分类问题^[1-5]。

报文分类问题可以用数学语言来描述^[6], 假定规则集包含 N 条规则, 每条规则 d 个字段域, 且规则的第 i 个域和数据包的第 i 个域对应。对任意的一个数据报文, 如果存在一条规则 R_k 对任意的字段 F_i 均有报文数据 P_i 与之匹配, 则数据报文与该规则匹配。

具有典型代表性的报文分类算法^[7-11]包括线性查找、Grid of Tries、Ternary CAM(三元内容寻找存储)、RFC(Recursive Flow Classification)、

HiCuts、Cross Producting 和 BV^[12]等。表 1 描述了部分报文分类算法的时间复杂度和空间复杂度, 其中 n 为规则数量, w 为待匹配报文数据的宽度(单位为比特), d 为每条规则的字段数量。可以看到, 基于高速硬件支持的 TCAM 方案时间复杂度最低, 但支持的规则字段维数少而应用场景有限。相比之下, RFC 算法能满足大多数应用场景空间复杂度要求, 且具有最低的时间复杂度, 非常适合应用于流量攻击防护设备的报文数据快速模式匹配场景。

表 1 不同类型报文分类算法的时间和空间复杂度对比表

报文分类算法名称	时间复杂度	空间复杂度
线性查找	n	n
Grid-of-Tries	dw	ndw
TCAM	1	n
RFC	d	n^d
HiCuts	d	n^d
Cross Producting	dw	n^d
Bit Vector	\log^n	n^2

针对 RFC 交叉积乘可能带来的“内存爆炸”问题, 文献[13]提出一种基于改变缩减树存储结

收稿日期: 2024-05-10 修稿日期: 2024-07-18

作者简介: 王瑞钦(2003—), 女, 重庆垫江人, 在读本科, 研究方向为信息管理与信息系统; *通信作者: 谭皇(1984—), 男, 四川成都人, 高级软件工程师, 研究方向为网络安全, E-mail: scu_hank@126.com

构的改进算法以去除交叉乘积表中的重复元素，在一定程度上实现了内存空间的压缩，但带来了配置平面设计复杂、预处理时间长，以及数据平面匹配性能降低等问题；文献[14]提出基于或运算压缩块比特位图向量表的Merge_RFC算法，解决了交叉乘积表中规则位图稀疏分布带来的冗余空间浪费，同样降低了配置平面和数据平面的处理性能；文献[15-19]本质是通过减少规则字段宽度来降低交叉乘积表的空间复杂度，同样是以降低数据平面处理性能为代价；针对预处理过程中稀疏规则位图匹配带来的低效匹配，文献[20]将规则位图按16个比特位为单位划分成若干个字，如果为0则直接跳过16个比特位规则位图的匹配过程，在一定程度上减少了配置平面的预处理时间；文献[21]针对具有非固定长度规则字段的应用场景，提出了一种RFC与Aho-Corasick相结合的报文数据模式匹配方案。

随着硬件成本不断降低，如何从软件层面降低时间复杂度、最大程度地同时兼顾配置平面和数据平面的处理性能变得愈加重要。现有的RFC改进算法大都将重点集中到如何降低空间复杂度。因此，本文提出一种RFC优化算法，实现配置平面和数据平面的处理性能同时提升，且能降低算法的部分空间复杂度。

1 RFC算法描述

RFC(Recursive Flow Classification, 递归流包分类)算法^[22-28]利用实际应用中规则的结构特性和冗余，将整个规则域分为若干个段，通过计算分段等价类和递归操作，实现空间的分步压缩映射。

假定报文数据分类规则有 d 维， S 表示报文中 d 个域形成的比特串数据总长度，报文分类的本质就是将报文数据中 S 个比特关键字信息映射成 $T(T = \log^d S, T \ll S)$ 比特的等价类标识(eqID)。如果通过一次内存访问就可获得匹配结果，则需要建立 2^S 个数到 T 个eqID之间的对应关系(至少需要 $2^S + 2^T$ 个内存空间来保存对应关系)，当 d 为源/目的IP地址、源/目的端口和协议域五元组时 $S = 104$ ，至少需要 2^{104} 比特空间。

RFC算法的基本思想是：将原来的一次映射转变为多级映射，将一个较大的集合映射成

若干个较小的集合；其中，每一阶段映射称为一次“缩减”，最终实现报文中 S 比特串数据到 T 位eqID的映射。

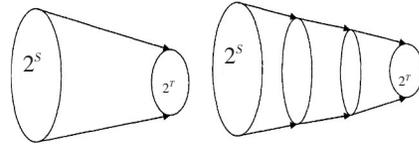


图1 RFC算法的基本思想示意图

RFC算法包括配置平面的预处理和数据平面的匹配查找，预处理过程中的每个阶段将一个较大的集合映射为一个较小的集合，这个过程称为一次“缩减”。整个过程可以分为 P 个阶段，每个阶段执行若干并行的查找过程，查找的结果用于计算下一阶段查找的索引值。

RFC算法的本质是在配置平面的预处理阶段建立一个“一一映射”的多级映射表，以保证数据平面的数据报文可以根据既定的数据字段按照“缩减”过程逐级搜索，最终索引到规则位图向量表。整个索引过程实际是根据索引值直接访问映射表数组，这也是RFC算法高性能的根本原因。

1.1 配置平面处理过程

预处理过程包含如下几个部分：

(1) 将规则表中的 n 元组规则按既定的宽度 w 划分为若干个chunk，如表2中的五元组规则将dip、sip、dp和sp等按16比特位宽划分为6个chunk，协议字段则单独为一个chunk。

表2 RFC报文数据模式匹配规则表

id	dip	sip	dp	sp	pro
0	125.73.166.3/32	125.73.166.0/25	*	4119	0x6/0xff
1	116.58.124.0/22	135.206.0.0/17	80	21	0x33/0xff
2	180.34.0.0/16	186.44.67.0/24	8080	3443	0x11/0xff
3	135.206.149.64/26	187.187.187.187/32	8191	*	0x11/0xff
4	0.0.0.0/0	0.0.0.0/0	23	1835	0x0/0x0

(2) 将chunk投影到数轴上形成一个(chunk equivalence set, CES)等价集，建立一个以eqID为索引和(chunk bitmap, CBM)比特位图为值的等价类表；再计算每个chunk在不同值对应的eqID值，建立一个以eqID为值的预处理索引向

量表。

(3) 根据步骤(2)计算得到的两个等价类表，对表中的每一项作交叉逻辑与运算得到交叉积向量表，且向量表的大小为两个等价类表的大小相乘，每一个向量存储eqID值，等价类表为去除重复后的CBM比特位图集合。

(4) 按步骤(3)进行递归操作，直到完成所有chunk的预处理过程。

以表2中的32位目的IP地址dip规则字段为例，将其以16比特为单位拆分为chunk0和chunk1(见表3)，并分别将chunk0和chunk1投影到数轴上形成两个等价集，结果见表4和表5。

对chunk0和chunk1作交叉逻辑“与”运算并对CBM作去重操作，得到如表6所示的交叉积向量表。

1.2 数据平面匹配查找

匹配查找实际是将报文数据按预处理规则既定的数据字段逐级“缩减”的过程，即：根据报文数据指定字段计算出chunkId，每两个chunkId作交叉积运算得到指向下一阶段向量表的索引值及eqID，递归该操作最终索引到满足匹配条件的CBM规则表位图。最后，按CBM规则表位图进行报文数据模式匹配。

表3 dip chunk投影到数组形成的映射关系表

id	dip		chunk0		chunk1	
0	125.73.166.30/32	125.73.166.30	125.73	32073	166.30	42526
1	116.58.124.0/22	116.58.124.0~116.58.127.255	116.58	29754	124.0~127.255	31744~32767
2	180.34.0.0/16	180.34.0.0~180.34.255.255	180.34	46214	0.0~255.255	0~65535
3	135.206.149.64/26	135.206.149.64~135.206.149.127	135.206	34766	149.64~149.127	38208~38271
4	0.0.0.0/0	0.0.0.0~255.255.255.255	0.0~255.255	0~65535	0.0~255.255	0~65535

表4 chunk0与eqID的映射关系表

chunkId	eqID0	CES	CBM
0~29753			
29755~32072			
32074~34765	0	其他	10000
34767~46213			
46215~65535			
29754	1	29754	10010
32073	2	32073	10001
34766	3	34766	11000
42614	4	42614	10100

表5 chunk1与eqID的映射关系表

chunkId	eqID1	CES	CBM
0~31743			
32768~38207			
38272~42525	0	其他	10100
42527~65535			
31744~32767	1	31744~32767	11010
32808~38271	2	32808~38271	11100
42526	3	42526	10101

表 6 chunk0和chunk1交叉逻辑积向量表

chunk0*chunk1	CES	eqID	CBM
0	其他, 其他	0	10000
1	其他, 31744~32767	0	10000
2	其他, 32808~38271	0	10000
3	其他, 42526	0	10000
4	29754, 其他	0	10000
5	29754, 31744~32767	1	10010
6	29754, 32808~38271	0	10000
7	29754, 42526	0	10000
8	32073, 其他	0	10000
9	32073, 31744~32767	0	10000
10	32073, 32808~38271	0	10000
11	32073, 42526	2	10001
12	34766, 其他	0	10000
13	34766, 31744~32767	3	11000
14	34766, 32808~38271	3	11000
15	34766, 42526	0	10000
16	42614, 其他	4	10100
17	42614, 31744~32767	0	10000
18	42614, 32808~38271	4	10100
19	42614, 42526	4	10100

2 RFC算法优化

2.1 问题提出

在RFC算法配置平面预处理阶段,对两个 chunk 作交叉积运算时,需要对CBM位图表进行去重以得到等价类表(总共包含 2^w 项)。可以看到,CBM位图的长度为模式匹配规则表的数量,当规则表数量较大时,如何实现CBM位图快速匹配将直接影响到配置平面的预处理时间。因此,本文采用哈希算法实现CBM匹配去重,可以较大程度地减少预处理时间。

在RFC算法匹配查找阶段,当最终索引到CBM规则表位图时,需要逐个匹配比特为1对应的模式匹配规则,直到规则匹配成功为止。通常情况下,这个CBM规则表位图是稀疏的,当规则表数量较大时,按比特位顺序匹配将会导致处理性能严重下降,也会带来存储空间的消耗。

因此,本文在配置平面预处理阶段对稀疏的CBM规则位图进行优化处理,用索引值数组代替CBM位图,可以有效降低匹配查找的时间复杂度和空间复杂度。

2.2 算法优化

在计算两个 chunk 交叉积过程中,对两个

chunk各自对应的CBM比特位图执行逻辑“与”运算,并将其加入到等价类表,等价类表中存储了不重复的CBM位图及eqID信息。因此,加入到等价类表前需要检查是否已存在相同的CBM项。

在构造等价类表的过程中,创建一个临时的、以CBM哈希值为关键字的哈希映射表。当通过逻辑与运算得到 chunk 交叉积时,计算CBM比特位图的哈希值,通过逐个与哈希节点比较CBM位图,以确定是否已存在相同的CBM项,如果不存在则新增CBM位图及eqID信息到等价类表。完成等价类表的构造后,销毁临时创建的哈希表。

RFC算法“缩减”过程的中间阶段,等价类表记录的是 chunkId 与 eqID 的映射关系,而“缩减”过程最后阶段则记录了 eqID 与 CBM 的映射关系表。用一个数组来记录CBM规则比特位图为1的规则所对应的索引值,不仅避免了稀疏CBM位图带来的空间浪费,更是将数据平面的报文数据匹配查找的时间复杂度由 $O(n^2)$ 降低到 $O(1)$ 。

如图2所示,当规则表大小为10000时,某个eqID对应一个稀疏的CBM比特位图(对应第2条规则和第10000条规则),将其转化为索引值数组后,存储空间由157*8字节缩减到4字节,匹配查找由157次匹配操作降低到2次,同时降低了对空间和时间复杂度的要求。

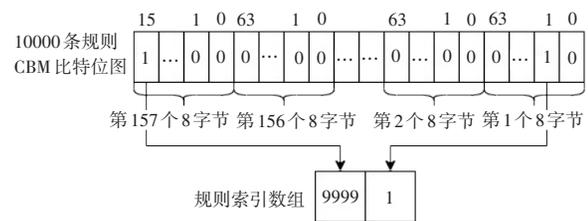


图 2 CBM比特位图转化为规则索引数组示意图

3 仿真实验

通过流量攻击生成工具随机生成多种攻击类型(如TCP/UDP/ICMP Flood等)的实时攻击报文,随机生成10000条报文数据模式匹配规则,对比不同情况下算法的时间和空间复杂度。

如表7所示,不同哈希算法对RFC配置平面的预处理过程时间性能影响不一。可以看到,

优良的哈希算法带来的哈希冲突数量相对较小,对整个预处理的时间要求更低。

表7 不同哈希算法对比数据表

	算法1	算法2	算法3	算法4
预处理时间/s	137	134	26	21
内存/KB	339935	339935	115151	127698
最大内存/KB	2398849	2398849	705450	698329
哈希冲突链平均长度	13	11	8	2
最大哈希链长度	3598	3179	1879	313

4 结语

本文提出了一种RFC优化算法,在配置平面预处理阶段生成等价类图过程中通过哈希算法对规则CBM比特位图进行去重匹配,有效降低了预处理过程的时间复杂度;另一方面,将“缩减”最后阶段的等价类规则CBM比特位图改为索引数组,不仅降低了算法的空间复杂度,更使得数据平面的报文数据匹配查找时间复杂度由 $O(n^2)$ 降低到 $O(1)$ 。实验结果表明,RFC优化算法有效地降低了时间和空间复杂度,已经达到100 Gbps的数据平面处理性能水平,可以应用到流量攻击防护的各种应用场景。

参考文献:

- [1] GUPTA P, MCKEOWN N. Algorithms for packet classification[J]. IEEE Network, 2001, 15(2):24-32.
- [2] GUPTA P, MCKEOWN N. Packet classification on multiple fields[J]. Proc. ACM SIGCOMM, 1999, 29(4):147-160.
- [3] SPITZNAGEL E W. Compressed data structures for recursive flow classification[EB/OL], 2003. <http://cse.seas.wustl.edu/techport-fields/getresport.asp?305>.
- [4] BABOESCU F, VARGHESE G. Scalable packet classification[J]. IEEE/ACM Transactions on Networking, 2005, 13(1):2-14.
- [5] QI Y X, XU L H, YANG B H, et al. Packet classification algorithms: from theory to practice[C]//Proc. of the IEEE INFOCOM. IEEE Press, 2009:648-656.
- [6] 元亚焯, 李军. 高性能网包分类理论与算法综述[J]. 计算机学报, 2013, 36(2):408-421.
- [7] 元亚焯. 多域网包分类算法研究[D]. 北京:清华大学, 2011.
- [8] 黄腾. 高性能报文分类算法的研究与实现[D]. 长沙:国防科学技术大学, 2016.
- [9] 柳孟跃. 基于FPGA的包分类算法的设计及实现[D]. 北京交通大学, 2015.
- [10] 吴层. 基于IXP2400的改进RFC算法的研究与设计[D]. 北京:北京交通大学, 2008.
- [11] 李学峰. IP报文分类算法研究[J]. 襄樊学院学报, 2010, 31(8):38-41.
- [12] 杨保华, 元亚焯, 薛一波. Bitmap结构在高性能网络算法设计中的应用[J]. 计算机工程与应用, 2009, 45(15):1-5.
- [13] 刘铎, 华蓓, 唐锡南, 等. Compact RFC:一种内存优化的RFC包分类算法[J]. 小型微型计算机系统, 2007, 28(3):482-487.
- [14] 曹婕, 陈兵. 一种内存优化的RFC包分类算法Merge_RFC[J]. 小型微型计算机系统, 2012, 33(4):865-868.
- [15] 刘胤, 杨世平. 基于RFC算法的快速多维数据包分类算法[J]. 计算机工程, 2008, 3(6):95-97.
- [16] 朱胜琼. 一种基于递归流的快速多维包分类算法[J]. 微计算机信息, 2006, 22(4):191-193.
- [17] 余虎, 黄宇. 递归流分类算法研究与改进[J]. 计算机工程, 2015, 41(12):44-47.
- [18] 田立勤, 林闯. 报文分类技术的研究及其应用[J]. 计算机研究与发展, 2003, 40(6):765-776.
- [19] 蔡艳婧, 孔苏鹏, 等. 基于多层迭代的递归数据流匹配改进算法[J]. 计算机应用与软件, 2021, 38(5):256-262.
- [20] 申震生, 龚向阳, 王文东, 等. 具有高速过滤算法的IP防火墙[J]. 计算机应用, 2001, 21(5):50-52.
- [21] 田珂, 朱清新, 向培素. 一种改进的多维高速报文分类算法[J]. 计算机应用研究, 2007, 24(2):27-29.
- [22] 杜静. 基于规则匹配的高效报文分类算法研究[D]. 银川:宁夏大学, 2018.
- [23] 徐晶. 基于IPv6的ACL系统设计及其报文匹配算法的优化[D]. 杭州:杭州电子科技大学, 2011.
- [24] 陈福明. 快速报文分类算法的研究与实现[D]. 北京:中国地质大学, 2006.
- [25] 孙毅, 刘彤, 蔡一兵, 等. 报文分类算法研究[J]. 计算机应用研究, 2007, 24(4):5-11.
- [26] 王艳秋, 赵昭灵, 兰巨龙, 等. 基于范围映射和定值映射的多域报文分类算法[J]. 四川大学学报(工程科学版), 2007, 52(6):52-57.
- [27] 万云凯, 嵩天, 刘苗苗, 等. 流量自适应的多维度包分类方法研究[J]. 计算机学报, 2017, 40(7):1543-1555.
- [28] TURNER D E T J S. ClassBench: a packet classification benchmark [J]. IEEE/ACM Transactions on Networking, 2007, 15(3):499-511.

Fast packet pattern-match based on optimized RFC algorithm

Wang Ruiqin¹, Tan Huang^{2*}

(1. College of Economics and Management, Beijing Jiaotong University, Beijing 100044, China;

2. Hefei Tynon Technology Corporation, Ltd., Hefei 230000, China)

Abstract: An optimized RFC algorithm is proposed to improve the speed of fast packet pattern match. This algorithm reduces the time complexity of preprocessing by using an excellent hash algorithm to deduplicate the CBM. In the final reduction stage, this algorithm also replaces the CBM with an index array, which reduces the complexity of packet matching time in the data plane from $\mathcal{O}(n^2)$ to $\mathcal{O}(1)$, and the consumption of memory space resource. The practical results show that the optimized RFC algorithm effectively reduces the time and memory space complexity and reaches the processing performance of 100 Gbps, which can be applied to various network traffic attack protection scenarios.

Keywords: RFC; CBM; Hash deduplication; index array; pattern match

(上接第 56 页)

A review of the development and detection technologies of DDoS attacks

Ma Lixin, Xue Zhanshuang, Liu Haiyan*

(School of Computer, North China Institute of Aerospace Engineering, Langfang 065000, China)

Abstract: In the field of network security, DDoS (Distributed Denial of Service) attacks are a major concern due to their severe threat to information systems. This study examines the evolution of DDoS attacks, from the flood attacks of the 1990s to the complex attack patterns utilizing botnets, Internet of Things (IoT) devices, and cloud services, highlighting the diversification of attack methods and techniques. It also provides a detailed classification of DDoS attacks, including methods based on attack vectors and traffic, as well as detection techniques based on traffic analysis, behavior analysis, machine learning, and deep learning, pointing out the advantages and limitations of these techniques. Furthermore, it introduces current mature DDoS detection tools and systems, such as Suricata, Snort, Bro(Zeek), and protection services like Cloudflare, and discusses future trends in DDoS attacks.

Keywords: cybersecurity; distributed denial of service; attack detection; deep learning

文章编号: 1007-1423(2024)20-0063-05

DOI: 10.3969/j.issn.1007-1423.2024.20.012

通过集成域对抗训练和最大平均差异的故障诊断

张宇^{1,2}, 张公政^{2,3*}

(1. 河钢数字技术股份有限公司, 石家庄 050035; 2. 河钢数字技术股份有限公司深圳分公司, 深圳 518000;
3. 雄安威赛博智能科技有限公司, 雄安 070001)

摘要: 目前, 工业物联网已成功应用于智能制造。物联网中的大量数据促进了基于深度学习的工业设备健康监测的发展。由于在不同工作条件或设备上收集的机械故障诊断监测数据存在域不匹配, 用训练数据训练的模型在实际应用中可能无法发挥作用。因此, 研究具有域自适应能力的故障诊断方法至关重要。提出了一种基于改进的域自适应方法的智能故障诊断方法。具体来说, 分别使用最大平均差异和域对抗训练关于特征空间距离和域不匹配的两个特征提取器来增强特征表示。由于单独的分类器训练特征提取器, 进一步利用集成学习获得最终结果。实验结果表明, 该方法在解决域不匹配故障诊断中是有效的, 且具有实际应用价值。

关键词: 域自适应; 域对抗性训练(DAT); 集成学习; 故障诊断; 最大均值差异(MMD)

0 引言

现在工业物联网已成功应用于智能制造。针对工厂车间的工业设备产生海量的数据已经进行了大量的研究, 这对促进数据驱动的设备健康管理^[1](PHM)系统的发展具有重要意义。故障诊断作为 PHM 的重要组成部分, 对保证生产安全至关重要。深度学习算法因能够解决复杂的问题, 产生高精度的结果, 已被广泛应用于机械故障诊断^[2]。

大多数深度学习算法都需要大量的训练数据, 并假设训练数据集和待测试的数据与相同的分布相关。但是, 由于现场工业环境的复杂性、工作条件(如机械负荷和速度)或待测试的设备可能会发生变化。因此, 训练良好的故障诊断模型在实际应用中可能具有较低的准确性。在此背景下, 先进的深度学习算法, 如迁移学习^[3]和域自适应^[4]已经出现并且发展比较成熟。文献[5-8]引入的方法同时关注源数据和目标数

据, 通过使用相应的损失函数, 如最大均值差异(MMD)^[9]和关联对齐距离(CORAL)^[8], 来学习这两个域的特征提取网络。该方法利用具有共享权值的体系结构来学习源域和目标域的不变特征, 并使用带注释的训练数据集来训练分类器。训练良好的特征提取器和分类器将用于预测目标域故障数据的标签。

上述基于域自适应的方法考虑了不变域特征, 但如果源域和目标域差异较大, 将数据投影到同一特征空间的方法会失去一些重要的特征表示。为此, 本文提出了一种基于改进的域自适应体系结构的智能故障诊断方法。具体来说, 使用两个特征提取器将数据投影到不同的特征空间中。一个特征提取器基于域对抗性学习来学习特征, 而另一个则通过考虑 MMD 作为损失函数来学习。在一个广泛应用的故障诊断基准上验证了该方法的有效性。结果表明, 该方法适用于不同工况下的故障诊断, 且采用人工数据训练和真实故障数据测试时的性能良好,

收稿日期: 2024-03-14 修稿日期: 2024-07-09

基金项目: 2023 年工信部工业互联网创新发展工程-基于工业互联网平台的重点行业生产设备智能运维系统项目(ZTZB-23-990-020); 河北省省级科技计划资助-重大科技成果转化专项(23280101Z)

作者简介: 张宇(1990—), 男, 广东深圳人, 博士, 研究方向为材料科学与工程; *通信作者: 张公政(1995—), 男, 广东深圳人, 硕士, 研究方向为信息与计算科学, E-mail: gzzhang_13@163.com

该方法优于现有的计算方法。

将先进的深度学习方法，如迁移学习和领域自适应集成到物联网应用程序中，可以提高性能。Zhang等^[10]将神经网络应用于基于迁移学习的故障诊断。他们使用大量的源数据来训练一个神经网络，之后使用少量的目标域数据共享神经网络的参数并且重新训练模型。但是，当所标记的目标数据不可用时，这种方法就失效了。因此，提出了基于源域标记训练数据和目标域无标记数据的无监督域自适应方法。考虑到这两个领域之间的特征空间差异，这些方法(如MMD、CORAL、EMD)通常被嵌入到损失函数中来提取域不变特征。在文献[7]和文献[11]中，MMD与迁移学习集成，训练具有领域适应和特征差异较小的模型。当样本数量较大时，MMD的计算复杂度较高，而在文献[5]中则采用了带有域自适应的EMD进行故障诊断。为了进一步学习不变特征，Guo等^[6]提出了一种通过同时最大化域分类误差和MMD距离来训练特征提取器的方法。

在文献[5, 7, 11]中提出的方法旨在最小化源域和目标域特征之间的差异，但它们可能会失去目标域的区别特征。为此，在文献[8]中提出的方法通过减少分布差异和保留鉴别信息来进行跨域故障诊断。特别是考虑到缺乏目标域的数据，提出的方法通过人工生成假样本进行域自适应，提供了可靠的跨域诊断方案。

1 相关工作

本文与文献[6, 7, 11]中的相关方法的区别主要是：①使用两个特征提取器，分别使用MMD和DAT来学习不变特征，它们比文献中的方法提供了更多的不变特征；②本文的工作对每个特征提取器使用一个分类器，这保证了两个特征提取器的参数学习过程的独立性；③本文的工作集成了使用MMD和DAT学习的不变特征，并集成了两个分类器用于故障诊断。

2 准备工作

2.1 问题公式化

本文的工作是根据以下假设进行的。

(1) 对于不同的域的故障诊断任务是相同的，即源域数据和目标域数据的标签被共享。

(2) 源域和目标域相互关联，但由于操作条件不同，分布不同。

(3) 来自源域的标记数据和来自目标域的未标记数据可用于训练。

(4) 在测试期间，可以获得来自目标域的未标记数据。

设 X 为输入空间， $Y = \{1, 2, \dots, N\}$ 为 N 个健康条件的集合。来自源域的 n_s 个标记样本为 $D_s = \{(x_i^s, y_i^s)\}_{i=1}^{n_s}$ ，其中 $D_t = \{(x_i^t)\}_{i=1}^{n_t}$ 表示来自目标域的未标记样本。 D_s 和 D_t 分别从联合分布 $P(X, Y)$ 和 $Q(X, Y)$ ($P \neq Q$)中采样。本文旨在设计一个深度神经网络 $y = f(x)$ ，它可以减少联合分布的跨域转移，学习域不变特征和分类器，并在源域监督下最小化目标风险 $R_t(f) = Pr_{(x, y) \sim Q}[f(x) \neq y]$ 。

2.2 领域对抗性训练

DAT^[12]在域适应过程中起着重要的作用。DAT的框架如图1所示。

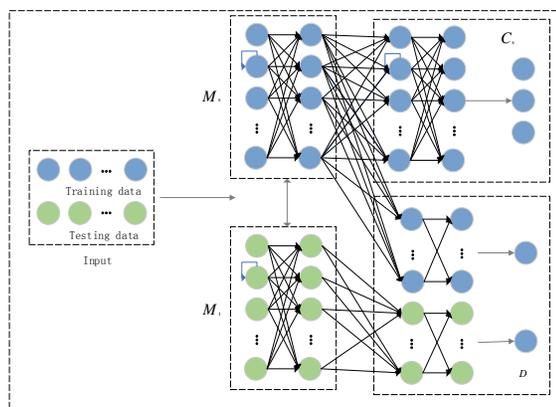


图1 DAT框架

设训练数据为 X_s ，训练标签为 Y_s ，测试数据为 X_t ，训练数据的特征提取器和测试数据分别为 M_s 和 M_t ，分类网络为 C_s ，域预测网络为 D 。在训练过程中，训练数据和测试数据首先输入特征提取器网络。然后，将训练数据的特征 $M_s(X_s)$ 作为分类网络的输入。利用反向传播算法训练 M_s 和 C_s 参数的分类，使得损失最小。

$$loss_{M_s, C_s}(X_s, Y_s) = - \sum_{x_i \in X_s, y_i \in Y_s} \sum_{k=1}^N t_{k, y_i} \cdot \log(C_s(M_s(x_i))) \quad (1)$$

式中： N 为类别的数量，且如果 $k = y_s$ ，那么 $t_{k,y_s} = 1$ ； $C_s(M_s(x_s))$ 表示正确的分类概率。然后，利用 M_s 的参数来初始化 M_t 。

域预测网络 D 和特征提取器 M_s 和 M_t 基于特征空间对齐训练。首先，域预测网络 D 的目的是区分目标域和源域的特征，并将 $M_s(X_s)$ 和 $M_t(X_t)$ 分为两类。让源域数据的特征标签为1，而目标域数据的特征标签为0。将域预测网络的损失函数定义为

$$\begin{aligned} \text{loss}_{D, M_t, M_s}(X_s, X_t) = & - \sum_{x_i \in X_s} \log(D(M_s(x_i))) - \\ & \sum_{x_i \in X_t} \log(1 - D(M_t(x_i))) \end{aligned} \quad (2)$$

在对抗性训练中， M_t 被认为是一个使用目标域数据作为输入的生成器。希望由 $M_t(X_t)$ 生成的数据可以欺骗域预测网络。也就是说，域预测网络将 $M_t(X_t)$ 作为标签为1的源域的特征。因此， M_t 通过最小化以下损失函数来训练。

$$\text{loss}_{D, M_t}(x_t) = - \sum_{x_i \in X_t} \log(D(M_t(x_i))) \quad (3)$$

基于DAT训练域自适应网络后，可以使用特征提取器 M_t 和分类网络 C_s 对测试数据进行预测，如图2所示。



图2 测试数据预测流程

在DAT中，源域数据和目标域数据被投影到同一个子空间中。特征提取器输出两个域的不变特征。在大多数情况下，DAT对于获得自适应模型是有效的和适用的。但是，如果这两个域之间存在较大的差异，那么当只考虑一个特征模型时，可能会失去一些具有代表性的特征。在此，本文引入了一种有效的结构来建模源数据和目标数据之间的不同特征表示，并将其集成到基于集成学习的领域自适应中。

3 提出的方法

3.1 数据过程

数据预处理涉及到数据分区和数据规范化。序列数据可在48 kHz或64 kHz范围下采集。为了获得适合深度学习的数据，应将原始数

据划分为相同大小 w 的序列。数据划分后，设训练数据集为 $X_s = \{x_{s,1}^T, x_{s,2}^T, \dots, x_{s,n_1}^T\}$ ，测试数据为 $X_t = \{x_{t,1}^T, x_{t,2}^T, \dots, x_{t,n_2}^T\}$ ，其中 n_1 和 n_2 表示 X_s 和 X_t 中的样本数。数据规范化实现如下：

$$\bar{x}_{s,i} = \frac{x_{s,i} - \text{mean}(x_{s,i})}{\text{std}(x_{s,i})}, (i = 1, 2, \dots, n_1) \quad (4)$$

$$\bar{x}_{t,j} = \frac{x_{t,j} - \text{mean}(x_{t,j})}{\text{std}(x_{t,j})}, (j = 1, 2, \dots, n_2) \quad (5)$$

式中： $\text{mean}(x_{s,i})$ 为 $x_{s,i}$ 的均值； $\text{std}(x_{s,i})$ 表示 $x_{s,i}$ 的标准差。

3.2 基于DAT的训练特征表示

基于特征提取器的DAT将一个特征提取器（特征提取器I）、一个分类器（分类网络I）和一个域预测网络合并训练。训练数据集 \bar{X}_s ，训练标签 Y_s ，测试数据集 \bar{X}_t ，特性提取器 M_t ，分类网络 C_t 和域预测网络 D 。 $M_t(\bar{X}_s)$ 和 $M_t(\bar{X}_t)$ 是 \bar{X}_s 和 \bar{X}_t 的特征表示。为了训练域预测网络，将源域数据 \bar{X}_s 标记为1，目标域数据 \bar{X}_t 标记为0，将域预测网络的损失函数定义为

$$\begin{aligned} \text{loss}_{D, M_t}(x_i) = & - \frac{1}{n} \sum_{j=1}^n [y_j \log(D(M_t(x_i))) + \\ & (1 - y_j) \log(1 - D(M_t(x_i)))] \end{aligned} \quad (6)$$

式中： n 为一个批次中的样本数； y_j 为域标签；源数据为0，目标数据为1； x_i 表示来自源域或目标域的数据。

采用对抗性训练来训练 M_t 和 D 的最优参数。 M_t 是通过最小化以下损失来进行训练的函数：

$$\text{loss}_{D, M_t}(\bar{X}_t) = - \sum_{\bar{x}_i \in \bar{X}_t} \log(D(M_t(\bar{x}_i))) \quad (7)$$

然后，使用以下分类损失函数来训练 M_t 和 C_t 中的参数：

$$\begin{aligned} \text{loss}_{M_t, C_t}(\bar{X}_t, Y_s) = & \\ & - \frac{1}{n} \sum_{j=1}^n \sum_{k=1}^N t_{k,y_s,j} \times \log \frac{\exp[(w_{l,k})^T M_t(\bar{x}_{s,j}) + b_l]}{\sum_{l=1}^N \exp[(w_{l,l})^T M_t(\bar{x}_{s,j}) + b_l]} \end{aligned} \quad (8)$$

式中： N 为故障类的数量； $w_{l,k}$ 和 b_l 为分类器 C_t 中的参数；如果 $k = y_{s,j}$ ，那么 $t_{k,y_s,j} = 1$ ； $M_t(\bar{x}_{s,j})$ 表示源域的特征。训练结束后，使用特征提取器 M_t 和分类网络 C_t 对测试数据进行预测。

3.3 基于MMD的训练特征表示

其他的特征表示和分类器则基于MMD进行训练。设特征提取器 Π 为 M_{Π} ，分类器 Π 为 C_{Π} 。 \bar{X}_s 和 \bar{X}_t 的特征可以分别表示为 $M_{\Pi}(\bar{X}_s)$ 和 $M_{\Pi}(\bar{X}_t)$ 。在本文中，使用基于RBF内核的MMD，它可以写为

$$\begin{aligned} \text{loss}_{M_{\Pi}}(\bar{X}_s, \bar{X}_t) = & \sum_{i,i'} \frac{G(M_{\Pi}(\bar{x}_{s,i}), M_{\Pi}(\bar{x}_{s,i'}))}{n^2} - \\ & 2 \sum_{i,j} \frac{G(M_{\Pi}(\bar{x}_{s,i}), M_{\Pi}(\bar{x}_{t,j}))}{n^2} + \\ & \sum_{j,j'} \frac{G(M_{\Pi}(\bar{x}_{t,j}), M_{\Pi}(\bar{x}_{t,j'}))}{n^2} \end{aligned} \quad (9)$$

式中： $G(\cdot)$ 表示RBF核，可以表示为 $G(x_1, x_2) = \exp(-\|x_1 - x_2\|^2/\tau)$ 。由于 M_{Π} 中的参数可以被学习以适应带宽 τ ，所以 τ 可以被设置为任意常数。在实验中，这个参数被设置为1.0。用 M_{Π} 同时训练分类器 C_{Π} 。用于训练 C_{Π} 的损失函数被定义为

$$\begin{aligned} \text{loss}_{M_{\Pi}, C_{\Pi}}(\bar{X}_s, Y_s) = \\ -\frac{1}{n} \sum_{j=1}^n \sum_{k=1}^N t_{k,y_s,j} \times \log \frac{\exp[(w_{\Pi,k})^T M_{\Pi}(\bar{x}_{s,j}) + b_{\Pi}]}{\sum_{l=1}^N [\exp((w_{\Pi,l})^T M_{\Pi}(\bar{x}_{s,j}) + b_{\Pi})]} \end{aligned} \quad (10)$$

式中： $w_{\Pi,k}$ 和 b_{Π} 为分类器 C_{Π} 中的参数。

4 实验结果

4.1 数据集

本文使用石家庄钢铁厂采集的轴承数据来评估所提方法的性能。首先将无线温振动传感器安装在轴承上，为了获取准确的数据，我们在设备的不同测量位置安装无线传感器，通过传感器采集时域上的振动加速度信号、温度信号、转速信号、电机电流、电机功率、电机负载等参数，并在现场机房安装一台服务器用于通讯和保存数据，为了便于深度学习训练，振动加速度信号采用64 kHz的采样频率，然后采集十个轴承近三个月的数据，包括正常和异常数据，故障类型分别为内圈故障、外圈故障、润滑不良。将其打标签并保存。

4.2 评价指标

定量性能的测量使用了分类精度，定义如下：

$$\rho = \frac{c_w}{c} \quad (11)$$

式中： c_w 为预测正确的样本数； c 为所有测试样本数。

4.3 实验过程

由于工业场景下传感器设备会产生数据偏移，本文用归一化的方式将数据映射到固定范围内，振动加速度信号按照3.1小节中的方法做数据归一化处理，得到 \bar{X}_s ，分别将 \bar{X}_s 输入到DAT模型和MMD模型中训练，训练的批次为100次。测试过程中为了融合两个模型的预测结果，对两个分类器输出的结果取均值。假设 \bar{X}_s 的预测结果分别为 $z_{L,s}$ 和 $z_{M,s}$ 。对 \bar{X}_s 预测的输出 z_s 为

$$z_s = \frac{z_{L,s} + z_{M,s}}{2} \quad (12)$$

4.4 结果和分析

从表1可以观察到该方法比基于DAT和基于MMD的域自适应具有更高的故障诊断精度，证明了融合不同特征空间和两种分类器集成的有效性。另一方面，实验结果表明：该方法可以获得较高的故障诊断预测；如果源域和目标域差异较大，如人为诱导损伤作为源域，真实损伤作为目标域，用伪标记目标数据对网络进行再训练将大大提高精度。

表1 实验结果

模型	MMD	DAT	本模型
训练集	0.9705	0.9712	0.9964
测试集	0.9616	0.9634	0.9824

5 结语

为了构建适用于不同设备或不同工作条件下采集数据的有效故障诊断模型，本文提出了一种改进的域自适应网络，以最小化多域的差异特征，同时使用DAT学习代表性特征。据此，基于MMD和域对抗训练，训练了两个独立的特征提取器及其相应的分类网络。该方法进一步集成了集成学习的最终预测结果。对不同工况下采集的故障数据进行了两次实验，结果验证

了该方法的有效性。未来将继续深化这两项工作。其一，本文将优化跨域故障诊断方法，以适应更困难的情况，如在不同的设备上发生的故障；其二，由于错误的数据比健康数据要少得多，因此必须在实际应用中开发处理不平衡数据的方法。

参考文献：

- [1] 刘冰扬,赵爽. 故障预测与健康管理的钢铁企业的应用[J]. 河北冶金,2022(11):81-86.
- [2] 王春燕,张旺,高二庆. 数据监测技术在轧钢设备故障诊断中的应用[J]. 河北冶金,2018(12):61-63.
- [3] PAN S J, YANG Q. A survey on transfer learning [J]. IEEE Transactions on Knowledge and Data Engineering, 2010, 22(10):1345-1359.
- [4] HAL CHARLES D, MARCU D C. Domain adaptation for statistical classifiers[J]. Journal of Artificial Intelligence Research, 2006, 26(1):101-126.
- [5] CHENG C, ZHOU B, MA G, et al. Wasserstein distance based deep adversarial transfer learning for intelligent fault diagnosis [J]. Neurocomputing, 2020, 409:35-45.
- [6] GUO L, LEI Y, XING S, et al. Deep convolutional transfer learning network: a new method for intelligent fault diagnosis of machines with unlabeled data [J]. IEEE Transactions on Industrial Electronics, 2019, 66(9):7316-7325.
- [7] LI X, ZHANG W, DING Q, et al. Multi-layer domain adaptation method for rolling bearing fault diagnosis [J]. Signal Processing, 2019, 157 (APR.): 180-197.
- [8] WANG X, HE H, LI L. A hierarchical deep domain adaptation approach for fault diagnosis of power plant thermal system [J]. IEEE Transactions on Industrial Informatics, 2019, 15(9):5139-5148.
- [9] BORGWARDT K M, GRETTON A, RASCH M J, et al. Integrating structured biological data by Kernel Maximum Mean Discrepancy [J]. Bioinformatics, 2006, 22(14):e49-e57.
- [10] ZHANG R, TAO H, WU L, et al. Transfer learning with neural networks for bearing fault diagnosis in changing working conditions [J]. IEEE Access, 2017, 5:14347-14357.
- [11] WEN L, GAO L, LI X. A new deep transfer learning based on sparse auto-encoder for fault diagnosis [J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2017(99):1-9.
- [12] TZENG E, HOFFMAN J, SAENKO K, et al. Adversarial discriminative domain adaptation [EB/OL]. arXiv.1702.05464, 2017.

Fault diagnosis through integrated domain adversarial training and maximum average difference

Zhang Yu^{1,2}, Zhang Gongzheng^{2,3*}

- (1. Hegang Digital Technology Co., Ltd., Shijiazhuang 050035, China;
2. Hegang Digital Technology Co., Ltd., Shenzhen Branch, Shenzhen 518000, China;
3. Xiongan Saibo Intelligent Technology Co., Ltd., Xiongan 070001, China)

Abstract: Currently, the industrial Internet of Things has been successfully applied in the intelligent manufacturing industry. The large amount of data in the Internet of Things has promoted the development of industrial equipment health monitoring based on deep learning. Due to domain mismatch in the mechanical fault diagnosis monitoring data collected under different working conditions or equipment, models trained with training data may not be effective in practical applications. Therefore, it is crucial to study fault diagnosis methods with domain adaptive capabilities. An intelligent fault diagnosis method based on an improved domain adaptive method was proposed in this paper. Specifically, two feature extractors for feature space distance and domain mismatch are trained using maximum average difference and domain adversarial training to enhance feature representation. Due to the separate classifier training feature extractor, further ensemble learning is utilized to obtain the final result. The experimental results show that this method is effective and has practical value in domain mismatch fault diagnosis.

Keywords: domain adaptation; domain adversarial training (DAT); ensemble learning; fault diagnosis; maximum mean difference (MMD)

文章编号: 1007-1423(2024)20-0068-05

DOI: 10.3969/j.issn.1007-1423.2024.20.013

基于在线评测的主从式数据库信息快速同步方法研究

王位明*

(江西现代职业技术学院, 南昌 330095)

摘要: 传统的同步方法缺乏实时评估和分析机制, 同步效果不佳。提出基于在线评测的主从式数据库信息快速同步方法。主数据库负责处理写操作并维护最新的数据, 在改动时通知从数据库进行数据同步。引入在线测评工具, 对同步的数据进行实时评估。以数据处理加速比作为评估指标, 调整数据处理方式, 设计数据库信息快速同步流程, 通过反馈机制将同步结果返回给主数据库, 确保同步的效果和质量。实验结果表明, 该方法的 CPU 使用率在合理范围内, 具备更佳的同步效果, 累积误差分布概率较小, 表明在信息同步过程中能够保持较高的可靠性。

关键词: 在线评测; 主从式数据库; 数据同步; 快速同步方法

0 引言

数据库作为信息系统的核心组成部分, 稳定性、高效性和数据一致性成为了关键指标^[1], 通过实现读写分离、提高系统的整体性能。然而, 这种架构也带来了数据同步的问题, 即如何确保主从数据库间的数据保持一致性^[2]。在主从式数据库架构中, 数据同步的效率和准确性是确保系统性能和数据一致性的关键^[3]。如何快速、准确地实现主从数据库间的信息同步, 成为了当前研究的热点和难点。

杨波等^[4]提出基于多核 CPU 的多源异构数据库信息快速同步方法, 使用多核 CPU 并行执行任务, 优化信息同步速度, 从而实现多源异构数据库信息的快速同步。由于缺乏实时评估和分析机制来监控同步过程, 容易造成数据同步错误或同步效果不佳。黄刘松等^[5]提出基于网闸的主从式数据库安全同步方法, 利用数据变化探知技术监测数据库变化, 结合数据加密压缩和编码融合方式实现网闸环境下异构数据库同步。该方法缺乏可以根据实时评估结果调整数据处理方式的机制, 影响了同步的准确性

和效率。郭良银^[6]提出多关系型数据库向图数据库转换与同步方法, 定义数据转换和同步方法。该方法缺乏实时评估和分析机制来监测同步过程, 无法即时检测到同步中的问题或错误。张万里^[7]提出分布式网络数据同步方法研究, 介绍了分布式网络数据同步技术, 构建了数据同步模型, 由此设计数据同步流程。该方法在分布式网络中进行数据同步涉及到大量的网络通信操作。当数据同步节点发生冲突的时候, 缺乏实时评估和分析机制来确定最佳的冲突解决方案, 会导致数据同步效果不佳或数据一致性问题。

鉴于上述研究背景和现状, 本文提出一种基于在线评测的主从式数据库信息快速同步方法。

1 在线评测的主从式数据库信息快速同步方法设计

1.1 主从式数据库整体架构设计

主从式数据库架构又被称为主-备数据库架构, 通常用于提高系统的可用性、性能和容错能力。在整体架构设计中, 有一个主数据库,

收稿日期: 2024-03-15 修稿日期: 2024-06-05

基金项目: 2022 年度江西省教育厅科学技术研究项目(GJJ2205123)

作者简介: *通信作者: 王位明(1994—), 男, 江西赣州人, 硕士研究生, 助教, 研究方向为大数据技术、软件技术, E-mail: sonagi0910@163.com

负责处理所有的写操作和数据更新；有一个或多个从数据库，负责读取操作。主数据库会将其更改同步到从数据库，以保持数据的一致性^[8]。主从式数据库整体架构见表1。

表1 主从式数据库整体架构表

数据库	组件/角色	描述
主数据库(Master)	写入操作	负责接收和处理所有写操作
	二进制日志	记录所有更改数据库数据的操作，用于从数据库的同步
	I/O线程	负责将二进制日志事件发送给从数据库
	读取操作	主要负责处理读请求，提供数据查询服务
从数据库(Slave)	I/O线程	从主数据库读取二进制日志事件
	SQL线程	执行从主数据库接收到的二进制日志中的SQL语句，更新从数据库的数据
	中继日志	存储从主数据库接收到的二进制日志事件，供SQL线程执行
网络	同步连接	主从数据库之间的网络连接，用于数据传输和同步
监控与管理	在线评测工具	监测同步性能、数据一致性，评估同步方法的效果
	管理工具	提供对主从数据库架构的配置、监控和管理功能

1.2 基于在线评测工具实现实时数据评估与分析

在线评测作为一种高效的自动化评估工具，其核心在于通过自动化处理和分析数据对象，实现对代码质量和性能的评估。这一过程中，数据对象的分类和过滤起着至关重要的作用，能够确保在线评测只针对符合特定标准或条件的数据进行，从而提高评测的准确性和效率。

处理者应首先分析收到的任务指令，这些指令明确了需要评估的代码范围、标准以及预期结果。随后，处理者会根据这些指令，将任务分配给独立的处理单元。这些处理单元具备

自动化分析和处理数据的能力，能够独立完成对代码的分析和评估。在处理过程中，数据对象的分类和过滤是不可或缺的一步。通过公式(1)和公式(2)，可以计算出数据集中每个样本点之间的距离和相似度，以及它们与噪声数据的关联性。根据这些计算结果，将数据集中的样本点分为不同的类别，并去除那些与所需数据不匹配或相关性较低的噪声数据。

$$\begin{cases} A_e(B_x, S_a) = \sin(S'_a, c_i) \times r \\ a = |S_a| \end{cases} \quad (1)$$

式中： S_a 表示一组数据集的中心点； B_x 表示主数据库或主节点需要判断其是否与噪声数据对象相关联。 $A_e(B_x, S_a)$ 表示 S_a 与 c_i 之间的相似度或距离度量； r 表示数据集中样本的数量。如果 p 被定义为数据集中所有样本点的平均距离，那么 p 可用于评估 (B_x, S_a) 之间的距离采样是否与所需的数据相匹配。

$$p = \frac{A_e(S'_a, c_i) \times A_e(B_x, S_a)}{x \times a} \quad (2)$$

式中： x 表示数据集中采样点的数量。如果 $A_e(S'_a, c_i)$ 的值小于预设阈值 p ，则确定存储该数据。如果 $A_e(S'_a, c_i)$ 的值大于 p ，则将其视为低相关性数据，并作为噪声数据去除。

经过分类和过滤后的数据对象，将被用于在线评测的核心分析过程。在这一阶段，处理单元利用公式(3)计算并行数据处理之间的加速比，以评估随着处理次数和任务规模的变化，任务执行速度的提升程度。同时，处理单元通过比较串联模式和并联模式下的处理时间函数，来确定在给定任务规模和处理次数下，哪种模式能够更快地完成任务。

$$g(f, h) = \frac{p[k(f) + b(f)]}{k(f) + b(f)/s + \gamma(f, h)} \quad (3)$$

式中： f 表示任务的处理规模； h 表示处理的次数； $g(f, h)$ 表示加速比函数，描述了随着处理次数 h 和任务规模 f 的变化，任务执行速度的提升程度。 $k(f)$ 表示串联模式下的处理时间函数，它依赖于任务的处理规模 f ； $b(f)$ 表示并联模式下的处理时间函数，同样受到任务处理规模 f 的影响； $\gamma(f, h)$ 表示在并联模式下，随着任务处理规模 f 和处理次数 h 的变化，完成任务所需的

总时间。

通过在线测评的自动化处理和分析，可以有效地减少同步前信息过滤的时间，提高信息同步的速度和准确性。此外，主从式数据库信息的集成也有助于防止数据操作不同步的问题，进一步提升了在线测评的效率和可靠性，为高速信息同步提供了坚实的基础。

1.3 数据库信息快速同步流程设定

数据库信息快速同步流程的设计是建立接口数据库的基础，这些接口在接收到指令后能立即执行相应的操作。性能说明中的每个对象都记录在安装程序日志表中。当多个异构信息源发生更改时，库表还会触发相应的活动更改，以避免在执行指令时发生信息同步的双重更改，从而保持多个主从式数据库源之间的同步，并减少所需的同步时间。图1展示了数据库信息快速同步流程。

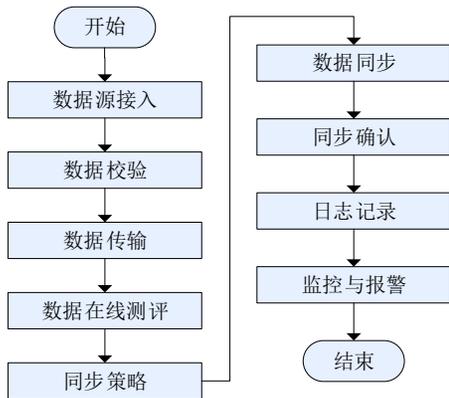


图1 数据库信息快速同步流程

数据库信息快速同步流程，提高了主从式数据库信息同步速度，该程序将定期扫描日志表，以便及时识别并删除那些已经完成同步的信息。

2 实验分析

为了验证本文提出的方法在实际应用中的有效性，将本研究中的信息同步方法用于同步数据库K中的各种类型数据。在对比实验中，将基于多核CPU的多源异构数据库信息快速同步方法作为传统方法1、基于网闸的主从式数据库安全同步方法作为传统方法2。实验参数见表2。

表2 实验仿真参数

参数	数值
起始节点	10
稳定性	100%
CPU使用率	15%或者30%
随机误差范围	-10~10μs

在本次实验中，选取K数据库中的数据和图像数据进行同步测试，目的是排除不同数据结构类型可能对实验结果的干扰，以确保结果的准确性和可靠性。

为了评估不同同步方法在处理大规模数据同步过程中对计算资源的消耗情况。CPU使用率作为一个重要的性能指标，可以反映出系统在执行任务时CPU的使用率，进而反映出系统的运行效率和负载情况。在实验过程中，采用主从式数据库信息进行十次的同步实验，并对不同方法的CPU使用率进行了分析。高清视频和一般视频的使用率，分别在30%左右和15%左右。根据常规经验，CPU使用率在15%左右或者30%左右是正常的，只要使用率不是长期超过50%，都属于正常范围内。三种方法在实验中的具体结果见表3。

表3 数据库信息CPU使用率结果对比(%)

实验序号	数据库信息	本文方法	传统方法1	传统方法2
1	视频	15.5	55.4	51.4
2	视频	14.7	56.4	51.3
3	视频	15.5	56.3	51.7
4	视频	15.6	56.7	50.5
5	视频	14.8	66.3	53.6
6	高清视频	30.5	57.2	52.3
7	高清视频	30.3	57.4	52.5
8	高清视频	30.4	57.5	51.3
9	高清视频	30.3	56.5	51.4
10	高清视频	30.4	57.3	52.5

观察表3中的实验数据可以发现，在进行数据库数据的同步时，本文方法的CPU使用率均在合理范围内，而传统方法1和传统方法2的CPU使用率较高，过高的CPU使用率可能会导致系统负载过重，甚至引发系统崩溃。实验结

果表明, 本文提出的数据库信息快速同步方法在CPU使用率方面相较于传统方法1和传统方法2具有优势, 同步效果更佳, 本文方法的优势在于通过在线测评工具的实时评估结果, 调整数据处理方式以提高同步效率。这有助于优化CPU使用率, 提高同步速度和同步效果。

在主从式数据库信息快速同步过程中, 网络传输和中转节点会引入一定的延迟和数据传输误差。因此, 测试不同节点跳数条件下的累积误差分布可以评估同步方法在实际网络环境中的稳定性和可靠性。在相同测试环境中展开500次主从式数据库信息快速同步节点同步测试, 在节点跳数分别为1跳、2跳、3跳、4跳的情况下, 测试上述方法的累积误差分布。测试结果如图2所示。

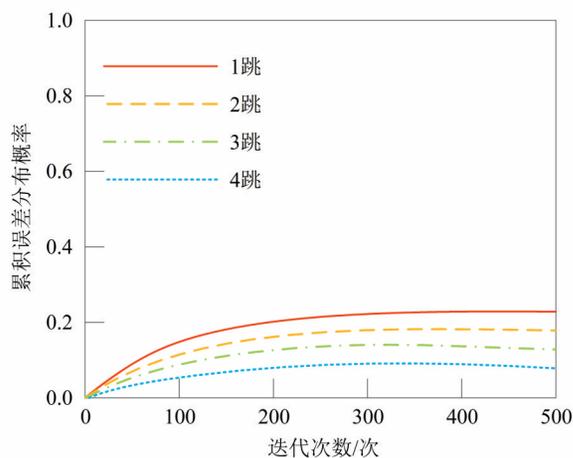


图2 不同节点跳数下的累积误差分布结果

对图2中的数据进行分析, 可以看到, 在不同节点跳数条件下, 本文提出的方法在信息同步过程中展现出较小的累积误差分布概率, 能够保持较高的可靠性, 更适用于主从式数据库信息快速同步。随着节点跳数的增加, 本文方法的累积误差分布概率略有上升, 但仍然保持在较低水平。即使在远程或复杂网络环境中, 本文方法依然能够有效地实现信息的快速同步。本文方法的优势在于在整体架构设计中, 主数据库负责处理写操作并维护最新的数据, 同时通过通知从数据库进行数据同步。这种设计方式能够保证数据同步的可靠性和一致性, 减少

数据传输过程中的信息丢失和错误。利用在线评测工具, 通过计算并比较并行处理和串行处理之间的加速比, 评估并行处理在提高同步效率方面的优劣势, 以此确定最佳的数据处理方式。这种灵活的数据处理方式能够更好地适应不同节点跳数条件下的变化, 以此提高了信息同步的可靠性和准确性。

3 结语

本文的创新之处在于将在线评测技术引入主从式数据库信息同步领域, 提出一种基于在线评测的快速同步方法。该方法不仅考虑了同步过程的效率和准确性, 还注重在实际应用环境中的动态评估和优化。通过理论分析和实验验证, 本文方法与传统方法相比, 在同步速度方面展现出显著的优势, 为数据库信息的快速同步提供了一种新的解决方案。

参考文献:

- [1] 谈天林, 贾文珏, 范延平, 等. 全国不动产登记数据库同步技术研究[J]. 自然资源信息化, 2023(5): 1-7.
- [2] 尹继璽. 电网调度控制系统数据库间同步故障分析及对策[J]. 现代工业经济和信息化, 2023, 13(8): 329-331.
- [3] 袁翔. 基于卫星遥感及数据库同步的气象灾害监测预警系统设计[J]. 计算机测量与控制, 2023, 31(7): 71-76.
- [4] 杨波, 邓欣, 王鑫章, 等. 基于多核CPU的多源异构数据库信息快速同步研究[J]. 中国信息化, 2022(9): 52-54.
- [5] 黄刘松, 张飞, 张学锋, 等. 基于网间的主从式数据库安全同步系统[J]. 大理大学学报, 2022, 7(12): 25-30.
- [6] 郭良银. 多关系型数据库向图数据库转换与同步方法[J]. 现代计算机, 2023, 29(10): 42-49.
- [7] 张万里. 分布式网络数据同步技术的应用研究[J]. 科技创新与应用, 2022, 12(21): 161-164.
- [8] 艾文凯, 江伟, 徐丹, 等. 基于访问趋势预测的时间序列数据库缓存管理研究[J]. 电工技术, 2022(6): 149-150, 153.

文章编号: 1007-1423(2024)20-0072-04

DOI: 10.3969/j.issn.1007-1423.2024.20.014

高并发 Web 应用系统性能优化研究

欧阳习彪¹, 徐宝林^{2*}

(广东白云学院大数据与计算机学院, 广州 510000)

摘要: 随着移动互联网的迅猛发展, 互联网用户急剧增加, 对于高并发的 Web 应用系统很容易出现响应时间过长、宕机等情况, 带来非常不好的用户体验, 存在用户流失的风险, 对网站进行性能优化的需求变得十分迫切。通过深入分析系统瓶颈与性能瓶颈, 从前端、后端、Nginx 服务器三个方面采用了一系列的优化策略, 有效提高了网站的性能, 显著提升了系统响应速度, 同时减少了资源浪费和减轻了服务器压力。

关键词: Web 应用系统; 性能优化; 高并发

0 引言

在当今数字时代, 高并发 Web 应用系统的广泛应用已经成为推动互联网发展的引擎之一。随着用户规模的迅猛增长, 这些系统在处理大规模并发请求时面临着前所未有的压力和挑战。系统性能的提升不仅关系到用户体验的好坏, 更直接影响着企业的竞争力和商业价值。因此, 如何有效地优化高并发 Web 应用系统的性能成为了业界研究的焦点之一。

高并发环境下, 系统往往面临着诸多性能瓶颈, 包括但不限于数据库负载过高、服务器响应时间过长、网络带宽受限等方面的问题。本文旨在通过深入分析这些瓶颈问题, 从前端到后端的多个关键领域, 包括使用图片服务器、CDN 加速、压缩合并资源、数据库优化、负载均衡技术、缓存策略、异步处理等方面的内容。通过综合运用这些手段, 能有效构建一个全面而实用的性能优化框架, 以适应高并发环境下的复杂场景。

1 高并发场景性能优化方案

1.1 前端优化

前端性能优化是提高网页加载速度和用户

体验的关键, 常用的前端优化方案如下:

(1) 压缩和合并资源: 压缩 CSS、JavaScript 和图像等前端资源, 以减小文件大小, 提高加载速度。此外, 合并多个文件为一个文件可以减少 HTTP 请求次数^[1]。

(2) 使用 CDN(内容分发网络): 利用 CDN 分发静态资源, 将内容(如网页、图像、视频等)分发到离用户更近的服务器上, 从而减少用户访问时的延迟和提高网站的性能。

(3) 异步加载和延迟加载: 将一些非必要且非实时展示的资源采用异步方式加载, 减少页面加载时间, 提高界面的响应时间。对于一些非关键资源, 可以延迟加载, 等主要内容加载完再加载这些资源^[2]。

(4) 优化图片: 使用适当的图像格式, 选择合适的分辨率, 利用图片懒加载技术, 以及采用图像压缩工具, 都可以减小图片文件大小, 提高页面加载速度。

(5) 减少 HTTP 请求: 减少 HTTP 请求是提升网站性能的有效手段之一。由于每个请求都会增加浏览器与服务器之间的通信次数, 因此, 减少不必要的资源请求可以显著降低加载时间。

收稿日期: 2024-04-08 修稿日期: 2024-05-07

基金项目: 广东白云学院 2022 年度广东省普通高校重点科研平台和项目(2022GCZX009)

作者简介: 欧阳习彪(1987—), 男, 湖南邵阳人, 硕士, 讲师, 研究方向为软件工程; *通信作者: 徐宝林(1972—), 男, 湖南邵阳人, 学士, 副教授, 研究方向为软件工程、信息系统集成, E-mail: XBL203@163.com

这可以通过图像优化实现，例如，选择适当的图像格式，仅在必要时使用图像，并通过 CSS Sprites 将多个图像合并为一个文件以减少请求数量。

(6) 使用浏览器缓存：浏览器缓存允许存储已访问网站的静态资源，当用户重复访问时可以直接从本地加载这些资源，而无需重新从服务器获取。通过设置合适的 HTTP 缓存头，可以控制缓存的时间和有效性，确保用户总是加载到最新的内容同时减少无谓的网络流量。

(7) 优化 CSS 和 JavaScript：避免使用过多的 CSS 和 JavaScript 文件，尽量减少 DOM 操作，压缩代码，移除不必要的注释和空白，以提高文件加载速度。

1.2 后端优化

1.2.1 数据库性能调优

在构建快速响应的高并发 Web 网站的后端性能优化策略中，数据库性能调优占据了核心地位。数据库作为存储与管理数据的关键系统组件，其性能直接影响到整个应用的响应速度和处理能力。因而，深入分析并实施针对性的调优措施，对保障网站能够高效、稳定地运行至关重要。

针对数据库性能调优，需从查询优化入手，关注 SQL 语句的执行效率。通过优化查询逻辑、使用索引、减少不必要的数据库访问和避免复杂的联接操作，可以显著提升查询速度。此外，对于频繁执行的查询，实施缓存策略可有效减少数据库的重复读取操作，进而提高整体性能。常用的 SQL 语句调优方法如下：

(1) 合理使用索引：在频繁用于查询的列上创建索引，特别是在 WHERE 和 ORDER BY 子句中使用的列。避免在经常变动的列上创建索引，因为索引的维护会增加额外开销。

(2) 避免全表扫描：确保 SQL 查询使用索引来定位数据，避免全表扫描。全表扫描会增加数据库负载并降低查询性能。

(3) 优化查询条件：确保 WHERE 子句中使用的条件是有效的，避免不必要的条件判断。

使用合适的条件顺序，将最具选择性的条件放在前面，以减少扫描的行数。

(4) 避免使用通配符查询：尽量避免使用通配符(如 %)在查询中开头位置，因为这会导致无法使用索引，需要进行全表扫描。

(5) 使用合适的 JOIN 类型：根据查询需求选择合适的 JOIN 类型，如 INNER JOIN、LEFT JOIN、RIGHT JOIN 等。避免笛卡尔积，确保 JOIN 条件正确，以避免返回过多的结果集。

(6) 适当使用子查询：在必要时使用子查询，但要注意子查询的效率，尽量避免嵌套过深的子查询。

(7) 批量操作优化：对于批量插入、更新或删除操作，使用批量操作语句(如 INSERT INTO ... VALUES、UPDATE ... SET、DELETE FROM ...)而不是逐条操作，以减少与数据库服务器的交互次数。

(8) 避免大量重复数据的返回：在查询结果中，尽量只返回需要的列，避免返回大量不必要的数据。

(9) 使用数据库执行计划分析：使用数据库提供的执行计划分析工具，分析 SQL 查询的执行计划，找出可能存在的性能瓶颈，并进行相应的优化。

1.2.2 缓存优化

缓存分为两大类：本地缓存和分布式缓存(redis、memcache等)。

(1) 本地缓存：本地缓存是将数据存储在应用程序的内存中，以提高数据访问速度和降低数据库压力。其工作原理是当应用程序第一次访问数据时，将查询结果缓存到本地内存中。下次再有相同的查询请求时，应用程序首先检查本地缓存，如果数据已经存在于缓存中，则直接从缓存中获取，避免了对数据库的查询操作^[3]。如果数据在缓存中过期或不存，则重新查询数据库，并将查询结果缓存到本地。定期或根据一定的策略更新或清理本地缓存，保证缓存数据的有效性和一致性。比如一个电子商务网站，首页需要展示热门商品列表。由于这些商品数据更新不频繁，可以将热门商品数

据缓存在本地内存中。当用户访问首页时，应用程序首先检查本地缓存，如果存在热门商品数据，则直接从本地缓存中获取，提高了页面加载速度。如果本地缓存中的数据已过期或不存，则从数据库中查询热门商品数据，并将查询结果缓存到本地，供下次访问使用。通过本地缓存，可以有效地减轻数据库的压力，提高网站的性能和用户体验。

(2) 分布式缓存：分布式缓存是将数据缓存在集群中的多个节点上，可以提供更高的扩展性和可靠性。其工作原理是将数据按照一定的规则分片存储在多个节点上，保证了数据的分布和负载均衡。当应用程序需要查询数据时，首先根据数据的键值计算出数据所在的节点，然后向该节点发送查询请求。如果数据在节点的缓存中存在，则直接返回数据；如果不存在，则从数据库中查询数据，并将查询结果缓存到节点中。分布式缓存通常支持数据的备份和复制机制，保证了数据的可靠性和一致性。假设一个社交网络应用，用户的个人信息和关注关系数据是频繁访问的热点数据。使用 Redis 或 Memcached 构建分布式缓存集群，将用户个人信息和关注关系数据缓存在多个节点上。当用户访问页面时，应用程序首先根据用户 ID 计算出数据所在的节点，然后从该节点中获取数据。通过分布式缓存，可以提高数据的访问速度和并发处理能力，保证了系统的性能和可靠性。

1.2.3 Nginx 服务器优化

在构建高并发 Web 应用时，Nginx 的性能优化至关重要。通过合理的配置参数调整和功能特性的利用，可以显著提升服务器的并发处理能力、响应速度和稳定性，从而提供更好的用户体验和系统可靠性^[4]。

(1) 配置参数调整

Nginx 的性能优化始于对其配置参数的调整。其中一个关键参数是 `worker_connections`，它决定了 Nginx 能同时处理的连接数。默认情况下，该值较低，通过增加这个参数可以提高 Nginx 的最大并发连接数。然而，调整时需要注意

服务器硬件资源的限制，确保不会过度消耗资源导致性能下降或系统崩溃。

另一个重要的参数是 `limit_req`，它用于限制客户端请求的速率，防止恶意请求导致服务器过载。此外，启用 `sendfile` 功能可以直接在磁盘和网络之间传输文件，减少内存拷贝，提高数据传输效率。

缓存机制的调整也是性能优化的重点。通过配置 `proxy_cache` 相关参数，可以将响应缓存到磁盘上，减少对后端服务器的请求，提高网站的访问速度。适当增加缓存大小和延长缓存时间有效期，可以进一步提升性能。

(2) 负载均衡与反向代理

负载均衡是通过将流量分配到多个服务器上来优化性能和可靠性的技术。Nginx 提供了多种负载均衡策略，如轮询、最少连接或基于 IP 哈希等。这些策略可以根据实际需求选择，确保每个服务器得到均等的请求处理机会，提高系统的吞吐量和可用性^[5]。

同时，Nginx 作为反向代理服务器，可以提供更高级的内容控制和安全保障。所有客户端请求首先经过 Nginx，再由其转发到后端的服务器。这种设计不仅可以提高缓存效率，减少后端服务器的直接暴露风险，还能通过修改请求或响应头信息实现更灵活的内容控制。

(3) keep-alive 参数的优化

`keep-alive` 参数允许在一个 TCP 连接上发送多个 HTTP 请求，减少了建立和关闭连接所消耗的时间和系统资源。然而，并非所有场景下提高 `keep-alive` 值都会带来性能提升。要根据客户端与服务器端的网络延迟情况、服务器的处理能力和应用的特点细致调整此参数。通过实际测试和监测工具评估不同设置对性能的影响，找到最佳平衡点。

综上所述，Nginx 的性能优化是一个复杂而细致的工作，需要深入理解其内部工作机制和网络流量模式。通过合理的配置参数调整、负载均衡和反向代理的实现以及 `keep-alive` 参数的优化，可以使 Nginx 在高并发环境下表现出色，为用户提供快速、稳定的访问体验。

2 结语

在数字化时代背景下，构建能够迅速响应并处理高并发请求的Web网站已成为信息技术领域的关键挑战之一。本研究针对此问题，从前端、后端及Nginx优化的角度出发，深入分析并提出了一系列性能优化策略。这些策略不仅涵盖了代码层面的优化，如减少HTTP请求、合理管理缓存和压缩文件等，还包括了系统架构调整，例如负载均衡和数据库分库分表等技术手段。经过综合分析，我们认识到前端优化的重要性，通过精简资源、使用CDN和数据异步加载等技术显著提升了页面加载速度。同时，后端优化中高效的数据库设计和查询优化，都对提高网站的并发处理能力起到了决定性作用。

此外，Nginx作为一个高性能的Web服务器和反向代理服务器，在处理静态内容、负载均衡和SSL终止等方面展现了极大的优势。

参考文献：

- [1] 钟宇飞. 移动互联网时代的Web性能优化实践[J]. 信息通信, 2020(1):267-268.
- [2] 陈洪雁, 万俊伟. 基于Web的企业网站性能优化方案研究与应用[J]. 物联网技术, 2018, 8(2):67-69, 73.
- [3] 王伟, 谭松荣. 高并发Web应用系统的性能优化分析与研究[J]. 现代信息科技, 2019, 3(3):22-23.
- [4] 李若兰. 基于Nginx的Web服务器优化的应用研究[J]. 科技风, 2021(9):119-120.
- [5] 刘振辉. 高性能网站安全建设探讨[J]. 计算机安全, 2012(10):45-48.

Research on performance optimization of high-concurrency web application system

Ouyang Xibiao¹, Xu Baolin^{2*}

(School of Guangdong Baiyun University, Guangzhou 510000, China)

Abstract: With the rapid development of the mobile Internet, the number of Internet users has increased dramatically. High-concurrency Web application systems are prone to long response times, downtime, and other issues, resulting in a very poor user experience and a risk of user churn. The need for performance optimization of websites has become urgent. Through in-depth analysis of system bottlenecks and performance bottlenecks, a series of optimization strategies have been adopted from the front-end, back-end, and Nginx server to effectively improve the performance of the website, significantly increase the system response speed, and reduce resource waste and server pressure.

Keywords: Web application system; performance optimization; high concurrency

开发案例

文章编号: 1007-1423(2024)20-0076-06

DOI: 10.3969/j.issn.1007-1423.2024.20.015

基于 MLP 的电动汽车充电桩共享推荐模型

王伟, 游凤芹*, 徐 晖

(南京理工大学紫金学院人工智能学院, 南京 210000)

摘要: 针对日益增长的电动汽车充电需求与公共充电桩分布不均、私人充电桩闲置率高之间的矛盾, 提出了一种基于多层感知器(MLP)的电动汽车充电桩智能推荐模型。该模型综合用户需求、充电设施特性及用户历史行为数据, 采用特征提取技术构建特征集, 并运用 MLP 算法进行训练与优化, 以实现个性化充电桩推荐。与随机森林算法相比, 该模型在用户个性化充电服务方面表现更优, 有效提升了充电资源利用率, 缓解了充电桩供需矛盾。

关键词: MLP; 充电桩; 智能推荐; 特征提取

0 引言

区别于传统能源汽车的充电方式, 新能源汽车充电桩的智能化推荐系统实现了一种基于用户需求和充电设施特性的智能匹配环境。然而, 由于充电桩的地理位置、充电速度、费用和可用性等因素的多样性, 给用户选择合适的充电桩带来了新的挑战。由此可见, 基于智能匹配的电动汽车充电桩共享系统, 既符合全球环保趋势和科技发展方向, 又能解决电动汽车发展中的关键问题, 具有重要的现实意义和广阔的应用前景。

目前, 围绕电动车充电桩共享领域的关键技术, 已有国内外多位学者进行了研究。Zhang 等^[1]通过多智能体强化学习方法, 研究了智能充电推荐问题, 主要针对如何通过智能体之间的协作来优化充电站的选择和充电时间的安排。该方法虽然提高了充电效率, 但是忽略用户个性化需求。Li 等^[2]采用动态规划方法, 针对电动汽车充电调度问题进行了研究, 旨在最小化充电成本的同时提高充电效率。他们的研究考虑

了电池状态、充电站可用性和电价等因素, 但在用户个性化信息的整合上可能不够充分。Ravi 等^[3]探讨了电动汽车在车联网服务中的应用, 特别是车辆到电网服务的潜力, 他们的研究主要关注于电网的整体效益。Zhang 等^[4]提出了一种基于量子粒子群优化算法的充电调度优化方法, 该方法利用量子计算的优势来提高充电调度的效率。然而, 这项研究可能同样缺乏对用户个性化信息的深入考虑。Lin 等^[5]研究了一种多阶段马尔可夫决策过程(MDP)模型, 用于充电站推荐。他们的模型考虑了电动汽车的行驶路线、电池状态和充电需求。吴敏^[6]研究了深度学习框架下的充电桩个性化推荐算法。考虑了用户与物品的交互行为, 但是未考虑用户和充电桩的内容信息。

为实现更高效、更个性化的充电桩推荐, 本文提供了一种基于多层感知器 (Multilayer Perceptron, MLP) 的智能推荐算法。通过分析用户的充电需求、行为偏好以及地理位置等信息, 为用户推荐最合适的充电桩, 从而提高用户的充电效率和满意度。

收稿日期: 2024-04-20 修稿日期: 2024-05-27

基金项目: 校级科研项目(2023ZRKX0401010)

作者简介: 王伟(2001—), 男, 江苏扬州人, 本科, 研究方向为算法设计与分析; *通信作者: 游凤芹(1983—), 女, 江苏姜堰人, 硕士, 高级工程师, 研究方向为嵌入式系统、算法设计与分析等, E-mail: irisyou1015@gmail.com; 徐晖(1992—), 女, 江苏镇江人, 硕士, 助教, 研究方向为计算机视觉、算法设计与分析等

1 系统设计

本文提出的智能充电桩推荐模型，其框架示意图如图1所示。该模型通过高效集成充电桩共享平台的实时数据，包括充电桩的状态和用户行为信息，利用特征提取技术深入分析并构建数据特征集。随后，采用多层感知器(MLP)模型对这些特征进行学习和训练，以生成个性化的充电桩推荐，从而显著提升了充电桩的利用率。

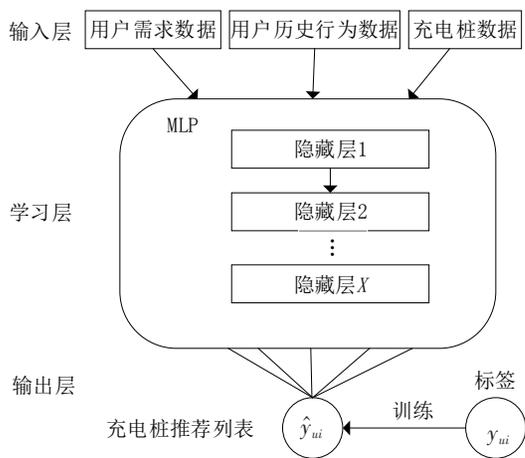


图1 基于MLP的智能推荐模型框架

2 系统实现

2.1 数据集设计

本文的数据集基于百度数据平台，整合了8万条充电桩及其用户行为的详细信息。表1展示的是数据集前三条数据。

数据集字段介绍：

charging_station_number(充电桩编号)：充电桩的唯一编号，用于标识每一个充电桩。

category(充电桩分类)：描述充电桩的类型，如快充、慢充等，反映了充电桩的充电速度和适用场景。

unit_price(单价)：充电的单价，单位通常

为元/度，反映了使用充电桩的成本。

total_cost(总金额)：充电订单结算总金额，单位通常为元，反映了使用充电桩的总金额。

userid：使用或查询充电桩用户的唯一标识符，可以用来跟踪用户行为和分析用户特征。

c_lo(经度)和c_la(纬度)：充电桩的地理坐标，提供精确的位置信息，对于地理位置分析和优化充电桩布局至关重要。

user_lo(经度)和user_la(纬度)：用户用车需求的地理坐标，提供精确的用户位置信息，对于距离计算以及充电桩推荐至关重要。

2.2 数据预处理

在本文中，算法的实现遵循以下专业化步骤：首先，从数据库中提取原始数据，经过一系列预处理操作，形成适合模型训练的数据集。在模型初始化阶段，我们采用了Xavier初始化策略，以促进网络参数的稳定学习。同时，选用ReLU作为激活函数，以增强模型在处理非线性问题时的表现力。随后，对构建的模型进行系统的训练和参数优化，以提升模型的预测准确性和泛化能力。数据预处理的具体操作细节已在图2中以流程图的形式详细展示。

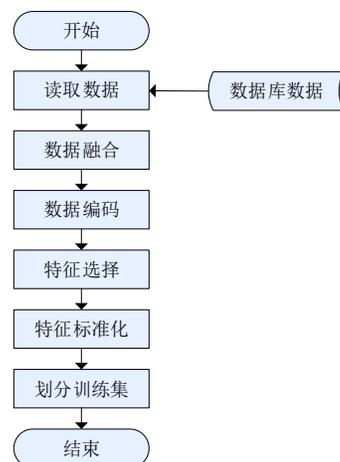


图2 MLP模型数据预处理流程图

表1 示例数据集

charging_station_number	category	unit_price	total_cost	userid	c_lo	c_la	user_lo	user_la
CJS-2024-0417-006	慢充	0.55	31.90	1708304370190	116.751	39.8756	118.908	32.0965
CJS-2024-0417-006	慢充	1.13	14.69	1708304370190	116.751	39.8756	118.908	32.0965
CJS-2024-0417-004	慢充	2.09	160.93	1708304370190	116.266	39.9576	118.908	32.0965

(1) 读取数据: 首先, 使用 pandas 库中的 read_sql 函数从数据库中读取数据, 并使用 head 函数获取前 100 条记录作为样本数据。

(2) 数据融合: 通过集成高德地图 API 获取的地理坐标点间距离信息, 作为额外的特征维度纳入数据集, 以增强模型对空间关系的捕捉能力。这一步是预处理中的一个创新点, 因为它结合了外部 API 来丰富原始数据集, 为模型提供了更多的上下文信息。

(3) 数据编码: 对于分类变量, 使用 Label-Encoder 进行编码, 将文本标签转换为数值, 以便模型能够处理。这是数据预处理中常见的步骤, 确保了模型能够理解分类特征。

(4) 特征选择: 选择了 category(充电桩分类)、total_cost(总金额)、distance(距离)和 unit_price(单价)作为特征, 而 charging_station_number(充电桩编号)作为标签。

(5) 特征标准化: 使用 StandardScaler 对特征进行标准化处理, 确保所有特征都在相同的尺度上, 这有助于提高模型的性能。

(6) 划分训练集: 将原始数据集 20% 的数据作为测试集, 剩余的 80% 将作为训练集。

2.3 模型的训练及优化

本文采用的 MLP 模型的训练通过调用 MLP-Classifer 来实现, 为了找到模型运行最佳的参数, 需要进行模型的优化。本算法采用的超参数优化技术是网格搜索。伪代码如下:

输入: 用户历史数据和充电桩数据

输出: 模型最佳参数组合和最佳准确率

1. 初始化参数网格 param_grid, 包含隐藏层大小和初始学习率的不同组合:

- 隐藏层大小: [100, 200, (100, 100), (50, 50), (100, 50, 100)]

- 初始学习率: [0.01, 0.001, 0.0001]

2. 创建 MLPClassifier 实例 mlp, 并设置随机状态为 42。

3. 创建一个 Pipeline 实例 pipeline, 它包含两个步骤:

- 第一步使用 StandardScaler 进行特征标准化, 命名为 'scaler'。

- 第二步使用之前创建的 MLPClassifier 实例 mlp, 命名为 'mlp'。

4. 初始化 GridSearchCV 实例 grid_search, 设置以下参数:

- pipeline: 之前创建的 Pipeline 实例。

- param_grid: 之前定义好的参数网格。

- cv: 交叉验证的折数, 设置为 5。

- scoring: 评分标准, 设置为 'accuracy'。

- n_jobs: 并行作业数, 设置为 -1 表示使用所有可用的 CPU 核心。

- verbose: 控制网格搜索的详细输出, 设置为 1 表示详细输出。

5. 使用 grid_search.fit() 方法在训练数据 X_train 和 y_train 上训练模型, 并执行网格搜索。

6. 打印最佳参数组合 grid_search.best_params_ 和对应的最高准确度 grid_search.best_score_。

通过网格搜索之后, 可以获取模型运行最佳的参数搭配, 如下所示:

```
Best parameters: {'mlp__hidden_layer_sizes': (200,),
                  'mlp__learning_rate_init': 0.01}
```

```
Best accuracy: 0.9875
```

这个结果表明, 在进行多层感知器 (MLP) 神经网络的超参数优化过程中, 找到了一组参数使模型在训练集上达到了最佳的准确度 0.9875。

3 实验与分析

3.1 结果可视化

为评估模型在实际应用中的性能, 本算法首先加载了特征缩放器 StandardScaler 对数据进行归一化处理, 确保了模型输入数据的一致性和可比性。从缩放后的数据中分离出特征和标签, 并将它们传递给 MLP 模型进行预测。之后使用模型的 predict 方法对缩放后的特征数据进行预测, 得到一个包含推荐充电桩 ID 的数组。使用 Counter 类统计预测结果中每个推荐值的出现次数, 并将其转换为一个排序后的字典, 以便进行可视化。模型通过创建一个条形图来可视化推荐结果的计数。显示出的条形图如图 3 所示, 代表着各个充电桩的推荐值。

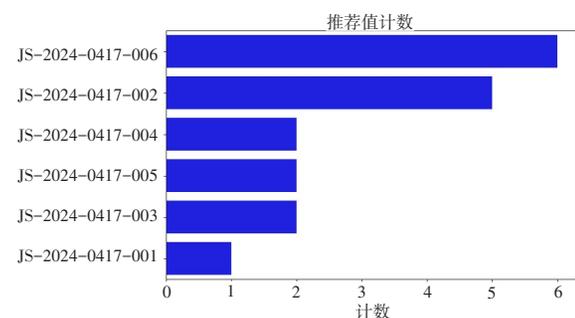


图 3 TopN 充电桩推荐列表

根据图表所示，充电桩CJS-2024-0417-006的推荐指数居于首位，这表明在综合评估了用户的充电需求、历史使用模式以及所有充电桩的属性数据后，智能推荐模型以优先级最高的形式推荐了该充电桩。此外，模型还对其他充电桩进行了推荐排序，形成了从高到低的推荐列表。这为用户提供了广泛的选择范围，但根据模型的评估，CJS-2024-0417-006充电桩在当前数据集和用户行为模式下，被认为是最优化的选择。此推荐结果旨在辅助用户做出更加明智的充电决策，同时提升充电基础设施的利用效率和用户满意度。

3.2 性能评估

为了评估算法的性能，可以通过准确度、精确度、召回度以及F1得分分析模型的性能，以下是这些参数的详细解释：

(1) 准确率 (Accuracy)：该指标直接反映了模型正确分类样本的比重，其中 TP (True Positives) 代表正确判定的正类样本数量， TN (True Negatives) 代表正确判定的负类样本数量。计算公式为

$$Accuracy = \frac{TP + TN}{Total\ Samples} \quad (1)$$

式中： TP 是真正例 (正确预测为正类的数量)， TN 是真负例 (正确预测为负类的数量)。

(2) 精确率 (Precision)：此指标衡量了被模型判定为正类的样本中，实际上是正类的比例。计算公式为

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

式中： FP 是假正例 (错误预测为正类的数量)。

(3) 召回率 (Recall)：召回度又称为真正例率 (True Positive Rate, TPR)，它表示在所有实际正类样本中，模型正确识别出来的比例。计算公式为

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

式中： FN (False Negatives) 是假负例 (错误预测为负类的数量)。

(4) F1分数 (F1 Score)：F1得分是精确度与召回度的调和平均，它试图在两者之间找到一个平衡点。F1分数对于那些不平衡类别分布的数据集尤其重要，因为它同时考虑了模型对少

数类的预测性能。计算公式为

$$F1_Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

为了体现本文提出的充电桩智能匹配算法的优越性，本文选取了随机森林算法作为性能比对的基准。随机森林算法作为一种集成学习技术，以其在分类和回归问题中的高效性和准确性而受到广泛关注^[7]。

本文旨在通过与随机森林算法的比较，进一步确定我们提出的智能匹配算法在不同场景下的表现及其潜在的优势和局限性，以确定各自的优缺点。以下分别是两个模型的构建代码：

```
rf = RandomForestClassifier(n_estimators=200,
                           random_state=42, max_depth=5, min_samples_
                           split=10, min_samples_leaf=5, max_features='sqrt',
                           bootstrap=False)
mlp = MLPClassifier(hidden_layer_sizes=(200,),
                   max_iter=1000, random_state=42, alpha=0.01)
```

为了保证两个模型参数上的公平性，随机森林的参数进行如下调整：

参数 $n_estimators=200$ 指定了组成森林的决策树数量。 $random_state=42$ 确保结果的可重复性。 $max_depth=5$ 限制了树的深度，有助于防止过拟合。 $min_samples_split=10$ 和 $min_samples_leaf=5$ 分别设置了分裂内部节点和位于叶节点的最小样本数，增加了模型的泛化能力。 $max_features='sqrt'$ 意味着在每个分裂点，随机森林将从所有特征中随机选择特征数量的平方根来考虑分裂。 $bootstrap=False$ 意味着在构建每棵树时使用整个数据集，而不是随机抽样，这通常用于减少过拟合。然后进行模型的训练和评估，之后各自计算性能指标，绘制出的折线图如图4所示。

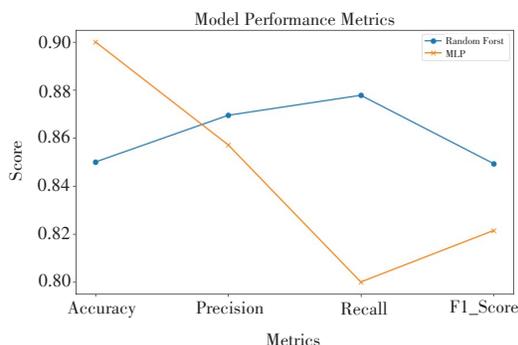


图4 模型性能对比折线图

通过对比两种模型的性能指标可以得出以下结论：在准确率方面，MLP模型比随机森林模型高出5个百分点(90% vs 85%)。在精确率方面，随机森林模型略高，领先MLP模型约1.23个百分点(86.94% vs 85.71%)。在召回率方面，随机森林模型也略胜一筹，比MLP模型高出7.78个百分点(87.78% vs 80%)。在F1分数方面，随机森林模型的性能略高于MLP模型，但两者相差不大(84.93% vs 82.14%)。

综上所述，尽管在精确率和召回率这两个重要指标上，随机森林展现出了更好的性能。但在准确率方面MLP具有更优的表现。

3.3 效率测试

为了评估多层感知器(MLP)算法在处理大规模数据集时的效率，本文设计了模拟实验以模拟不同数据量下的预测性能。实验从数据集中分别取出20、30、40、90、100个样本，从而观察模型的运行效率，结果如图5所示。

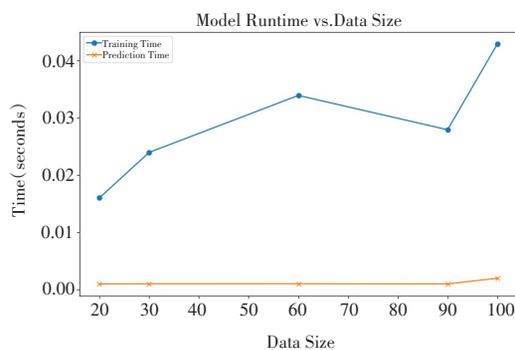


图5 MLP算法效率折线图

可以看出，随着数据集样本的不断增大，训练时间也逐渐增大(数据集的复杂性不同导致训练时间的波动，但训练时间呈现总体上升趋势)，预测时间还是维持在很短的时间内。实验表明本文展示的模型具有良好的运行效率。

4 结语

本文提出的基于MLP的充电桩推荐共享模型，通过精准的、个性化的充电桩推荐，用户可以更快速地找到符合需求的充电桩，减少寻

找充电桩的时间，提高充电效率。系统充分考虑了用户需求和充电桩的特性，并通过MLP模型的学习和训练，实现了高效的推荐机制。与随机森林算法的比较分析进一步证实了MLP模型在准确度上的优势。此外，效率测试显示了MLP模型在大规模数据处理上的高效性，验证了其在实际应用中的潜力。综上，本课题通过提升用户体验、优化资源利用率，为电动汽车共享充电服务的发展提供技术支持。未来的工作可以根据用户的实时需求和环境变化动态调整推荐结果。

参考文献：

- [1] ZHANG W, LIU H, WANG F, et al. Intelligent electric vehicle charging recommendation based on multi-agent reinforcement learning [C] // Proceedings of the Web Conference 2021, Ljubljana: ACM, 2021: 1856-1867.
- [2] LI K, ZHANG Y, DU C, et al. Dynamic programming-based optimal charging scheduling for electric vehicles [C] // Proceedings of the 2022 IEEE 7th International Conference on Intelligent Transportation Engineering (ICITE), Beijing: IEEE, 2022: 545-550.
- [3] RAVI S S, AZIZ M. Utilization of electric vehicles for vehicle-to-grid services: progress and perspectives [J]. Energies, 2022, 15(2): 589.
- [4] ZHANG G, DAI M, ZHAO S, et al. Dynamic charging scheduling optimization for electric vehicles based on quantum SWARM algorithm [C] // Proceedings of the 2022 7th Asia-Pacific Conference on Intelligent Robot Systems (ACIRS), Tianjin: IEEE, 2022: 154-157.
- [5] LIN H, LIN X, LABIOD H, et al. Toward multiple-phase MDP model for charging station recommendation [J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(8): 10583-10595.
- [6] 吴敏. 深度学习框架下充电桩的个性化推荐算法研究[D]. 武汉: 湖北工业大学, 2020.
- [7] 邵佳佳, 陈小毅, 戴媛媛, 等. 基于随机森林的纯电动汽车公共充电需求研究[J]. 供用电, 2023, 40(9): 58-64.

An MLP-based smart recommendation model for electric vehicle charging pile sharing

Wang Wei, You Fengqin*, Xu Hui

(Zijin College, Nanjing University of Science and Technology, Nanjing 210000, China)

Abstract: Addressing the growing demand for electric vehicle (EV) charging and the contradiction between the uneven distribution of public charging piles and the high vacancy rate of private charging piles, the author proposes an intelligent recommendation model for electric vehicle charging piles based on the MultiLayer Perceptron (MLP). The model integrates user requirements, characteristics of charging facilities, and historical user behavior data, employing feature extraction techniques to construct a feature set, and utilizes the MLP algorithm for training and optimization to achieve personalized charging pile recommendations. Compared with the Random Forest algorithm, this model demonstrates superior performance in providing personalized charging services for users, effectively enhancing the utilization rate of charging resources and alleviating the contradiction between the supply and demand of charging piles.

Keywords: MLP; charging pile; intelligent recommendation; feature extraction

(上接第71页)

Research on the rapid synchronization method of master-slave database information based on online evaluation

Wang Weiming*

(Jiangxi Modern Polytechnic College, Nanchang 330095, China)

Abstract: The traditional synchronization method lacks real-time evaluation and analysis mechanism, and the synchronization effect is not good. Therefore, this paper proposes a fast synchronization method of master-slave database information based on online evaluation. The master database is responsible for handling the write operation and maintaining the latest data, and notifying the slave database to synchronize the data when it changes. The online evaluation tool is introduced to evaluate the synchronized data in real time. Taking the data processing acceleration ratio as the evaluation index, the data processing mode is adjusted, and the rapid synchronization process of database information is designed. The synchronization results are returned to the main database through the feedback mechanism to ensure the synchronization effect and quality. The experimental results show that the CPU utilization rate of this method is within a reasonable range, and it has better synchronization effect, and the probability of cumulative error distribution is small, which indicates that it can maintain high reliability in the process of information synchronization.

Keywords: online evaluation; master database; data synchronization; fast synchronization method

文章编号: 1007-1423(2024)20-0082-05

DOI: 10.3969/j.issn.1007-1423.2024.20.016

“拾忆长征”红色旅游小程序设计与开发

沈 楨, 李丽娟*, 徐 妍

(江苏第二师范学院教育科学学院, 南京 211200)

摘要: 互联网+时代文化旅游产业资源呈现多样化发展, 旅游业和文化产业高度融合, 红色旅游发展态势尤为迅猛, 但对长征路线旅游资源进行智媒化开发和宣传的研究仍屈指可数。以贵州省为例, 采用数字化手段, 以让用户了解贵州省的长征旅游路线及其背后蕴含的红色精神为目的, 设计具有个性化旅游路线选择、3D 景点漫游、答题交互、积分兑换等功能的微信小程序, 旨在为红色旅游新模式的开发做出贡献, 并为用户学习长征精神提供一个平台。

关键词: 微信小程序; 红色旅游; 长征精神; 互联网+; 全景漫游

0 引言

2023 年召开的全国宣传思想文化工作会议中, 习近平总书记强调文旅工作者应以奋斗姿态和永不懈怠的精神状态推动文化和旅游高质量发展, 实现创造性转化和创新性发展。在互联网+时代文化旅游产业资源呈现多样化发展, 贵州作为很多学者关注的红色革命基地, 围绕其资源共生系统、可持续发展、媒体化、扶贫政策以及旅游形象开发等方面有大量的研究, 其中对生态可持续发展和旅游形象整合发展的研究占据主导地位, 然而, 对贵州长征路线的旅游资源进行智媒化开发和宣传的研究仍然较为稀缺。

在互联网+新媒体时代, 小程序因其便利快捷的优点, 深受大众欢迎^[8]。对贵州长征路线的旅游资源进行智媒化开发, 这在形式上与信息时代的特点更契合, 可以让人们更好地学习长征历史, 了解长征路线上城市的红色文化底蕴, 传播发扬红色精神, 对人们积极学习党史, 提升民族自豪感有促进作用。

红色旅游市场发展兴旺, 但红色景点的宣

传和可持续发展还存在一些困境, 例如, 一些景点真正的精神内核在游客游览的过程中很难去深度挖掘, 千篇一律的“拍照, 打卡, 走人”模式是否会引起游客的深度思考, 这也是开发本小程序的目的。在红色旅游百花齐放的大背景下, 为了提升游客的满意度和体验质量, 红色旅游产品的创新就显得尤为重要。因此, 需要从立意点、深度、用户的参与度和体验质量等方面进行创新, 以满足游客的文化需求, 从而提高竞争力, 进而让红色基因植根于人们心中^[1]。

1 “拾忆长征”小程序设计思想

1.1 内容设计

本研究深入挖掘贵州地区的长征红色景点及其历史背景, 基于信息时代下人们对于电子设备的青睐, 通过小程序与公众号等数字媒介进行宣传与展示。同时还将重点宣传贵州省的红色旅游资源, 利用 3D 立体图片技术, 让用户沉浸于长征路线与特色景区的场景中, 提供个性化的旅游路线建议, 使游客能够身临其境地

收稿日期: 2024-04-05 修稿日期: 2024-05-07

基金项目: 2023 大学生创新创业训练计划江苏省级项目(014DC23025); 2021 江苏省高校哲学社会科学研究项目(2021SJA0592)

作者简介: 沈楨(2003—), 女, 江苏无锡人, 在读本科, 专业为教育技术学; *通信作者: 李丽娟(1981—), 女, 河南许昌人, 硕士, 讲师, 研究方向为数字化教育资源、信息化教育, E-mail: llj2009@163.com; 徐妍(2003—), 女, 江苏南通人, 在读本科, 专业为教育技术学

感受红色旅游的独特魅力。

“拾忆长征”小程序考虑到现代旅游攻略的繁杂性，精简了小程序的功能，让更多用户体会到一键生成路线和沉浸式游览的便捷性，因此该小程序提供了路线和文化特色板块。在路线模块，为了充实用户的选择和学习的趣味性，用交互式动画呈现出多条红色旅游路线，尽力开发出多种贵州长征红色旅游资源，使用户在旅游前对路线的规划和各个景点的介绍有提前的了解。

本小程序内容上，为了让用户全面了解贵州省的文化习俗和深入学习贵州省的长征故事，均以视频和答题的形式呈现。在民俗文化模块，用视频的形式介绍贵州的历史、地理、习俗、节日、特色美食等方面，用户可以选择自己想要了解的板块观看视频，视频播放结束可以进入答题界面，根据视频内容答题，答对即可获得积分。另一个长征故事模块用动画视频展示一些长征时期发生在贵州的历史故事，以及长征中涌现出的精神，如长征精神、遵义会议精神等，以此来达到红色教育的目的^[2]，实现路径如图1所示。

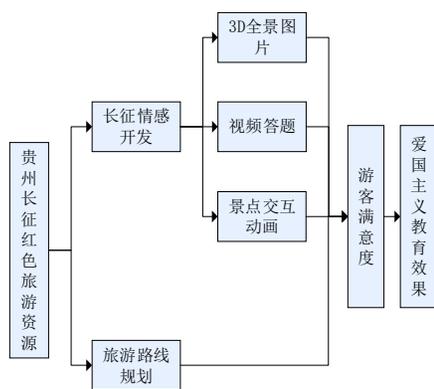


图1 “拾忆长征”小程序爱国主义教育目的实现路径

在红色旅游景点中，游客通过亲身体验和感受红色文化，将红色精神与文化价值观内化为自身的信念，这一过程被定义为游客的红色精神升华。对于这一过程中的内在规律和明确的路径，仍需进行深入的探讨和研究。“拾忆长征”让游客在出行之前有规划，对长征精神有深入挖掘，进一步提升新时代红色旅游的发展活力，推动红色旅游的高质量发展，能够为用户提供更为便捷、直观的红色长征历史文化体验，激发大众对长征精神的学习与传承的热情^[3]。

1.2 长征情感开发

(1) 3D全景图片：一般景点介绍只局限于2D图片以及文字相结合，“拾忆长征”小程序突破传统观念，将现下新兴的全景拍摄技术与景点导览相结合，使用户身临其境地感受到贵州长征旧址曾经发生的故事。用户通过拖拽和放大游览景点，意识到旧址不仅是缅怀革命先烈的地方，更是接受爱国主义教育的圣地^[4]。

(2) 视频答题：视频呈现一个系列，其中包括了四渡赤水、飞夺泸定桥等著名的战役故事，视频中用纪实与动画相结合的方式呈现出红军坚定的革命信念和艰苦奋斗精神，在不知不觉中将红色长征精神深入用户心中。视频结束后也有答题环节，这样不仅可以巩固对长征故事的了解，更可以引导用户深入思考和体会红军长征精神的内涵和价值。

(3) 景点交互动画：长征军队的众多导致旅游路线也呈现百花齐放的态势，“拾忆长征”小程序通过交互动画的形式呈现出长征的路线和历程，这不仅可以增加用户对长征的认识，更可以激发用户的爱国情感和民族自豪感。

2 “拾忆长征”小程序具体设计

2.1 总体设计

“拾忆长征”微信小程序包括四个主要页面，如图2所示。

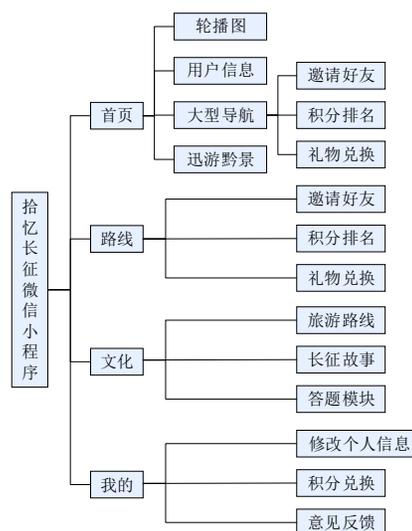


图2 “拾忆长征”小程序总体设计结构图

其中，“首页”帮助用户快速了解该小程序的主

体功能；“路线”页面呈现多条贵州红色旅游路线供用户选择、游览；“文化”页面为用户提供贵州省的文化习俗、长征精神两方面的视频让用户深度学习、获得切实的学习体验；“我的”页面包括用户个人信息、积分兑换、意见反馈等，帮助用户提高使用体验感和获得感。

2.2 小程序功能模块

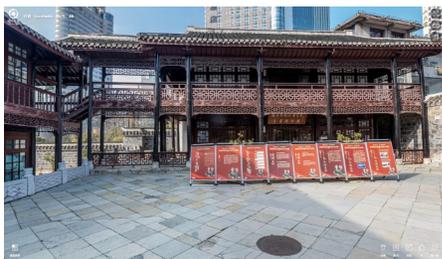
(1) 功能信息速递：该小程序首页的按钮可以帮助用户快速跳转到各个功能界面，为用户提供清晰的功能指引和方便快捷的操作。用户在轮播图区域可以看到长征相关的宣传海报，页面下方的迅游黔景模块为用户呈现了多个贵州著名红色景点的图片，用户点击即可看到景点的详细信息。

(2) 旅游路线选择：“路线”页面是该小程序向用户介绍贵州红色景点及长征精神的重要途径之一。该页面呈现一张贵州的地图(如图3)，上面标志着著名的景点以及地形地貌，地图下面是多条贵州红军长征路线，用户点击某个路线按钮，该路线上的著名景点就会以卡通图标的形式呈现在地图上，点击每个卡通图标都会跳出一个介绍该景点的新页面，包括景点的文字介绍和图片，该功能让用户全面了解各个贵州省长征旅游路线，为用户提供个性化旅游路线定制。



图3 小程序“定制路线”功能展示图

(3) 3D全景漫游：景点详情页面最上方是一张该景点的全景照片，用户通过手动旋转和拖动可以从不同角度观察该景点，该小程序的全景漫游还对景点的各个标志性建筑设置了交互，用户点击交互按钮即可进入或近距离观察标志性建筑，对于一些石碑上的碑文也进行了相关处理，用户通过交互可以清晰地浏览石碑上的文字。该功能让用户在线上也可以身临其境地游览贵州的红色景点^[7]。



(a)



(b)

图4 小程序全景漫游部分展示图

(4) 答题积分兑换：“文化”页面是该小程序让用户深入学习红色文化而设置的页面，用户在此页面观看视频学习贵州的文化习俗和长征故事，看完视频后进入答题界面，根据视频内容回答问题，答对获得积分，积分积累到一定数额可以在礼物兑换界面兑换礼品，以此手段作为动力机制，一定程度上为用户了解贵州红色文化、学习长征精神起到了激励作用。

(5) 个性化设置：“我的”页面主要满足用户个性化设置的需求，用户可根据自身喜好自由切换小程序主题背景；修改个人信息；查看个人积分及排名，利用已获得的积分兑换有关奖品；用户还可以在意见反馈模块提出使用问题或建议，本团队会及时处理反馈信息，有针对性地小程序进行优化，以更好地满足用户需求。

3 小程序技术实现

3.1 前后端技术

在整体框架搭建和运行机制上，采用了具有跨平台开发能力的 uni-app 框架，结合 HTML、CSS 和 JavaScript 技术，实现了高效的前端应用开发。后端则使用 Java 语言，负责处理用户请求、数据存储和业务逻辑，确保系统的稳定运行。

3.2 音视频处理技术

在内容实现上，小程序重点运用了音视频处理技术，包括音视频剪辑与合成、音频降噪和混响、视频特效制作等，以满足用户在视觉和听觉上的个性化需求。

3.3 全景拍摄技术

全景照片拍摄技术为小程序的 3D 图片展示功能提供了重要支持，通过图像处理软件对全景照片进行后期处理，用户可以滑动手机屏幕浏览景点的各个角度，获得身临其境的体验。

3.4 数据库技术

数据库技术作为小程序的核心之一，该小程序使用云数据库存储用户信息和答题积分数据，生成实时积分排名，并收集用户反馈以优化服务。

表 1~表 3 为小程序云数据库中资源类数据表字段集合，小程序中提供的资源均存储在这些表中，包括所有的景点、视频资源信息，以及积分兑换模块的商品信息。

表 1 Attractions 表

字段	类型	说明
id	int	景点序号
name	varchar	景点名称
image_url	varchar	景点图片 URL
detail_text	longtext	景点详情(附文本)

表 2 cmscustomculture 表

字段	类型	说明
id	int	id
categoryid	int	类目 id
title	varchar	标题
context	text	内容
resource	varchar	视频或图片
iSanswerQuestions	int	1 表示是，0 表示否

表 3 wechat_resources

字段	类型	说明
id	int	资源序号
name	varchar	资源名称
path	varchar	资源路径
url	varchar	资源 URL
type	varchar	资源类型
size	int	资源大小(字节)
creator	varchar	创建者
create_time	datetime	创建时间
comments	varchar	备注

表 4 为云数据库中关于用户信息管理的数据字段集合，包括最常用的用户 ID、账号、密码以及与小程序答题模块相匹配的积分和反馈内容相关信息等字段。

表 4 云数据库用户信息设计

字段	类型	说明
user_id	bigint(20)	用户 ID
user_name	varchar(30)	用户账号
nick_name	varchar(30)	用户昵称
email	varchar(30)	用户邮箱
phonenummer	varchar(11)	手机号码
sex	char(1)	用户性别 (0 男 1 女 2 未知)
avatar	varchar(100)	头像地址
password	varchar(100)	密码
integral	int	积分
title	varchar(100)	反馈标题
context	varchar(100)	反馈内容
data	datetime	发布时间

这些技术的综合应用，使得“拾忆长征”小程序能够提供丰富、多样的学习资源，为用户带来全新的学习体验。

4 功能测试与用户反馈

4.1 功能测试

为确保“拾忆长征”小程序的功能正常、用户体验良好，本团队设计了相关测试用例，包括注册登录、页面跳转、数据加载、答题测试、反馈测试。测试结果显示该小程序页面跳转顺畅，数据加载速度较快，各模块功能均可通过测试。

4.2 用户体验

(1) 使用体验: 该微信小程序得到了体验用户的认可, 大部分用户认为小程序界面简洁明了、美观大方, 对布局、栏目和整体设计比较满意; 认为该小程序提供的信息内容丰富准确, 对该小程序的互动功能非常满意; 从流畅度来看, 用户希望该小程序在网络流畅性方面可以进一步完善; 从教育功能来看, 多数用户喜欢该小程序的路线选择模块, 并且表示会因为该小程序而增加对贵州省和长征精神的关注和了解。

(2) 改进建议: 通过体验该小程序, 大多数用户认为该小程序具有实用性和教育性, 同时也提出了小程序当前存在的问题, 如界面友好性方面, 用户注册时可以提示密码位数、加强引导设置; 内容趣味性方面, 介绍文字、视频答题模块的趣味性有待提高, 可以扩充景点介绍和学习视频的趣味性。

5 结语

“拾忆长征”以贵州省为例, 通过数字化手段让用户了解贵州省的长征旅游路线及其背后的红色精神, 不仅有助于提升贵州地区红色旅游的知名度与吸引力, 同时也为该地区的经济发展提供新的动力。该小程序的开发, 拥有无可比拟的教育价值, 承载着继承与发扬贵州长征文化的重大责任, 设计了旅游路线选择、线上参观红色景点、深入学习长征精神等功能, 为红色旅游新模式的开发做出贡献的同时, 也为用户提供一个学习红色精神平台。本团队会根据用户反馈对小程序界面的友好性进行完善。

参考文献:

- [1] 曹宇, 卑施源. 红色旅游智慧导览平台构建研究: 以“极智旅游 Lite”小程序为例[J]. 无线互联科技, 2022, 19(13): 57-59.
- [2] 吴若山. 红色旅游发展的“破”与“立”[J]. 旅游学刊, 2021, 36(6): 1-3.
- [3] 周慧玲, 张倩茹, 蔡逸伟, 等. 红色旅游对游客精神升华的影响路径研究: 以红军长征湘江战役纪念馆为例[J/OL]. 桂林理工大学学报: 1-15 [2024-01-30]. <http://kns.cnki.net/kcms/detail/45.1375.N.20240119.1733.002.html>.
- [4] 郝天齐. 龙州红色旅游爱国主义教育目的实现路径优化研究[D]. 南宁: 广西民族大学, 2023.
- [5] 宋彩虹. 文旅融合背景下山西红色旅游创新模式平台探析: 以“易游”微信小程序为例[J]. 旅游与摄影, 2022(2): 76-78.
- [6] 缪伟捷, 乔付. 湛江市红色旅游景点线上线下视听系统设计及实现[J]. 无线互联科技, 2022, 19(10): 57-59.
- [7] 鲁钺山, 刘佳宁. 互联网视野下的红色文化传播: “打卡广东红”小程序探索党史学习教育新模式[J]. 南方传媒研究, 2021(3): 18-23, 2.
- [8] 孙华林, 仇彬, 杜文帝. 基于微信小程序的党建红色微平台的设计与开发[J]. 电脑知识与技术, 2019, 15(5): 86-88.
- [9] 赵鑫宇. “互联网+”新常态下黑龙江红色旅游开发问题研究: 以第四野战军前线指挥部旧址为例[J]. 旅游纵览(下半月), 2018(10): 99.
- [10] 平欣, 宋育斌, 孔维宾, 等. 基于GPS与微信小程序的智能宠物管理系统设计[J]. 软件导刊, 2023, 22(2): 121-126.
- [11] 林远棋, 严仙荣, 丁嘉欣. 基于微信云开发的趣味游戏策划与设计: “大词吃小词”背词轻游戏[J]. 软件, 2023, 44(1): 35-38.

Design and development of a red tourism mini program for “recalling the Long March”

Shen Zhen, Li Lijuan*, Xu Yan

(School of Education Science, Jiangsu Second Normal University, Nanjing 211200, China)

Abstract: The integration of tourism and cultural industry has become the development trend of local economic market, and the development of red tourism is especially rapid. In today's information age, there are still only a few studies on the intelligent media development and promotion of the tourism resources of the Long March routes. Taking Guizhou Province as an example, this study aims to make users understand the Long March tourism routes in Guizhou Province and the red spirit behind them by digital means, and designs WeChat applets with the functions of personalized tourism routes selection, 3D attractions roaming, question-answer interactions, and points redemption, etc., with the aim of contributing to the development of a new mode of red tourism and providing a platform for users to learn about the spirit of the Long March.

Keywords: WeChat mini program; red tourism; the spirit of the Long March; Internet; panoramic camera

文章编号: 1007-1423(2024)20-0087-05

DOI: 10.3969/j.issn.1007-1423.2024.20.017

情感陪伴智能桌面萌宠系统的研究与设计

李明亮, 谢桂芳*, 刘灵丽, 李昱强

(湘南学院计算机与人工智能学院, 郴州 423000)

摘要: 对具有情感交互安慰功能的智能桌面萌宠系统进行研究, 旨在解决孤独问题。通过分析需求、界面、行为和交互, 使用 Kotlin 开发, 结合 Android 界面和人工智能。核心功能包括萌宠外观、行为模拟和智能响应, 结合机器学习和自然语言处理技术, 提供互动体验。经用户调查和实验评估, 证明系统满足期望, 提升体验。结果显示, 智能桌面萌宠系统有效缓解孤独感, 为未来系统发展提供思路和方案。

关键词: 情感陪伴; 桌面萌宠; 交互式; 人工智能

0 引言

在当今数字化时代, 人们对于情感满足和个性化的需求日益增长。可交互式桌面萌宠服务作为一种新兴娱乐形式, 为用户提供了与虚拟萌宠进行互动和陪伴的机会, 满足了人们在快节奏生活中对于陪伴和放松的渴望。然而, 要实现这样的愉快体验并非易事。本研究旨在解决桌面萌宠服务领域中所面临的技术挑战, 包括提升用户的情感体验、加强社交互动等方面的问题^[1]。

本文以情感陪伴智能桌面萌宠系统为研究对象, 通过系统设计和实现, 深入探讨相关关键技术和框架, 旨在探索提高桌面萌宠服务品质和用户体验的途径。

希望本研究的努力和创新能够为用户带来更多的乐趣, 为情感化智能陪伴领域的发展贡献力量, 以及成为数字化时代中不可或缺的一部分。

1 系统研究

1.1 可交互桌面萌宠的挑战

可交互式桌面萌宠服务领域虽然为用户展

现出广阔的市场前景, 但同时也伴随着如下一系列的需求和挑战。

(1) 用户体验

用户期待桌面萌宠服务能提供直观、流畅且富有创意的操作界面。因此, 挑战在于如何巧妙设计界面, 使之既符合用户审美, 又能提供丰富多样的互动体验, 让用户在与萌宠的互动中感受到乐趣与温暖。

(2) 萌宠行为模拟

用户渴望桌面萌宠能够展现出真实、生动的动作与行为, 并能根据用户的互动做出智能反应。这要求研究人员在算法设计和开发上不断创新, 使萌宠能够灵活应对各种情境, 展现出个性化的动作和表情。

(3) 社交互动

为了满足用户与他人分享和交流的愿望, 桌面萌宠服务应提供丰富的社交功能。这包括好友列表的建立、私聊功能的实现以及动态发布等, 旨在为用户打造一个温馨、有趣的社交空间。

(4) 安全和隐私

在提供桌面萌宠服务的过程中, 研究人员

收稿日期: 2024-04-11 修稿日期: 2024-05-30

基金项目: 湖南省大学生创新创业训练计划项目(3728)

作者简介: 李明亮(2002—), 男, 湖北潜江人, 本科生, 研究方向为软件开发、人工智能; *通信作者: 谢桂芳(1973—), 女, 湖南邵阳人, 副教授, 硕士, 研究方向为无线网络、信息处理, E-mail: 18864401@qq.com; 刘灵丽(1977—), 女, 湖南郴州人, 副教授, 硕士, 研究方向为无线网络; 李昱强(2002—), 男, 广西贵港人, 本科生, 研究方向为人工智能

必须高度重视用户的安全与隐私。通过采取先进的安全措施和加密机制，可以有效保护用户数据的安全，防止恶意攻击和数据泄露事件的发生。

综合以上需求和挑战，传统可交互式桌面萌宠服务在多个方面仍有待提升。因此，本文将通过智能化的研究和创新，有望解决这些挑战，进一步提升桌面萌宠服务的品质和用户体验。

1.2 已有的相关研究

当前桌面萌宠的研究主要集中在萌宠的交互性技术，它主要涉及到萌宠的行为模拟、音视频处理与感知、自适应学习与个性化设计、情感交互设计、社交互动与用户共享以及用户研究和评估等多个方面。高雪^[2]通过精心观察真实动物行为特征，构建算法以模拟萌宠行为，同时引入先进的音视频处理技术，增强萌宠的感知与响应能力。邓佳^[3]根据用户数据设计自适应学习算法，实现个性化互动，并通过情感交互设计增强用户与萌宠的情感连接。彭俊杰^[4]从情感的定义与分类入手，对情感分析研究需要解决的问题进行了讨论，并且刘云松等^[5]分析了软件系统交互性的发展情况，提出了提升软件系统交互性设计的思路，并对未来发展趋势进行了展望。虽然当前桌面萌宠的研究在这些方面取得了进展，但在增强用户体验、深化社交功能、提升资源效率和丰富行为真实性方面仍有挑战。本研究将针对这些领域的不足，探索创新的交互设计理念，以期提供更贴心的用户服务，带动软件系统交互性设计的新风潮。

2 系统设计

2.1 相关核心技术

(1) 自然语言处理

通过先进算法和模型深度解析文本，实现文本处理、信息提取和机器翻译等功能。语言理解涉及词汇把握、语法分析、语义探索和语境运用，而语言生成则基于指令或语义信息生成符合规范的文本，实现自然流畅的人机交互^[6]。

(2) 行为模拟与智能决策

模拟实体行为，依据环境和任务需求做出

明智决策。行为模拟基于模型和算法模拟行为模式，智能决策则基于知识和信息做出合理化选择。两者相辅相成，使实体在复杂环境中以最优方式行动，实现预定目标^[7]。

(3) 多模态交互融合

综合运用多种感知模态实现信息深度交流和高效交互^[8]。多模态交互突破传统局限，允许用户以自然多样方式与计算机交流；融合则对不同模态信息进行整合，形成全面准确理解，提升系统智能化程度和响应速度。最终目标是实现更自然、高效和智能的人机交互体验。

2.2 系统设计框架

系统设计分硬件、软件集成两部分，其中软件分为前端和后端两部分，其系统架构如图1所示。

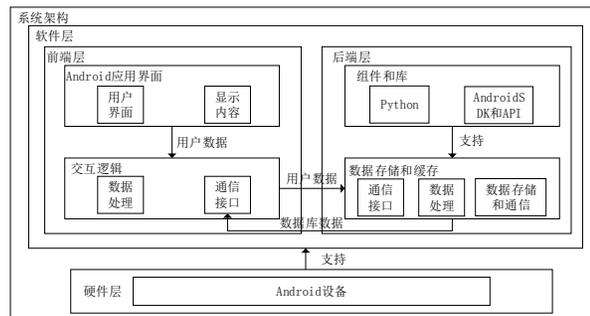


图1 系统架构图

(1) 前端层设计框架

首先，软件层前端设计在整个系统架构中占据了核心位置，它与用户界面紧密相连，直接影响到用户体验的优劣。因此，前端设计的目标主要是确保界面友好、操作便捷，同时能够高效地展示和处理数据。

在软件层前端设计中，Android应用界面是其中的重要组成部分。需要考虑到不同屏幕尺寸、分辨率和设备特性的兼容性，确保界面能够在各种Android设备上正常运行且保持良好的视觉效果。

此外，前端设计还需要关注用户显示内容的设计。这包括界面布局、颜色搭配、字体选择等方面，同时，精心布置用户界面元素(如按钮、滑块等)，为用户和萌宠之间的亲密互动提

供直观且便捷的接口^[9]。Kotlin 语言实现 API 在前端设计中扮演着重要的角色，它们负责处理前端与后端之间的数据交互，确保数据的准确传递和实时更新。

(2) 后端层设计框架

后端层是整个系统架构中的核心部分之一，负责处理数据支持、交互逻辑、数据存储和缓存等关键功能。这些功能对于确保整个系统的稳定运行和高效性能至关重要。

3 系统实现

系统实现主要划分为客户端与后端两大核心部分。客户端使用 Android，并选用 Kotlin 开发语言以确保系统界面的友好性和用户操作的便捷性。

3.1 萌宠外观设计

萌宠的外观设计融合了精心雕琢的 XML 布局文件与 Kotlin 代码的艺术性实现。借助 Android Studio 的图形界面设计利器，创建了丰富多彩的萌宠外观 XML 布局，伪代码示例如下：

```
// 萌宠外观定义的伪代码
创建萌宠外观的 XML 布局
    创建 LinearLayout
        设置宽度为整个父控件
        设置高度为自适应内容
        设置垂直方向排列
        创建 petBody 的 ImageView
            设置图片源为 pet_body
            设置宽度为自适应内容
            设置高度为自适应内容
        创建 petEyes 的 ImageView
            设置图片源为 pet_eyes
            设置宽度为自适应内容
            设置高度为自适应内容
    // 可以继续描述其他部件的定义
// Kotlin 代码中设置和更新外观的伪代码
获取 petBody 的引用
使用 custom_pet_body 图片替换 petBody 当前的图像
```

在萌宠外观设计中，使用自定义视图和控件能够更好地控制外观和交互。在 Android Studio 中，创建新的 Java 类并继承现有的视图或控件，然后重写相关方法以实现自定义功能。

3.2 行为模拟

行为模拟功能是通过结合动画框架与 Kotlin 协程的精细配合而实现的。精心设计了一系列动画序列，用以细腻地模拟萌宠的多样化行为。同时，利用 Kotlin 协程的强大功能，有效地管理了动画的播放与暂停，确保萌宠的行为表现既生动又流畅。伪代码示例如下：

```
// 定义萌宠行为动画
定义 walkAnimation 为将 petView 在 X 轴上从 0 移动到 100 的动画
设置 walkAnimation 持续时间为 1000 毫秒
// 持续时间 1 秒
// 使用协程播放动画
在生命周期范围内启动新协程 {
    启动 walkAnimation
    等待 walkAnimation 的持续时间完成 // 等待动画结束
// 接下来可以播放其他动画或进行其他行为 }
```

本研究采用了 Android 动画框架（如 Property Animation）或其他第三方动画库，以创建并控制各种动画效果，以展现萌宠的行为模拟。

3.3 智能响应机制

智能响应机制巧妙地融合了自然语言处理库与事件监听器，为用户带来了更加智能与互动的体验。运用了 TensorFlow Lite 模型进行精准的情感分析，使得萌宠能够深入理解用户的情绪。同时，借助事件监听器，精确地捕捉了用户的交互动作，实现了萌宠与用户之间的即时互动与反馈。伪代码示例如下：

```
// 初始化自然语言处理模型
定义 nlpModel 为使用 TensorFlowLite 在给定上下文中
    创建自然语言处理模型 “nlp_model.tflite”
// 定义事件监听器来响应用户交互
petView.设置触摸监听器 { _, motionEvent ->
    当 motionEvent.动作为
        当触摸动作为按下时 -> {
// 用户触摸萌宠, 可以触发萌宠的某种反应
定义 userEmotion 为分析情感函数 analyzeEmotion
    (nlpModel, userInput)
// 假设有分析情感的函数
petView.根据情感做出反应(userEmotion)
// 根据情感做出反应
返回@setOnTouchListener true }
// 其他动作处理
```

```
返回 false }
```

为了保证情感分析的高效性能和准确性，我们对 TensorFlow Lite 模型进行了集成和部署。这一过程包括模型的导入、在移动应用中的集成以及性能优化。通过对模型的轻量化化和定量化量化，我们确保了模型在移动设备上的高效运行，从而实现了对用户情感进行快速准确的分析。

3.4 交互方式

交互方式的设计充分考虑了用户的便捷性和体验性，主要通过界面组件的点击事件以及语音识别库的回调功能来实现。用户可以通过简单的点击操作与萌宠进行互动，同时，借助语音识别技术，用户的语音指令也能被准确识别并转化为相应的操作^[10]，从而实现了更加自然和智能的交互体验。伪代码示例如下：

```
// 设置点击事件监听器
petView.设置点击监听器 {
// 处理点击事件,如显示菜单或播放声音}
// 初始化语音识别
定义 speechRecognizer 为创建语音识别器
SpeechRecognizer.createSpeechRecognizer(context)
speechRecognizer.设置识别监听器(匿名对象:
RecognitionListener {
重写函数 onResults(results: Bundle?) {
定义 data 为 results.getStringArrayList(SpeechRecogn-
nizer.RESULTS_RECOGNITION)
如果 (data != null 且 data.size > 0) {
定义 userCommand 为 data[0]执行用户命令的函数
executeUserCommand(userCommand)} }
// 实现其他 RecognitionListener 回调方法 }
```

在 Android 平台，可以通过 findViewById() 方法或 View Binding 来获取界面上的组件，然后使用 setOnClickListener() 方法为组件绑定点击监听器。当用户与界面交互时，Android 系统会捕获用户的点击事件，然后触发相应的监听器方法。在处理点击事件时，可以进行 UI 更新，执行数据传递或者其他相关操作。例如，当用户点击时，可以触发数据更新操作，显示菜单、播放声音等。这些操作需要通过 Android 的 UI 线程进行，并确保不阻塞用户界面的响应。

代码逻辑示例：

```
button.setOnClickListener {
// 处理点击事件，例如显示菜单
showMenu()
}
```

3.5 隐私保护

在前后端功能实现中，特别重视数据的隐私与保护。采用加密技术对用户信息进行加密处理，并严格遵守数据保护法规。在数据传输和存储方面，运用安全协议和加密技术，确保数据的安全性。同时，设置权限管理机制，限制数据访问权限，防止滥用和泄露^[11]。总之，致力于为用户提供安全、可靠的服务，保障用户数据的隐私与安全。

4 系统评估

本文从以下两个维度对系统进行综合评估，以揭示它的优势与不足。

(1) 功能性方面，系统满足用户需求与设计规范，功能完备且易用。用户界面简洁明了，操作逻辑清晰，用户体验良好。同时，系统具备良好的跨平台与设备适应性，但在满足定制化需求方面仍有待提升^[12]。

(2) 性能方面，系统响应速度快，资源利用相对合理。在多用户同时访问的情况下，系统能稳定支持并发操作，提供流畅且高效的用户体验。但仍需在某些方面持续改进，以进一步提升整体性能^[13]。

5 结语

情感陪伴智能桌面萌宠系统的研究与设计，无疑是一个极具前瞻性和潜力的领域。随着人工智能、人机交互以及娱乐设计技术的飞速发展，此系统展现出了无比广阔的应用前景与发展空间。将继续深入研究与创新，不断提升系统的可靠性、可扩展性以及用户体验，力求为孤独人群带来更多的惊喜与快乐。同时也坚信，该系统将在数字化时代为人们带来更多的乐趣与便利，让每一个孤独的心灵都能找到属于自己的温暖陪伴。

参考文献：

- [1] 冯扬, 骆德汉. 简析 Android 语音助手开发的一种实现方案[J]. 现代计算机, 2020, 26(6): 102-108.

- [2] 高雪. 语音识别技术在人机交互中的应用研究[D]. 北京:北方工业大学,2017.
- [3] 邓佳. 基于机器学习的手势与表情识别技术及其在人机交互中的应用[D]. 杭州:浙江大学,2020.
- [4] 彭俊杰. 面向机器智能的情感分析[J/OL]. 自然杂志: 1-7 [2024-03-23]. <http://kns.cnki.net/kcms/detail/31.1418.N.20240229.0959.002.html>.
- [5] 刘云松,张尚伟,杨正琳. 软件系统的交互研究[J]. 软件,2021,42(5):42-45.
- [6] 陈伟. 人工智能在自然语言处理中的研究[J]. 信息记录材料,2023,24(10):92-94.
- [7] 曹奕. 虚拟环境中鱼群行为模拟研究[D]. 上海:上海海洋大学,2023.
- [8] 陶建华,巫英才,喻纯,等. 多模态人机交互综述[J]. 中国图象图形学报,2022,27(6):1956-1987.
- [9] 赵晓影. Android应用开发中的UI设计[J]. 劳动保障世界(理论版),2013,(12):111.
- [10] 郑国飞. 服务机器人语音交互技术研究及实现[D]. 广州:华南理工大学,2018.
- [11] MUSA S H, KRICHEN M, ALTUN A A, et al. Survey on blockchain-based data storage security for android mobile applications [J]. Sensors, 2023, 23(21): 1-4.
- [12] YAVUZ K, ALPER S. Functional test generation from UI test scenarios using reinforcement learning for android applications [J]. Software Testing, Verification and Reliability, 2020, 31(3): 2-3.
- [13] FU J, WANG Y, ZHOU Y, et al. How resource utilization influences UI responsiveness of Android software [J]. Information and Software Technology, 2022, 141: 1-3.

A study and design of emotionally supportive intelligent desktop pet system

Li Mingliang, Xie Guifang*, Liu Lingli, Li Yuqiang

(School of Computer and Artificial Intelligence, Xiangnan University, Chenzhou 423000, China)

Abstract: The intelligent desktop cute pet system with emotional interaction and comfort function was studied, aiming to solve the problem of loneliness. By analyzing requirements, interfaces, behaviors, and interactions, developed using Kotlin, combined with Android interfaces and artificial intelligence. Core features include pet appearance, behavior simulation and intelligent responses, combined with machine learning and natural language processing technologies to provide interactive experiences. Through user survey and experimental evaluation, it is proved that the system meets the expectation and improves the experience. The results show that the smart desktop pet system effectively alleviates loneliness and provides ideas and schemes for future system development.

Keywords: emotional companionship; desktop pets; interactive; artificial intelligence

文章编号: 1007-1423(2024)20-0092-05

DOI: 10.3969/j.issn.1007-1423.2024.20.018

基于云计算的医药信息集成共享平台设计与应用

范颖颖*

(河南应用技术职业学院信息工程学院, 郑州 450042)

摘要: 现有的医药信息集成共享平台往往采用较为简单的数据处理方式, 这种方式在处理小规模数据时可能有效, 但面对大规模数据时, 特别是在数据分类和整合方面, 由于缺乏智能化和自动化的有效手段, 往往导致数据分类不够精确, 整合过程中也容易出现错误。这些问题不仅影响了数据的准确性和完整性, 更使得平台的吞吐性能受到严重制约, 难以高效、顺畅地处理海量的医药数据。为解决这一问题, 提出了基于云计算的医药信息集成共享平台设计与应用方案。在该方案中, 选择 SOM-3568-SMARC 核心板作为医药信息集成共享平台的硬件装置。这款核心板以其卓越的性能和丰富的功能接口, 能够高效地采集和处理数据。在软件方面, 利用云计算技术对医药信息进行分类整合, 将具有相同属性标签的信息进行集成共享。最后应用对比实验证明所设计平台的有效性。测试结果表明, 与其他对照组相比, 文章所设计的平台在不同应用场景下均展现出更高的吞吐量。这充分证明了文章所设计的平台能够更好地适应实际应用场景, 实现对数据的高效集成与共享, 应用效果较好。

关键词: 云计算; 医药信息; 集成共享; SOM-3568-SMARC 核心板; 数据分析; 属性标签

0 引言

随着医学研究的日益深入和技术的不断革新, 新的药物、治疗方法和临床研究成果层出不穷, 使得医药信息呈现出井喷式的增长态势。这种规模庞大的信息对于医药领域的研究、开发、生产及监管等各个环节都至关重要。同时, 由于医学研究的持续进步和技术的不断创新, 医药信息需要与时俱进, 以满足日益变化的临床和市场需求。这种信息的快速更新要求医药领域的从业人员能够迅速获取、精准分析并有效应用最新的研究成果和临床数据, 以便为医疗实践提供更为精准的指导, 进一步提升医疗服务质量。基于此, 开展医药信息集成共享管理具有重要的实际意义, 其不仅可以促进医药信息的高效流通与利用, 还能为医药行业的持续健康发展提供有力支撑。

近年来, 众多学者开展相关研究, 并已经取得了一定研究成果。如文献[1]中提出了一种

基于区块链的医疗数据分级访问控制与共享平台。该平台先对医疗数据进行敏感分级及加密, 然后基于许可区块链解决集中式共享平台存在的中心化信任问题, 实现医疗数据安全共享。由于缺乏智能化和自动化手段, 该平台在进行医疗数据分类时可能无法充分考虑复杂的数据特征和变化, 导致分类不够精确。这会导致数据被错误地归类, 影响后续访问控制和共享过程的准确性。文献[2]中提出了一种基于 Spring-Cloud 框架的医疗信息共享平台设计与实现研究。该平台主要基于 Spring Cloud 微服务思路, 将医疗信息共享平台功能分为人事管理微服务(SMS)、患者管理微服务(PMS)、诊断管理微服务(DMS)和财务管理微服务(FMS)四部分, 并分别设计不同的医疗信息共享功能。在该平台中, 不同微服务对数据的处理方式可能不尽相同, 缺乏综合的智能化分类机制, 可能会导致数据分类不够准确, 造成数据冗余或不一致的情况, 影响数据完整性和准确性。文献[3]设计

收稿日期: 2024-03-19 修稿日期: 2024-04-25

基金项目: 河南省开封市科技发展计划项目(2304049)

作者简介: *通信作者: 范颖颖(1979—), 女, 河南开封人, 本科, 讲师, 研究方向为计算机应用, E-mail: fyy450155@yeah.net

一个基于 MongoDB 的信息集成平台。该平台采用 MongoDB Cluster 集群的分布式架构，以应对大规模数据存储需求，确保数据的高可用性与扩展性。通过此设计，系统不仅能够实现数据的统一管理，还能提升数据处理效率，为企业或组织提供强大的数据支持。由于该平台主要关注于处理大规模数据存储需求，但未针对数据的智能化分类和自动化处理进行充分考量，可能会影响数据分类的准确性，使得数据低效、混乱而难以管理。文献[4]设计了一种基于 ODI 技术的系统集成方案。在该方案中，设计并实现了 Get 接口和 Post 接口，用于数据的交互与传输。通过一系列实验验证，实现了从人事系统数据库抽取教师基础数据，再将其推送至共享数据中心数据库，并最终将数据同步至科研系统数据库的过程。该平台在数据交互和传输过程中，如果缺乏智能化的数据分类和校验机制，容易导致数据的一致性问题，例如信息丢失、重复传输等，降低数据的准确性和完整性。

上述提到的医疗信息集成共享平台中存在着数据分类不够精确、整合过程容易出错的问题，造成平台吞吐性能受限，难以处理大规模医药信息的挑战。基于此背景，本研究提出了一种基于云计算的医药信息集成共享平台，并通过对比测试环境下的深入分析，探讨了该平台的应用价值。

1 医药信息集成共享平台硬件设计

在医药信息集成共享平台的硬件设计环节，鉴于医药信息具有规模庞大且更新迅速的特点，需要高效的硬件作为保障。为此，本文选择金亚太 SOM-3568-SMARC 工业级核心板作为设计平台的中心装置。SOM-3568-SMARC 核心板搭载了 RK3568 四核 64 位 Cortex-A55 处理器，确保在平台运行过程中主频能达到 2.0 GHz，从而提供强大的运算能力。此外，该核心板还集成了双核心架构的 GPU (ARM G52 2EE) 和高能效 NPU (1Tops@INT8，集成高性能 AI 加速器 RKNN NPU)，能实时处理最大 8 GB 内存的数据，满足医药信息处理的复杂需求。考虑到医药信息类型的多样性，SOM-3568-SMARC 核心板还配备了多种数据输入输出接口，具体配置

见表 1，以确保数据的顺畅传输与共享。

表 1 核心板数据接口配置

接口类型	配置
视频输入	MIPI CSI (4 Lane)*1
	HDMI2.1 (4K@60Hz)*1
文本输入	MIPI-DSI*1
	LVDS*1
视频输出	I2S0 with 8 channel*1
	I2S1 with 8 channel*1
	I2S2/I2S3 with 2 channel*1
文本输出	PDM*2
	TDM*2
	Digital Audio Codec*1

此外，SOM-3568-SMARC 核心板在接口配置的基础上，还集成了 Wi-Fi 和双千兆 RGMII 高速网络功能，这使得平台能够根据实际数据请求实现实时通讯，保障在各种应用场景下都能迅速响应，从而满足实际使用中的性能需求。

2 医药信息集成共享平台软件设计

以上述硬件架构为基础，完成医药信息集成共享软件设计，具体如下：

2.1 基于云计算的医药信息深度分类

原始的医药信息规模较大、种类繁多，而且构成复杂、更新迅速。这种特性使得直接进行集成管理时需要进行大量的冗余计算，导致集成效果和效率难以达到实际应用的要求。云计算技术具有强大的计算能力和存储能力，能够处理海量的医药信息数据。通过云计算平台，可以对医药信息进行快速的分析和处理，从而实现信息的有效分类和整合。这不仅可以提高信息的处理效率和准确性，还可以促进信息的共享和协同工作，为医疗行业的发展提供有力的支持。因此，本文在具体实施集成管理之前，先利用云计算技术对原始的医药信息进行深入分析。基于数据信息的属性分析结果，为每一条信息赋予相应的标签。

在具体的执行过程中，首先对信息的类别进行初级划分，具体划分结果可表示为

$$x_i = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\} \quad (1)$$

式中： x_i 表示医药信息初始分类集合； x_1 表示药品信息值； x_2 表示价格信息值； x_3 表示库存信息值； x_4 表示销售信息值； x_5 表示客户信息值； x_6 表示供应商信息值； x_7 表示医药信息中的物流信息值； x_8 表示医药信息中的质量信息值。

根据式(1)，将相应的分类结果作为医药信息的标签信息。需要特别注意的是，一些信息可能会存在不同类别之间的交叉^[5]。为了解决这个问题，本文在设置医药信息标签时，构建了非定量结构模式，即每个数据信息对应的标签不唯一。通过这一方式，最大限度提高后续集成的可靠性和准确性^[6]。结合初级分类结果，利用云计算的方式对不同分类中的数据进行二次分类。在该阶段，本文主要根据数据的类型实现，具体计算方式可表示为

$$T_n^L(x_i) = \frac{(1 - \alpha_n(x_i))D_n(x_i)L_n}{f_n^L(x_i)} \quad (2)$$

式中： $T_n^L(x_i)$ 表示 x_i 类医药信息的二级属性识别结果； $D_n(x_i)$ 表示在 n 维裕度上， x_i 类医药信息的覆盖范畴，该信息主要体现了医药信息的内在属性； $\alpha_n(x_i)$ 表示在 n 维裕度上， x_i 类医药信息的来源范畴，该信息主要体现了医药信息的外在属性； L_n 表示在 n 维裕度上， x_i 类医药信息的大小， $f_n^L(x_i)$ 表示 x_i 类医药信息的模式参数，即其输入到平台中的形式^[7]。以此，利用云计算的方式实现对医药信息的分析，即通过构建的属性标签，以为后续的成分分享提供更有效的基础保障。

2.2 基于分类属性标签的医药信息集成共享

结合2.1部分构建的医药信息分类属性标签，本文主要依据标签的一致性与重叠关系来实现信息的集成管理。在进行信息共享管理时，则主要依赖于对共享请求的匹配结果。这两种管理方式的具体实现方式可以表示为

$$P(A) = \sum sim(x \rightarrow T_n^L(x_i)) \quad (3)$$

$$Q(A) = MAXsim[q(x_i) \rightarrow T_n^L(x_i)] \quad (4)$$

式中： $P(A)$ 表示对医药信息的集成结果； $Q(A)$ 表示对医药信息的共享结果； sim 表示相似函数； $q(x_i)$ 表示共享请求对应的属性参数^[8]。

结合式(3)和式(4)可以看出，本文在进行医药信息的集成管理时，主要基于信息属性的

相似性进行，而并非强调标签的一致性程度；而在进行信息共享管理时，同样以信息属性的相似性为基础，但更加注重标签的一致性程度，确保选择相似度最高的信息作为最终的共享输出。通过这种方式，确保所设计的平台性能输出能够充分满足实际应用的需求。

3 测试与分析

3.1 测试准备

为验证设计的医疗信息共享平台的先进性，本研究选取了文献[1]中提出的基于区块链的医疗数据分级访问控制与共享平台以及文献[2]中提出的基于SpringCloud框架的医疗信息共享平台作为对比平台，进行对比实验。实验旨在全面验证所设计的医疗信息共享平台的先进性。本研究以某医药连锁企业为具体测试对象，从其公司数据库中提取医药信息，其涵盖内容见表2。

表2 测试医药信息构成

信息类别	信息类型构成	信息来源构成
药品信息	药品名称、规格、剂型、生产厂家、批准文号	国家药品监督管理局、药品生产厂家、供应商
价格信息	采购价格、零售价格、促销价格、价格有效期	供应商、市场调研、竞争对手分析
库存信息	库存数量、库存地点、库存状态(正常/预警/缺货)	仓库管理系统、门店库存报告
销售信息	销售数量、销售额、销售渠道、销售区域	门店销售报告、电子商务平台、第三方销售渠道
客户信息	客户姓名、联系方式、购买记录、会员等级	门店客户管理系统、会员管理系统
供应商信息	供应商名称、联系方式、供应产品、合作条款	供应商合作协议、采购部门联系记录
物流信息	发货地点、收货地点、运输方式、预计到达时间	物流合作伙伴、仓库发货记录
质量信息	药品检验报告、不良反应报告、药品召回信息	质量管理部门、药品监管部门通知

结合表2所示的信息可以看出，X医药连锁公司根据自身的业务需求，对应的医药信息构成涵盖了药品信息、价格信息、库存信息、销售信息、客户信息、供应商信息、物流信息和

质量信息等关键信息类别，获取数据所占内存达到 5.3 TB。为了便于公司的日常运营、采购决策、销售分析以及客户关系管理，需要对上述信息进行集成管理。以此为基础，分别采用三个系统进行应用测试。

3.2 测试结果与分析

在对不同系统的性能进行深入分析时，本研究选取了信息集成共享操作成功率作为关键的性能评价指标，通过一系列实验测试，得到结果如图 1 所示。

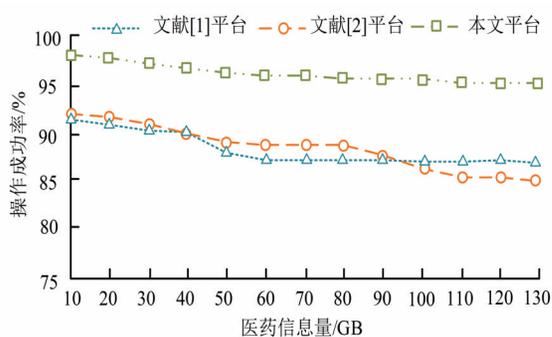


图 1 操作成功率对比

从图 1 可以清晰地看到，随着医药信息量的逐渐增长，三种平台的操作成功率均呈现出下降的趋势。特别值得一提的是，当采用本文所提出的平台进行处理时，即使在医药信息量达到 130 GB 的高负荷情况下，其操作成功率能够保持在 96.1% 的较高水平。相比之下，另外两种平台在相同的信息量下，其操作成功率均低于 93.0%，尤其是在达到 130 GB 时，操作成功率更是低于 88.0%。这一结果充分证明了本文所设计的平台在处理大规模医药信息时具有显著的优势和稳定性。本文所设计的平台主要优势在于选择 SOM-3568-SMARC 核心板作为医药信息集成共享平台的硬件装置。这款核心板以其卓越的性能和丰富的功能接口，能够高效地采集和处理大规模的医药信息。通过硬件的优化配置，平台能够保持较高的操作成功率，并展现出对高负荷情况的稳定性。

为了进一步验证平台的性能，以数据吞吐量作为另一个重要的评价指标进行了测试。如图 2 所示，三种平台在不同信息规模下的数据吞吐量对比结果一目了然。

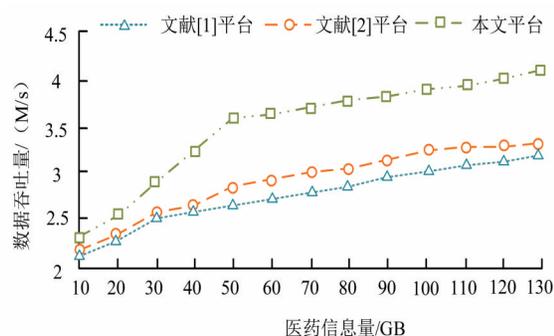


图 2 数据吞吐量对比

根据图 2 中的测试结果，可以清晰地观察到不同集成平台在业务执行阶段的数据吞吐量随着信息规模增加而逐渐提升。但是相比之下，本文所设计的平台在吞吐量方面展现出了显著的优势。特别是在处理大规模医药信息时，当信息规模达到 130 GB 时，本文设计平台的吞吐量高达 4.12 M/s，这一数值明显高于其他两种对比平台。这一结果表明，本文设计的平台在应对实际应用场景时具有更好的适应性，能够高效地实现数据的集成与共享，从而满足实际业务需求。具体分析可知，本文所设计平台的优势在于选择了基于云计算技术对医药信息进行分类整合。通过云计算技术，平台能够实现高效的数据处理和传输，提升了数据吞吐量。云计算平台可以弹性扩展资源，有效处理大规模数据的需求，同时提供高可用性和稳定性。这种技术选择为平台提供了强大的计算能力和灵活性，确保了在处理大规模医药信息时的高效性和性能表现。

4 结语

无论是从提升医疗服务质量的维度考量，还是从推动医学研究的深度与广度探索，亦或是着眼于医疗创新与患者安全保障的需求，开展医药信息集成共享管理都显得至关重要。在该背景下，本文致力于设计与研究基于云计算的医药信息集成共享平台。以 SOM-3568-SMARC 核心板为硬件基础，结合云计算技术进行数据分析和信息集成共享，构建了一个高效、先进的数据处理平台。为了验证平台的性能与效果，进行了一系列对比实验。实验结果表明，当医药信息量达到 130 GB 时，本文设计的平台在操作成

功率和吞吐量方面均表现出色。具体来说,其操作成功率高达96.1%,吞吐量可达4.12 M/s,均优于其他对比方法,充分证明了本平台的优越性能和应用效果。本文提出的基于云计算的医药信息集成共享平台,不仅为医药信息领域的发展提供了有力支持,也为提升医疗服务质量、推动医学研究和保障患者安全等方面做出了积极贡献。

参考文献:

- [1] 曹萌,余孙婕,曾辉,等.基于区块链的医疗数据分级访问控制与共享系统[J].计算机应用,2023,43(5):1518-1526.
- [2] 王蓉,李晗,周国海,等.基于SpringCloud框架的医疗信息共享平台设计与实现[J].中国医学装备,2022,19(5):133-137.
- [3] 房萍.基于MongoDB的信息集成系统的设计研究[J].信息记录材料,2023,24(10):104-106,109.
- [4] 吕忠亭,张婕,张庭姣.基于ODI技术的高职院校信息系统集成设计与实现[J].延安职业技术学院学报,2023,37(3):96-99.
- [5] 张兰华,赵鑫,王玫,等.阶梯性区域医疗卫生资源信息一体化共享集成平台的建设与对策研究[J].中国软科学,2022(S1):187-192.
- [6] 吴强,李啸,廖达俊,等.居家血糖测量数据在医院信息系统中的集成应用研究[J].中国数字医学,2022,17(11):27-31,88.
- [7] 张远林,周后签,卢熙.基于集成平台的移动护理信息系统的设计与应用[J].当代护士(下旬刊),2022,29(6):162-166.
- [8] 张丽霞,李国良,丛明.基于集成学习的电子政务信息共享平台设计[J].现代电子技术,2022,45(8):101-105.

Design and application of medical information integration and sharing platform based on cloud computing

Fan Yingying*

(College of Information Engineering, Henan Technical Institute, Zhengzhou 450042, China)

Abstract: the existing medical information integration sharing platform often adopt relatively simple data processing way, this way may be effective when dealing with small data, but in the face of large-scale data, especially in the aspect of data classification and integration, due to the lack of intelligent and automatic effective means, often lead to data classification is not accurate, also prone to errors in the process of integration. These problems not only affect the accuracy and integrity of the data, but also make the throughput performance of the platform seriously restricted, making it difficult to process massive medical data efficiently and smoothly. To solve the problem, this study proposes the design and application scheme of the medical information integration and sharing platform based on cloud computing. In this scheme, the SOM-3568-SMARC core board is selected as the hardware device of the medical information integration and sharing platform. With its excellent performance and rich functional interface, this core board can collect and process data efficiently. In terms of software, cloud computing technology is used to classify and integrate medical information, and the information with the same attribute labels is integrated and shared. Finally, comparative experiments are applied to prove the effectiveness of the designed platform. The test results show that compared with other control groups, the designed platform shows higher throughput in different application scenarios. This fully proves that the platform designed in this paper can better adapt to the actual application scenarios, realize the efficient integration and sharing of data, and have a better application effect.

Keywords: cloud computing; medical information; integration and sharing; SOM-3568-SMARC core board; data analysis; attribute label

文章编号: 1007-1423(2024)20-0097-05

DOI: 10.3969/j.issn.1007-1423.2024.20.019

基于B/S架构的在线实验系统的设计与实现

魏武华*, 黄业坤, 于豪, 李森鑫, 郝茵茵, 梁静, 刘少华

(西安文理学院信息工程学院计算机系, 西安 710065)

摘要: 针对高校现有计算机机房学生传统的上机方式存在的问题, 基于B/S架构, 设计并开发了计算机类课程在线实验系统。教师和学生使用浏览器访问部署在服务器上的在线实验系统, 通过该平台, 学生可以在线打开嵌入在网页中的实验环境, 在线进行实验并提交实验报告; 教师可以定制课程实验的内容, 在线评阅实验报告等。本系统的开发和使用, 为改进计算机类课程实验的模式探索出了一条新途径。

关键词: B/S架构; 在线实验系统; Vue; SpringBoot

0 引言

作为高等院校的实践教学中心, 计算机机房扮演着重要的角色。目前, 我国各高校设立的计算机机房已经普遍将计算机连接成了局域网, 再连入校园网进而接入互联网。在实验课上, 每个学生都配备一台个人计算机。随课程的不同, 需要在计算机上安装不同的系统软件和应用软件。

高校现有的计算机机房, 学生传统的上机方式存在以下几个方面的问题:

(1) 不同软件安装在同一台计算机上, 有时会因为兼容性问题而导致安装失败。

(2) 为了便于管理, 机房管理人员将计算机都设置为重启还原方式。这就导致有些本次没有完成的实验(如服务器配置等), 其阶段性结果无法保存, 下次实验需要重新开始。

(3) 由于硬件资源的单一性, 像大数据类课程, 在进行实验教学时遇到较大的困难。

本文探索了解决以上问题的有效途径。通

过构建计算机类课程在线实验系统, 将其部署于服务器端, 学生和教师通过终端设备进行访问。学生能够查看实验内容, 在线打开实验环境开展在线实验, 并提交实验报告; 教师能够定制实验内容、在线评阅实验报告等。

1 系统架构

在线实验系统的构建是基于B/S模式, 学生和教师通过客户端浏览器访问部署在服务器上的在线实验系统^[1-3], 系统架构如图1所示。

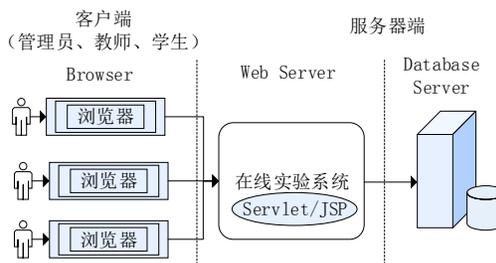


图1 在线实验系统架构

B/S被称为浏览器/服务器架构, 是随着互

收稿日期: 2024-05-05 修稿日期: 2024-06-03

基金项目: 陕西省教育科学“十四五”规划2023年度立项一般课题(SGH23Y2449)

作者简介: *通信作者: 魏武华(1968—), 女, 陕西西安人, 硕士, 教授, 主要研究方向为智能信息处理、云计算技术及应用, E-mail: 296812546@qq.com; 黄业坤(2002—), 男, 陕西安康人, 在读本科, 研究方向为软件开发; 于豪(2003—), 男, 湖南常德人, 在读本科, 研究方向为软件开发; 李森鑫(2003—), 男, 湖南耒阳人, 在读本科, 研究方向为软件开发; 郝茵茵(2003—), 女, 陕西商洛人, 在读本科, 研究方向为软件开发; 梁静(2003—), 女, 陕西汉中, 在读本科, 研究方向为软件开发; 刘少华(2002—), 男, 湖南衡阳人, 在读本科, 研究方向为软件开发

联网的兴起而发展起来的一种网络模式。在这一模式中, Web浏览器扮演着客户端的角色, 而系统的核心功能则被集中在服务器端。用户只需在终端设备上安装一个网络浏览器, 诸如Chrome、Safari、Microsoft Edge、Netscape Navigator 或 Internet Explorer(已被 Microsoft Edge 取代), 而服务器端需安装 Web 服务器软件和数据库软件, 如 SQL Server、Oracle、Sybase 或 MySQL 等, 网络浏览器与网络服务器之间进行数据交互^[4]。这种模式简化了系统的设计、维护和操作, 有效降低了客户端设备的负荷以及系统维护和升级所需的成本和工作量。

2 系统功能设计

通过深入调研和分析, 确定该系统包含三种不同的用户身份: 管理员、教师和学生^[5]。系统的总体功能设计如图2所示。

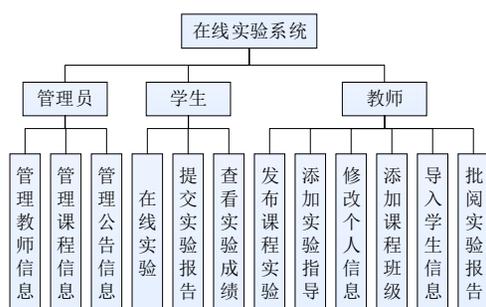


图2 系统总体功能结构图

该系统管理员功能模块包括对教师信息、实验课程信息和公告内容进行查询、添加、删除和修改的功能; 教师功能模块包括查看公告、查看和修改教师个人信息、发布和修改课程实验、上传和查看实验指导、添加课程班级并导入学生信息、查看并批阅实验报告等功能; 学生功能模块包括查看实验指导、在线实验、提交实验报告、查看实验报告成绩等功能。

3 数据库设计

该系统的数据存储在数据库表中, 主要的表结构设计如下。

(1) experiment table: 包括 id、course、ex_id、ex_name、ex_goal、ex_progress、ex_requirement、x_announcement、ex_tid 和 ex_guidwordurl 等字段。

(2) administrator table: 包括 ad_id、ad_username、ad_pwd、ad_name、ad_sex、ad_tel 等字段。

(3) announcement table: 包括 announce_id、announce_time、announce_title 和 announce_content 等字段。

(4) couclass table: 包括 id、course_id 和 exclass_id 等字段。

(5) course table: 包括 course_id、course_name、course_teacher、course_tid 和 course_imgurl 等字段。

(6) ex_class table: 包括 exclass_id、exclass_name 和 exclass_tnum 等字段。

(7) ex_guide table: 包括 guid_id、guid_title、guid_content、guid_notice、ex_number 和 course_id 等字段。

(8) stu_ex_grade table: 包括 id、stu_id、ex_id 和 grade 等字段。

(9) student table: 包括 stu_id、stu_name、stu_number、stu_pwd、stu_class、stu_sex 和 stu_college 等字段。

(10) studentuploaddocx table: 包含 id、url、date、stu_id 和 experiment_id 等字段。

(11) teacher table: 包含 tea_id、emp_number、tea_pwd、tea_age、tea_sex、tea_college、tea_name 和 tea_photourl 等字段。

(12) ex_type table: 包含 id、ex_typename 和 ex_url 等字段。

4 开发技术介绍

4.1 前后端分离技术

该系统采用前后端分离架构进行开发, 将前端用户界面和交互逻辑、后端数据处理和业务逻辑分别独立开发、部署和维护。前端负责在网页上展示内容和与用户互动, 而后端则负责处理业务逻辑和存储数据。通过 API 实现前后端间的数据交互, 这种架构模式有助于提高开发速度, 增强系统的易维护性和可扩展性。

在使用前后端分离架构时, 前端的职责在于设计和呈现页面, 程序员会利用 HTML、

CSS、JavaScript等前端工具以及可能的前端框架进行开发，如该系统基于Vue框架。而后端则致力于处理和储存数据，通常会通过RESTful API与前端进行数据交互，后端开发者使用不同的技术栈，像Java、Python、Ruby等，相应的数据库技术，以及可能的后端框架，该系统使用Java语言和MySQL数据库，基于SpringBoot框架。

4.2 Vue框架

Vue是一款备受欢迎的JavaScript框架，其独特的声明式编程模型和组件化结构，使开发者可以更高效、更便捷地设计各种用户界面。无论是简约清新的设计还是复杂多变的界面布局，Vue都能够轻松胜任^[6]。

4.3 Spring Boot框架

Spring Boot是一款备受推崇的Java应用构建框架，旨在简化开发流程、提高效率。它基于Spring框架，以自动配置和约定高于配置的原则为基础，有效减轻了程序员的工作负担。此外，Spring Boot还附带了大量实用功能和插件，如内置Web服务器、安全性、数据库访问以及缓存等，有助于开发者专注于业务逻辑。总体而言，Spring Boot是一个强大的工具，能够帮助开发者快速打造可靠、高效的Java应用^[7]。

5 系统的实现

在线实验系统采用B/S架构，以IDEA为开发软件，采用HTML+CSS+JavaScript为主要前端开发语言，完成基本页面开发^[8]，前端框架使用Vue；后端编程语言为Java，开发基于SpringBoot框架，使用MySQL数据库，包管理工具采用npm。系统部署在Centos7上，Web服务器使用Nginx，采用MySQL作为数据库服务器，系统已成功实现所有必要功能，具备良好的适应性和可扩展性。用户可以利用当今流行的各种网络浏览器，如谷歌Chrome浏览器访问部署在服务器上的系统。

5.1 登录系统

用户在浏览器地址栏输入要访问的系统的网址后，进入系统登录界面，如图3所示。



图3 登录界面

用户在该界面选择管理员、教师或学生身份后，输入正确的账号和密码，点击登录按钮，即可登录相应角色的主界面。

5.2 管理员功能界面

管理员成功登录系统后，即可访问管理员主界面，如图4所示。

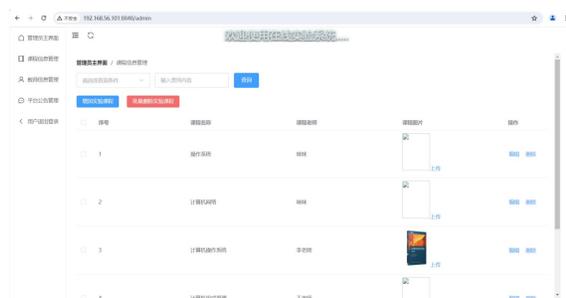


图4 管理员主界面

在该界面，管理员有权限执行添加、删除、修改和查询教师信息、实验课程信息以及公告内容等操作。

5.3 教师功能界面

教师登录系统后，便可进入教师主界面，具体展示如图5所示。



图5 教师主界面

在该界面，教师可以发布和修改课程实验内容及要求、上传和查看实验指导、导入学生信息、查看并批阅学生提交的实验报告等。

5.4 学生功能界面

学生成功登录系统后，便可进入与学生相关的页面，如图6所示。



图6 学生主界面

在该界面，学生可以选择课程实验，点击进入在线实验环境，完成实验内容，如图7所示。

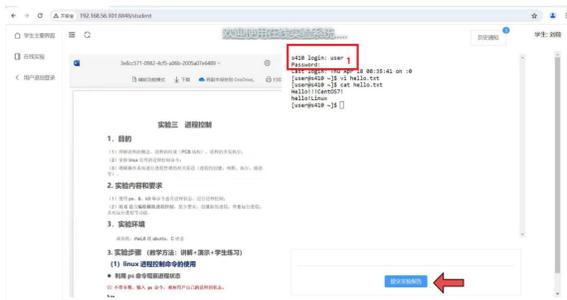


图7 在线实验界面

图7所示的在线实验页面，左边展示实验要求、右边为实验环境。学生按照实验要求，在对应的实验环境中完成实验。随后提交完成的实验报告，教师批阅。教师批阅后，学生可在如图6的界面中查看到实验成绩。

6 结语

在线实验系统的开发和应用，为高校开展计算机类课程实验提供了一种新途径。该系统基于目前流行的B/S架构，使用户只要使用浏览

器即可访问服务端的在线实验系统。该系统将在在线实验环境嵌入在网页中^[9-10]，使学生能够边查看实验要求，边在线做实验；教师能够使用该系统定制实验内容，在线批阅学生提交的实验报告。该系统的开发采用了前后端分离技术，前端基于Vue框架，而后端则采用了Spring Boot框架，所使用的数据库为MySQL。软件部署在CentOS7平台上。随着云计算技术的不断发展和应用，软件部署逐渐云端化，构建高校的私有云，将在线实验系统部署在云端将是进一步研究的课题。

参考文献:

- [1] 李梦雪,宋艳.基于数据库的数据挖掘实验平台设计与实现[J].科技创新与应用,2023(24):46-50,55
- [2] 杨力,崔江涛,黄伯虎,等.面向资源共享的计算机类虚拟在线实验建设与实践[J].计算机教育,2020(10):42-46.
- [3] 杜鹏举.基于B/S架构的高校英语远程教学系统研究[J].自动化技术与应用,2022,41(11):180-183.
- [4] 网讯.BS架构[EB/OL].[2024-04-24].<https://aistudy.baidu.com/okam/pages/article/index?categoryLv1=%E6%95%99%E8%82%B2%E5%9F%B9%E8%AE%AD&ch=54&srcid=10004&strategyId=136514066305736&source=natural>.
- [5] 潘齐心,高若铭,彭梅芳,等.Java在线实验系统的设计与开发[J].福建电脑,2019,35(5):86-87.
- [6] 什么是Vue[EB/OL].[2024-04-24].<https://cn.vuejs.org/guide/introduction.html>.
- [7] Spring boot [EB/OL].[2024-04-24].<https://docs.spring.io/spring-boot/index.html>.
- [8] 吕继东,赵松明,李龙龙,等.单片机线上实验及管理平台的设计构建[J].实验科学与技术,2023,21(6):139-145.
- [9] Xiaodongsu.Web终端之使用shellinabox在浏览器进行ssh登录[EB/OL].[2024-04-24].<https://www.jianshu.com/p/afec77178b67>.
- [10] 何洪磊.一种高性能的在线Python实验平台的研究[J].现代信息科技,2022,6(1):11-13.

(下转第105页)

文章编号: 1007-1423(2024)20-0101-05

DOI: 10.3969/j.issn.1007-1423.2024.20.020

手机端自习室预约系统的设计与实现

王 菲, 张凤涛*, 热依丽罕·图尔荪, 朱芳瑶

(长春师范大学工程学院, 长春 130032)

摘要: 针对传统自习室管理效率低下、用户预约不便的问题, 基于JavaWeb和SpringBoot数据库技术, 分析了自习室预约管理的现状和需求, 研究了相关技术实现方案。通过提出自习室预约系统的设计方案, 改进了传统管理方式, 给出了系统架构和功能模块, 设计了用户界面和数据库结构, 实现了自习室预约、查询和管理功能。结果表明, 该系统提高了自习室的使用效率和管理水平, 为人们的学习提供了更加便捷、高效的服务。

关键词: 自习室预约系统; 云计算; 互联网技术; SpringBoot框架

0 引言

现代生活节奏的快速发展、5G网络的普及, 给人们生活带来极大便利的同时也改变了人们的生活习惯。现代人不仅仅追求物质上的丰富, 更开始追求精神方面的富足, 人们对自己的内在有着高水准和高要求, 而学习往往是人们的第一选择。

在人才激烈竞争的当代, 面对当今培育人才计划的压力, 人们需要汲取更多不同领域的知识来扩充自己的知识层面, 自习室作为安静高效率的学习宝地, 在人们心中有着不可替代的地位。为此, 开发了一款基于微信小程序的自习室预约系统, 旨在解决传统预约方式的不便和繁琐。该系统让用户可以随时随地方便地预约自习室, 并实时查看座位情况和预约状态, 避免了排队和等待的麻烦。用户还可以根据个人需求选择合适的学习环境, 提高学习效率^[1]。这个自习室预约系统将为用户带来更便捷、高效的学习体验, 让他们能够更好地利用时间, 提升学习成果。

1 相关技术介绍

该系统运用了SpringBoot、Java、JSP、WXML

和MySQL技术。

SpringBoot简化了大部分基于Spring的应用开发, 在这个条件下, 开发者只需“run”就可以初步建立一个独立的、生产级别的Spring应用。多数的SpringBoot只需很少的Spring配置。例如开发者使用SpringBoot创建Java应用, 只需使用Java-jar进行启动就可以得到一个生成级别的Web工程。

Java作为一门面向对象程序设计语言。它不仅吸收了C++语言的优点, 还将C++难以理解的多继承、指针等概念摒弃、替换, 所以Java语言同时具备了功能强和操作简单两个特征。Java语言作为静态面向对象编程语的代表, 允许程序员以优雅的思维方式进行复杂的编程, 实现了面向对象理论。

采用JSP开发的项目, 通常使用HTML语言来设计和格式化静态页面内容, 而使用JSP标签来实现动态部分, 业务代码通常使用servlet、struts、springmvc等业务控制层来处理, 从而实现业务层和视图层分离。这样JSP只负责显示数据即可, 修改业务代码不会影响JSP页面代码。

WXML是框架设计的一套标签语言, 结合基础组件和事件系统可以构建出页面的结构,

收稿日期: 2024-04-22 修稿日期: 2024-05-30

基金项目: 吉林省大学生创新创业训练项目国家级(202310205019)

作者简介: 王菲(2003—), 女, 吉林长春人, 本科生, 专业为智能制造工程; *通信作者: 张凤涛(1978—), 男, 吉林长春人, 副教授, 研究方向为机械自动化技术, E-mail: 175986366@qq.com; 热依丽罕·图尔荪(2001—), 女, 新疆喀什人, 本科生, 专业为智能制造工程; 朱芳瑶(2002—), 女, 吉林公主岭人, 本科生, 专业为智能制造工程

最常用于小程序上的开发。与页面开发语言 HTML 不同, WXML 指定了界面的框架结构, 也逐渐演变成开发小程序最常用的技术。在本系统中利用 WXML 技术结合 SpringBoot 以及数据库建立的数据将基于微信小程序的自习室预约系统初步建立。

MySQL 属于 Oracle 旗下产品。MySQL 是一款安全、跨平台、高效的, 并与 PHP、Java 等主流编程语言紧密结合的数据库系统, 是当下最流行的关系型数据库管理系统之一。关系数据库将数据保存在不同的表中, 而不是将所有数据放在一个大仓库内, 由于其积小、速度快、总体拥有成本低, 尤其是开放源码这一特点, 一般中小型和大型网站的开发都选择 MySQL 作为网站数据库, 这样就增加了速度并提高了灵活性。

2 系统分析

可行性分析的目的是从多角度出发分析手机端自习室预约系统是否能进行开发, 在已有的硬件设施以及软件条件下能否支持自习室预约系统的发行与实现, 以及推出后在操作上是是否获得大众认可, 在法律上能否获得承认, 接下来主要从技术可行性、操作可行性两个方面进行介绍^[2]。

2.1 技术可行性分析

需要使用一种稳定可靠的服务器环境来支撑系统的正常运行和数据存储。其次, 需要使用合适的编程语言和开发框架, 如 Java、Python 和 SpringBoot 流行的开发框架, 以快速构建高效的系统。系统的核心功能是预约自习室, 因此需要具有对自习室信息的展示和管理功能。

为实现自习室预约功能, 需要使用数据库进行数据存储, 并实现相应的后端接口和前端页面。在处理用户账号和密码时, 需要保障安全可靠, 通常采用加密技术来加强数据的安全性。为了提升用户体验, 应用程序还需要设计友好的界面, 使用户能够通过简单直观的操作, 轻松完成预约自习室的流程。

为提高系统的稳定性和性能, 需要采用一些技术手段, 如负载均衡、缓存等技术来优化系统性能, 并使用日志和监控系统进行异常监测和预警, 发现并解决潜在的问题。

2.2 操作可行性分析

在系统开发阶段, 应注重设计简洁明了、易于操作的用户界面。用户可以通过查看自习室平面图来选择合适的自习室。系统还提供自习室详情、预约时间、预约人数等信息, 让用户更加清楚地了解想要预约的自习室的具体情况。

在进行预约时, 系统需要提供方便的预约流程, 用户只需输入相应的预约信息即可完成自习室预约操作。在进行预约时也应做一些基本的校验, 防止用户输入错误或数据不合法。如果用户由于种种原因需要取消预约, 则同样可以通过系统操作取消预约, 让用户的操作更加灵活高效。

在系统整体运行过程中, 需要保证系统的稳定性和响应速度。系统需要采用合理的技术手段来保证服务器稳定, 并能够处理大量信息并发请求。此外, 还需要设置日志系统和监控系统, 以便管理人员随时了解系统的运行情况, 及时发现并解决问题。

3 系统功能详细设计

自习室预约系统主要分为两个功能模块: 用户模块和管理员模块。用户登录前端自习室预约小程序进行课程预约, 管理员登录后端管理前端的用户以及自习室预约情况等信息^[3-5]。

3.1 用户模块

用户在微信小程序上进行相关操作, 前端用户功能具体如图 1 所示。

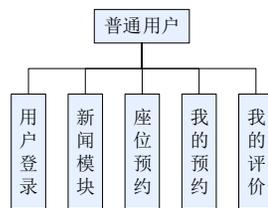


图 1 普通用户模块图

(1) 用户登录: 如初次使用自习室预约小程序, 要先进行账号注册, 其中账号注册位数为 12 位, 密码采用加密方式显示, 提高系统和用户信息的安全性。只有合法的用户才能登录此系统, 进行相关操作, 在登录时可直接和微信用户进行绑定, 规范收集用户个人信息并保

密，预约平台的用户在同意相关协议后，即可开始使用本预约平台的相关功能。

(2) 新闻模块：用户可以在此模块中查看管理员所发布的内容，其中包括自习室信息的更新和自习室信息的更改，用户也可以在新闻模块进行搜索关键词来获取信息。

(3) 座位预约：用户可以通过提供的自习室平面图来了解上座情况，预约系统用不同的符号标注自习室座位的预约情况，如某一个时间段上该位置的人员预留情况，来选择自习室，同时进行座位选择。

(4) 我的预约：此模块显示该用户预约的全部历史记录，可以显示每次预约的详细信息。如果用户需要修改自习时间，可以在此模块选择修改预约时间完成操作；如用户因特殊情况需取消预约，可以在此模块进行操作。但为了避免自习室资源浪费，用户需要在预约时间前十小时内取消预约。

(5) 我的评价：用户可在学习结束后对自习室预约平台发布评价，与他人交流想法。在此功能模块，可以查看、删除、修改自己所发布的评价。

3.2 管理员模块

管理员在浏览器网页上对用户和自习室进行管理操作，后端管理员功能具体如图2所示。

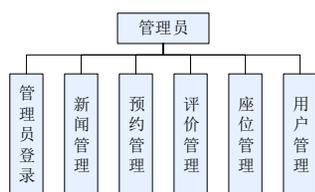


图2 管理员模块图

(1) 管理员登录：在浏览器在地址栏输入预约管理系统网址，进入自习室预约平台后端。输入正确的管理员账号和密码，输入完毕后即可进入管理员后端；输入错误将提示“用户不存在，重新输入”字样。

(2) 新闻管理：发布关于自习室信息，包括营业时间、预约情况、是否有无座位等信息。

(3) 预约管理：查看用户预约的详细信息，包括预约座位信息、预约时间信息、预约日期

信息等。

(4) 评价管理：为营造良好的自习室学习氛围，将对用户所发表的评价进行审核，对于一些影响公共秩序的评价采取删除，严重者对其拉黑操作。符合标准的良好言论将通过审核，发布到平台上，供用户交流。

(5) 座位管理：在此模块管理员可以查看自习室每个座位的状态，是否有人预约情况。

(6) 用户管理：此模块可以显示注册用户的基本信息，包括用户名、昵称、联系方式、用户简介、注册时间、历史预约。通过用户信息对用户进行检索，也可以对用户信息进行添加、删除和修改。

4 系统实现

4.1 开发工具

自习室预约系统是一个综合性的服务平台，它集成了服务管理、服务监控、数据管理和数据分析等多种功能，旨在提供全面、高效、便捷的服务体验。平台层次架构分为三个层次：控制层、业务层和持久层。

这种架构设计使得平台能够更好地进行控制、处理业务逻辑和持久化数据。控制层在平台层次框架中负责集中控制和调整；业务层是系统框架中的核心部分，它主要负责处理具体的业务逻辑；持久层是指将业务数据存储到磁盘，并具备长期存储的功能。控制层通过参数和对象，利用 Ajax 异步处理方式进行数据交互，然后将请求传递到业务层(server层)，业务层接收到请求后，通过代码来实现业务功能，并将请求参数传递持久层(DAO层)进行处理。持久层采用 MyBatis 作为 ORM 框架实现数据库访问，通过 SQL 语句增删查改等功能，并将结果返回给 DAO 层，DAO 层再返回到业务层，业务层进一步处理后将结果通过 Ajax 技术传递给控制层，最终将数据展示在用户界面上^[6]。

4.2 系统的软件环境

(1) 用户端

操作系统：Windows 11 或更新版本；

浏览器：Chrome、Firefox。

(2) 应用程序服务端

硬件环境：128 GB 内存，4 千兆电口；

操作系统: Windows 11或更新版本;
应用服务器: SpringBoot 或更新版本;
数据库访问: JDBC。

(3) 数据库服务器端

硬件环境: 64/128 G内存, 4万兆光口;
操作系统: Windows 11或更新版本;
数据库系统: Oracle 5g。

4.3 用户端安装与配置

首先需要下载微信开发者工具, 其次安装微信开发者工具, 然后创建测试账户, 接着进行前端页面设计, 最后完成编译与执行。

4.4 管理员端安装与配置

在服务器上安装微信开发者工具和 IntelliJ IDEA 后, 安装 JDK, 再安装 Navicat Premium5、MySQL 5g 数据库, 连接此数据库, 新建自习室预约系统数据库, 导入用户、座位和管理员等数据, 并在 MySQL 中建立自习室预约系统的数据表。接着在 IntelliJ IDEA 中的 Web 应用程序中创建后端基本信息, 最后输入管理员名称和密码后, 进入管理员后端, 进行管理员密码修改、新闻管理、座位管理和用户管理等操作。

5 系统测试

5.1 测试方法

应用黑盒测试将程序当作一个看不见内部的黑盒, 在完全不考虑程序内部结构和处理过程的情况下, 测试者仅依据程序功能的需求规范考虑, 确定测试用例和推断测试结果的正确性。

应用白盒测试将程序当作一个透明的盒子, 测试者可以清楚了解盒子内部的东西以及内部是如何运作的, 按照程序内部的逻辑测试程序。

5.2 测试结果

5.2.1 管理员登录测试用例

- (1) 测试项目: 测试管理员登录功能;
- (2) 测试标题: 非本系统中含有的账户;
- (3) 重要级别: 高;
- (4) 预置条件: 项目正常运行, 并且可以访问到系统登录页面;
- (5) 输入: 输入一个本系统账号和密码;

(6) 预期输出: 提示用户或密码错误, 请重试;

(7) 实测结果: 与预期效果相同。

5.2.2 用户预约测试用例

- (1) 测试项目: 测试用户预约测试功能;
- (2) 测试标题: 测试用户预约;
- (3) 重要级别: 高;
- (4) 预置条件: 项目正常运行, 并且可以访问到系统登录页面;
- (5) 输入: 输入系统中用户的账号密码;
- (6) 预期输出: 预约成功;
- (7) 实测结果: 与预期效果相同。

6 结语

商业化自习室设立的目的是方便社会人群闲暇之余和在校学生课余之时进行自我提升, 为学习者提供空间和服务, 满足用户的需求, 把用户体验感放在首位。自习室能够充分营造学习氛围, 学习者能够受其自习室氛围影响, 自觉或不自觉做出与室内其他学习者相一致的态度或行为, 随着互联网技术的不断进步, 在数据爆炸的时代, 人们对于能在空余时间学习的需求日益突出, 尤其是国家提倡的“终身学习”理念, 让学习者对于寻找安静方便的学习场所变得更为重要。

参考文献:

- [1] 马君. 基于微信小程序的微课堂设计与应用[J]. 西北民族大学学报(自然科学版), 2019, 40(3): 9-14.
- [2] 王超斌. 基于 Android 的移动点菜系统的研究与实现[D]. 昆明: 昆明理工大学, 2013.
- [3] 林琴, 王庆伟. 基于 JavaWeb 的图书馆座位预约系统设计及实现[J]. 电脑知识与技术, 2019, 15(5): 78-79.
- [4] 李霞. 基于微信小程序的计算机基础测试系统[J]. 现代计算机, 2021, 27(5): 116-120.
- [5] 刘绍婕. 基于微信小程序的考勤签到系统的设计与实现[J]. 信息与电脑(理论版), 2021, 33(7): 158-162.
- [6] 柏超宇, 顾怡, 杨丽雯, 等. 智慧校园微信小程序云服务开发与构建[J]. 电子技术与软件工程, 2018(19): 72-73.

Design and implementation of self-study room reservation system on mobile phone

Wang Fei, Zhang Fengtao*, Jerihan Tursun, Zhu Fangyao

(College of Engineering, Changchun Normal University, Changchun 130032, China)

Abstract: Based on JavaWeb and SpringBoot database technology, this paper conducts an analysis of the current status and requirements of self-study room reservation management. Furthermore, it investigates the implementation scheme related to the pertinent technology. By proposing a design scheme for the study room reservation system, we aim to enhance the traditional management approach. The system architecture and functional modules are outlined, along with the design of the user interface and database structure. The system is equipped with functions for study room reservation, inquiry, and management.

Keywords: study room reservation system; cloud computing; Internet technology; SpringBoot framework

(上接第100页)

Design and implementation of online experiment system based on B/S architecture

Wei Wuhua*, Huang Yekun, Yu Hao, Li Senxin, Hao Yinyin, Liang Jing, Liu Shaohua

(Computer Department of Information Engineering College, Xi'an University, Xi'an 710065, China)

Abstract: This paper addresses the issues present in traditional computer rooms for college students and proposes the design and development of an online experiment system for computer courses based on B/S architecture. Both teachers and students can access the online experiment platform deployed on the server using web browsers. Through this platform, students are able to open the embedded experiment environment within a web page, conduct experiments online, and submit their reports. Meanwhile, teachers have the ability to customize course experimental content and review student reports through the same platform. The development and use of this system has paved a new way to enhance the model of computer course experiments.

Keywords: B/S architecture; online experiment system; Vue; Spring Boot

文章编号: 1007-1423(2024)20-0106-04

DOI: 10.3969/j.issn.1007-1423.2024.20.021

智能疫苗接种系统的设计

张琳婧, 李庆玲*

(济宁医学院医学信息工程学院, 日照 276826)

摘要: 在对疫苗接种现状进行调研的基础上, 构建智能疫苗接种系统, 对平台构架及其具体功能展开研究。利用数据挖掘技术分析用户所需疫苗种类, 提供注射疫苗前后注意事项, 发现注射后不良反应规律, 提高平台预约疫苗注射成功率, 解决用户预约难、排队时间长等相关问题。

关键词: 智能; 疫苗接种; 预约机制

0 引言

随着全球环境的恶化和人口数量的增加, 传染病的发生率呈上升趋势。尤其最近几年新冠疫情的泛滥, 使得如何有效控制流行性病毒的传播以及保护弱势群体不被感染成为全民关注的话题。注射疫苗是预防传染病最有效的措施^[1]。为了有效控制注射地点人流量, 降低传染病的传播速度, 更好地管理疫苗, 提高疫苗利用率^[2], 线上预约疫苗接种系统开始出现。目前有多平台疫苗查询管理和疫苗知识推广系统, 能提高疾控中心对疫苗的管理水平, 引导民众正确认识疫苗接种与预防疾病^[3], 但是预约机制和信誉机制不够完善。InterSystems 推出新冠肺炎疫苗接种自动预约解决方案, 为英国国家医疗服务体系节省了 500 多个小时的电话呼叫时间, 该系统可以确保在规定时间内完成第二次接种预约并发送提醒, 但是忽略了接种人的突发且必要的行程安排等问题^[4]。其它在线系统存在预约机制单一、无法高效解决用户因出现突发情况而未按时注射疫苗等问题。系统仅具备疫苗接种登记功能, 没有有效疫苗接种接管、溯源跟踪等功能^[5-6]。

为改善现有线上预约疫苗系统的功能, 开发一款智能化的疫苗预约接种系统是十分必要的。通过系统的使用, 可以使疫苗注射机构能够及时准确地掌握疫苗接种者的相关信息, 便于疫苗的监管及追溯, 拥有更加明确的解决措施方案。在新的疫苗预约机制以及信誉机制下, 用户节省预约的时间成本等。另外, 动态的疫苗预约和注射数量, 利于疫苗生产企业及时调整生产规模和数量。

1 系统设计

1.1 系统架构

平台由数据采集、数据传输和数据服务器三部分组成, 如图 1 所示。

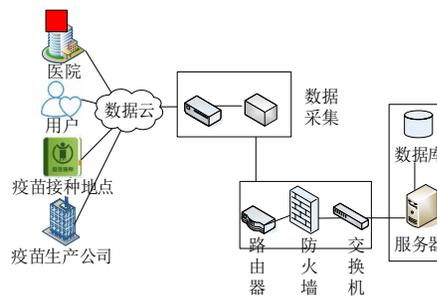


图 1 平台架构图

收稿日期: 2023-12-20 修稿日期: 2024-04-16

基金项目: 济宁医学院大学生创新创业项目(cx2022088)

作者简介: 张琳婧(2002—), 女, 山东青岛人, 在读本科, 研究方向为智慧医疗; *通信作者: 李庆玲(1979—), 女, 山东日照人, 硕士研究生, 副教授, 研究方向为计算机教育应用、智慧医疗等, E-mail: ling2001er@163.com

1.1.1 数据采集

数据采集部分负责将用户查询关注疫苗的数据、不良反应提供的描述进行采集，并将数据汇总上传至服务器。用户通过手机网络访问该系统，实现线上疫苗预约以及咨询。

1.1.2 数据传输

将原始的有效信息封装，通过传输介质发送至路由器，分析IP数据包后重新封装，路由表将数据转发至服务器。

1.1.3 数据服务器

服务器接收到数据后对信息进行有效检验，信息不符合常理将其抛弃，有效信息达标将正确封装写入数据库。数据封装图如图2所示^[7]。

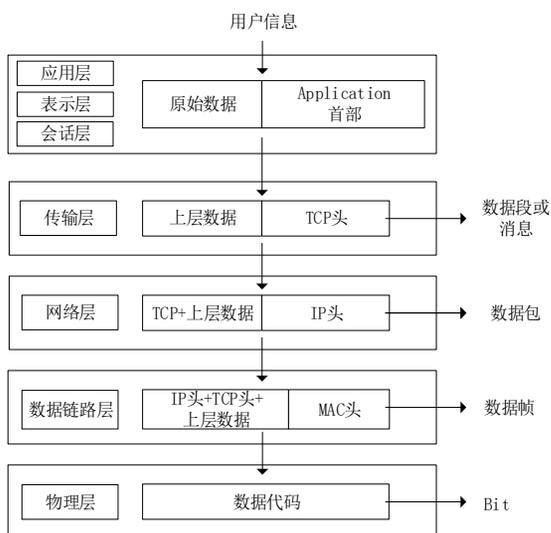


图2 数据封装图

1.1.4 数据共享

医院、接种点、疫苗生产公司和用户拥有访问除隐私外的信息权限。医院、接种点、疫苗生产公司固定的管理人员拥有数据完善权限。该部分信息共享基于区块链技术，实现用户不良反应及具体情况的公开透明。

1.2 数据库设计

数据库的设计从基本信息、服务需求、健康档案和疫苗信息四个部分进行。

1.2.1 基本信息

对用户基本情况信息进行登记，如姓名、年龄、性别、联系方式、紧急联系人等。

1.2.2 服务需求

记录用户所关注的疫苗生产、库存情况，以及周围各注射点距离，预测关注疫苗成功预约的时间。若家中有小孩，则及时推送注射疫苗年龄表，若有老年人重点推送注射疫苗副作用与年龄之间的关系以及注射疫苗前后的注意事项。

1.2.3 健康档案

为用户及其家人建立云健康档案，记录用户所注射的疫苗及时间，手术内容、用药情况、基础疾病等，以及药物过敏史，尤其是最近的生病情况及用药情况等相关信息。

1.2.4 疫苗信息

记录疫苗生产公司、疫苗运输位置、疫苗编号、生产日期以及疫苗在各个注射点的存储情况，并记录可能发生的注射后的不良反应以及应对措施。

2 系统详细设计

2.1 用户端功能设计

为了充分利用医疗资源，更好地为患者提供距离近、人少并有疫苗存货的注射点，开发一套能提供多样化医疗服务的智能疫苗接种系统具有一定的实际意义。根据智能疫苗接种系统的研究背景、意义以及现状，深入分析了系统的业务流程，最终设计一套集疫苗信息查询、接种档案查询、问题反馈和信誉机制功能为一体的智能疫苗接种系统。如图3所示。

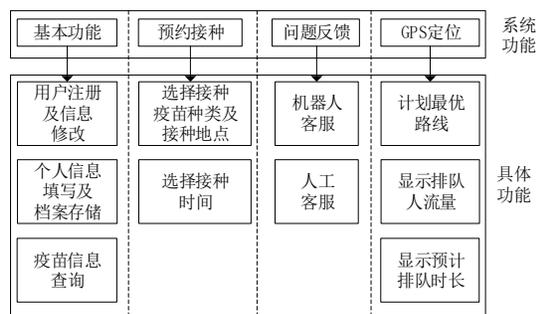


图3 功能模块图

2.1.1 基本功能

包括用户信息及档案的存储、个人信息修

改、疫苗信息查询等。系统会定期备份用户的数据并实时更新，防止用户的数据丢失，保证了系统的稳定性。在系统中用户可以查询疫苗的功能、注意事项及存货量，并对疫苗的存货量进行实时更新，便于用户进行疫苗注射的预约^[8]。

2.1.2 接种预约

用户可提前对想要接种的疫苗进行预约，通过对疫苗有关信息的查询以及接种点的选择后即可预约。预约时需要在预约窗口处填写接种预约表，如果计划有变，可在系统中取消预约，等待系统判定成功即可，但是会计入信誉机制。当用户出现无理由失约时，系统会根据实际情况扣信誉分。当信誉分下降到一定值时，将会出现预约冻结或者疫苗预约排序延后等情况。如图4所示。

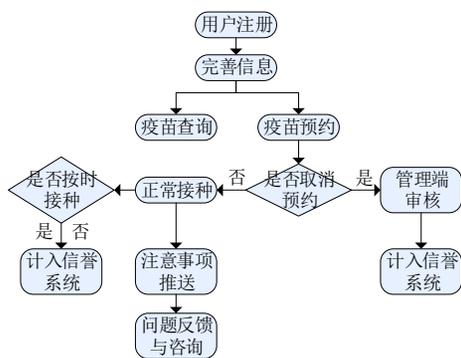


图4 用户预约流程图

2.1.3 问题反馈

本系统设置了机器人客服和人工客服两种形式，机器人客服会根据用户提问的较官方问题进行标准系统性的回答，并且24小时开放；如果用户的问题比较实际，机器人客服无法满足，用户可以选择转为人工客服进行咨询，会有专门的工作人员为用户解答疑问。通过标签匹配的方法分析患者的询问意图，再使用算法对患者的问题进行反馈。当用户注射完疫苗后出现不良反应时，及时对用户询问的不良反应做出回复，避免用户身体健康出现巨大问题。

2.1.4 GPS定位系统

在紧急情况下，用户可以通过信息查询索引表进行查询，并预选接种中心集合的信息。GPS定位系统会自动启动为用户提供相关数据，让用户了解周围医院及接种点中现有疫苗的数量

和到达接种地所需的时间，并且会显示接种地人流量情况和预测排队时间。该模块精准地为用户推荐最近并且人少、有疫苗的注射点，能给予用户最优推荐。

2.2 管理端功能设计

该模块包括用户信息管理、疫苗信息管理、预约接种管理、接种咨询管理模块。该部分由疫苗生产商管理层、医院以及疫苗注射点共同管理与使用。如图5所示。

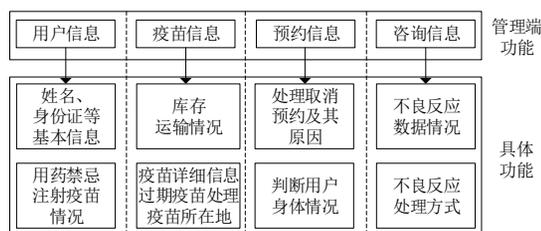


图5 管理端功能设计

2.2.1 用户管理模块

用户管理模块可查看用户在注册平台账号时填写的相关信息，包括姓名、身份证号、过往疫苗注射情况，以及填写病史(尤其是过敏史)、用药禁忌等。这些数据可以用来分析用户近期是否适合接种所预约疫苗，保障其身体的健康以及预测注射后的不良反应在可控的范围之内。同时接种完之后，用户可以填写注射后的不良反应和发病记录，这部分数据可以让注射点分析不良反应发生剧烈时的必要条件。并且以年龄为基础，根据用户所提供的信息将人群进行分类，建立完善的用户疫苗接种数据系统。

2.2.2 疫苗管理模块

该部分主要的功能为当用户关注或者预约注射某一疫苗时，会提醒用户疫苗存货量有多少；若无存货，会及时告知用户疫苗的运送情况和预计送达时间；若其他临近注射点有存货，会及时推荐给用户，方便用户申请预约。接种点可根据实际情况对系统中的疫苗数量进行修改，疫苗现有量实时更新，方便用户查询预约。该模块也与各医院及接种点建立连接，可以对运输至接种点的疫苗进行批量录入，方便疫苗数量等信息的导入，以便于计算设置疫苗预约的种类以及每日预约注射疫苗的人次。接种点

的疫苗难免会出现剩余过期的情况，过期疫苗需要通过正规途径处理掉。过期疫苗由专门负责的人员进行统计管理，避免出现过期疫苗被预约而无法接种的情况^[8]。

2.2.3 接种管理模块

接种管理模块包含预约管理与信息咨询两部分功能。疫苗预约需要在用户完善完过敏史、基础疾病记录的情况下进行预约。系统会根据用户所填写的信息进行统计分析，以疫苗接种明细表和接种预约明细表的形式出现，并存入用户个人信息库中。当用户将预约信息上传到平台后，系统会将用户的预约信息发送给用户所预约的接种单位，接种单位会通过用户的数据判断该用户预约的时间段身体状况是否能够接种疫苗，并对平台进行反馈。如果合适，该单位将接受预约，并且平台会提醒用户预约成功；如果不合适，平台将会根据反馈给用户发送信息和医生建议，提供给用户原因以及调理方法。当用户申请取消预约时，系统需要在24 h内反馈用户该申请是否同意，并将相关记录计入信誉系统。专门的管理员账号也可以登录增加数据或修改数据。

用户接种完成后，系统会实时推送注射后可能出现的不良反应以及处理方式，使用户感到更加安全与放心。若有疑问可以在公众号选择问题分析功能，进行机器人或者人工进行询问并得到解答。

3 结语

随着生活节奏的加快，病毒也在不断进化，人类感染病毒的几率也在大幅上升，一款有效的疫苗接种系统利于疫苗接种接管、溯源跟踪，

方便接种点控制人流量的同时，满足大部分用户疫苗接种需求。本文通过智慧医疗系统的应用，创建一个智能化的预约疫苗系统，提高人们注射疫苗的速度，减轻某些接种点因人过多而疫苗数量不足的压力。通过设计用户端，用户可以实现在线填写自身信息、在线问询、查询附近接种点及各地疫苗存货、在线预约疫苗等功能。通过用户端实现在线预约等功能，力求实现一个接种点与患者交互的、实时的、自动的、便利的疫苗接种系统，保证疫苗接种工作更加科学、规范。

参考文献：

- [1] 陈盼,杨剑,肖永红. 新型冠状病毒疫苗的研究进展[J/OL]. 医药导报: 1-18[2023-11-17]. <http://kns.cnki.net/kcms/detail/42.1293.R.20230927.1046.002.html>.
- [2] 陈煜,李园园,张欣羽. 新冠疫苗预约接种系统的设计与实现[J]. 无线互联科技,2022,19(13):86-88.
- [3] 华伟,杨菲,韩斌. 多平台疫苗查询管理和疫苗知识推广系统的开发与应用[J]. 电子技术与软件工程,2018(19):73-74.
- [4] 王铎霖,刘钧玉,陈新,等. 疫苗接种全流程管理系统的设计与应用[J]. 中国数字医学,2022,17(12):72-76.
- [5] 陈宇斌. 疫苗接种管理系统的设计与实现[D]. 厦门:厦门大学,2014.
- [6] 王忠周. 千佛山医院疫苗接种管理系统的设计与实现[D]. 济南:山东大学,2012.
- [7] 马金凤,林坤,李佳琦,等. 社区医院疫苗接种管理系统设计[J]. 福建电脑,2022,38(7):59-65.
- [8] 吴秀芹,徐义霞. 人用狂犬病疫苗接种管理系统的设计与应用[J]. 中国医疗设备,2016,31(8):110-112.

Design of intelligent vaccination system

Zhang Linjing, Li Qingling*

(School of Medical Information Engineering, Jining Medical College, Rizhao 276826, China)

Abstract: Based on the investigation of the current situation of vaccination, an intelligent vaccination system is constructed, and the platform architecture and its specific functions are studied. Data mining technology is used to analyze the types of vaccines required by users, provide precautions before and after vaccine injection, find the law of adverse reactions after injection, improve the success rate of vaccine injection on the platform, and solve related problems such as difficult appointment and long queuing time for users.

Keywords: intelligent; vaccination; appointment mechanism

文章编号: 1007-1423(2024)20-0110-04

DOI: 10.3969/j.issn.1007-1423.2024.20.022

基于自动跟随的智能购物车设计与研究

李文豪*

(郑州科技学院信息工程学院, 郑州 450064)

摘要: 针对传统的零售行业由于购物车而导致的运转缓慢这一问题, 提出了一种智能化购物车方案, 以提升顾客的消费体验。国内对智能购物车的设计起步较晚, 且由于成本以及设计缺陷等问题智能购物车始终没有普及, 因此设计一种新的基于STM32跟随算法, 在此基础上设计一款能够实现顾客跟随、自主归位、自动避障等功能的智能购物车, 保证低成本的同时提高泛用性。这样的购物车既能节约顾客的时间, 也能优化购物体验, 从而进一步推动商家的营业收入的增长。

关键词: STM32; 智能购物车; 自动跟随

0 引言

人们的大部分生活用品源于超市, 同时超市也成为人们休闲、放松的场所之一。随着时代的进步与发展, 人们对生活中的各种体验有了更高的要求。超市客流量逐日上升的同时, 也伴随着不少问题的出现。例如: 在购物较多时推动购物车需要耗费很大力气; 在推拉购物车的同时不能很好地照顾孩子等。在电商平台发展越来越完善的时代背景下, 这些问题无疑会消磨消费者的购物耐心, 会使超市的客流量减少, 最终影响超市的盈利和长远的发展。因此研发智能购物车对于提高超市服务水平和顾客购物体验具有重要意义^[1]。

在国内, 智能购物车的研究大多集中在商品识别和商品结算, 对于跟随式购物车研究较浅显, 大多只停留在表面。本文设计了一种基于STM32的超声波跟踪算法, 在此基础上设计一款具有跟随功能和归位功能的智能购物车, 保证购物车能够做到路线跟踪式的跟随模式。本文设计的智能购物车成本低廉、泛用性高, 国内各类零售行业都能用的上、用的起。

1 硬件需求

1.1 STM32 单片机

智能购物车设计中, 我们将STM32单片机用作核心计算器件, 负责接收外部信息并对其进行管理与操控。STM32是一款由意法半导体(ST)公司推出的32位微控制器。其内核基于Cortex-M3, 采用LQFP48封装, 属于ST公司微控制器系列中的STM32产品之一。该微控制器融合了强大的处理能力、丰富的外设以及广泛的应用领域支持, 特别适用于工业控制、消费电子、通信和汽车等领域。STM32的系统结构可以分为ARM设计的Cortex-M3内核和ST在此基础上优化的总线矩阵、DMA、AHB总线, 以及APB1和APB2接口上的外设等组成部分^[2]。

1.2 超声波传感器

所选用的超声波模块为市面上常见的传感器, 可在网上购买, 其接收范围为4~500 cm, 工作电压为5 V, 采用了串口通信的工作方式, 发送频率为50 Hz。这种超声波传感器可以快速测量, 且便捷、易操作。超声波传感器相对于其他距离测量技术(如激光雷达)来说成本较低,

收稿日期: 2024-03-09 修稿日期: 2024-05-03

基金项目: 郑州科技学院2023年大学生创新创业训练计划项目(DC202353)

作者简介: *通信作者: 李文豪(2003—), 男, 河南周口人, 在读本科, 研究方向为计算机科学与技术、JavaWeb前端开发, E-mail: 1748506890@qq.com

这使它在智能车辆中得到广泛应用。低成本的特点使超声波模块更容易被集成到智能车辆系统中，以提供更好的安全性和便利性。

1.3 红外线避障传感器

红外传感器，也被称为红外避障传感器，是由两个红外管组成，一个用于发送红外光线，另一个用于接收红外光线。这种传感器具有很强的环境适应性和较小的干扰，并且安装起来非常方便。它主要包括一对红外光线发射管和一个红外光线接收管以及相关的传感器电路。本文采用的是红外线避障传感器中的直式避障传感器模块，它的距离检测范围通常在几厘米到几米之间，电压一般为3.3~5 V。红外线避障传感器通常具有低功耗的特点，能够在长时间的使用中保持稳定性能。这对于智能车辆来说非常重要，可以延长电池寿命，并减少能源消耗。

1.4 TCRT500 循迹传感器

在TCRT5000传感器的内部，红外发射二极管始终处于工作模式并不断释放出红外光波。如果未有任何红外光波返回或者其反射程度不足以触发红外接收管开启，那么该接收管会维持关断的状态，从而使输出的信号是高电压且二极管停止发亮。然而，若存在某个对象进入了探测范围并且成功地反射了足够的红外光波，使红外接收管达到了饱和点，则它将会切换至开机状态并将输出信号转换成低电压，同时让二极管恢复到正常的工作状态。

2 技术解析

2.1 对波式超声波

在跟随购物车框架设计上本文采用的是对波式超声波，相对于市场上常见的OpenCV机器视觉跟随模式，对波式超声波价格较低且对顾客的追随准确率更高，通常为100%。

超声波传感器不具有识别功能，那么如何让购物车跟随指定的顾客呢？只需让用户携带一个超声波传感器实现对波式超声波就能实现。对波式超声波是指将两个超声波相对放置，分别为发射端和接收端，可根据不同的应用场景来设置用法。

在购物车上的应用如图1所示，购物车A上的超声波为发射端、顾客B携带的超声波为接收端，两个超声波通过蓝牙进行信号传输。购物车运行后发射端发出信号波，信号波的识别情况一共分为三种情况：①信号波没有识别到物体且被接收端接受，购物车启动跟随算法；②信号波在被接受之前检测到物体，启动避障模块；③信号波没有识别到物体但没有被接收端接收，购物车发出未检测到接收端的提示。

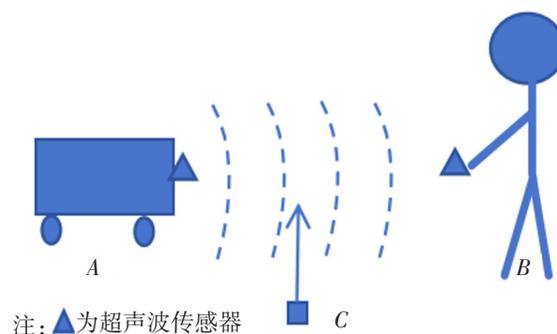


图1 对波式超声波

2.2 三边定位算法

在定位上本文采用超声波三边定位算法。参照图2，假设购物车的中心坐标为 $(0, 0, h)$ ，在购物车上放置了三个不同点的超声波传感器，坐标分别为 $A(x_a, y_a)$ ， $B(x_b, y_b)$ ， $C(x_c, y_c)$ ，定位目标的坐标为 $P(x_p, y_p, H)$ ，其中 h 和 H 分别为超声波接收端和定位目标上超声波发射端相对于地面的高度。

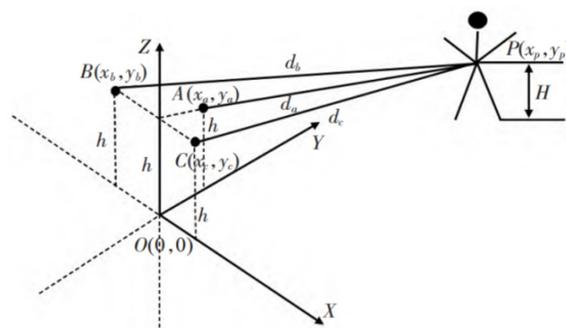


图2 三边定位算法

由超声波测距原理可得出接收端相距移动目标的距离分别为 d_a ， d_b ， d_c ，根据三边定位算

法原理，得出跟随目标的位置坐标计算公式，如式(1)：

$$\begin{cases} (x_p - x_a)^2 + (y_p - y_a)^2 + (H - h)^2 = d_a^2 \\ (x_p - x_b)^2 + (y_p - y_b)^2 + (H - h)^2 = d_b^2 \\ (x_p - x_c)^2 + (y_p - y_c)^2 + (H - h)^2 = d_c^2 \end{cases} \quad (1)$$

因为是平面定位，忽略垂直坐标，便可得到跟随目标的位置坐标，如式(2)：

$$\begin{aligned} \begin{cases} x_p \\ y_p \end{cases} &= \begin{cases} 2(x_a - x_c)2(y_a - y_c) \\ 2(x_b - x_c)2(y_b - y_c) \end{cases}^{-1} \\ &= \begin{cases} x_a^2 - x_c^2 + y_a^2 - y_c^2 + d_c^2 - d_a^2 \\ x_b^2 - x_c^2 + y_b^2 - y_c^2 + d_c^2 - d_b^2 \end{cases} \end{aligned} \quad (2)$$

此时跟随目标 $P(x_p, y_p)$ 的位置坐标便可确定。

2.3 跟随算法

STM32传统的跟随算法在现实应用中有很大的缺陷，因此本文设计了一种基于STM32的新型跟随算法，以此来确保购物车在跟随过程中能够稳定、高效。

STM32传统的跟随算法是两点一线式跟随，通过超声波定位到跟随物体然后以直线的形式前进，当定位点更新后改变原来运动轨迹向新的定位点直线前进。假设A为购物车， b, c, d, e 为定位点，传统跟随算法很容易出现图3这样的弊端：当定位点由 b 变成 e 后，由于传统的跟随算法是以直线的方向前进，这样购物车便会撞上障碍物，即使加上避障也无法保证购物车能够稳定地运行。

本文设计了一种基于STM32的新型跟随算法，采用多点多线的模式，如图4所示。假设A为购物车， b, c, d, e 为定位点，当A定位到 b 点后A开始向 b 运动并将 b 的位置存储在STM32存储器中，当定位到 c 后连接 c 与 d 作为运动路线，将 d 和 cd 运动路线存储在存储器中，同理就可以得到一条由 b 到 e 且经过 b, c, d, e 的运动路径。只要STM32的刷新率高就能保证定位点更加密集，对运动路径的描绘就更加细致，对顾客的跟随也更加稳定、高效。

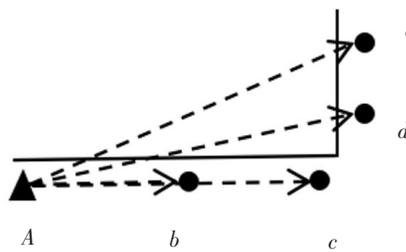


图3 传统跟随算法

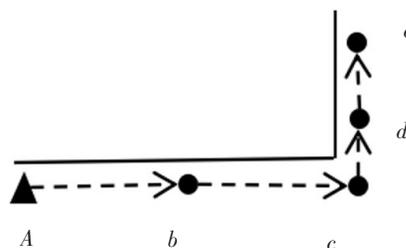


图4 新型跟随算法

3 智能购物车设计

3.1 智能跟随模块

智能跟随模块主要依赖于对波式超声波和跟随算法。其流程图如图5所示，购物车启动后，STM32操控对波式超声运转，得出购物车与顾客的距离数据。在STM32中设置一个初始距离值，当实际距离值大于初始距离时购物车启动跟随算法，反之则停留在原地等待顾客运动，在整个过程中STM32会保持极高的刷新率，确保对顾客的实时监控。

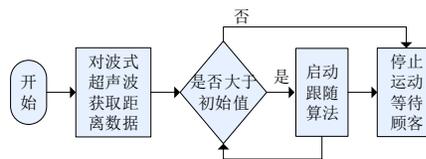


图5 跟随模块流程图

3.2 避障模块

避障模块主要依赖于超声波传感器和红外线检测来实现。避障模块启动后，购物车利用最佳路径算法来判断如何高效地躲避障碍物。

该模块的设计流程可参考图6，若红外线探测仪发现有障碍物存在，会触发中断，在此期间不会对由超声波引起的PWM值做出响应，并且朝着障碍物的相反方向移动以防止碰撞。当未检测到任何障碍物时，利用STM32先定义一组PID参数，然后计算超声波两侧间的距离差，并在STM32中预设一个距离差值。如果距离差为正，购物车选择右转；若是负，选择左转；而如果两者相同，购物车将继续按照既定轨迹运行。在计算PWM的过程中必须注意距离的大小问题，只有满足条件才能将其与转向控制参数相乘，从而得出所需的PWM值，并将此结果传递至购物车，使之能提前或自动完成转向动作。

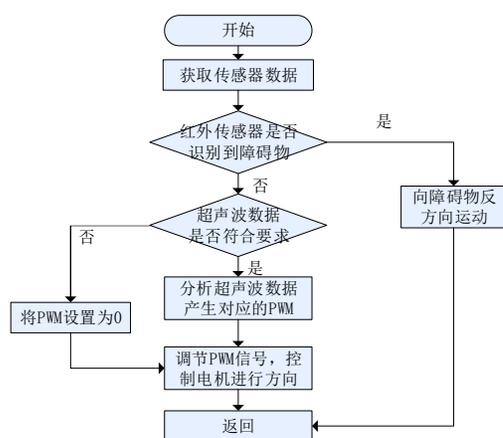


图6 避障模块流程图

3.3 循迹归位模块

在顾客购物结束后，等待顾客将购物车中的物品全部拿出，购物车启动循迹归位模块。首先启动安装在底部的循迹传感器，通过循迹传感器对超市地面上铺设的特定路线标志进行识别，循迹传感器识别到特定标志后启动电机，购物车开始向铺设的路线运动。通过循迹传感器对地面标识的持续识别，能确保保证购物车按照规划好的特定路线返回购物车的停放区域。

4 结语

本文对当前零售店内遇到的手推式购物车不便使用和摆放混乱等难题进行了深入探讨及研究，提出了以传感器技术为主导的智能购物车的解决方案并对其进行了详细的设计，该项设计能有效应对现阶段零售店铺面临的问题，同时也为我们构建未来的智慧型商店奠定了坚实的基础。然而，这项研究还面临许多待解决的问题，如自助结账功能的购物车、商品导航等，这是本团队未来的研究方向。

参考文献：

- [1] 项志娟. 超市购物车改良设计研究[D]. 长沙: 中南林业科技大学, 2014.
- [2] 刘东阳, 岳雪亭. 基于STM32的智能购物车设计[J]. 工业控制计算机, 2021, 34(7): 160-161.

Design and research of smart shopping cart based on auto-following

Li Wenhao*

(School of Information Engineering, Zhengzhou University of Science and Technology, Zhengzhou 450064, China)

Abstract: In order to solve the problem of slow operation caused by shopping carts in the traditional retail industry, an intelligent shopping cart scheme was proposed to improve the consumption experience of customers. The design of smart shopping carts in China started late, and due to problems such as cost and design defects, smart shopping carts have not been popularized, so a new STM32 tracking algorithm is designed, and on this basis, a smart shopping cart that can realize functions such as customer following, autonomous homing, and automatic obstacle avoidance is designed to ensure low cost and improve versatility. Such a shopping cart saves customers time and optimizes the shopping experience, which in turn further drives the merchant's bottom line.

Keywords: STM32; smart shopping cart; follow algorithm

文章编号: 1007-1423(2024)20-0114-04

DOI: 10.3969/j.issn.1007-1423.2024.20.023

基于动作识别的智能家居控制系统研究

方亮*, 江权威, 黎思超, 刘伟

(南华大学土木工程学院, 衡阳 421000)

摘要: 针对当前智能家居控制系统存在的问题, 提出了一种基于 Wi-Fi 通道状态信息(CSI)的动作识别方法。该方法通过采集和分析 Wi-Fi 信号中的 CSI 数据, 实现对家居环境中人员动作的有效识别。在此基础上, 设计了基于动作识别的智能家居控制系统解决方案, 能够根据识别出的动作信息智能化地控制家居设备, 满足用户对于智能家居控制的多样化需求。

关键词: 动作识别; 智能家居; 通道状态信息; 无线感知

0 引言

随着物联网技术的飞速发展, 智能家居系统已成为改善居住环境和提高生活质量的关键技术之一。在众多智能家居应用中, 基于动作识别的交互方式因其自然性和便捷性而备受关注。通过捕捉和分析用户的动作和手势, 智能家居系统能够无缝响应用户的需求, 从而实现更加智能化和个性化的居住环境。

然而, 现有的动作识别技术面临着一系列的挑战, 诸如传感器的依赖性、环境的干扰以及数据处理的复杂性等, 均可能影响系统的整体性能和用户体验。为了解决这些问题, 研究者开始探索利用无线信号, 如 Wi-Fi Channel State Information (CSI) 的特性进行动作识别。CSI 数据能够提供比传统无线信号强度更丰富的信息, 反映信号在多径传播过程中的细微变化, 从而捕捉到人体活动的微小影响。

深度学习作为近年来兴起的一项强大的数据分析工具, 其在图像识别、语音处理等领域取得了显著的成功。结合深度学习对 CSI 数据进行分析, 有望显著提升动作识别的准确度和

系统的鲁棒性。本研究旨在探索结合 Wi-Fi CSI 信息和深度学习技术来实现一个高效、准确的基于动作识别的智能家居控制系统。

1 动作识别技术概述

目前大多数智能家居的人机交互方式采用语音识别技术或触摸式交互技术, 此类技术方案虽然应用较多, 但往往面临着指令易受干扰、需实体接触等缺点。尤其对于聋哑人士或传染病医院等不适宜语音交互和直接接触的使用场景, 语音或触摸式技术的局限性更为突出。

动作识别技术在智能家居中可以应用于多个方面。例如, 通过人体动作识别可以实现智能灯光的调节, 还可以实现智能家居设备的智能操控, 通过识别人体动作来控制家电开关、窗帘的升降等; 此外, 动作识别还可以应用于智能安防系统中, 识别异常的人体动作并及时报警等。

信道状态信息(CSI)是指接收端感知到发送端的发送信号在信道中传输的状态信息, 包括信道的增益、幅值和相位等参数^[1], 携带了反映环境特征的信息, 如人的位置、特征、姿势、

收稿日期: 2024-04-30 修稿日期: 2024-05-30

基金项目: 国家级大学生创新创业训练计划项目(202210555047); 2023 年度南华大学大学生创新创业训练计划项目(D202305211229021701)

作者简介: *通信作者: 方亮(2000—), 男, 四川内江人, 在读本科生, 专业为建筑电气与智能化, E-mail: 3325035137@qq.com; 江权威(2002—), 男, 江西上饶人, 在读本科生, 专业为建筑电气与智能化; 黎思超(2003—), 男, 湖南常德人, 在读本科生, 专业为建筑电气与智能化; 刘伟(2000—), 男, 内蒙古巴彦淖尔人, 在读本科生, 专业为建筑电气与智能化

动作等。使用CSI做动作识别的无线感知技术具有一些优点，如无需佩戴任何设备、非侵入式、适用范围广等，同时也有信号干扰、多径效应、人体遮挡等问题。通过将动作识别技术与智能家居系统相结合，为用户提供更高度个性化的交互方式、更好的隐私保护，推动智能家居技术的进一步发展，为未来智能家居系统的设计和应用提供有益的指导。

2 手势识别系统设计

2.1 Wi-Fi CSI采集模块

本研究选用的是ESP32单片机，由乐鑫信息科技有限公司推出的一系列低成本、低功耗的微控制器，集成了Wi-Fi和双模蓝牙，ESP32系列采用Tensilica Xtensa LX6微处理器，内置天线开关、RF射频模块、功率放大器、低噪声接收放大器、滤波器和电源管理模块，广泛应用于物联网领域。

Wi-Fi CSI采集模块只需在最小系统板上外接一个天线，如图1所示。在ESP-IDF开发框架下，使用ESP-CSI组件的API即可实现对Wi-Fi CSI的信息采集并通过串口上传到动作预测上位机。相比较于一些网卡模块，成本更低、灵活性更高。



图1 CSI采集模块

CSI数据采集过程如下：

(1) 连接路由器。为避免其他网络对CSI信息产生影响，采集模块需要连接单独的Wi-Fi网络建立通信链路，采集模块工作在Wi-Fi STA模式，用于接收和采集CSI信息。路由器可由模块工作在Wi-Fi AP模式代替，用于接受Wi-Fi连接和发送数据包。

(2) 采集Wi-Fi CSI数据。采集模块节点首先发送一个ping指令数据包给路由器，路由器会返回一个数据包，采集节点通过接收并分析返回数据包的子载波的信息。

(3) 上传CSI数据。将采集到的CSI数据暂

存，为避免对下次CSI数据产生影响，使用有线串口将数据发送到动作识别上位机。

Wi-Fi CSI子载波幅度和RSSI如图2所示。

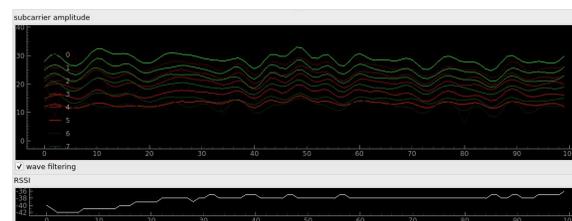


图2 CSI子载波和RSSI

2.2 手势识别平台

本研究采用深度学习技术进行动作识别，选择一台PC上位机做动作识别平台，首先搭建一个Wi-Fi CSI的动作识别系统，比如跨域手势识别系统：Widar 3.0^[2-3]。

Widar3.0的主要部分包括BVP(body-coordinate velocity profile)生成模块和手势识别模块。BVP生成模块为每个CSI段生成BVP，将BVP作识别模型的输入。Widar3.0在监控区域周围部署了多个无线链路，并记录用户扭曲的无线信号的CSI测量值进行预处理，以消除振幅噪声和相位偏移。然后，Widar3.0设计了一个学习模型来捕捉手势的时空特征，并最终实现手势分类。此外，Widar3.0还开发了一个只需一次训练但是可以适应不同数据领域的模型，这大大提高了其适应性和实用性。

该项目的开源数据集包含三种室内场景及22种动态手势，如图3所示，数据集的数据类型有：CSI、BVP和DFS(Doppler Frequency Spectrum)三种，使用开源的csi_tool_box可实现这三种数据格式的相互转换。由于无线感知对场景的依赖很高，在数据采集时发射机和接收机摆放方式等都会影响数据集质量，所以针对具体场景可采集大量的数据集重新训练识别模型，以达到更好的识别效果。具体为将采集到的CSI数据进行预处理操作，包括去噪、滤波、提取相位和振幅信息，动作起止区间检测，数据格式转换等^[4]，转换后的DFS数据划分为训练集与验证集，设计出基于tensorflow框架的手势识别模型，最后进行不断的模型训练与验证评估，达到良好的手势识别精度和应用部署要求。

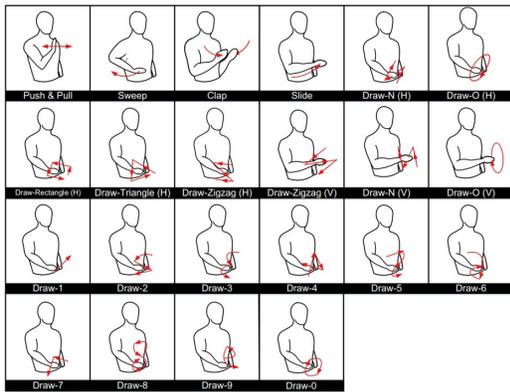


图3 手势类别

在应用部署阶段，将动作识别模型部署在上位机中，搭建出动作识别系统，通过连接采集节点的串口，接收采集模块发送的CSI数据，处理和转换为DFS数据，由手势识别模型完成手势动作预测，最终输出手势类别结果，完成基于Wi-Fi CSI的手势动作识别流程。

3 智能家居控制系统设计

将动作识别系统布置在家庭区域内，识别感应区内的动作，实现实时控制智能家居的设备。动作识别设备平面布置如图4所示。

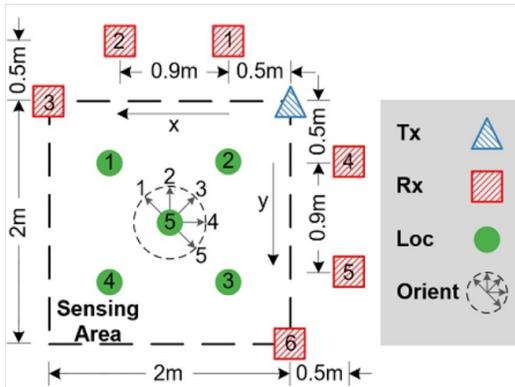


图4 动作识别设备布置

设备包括有6个CSI采集节点、1个路由器信号发送点，除此之外还包含1个上位机和智能家居控制系统。在这样的设备布置中，多个采集节点可以上传更多的人体反射信号的信息，减少了信号采集误差和提高了识别准确率，在5号位置可以识别用户的朝向，丰富了用户动作的多样性和互动性。

动作识别的智能家居控制流程如图5所示。用户在感应区内做出特定的动作手势即可被系统识别并控制家居设备的状态，如在客厅场景的感应区内，用户做出推拉的手势，客厅的灯便会打开，实现了无接触的控制智能家居设备。

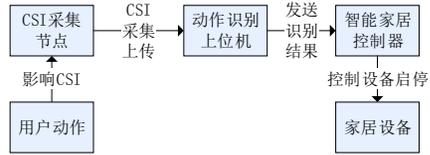


图5 智能家居控制流程图

4 结语

本文提出了使用Wi-Fi CSI的动作识别智能家居控制系统，为在智能家居中的应用提供了更好的隐私保护和高度个性化的交互方式。然而，当前的研究还存在一些不足之处。首先，Wi-Fi CSI识别的准确性和稳定性有待提高，特别是在复杂环境下的准确度有所欠缺。其次，动作识别的实时性需要进一步改进，目前存在一定的延迟问题。此外，对于多人场景中的动作识别，还缺乏有效的处理方法和算法。

未来的发展趋势方向主要包括以下几个方面：一是使用更先进的计算机深度学习技术，实现更准确的动作识别；二是加强多人动作识别的研究，实现对多人场景的有效分析和识别；三是提高动作识别系统的实时性和互动性，使其能够更好地与用户进行交互。

参考文献：

- [1] 马子或. 基于Wi-Fi的动态手势识别研究与应用[D]. 北京:北京邮电大学,2023.
- [2] YI Z, YUE Z, KUN Q, et al. Widar 3.0: zero-effort cross-domain gesture recognition with Wi-Fi [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44(11): 8671-8688.
- [3] YUE Z, YI Z, KUN Q, et al. Zero-effort cross-domain gesture recognition with Wi-Fi [C] //Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services, 2019: 313-325.
- [4] 李江安. 基于WiFi信道状态信息的跨域手势识别研究[D]. 合肥:合肥工业大学,2022.

文章编号: 1007-1423(2024)20-0117-04

DOI: 10.3969/j.issn.1007-1423.2024.20.024

基于深度学习的中式菜品检测识别系统

李娟*

(福州理工学院计算与信息科学学院, 福州 350506)

摘要: 菜品检测识别系统利用计算机视觉技术实现菜品的自动辨识和定价结算, 从而降低人力成本、提高结算效率、减少人工错误带来的损失。然而中式菜品由于切菜方式不同、烹饪方法多样等复杂性增加了检测识别难度。系统分别采用YOLOv8实现菜品定位和ViT实现菜品识别, 并设计实现了中式菜品检测识别和结算。系统具有较高的准确率和快速响应能力, 可应用于中餐馆或中式快餐的智慧餐厅。

关键词: 中式菜品定位; 中式菜品识别; 菜品结算; YOLOv8; ViT

0 引言

菜品检测识别与结算系统利用计算机视觉技术自动检测、识别菜品并进行价格结算, 适用于餐饮行业的多个场景。尽管CNN在菜品检测识别中取得了成功^[1-4], 但中式菜品的检测与识别仍存在一些挑战。首先, 中国厨师注重切菜技巧, 使菜品的形状多样, 如块、条、立方体、片、丝、颗粒等。其次, 中国烹饪方法多样, 如煎、煮、蒸、焖、烤、炖、炸等, 增加了菜品调味的复杂性。此外, 中式菜品的八大菜系涵盖了数百种菜肴, 其成分和颜色各异, 增加了检测识别难度。故基于深度学习设计开发了一款中式菜品检测识别系统, 该系统使用OpenCV对视频流数据进行处理, YOLOv8^[5]对视频流中帧图像进行菜品定位, 并对菜品位置图像依次使用ViT^[6]进行菜品识别, 最后根据识别结果生成相应的账单。该系统可应用于中餐馆或中式快餐的智慧餐厅, 从而提升信息化水平的同时为决策提供准确的数据支持。

1 系统总体设计

1.1 系统整体架构

设计开发的中式菜品检测识别系统整体结

构设计如图1所示, 包括前端应用层、后端服务层和数据库层。前端应用层使用PyQt5^[7]开发页面, 方便用户与系统的交互, 包括菜品定位、菜品识别、账单结算、用户登录注册、管理员数据管理等; 后端服务层使用PyTorch、Python QThread进行开发, 主要聚焦菜品定位、菜品识别模块, 包括OpenCV获取帧图像、YOLOv8菜品定位、ViT菜品识别、多线程等, 后端模块还负责权限管理和接口封装, 切实保障系统的可靠性和稳定性。数据库层采用MySQL存储用户的系统信息, 实现对用户信息、菜品图片、结算等数据的管理和维护。

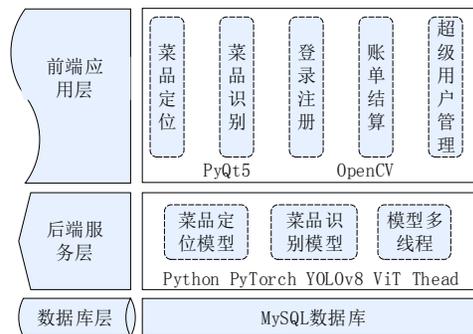


图1 系统整体结构

收稿日期: 2024-07-29 修稿日期: 2024-08-26

基金项目: 2022年移动通信和物联网福建省高校工程研究中心课题(Kfkt202202); 2021年度福建省中青年骨干教师教育科研项目(JAT210607); 2024年校级科研基金项目(FTKY2024001)

作者简介: *通信作者: 李娟(1989—), 女, 四川南充人, 硕士, 工程师, 讲师, 研究方向为机器学习与计算机视觉, E-mail: lijuan12@lzu.edu.cn

1.2 系统功能结构

设计开发的中式菜品检测识别系统包括数据管理系统后台和菜品识别与结算前台两部分，具体功能模块如图2所示。其中数据管理系统后台主要有四个子模块：用户管理主要包括超级管理员的登录验证，这里所针对的超级管理员为餐厅的管理经营者。菜品信息管理模块包含对菜品价格、菜品名称和菜品编号修改和更新。账单数据管理模块主要的功能是对已存在的账单数据删除与打印，计算近一月或者近一周账单营业额。菜品信息管理模块主要的功能是对管理员和用户信息的修改更新。而菜品识别与结算前台主要有四个子模块：用户管理模块、菜品定位模块、菜品识别模块、账单结算模块。用户管理模块和数据管理后台里的用户管理模块相似，但是针对的是餐厅员工。菜品定位模块主要的功能是通过YOLOv8模型对视频流实现实时菜品定位。菜品识别模块主要的功能是对来源于菜品定位模块定位数据通过ViT模型识别菜品。账单结算模块主要的功能是对菜品识别模块识别出的菜品从数据库中查询对应价格计算总价生成账单，消费者可以看到菜品单价、菜品名称和菜品实时图片以便消费者确认订单信息，最后完成支付订单。

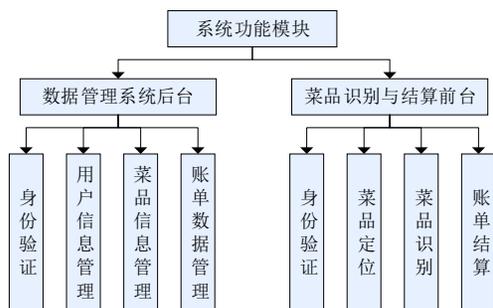


图2 系统总体功能模块图

2 主体功能的实现

系统工作流程上，中式菜品检测识别系统的超级管理员需要通过管理后台对菜品信息进行管理，包括菜品价格、名称等。并在系统使用过程中可随时调整系统中的菜品信息。以食堂场景下的中式菜品检测识别为例，中式菜品

的检测与识别过程分为菜品定位与菜品识别结算两个主要部分。图片经过菜品位置YOLOv8检测模型进行菜品位置定位，然后再把定位好的菜品图片输入到ViT菜品识别模型中进行识别分类，最后结算模块根据输出的菜品类别查询价格生成账单以供消费者结算。其具体流程如图3所示。

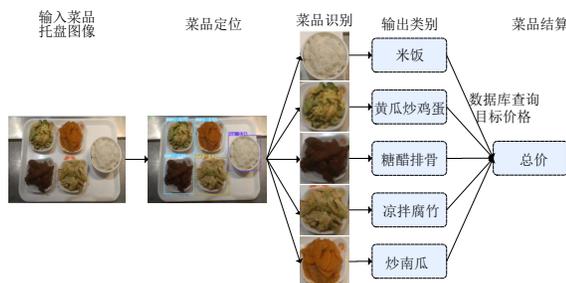


图3 中式菜品检测识别流程

2.1 中式菜品检测

首先利用2K高清摄像头在FIT食堂拍摄了约2000张分辨率为1080P的餐盘实物图片组成菜品检测数据集，称为FIT-DISH2000。其拥有5种不同形状的餐盘，并将数据集划分为1769个训练样本、160个测试样本和50个验证样本。再使用YOLOv8的预训练模型进行微调训练。实验证明，在FIT-DISH2000数据集上达到95.2%准确率。相比于YOLOv3、YOLOv5，在推理速度、召回率、准确率和泛化能力方面具有显著优势。

2.2 中式菜品识别

首先从开源数据集ChineseFoodNet中选取FIT食堂中的10类菜品样本组成菜品识别数据集ChineseFood10，并将数据集80%作为训练集，10%作为验证集，10%作为测试集。再使用ViT模型训练菜品识别，其中ViT模型需要将菜品图像分解成多个小块，然后通过嵌入层将每个小块转换为特征向量序列。再通过Transformer编码器将这些特征向量序列进行编码和交互，以捕获图像中不同区域之间的关系和上下文信息。最后通过全局池化或添加额外的分类头来获取整个图像的代表，并进行菜品识别。ViT模型框架如图4所示。

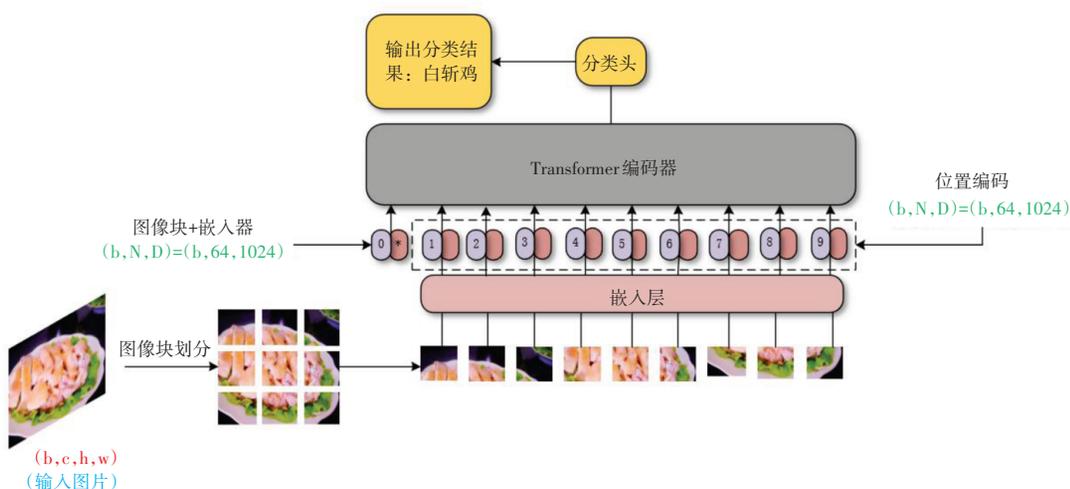


图4 中式菜品检测ViT识别流程

2.3 中式菜品结算

账单生成将菜品识别的菜品类别送入数据库进行查询，以获取相应菜品的价格信息。一旦查询到价格，系统将价格信息反馈到主界面上，生成完整的账单，包括菜品名称和价格等信息，以使用户查看。然后将生成的账单传递给结算页面，系统生成二维码供消费者使用。消费者可以通过扫描二维码完成支付操作，实现快速便捷的结算流程。

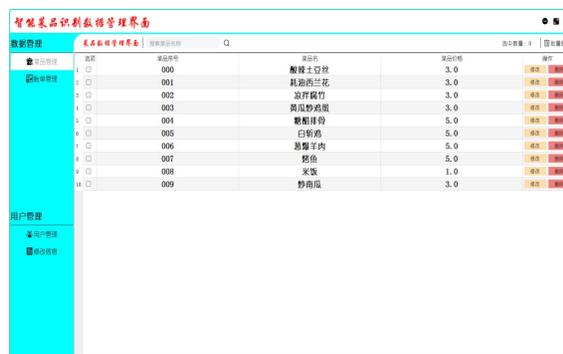


图6 菜品菜单界面

3 系统测试运行

系统的核心模块为中式菜品的检测与识别结算，经过测试系统能在轻量级的算力平台上实现视频实时检测识别，部分系统运行界面如图5、图6所示。



图5 菜品检测、识别、结算界面

4 结语

基于深度学习设计实现了一款中式菜品检测识别系统，该系统通过YOLOv8对摄像头采集的中式菜品视频实现菜品定位，并通过ViT对定位的菜品进行识别和结算；系统还支持对菜品信息的高效化管理，可为后续包括营养学分析等各种数据分析提供便利，可应用于中餐馆或中式快餐的智慧餐厅。

参考文献：

[1] 邓志良,李磊.基于改进残差网络的中式菜品识别模型[J].激光与光电子学进展,2021,58(6):264-272.

- [2] CHEN X, ZHU Y, ZHOU H, et al. ChineseFood-Net: a large-scale image dataset for Chinese Food recognition[EB/OL]. arXiv:1705.02743, 2017.
- [3] HASSANNEJAD H, MATRELLA G, CIAMPOLINI P, et al. Food image recognition using very deep convolutional networks[C]//Proceedings of the 2nd International Workshop on Multimedia Assisted Dietary Management. Amsterdam, The Netherlands, Association for Computing Machinery, 2016: 41-49.
- [4] MIN W, JIANG S, LIU L, et al. A survey on food computing[J]. ACM Comput Surv, 2019, 52(5): 92.
- [5] Ultralytics. YOLOv8 [Z], 2023. <https://github.com/ultralytics/ultralytics>.
- [6] ALEXEY D, LUCAS B, ALEXANDER K, et al. An image is worth 16×16 words: transformers for image recognition at scale[C]//Proceedings of the International Conference on Learning Representations (ICLR), 2021.
- [7] WILLMAN J. Modern PyQt: create GUI applications for project management, computer vision, and data analysis[M]. Berkeley, CA, Apress, 2021.

Chinese food detection recognition system based on deep learning

Li Juan*

(School of Computing and Information Science, Fuzhou Institute of Technology, Fuzhou 350506, China)

Abstract: Food detection and recognition system uses computer vision technology to achieve automatic identification and pricing settlement of dishes, thus reducing labour costs, improving settlement efficiency and reducing losses caused by manual errors. However, the complexity of Chinese food increases the difficulty of detection and identification due to the different cutting methods and cooking methods. The system adopts YOLOv8 to locate food and ViT to identify the dishes, and designs and implements the detection, identification and settlement of Chinese food. The system has high accuracy and fast response ability, and can be applied to smart restaurants, such as Chinese restaurants or Chinese fast food.

Keywords: Chinese food detection; Chinese food recognition; food settlement; YOLOv8; ViT

(上接第 116 页)

Research on smart home control system based on motion recognition

Fang Liang*, Jiang Quanwei, Li Sichao, Liu Wei

(School of Civil Engineering, University of South China, Hengyang 421000, China)

Abstract: Addressing the issues existing in current smart home control systems, this study proposes a motion recognition method based on Wi-Fi Channel State Information (CSI). This method effectively identifies the movements of individuals within the home environment by collecting and analyzing CSI data from Wi-Fi signals. Based on this, a smart home control system solution is designed, utilizing motion recognition. This solution intelligently controls home devices based on recognized motion information, meets their diverse needs for smart home control.

Keywords: motion recognition; smart home; channel state information; wireless perception