

M 现代计算机

XIANDAI JISUANJI

第29卷（卷终）第24期（总第792期）

半月刊（1984年创刊）

2023年12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 广州一龙印刷有限公司
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



M 现代计算机

第29卷（卷终）第24期（总第792期）

2023年12月

2023年12月 第29卷（卷终）

第24期（总第792期）



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

基于机器学习的速冻水饺缺陷检测·····	刘一阳, 马子领	(1)
基于Improving-SRGAN算法的岩心图像超分辨率重建 ·····	罗旺健, 张豫堃, 吴晓红, 何小海, 熊淑华	(10)
基于用户长短期偏好及时空场景的下一个兴趣点推荐·····	郑宏洲, 曾国荪	(18)
基于深度学习的苹果叶面病害缺素识别系统 ·····	郑淋萍, 李万益, 李南健, 黄灿敏, 林祎琦, 陈津毅, 邱 芸, 郭小芸	(26)
基于改进BiseNet的轻量化上颌前牙分割算法 ·····	吴 杰, 陶青川	(33)
改进YOLOv5的安全帽佩戴检测算法 ·····	王 坡, 罗红旗	(40)
一种聚类矩阵和剪枝策略结合的关联规则算法·····	许珂乐	(46)
多维攻击下的网络入侵信号检测研究·····	岳 茹	(51)
基于K-means和Dlib的脸型分类算法设计与实现 ·····	李东阳, 项辉宇, 冷崇杰, 张 勇	(56)
基于多尺度特征的遥感图像检测方法研究·····	陈 芳	(61)

实践与经验

基于藏成药数据的知识图谱可视化构建·····	樊 森, 高 屹	(64)
基于AutoRec的推荐系统模型研究 ·····	严武军, 刘守业, 贺娇娇	(69)
基于BERT+BiLSTM+CRF的中国非物质文化遗产知识图谱构建 ·····	黄 薇, 周 卫, 李宏杰, 张 涛	(74)
基于多线程对TCP端口扫描技术的研究·····	石志强, 石美静	(79)

开发案例

蔬菜大棚温度可视化控制系统的设计与实现·····	郭金文	(83)
基于JavaWeb的线上考试系统的设计与实现 ·····	操心慧, 温智聪	(90)
基于GPR模型的草地地上生物量估算软件设计 ·····	王志飞	(97)
基于无纸化策略的胎儿医学中心信息平台改造·····	莫远明, 乔梦薇, 林思阳, 黎鹏安	(103)
基于Arduino的智能垃圾分类系统的设计与开发 ·····	张银霞, 张新雨, 王冀伟, 聂 鑫, 王鸣政, 王 硕	(108)
一种微服务框架的浙江气象应用开放服务技术研究与实现·····	曾 悠, 封秀燕, 杨 明	(112)
基于SeetaFace6的人脸识别信息管理系统 ·····	李 娟, 杨溢超	(117)

Modern Computer

(Vol. 29, No. 24; Dec. 25, 2023)

CONTENTS

Research and Development

Defect detection of quick-frozen dumpling based on machine learning	(1)
Super-resolution reconstruction of core images based on the improving-SRGAN algorithm	(10)
Next point of interest recommendation based on user's long and short term preference and spatio-temporal context	(18)
Apple leaf disease deficiency based on deep learning	(26)
Lightweight maxillary anterior tooth aegmentation algorithm based on improved BiseNet	(33)
Improved helmet wearing detection method of YOLOv5	(40)
A data mining algorithm based on clustering matrix and pruning strategy	(46)
Study on network intrusion signal detection under multidimensional attack	(51)
Design and implementation of face shape classification algorithm based on K-means and Dlib	(56)
Research on remote sensing image detection method based on multiscale features	(61)

Practice and Experience

Visualization construction of knowledge graph based on Tibetan medicine data	(64)
Research on recommendation system model based on AutoRec	(69)
Construction of a knowledge map of China's intangible cultural heritage based on BERT+BILSTM+CRF	(74)
Research on TCP port scanning technology based on multi-threading	(79)

Development Solution

Design and implementation of a visualized temperature control system for vegetable greenhouses	(83)
Design and implementation of an online examination system based on JavaWeb	(90)
Design of grassland aboveground biomass estimation software based on GPR model	(97)
Reconstruction of fetal medical center information platform based on paperless office strategy	(103)
Design and development of an intelligent garbage classification system based on Arduino	(108)
Research and implementation of Zhejiang open service technology for meteorological application based on microservice framework	(112)
Face recognition information management system based on SeetaFace6	(117)

研究与开发

文章编号: 1007-1423(2023)24-0001-10

DOI: 10.3969/j.issn.1007-1423.2023.24.001

基于机器学习的速冻水饺缺陷检测

刘一阳*, 马子领

(华北水利水电大学机械学院, 郑州 450045)

摘要: 速冻水饺在冷冻过程中, 容易造成冻裂、破损的现象, 采用人工方法对破损水饺进行剔除耗时耗力, 因此, 设计一种基于机器学习的速冻水饺缺陷检测方法。首先, 分别提取 R, G, B, H, S, V 颜色空间的灰度共生矩阵 (GLCM) 和颜色矩共 20 维特征向量, 其次采用主成分分析法 (PCA) 对特征向量进行优化, 得到 6 个最优特征向量作为支持向量机输入, 采取交叉验证超参数寻优确定支持向量机参数, 最后用训练好的支持向量机模型对缺陷饺子进行分类识别, 对有缺陷的饺子外接最小矩形, 并输出外接矩形的中心点坐标。结果证明, 该方法可以对有缺陷的饺子进行快速识别, 识别准确率为 95.5%, 平均每个饺子的检测耗时为 13.13 ms。

关键词: 速冻食品缺陷检测; 灰度共生矩阵; PCA; 支持向量机

0 引言

速冻水饺在生产过程中, 由于受到材质和生产流程的原因, 表面会受到一定程度的破损, 产生破肚现象, 这些缺陷将严重影响速冻水饺的品质以及食用时的口感。对此, 速冻加工产业多采用人工肉眼识别的模式, 对缺陷速冻水饺进行剔除。随着信息时代的飞速发展, 图像处理技术和机器学习技术的不断完善, 采用机器学习的方法对速冻水饺破损进行在线检测代替传统人工肉眼识别成为了可能。

近年来国内外许多学者对食品缺陷的检测进行了研究, 并取得一定程度的进展。田有文等^[1-2]采用计算机图像处理技术和支持向量机分类, 通过创建 CIE (Commission Internationale de l'Eclairage/International) 颜色空间的色度矩集作为玉米和葡萄纹理特征的特征向量, 采用支持向量机模型对玉米和葡萄病害进行识别, 平均识别准确率达到 87.5% 和 86.7%。陈荣等^[3]采用茶叶病害的形状特征作为支持向量机模型的特征向量, 对茶叶的茶炭疽病、茶饼病、茶白星病进行检测, 平均识别率可以达到 90%。刘军等^[4]

采用机器视觉技术对核桃外部缺陷进行了研究, 筛选出 9 维特征组合 (长宽比、分散度、欧拉数、相同二阶矩椭圆的偏心率、最小多边形顶点数、面积/凸壳面积、颜色矩的二阶矩), 分别采用支持向量机、BP 神经网络等 15 种识别模型进行了准确率检测, 经过对比发现, 支持向量机模型的总识别率较高。朱丽娟^[5]采用灰度共生矩阵的方法对红枣提取纹理特征, 对红枣的分级起到了效果。李亚文等^[6]采用聚类算法对病害区域进行切割, 对区域内的纹理特征进行提取, 通过均值范围来判断黄瓜叶根部是否有炭疽病, 取得了较好的效果。Singh 等^[7]对植物叶片的颜色、纹理、HOG 特征进行分类研究, 提出贝叶斯优化支持向量机, 检测准确率有所提高。Kumar 等^[8]提出的 HRF-MCSVM 能够准确分类植物叶片的疾病, 测试精度达到 98.9%。

当前, 国内外对于将支持向量机应用于速冻水饺缺陷检测识别的研究未见报道, 本研究采用图像处理技术对速冻饺子缺陷的原始图片进行预处理, 缺陷部位进行分割, 灰度共生矩阵和颜色矩的提取, 利用 PCA 对提取的特征向量进行降维, 然后利用粒子群算法的支持向量

收稿日期: 2023-03-28 修稿日期: 2023-09-02

作者简介: *通信作者: 刘一阳 (1998—), 男, 河南商丘人, 在读硕士, 研究方向为机器学习, E-mail: 1875216467@qq.com; 马子领 (1983—), 男, 河南新乡人, 博士, 副教授, 研究方向为机器视觉

机训练模型,并对模型的准确率进行检测,最后对支持向量机的输出数据进行识别,并对有缺陷的饺子进行定位,输出缺陷饺子最小外接矩形中心点的像素坐标,方便后续可编程逻辑处理器(PLC)控制处理。这将为速冻食品缺陷检测提供参考依据,促进速冻食品产业智能化,对机器学习在速冻食品领域的应用有重大的实际意义。

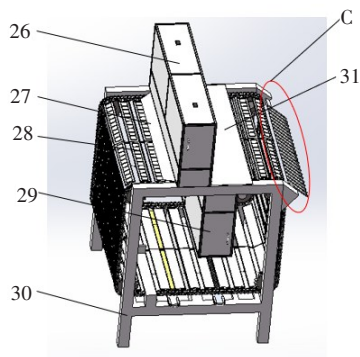
1 材料与方法

1.1 试验材料

根据表面是否具有破肚、裂纹等特征将速冻饺子划分为完整水饺和缺陷水饺,训练样本为完整水饺和缺陷水饺各300个,测试样本为剩余的完整水饺和缺陷水饺各100个。

1.2 饺子图像采集机构设计

饺子采集装置的设计拟采用上下两端架设拍摄装置(如图1所示),上下相机框架内部设置有LED照明装置,水平遮光板可以避免外部光源干扰,共设计4个视觉检测箱体,1个箱体负责检测10个方格内的饺子,提高检测效率,由于拍摄饺子图片共涉及10个饺子,后续将通过Roi区域提取,将10个饺子分成10份去处理,方便后续剔除装置对缺陷饺子进行剔除。



C—视觉部分出料装置; 26—上相机框架;
27—斜遮光板; 28—玻璃板; 29—下相机框架;
30—链传动支架; 31—水平遮光板;

图1 视觉检测机构设计

1.3 饺子图片预处理

由于所采集的饺子图片尺寸过大,且图片

中其他干扰因素过多,为了加速检测效率,需要对图片进行感兴趣区域提取,原图像尺寸为800×600 px,经提取后的图片(如图2所示)尺寸为256×256 px,提升后续的检测效率。



图2 感兴趣区域提取

感兴趣表面图像灰度化和滤波是表面图像特征提取的重要前提^[9]。灰度化可以将RGB三通道彩色图像转化为单通道灰白图像,加快计算机对于图片的处理速度,滤波可以过滤不必要的噪声,由于速冻水饺为面食制品,在感兴趣区域内会有面粉点的残留,经过实验,采用中值滤波可以消除大部分的噪声,使后期的特征提取更加精准。对于水饺缺陷区域的提取,首先将缺陷饺子图像转化为二值图像,转化方法采用最大类间方差法,其次进行形态学开运算和轮廓填充,最后经过图片的绝对差来得到缺陷区域的二值图片。

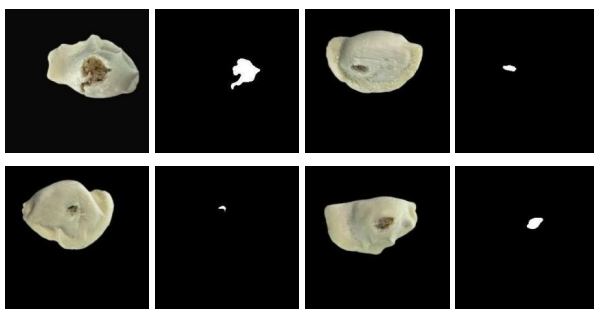


图3 阈值分割缺陷区域提取图

2 饺子特征提取及优化

图像特征提取是数字图像处理的关键步骤之一,关系到分类器的效率与准确度^[10],查阅文献发现,常用的特征分类包括颜色特征、形态特征、纹理特征,选择合适的特征对于支持向量机的分类预测起着决定性作用。

对饺子缺陷图片进行分析可知,颜色特征

具有平移不变性, 纹理特征具有旋转不变性以及较强的噪声抵抗性, 两者作为支持向量机的特征向量比较理想。

2.1 颜色矩提取

颜色矩主要表现颜色的分布, 通常采用颜色的一阶矩、二阶矩、三阶矩就足以表达颜色的分布, 为了寻找最优的特征向量, 本文采用RGB颜色空间和HSV颜色空间的一阶矩(均值 μ_i)和二阶矩(方差 σ_i), 共计 $3 \times 2 \times 2$ 个特征向量, 其数学定义:

$$\mu_i = \sum_{j=1}^N \frac{1}{N} P_{ij} \quad (1)$$

$$\sigma_i = \left(\frac{1}{N} \sum_{j=1}^N (P_{ij} - \mu_i)^2 \right)^{\frac{1}{2}} \quad (2)$$

式中: μ_i 为灰度均值; σ_i 为灰度方差; N 为像素个数; P_{ij} 为第 i 个颜色通道中第 j 个像素值。

为验证颜色矩的可行性, 将10组完整饺子和缺陷饺子的颜色矩提取数据制成折线图, 由图4可知, RGB通道颜色矩对于饺子缺陷的区分度不明显, 因此以RGB通道的颜色矩特征向量对缺陷饺子进行分类的效果不理想, HSV通道中S方向的均值对于完整饺子和缺陷饺子有着较好的区分效果, 因此选择S通道的均值作为特征向量可以提高整体检测的准确率。

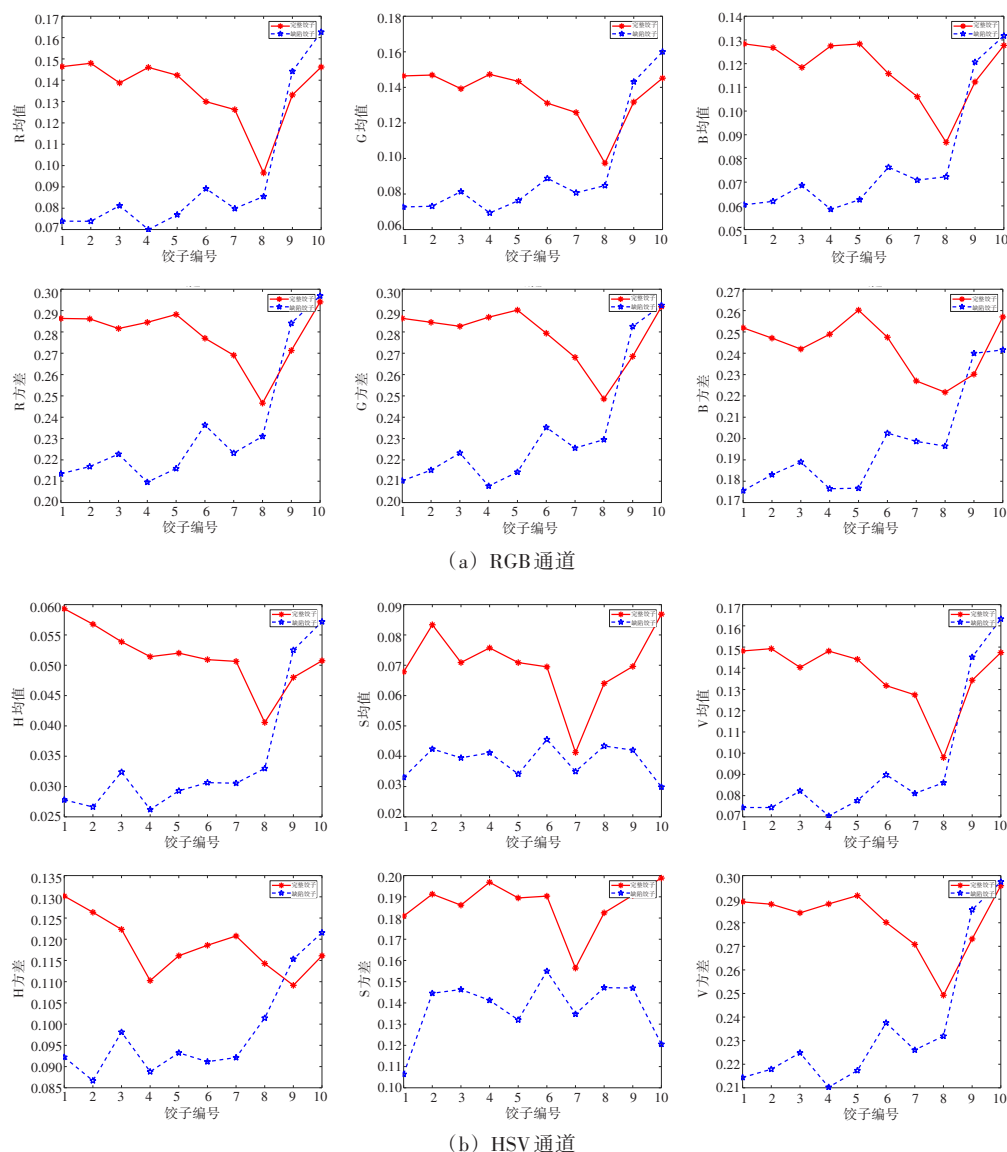


图4 颜色矩对比折线图

2.2 纹理特征提取

纹理是一种反映图像中同质现象的视觉特征,它体现了物体表面的具有缓慢变化或者周期性变化的表面结构组织排列属性^[11]。图片的纹理特征常用空间灰度共生矩阵来表示,本文采用MATLAB R2022a中的graycomatrix函数对饺子图片进行空间灰度共生矩阵提取,一共可提取出0°,45°,90°,135°共计四个方向的对比度、相关性、同质性,能量均值的纹理特征。数学公式如下:

(1) 对比度:反应像素与其周边像素之间的强度对比,反应图片的清晰和纹理沟纹深浅的程度,反应灰度共生矩阵的局部变化。

$$C = \sum_{i,j=0}^{N-1} |i-j|^2 P(i,j) \quad (3)$$

式中: C 为对比度; N 为灰度级数; $P(i,j)$ 为灰度共生矩阵的第*i*行,第*j*列的元素。

(2) 同质性:灰度共生矩阵中元素与对角线元素的接近程度,反映了纹理的局部变化。

$$H = \sum_{i,j=0}^{N-1} \frac{P(i,j)}{|i-j|+1} \quad (4)$$

式中: H 为同质性。

(3) 相关性:图像像素与其周边像素的相关程度,反映了图像中局部灰度相关性。

$$f = \sum_{i,j=0}^{N-1} \frac{P(i,j)}{(i-j)^2 + 1} \quad (5)$$

式中: f 为相关性。

(4) 能量:图像灰度共生矩阵中各元素的平方和,反应图像灰度分布均匀程度和纹理的粗细程度。

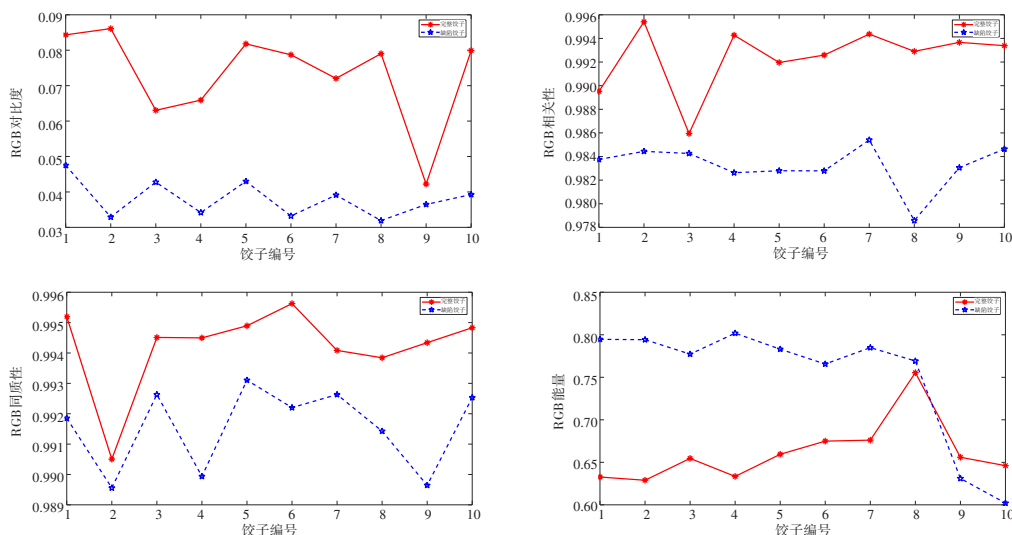
$$E = \sum_{i,j=0}^{N-1} P^2(i,j) \quad (6)$$

式中: E 为能量。

为验证其可行性,分别将10组完整饺子和缺陷饺子的纹理特征数值数据制作成折线图,由图5分析可知,缺陷饺子和完整饺子在RGB通道的对比度、相关性、同质性有着明显的差异,能量有小范围的重叠,故采用RGB通道的灰度共生矩阵分类效果较优。

2.3 PCA特征向量优化

为了提高检测的准确率,尽可能多的提取饺子不同颜色空间的图像特征,构建了20维度的特征向量,但这些特征信息量复杂,存在信息交叉的现象,为了提高分类器的效率,需要对数据进行降维。主成分分析(PCA)是通过降维的方法将多个变量转为少数几个综合变量的多元统计分析方法,可将众多信息中相互重叠的信息去除^[12],采用降维之后的特征向量既可以保证分类器检测的准确率,又可以缩短检测时间,提高缺陷检测的效率。



(a) RGB通道

图5 灰度共生矩阵对比折线图

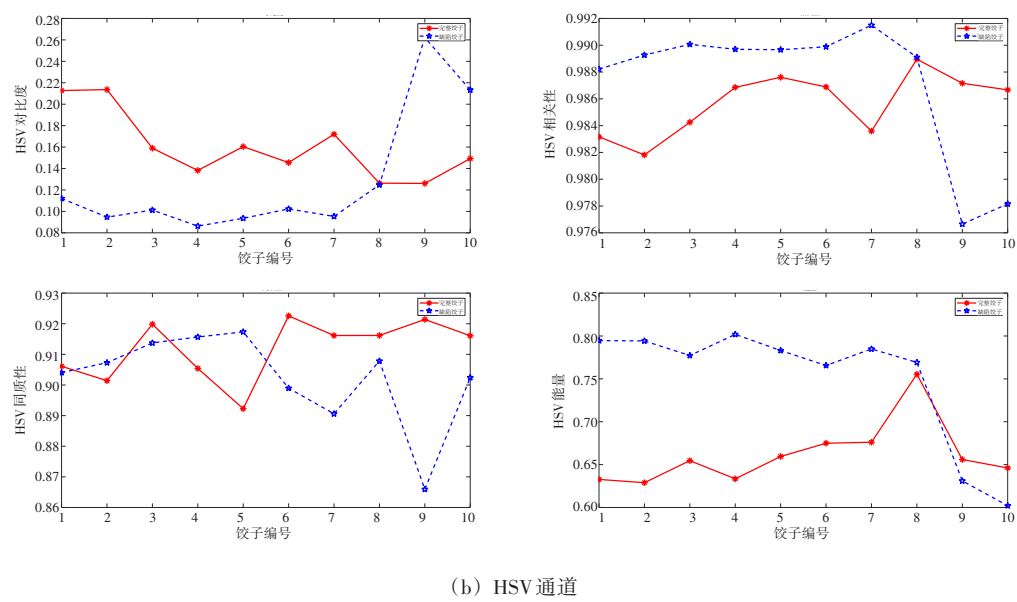


图 5 灰度共生矩阵对比折线图（续）

根据前人的经验判断，主成分分析需要进行一定的前置条件检测：KMO 值和巴特利特球形度检验^[13]。KMO 值计算公式如下：

$$KMO = \frac{\sum_{i \neq j}^2 r_{ij}^2}{\sum_{i \neq j}^2 r_{ij}^2 + \sum_{i \neq j}^2 \alpha_{ij}^2} \quad (7)$$

式中： r_{ij} 为简单相关系数； $\alpha_{ij, 1, 2, 3, \dots, k}^2$ 为偏向关系数。

KMO 的计算结果反馈见表 1。本文通过 SPSS27 进行分析计算，结果见表 2，KMO 取样适切性量数为 0.859，适合做主成分分析。

表 1 KMO 取值与主成分分析的关系	
KMO 取值	PCA 主成分分析结果
<0.5	不能进行主成分分析
0.5~0.6	可以进行主成分分析，效果较差
0.6~0.7	主成分分析效果可以接受
0.7~0.8	主成分分析效果一般
0.8~0.9	适合主成分分析
>0.9	比较适合主成分分析

表 2 KMO 和巴特利特检验结果	
KMO 取样适切性量数	0.859
巴特利特球形度检验	近似卡方 60281.231
	自由度 190.000
	显著性 0.000

对 20 维度的特征向量进行主成分分析，各特

征解释方差见表 3。由表 3 可知，当特征向量为 3 时，累计方差贡献率达到 95.193%。结合图 6，当特征向量为 6 时，碎石图变化趋于平缓，且累计方差贡献率达到 99.365%，综合上述结果，取 6 维特征向量能很好地反映饺子缺陷的特征。

表 3 总方差解释表			
成分	初始特征值		
	总计	方差百分比	累积/%
1	16.015	80.073	80.073
2	1.862	9.311	89.384
3	1.162	5.809	95.193
4	0.602	3.012	98.204
5	0.125	0.627	98.832
6	0.107	0.534	99.365
7	0.059	0.295	99.660
8	0.028	0.140	99.800
9	0.014	0.070	99.870
10	0.008	0.042	99.913
11	0.006	0.030	99.943
12	0.005	0.024	99.967
13	0.004	0.019	99.986
14	0.002	0.010	99.995
15	0.001	0.003	99.998
16	0.000	0.001	100.000
17	5.517E-5	0.000	100.000
18	3.836E-5	0.000	100.000
19	3.514E-5	1.757E-5	100.000
20	4.545E-5	2.273E-6	100.000

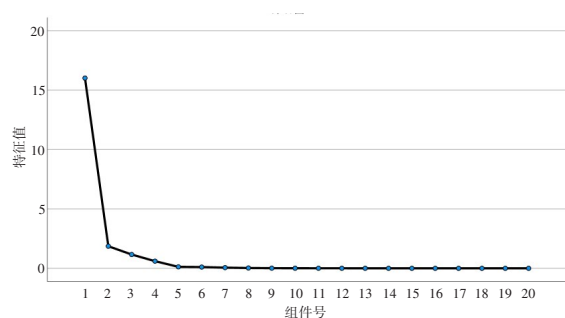
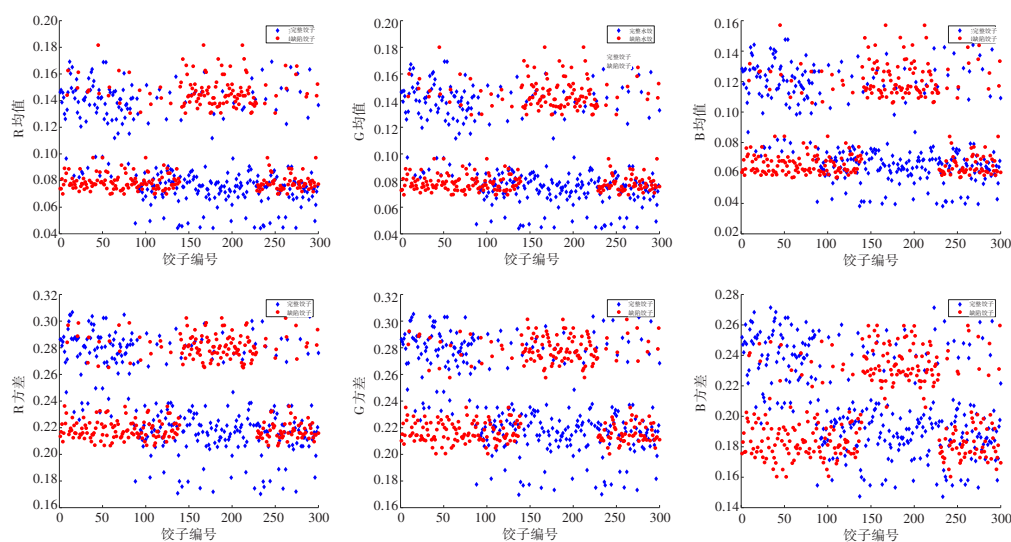


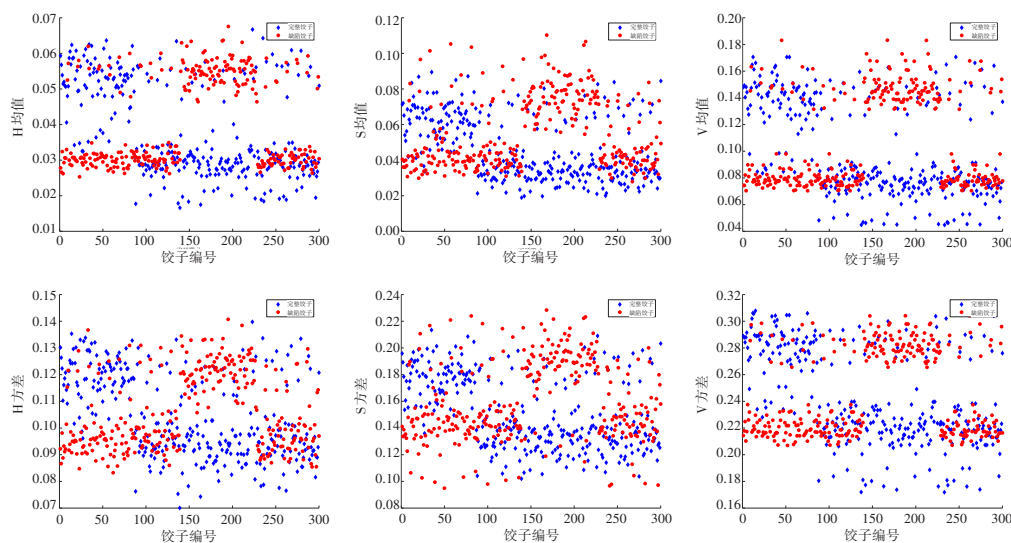
图6 特征向量碎石图

2.4 特征向量的选取

利用 MATLAB R2022a 制作训练集 600 张饺子 20 维特征向量的散点图(如图 7 所示), 观察散点图中红点和蓝点分布情况, 结合图 4 和图 5 折线图, 选取 RGB 通道中图片纹理同质性、能量、HSV 通道中图片纹理对比度、相关性和 HSV 颜色矩的方差共 6 个特征作为支持向量机的输入向量。

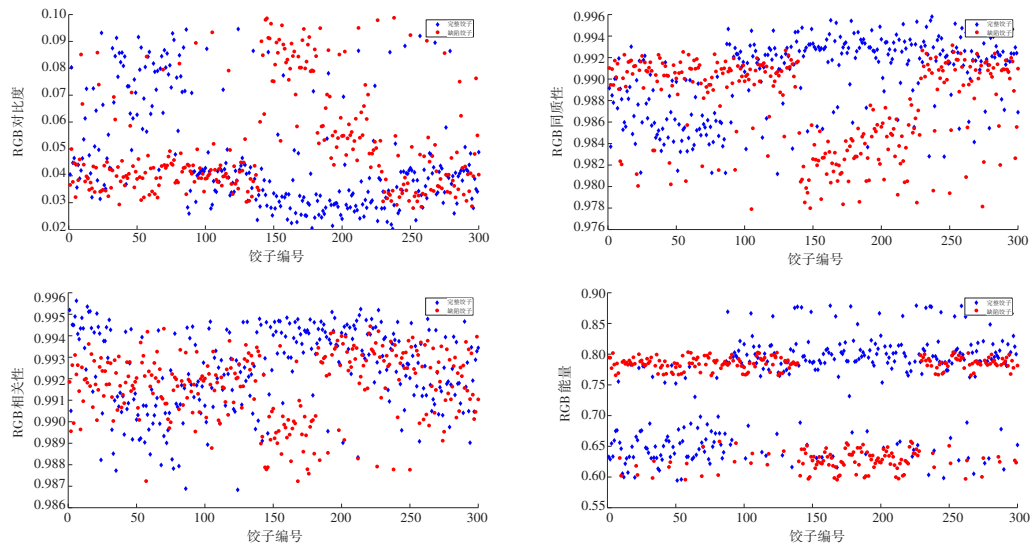


(a) RGB通道 (结合图4折线图)

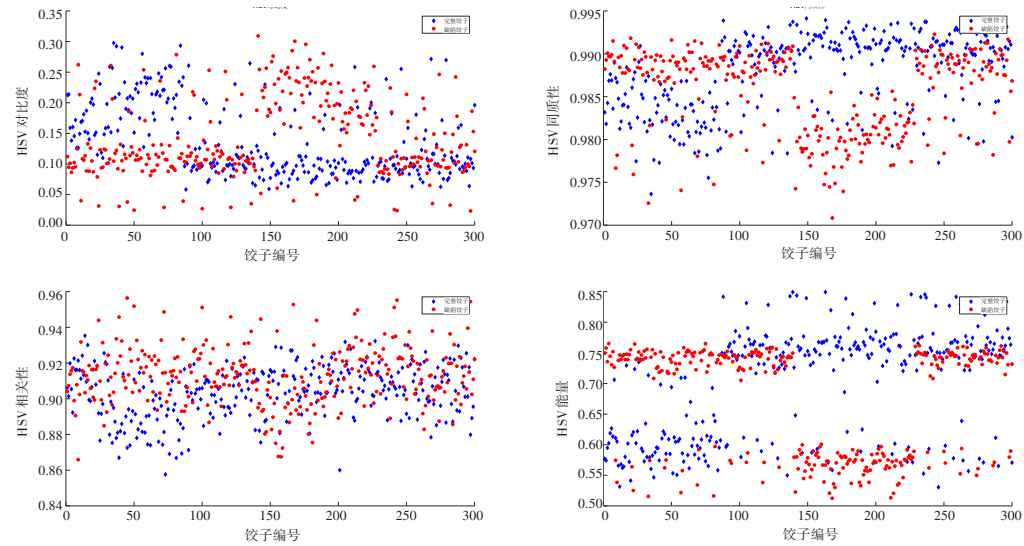


(b) HSV通道 (结合图4折线图)

图7 20维特征向量散点对比图



(c) RGB通道 (结合图5折线图)



(d) HSV通道 (结合图5折线图)

图7 20维特征向量散点对比图 (续)

3 基于支持向量机的饺子缺陷识别

支持向量机是常用的二分类模型，能够高效快速地处理小样本数据。对于非线性可分的问题，常采用高斯核函数，针对最优核函数参数 g 和惩罚系数 c 对支持向量机模型进行训练，通过训练好的模型来对饺子进行缺陷识别。

3.1 支持向量机

支持向量机的数学意义是找到最大边距的超平面，作为其决策边界，得到下述的优

化目标：

$$\begin{cases} \max \left\{ \frac{1}{\|\omega\|} - C \sum_{i=1}^m \gamma_i \right\} \\ \text{s.t. } \gamma_i (\omega^T x_i + b) \geq 1 - \beta_i \\ \beta_i \geq 0 \\ i = 1, 2, \dots, m \end{cases} \quad (8)$$

式中： m 为样本个数； i 为样本索引； x_i 为训练样本； ω 、 b 为主问题参数； y_i 为样本标签； C 为惩罚因子； γ_i 为松弛变量。

针对上述条件最值问题,构造拉格朗日函数:

$$L(\omega, \gamma, b, a, \mu) = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^m \gamma_i + \sum_{i=1}^m a_i \{1 - \gamma_i - y_i [\omega^T \phi(x_i) + b]\} - \sum_{i=1}^m \mu_i \gamma_i \quad (9)$$

式中: a_i, μ_i 为拉格朗日乘子。

对拉格朗日函数自变量 ω, b, γ_i 求偏导为0, 可得:

$$\nabla_{\omega, b, \gamma_i} L = \begin{cases} \frac{\partial L}{\partial \omega} = \omega - \sum_{i=1}^m a_i y_i \phi(x_i) = 0 \\ \frac{\partial L}{\partial b} = -\sum_{i=1}^m a_i y_i = 0 \\ \frac{\partial L}{\partial \gamma_i} = C - a_i - \mu_i = 0 \end{cases} \quad (10)$$

将上式代入拉格朗日函数得最终条件终值问题:

$$\begin{cases} \min \left\{ \frac{1}{2} - \sum_{i=1}^m \sum_{j=1}^m a_i a_j y_i y_j \phi(x_i)^T \phi(x_j) - \sum_{i=1}^m a_i \right\} \\ \text{s.t. } \sum_{i=1}^m a_i y_i = 0 \\ 0 \leq a_i \leq C \\ i = 1, 2, \dots, m \end{cases} \quad (11)$$

由于涉及到 $\phi(x_i)^T \phi(x_j)$ 向量内积映射到高维空间的问题, 所以引入了四种常见的核函数。

(1) 线性核函数, 适用于线性可分的问题。

$$k(x_i, x_j) = x_i^T x_j \quad (12)$$

(2) 高斯核函数, 适用于线性不可分问题。

$$k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (13)$$

(3) 多项式核函数, 适用于没有先验经验的分类。

$$k(x_i, x_j) = (a x_i^T x_j + c)^d \quad (14)$$

(4) Sigmoid 核函数, 此时支持向量机是一种多层感知器神经网络。

$$k(x_i, x_j) = \tanh(ax_i^T x_j + c) \quad (15)$$

3.2 交叉验证超参数寻优

首先将数据集分成 n 份, 每次将 $n - 1$ 份数据集作为支持向量机模型的训练集, 将剩

余的 1 份作为测试集对测试样本进行分类, 共训练 n 个模型, 将 n 个模型验证准确率作为评价指标, 找到模型验证准确率最高的惩罚因子 C 和核函数参数 g , 作为支持向量机的优化参数。

4 实验结果和分析

实验是在 Intel 12400F 处理器, 运行内存 16 GB 的电脑上进行的。编程环境是 MATLAB R2022a。采用 600 万像素工业 CCD 相机对饺子进行图像采集。

对采集的 800 张饺子图片 RGB 通道的同质性、能量、HSV 通道的同质性、相关性、能量, 和 S 通道上的均值进行提取, 制作支持向量机的训练集和测试集。粒子群优化算法的种群数量设置为 10, 最大迭代次数设置为 20, 得到支持向量机的惩罚参数 $C=10$, 核参数 $g=6.209$ 。

将 600 组 6 维的特征向量作为训练集对支持向量机模型进行训练, 200 组 6 维的特征向量作为测试集对训练好的支持向量机模型进行测试, 选择不同的核函数进行分类准确性试验, 准确率见表 4。

表 4 核函数准确率对比表

核函数选择	分类准确率/%
线性核函数	87.0
多项式核函数	92.5
高斯核函数	95.5
Sigmoid 核函数	50.0

高斯核函数的准确性最高, 达到了 95.5%, 对支持向量机分类的结果进行提取, 判别为有缺陷的部分饺子外接最小矩形如图 8 所示, 并输出最小外接矩形的中点坐标(见表 5), 以达到缺陷检测定位的目的。

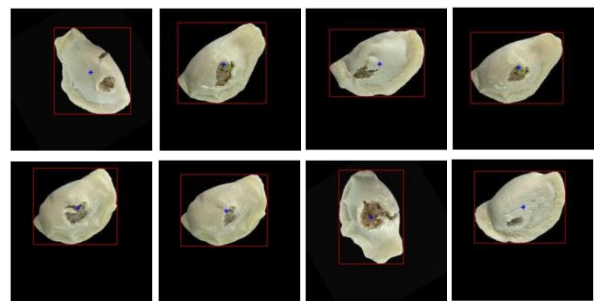


图 8 部分缺陷饺子最小外接矩形

表5 缺陷饺子外接矩形中心点坐标

标号	X	Y
1	116.5	100.0
2	105.5	103.5
3	109.5	102.0
4	102.0	103.5
5	113.5	116.5
6	111.5	110.0
7	102.5	99.5
8	97.0	124.5

经过对不同核函数准确性的分析, RGB 空间的颜色矩对于饺子缺陷检测的效果并不理想, HSV 空间的纹理特征对于饺子缺陷的特征有一定的区分度, 采用 PCA 对数据进行降维, 选取最优特征向量作为支持向量机的输入集, 能够避免模型出现过拟合的情况, 还可以提高检测效率, 最终准确检测率为 95.5%。影响饺子检测准确率的因素可能为实验时光照强度不足, 对于饺子颜色的特征不明显。

5 结语

为了使饺子加工生产更加智能化, 提出了一种以支持向量机为主, PCA 主成分分析优化特征向量为辅的饺子缺陷检测方法, 将降维之后的特征向量应用于支持向量机来对完整饺子和缺陷饺子进行分类训练, 并对支持向量机输出的预测向量中有缺陷的饺子外接最小矩形, 并输出中心点坐标, 以达到缺陷饺子定位的效果, 最终饺子检测的准确率为 95.5%, 每个饺子检测的消耗时间为 13.13 ms, 证明了该方法的可行性。该研究对于机器学习在速冻食品质量检测领域的应用提供重要的参考价值。

参考文献:

[1] 田有文, 王立地, 姜淑华. 基于图像处理和支撑向量机的玉米病害识别[J]. 仪器仪表学报, 2006 (S3): 2123-2124.

- [2] 田有文, 李天来, 李成华, 等. 基于支持向量机的葡萄病害图像识别方法[J]. 农业工程学报, 2007(6): 175-180.
- [3] 陈荣, 李旺, 周文玉. 基于形状特征和支持向量机(SVM)的茶叶病害识别方法[J]. 贵州农业科学, 2021, 49(4): 53-59.
- [4] 刘军, 郭俊先, 帕提古丽·司拉木, 等. 基于机器视觉与支撑向量机的核桃外部缺陷判别分析方法[J]. 食品科学, 2015, 36(20): 211-217.
- [5] 朱丽娟. 基于灰度共生矩阵的红枣纹理特征提取[J]. 科技风, 2022(25): 78-80.
- [6] 李亚文, 刘爱军, 陈焱. 基于 GLCM 纹理特征提取的黄瓜叶部病害检测算法研究[J]. 湖北农业科学, 2022, 61(9): 141-145.
- [7] SINGH A K, SREENIVASU S, MAHALAXMI U S B K, et al. Hybrid feature-based disease detection in plant leaf using convolutional neural network, Bayesian optimized SVM, and random forest classifier[J]. Journal of Food Quality, 2022, 2022: 1-16.
- [8] KUMAR S S, PANDEY M. An optimal hybrid multiclass SVM for plant leaf disease detection using spatial Fuzzy C-Means model[J]. Expert Systems With Applications, 2023, 214: 118989.
- [9] 魏万珍, 田建艳, 姜茜, 等. 基于图像技术的零件表面粗糙度支持向量机检测模型[J]. 现代制造工程, 2022(5): 100-109, 114.
- [10] 楚松峰, 赵凤霞, 方双, 等. 基于 PCA-SVM 的红枣缺陷识别方法[J]. 食品与机械, 2021, 37(1): 156-160, 198.
- [11] 刘丽, 匡纲要. 图像纹理特征提取方法综述[J]. 中国图象图形学报, 2009, 14(4): 622-635.
- [12] 郭刚, 马云龙, 徐铭辰. 不同品种抹茶营养物质和抗氧化活性的测定及主成分分析[J/OL]. 食品与发酵工业: 1-10 [2023-02-23]. DOI: 10.13995/j.cnki.11-1802/ts.033644.
- [13] 周寅聪, 宋淑鸿. 基于 PCA 与因子分析法的财务绩效评价: 以道路运输业为例[J]. 物流工程与管理, 2022, 44(11): 116-119.

(下转第 39 页)

文章编号: 1007-1423(2023)24-0010-08

DOI: 10.3969/j.issn.1007-1423.2023.24.002

基于 Improving-SRGAN 算法的岩心图像超分辨率重建

罗旺健¹, 张豫堃², 吴晓红¹, 何小海¹, 熊淑华^{1*}

(1. 四川大学电子信息学院, 成都 610065; 2. 成都西图科技有限公司软件部, 成都 610065)

摘要: 岩心图像是油气勘探、地质分析的宝贵资料, 由于部分旧设备、低分辨率设备采集到的图像细节不够, 或因压缩过程中造成的图像细节丢失, 容易影响地质工作者的判断, 从而影响工作进度。为了获得更加真实、重建视觉效果更好的岩心图像, 提出了一种改进的岩心图像超分辨率重建算法 Improving-SRGAN。该算法采用生成对抗网络框架, 将原生成器模型中的残差模块替换为残差注意力 RAM 模块。RAM 模块可以对输入的特征图进行空间域和通道域加权, 以更好地提取图像的高层次语义信息。同时在判别器中嵌入空洞空间金字塔池化 ASPP 模块, 提升网络对岩心图像多尺度信息的感知能力, 从而更准确地评估生成器生成图像的质量。相比 SRGAN 算法, 所提出的算法在自建岩心数据集上的实验数据表明, 其 PSNR 和 SSIM 分别提高了 1.469、0.074; 在 DIV2K_Valid 数据集上的 PSNR 和 SSIM 分别提高了 0.721、0.017, 且算法的 MOS 得分更高, 重建视觉效果优于其他算法。

关键词: 岩心图像; 超分辨率; 空洞卷积; 残差注意力

0 引言

岩心图像是油气勘探的宝贵分析资料, 随着工业相机的迅速发展以及图像处理算法的不断创新, 岩心图像分析工作对岩心图像的质量提出了更高的要求。目前的岩心图像采集系统一般使用面阵或线阵相机来采集岩心荧光或白光图像^[1], 由于部分旧设备、低分辨率采集设备采集到的图像细节不够, 或因压缩过程中造成的图像细节丢失, 不利于研究人员后续研究工作的深入开展。因此对获取的岩心图像进行适当的超分辨率重建, 有助于研究人员获得更加丰富、纹理更加清晰的高分辨率岩心图像。

超分辨率重建是图像处理领域的一个重要研究方向, 目前已广泛应用在工业制造、医学影像和遥感图像处理等领域^[2]。2014 年, Dong 等^[3]提出了一种基于深度卷积神经网络的超分

辨率重建算法 SRCNN, 虽然优于传统方法, 但是由于网络较浅, 感受野较小, 无法有效提取图像的全局信息。2017 年, Lim 等^[4]提出了增强深度超分辨率网络 EDSR。其删除了传统网络不必要的模块, 模型紧凑的同时提升了模型效果, 不仅比当时其他方法具有更好的性能, 还大大降低了模型的复杂度。2017 年, Ledig 等^[5]首次提出一种基于生成对抗网络的超分辨率重建算法 SRGAN, 其主要思想是通过对抗训练来生成高质量的超分辨率图像。2018 年, Wang 等^[6]提出了一种增强型超分辨率生成对抗网络 ESRGAN, 通过引入更深的网络和残差注意力来提高网络性能。2021 年, Kai 等^[7]提出了一种盲超分辨率生成对抗网络 BSRGAN, 设计了一种退化模型, 使用各种退化策略来覆盖现实场景中的退化情况, 为解决盲 SR 问题提供了一种

收稿日期: 2023-08-16 修稿日期: 2023-09-12

基金项目: 国家自然科学基金(62071315)

作者简介: 罗旺健(1999—), 男, 河南上蔡人, 硕士, 研究方向为计算机视觉、图像处理与网络通信; 张豫堃(1986—), 男, 四川成都人, 学士, 研究方向为智能控制、图像处理及应用; 吴晓红(1970—), 女, 四川成都人, 博士, 副教授, 硕士生导师, 研究方向为图像处理与模式识别、计算机视觉、无线通信; 何小海(1964—), 男, 四川成都人, 博士, 教授, 博士生导师, 研究方向为图像处理与网络通信、人工智能与大数据分析、图像通信; *通信作者: 熊淑华(1969—), 女, 四川成都人, 博士, 副教授, 硕士生导师, 研究方向为多媒体通信, E-mail: xiongsh@scu.edu.cn

新的方法。2021年, Liang等^[8]提出了一种基于Swin变换的强基线模型SwinIR。SwinIR在重建模块中融合浅层和深层特征, 进行高质量的图像重建, 但生成的图像过于平滑。

基于以上分析, 为解决岩心图像超分辨率过程中重建图像细节不够清晰、依赖配对数据集等问题, 对SRGAN网络进行探索, 提出了基于Improving-SRGAN的岩心图像超分辨率重建网络。首先采用残差注意力RAM(residual attention module)模块替代SRGAN网络中的残差模块, 该模块的主要作用是对输入的特征图进行通道和空间注意力加权, 以便更好地提取图像的高层次语义信息。同时在模型的判别器中加入了空洞空间金字塔池化ASPP(atrous spatial pyramid pooling)模块, 以提升模型对岩心图像特征多尺度信息的感知能力, 使判别器更好地理解输入图像的全局结构, 从而更准确地评估生成器生成图像的质量。最后在自建岩心图像数据集和公共数据集上对所提算法进行验证, 并采用客观和主观评价指标对重建图像的质量进行了评价。

1 基于 Improving-SRGAN 的图像超分辨率重建算法

1.1 Improving-SRGAN 网络结构

与传统的重建方法不同, SRGAN算法可以

生成更加真实和细节丰富的高分辨率图像, 增强了图像感官上的真实感, 有效解决了以往算法重建中图像过于平滑的问题, 其核心思想是通过对抗训练来生成高质量的超分辨率图像。

为了更好地实现岩心图像的超分辨率重建, 提出了一种重建视觉效果更加优秀的超分辨率网络Improving-SRGAN。该网络以生成对抗网络SRGAN为基础, 首先将生成器中的残差块替换为RAM模块, 以便更好地提取图像的高层次语义信息; 其次在判别器中加入了ASPP模块, 提升网络对岩心图像多尺度信息的感知能力。Improving-SRGAN的网络结构如图1所示, 其中生成器中的黄色部分为RAM模块, 判别器中的青色部分为ASPP模块。

1.2 残差注意力模块(RAM)

残差注意力RAM模块是一种新的注意力模型^[9], 可以对输入的特征图进行加权, 以便更好地提取图像的高层次语义信息, 专门用来针对SR问题的优化。该模块由通道注意力机制CA和空间注意力机制SA以及结合这两者的融合注意力机制FA组成, 并将融合注意力机制FA和残差网络结构ResNet相结合来提出RAM模块, 其网络结构如图2所示。

设 F_{r-1} 和 F_r 为第 r 个RAM的输入和输出特

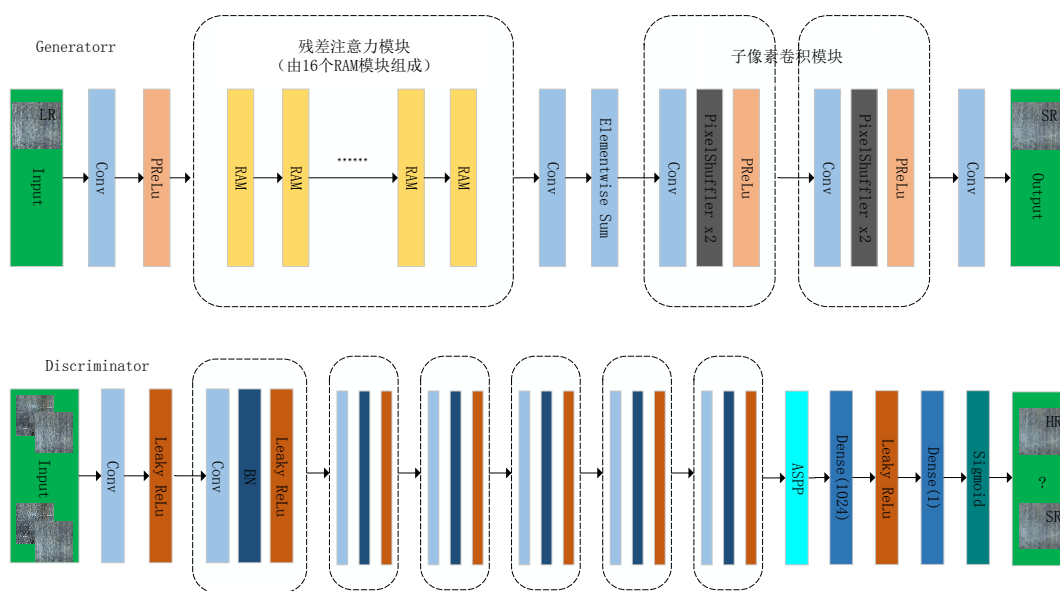


图1 Improving-SRGAN网络结构

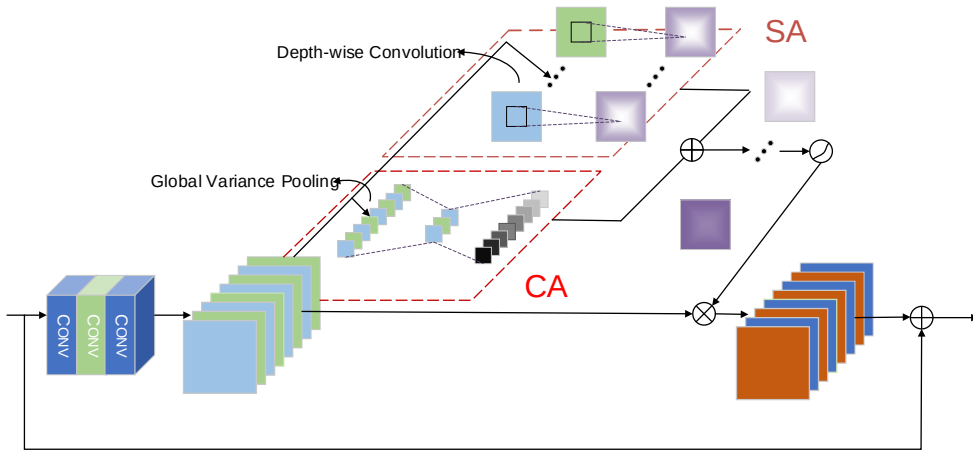


图2 残差注意力模块网络结构

征图。 F_r 可以公式化为

$$F_r = R_r(F_{r-1}) = F_{r-1} + f^{\text{FA}}(f_{\text{trans}}(F_{r-1})) \quad (1)$$

其中 $f_{\text{trans}}()$ 表示由卷积和ReLU组成的函数， $f^{\text{FA}}()$ 指FA机制。

1.2.1 通道注意力(CA)

在RAM模块中通道注意力CA通过计算每个通道大小为 $W \times H$ 的特征图的方差来获取通道的注意力权重，从而提升超分辨率图像的质量。这种选择方差作为池化方法的通道注意力方法可以更好地捕捉图像的高频细节信息，从而产生更好的超分辨率结果。

$$S_{r-1} = \text{pool}_{\text{var}}(X_{r-1}) \quad (2)$$

$$X_{r-1} = f_{\text{trans}}(F_{r-1}) \quad (3)$$

其中 $\text{pool}_{\text{var}}()$ 表示方差池， $S_{r-1} \in \mathbf{R}^{1 \times 1 \times C}$ 表示输出统计量。

1.2.2 空间注意力(SA)

在空间注意力机制SA中，针对超分辨率问题中的不同空间区域，其通道的重要性可能会有所不同。例如，在边缘或复杂纹理特征图中，复杂滤波器的信息更加重要，而在几乎没有高频分量的区域，相对较不详细的信息更加重要。通过使用深度卷积可以为每个通道生成不同的SA映射。

$$M_{r-1}^{\text{SA}} = \text{conv}_{3 \times 3, \text{depth}}(X_{r-1}) \quad (4)$$

其中 $\text{conv}_{3 \times 3, \text{depth}}()$ 表示 3×3 深度卷积。

1.2.3 融合注意力(FA)

CA和SA机制可以分别捕捉输入特征映射中的信道间与信道内关系，通过添加CA和SA映射将它们结合起来，并使用Sigmoid函数进行缩放处理，以重新校准特征映射 X_{r-1} ：

$$\bar{X}_{r-1} = f^{\text{FA}}(X_{r-1}) = \sigma(M_{r-1}^{\text{CA}} \oplus M_{r-1}^{\text{SA}}) \otimes X_{r-1} \quad (5)$$

其中， $\oplus$ 和 $\otimes$ 分别表示元素相加和乘积。

最后，将通道和空间注意力图相加，并经过Sigmoid函数处理后将其与原始特征相乘，再与残差网络结合便得到图2所示的残差注意力RAM模块的网络结构。本文将原生成器模型中的残差块替换为残差注意力RAM模块，可以有效提取岩心图像的高频特性，使得生成器生成更加清晰、真实的高分辨率岩心图像。

1.3 空洞空间金字塔池化(ASPP)模块

空洞卷积(dilated convolution)是一种卷积神经网络中常用的卷积操作。与传统的卷积操作不同的是，空洞卷积在卷积核中插入了一些空洞，可以在不增加卷积核大小的情况下，增加卷积操作的感受野，从而提高网络特征提取的能力。假设卷积核尺寸为 k ，空洞率为 $rate$ ，则空洞卷积感受野 N 的计算如式(6)所示：

$$N = (rate - 1) \times (k - 1) + k \quad (6)$$

图3中展示了4个卷积核尺寸为3，空洞率分别为1, 2, 3, 4的空洞卷积核。图中红色是目标像素，黄色是需要进行卷积运算的像素。

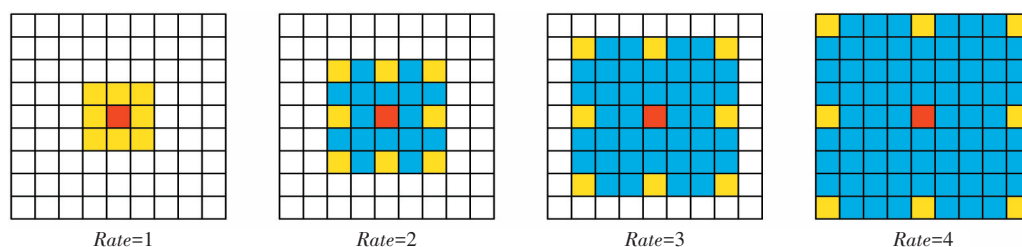


图3 不同空洞率的空洞卷积

ASPP模块的主要思想是利用多个具有不同的空洞率的并行空洞卷积层，对输入特征图进行采样，以获取不同尺度的上下文信息^[10]。同时还引入了空间金字塔池化，对输入特征图进行全局池化操作，以得到整个图像的全局信息。最后将这两者的特征图进行融合，形成最终的ASPP特征表示。ASPP模块可以帮助网络获取更多的上下文信息，提高图像处理的准确性和效果。ASPP模块的网络结构如图4所示。

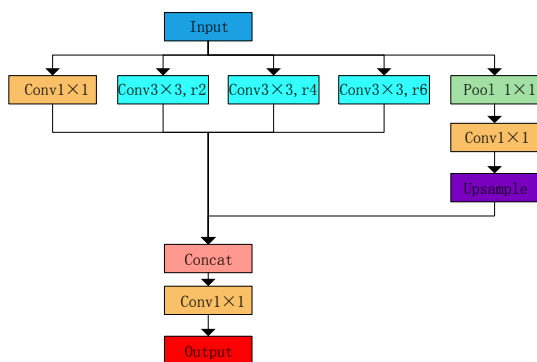


图4 ASPP网络结构

在SRGAN网络的判别器中添加ASPP模块，可以增加网络感受野，使判别器更好地理解输入岩心图像的全局结构，从而更准确地评估生成器生成的岩心SR图像的质量。但将ASPP模块添加到判别器中也会导致训练时间增加，因为该模块需要更多的计算资源。同时，经过多次实验调整，将ASPP空洞卷积的空洞率分别设置为2, 4, 6，以确保网络的性能得到最大化。

2 实验结果与分析

为了验证所提算法的有效性，在自建岩心数据集上进行了不同算法的对比试验；同时为

了评估所提算法的泛化能力，在DF2K数据集上进行了不同算法的对比试验。在这2组实验中，对比的算法包括Bicubic、SRCNN、EDSR、SwinIR、SRRESNet、SRGAN和ESRGAN。另外，为了验证所加入的各个模块的有效性，进行了消融实验。

2.1 实验环境及参数设置

所有实验都在Ubuntu 18.04.1系统下进行，CPU为Intel(R) Core(TM)i7-9700，主频2.30 GHz，显卡为NVIDIA GeForce RTX3090，显存为24 GB。实验采用的深度学习框架为PyTorch1.8.0。训练时使用Adam优化算法更新网络参数，初始学习率为0.0001，批大小为32，迭代次数为100次，网络放大倍数为4倍。

2.2 数据集制作

自建岩心数据集使用数字化岩心图像采集系统所拍摄，从拍摄的图像中剔除纹理不清晰、色彩偏暗的图像，然后裁剪成1000 × 1000大小的图像。接着人工筛选出岩心占比大于80%区域的图像，最后得3450张岩心GT图像，其中3360张作为训练集，90张用作测试集。

DF2K数据集由DIV2K和Flickr2K合并得到的。DIV2K数据集包含900张2K分辨率的图片，Flickr2K数据集包含2650张2K分辨率的图片。训练时联合使用DIV2K_Train和Flickr2K两部分数据，验证时使用DIV2K_Valid数据集进行验证。

上述两个数据集训练时从训练集HR图像中随机裁剪128 × 128的图像块作为原始HR图像，使用bicubic插值将其下采样4倍，生成的32 × 32的图像块作为原始LR图像。在验证时将验证集HR图像使用bicubic插值下采样4倍为LR图像进行验证。

2.3 评价指标

为验证所提算法的有效性,采用峰值信噪比 $PSNR$ 和结构相似性 $SSIM$ 两种客观评价指标以及平均意见得分 MOS 测试一种主观评价指标来进行验证。

2.3.1 峰值信噪比($PSNR$)

峰值信噪比 $PSNR$ 是衡量生成图像质量的客观评价指标之一。它通过计算原始和重建图像之间的均方误差来衡量它们的相似度。 $PSNR$ 的值越高,表示重建与原始图像的相似度越高,其计算公式为

$$PSNR = 10 \times \log_{10} \left(\frac{\text{Max}I^2}{MSE} \right) = 20 \times \log_{10} \left(\frac{\text{Max}I}{\sqrt{(MSE)}} \right) \quad (7)$$

$$MSE = \frac{1}{M \times N} \sum_{i=1}^N \sum_{j=1}^M (f_{ij} - f'_{ij})^2 \quad (8)$$

$PSNR$ 表达式中均方误差 MSE 的计算如式(8)所示, M 和 N 为图像长和宽, I 为参考图像, K 代表重建图像, Max 是图像像素点颜色最大值,在本文中等于 255。

2.3.2 结构相似度($SSIM$)

结构相似度 $SSIM$ 是另一种常用的图像质量评价指标,从亮度、对比度和结构三个方面衡量图像的相似性, $SSIM$ 数值越高,表示重建与原始图像的相似度越高。假设输入的两张图像分别是 x 和 y , 那么其计算公式为

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (9)$$

其中, μ_x 和 μ_y 为 (x, y) 的平均值, σ_x 和 σ_y 为 (x, y) 的标准差, σ_{xy} 为 (x, y) 的协方差, c_1, c_2, c_3 分别为常数。

2.3.3 平均意见得分(MOS)测试

平均意见得分 MOS 测试,是人们通过自己的视觉感受对重建图像进行量化评估,为验证重建图像的质量,邀请了两名具有多年岩心地质工作经验的专家通过视觉感受对重建图像进行评估打分,其图像质量打分结果见表 1。

表 1 图像质量 MOS 测试打分表

图像质量评估标准	得分
该图像质量最差	1
该图像质量低于平均水平	2
该图像质量处于平均水平	3
该图像质量高于平均水平	4
该图像质量最好	5

在该主观评价任务中,开发了一个用于 MOS 测试的评分系统界面,每次待评价图像包括 LR、HR 图像和不同算法重建的 SR 图像。测试过程中图像的秩序被打乱,测试人员依据个人视觉感受对不同算法重建的 SR 图像进行打分,测试完成后计算不同算法最终的平均 MOS 得分,系统界面如图 5 所示。

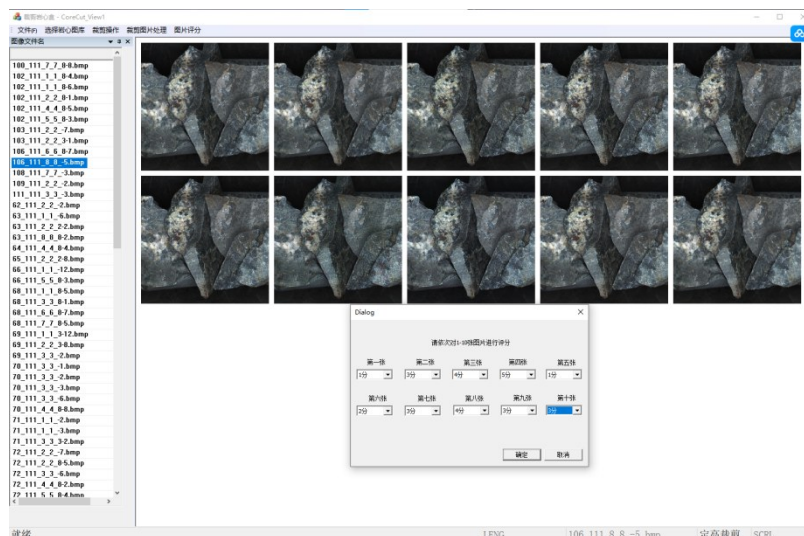


图 5 MOS 评分系统界面

2.4 实验结果及分析

2.4.1 Improving-SRGAN 的有效性验证

为了验证所提算法的有效性，在岩心数据集上与主流超分辨率算法进行了对比。图6为各个算法的重建可视化结果，表2为各个算法的重建指标结果。实验结果表明，在岩心图像数据集上，Bicubic、SRCNN、EDSR、SWINIR 以及 SRRESNet 算法取得了较高的 *PSNR* 和 *SSIM* 指标，但其重建图像的可视化效果较差，原因在于这类算法一般将 *MSE* 作为损失函数，导致其重建图像与原图像相比缺乏细节和真实感。总体来看，以生成对抗网络为主体的 SRGAN 和 ESRGAN 算法的重建视觉效果要优于其他网络，但其重建图像与原图像相比具有一定的色彩差异。对于岩心图像细节的恢复，所提出的 Improving-SRGAN 算法重建的岩心图像具有更高的视觉效果，图像细节更加丰富，*MOS* 得分最高。原因在于在生成器中加入了 RAM 模块，其能有效提取岩心图像的高层次语义信息；在判别器中加入的 ASPP 模块能使判别器更准确地评估生成器生成图像的质量。同时其 *PSNR* 指标比 SRGAN、ESRGAN 分别提高了 1.469、1.238，

SSIM 指标分别提高了 0.074、0.069。综合主观和客观结果表明，所提出的 Improving-SRGAN 算法在岩心图像超分辨率重建上表现出了较优的性能。

表 2 岩心图像数据集的评价指标对比结果

Algorithm	<i>PSNR</i> /dB	<i>SSIM</i>	<i>MOS</i>
Bicubic	28.516	0.731	1.43
SRCNN	31.137	0.793	2.52
EDSR	31.307	0.798	3.45
SRRESNET	31.241	0.795	3.20
SWINIR	30.821	0.777	3.58
SRGAN	27.401	0.638	4.14
ESRGAN	27.632	0.643	4.28
Improving-SRGAN	28.870	0.712	4.54

2.4.2 Improving-SRGAN 的泛化性能评估

为了评估所提算法的泛化性，在 DIV2K_Valid 数据集上对比了不同算法的重建结果。图7为各个算法的重建可视化结果，结果表明所提算法重建的图像较为真实、细节丰富，重建效果更好，*MOS* 得分更高。

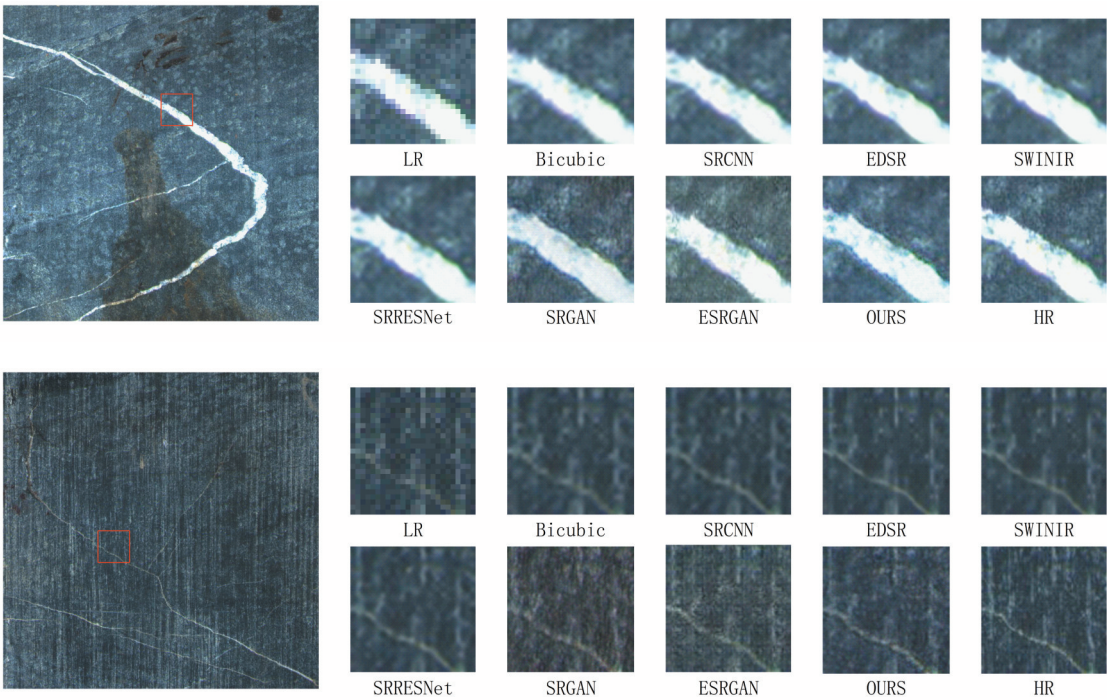


图 6 4倍放大因子下各算法 SR 重建效果的可视化对比



图7 4倍放大因子下各算法SR重建效果的可视化对比

表3是各个算法在DIV2K_Valid数据集上的指标结果,相比于重建效果较好的SRGAN和ESRGAN,其 $PSNR$ 指标分别提高了0.721、0.888, $SSIM$ 指标分别提高了0.017、0.052,综合考虑主观和客观结果,验证了所提算法具有一定的泛化性。

表3 公共图像数据集的评价指标对比结果

Algorithm	$PSNR/dB$	$SSIM$	MOS
Bicubic	26.695	0.766	1.45
SRCNN	29.121	0.818	2.56
EDSR	29.881	0.837	3.63
SRRESNET	29.486	0.830	3.15
SwinIR	30.589	0.841	3.72
SRGAN	26.844	0.762	4.25
ESRGAN	26.677	0.727	4.47
Improving-SRGAN	27.565	0.779	4.62

2.5 消融实验

为了验证Improving-SRGAN网络各个模块的有效性,在岩心图像数据集上进行了四组消融实验,第一组实验使用的网络为SRGAN;第二组在第一组网络的基础上将生成器的残差块替换为RAM模块;第三组在第一组的基础上在判别器中加入ASPP模块;第四组在第二组的基础

上在判别器中加入ASPP模块,实验的训练参数均保持一致。消融实验的具体结果见表4,可以看出两种模块均使得网络的重建指标有了一定的提升。

表4 不同模块的消融实验评价指标对比结果

分组	RAM	ASPP	$PSNR/dB$	$SSIM$	单张所用时间/s	模型大小/MB
第一组	N	N	27.401	0.638	0.153	107.87
第二组	Y	N	28.450	0.701	0.179	108.94
第三组	N	Y	27.590	0.657	0.143	141.89
第四组	Y	Y	28.870	0.712	0.187	142.95

3 结语

针对SRGAN网络在岩心图像重建中不能有效恢复岩心图像细节的问题,将生成器中的残差块替换为RAM模块,以便更好地提取图像的高层次语义信息。在判别器中加入了ASPP模块以融合不同尺度的图像特征,使判别器更好地理解输入图像的全局结构,从而更准确地评估生成器生成图像的质量。在岩心数据集上进行的实验表明,对于岩心图像细节的恢复,所提算法重建的岩心图像具有更好的视觉效果,在DIV2K_Valid数据集上进行的实验则验证了所提算法具有一定的泛化性。同时进行的消融实验验证了所提出的各个模块的有效性。但由于

RAM 和 ASPP 模块的引入, 导致模型在训练和测试时会占用大量的显存和时间, 在后续的工作中, 将致力于在保证重建质量的前提下探索模型的轻量化研究。

参考文献:

- [1] 胡明. 一体化岩心采集系统图像后处理关键技术的研究与实现[D]. 成都: 四川大学, 2015.
- [2] SHAH H Z, MÜLLER M, WANG T, et al. Deep-learning based denoising and reconstruction of super-resolution structured illumination microscopy images[J]. *Photonics Research*, 2021, 9(5): 570-583.
- [3] CHAO D, CHEN C L, HE K, et al. Learning a deep convolutional network for image super-resolution [C]//*Proceedings of the European Conference on Computer Vision*, Springer, Cham, 2014: 184-199.
- [4] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution [C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2017: 136-144.
- [5] LEDIG C, THEIS L, F HUSZAR, et al. Photo-realistic single image super-resolution using a generative adversarial network [C]//*Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 105-114.
- [6] WANG X, YU K, WU S, et al. ESRGAN: enhanced super-resolution generative adversarial networks [C]//*Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, 2018: 63-79.
- [7] ZHANG K, LIANG J, VAN GOOL L, et al. Designing a practical degradation model for deep blind image super-resolution [C]//*Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021: 4791-4800.
- [8] 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.
- [9] KIM J H, CHOI J H, CHEON M, et al. MAMNet: multi-path adaptive modulation network for image super-resolution [J]. *Neurocomputing*, 2020, 402: 38-49.
- [10] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 40 (4) : 834-848.

Super-resolution reconstruction of core images based on the improving-SRGAN algorithm

Luo Wangjian¹, Zhang Yukun², Wu Xiaohong¹, He Xiaohai¹, Xiong Shuhua^{1*}

(1. College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China;

2. Software Department, Chengdu Xitu Technology Co. Ltd., Chengdu 610065, China)

Abstract: Core images are valuable data for oil and gas exploration and geological analysis, and due to insufficient image details collected by some old equipment and low-resolution acquisition equipment, or the loss of image details caused by the compression process, it is easy to affect the judgment of geologists, thereby affecting the progress of work. In order to obtain more realistic and better reconstructed high-resolution core images, an improved core image super-resolution reconstruction algorithm Improving-SRGAN is proposed. The algorithm adopts the generative adversarial network framework, and replaces the residual attention module in the original generator model with the residual attention module (Residual Attention Module). The RAM module can weight the spatial domain and channel domain of the input feature map to better extract the high-level semantic information of the image. At the same time, the Atrous Spatial Pyramid Pooling ASPP module is embedded in the discriminator to improve the network's perception of the multi-scale information of core images, so as to more accurately evaluate the quality of the images generated by the generator. Compared with the SRGAN algorithm, the experimental data of the proposed algorithm on the self-built core dataset show that the PSNR and SSIM are increased by 1.469 and 0.074, respectively, while the PSNR and SSIM on the DIV2K_Valid dataset are increased by 0.721 and 0.017, respectively, and the MOS score of the algorithm is higher, and the reconstruction visual effect is better than that of other algorithms.

Keywords: core images; super-resolution; dilated convolution; residual attention

文章编号: 1007-1423(2023)24-0018-08

DOI: 10.3969/j.issn.1007-1423.2023.24.003

基于用户长短期偏好及时空场景的下一个兴趣点推荐

郑宏洲, 曾国荪*

(同济大学计算机科学与技术系, 上海 200092)

摘要: 在现实生活中, 用户对兴趣点的偏好会受到时空场景的影响, 用户希望获得匹配当前时间的推荐结果。由此, 提出基于用户长短期偏好及时空场景的下一个兴趣点推荐模型。该模型围绕实时兴趣点推荐这一问题, 从用户的长短期偏好两方面来挖掘用户的实时兴趣偏好。对于长期偏好, 从历史数据中收集与当前时空场景最相关的信息。对于短期偏好, 在序列影响的基础上考虑时间推移影响。在公开数据集上的实验结果证明了方法的有效性。

关键词: 兴趣点推荐; 长短期偏好; 时空场景信息; 偏好推荐模型

0 引言

近年来, 随着移动互联网的发展, 基于位置的社交网络(location-based social network, LBSN)逐渐成为人们签到、发现兴趣点的重要平台。尽管用户可以在LBSN上获得海量的兴趣点信息, 但是要从中找到可能感兴趣的兴趣点作为下一步决策, 却是十分困难且耗时的。缓解这一问题的解决方案之一是根据用户的历史签到数据建立用户偏好模型, 为用户推荐下一个兴趣点。在现实生活中, 用户对兴趣点的偏好会受到当前情境的影响, 其中时间是一个很重要的影响因素。例如, 用户在健身后, 在一段时间内对附近的餐饮类地点兴趣较高, 随着时间推移兴趣逐渐下降, 这体现了用户受时间推移影响的短期偏好; 用户在上午对博物馆、茶馆等兴趣点比较感兴趣, 而晚上更喜欢去酒吧消遣, 体现了用户在不同时间段的长期偏好。因此推荐结果要能够匹配用户当前时间的兴趣。

为此, 本文在现有研究的基础上, 探索基于用户长短期偏好及时空场景的下一个兴趣点推荐模型(Next Point of Interest Recommendation

Based on User's Long and Short Term Preference and spatio-temporal context, NP-ULSP)。该模型围绕实时兴趣点推荐这一问题, 从用户的长短期偏好两方面来挖掘用户的实时兴趣偏好。在建模用户的长期偏好时, 考虑了时空影响, 通过时间段相似性加权操作和注意力机制来对用户历史签到行为对当前时空场景的影响进行有效表示。在建模用户的短期偏好时, 在考虑序列影响的基础上引入时间转移矩阵来表示最后一次签到至当前时间的时间推移影响。最后综合两种偏好为用户进行实时兴趣点推荐。

1 相关工作

1.1 传统的兴趣点推荐

在兴趣点推荐的早期工作中, 根据兴趣点的流行度^[1]为用户进行推荐, 为用户提供热门兴趣点的参考。然而, 这种方法可能会造成数据倾斜, 并且无法满足用户的个性化需求。因此, 人们开始尝试将传统的推荐方法应用到兴趣点推荐领域, 例如协同过滤^[2-3], 矩阵分解^[4]等。但是这些方法将用户的历史签到记录视为

收稿日期: 2023-09-13 修稿日期: 2023-10-16

作者简介: 郑宏洲(1997—), 男, 广东汕头人, 硕士研究生, 研究方向为推荐系统、数据挖掘; *通信作者: 曾国荪(1964—), 男, 江西吉安人, 博士, 教授, 博士生导师, 主要研究方向为并行分布计算、可信软件、信息安全, E-mail: gszeng@tongji.edu.cn

一个集合，没有深入挖掘签到记录之间的顺序关系。研究表明，用户下一个签到的兴趣点会受到之前的签到行为的影响，具有一定的顺序关系。Rendel等^[5]结合矩阵分解和马尔可夫链来进行推荐，矩阵分解学习用户的长期偏好，马尔可夫链通过学习兴趣点之间的转移矩阵模拟用户的顺序行为。

1.2 基于深度学习的兴趣点推荐

随着深度学习技术的发展，基于深度学习的兴趣点推荐逐渐成为研究的主流，例如循环神经网络。Hidasi等^[6]利用门控循环单元进行短序列推荐。Liu等^[7]将连续两次签到之间的时空间隔编码为嵌入向量，参与循环隐藏单元状态的计算，捕获序列中复杂的时空依赖影响。Zhu等^[8]在LSTM中增加了时间门，用以表示连续签到之间的时间间隔影响。Sun等^[9]采用两个LSTM学习用户的长短期偏好，包括一个用于长期偏好建模的非局部网络和一个用于短期偏好学习的地理扩张网络。

2 问题描述

用 $U = \{u_1, u_2, \dots, u_{|U|}\}$ 表示 LBSN 中的用户集合， $L = \{l_1, l_2, \dots, l_{|L|}\}$ 表示兴趣点集合，每一个兴趣点有一个经纬度坐标 (lon_l, lat_l) 。对于每个用户 $u \in U$ ，将时间戳处于同一天内的签到记录编入一个签到轨迹中，可以得到一个签到轨迹集合 $S = \{S_1, S_2, \dots, S_n\}$ ，其中下标 n 表示当前天。集合中的签到轨迹 $S_m = \{l_1, l_2, \dots, l_{|S_m|}\}$ 是用户某一天签到的兴趣点序列。我们将 $S_i \in \{S_1, S_2, \dots, S_{n-1}\}$ 称为历史签到轨迹，将 S_n 称为当前签到轨迹。特别地，为了区分历史签到轨迹和当前签到轨迹，将当前签到轨迹表示为 $S_n = \{l_1, l_2, \dots, l_{p-1}\}$ ，对应的时间戳序列为 $T_n = \{t_1, t_2, \dots, t_{p-1}\}$ 。

下一个兴趣点推荐问题可以描述为：对于目标用户 $u \in U$ ，给定签到轨迹集 S ，为用户推荐当前时间 t_p 可能感兴趣的兴趣点。

3 基于用户长短期偏好及时空场景的下一个兴趣点推荐模型 NP-ULSP

本节将介绍本文提出的 NP-ULSP 模型，模

型的总体架构如图1所示。

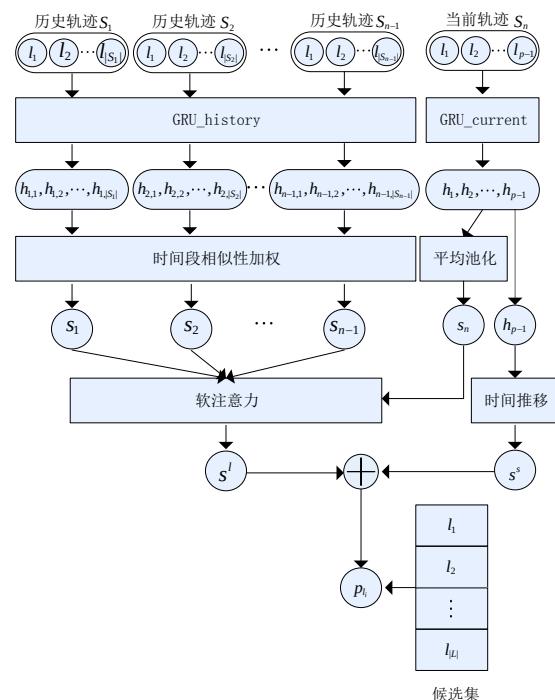


图1 NP-ULSP 总体架构

3.1 门控循环单元网络

本文利用门控循环单元(gated recurrent unit, GRU)提取签到轨迹的序列特征。GRU是一种改进型的循环神经网络，Hidasi等^[6]证明了GRU在序列推荐任务上比RNN、LSTM具有更好的性能。本文采用的GRU表达式如下：

$$z_t = \text{Sigmoid}(\mathbf{W}_z x_t + \mathbf{U}_z h_{t-1} + b_z) \quad (1)$$

$$r_t = \text{Sigmoid}(\mathbf{W}_r x_t + \mathbf{U}_r h_{t-1} + b_r) \quad (2)$$

$$\tilde{h}_t = \text{Tanh}(\mathbf{W}_{\tilde{h}} x_t + \mathbf{U}_{\tilde{h}} (r_t \odot h_{t-1}) + b_{\tilde{h}}) \quad (3)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (4)$$

其中， x_t 是当前输入。 b_* 为偏置， $\mathbf{W}_*$ 、 $\mathbf{U}_*$ 为权重矩阵参数。 $\odot$ 表示哈达玛积，也就是对应元素相乘。 r_t 和 z_t 分别是重置门和更新门的输出。 h_t 是当前时间步的隐藏状态。

3.2 长期偏好建模

用户的历史签到轨迹中蕴含着用户的长期偏好信息，例如用户习惯在某个时间段访问某些特定兴趣点。一个直观的想法是根据当前的情境从历史签到轨迹集中选择性地收集信息，因此我们利用基于时间段相似性的加权操作来

学习历史签到轨迹的表示。然后,根据历史签到轨迹和当前签到轨迹在空间上的距离关系将这些签到轨迹的表示整合成用户的长期偏好表示。做法如下:

(1) 首先,对于用户的每个历史签到轨迹 $S_m \in \{S_1, S_2, \dots, S_{n-1}\}$, 我们将其输入到 GRU 网络, 记为 GRU_history, 提取序列特征:

$$h_{m,i} = \text{GRU_history}(e_{m,i}, h_{m,i-1}),$$

$$m \in \{1, 2, \dots, n-1\}, i \in \{1, 2, \dots, |S_m|\}$$
(5)

其中, $h_{m,i} \in \mathbf{R}^{d \times 1}$ 表示用户第 m 个签到轨迹中第 i 次签到对应的隐藏状态, 可作为用户在这个时间步的偏好表示; $e_{m,i} \in \mathbf{R}^{d \times 1}$ 表示用户第 m 个签到轨迹中第 i 次签到的兴趣点对应的嵌入向量。这样, 签到轨迹 S_m 被编码为 $\{h_{m,1}, h_{m,2}, \dots, h_{m,|S_m|}\}$ 。

(2) 接下来, 从每一个历史签到轨迹中整合信息, 获得对应的总体表示。在现实生活中, 人们对兴趣点的偏好会随着所处时间段的不同而发生变化。例如, 人们可能在上午对博物馆、茶馆等兴趣点比较感兴趣, 而晚上更喜欢去酒吧消遣。因此, 每个签到轨迹 S_m 的总体表示 s_m 在不同时间段应有不同的表示, 以捕获时间敏感性。为了实现这一点, 我们使用一种基于时间段相似性的加权操作。

具体来说, 将一天的时间划分为 24 个时间段。对于每个时间段 q , 可以构造一个签到集合 $H_q = \{l_1, l_2, \dots, l_{l_q}\}$, 其中 $l \in L$ 是在时间段 q 内所有用户签到过的兴趣点。时间段 q 和 w 的相似度可以计算如下:

$$\tau_{q,w} = \frac{|H_q \cap H_w|}{|H_q \cup H_w|}$$
(6)

两个时间段签到的兴趣点重叠越多, 则两个时间段的相似度就越高。对于某个历史签到轨迹 S_m , 可以根据每个兴趣点的签到时间推导出对应的签到时间段序列 $\{seg_1, seg_2, \dots, seg_{|S_m|}\}$, $seg \in \{1, 2, \dots, 24\}$ 。根据用户的当前时间可推导出当前时间段 c 。某个历史签到轨迹 S_m 的总体表示 s_m 为

$$s_m = \sum_{i=1}^{|S_m|} w_i h_{m,i}, w_i = \frac{\exp(\tau_{c, seg_i})}{\sum_{j=1}^{|S_m|} \exp(\tau_{c, seg_j})}$$
(7)

其中, τ_{c, seg_j} 为当前时间段 c 与 S_m 中第 j 次签到的时间段相似性。通过基于时间段相似性的加权操作, 用户的某个历史签到轨迹 S_m 与用户当前时间段建立联系。我们用 $\{s_1, s_2, \dots, s_{n-1}\}$ 表示 $n-1$ 个历史签到轨迹的总体表示。

(3) 对于当前轨迹, 利用不同的 GRU 网络, 记为 GRU_current, 将当前签到轨迹输入到网络中得到对应的偏好向量:

$$h_i = \text{GRU_current}(e_i, h_{i-1}), i \in \{1, 2, \dots, p-1\}$$
(8)

其中, $h_i \in \mathbf{R}^{d \times 1}$ 表示用户当前签到轨迹中第 i 次签到对应的隐藏状态; $e_i \in \mathbf{R}^{d \times 1}$ 表示用户当前签到轨迹中第 i 次签到的兴趣点对应的嵌入向量。这样, 当前签到轨迹 S_n 就被编码为 $\{h_1, h_2, \dots, h_{p-1}\}$ 。使用式(9)获得当前签到轨迹的总体表示:

$$s_n = \frac{1}{|S_n|} \sum_{i=1}^{|S_n|} h_i$$
(9)

对于当前签到轨迹, 我们并没有使用基于时段相似性的加权操作, 而是采用平均池化。这是考虑当前签到轨迹的签到信息对用户偏好的影响较大, 加权操作可能会丢失信息, 因此这里采用了平均池化的操作。

(4) 在已经得到 $n-1$ 个历史签到轨迹的总体表示 $\{s_1, s_2, \dots, s_{n-1}\}$ 以及当前签到轨迹的总体表示 s_n 之后, 接下来需要将这些轨迹的总体表示进行整合, 得到用户长期偏好的表示。计算公式如下:

$$s^l = \sum_{m=1}^{n-1} \varphi_m s_m$$
(10)

φ_m 是对过去签到轨迹的加权, 权重的计算考虑了当前签到轨迹和历史签到轨迹之间的空间关系。计算公式如下:

$$\varphi_m = \frac{\exp(W^Q s_n (W^K s_m + e_{m,n}^D))}{\sum_{j=1}^{n-1} \exp(W^Q s_n (W^K s_j + e_{j,n}^D))}$$
(11)

其中, $W^Q \in \mathbf{R}^{d \times d}$ 、 $W^K \in \mathbf{R}^{d \times d}$ 分别是查询和键的权重矩阵。 $\Delta_{m,n}^D$ 表示当前签到轨迹与历史签到轨迹的活动范围中心坐标之间的距离, $\Delta_{m,n}^D$ 的计

算方式如下：

$$\text{lon}_{s_m} = \frac{\text{lon}_{l_1} + \text{lon}_{l_2} + \dots + \text{lon}_{l_{|S_m|}}}{|S_m|} \quad (12)$$

$$\text{lat}_{s_m} = \frac{\text{lat}_{l_1} + \text{lat}_{l_2} + \dots + \text{lat}_{l_{|S_m|}}}{|S_m|} \quad (13)$$

$$\Delta_{m,n}^D = 2r \arcsin \left(\sqrt{\sin^2 \left(\frac{\text{lat}_{s_n} - \text{lat}_{s_m}}{2} \right) + \cos(\text{lon}_{s_n}) \sin^2 \left(\frac{\text{lon}_{s_n} - \text{lon}_{s_m}}{2} \right)} \right) \quad (14)$$

公式(12)和(13)对签到轨迹中的兴趣点经纬度坐标求平均值，得到每个签到轨迹的活动范围中心坐标。 r 是地球半径， $\Delta_{m,n}^D$ 是两个中心坐标之间的球面距离。 $\Delta_{m,n}^D$ 进行 $\Delta_{m,n}^D = \lfloor \Delta_{m,n}^D / 100 \rfloor$ 缩放和 $\Delta_{m,n}^D = \min(a, \Delta_{m,n}^D)$ 裁剪处理， a 是空间间隔取值上限的超参数， $e_{m,n}^D$ 是 $\Delta_{m,n}^D$ 对应的嵌入向量。

3.3 短期偏好建模

在公式(8)中，已经得到当前签到轨迹的隐藏状态 $\{h_1, h_2, \dots, h_{p-1}\}$ 。 h_{p-1} 编码了最近签到的序列影响，可以作为用户的短期偏好表示，但是 h_{p-1} 并没有包含当前时间的信息，也就是说忽略了时间推移的影响。以 h_{p-1} 作为偏好向量只能够预测未来不确定时间可能感兴趣的兴趣点，这无法满足实时兴趣点推荐的需求。因此，为了表示时间推移的影响，我们为每个可能的时间间隔创建时间转移矩阵 $\mathbf{W}_{\text{trans}}^b$ ， $b \in \{0, 1, \dots, 23\}$ ，时间间隔以小时为单位。短期偏好最终可表示为

$$s^b = \mathbf{W}_{\text{trans}}^{c - \text{seg}_{p-1}} \cdot h_{p-1} \quad (15)$$

3.4 预测层

预测层利用用户的长短期偏好计算用户在目标时间对兴趣点的偏好分数，计算公式如下：

$$p_{l_i} = \left(\mathbf{W}_o (s^l \oplus s^s) \right)^T e_{l_i} \quad (16)$$

其中， $\oplus$ 表示连接操作，将用户的长短期偏好向量通过连接操作组成用户偏好向量； $\mathbf{W}_o \in \mathbf{R}^{d \times 2d}$ 是用于转换维度的权重矩阵； e_{l_i} 是兴趣点 $l_i \in L$ 的嵌入表示； p_{l_i} 表示用户对兴趣点 l_i 的偏好分数。最后，可以获得按分数高低排序的top-k兴趣点推荐列表。

3.5 模型训练

本文使用二元交叉熵作为损失函数对模型进行训练，如下式：

$$\text{loss} = - \sum_{r=1}^R \log(\sigma(p_{\text{pos}})) + \log(1 - \sigma(p_{\text{neg}})) \quad (17)$$

其中， R 表示所有样本数量， $\sigma(\cdot)$ 是Sigmoid函数。 p_{pos} 表示用户实际访问的下一个兴趣点，作为正样本。 p_{neg} 是从用户从未访问过的兴趣点集合中随机采样的兴趣点，作为负样本。

4 实验数据和评价指标

4.1 数据集及预处理

本文在 Gowalla 和 Yelp 数据集上进行了模型评估。这两个数据集是典型的基于位置的社交网络数据集，其中每条记录包括用户、兴趣点及对应的经纬度坐标、访问时间戳、类别等信息。

对数据集的分析可以发现，用户的偏好存在时间敏感性。本文选取了咖啡馆和音乐演出场所这两类兴趣点为代表，统计了它们在不同时间段内被用户访问的次数，如图2、图3所示。可以发现咖啡馆类的兴趣点访问次数于午后较为密集，音乐演出场所类的兴趣点的访问多集中于深夜。

此外，我们还选取了披萨店这类兴趣点为代表，统计了在一天之中，用户签到披萨店后的1~24小时内访问次数前三的兴趣点，如图4所示。可以发现，在1~2小时内，用户明显更青睐签到酒吧、饮品店等能够解渴的兴趣点。而随着时间的推移，用户的兴趣也逐渐与披萨店相关的兴趣点无关。

本文对数据集进行了预处理：①删除了签到记录少于10条的用户以及被签到次数少于10次的兴趣点。②对用户的签到序列作分割处理，如果签到记录的时间处于同一天，则编入同一签到轨迹中。预处理后的数据集的统计数据见表1。完成预处理步骤后，将每个用户的签到轨迹按照时间排序，前80%的签到轨迹作为训练集，后20%的签到轨迹作为测试集。

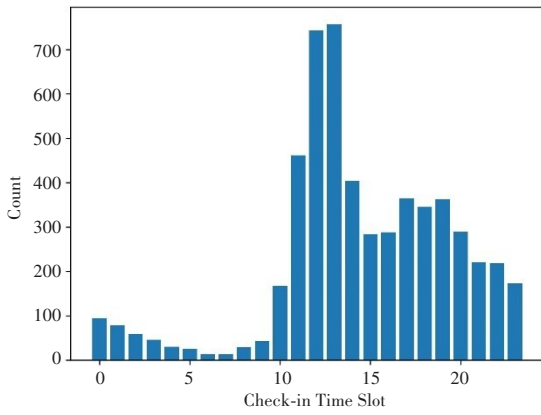


图 2 签到咖啡馆的时间段统计

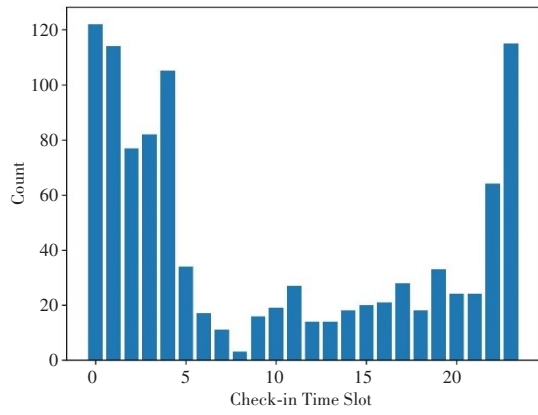


图 3 签到音乐演出场所的时间段统计

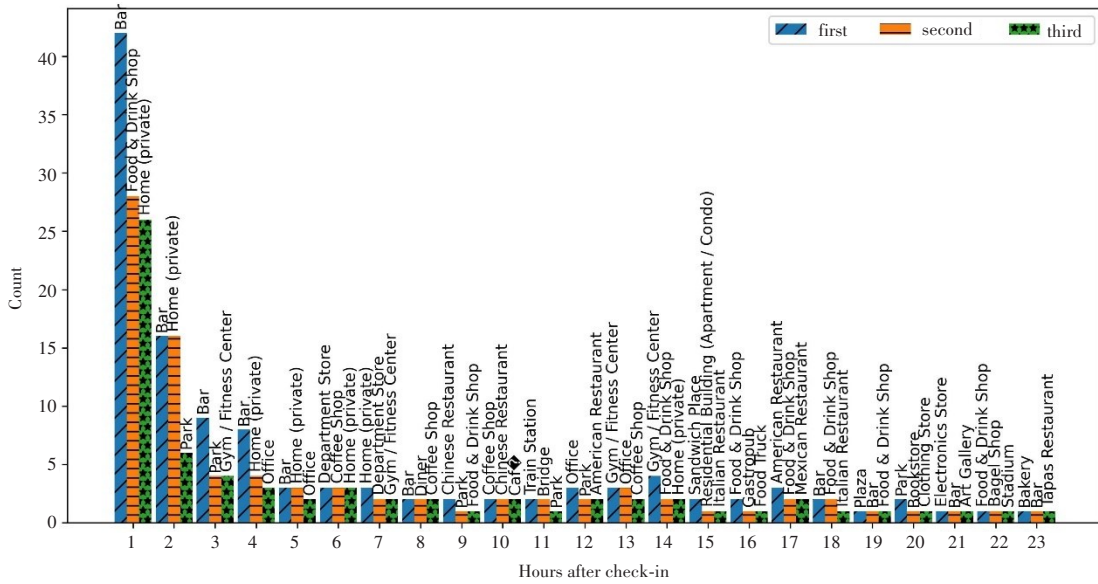


图 4 披萨店的下一个签到时间段统计

表 1 Gowalla和Yelp数据集的统计信息

统计信息	数据集	
	Gowalla	Yelp
用户数	7856	30514
兴趣点数	12543	18932
签到记录	515600	860910
用户平均签到	65.63	28.21
稀疏度	99.48%	99.85%

4.2 评价指标

本文采用的评价指标是命中率 $Hit@k$ 和归一化折损累计增益 $NDCG@k$ 。

$Hit@k$ 是一种衡量召回率的指标，计算公式

如下：

$$Hit@k = \frac{1}{R} \sum_{r=1}^R I(r) \quad (18)$$

其中， R 是样本数量， $I(r)$ 是一个指示函数，表示样本 r 是否被模型正确推荐。如果用户在样本 r 中实际签到的下一个兴趣点出现在模型的前 k 个推荐项中，则 $I(r) = 1$ ，表示命中；否则 $I(r) = 0$ ，表示未命中。 $Hit@k$ 的值越大表示召回率越高。

$NDCG@k$ 是一种衡量排序质量的指标，计算公式如下：

$$NDCG@k = \frac{1}{R} \sum_{r=1}^R \frac{DCG@k}{IDCG@k} \quad (19)$$

$$DCG@k = \sum_{i=1}^k \frac{2^{rel_i} - 1}{\log_2(i + 1)} \quad (20)$$

$$IDCG@k = \sum_{i=1}^{|REL|} \frac{2^{rel_i} - 1}{\log_2(i + 1)} \quad (21)$$

其中, rel_i 指第 i 个结果的相关性分数, 取值为 0 或 1, 如果在推荐列表中第 i 个兴趣点是用户实际签到的下一个兴趣点, 取值为 1, 否则为 0。|REL| 代表用户签到过的兴趣点个数。 $DCG@k$ 是折损累计增益, 它是让相关性越好的结果排名越靠前。 $IDCG@k$ 表示理想情况下最大的 DCG 值。 $NDCG@k$ 的值越大表示排序质量越好。

4.3 评价对比方法

本文选取了六种具有代表性的推荐方法进行对比, 其中 POP, FPMC-LR 是传统的兴趣点推荐算法, GRU4Rec、ST-RNN、Time-LSTM、LSTPM 是基于循环神经网络的推荐算法, 详细介绍如下:

(1) POP^[1]: 基于兴趣点的流行度进行推荐。

(2) FPMC-LR^[5]: 基于矩阵分解和马尔可夫链的兴趣点推荐模型, 依据矩阵分解学习所有用户的整体偏好, 通过兴趣点转移矩阵的马尔可夫链来挖掘人们在时间维度的移动模式。

(3) GRU4Rec^[6]: 利用 GRU 提取签到序列的序列特征, 以预测用户下一个行为。

(4) ST-RNN^[7]: 在 RNN 的隐藏状态计算过程中增加了时空转移矩阵, 捕获相邻签到记录之间的时空间隔对下一个 POI 推荐的影响。

(5) Time-LSTM^[8]: 在标准的 LSTM 单元中增加一个时间门, 用来调节上一个状态对当前状态的影响, 根据时间间隔的大小, 增强或减弱短期记忆。

(6) LSTPM^[9]: 基于 LSTM 分别建模用户的长短期偏好, 综合两种偏好为用户进行推荐。

5 实验结果分析

5.1 NP-ULSP 模型与其他推荐方法的性能对比

在本实验中, 我们在 Gowalla 和 Yelp 数据集上, 将 NP-ULSP 与一些主流的兴趣点推荐方法

进行性能对比, 考察了模型在推荐 5、10 个兴趣点时的性能, 结果如图 5~图 8 所示, 图 5 和图 7 是在 Gowalla 数据集上的实验结果, 图 6 和图 8 是在 Yelp 数据集上的实验结果。

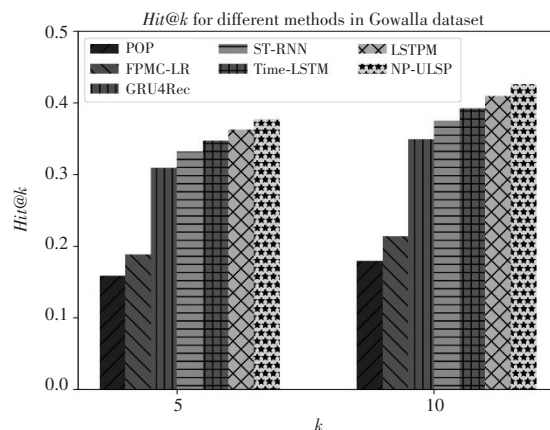


图 5 Gowalla 上的命中率

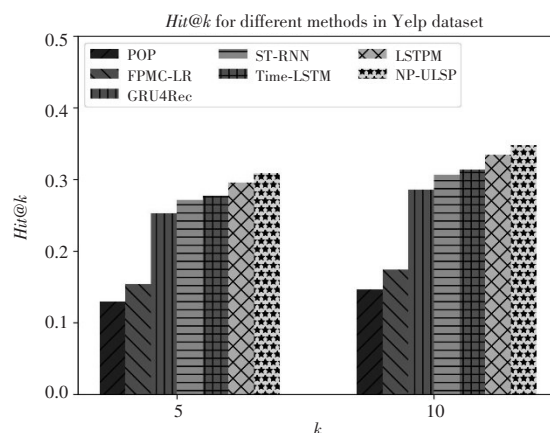


图 6 Yelp 上的命中率

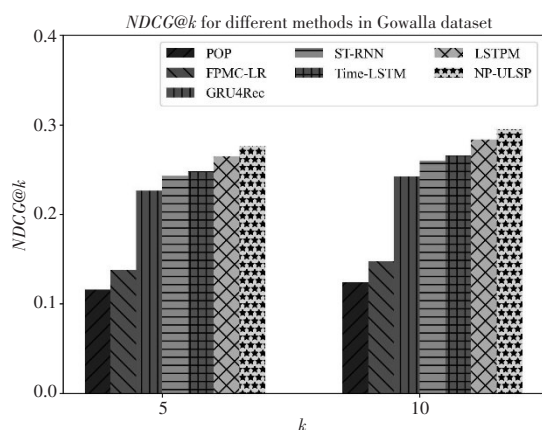


图 7 Gowalla 上的归一化折损增益

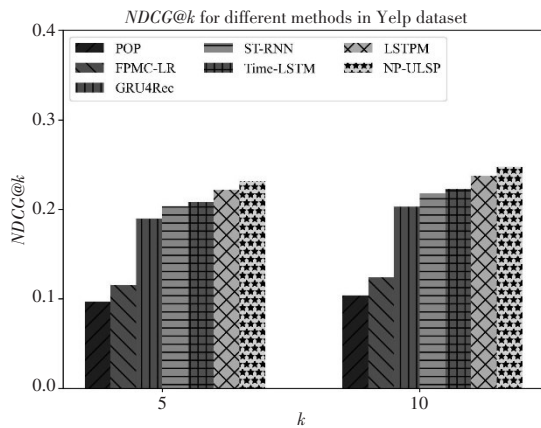


图8 Yelp上的归一化折损增益

从图中可以观察到：

(1) 本文提出的模型 NP-ULSP 在两个数据集上的命中率和归一化折损累计增益指标都要优于对比方法，这证明了模型的有效性。

(2) POP 模型在以上对比方法中表现最差，因为该方法是基于流行度的兴趣点推荐方法，它只考虑了物品的流行程度，没有考虑到用户的动态偏好。FPMC-LR 的性能优于 POP，说明下一个兴趣点推荐中兴趣点的辅助信息能够提升推荐效果。GRU4Rec、ST-RNN、Time-LSTM 的性能优于 FPMC-LR，这表明循环神经网络在提取签到序列特征，捕获用户动态偏好方面是有效的。

(3) LSTPM 模型是基于用户长短期偏好来进行兴趣点推荐的。对于长期偏好，使用一种非局部的神经操作聚合序列的隐藏状态。对于短期偏好，引入联合方式的 geo-dilated LSTM，来考虑地理因素的影响。因此，LSTPM 能够全面地考虑长短期偏好的影响，实现性能更优的

兴趣点推荐。本文提出的 NP-ULSP 同样对用户的长期偏好和短期偏好进行建模，但仍然存在较大区别。首先，LSTPM 在对短期偏好进行建模时，并没有当前时间与最近一个签到之间的时间推移影响，只是预测序列因素影响下最有可能的结果。本文的模型将时间推移的影响融入短期偏好建模中，捕获了实时偏好变化。此外，LSTPM 对长期偏好进行建模时，采用了简单的时空距离加权方式聚合签到轨迹的总体表示，本文采用了注意力机制的加权方式，能够表示更复杂的时空关系。最终，NP-ULSP 模型能够实现更好的兴趣点推荐结果。

5.2 用户长短期偏好对模型的影响

为了验证 NP-ULSP 模型中长期、短期偏好建模对性能的影响，我们设计了四个简化版模型在 Gowalla 和 Yelp 数据集上的消融实验，结果见表 2。

(1) NP-ULSP_S'：此版本不对长期偏好进行建模。

(2) NP-ULSP_S：此版本不对长期偏好进行建模，且不考虑时间推移对短期偏好的影响。

(3) NP-ULSP_L'：此版本不对短期偏好进行建模。

(4) NP-ULSP_L：此版本不对短期偏好进行建模，且使用平均池化操作替代基于时间段相似性的加权操作。

从表 2 可以观察到：

(1) NP-ULSP_L' 总是比 NP-ULSP_S' 表现更优，说明用户在目标时间段的长期行为习惯对推荐结果影响更大。

(2) NP-ULSP_L' 相比 NP-ULSP_L 进一步提

表2 消融实验

数据集	方法	Hit@5	Hit@10	NDCG@5	NDCG@10
Gowalla	NP-ULSP_S'	0.2743	0.3154	0.2194	0.2347
	NP-ULSP_S	0.2646	0.2990	0.2154	0.2305
	NP-ULSP_L'	0.3431	0.3877	0.2545	0.2683
	NP-ULSP_L	0.3260	0.3683	0.2418	0.2587
	NP-ULSP	0.3774	0.4264	0.2760	0.2953
Yelp	NP-ULSP_S'	0.2193	0.2522	0.1664	0.1780
	NP-ULSP_S	0.2117	0.2413	0.1587	0.1706
	NP-ULSP_L'	0.2745	0.3156	0.2058	0.2202
	NP-ULSP_L	0.2608	0.2973	0.1956	0.2093
	NP-ULSP	0.3084	0.3484	0.2313	0.2474

高了推荐性能,原因是基于时间段相似性的加权操作可以根据用户的当前所处时间段从用户历史数据中收集更有用的信息。

(3) NP8-ULSP_S'相比NP-ULSP_S略微提升了推荐性能,这说明考虑时间推移影响确实能够提升推荐性能。

6 结语

兴趣点推荐近年来受到了学术界的广泛关注,但是大多数研究工作忽略了有效地利用时空信息。在现实生活中,人们往往需要推荐系统根据用户目前所处的时空场景,给出一个实时的推荐结果。本文围绕这一问题,同时利用用户的长短期偏好为用户进行兴趣点推荐。对于长期偏好,从用户历史信息中筛选出与当前时空场景最相关的信息;对于短期偏好,考虑了最后一次签到至当前时间的时间推移影响。实验结果证明了本文方法的有效性。但是,本文的方法仅利用了时空信息作为辅助信息,未考虑其他模态的信息,如语义、社交关系等。如果能够结合这些信息,可能会进一步提高信息增益。

参考文献:

- [1] YE J, ZHU Z, CHENG H. What's your next move: user activity prediction in location-based social networks[C]//Proceedings of the 2013 SIAM International Conference on Data Mining, Austin, Texas, USA, 2013: 171-179.
- [2] YIN H, WANG W, WANG H, et al. Spatial-aware hierarchical collaborative deep learning for POI recommendation[J]. IEEE Transactions on Knowledge and Data Engineering, 2017, 29(11): 2537-2551.
- [3] JIANG S, QIAN X, SHEN J, et al. Author topic model-based collaborative filtering for personalized POI recommendations[J]. IEEE Transactions on Multimedia, 2015, 17(6): 907-918.
- [4] LUAN W, LIU G, JIANG C, et al. Partition-based collaborative tensor factorization for POI recommendation[J]. IEEE/CAA Journal of Automatica Sinica, 2017, 4(3): 437-446.
- [5] RENDLE S, FREUDENTHALER C, SCHMIDT-THIEME L. Factorizing personalized Markov chains for next-basket recommendation[C]//Proceedings of the 19th International Conference on World Wide Web, Raleigh, North Carolina, USA, 2010: 811-820.
- [6] HIDASI B, KARATZOGLOU A, BALTRUNAS L, et al. Session-based recommendations with recurrent neural networks[EB/OL]. arXiv:1511.06939, 2015.
- [7] LIU Q, WU S, WANG L, et al. Predicting the next location: a recurrent model with spatial and temporal contexts[C]//Proceedings of the AAAI Conference on Artificial Intelligence, Phoenix, Arizona USA, 2016.
- [8] ZHU Y, LI H, LIAO Y, et al. What to do next: modeling user behaviors by Time-LSTM[C]//Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-2017), Melbourne, Australia, 2017: 3602-3608.
- [9] SUN K, QIAN T, CHEN T, et al. Where to go next: modeling long-and short-term user preferences for point-of-interest recommendation[C]//Proceedings of the AAAI Conference on Artificial Intelligence, Hilton New York Midtown, New York, USA, 2020: 214-221.

Next point of interest recommendation based on user's long and short term preference and spatio-temporal context

Zheng Hongzhou, Zeng Guosun*

(Department of Computer Science and Technology, Tongji University, Shanghai 201804, China)

Abstract: In real life, users' preferences for points of interest are influenced by spatio-temporal context, and users hope to get recommendation results that match the current time. Therefore, a next point of interest recommendation model based on users' long-term and short-term preferences and spatio-temporal context is proposed. This model revolves around the issue of real-time point of interest recommendation, and explores users' real-time interest preferences from both long-term and short-term preferences. For long-term preferences, information most relevant to the current spatio-temporal context is collected from historical data. For short-term preferences, the influence of time progression is considered on the basis of sequence influence. The experimental results on public datasets prove the effectiveness of the method.

Keywords: point of interest recommendation; long and short term preference; spatio-temporal information; preference recommendation model

文章编号: 1007-1423(2023)24-0026-07

DOI: 10.3969/j.issn.1007-1423.2023.24.004

基于深度学习的苹果叶面病害缺素识别系统

郑淋萍, 李万益*, 李南健, 黄灿敏, 林祎琦, 陈津毅, 邝 芸, 郭小芸

(广东第二师范学院计算机学院, 广州 510303)

摘要: 针对叶面病害对苹果单产水平的提升产生负面影响的问题, 将斑点落叶病、花叶病、褐斑病、灰斑病和锈病等五种作为重点。主要研究内容及结果如下: 在苹果叶面存在差异以及图像环境不一的影响下, 对病害图像进行去噪操作后, 通过对关注主体的图像增强及 Ostu 法进行彩色图像分割, 得到病斑特征与背景灰度的强度比图像。记下分割后的二值化图像的像素灰度值为 0 的坐标, 并使原图在对应位置的像素灰度值为 0。通过对卷积神经网络的特征提取, 得出颜色、形状、纹理三种病斑特征参数。运用数学方法提取病斑的 H 方差后, 得到 H-S 直方图为区分病斑的主要参照, 采用一定的权重对特征参数建立病叶特征判别函数。根据判别函数对苹果叶片病害进行有效的统计学归类, 建立 Swim Transformer 识别模型, 将图像从测试数据集精准分类到特定的叶面病害类型。对卷积神经网络模型进行改进后, 训练获得 AlexNet-F 苹果病害识别模型。采用 Python 代码运行程序, 使用 Python 语言的 Django 框架, 将苹果叶面病害识别模型及知识图谱作为核心, 融合前端、数据库技术来开发系统。结果表明, 上述分类操作能够对重点的 5 种苹果叶面病害图高效识别, 满足项目作为识别系统的需要。

关键词: 苹果叶面病害; 二值化图像; 卷积神经网络; Swim Transformer 识别模型

0 引言

推进苹果产业高水平发展已经成为实现乡村振兴的产业基础。在进行苹果种植过程中, 各种屡见不鲜的叶面病害对苹果的产量和品质造成不同程度的影响。据联合国粮农组织的统计, 叶面病害可对苹果产业造成约为总产量 10%~15% 的损失, 这就迫切需要苹果叶面病害识别这项艰苦且繁琐的工程由科技代替^[1]。

智能化时代下, 卷积神经网络、人工智能技术和深度学习随之发展, 使其在叶面病害检测中得到广泛的应用。苹果叶面病害缺素识别与分类系统能够实时性地检测并精准识别苹果叶面病害与缺素情况, 并对此提出有针对性的防治措施, 从本质上提高苹果产量和作物品质。

如今随着智能手机的广泛普及, 图像处理技术也随之渗透到社会生活和生产的不同领域。

收稿日期: 2023-06-17 修稿日期: 2023-12-07

基金项目: 广东省大学生创新训练计划资助项目国家级大学生创新训练计划资助项目(S202314278042); 广东省教育科学规划课题(2022GXJK073、2023GXJK125); 广东省高等学校党的建设研究会 2022 年党建研究课题(2022BK024); 广州市哲学社会科学发展“十四五”规划共建课题(2023GZGJ171); 广州市基础与应用基础研究项目(202002030232); 广东第二师范学院高等教育教学改革项目(2022jxgg33); 广东第二师范学院高层次人才培养特殊支持计划项目(2022 年优秀青年教师培养对象: 李万益); 广东第二师范学院思政专项研究项目(2022SZZX13); 广东第二师范学院 2023 年校内科研项目(学生工作专项)(2023XSGZ05)

作者简介: 郑淋萍(2002—), 女, 江西瑞金人, 本科生, 研究方向为图像理解、神经网络; *通信作者: 李万益(1985—), 男, 广西南宁人, 博士, 讲师, 从事领域为图像理解、机器学习、深度学习, E-mail: luther1212@163.com; 李南健(2003—), 男, 广东雷州人, 本科生, 研究方向为自然语言处理、数据挖掘; 黄灿敏(2002—), 男, 广东潮州人, 本科生, 研究方向为图像处理、人工智能; 林祎琦(2003—), 女, 广东潮州人, 本科生, 研究方向为图像理解、神经网络; 陈津毅(2003—), 女, 广东梅州人, 本科生, 研究方向为深度学习、图像处理; 邝芸(1996—), 女, 广东韶关人, 硕士, 讲师, 从事领域为思想政治教育、党建工作; 郭小芸(1996—), 女, 广东汕头人, 硕士, 研究方向为教育测量与评价

通过实时拍摄照片上传到苹果叶面病害缺素识别系统进行智能剖析,通过辅助决策系统提供解决措施,从根本上解决农民缺乏专家的专业性指导的问题,也能通过系统对病叶实施痛点式的应对措施^[2]。

综合上述,前期构建苹果叶面病害知识图谱。通过网络爬虫、查阅专业书籍等途径获取苹果病害相关数据,进行数据处理,采用知识抽取手段构建实体、关系和属性的三元组,并导入 Neo4J 图数据库,完成知识图谱可视化^[3]。设计病害问题数据集,结合机器学习和图数据库相关技术完成苹果病害知识问答功能。

上世纪90年代,与植物病害相关的图像处理技术的研究开始兴起,并被大量引入农业工程。在图像分割方面,采取中值滤波算法进行预操作,再用K均值聚类算法分割图像,辅以SVM算法,使得准确率上升至99.33%。

在特征提取方面,HIS模型发现病叶和健康叶之间存在区别很大的色调H分量,通过色调直方图将参考到的合适阈值将病叶黄瓜和健康黄瓜进行区分,将色调的影响力推上一个新高度。

近代发展以来,遗传算法和神经网络的结合建立了模糊的神经网络图像识别系统,并优化其中的可调参数,加快网络训练度,使得传统BP神经网络算法中常出现的局部最小值现象显著减少,提高精确度,且受外界因素影响程度小。

目前为止,国内虽在计算机图像分析病叶方面进行了不少研究,但多是停留在探究算法可行性这一层面,识别率有待进一步提高。且大多以颜色、形状、纹理作为特征参数,识别的手段有待拓宽^[4]。由此本研究具有较大的实际意义。

21世纪初,利用人工神经网络分类算法,对黄瓜的白粉病和霜霉病进行识别,并开发自动化病斑检测系统。而后,以蝴蝶兰秧苗为提取对象的研究将十个参数作为输入变量,借以前馈神经网络算法将软腐病和褐斑病区分出来,且拥有97.2%的识别率。大约在同期,利用RGB、HIS和YCbCr等彩色模型对植物彩色病害图像进行色彩转换,从其中的病斑图像提取多

个颜色分量,在此基础上采用局部优化阈值法分割病斑部位,发现颜色分量信息在阈值分割中的较大影响力。

直至近代以来,边缘检测算法对大麦病斑图像提取出病斑区域,使得大麦菌斑终于完好地从背景中分离出来。

本项目所用的Otsu算法被用于计算病叶图绿色像素的屏蔽阈值,消除置零的RGB以及感染的边缘特征,K均值聚类法对病叶图像的分割十分成功。并将多种植物叶片用于研究,基于二叉树结构提出了多支持的向量机方法识别,不管是在识别准度还是允许识别复杂度方面,均远远超过BP神经网络分类器^[5]。

本研究基于深度学习苹果叶面病害缺素识别,将苹果叶面作为研究对象,针对花叶病、斑点落叶病、锈病、灰斑病和褐斑病等五种症状明显的叶面病害进行深入研究。

(1) 基于分割算法的图像处理:本项目数据集来自苹果种植园的自然条件下获取的样本,室外环境下采集的样本的RGB颜色空间图像对光照敏感度较高,选择采用颜色空间的转换^[6]。考虑到复杂背景下的复杂分割过程,选择分布来进行图像分割,即结合K均值聚类算法以及最大类间方差法^[7]。

(2) 提取苹果叶面病害特征参数:预处理操作后,会初步得到基本完整的病斑,对病叶颜色、纹理、形状等特征进行处理,根据所提出的区分性、可靠性和复杂度的三原则,筛选出最佳特征参数,作为识别分类系统的输入参数对机器样例进行训练^[8]。

(3) 根据判别函数对苹果叶片病害进行有效的统计学归类:通过对卷积神经网络的特征提取,并采用一定的权重对特征参数建立病叶特征判别函数,辅以池化层降低计算复杂度^[9]。而后建立Swim Transformer识别模型,将图像从测试数据集精准分类到特定的叶面病害类型。

(4) 苹果叶面病害缺素识别系统设计:綜上研究内容,基于使用Python语言的Django框架,结合前端、数据库技术等相关技术完成系统的开发,辅助苹果种植农人高效识别苹果叶面病害,并给出合理有效的施药防治措施方案。

1 相关理论技术

1.1 相关提取卷积神经网络

卷积神经网络主要是运用其神经元进行局部连接和权值共享,通过这些特点来减少参数,与此同时,池化层也利用到图像的局部相关性,再对图像进行子抽样,对数据进行减少处理的同时也保留了有用的信息,这也是现在图像识别最重要的一种^[10]。

卷积层是卷积神经网络中最重要的结构,其作用主要是对输入的数据进行特征提取。BP神经网络的过程可以分为两个阶段:第一阶段是信号的前向传播,让信号从输入层经过隐含层,最后再到达输出层;第二阶段是误差的反向传播,让信号从输出层到隐含层,最后再到达输入层,再依次调节隐含层到输出层、输入层到隐含层的权重和偏置。神经网络是利用当前拥有的数据找出输入和输出之间的权值关系,然后再利用这样的权值关系来进行仿真。

1.1.1 池化层

池化层主要是在卷积层之后的,其主要作用就是降采样,在其保留重要特征的情况下对特征进行降维,再通过压缩然后输入特征图,来降低下一层需要处理的数据量,对网络的计算复杂度进行简化。池化的操作主要可以分为最大池化(Max Pooling)和平均池化(Avg Pooling),其中最大池化使用场景相对较多。最大池化是将输入进来的图像划分为多个矩形区域,然后再选取每一个区域中的最大值来进行输出。

1.1.2 激活函数

在现实应用中,数据大多都是非线性发布的,但是神经网络中是线性的计算,所以为了使神经网络能够较好地解决非线性问题,我们在神经网络中使用了激活函数。激活函数的作用是将非线性因素加到神经网络中,从而使神经网络的表达能力得到增强,进而能够让神经网络解决更多、更复杂的问题。在人工神经网络中,每一个神经元的节点都对应相对的输入和输出。

图1为神经网络中一个神经元数据的传输过程,神经元的节点将接受上一层神经元的输出作为自己节点的输入,然后激活函数再对其进

行处理,随后将输出值传递到下一层。不同的场景需要应用到的激活函数也不用,他们都有不用的特点。一般常用的激活函数有Sigmoid函数、ReLU函数等。Sigmoid函数需要将输出限制在(0, 1)的区间,假如是一个非常大的负数,则输出就是0;如果是一个非常大的正数,则输出就是1,这一激活函数使数据在传递的过程中不容易发散,适合用来进行概率值处理。ReLU函数是当输入的值小于0时,输出就是0;反之输入值大于0时,输出等于输入,其本质也就是一个取最大值的函数,非常适合用到深层网络的训练中。在不同的场景中选择相对合适的激活函数,可以在训练神经网络时加快模型的收敛速度,更有效地改善模型的学习能力。

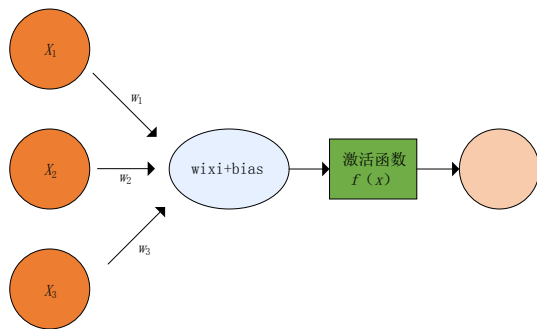


图1 神经元数据传输过程

1.2 集成创新

本系统融合了计算机视觉模型和Swin Transformer等模型和算法,在结合了苹果的各项特性信息后,设计出来的苹果页面病害缺素图像识别系统。

1.3 Swin transformer 模型

Swin Transformer是一个全新的Transformer模型,主要使用了CNN结构设计的一些理念(降采样、局部dependency等)来重新设计Transformer。

1.4 OpenCV 开源计算机视觉

图像处理主要运用了计算机技术对图像进行分析,从而达到预期的效果,这一过程包括了图像压缩、增强和复原、匹配描述和识别这三个方面。主要通过系统对病害图像的数据进行数据增强来使数据得到使用。这一过程主要

采取的扩展手段包括翻转变换、尺度变换、剪切变换、噪声扰动。

2 基于改进 AlexNet 的苹果病害识别模型构建

通过将收集来的苹果叶部病害图像和网上开源病害图像数据集进行整理、数据扩充,获得苹果叶部病害图像数据集,并通过对卷积神经网络模型 AlexNet 进行改进,训练获得 AlexNet-F 苹果病害识别模型,如图 2 所示,改进后的 AlexNet 模型包含 5 个卷积层、3 个池化层及最后一层全局平均池化层。

模型 AlexNet-F 使用全局平均池化层代替全连接层,减少网络中大量参数,避免过拟合;使用批归一化算法提高网络的泛化能力。并改变激活函数,使用 Swish 激活函数提高模型性能。改进后的 AlexNet-F 模型在验证集上准确率为 97.76%,与原始 AlexNet 模型相比识别精度有较大的提升。

在构建模型时,使用 keras 深度学习框架,以 Tensorflow 为后端进行模型训练,并使用 model.save 命令将训练模型的结构和权重都进行保存,保存为 AlexNet-F.hdf5 文件,之后可直接加载此文件进行病害识别。图 3 为 AlexNet-F 在扩充后数据集上识别测试集准确率与损失曲线图。

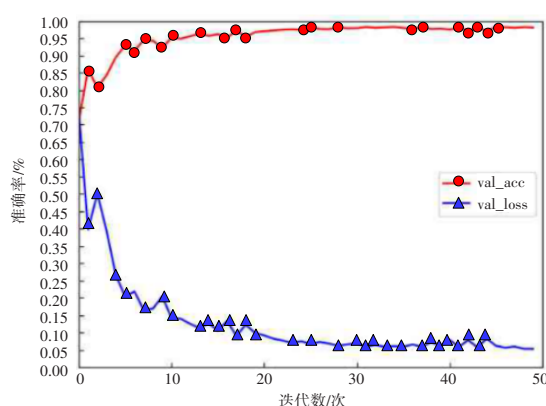


图 3 AlexNet-F 验证集准确率与损失曲线变化

表 1 为 AlexNet-F 在苹果病害原数据集与扩充后数据集上对不同病害的识别率。

3 苹果病害识别功能设计与实现

本节设计苹果叶片病害的识别模型,确定特征辨别函数,以 Matlab 为开发工具,根据各个模块功能要求编写程序,实现苹果叶片病害图像识别系统相关模块的基本功能。

3.1 图像识别基本流程

本研究中苹果病害叶片图像的识别系统包含四个主要部分,第一部分是取样并读取图像,第二部分是对图像进行预处理,第三部分是图

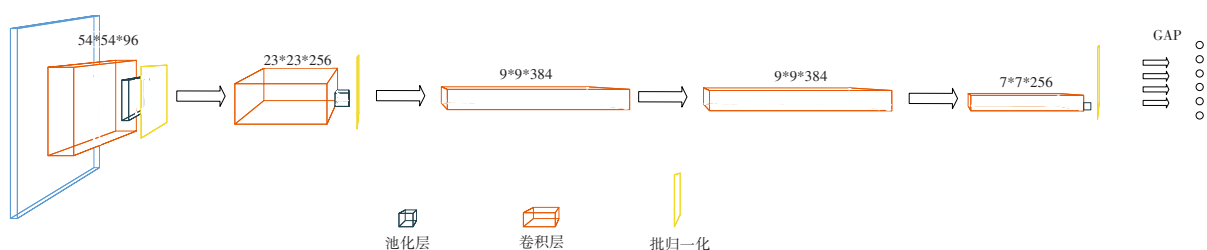


图 2 改进后的 AlexNet 模型结构示意图

表 1 苹果各类病害在 AlexNet-F 上的识别率(%)

病害种类 Disease Type	炭疽叶枯病 Anthracnose leaf blight	花叶病 Mosaic disease	锈病 Rust	黑星病 Scab	斑点落叶病 Spotted leaf disease	健康叶片 Healthy leaves
原始数据集 Original Data Set	85.86	85.86	70.18	88.51	82.09	95.18
扩充后数据集 Expanded Data Set	96.98	99.74	97.22	94.67	96.52	99.49

像信息的加工与特征提取，第四部分是判断与识别^[11]。识别流程如图 4 所示。

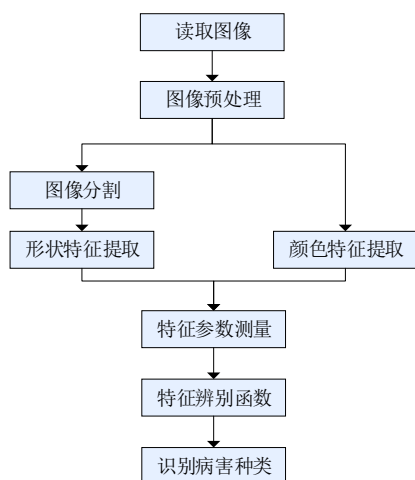


图 4 苹果叶面病害识别系统研究方案流程

3.2 图像识别基本过程

在本识别系统中，苹果叶片病害检测方法部分主要有四步：第一步是图像增强；第二步是图像分割，通过对比研究，采用彩色图像分割方法来分割图像；第三步是特征提取，对分割后的病害部位图像提取颜色、纹理和病斑形状三方面的特征参数；第四步是模式识别系统算法的设计与实现^[12]。

3.2.1 分析苹果病害、缺素叶片病斑的特征

针对其预设特征提取的方案，对图像预处理算法进行选择与优化，针对每一步处理过程优化选择最佳算法，最终得到病斑与背景灰度对比强烈的图像。从颜色、形状、纹理三种病斑特征进行图像特征提取算法的研究，最后分析这些特征参数。

3.2.2 苹果叶片病害识别模型建立

使用 Matlab 软件，通过对大量的苹果锈病、花叶病和斑点落叶病等病害图像进行分析和处理，提取出不同病害的相关特征参数，对原始图像采用本方法实现叶片提取，效果如图 5 所示。并计算这些特征参数的均值和标准差，然后采用一定的权重对特征参数建立特征判别函数，从而对苹果叶片病害进行有效的统计学归类^[13]。

3.2.3 苹果叶片病害的识别

采集待识别的苹果叶片图像，运用 Matlab 软件对该图像进行分析及处理，得到相应的特征参数，根据已建立起来的特征判别函数，可以有效识别出该苹果叶片的病害种类^[14]。

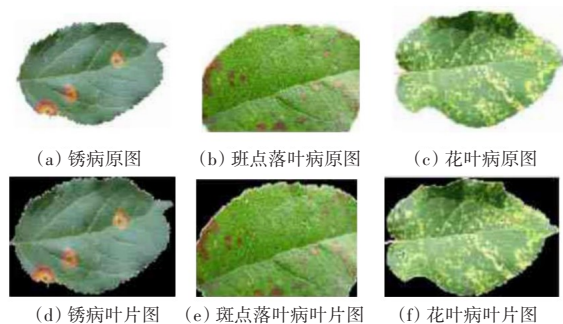


图 5 叶片提取效果图

4 苹果病害施药辅助决策系统设计与实现

4.1 系统概述

苹果病害缺素施药知识问答辅助决策系统，运用集成机器学习、人工智能领域专业技术、深度学习、网络爬虫、图像处理、数据处理、数据库技术、Web 开发等多方面高新技术集成苹果病害相关领域专业知识。系统通过神经网络构建苹果叶面病害缺素识别模型，精准、快速识别苹果叶部病害以及缺素情况，同时构建苹果病害缺素知识图谱，并以此为基础建立苹果病害缺素知识问答系统。

在施药辅助决策方面，系统将图像识别与知识问答功能相结合，以专业化、智能化手段帮助苹果种植人员快速判断当前苹果叶面病害情况，降低种植人员的学习成本，为种植人员制定及时有效的解决方案并提供辅助决策，在一定程度上能够提高苹果种植人员的施药准确率，减少农药不必要的浪费，降低病害对果实造成的危害，提升苹果的产量和品质^[15]。

4.2 知识问答功能流程设计

问答功能以构建的苹果病害知识图谱为基础，结合图数据库、机器学习技术完成苹果病害知识问答功能设计。知识问答功能流程设计如图 6 所示，大体流程如下：首先整理常出现的

苹果病害缺素相关问题，制作成问题数据集，使用机器学习技术训练得到问题分类器。用户提问时，输入的问题为自然语言，系统对问题语句进行处理，选取其中的关键信息，结合之前得到的问题分类器弄清楚用户所提问的问题的目的，然后调用相对应的语句，在数据库中进行查询，获取问题答案，将结果反馈给用户。

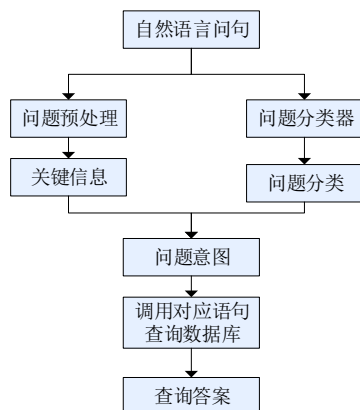


图6 病害缺素知识问答功能流程

4.3 病害知识问答功能实现

实现病害知识问答功能需要以下步骤(如图7所示)。

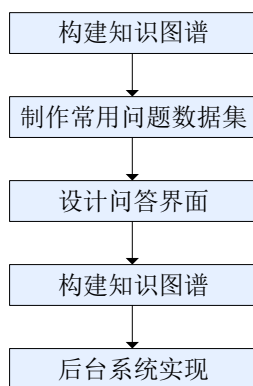


图7 问答功能步骤

(1) 构建苹果病害知识图谱，并将其存储到图数据库中。

(2) 制作常用问题数据集，并将其用于训练问题分类器。

(3) 设计问答界面，包括问题搜索框和结果展示框。

(4) 在后台系统中实现问题分类和答案查询

功能。

当用户输入问题后，系统会先使用训练好的问题分类器对问题进行分类，然后根据分类结果，在图数据库中执行相应的查询语句，最后将查询结果返回给用户。

5 总结与展望

5.1 总结

本研究基于深度学习技术对苹果叶面病害进行了缺素识别，主要针对花叶病、斑点落叶病、锈病、灰斑病和褐斑病等五种病害进行了重点研究。通过对病斑特征进行颜色、形状、纹理三个方面的参数提取，建立了病叶特征判别函数，辅以池化层降低计算复杂度，最终建立了 Swim Transformer 识别模型和 AlexNet-F 苹果病害识别模型。在此基础上，开发了施药辅助决策系统，有效地帮助种植人员识别叶面病害并提供合理的防治措施。

5.2 展望

未来可以进一步拓展研究范围，将模型应用于更多的果树叶面病害识别中，提高果树病害防治水平，同时也可以探索更加高效、准确的特征提取和模型建立方法，提高识别准确率和效率。另外，也可以加强与实际种植业生产实践的结合，根据实际情况进行优化和改进，让识别系统更加贴近实际需求，为果树种植业的健康发展做出更大贡献。

参考文献：

- [1] 陈聪,于啸,宫琪.基于改进残差网络的苹果叶片病害识别研究[J].河南农业科学,2023,52(4):152-161.
- [2] 姜红花,杨祥海,丁睿柔,等.基于改进ResNet18的苹果叶部病害多分类算法研究[J].农业机械学报,2023,54(4):295-303.
- [3] 郭董菊.基于图像处理的苹果表面及叶面病害检测技术研究[D].杭州:浙江科技学院,2022.
- [4] 郑新桥.图像处理在农业领域的应用研究:以LeNet-5模型的苹果叶部病害研究为例[J].南方农机,2022,53(15):36-39.
- [5] 孙丰刚,王云露,兰鹏,等.基于改进YOLOv5s和迁移学习的苹果果实病害识别方法[J].农业工程学报,2022,38(11):171-179.

- [6] CHEN R, XUE X M, WANG G P, et al. Determination and dietary intake risk assessment of 14 pesticide residues in apples of China[J]. Food Chemistry, 2021, 351: 129266.
- [7] LI C J, YUAN S F, XIE Y F, et al. Transformation of fluopyram during enzymatic hydrolysis of apple and its effect on polygalacturonase and apple juice yield[J]. Food Chemistry, 2021, 357: 129842.
- [8] LI X P, LI S Q. Transformer help CNN see better: a lightweight hybrid apple disease identification model based on transformers[J]. Agriculture, 2022, 12(6): 12060884.
- [9] 丁睿柔. 基于卷积神经网络的苹果叶部病害识别与分割研究[D]. 山东: 山东农业大学, 2022.
- [10] 刘斌, 张嘉琦, 田靓靓, 等. 一种基于卷积神经网络的苹果叶部病害诊断方法[D]. 咸阳: 西北农林科技大学, 2022.
- [11] 王云露, 吴杰芳, 兰鹏, 等. 基于改进 Faster R-CNN 的苹果叶部病害识别方法[J]. 林业工程学报, 2022(1): 153-159.
- [12] 晁晓菲. 基于深度学习的常见苹果叶片病害识别与病斑分割方法研究[D]. 咸阳: 西北农林科技大学, 2021.
- [13] 罗元秋. 基于深度学习的复杂背景下苹果叶面病害细粒度图像识别算法研究[D]. 江苏大学, 2023.
- [14] LI P, JING R Z, SHI X L. Apple disease recognition based on convolutional neural networks with modified softmax[J]. Frontiers in Plant Science, 2022, 13: 820146.
- [15] 张华鹏. 基于深度学习的苹果病害识别与施药辅助决策系统研发[D]. 泰安: 山东农业大学, 2021.

Apple leaf disease deficiency based on deep learning

Zheng Linping, Li Wanyi*, Li Nanjian, Huang Canmin, Lin Yiqi, Chen Jinyi,
Kuang Yun, Guo Xiaoyun

(School of Computer Science, Guangdong University of Education, Guangzhou 510303, China)

Abstract: In response to the negative impact of leaf diseases on the improvement of apple yield, five key diseases were identified, including leaf spot disease, mosaic disease, brown spot disease, gray spot disease, and rust disease. The main research contents and results are as follows: Under the influence of the difference of apple leaf surface and the different image environment, after the denoising operation of the disease image, the intensity ratio image of the disease spot feature and the background gray level is obtained by image enhancement and Ostu method of color image segmentation. Write down the coordinates where the pixel gray value of the binary image after segmentation is 0, and make the pixel gray value of the original image at the corresponding position be 0. Through the feature extraction of convolutional neural network, three characteristic parameters of color, shape and texture were obtained. The H-S histogram was obtained as the main reference for distinguishing diseased spots, and the characteristic discriminant function of diseased leaves was established with certain weights for the characteristic parameters. According to the discriminant function, the apple leaf diseases are effectively classified statistically, and a Swim Transformer recognition model is established to accurately classify images from the test data set to specific leaf disease types. After improving the convolutional neural network model, AlexNet-F apple disease recognition model was obtained. Use Python code to run the program, use the Django framework of Python language, take the identification model of apple leaf disease and Knowledge graph as the core, and integrate front-end and database technology to develop the system. The results show that the above classification operation can effectively identify the key 5 kinds of apple leaf surface disease maps, and meet the needs of the project as an identification system.

Keywords: leaf disease of apple; binary image; convolutional neural network; Swim Transformer identification model

文章编号: 1007-1423(2023)24-0033-07

DOI: 10.3969/j.issn.1007-1423.2023.24.005

基于改进 BiseNet 的轻量化上颌前牙分割算法

吴 杰, 陶青川*

(四川大学电子信息学院, 成都 610000)

摘要: 针对移动终端等低算力设备, 提出了一种应用于自然彩色图像的轻量化上颌前牙分割算法, 选用了在自建数据集上表现较好的 BiseNet-ResNet18 作为基准模型, 然后对模型进行了改进。实验结果显示, 改进后的 BiseNet 在自建数据集上的 MIoU 达到了 90.7%, 较 BiseNet 提升了 1.2 个百分点。在轻量化和实时性方面, 模型大小只有 BiseNet 的 1/8, 并且每秒帧数达到了 63.1。该项研究成果有望在牙科医学和医疗影像学领域得到应用, 推动牙齿分割、微笑美学自动分析等技术的发展。

关键词: 牙齿分割; 语义分割; BiseNet

0 引言

随着深度学习技术的快速发展, 越来越多的人工智能算法被应用于口腔医学图像领域, 实现了智能化分析处理口腔医学图像。例如, Estai 等^[1]结合 U-Net 和 Faster R-CNN 网络, 提出了一种应用于牙齿全景 X 线片上的牙齿自动检测分类系统。Im 等^[2]提出一种基于动态图卷积神经网络(DGCNN)的算法, 对数字牙齿模型进行牙齿分割, 显示出了较高的分割准确性和效率, 提出了深度学习可以用于正畸诊断和矫治器制造。Zheng 等^[3]提出了一种新的基于图神经网络的三维牙齿分割方法 TeethGNN, 解决了以前方法中普遍存在的牙齿拥挤问题和不完全分割问题。然而, 现有牙齿分割相关研究大多基于良好的临床环境和专业设备。随着移动设备的发展, 我们需要一种能实时监测健康状况的设备, 而不是花费大量时间预约仪器。自然彩色微笑图像是口腔正畸诊断中的重要图像, 其相较于 X 线片、锥形束 CT(CBCT), 具有获取方法简便、视觉效果好以及患者接受度高等优点。然而, 目前对自然彩色微笑图像上颌前牙分割

的研究较少。在自然彩色微笑图像中, 上颌前牙区牙齿显露量较少, 对整个面部的协调美观性起着至关重要的作用, 也是微笑美学研究中的重点内容^[4-5]。因此, 本文提出了一种能够在移动设备上快速运行的轻量化上颌前牙分割算法。

这项工作的意义在于:

- (1) 将轻量化语义分割算法引入到微笑图像前牙分割当中, 并取得了较好的效果;
- (2) 提出了基于深度学习的自然彩色图像上颌前牙分割 workflow;
- (3) 改进了 BiseNet, 使其在低算力设备上也能较好地进行牙齿分割。

1 资料与方法

1.1 数据集

在本文中, 研究微笑美学的图像是恒牙期正面微笑图像, 且微笑为姿势位微笑。Rigsbee 等^[6]提出姿势位微笑的概念, 指在社交场合或被人要求微笑时产生的微笑, 其重复性高、维持性好且与情绪无关, 大多数微笑美学的研究都建立在此

收稿日期: 2023-10-07 修稿日期: 2023-11-15

作者简介: 吴杰(1998—), 男, 四川资阳人, 硕士研究生, 研究方向为智能信息系统; *通信作者: 陶青川(1972—), 男, 四川南充人, 博士, 副教授, 研究方向为模式识别与智能系统, E-mail: taoqingchuan@scu.edu.cn

之上^[7]。微笑时下颌前牙牙齿一般不显露或较少显露^[8]，并且美学区一般指上颌前牙区^[9]，在本文研究中需要被分割的牙齿为上颌前牙，上颌前牙采用国际牙科联合会 FDI 牙位表示法进行标记，便于后续的牙齿标注。



图1 牙位标记示意图

由于暂无公开的自然彩色微笑图像牙齿分割数据集，本次实验使用自制数据集，数据集总张数为 3945 张。数据集使用标注工具

LabelMe 进行标注，由经过培训的人员完成。标注示例如图 2 所示。数据集按照 8 : 2 的比例划分为训练集和测试集。

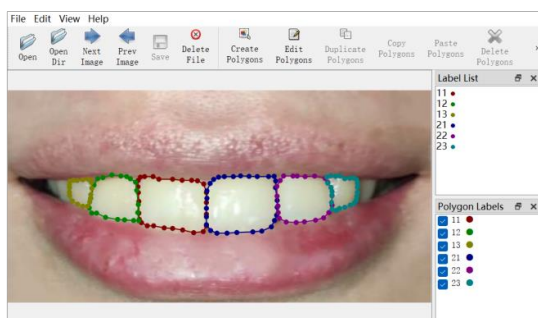


图2 语义分割标注示例图

1.2 研究方法

1.2.1 算法流程

由于口腔部位在图像中所占像素较少，为了更精确地分割牙齿，我们采用了分级的方式，即首先通过人脸识别技术对人脸进行识别，识别到人脸之后应用人脸关键点检测算法和人脸对齐算法，得到口腔部位的关键点，随后裁切口腔部位图像送入语义分割网络。算法的整体流程如图 3 所示。

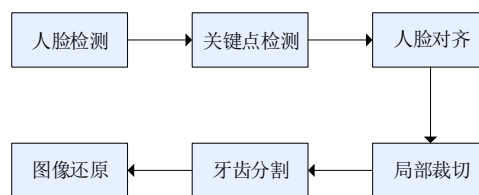


图3 算法整体流程图

1.2.2 BiSeNet 网络

BiSeNet^[10]网络不同于 Unet 的编码器解码器结构，使用双分支结构进行空间信息编码和语义信息提取。其模型体积较小、运行速度较快，在工业界得到了广泛的应用。在空间路径模块 (spatial path)，通过 3 次下采样输出特征图为原来的 1/8，通过保留较大的特征图，可以保证编码较丰富的空间信息；在上下文路径模块 (context path)，使用注意力细化模块 (attention refinement module, ARM) 优化该阶段的特征，然后通过特征融合模块 (feature fusion module, FFM) 对语义信息和空间信息融合，最后经过上采样完成语义分割。

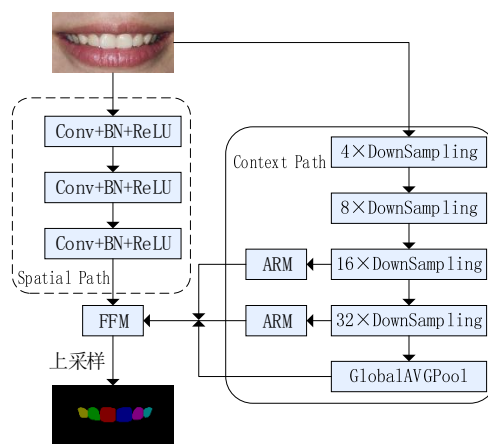


图4 BiSeNet 网络结构图

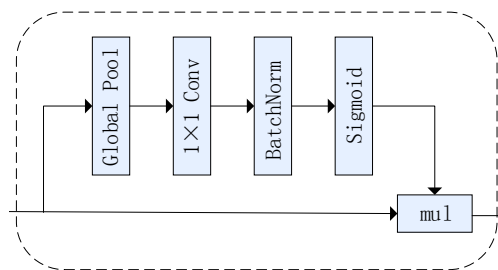


图5 ARM 网络结构图

坐标信息生成两部分组成,其结构如图9所示。

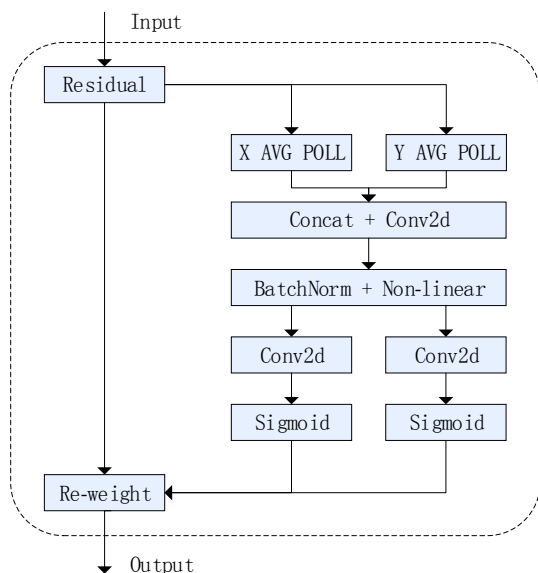


图9 CA注意力机制网络结构图

在引入 MobileNetV3 轻量化主干网络之后, MIOU 存在一定的下降, 考虑尝试在空间路径中引入注意力机制, 经过实验验证, 在空间路径第二层卷积之后嵌入坐标注意力模块, 能够提升牙齿分割模型的准确性。调整后的特征图不仅在通道层面得到了提升, 同时通过引入坐标信息, 使得牙齿的空间上下文信息挖掘更加充分。改进后的空间路径模块如图10所示。

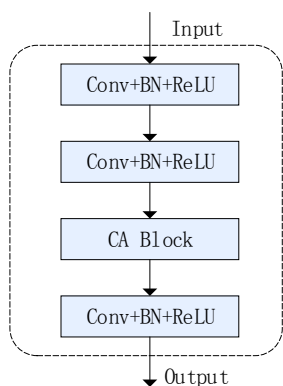


图10 改进后的空间路径模块网络结构图

2 实验与分析

2.1 实验设置

本文实验机器配置为 64 位 Ubuntu22.04,

CPU 型号为 Inter Core i7, 显卡型号为 NVIDIA GeForce GTX 3090 Ti 24 G, 深度学习框架使用的是 PyTorch 1.11.0 版本, CUDA 版本为 11.3, 优化器选择 SGD, 训练超参数设置见表2。

表2 训练超参数设置

参数	值
epochs	100
batch-size	16
lr	1e-2
momentum	0.9
weight-decay	1e-4
warmup-factor	1.0 / 3
warmup-method	linear

2.2 评价指标

为了客观准确地评估分割模型性能, 使用平均交并比(MIoU)、浮点运算数(Flops)、模型参数量(Params)、每秒帧数(FPS)和模型大小(Size)对分割模型进行评价。为了满足后续移动设备部署, 评价设备选择 CPU。

$$MIoU = \frac{1}{n+1} \sum_{i=0}^n \frac{P_{ii}}{\sum_{j=0}^n P_{ij} + \sum_{j=0}^n P_{ji} - P_{ii}}$$

其中, P_{ii} 表示实际为 i 类且预测为 i 类的像素数量, P_{ij} 表示属于 i 类却预测为 j 类的像素数量, P_{ji} 表示属于 j 类却预测为 i 类的像素数量。

2.3 结果与分析

2.3.1 不同改进方法的模型性能评估

表3 不同改进方法模型性能评估

实验	MIoU/%	Flops/G	Params/M	FPS/f·s ⁻¹	Size/MB
BiseNet	89.5	3.35	12.72	43.5	51.0
BiseNet+v3	88.9	0.52	1.45	63.9	6.1
BiseNet+ca	91.1	3.35	12.72	42.2	51.0
BiseNet+v3+ca	90.7	0.52	1.45	63.1	6.1

表中: 1 M = 1×10^6 , 1 G = 1×10^9 , v3代表 MobileNetV3

在单独引入 MobileNetV3 后, BiseNet-MobileNetV3 的 MIoU 达到了 88.9%, 相比于 BiseNet, 模型参数量下降到 1.45 M, FPS 提升到了 63.9, 说明 MobileNetV3 的 Benck 的线性瓶颈倒残差结构、深度可分离卷积和 SE 注意力机制

的有效性得到了验证。在单独引入CA注意力机制后，相比于 BiSeNet，MIoU 提升了1.6个百分点，说明CA注意力机制在空间路径模块中起到了特征提取增强的作用。在同时引入 MobileNetV3 和 CA 注意力机制后，相比于 BiSeNet，MIoU 提升了1.2个百分点，较好地平衡了分割的准确性和速度。

2.3.2 不同语义分割模型性能评估

本文选取了 DeepLabV3^[17]、DUnet^[18]、FCN-32S^[19]、BiSeNet-ResNet18 语义分割模型进行性能对比，所有模型按照相同的实验条件进行训练，并使用相同的测试集进行测试，实验结果见表4。

表4 牙齿分割模型性能评估

实验	MIoU/%	Flops/G	Params/M	FPS/ $f \cdot s^{-1}$	Size/MB
DeepLabV3	85.2	48.17	39.76	4.4	167.6
DUnet	87.9	36.78	29.35	5.9	125.9
FCN-32S	76.1	22.59	15.31	10.2	61.2
BiSeNet	89.5	3.35	12.72	43.5	51.0
本文模型	90.7	0.52	1.45	63.1	6.1

表中：1 M = 1×10^6 ，1 G = 1×10^9

由表4所知，在精度方面，本文模型的MIoU 为 90.7%，相比于 DeepLabV3、DUnet、

FCN-32S、BiSeNet 分别提升了 5.5、2.8、14.6 和 1.2 个百分点。在轻量化方面，通过引入 MobileNetV3 作为上下文路径主干并缩减通道数量，减少了计算开销和模型体积，本文模型浮点运算数 Flops 相比于 DeepLabV3、DUnet、FCN-32S、BiSeNet 分别减少了 47.65 G、36.26 G、22.07 G、2.83 G；模型参数量 Params 分别减少了 38.31 M、27.90 M、13.86 M、11.27 M；模型大小只有 6.1 MB，约为 DeepLabV3 的 1/27，BiSeNet 的 1/8。在实时性方面，本文模型每秒帧数 FPS 分别是 DeepLabV3、DUnet、FCN-32S、BiSeNet 的 14.3 倍、10.7 倍、6.2 倍、1.5 倍，使其能够以较高帧率运行，适用于基于视频分析的场景。综上所述，本文模型兼顾了精度和实时性，在低算力设备上有着良好的性能。

为了比较不同模型牙齿分割的效果，选取了效果较好的 DUnet 和效果较差的 FCN-32S 进行分割效果对比，不同模型的分割结果如图 11 所示。

可以看到，FCN-32S 模型没能准确地将每颗牙齿分割出来。DUnet 虽然能够较为准确地分割出每颗牙齿，但是牙齿边缘分割的不够细腻，部分牙齿分割不完全。本文模型能够较好地将每颗牙齿分割出来，预测图像最接近标签图像。

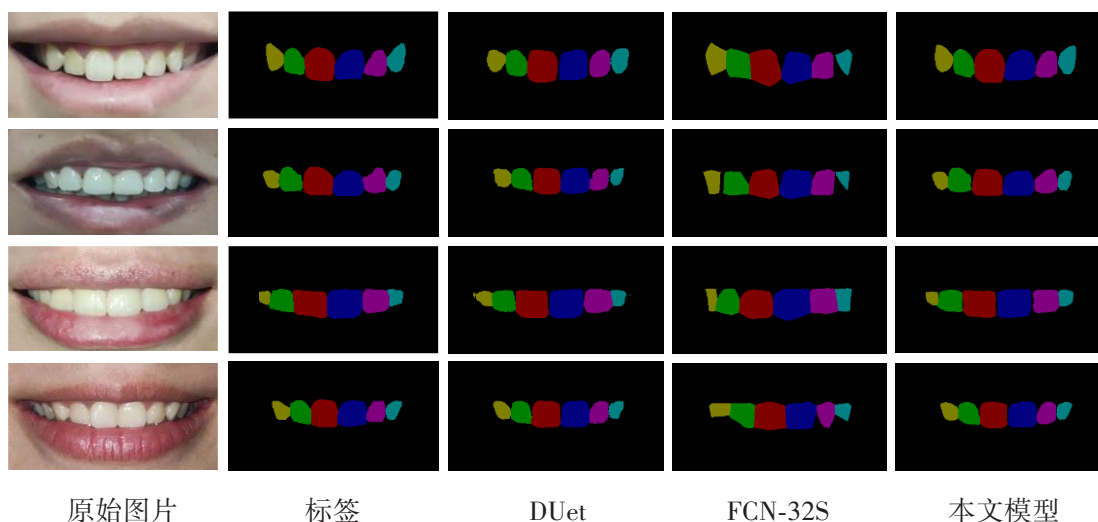


图 11 不同网络模型的分割结果

3 结语

本文提出了一种基于改进 BiSeNet 的轻量化上颌前牙分割算法, 使用 MobileNetV3 替换原 ResNet18 网络, 并在下采样过程中缩减通道数提升检测速度, 在空间路径模块中嵌入 CA 注意力机制, 提升模型精度。本文分割模型相比于其他模型在精度上有一定程度的提高, 并且推理速度较快, 模型大小、参数量、浮点运算数都有一定程度的降低, 每秒帧数 FPS 达到了 63.1。本文结果表明, 语义分割适用于自然彩色微笑图像的牙齿分割, 所提出的牙齿分割方法对牙科医学和医疗影像学的发展具有一定的促进作用。在之后的工作中, 可以结合不同移动设备的硬件特性, 对网络结构进行分析和优化, 进一步提高推理速度和精度。

参考文献:

- [1] ESTAI M, TENNANT M, GEBAUER D, et al. Deep learning for automated detection and numbering of permanent teeth on panoramic images[J]. Dentomaxillofacial Radiology, 2022, 51(2): 20210296.
- [2] IM J, KIM J Y, YU H S, et al. Accuracy and efficiency of automatic tooth segmentation in digital dental models using deep learning[J]. Scientific reports, 2022, 12(1): 9429.
- [3] ZHENG Y, CHEN B, SHEN Y, et al. TeethGNN: semantic 3D teeth segmentation with graph neural networks[J]. IEEE Transactions on Visualization and Computer Graphics, 2022, 29(7): 3158-3168.
- [4] 项闫颜, 宋东升, 王绍泰, 等. 口腔治疗中前牙区美学的考量因素[J]. 口腔医学杂志, 2022, 42(9): 838-841.
- [5] 闻静, 许妍. 前牙美学和微笑美学相关因素的研究进展[C]//中华口腔医学会口腔正畸专业委员会. 2017年国际正畸大会暨第十六次全国口腔正畸学术会议论文汇编, 上海, 2017: 885-886.
- [6] RIGSBEE O H, SPERRY T P, BEGOLE E A. The influence of facial animation on smile characteristics[J]. The International Journal of Adult Orthodontics and Orthognathic Surgery, 1988, 3(4): 233-239.
- [7] 谭学莲, 甯佳丽, 陈河林, 等. 拔牙和非拔牙矫治对正面微笑和静态侧貌的美学影响[J]. 国际口腔医学杂志, 2015, 42(2): 147-151.
- [8] 凌均荣, 傅海君. 前牙美学不同观点及研究现状[J]. 中华临床医师杂志(电子版), 2009, 3(9): 1449-1456.
- [9] 张馨月, 王柏翔, 王慧明. 数字化技术在美学区种植研究进展及应用前景[J]. 中国口腔种植学杂志, 2021, 26(3): 206-210.
- [10] YU C, WANG J, PENG C, et al. BiSeNet: bilateral segmentation network for real-time semantic segmentation[C]//Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 2018: 325-341.
- [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, Las Vegas, USA, 2016: 770-778.
- [12] HOWARD A, SANDLER M, CHU G, et al. Searching for mobilenetv3[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Korea(South), 2019: 1314-1324.
- [13] HOWARD A G, ZHU M, CHEN B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications [EB/OL]. arXiv: 1704.04861, 2017.
- [14] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: inverted residuals and linear bottlenecks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018: 4510-4520.
- [15] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018: 7132-7141.
- [16] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, virtual, 2021: 13713-13722.
- [17] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation[EB/OL]. arXiv: 1706.05587, 2017.
- [18] JIN Q, MENG Z, PHAM T D, et al. DUNet: a deformable network for retinal vessel segmentation[J]. Knowledge-Based Systems, 2019, 178: 149-162.
- [19] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, USA, 2015: 3431-3440.

Lightweight maxillary anterior tooth segmentation algorithm based on improved BiseNet

Wu Jie, Tao Qingchuan*

(College of Electronics and Information Engineering, Sichuan University, Chengdu 610000, China)

Abstract: A segmentation algorithm is proposed for lightweight maxillary anterior tooth segmentation on low-computing power devices, such as mobile terminals. The algorithm is applied to natural color images. The benchmark model chosen for comparison is BiseNet-ResNet18, which performs better on the self-built dataset. The model is then further improved. The experimental results show that the mean intersection over union (MIoU) of the improved BiseNet on the self-built dataset reaches 90.7%, which represents a 1.2 percentage improvement over BiseNet. In terms of lightweight and real-time performance, the model size is only 1/8 of BiseNet, and it achieves a frame rate of 63.1 frames per second. The research results are expected to be applied in the fields of dental medicine and medical imaging, promoting the development of technologies such as tooth segmentation and automated analysis of smile aesthetics.

Keywords: tooth segmentation; semantic segmentation; BiseNet

(上接第9页)

Defect detection of quick-frozen dumpling based on machine learning

Liu Yiyang*, Ma Ziling

(Mechanical College of North China University of Water Resources and Electric Power, Zhengzhou 450045, China)

Abstract: In the freezing process, quick-frozen dumplings are easy to cause freeze cracking and damage, and it is time-consuming and labor-intensive to manually remove broken dumplings, so designing machine learning-based defect detection method for quick-frozen dumplings. First, the grayscale coexistence matrix and color moment of R, G, B, H, S, V color space are extracted with a total of 20-dimensional eigenvectors, which are optimized by the method of PCA. Six optimal eigenvectors are obtained as the input of the support vector machine. The parameters of the support vector machine are determined by the cross-algorithm hyper-parametric optimization, and the defective dumplings are classified and identified by the trained support vector machine model. The defective dumplings are externally connected with the minimum circumscribed rectangles, and the center point coordinates of the minimum circumscribed rectangle are output. The results show that the method can quickly identify the defective dumplings with the recognition accuracy of 95.5% and the time-consuming of 13.13 ms per dumping.

Keywords: defect detection of quick-frozen food; GLCM; PCA; support vector machine

文章编号: 1007-1423(2023)24-0040-06

DOI: 10.3969/j.issn.1007-1423.2023.24.006

改进 YOLOv5 的安全帽佩戴检测算法

王 坡, 罗红旗*

(北京工商大学人工智能学院, 北京 100048)

摘要: 确保工业生产中的操作人员佩戴安全帽以降低风险至关重要。然而, 传统检测方法存在小目标漏检和精度低的问题。故在原 YOLOv5s 模型上改进, 引入多项技术提升性能。首先, 融合 CBAM 注意力和 DB_CSP 模块, 加强模型的关键信息提取能力。DB_CSP 增强模型的特征多样性。其次, 使用 α -CIoU 损失替代传统 GIoU 损失, 强化目标定位和边界回归的精度, 提高鲁棒性。实验表明, 改进模型平均精度达 96.3%, 较原 YOLOv5s 提升 1.2 个百分点。本研究方法满足车间作业需求, 提高安全帽检测效果。

关键词: 安全帽佩戴; YOLOv5s; CBAM 注意力机制; α -CIoU 损失函数

0 引言

钢铁、化工等行业的机械加工车间内环境闷热, 操作工人作业时常不按规定佩戴安全帽, 为安全生产埋下巨大的隐患。例如起重机械吊运重物时, 有的工人安全意识淡薄, 在吊装区域长时间停留, 被意外滑落的物体击中头部身亡; 某些女工弯腰检修机床故障时, 头发被卷入运转中的设备, 导致头部被猛烈撞击身亡。因此, 安全帽的使用能够高效地保护工人的头部, 预防或减少因高空坠物或碰撞造成的头部伤害。传统的监管模式往往局限于静态监控和人工干预, 不能够及时响应和处理突发事件, 无法满足对高危行业安全管理的实时性和精确性要求。针对以上问题, 本研究选用 YOLOv5 目标检测算法, 该算法相较于传统检测方法, 在精度和速度上有所提高, 但对于小目标和密集目标的检测仍有不足之处, 本文将对 YOLOv5 检测算法进行改进, 使其在小目标和密集目标检测方面取得更好的效果^[1]。

近年来, 随着深度学习目标检测技术的持续发展, 国内外研究人员逐渐将深度学习技术应用于安全帽的检测中。基于深度学习技术的目标检测算法主要分为基于候选框的两阶段检测算法和基于回归的一阶段检测算法^[2]。一些研究者对目标检测算法进行了改进, 以提升性能和鲁棒性。刘晓慧等^[3]通过分析工人的面部特征和肤色相关信息来实现安全帽佩戴检测。将头发与安全帽颜色进行二分类判断, 来确定工人是否佩戴了安全帽。李鹏^[4]在 Faster R-CNN 的基础上进行了改进, 在模型中引入可变形卷积和可转换的空洞卷积, 以构建一个自适应目标尺度转换的多尺度安全帽佩戴检测网络结构, 这一改进在平均精确度方面相较之前的最优模型有所提高, 但对于小目标安全帽和被遮挡目标的检测效果并不理想。许凯等^[5]在 YOLOv3 的基础上引入了 K-means 聚类算法并增加特征图, 有效地缓解了正负样本不均衡问题, 提高了模型的性能和准确性。然而增加特征图使得网络结构变得更为复杂, 降低了检测速度。上述这些方法旨

收稿日期: 2023-08-13 修稿日期: 2023-09-24

基金项目: 国家重点研发计划项目(2017YFF0207201)

作者简介: 王坡(1998—), 男, 北京人, 硕士研究生, 研究方向为机械安全; *通信作者: 罗红旗(1977—), 男, 湖南人, 副教授, 博士, 研究方向为机械安全, E-mail: lhqfrq@126.com

在保留小目标特征，加强特征提取过程。

智能工业安全管理中的安全帽检测问题，还能够实时追踪未佩戴安全帽的工人。本文通过融合 CBAM 注意力模块、DB_CSP 模块和 α -CIoU 损失函数等前沿技术，有效地提升了模型的检测能力并增强模型的鲁棒性和稳定性。

1 YOLOv5 模型

YOLOv5 目标检测网络共有四种不同型号，其中 YOLOv5s 模型的深度和宽度最小，检测速度也最快，因此本文选择 YOLOv5s 作为基本网

络^[6]。与 YOLOv4 相比，骨干网络增加了 Focus 结构来执行切片操作，减少了计算量的同时提高了特征提取的速度。同时，骨干和颈部采取了两种不同的 CSP 结构。SPP 模块是一个空间金字塔池化模块，它使用不同尺寸的卷积核来解决输入图像的大小问题。

考虑到有效性和效率之间的平衡，本文选用主流的一阶段实时目标检测算法 YOLOv5s 网络模型作为该模型的主干。YOLOv5s 的网络结构包括输入端、主干网络、颈部网络、输出端四部分^[7]，其网络结构如图 1 所示。

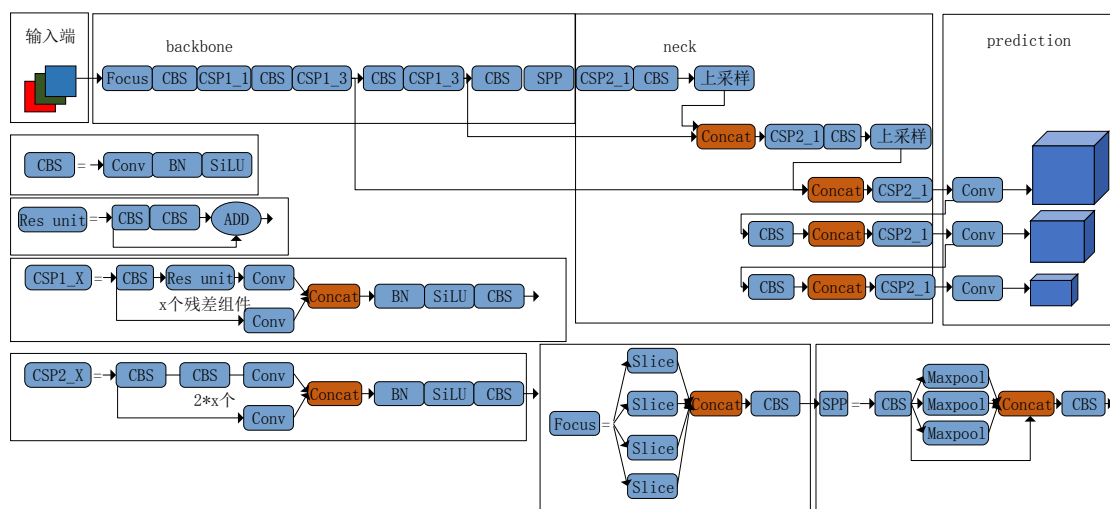


图1 YOLOv5s 网络结构图

2 模型的改进与优化

2.1 引入DB_CSP模块

YOLOv5 的重要改进是使用了两种不同的 CSP 结构，分别用于骨干网络和颈部网络。不同之处在于，骨干网络采用 BottleneckCSP (True)，颈部网络采用 BottleneckCSP (False)。其目的是加强网络的特征提取能力。

为改善梯度消失问题，有人提出了 DenseNet 网络^[8]。它吸取了 ResNet 和 GoogLeNet 的优点，使每两个卷积层之间都有连接，并直接将所有特征输入下一层。深层和浅层特征都能得到充

分利用，提取的信息丰富多样。DenseNet 计算公式如公式(1)所示。

$$xl = hl([x_0, x_1, \dots, x_{l-1}]) \quad (1)$$

式(1)中， xl 表示 l 层网络的输出特性， $x_0, x_1, \dots, x_{l-1}$ 表示 0 至 $l-1$ 层的特征，代表特征图的卷积、池化、归一化、激活函数等处理方法。

在此基础上，本文提出了一种新的网络连接模块 DB_CSP 替代原有的 BottleneckCSP (True)。其结构如图 2 所示，在 BottleneckCSP (True) 的各个 CBL 模块之间交叉连接，更好地完成特征融合，解决梯度消失带来的各种问题。

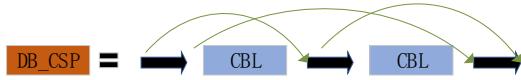


图2 DB_CSP结构图

2.2 CBAM注意力机制

卷积注意力模块(CBAM)^[9]是一种用于强化神经网络的紧凑型、适应性强的模块。本研究中, YOLOv5s的骨干和颈部跨阶段部分(CSP)模块的最后一层包含了CBAM模块。这种整合增强了模型提取特征的能力, 同时也降低了计算复杂度。

通道注意力模块和空间注意力模块是构成CBAM模块的两个子模块。首先从深度网络中获取中间特征图, 而后利用CBAM模块自适应地改进这些图片。因此, 注意力图就沿着通道和空间维度推断出来。图3即为CBAM注意力模块的细节。

中间特征图 $F \in \mathbf{R}^{C \times H \times W}$ 是CBAM的输入, 该模块通过顺序推理得到一维通道注意力图 $M_c \in \mathbf{R}^{C \times 1 \times 1}$ 和二维空间注意力图 $M_s \in \mathbf{R}^{1 \times H \times W}$, 注意过程的数学表达式为

$$F' = m_c(F) \otimes F \quad (2)$$

$$F'' = m_s(F') \otimes F' \quad (3)$$

式(3)中, 符号 $\otimes$ 表示注意过程中的元素乘法, 空间维度与通道维度一起传递, F'' 表示细化输出。

值得注意的是, 通过通道注意力机制时, 特征图在空间维度上被压缩, 生成一维向量, 相应的计算公式为

$$\begin{aligned} M_c(F) &= \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{Maxpool}(F))) \\ &= \sigma(W_1(W_0(F_{\text{avg}}^c)) + W_1(W_0(F_{\text{max}}^c))) \end{aligned} \quad (4)$$

空间注意力机制通过在信道维度上采用平均池化和最大池化来压缩信道, 其计算公式为

$$\begin{aligned} M_s(F) &= \sigma(f^{7 \times 7}[\text{AvgPool}(F); \text{MaxPool}(F)]) \\ &= \sigma(f^{7 \times 7}[F_{\text{avg}}^s; F_{\text{max}}^s]) \end{aligned} \quad (5)$$

通过将CBAM注意力模块嵌入骨干backbone和颈部CSP的最后一层, 深度网络的每个卷积块都会经过CBAM模块并对特征图进行自适应细化。这一过程降低了模型计算的复杂度, 并建立高维空间特征的相关性, 加强了网络对于重要特征的关注, 增强特征表达能力, 改善了模型的性能。表1即为改进后的YOLOv5s的网络结构图。

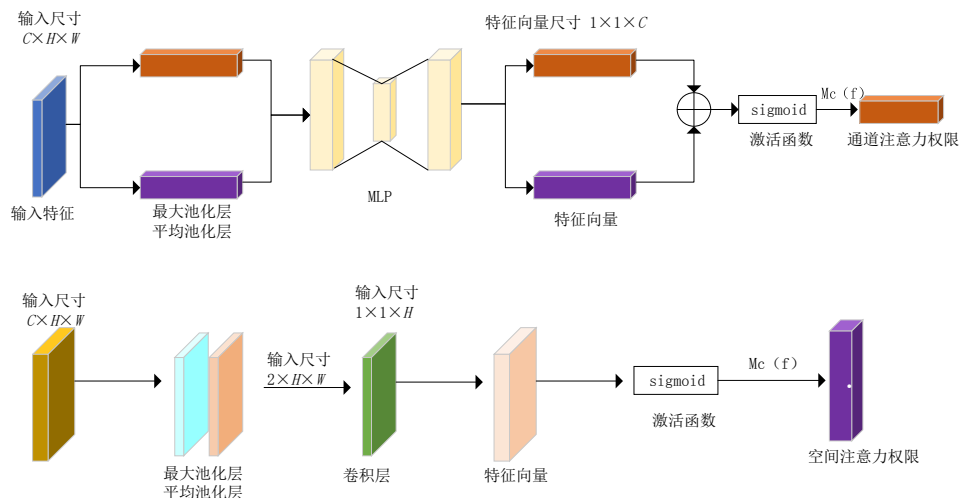


图3 CBAM注意力模块

表 1 改进后YOLOv5s网络结构图

Number	From	Params	Module	Arguments
0	-1	3520	Focus	[3,32,3]
1	-1	19560	Conv	[32,64,3,2]
2	-1	18826	C3	[64,64,1]
3	-1	73956	Conv	[64,128,3,2]
4	-1	156928	C3	[128,128,3]
5	-1	295424	Conv	[128,256,3,2]
6	-1	625152	C3	[256,256,3]
7	-1	1180672	Conv	[256,512,3,2]
8	-1	656896	SPP	[512,512,[5,9,13]]
9	-1	1215586	CBAMC3	[512,512,1,False]
10	-1	131584	Conv	[512,256,1,1]
11	-1	0	Upsample	[None,2,'nearest']
12	[-1,6]	0	Concat	[1]
13	-1	361984	DB_C3	[512,256,1,False]
14	-1	33024	Conv	[256,128,1,1]
15	-1	0	Upsample	[None,2,'nearest']
16	[-1,4]	0	Concat	[1]
17	-1	90880	DB_C3	[256,128,1,False]
18	-1	147712	Conv	[128,128,3,2]
19	[-1,15]	0	Concat	[1]
20	-1	296448	DB_C3	[256,256,1,False]
21	-1	590346	Conv	[256,256,3,2]
22	[-1,11]	0	Concat	[1]
23	-1	1215684	CBAMC3	[512,512,1,False]

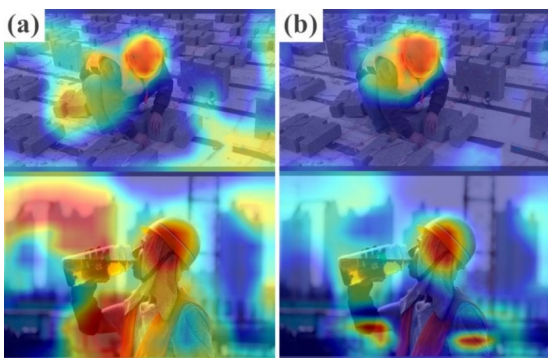


图 4 添加注意力机制前后模型热图特征对比

图4中，a是原始YOLOv5s模型下的测试结果，b是在原模型基础上添加关注机制后的结

果，可以看出添加CBAM注意力机制对于模型的检测结果有很大改善，能够更加准确的检测出头盔这类小目标。

2.3 边框损失函数的改进

传统YOLOv5s模型所采用的GIoU损失函数存在缺陷，即当预测框与真实框出现包含现象时，GIoU会退化成IoU，很难找到理想的预测框。为了解决这个问题，本文采用 α -CIoU损失函数代替GIoU损失函数来优化锚点框^[10]。

α -IoU损失为现有的IoU损失引入了功率变换，其定义如式(6)所示：

$$loss_{\alpha-IoU} = \frac{1 - IoU^\alpha}{\alpha}, \alpha > 0 \quad (6)$$

本文采用基于CIoU损失的 α -CIoU损失，计算式如下：

$$loss_{\alpha-CIoU} = 1 - IoU^\alpha + \frac{\rho^{2\alpha}(b, b^{gl})}{C^{2\alpha}} + (\beta v)^\alpha \quad (7)$$

式(7)中， c 表示可包含两者的最小封闭区域的对角线距离， β 是正权衡参数， v 衡量方面的一致性比率。

3 实验与结果分析

3.1 实验环境和数据集

本文的实验环境建立在RTX3060GPU和6GB内存环境上，包括CUDA和cuDNN，使用PyTorch和ResNet35模型，基于Python3.9实现。

针对目前数据集数量少、模型无法完全拟合特征等问题，本实验选取的数据集由实际工厂场景组成，检测模型数据集共计4965个，目标9720个，分类模型的数据集为30000个。数据集采用VOC格式。数据集按照7:3的比例划分，检测模型训练集为3475个，测试集为1490个。分类模型的训练集为20000个，其中不戴头盔的训练集为10000个，戴头盔的训练集为10000个，测试集为10000个。

3.2 评价指标

本文使用了三个主要指标衡量YOLOv5算法的有效性：准确率(P)、召回率($Recall$)、平均精确度均值(mAP)，分别由式(8)~式(11)表示。准确率是衡量模型预测正确性的指标。它衡量的是模型预测为正样本中实际为正样本的

比例,准确率高意味着模型预测正确的样本数量较多。召回率是衡量模型查全率的指标,它衡量的是模型能够正确检测到的正样本数量占所有正样本数量的比例。召回率高意味着模型能够更全面地检测出正样本。平均精确度考虑了模型在不同阈值下的表现,能够综合评估模型的检测效果。其中召回率和平均精确度的计算公式为

$$Precision = \frac{TP}{TP + FP} \quad (8)$$

$$R = \frac{TP}{TP + FN} \quad (9)$$

$$AP = \frac{TP + TN}{TP + TN + FP} \quad (10)$$

式(8)中: TP 代表佩戴了安全帽并且被正确检测出来的数目, FP 代表了佩戴安全帽但错误检测的数目。式(10)中: FP 代表平均精度,即所有召回率下得到的所有精度的平均值,平均精度均值(mAP)是在所有类别下 AP 值的平均值。平均精确度均值的计算公式为

$$mAP = \frac{1}{Q} \sum_{q \in Q} A(q) \quad (11)$$

式(11)中: Q 为总类别数量。

3.3 对比实验

为研究各种注意力机制模块对于 YOLOv5s 算法的影响,本文使用常用的注意力模块进行实验,如 SE 注意力机制模块、CBAM 注意力机制模块、TAM 自适应模块,表 2 列出了不同模块下 YOLOv5s 算法获得的训练精度。

表 2 类似算法比较

Models	P/%	R/%	mAP/%
YOLOv5	94.03	92.49	95.10
YOLOv5+CBAM	95.40	94.50	96.30
YOLOv5+SE	91.50	94.30	95.30
TAM-YOLOv5	92.60	94.20	94.60

从表 2 的结果可以看出, YOLOv5+CBAM 模型在准确率、召回率和平均精度上均优于原模型。相较于 YOLOv5+SE 模型,其精度、召回率、平均精度依次提高了 3.9、0.2 和 1.0 个百分点;与 TAM-YOLOv5 相比,其精度、召回率、平均精度依次提高了 2.8、0.3 和 1.7 个百分点。

上述结果表明 CBAM 注意力机制能够加强模型对输入数据的适应能力,有助于模型在各种变化和干扰下保持准确性。

3.4 消融实验

为评估不同组合提高算法性能的程度,本文设计了消融实验。消融实验中所有超参数在整个训练过程都是一致的,参数设计见表 3。

表 3 消融实验参数

类别	数值
轮数	300
批量大小	16
输入图片大小	640×640
优化器	SDGM
初始学习率	0.01

在消融实验中,改进后的 YOLOv5 网络在安全帽数据集上呈现了如下结果(见表 4)。实验结果显示,每项策略的引入都在不同程度上优化了模型结构。实验 2 表明,当网络中引入 DB_CSP 模块后, $mAP_{0.5}$ 和 $mAP_{0.5:0.95}$ 仅下降了 1.5% 和 1.9%。这说明 DB_CSP 模块改进了复杂性,略微降低了准确性。实验 3 显示, CBAM 注意力机制的引入增强了骨干网络的特征提取能力,进而提高了模型检测的准确性。实验 4 中, α -CIoU 损失函数在不增加计算复杂度的前提下,将 mAP 与 $mAP_{0.5:0.95}$ 提升了 1.7% 和 1.9%。上述结果表明 α -CIoU 损失函数能够提高模型的定位精度、增强边界框回归精度,提升了模型的鲁棒性,在小目标识别检测过程中表现更优,使得模型能够更准确地定位和识别目标,从而提升模型整体的检测性能。

表 4 消融实验结果

	DB_CSP	CBAM	α -CIoU	$mAP_{0.5}/\%$	$mAP_{0.5:0.95}/\%$
1				75.2	47.6
2	Δ			74.1	46.7
3		Δ		75.7	47
4			Δ	76.5	48.5
5		Δ	Δ	78.3	49.1
6	Δ	Δ		76.5	46.3
7	Δ		Δ	76.5	47.6
8	Δ	Δ	Δ	76.9	48.3

不同策略的组合对网络结构产生积极优化效果。通过比较实验5、实验2和实验3的数据,可观察到使用DB_CSP模块和 α -CIoU损失的网络在mAP0.5和mAP0.5:0.95上分别比仅使用DB_CSP模块的网络提升了1.7%和2.1%,比仅使用 α -CIoU损失的网络分别提升了0.7%和0.3%。实验6和实验7与实验2的对比显示,在引入DB_CSP模块后,CBAM注意机制和 α -CIoU损失均适用于网络。加入CBAM注意机制后,网络的mAP0.5和mAP0.5:0.95分别提高了1.1%和0.8%;加入 α -CIoU损失机制后,网络的mAP0.5和mAP0.5:0.95分别提高了2.0%和2.1%。这表明同时采用这三种策略会在一定程度上减缓单一策略的优化效果,但仍能保持良好的模型精度和网络结构深度。

4 结语

本文提出一种改进YOLOv5的安全帽佩戴检测算法,首先在原模型的骨干网络和颈部网络DB_CSP模块,而后在骨干网络和颈部网络连接处的CSP模块最后一层融合了CBAM注意力机制,加强了模型的特征提取能力,降低了计算的复杂度,使得模型对于头盔等小目标的检测能力大大增强。对于原模型锚点框的优化,本文引入了 α -CIoU损失函数,提高了模型目标检测精确度。实验表明,与原始模型相比,本文所提出的算法在精确度和召回率方面都有一定程度的提升,可以满足车间现场作业人员穿戴安全帽检测的实时性。

参考文献:

- [1] 刘欣,张灿明.基于卷积神经网络的矿井安全帽佩戴检测[J].电子技术应用,2020,46(09):38-42,46.
- [2] 陈勇,黄涛,张煜昊,等.工地现场多目标危险行为检测算法的研究与实现[J].计算机应用研究,2020,37(S2):313-315.
- [3] 刘晓慧,叶西宁.肤色检测和Hu矩在安全帽识别中的应用[J].华东理工大学学报(自然科学版),2014,40(3):365-370.
- [4] 李鹏.基于目标检测与深度估计的施工现场安全预警关键技术研究[实现[D].成都:电子科技大学,2021.
- [5] 许凯,邓超.基于改进YOLOv3的安全帽佩戴识别算法[J].激光与光电子学进展,2021,58(6):300-307.
- [6] 杨永波,李栋.改进YOLOv5的轻量级安全帽佩戴检测算法[J].计算机工程与应用,2022,58(9):201-207.
- [7] 王玲敏,段军,辛立伟.引入注意力机制的YOLOv5安全帽佩戴检测方法[J].计算机工程与应用,2022,58(9):303-312.
- [8] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017:4700-4708.
- [9] HE J, ERFANI S, MA X, et al. α -IoU: a family of power intersection over union losses for bounding box regression[J]. Adv Neural Inf Process Syst 2021, 34:20230-20242.
- [10] 李冰涛,李大海.面向小目标的YOLOv5s安全帽佩戴检测[J].计算机系统应用2023,32(8):221-229.

Improved helmet wearing detection method of YOLOv5

Wang Po, Luo Hongqi*

(School of Artificial Intelligence, Beijing Technology and Business University, Beijing 100048, China)

Abstract: It is important to ensure that operators in industrial production wear safety helmets to reduce risks. However, the traditional detection methods have the problems of missing small targets and low precision. Therefore, the original YOLOv5s model is improved and a number of technologies are introduced to improve the performance. Firstly, CBAM attention and DB_CSP modules are integrated to enhance the key information extraction capability of the model. DB_CSP enhances the feature diversity of the model. Secondly, α -CIoU loss is used to replace the traditional GIoU loss, which strengthens the precision of target positioning and boundary regression, and improves the robustness. Experiments show that the average accuracy of the improved model is 96.3%, which is 1.2 percentage higher than that of the original YOLOv5s. This research method can meet the requirements of workshop operation and improve the effect of safety helmet detection.

Keywords: helmet wearing; YOLOv5s; CBAM attention mechanism; α -CIoU loss function

文章编号: 1007-1423(2023)24-0046-06

DOI: 10.3969/j.issn.1007-1423.2023.24.007

一种聚类矩阵和剪枝策略结合的关联规则算法

许珂乐*

(咸阳职业技术学院信息工程学院, 咸阳 712000)

摘要: 在医学领域, Apriori 关联规则算法在慢性病等方面有广泛应用。然而, 该算法存在一些问题, 比如需要多次扫描病例数据库、产生大量数据集和冗余规则。针对出现的相关问题, 提出一种结合聚类矩阵和剪枝策略的关联规则数据挖掘算法。通过添加约束条件, 并采用聚类矩阵方法来压缩存储事务数据库。同时, 引入前剪枝和后剪枝策略优化算法。实验证明, 这种优化算法显著减少了数据库的扫描次数, 降低了候选项目集的数量, 缩短了算法的执行时间, 大大提高了运行效率。

关键词: 数据挖掘; 关联规则; Apriori 算法; 候选项集; 聚簇矩阵; 剪枝策略

0 引言

随着信息化技术的进步, 数据库和计算机技术各个领域得到广泛应用, 产生了海量数据。如何快速从这些数据中找到有用的信息成为重要问题, 数据挖掘技术应运而生。通过数据挖掘技术, 我们可以发现并探索海量数据背后隐藏的有用信息。商业模式的不断涌现, 数据挖掘技术的商业价值愈发凸显, 成为关注焦点。此外, 医疗行业也迎来了大数据时代, 医疗系统收集的数据对医学研究和慢性病诊治具有重要价值。然而, 目前大多数医院只是收集和存储患者数据, 没有深入分析和有效利用。如何利用数据挖掘技术从医疗数据中发现隐藏的关系, 并为临床诊断提供支持成为研究人员关注的问题。关联规则在数据挖掘技术中能帮助医生发现慢性病患者的体征数据, 为精准诊断提供科学依据。因此, 数据挖掘技术在医学研究和医院评估中的作用日益重要, 已成为现代医学的不可或缺部分。

关联规则^[1]、粗糙集理论^[2]、机器学习、神经网络^[3]等数据挖掘方法在医疗数据挖掘中各

有优缺点^[4]。在数据挖掘领域中, 关联规则的数据挖掘模式是一个重要的研究方向, 已经被广泛应用到电子商务、慢性病的预防^[5]、电信和保险等行业。关联规则是数据挖掘的重要研究方向, 通过数据挖掘得到的关联规则信息对研究挖掘出的效果有着非常重要的作用。所以研究关联规则算法是数据挖掘中的重要部分。

关联规则算法数据挖掘被应用到慢性病医疗数据分析中已具有较丰富的实践经。文献[6]利用聚集算法数值型属性的特点将数据集进行离散, 分为若干个优化数据集, 挖掘出肿瘤诊断数据中有用的关联规则, 为肿瘤慢性病的临床诊断提供重要的参考价值; 文献[7]应用人工神经网络和 Logistic 回归分析技术建立了慢性病患者的预测模型, 对高血压、糖尿病等慢性病提供了较好的预测效果; 文献[8]使用分类算法对 I 型糖尿病患者的数据进行分析, 发现所产生的分类关联规则与医学研究结果高度吻合, 为研究糖尿病患者的病情提供了较好的理论研究基础; 文献[9]对传统的 Apriori 算法进行了综述, 找出其最大的缺陷是需要对数据库进行多

收稿日期: 2023-07-09 修稿日期: 2023-11-14

基金项目: 咸阳职业技术学院科研基金项目(2021KJB05)

作者简介: *通信作者: 许珂乐(1983—), 女, 陕西咸阳人, 硕士, 副教授, 主要研究方向为计算机技术以及大数据分析与应用, E-mail: 18992009262@189.cn

次扫描后才能得到频繁项集，多次扫描数据库必将影响数据挖掘的效率和占用较大的内存；文献[10]提出了一种矩阵和排序索引的关联规则数据挖掘算法，该算法通过删除事务数据库中无用的事务和项，然后通过矩阵的乘积得到频繁的二项式集合，最后得到频繁 k -项集，该算法的最大优势是可以直接查找频繁项集并对数据库进行扫描，数据挖掘的效率从一定程度上获得了大幅度的提高，但是该算法却需要对数据库进行实时更新。虽然关联规则数据挖掘技术的理论研究已经非常广泛，但是现有的算法不能准确快速地获取规则，挖掘效率低下，对数据挖掘系统的准确性和有效性有非常大的影响，所以，提升关联规则算法的挖掘效率具有较好的发展前景。

1 关联规则技术

(1) 关联规则可以通过以下定义描述：

定义 1： 设项目集合 $I = \{I_1, I_2, I_3, \dots, I_n\}$ ，数据 D 组成事务集合。假设 X 和 Y 是两个项集，且 X 和 Y 均为 I 的非空子集，关联规则是形如 $X \rightarrow Y$ 的一种蕴含式，在蕴含式中 $X \subset I$ ， $Y \subset I$ ，且 $X \cap Y = \emptyset$ 。若 X 集合中的元素个数为 k ，则称项集 X 为 k -项集。

定义 2： 支持度 $(\text{support})(X \rightarrow Y)$ 表示包含 X 和 Y 的元组数/元组总数^[11]，也即是在数据库中满足规则的元组数占元组总数的百分比，如公式(1)所示：

$$\text{support}(X \rightarrow Y) = P(X \cup Y) = \frac{\text{count}(X \cap Y)}{\text{count}(D)} \times 100\% \quad (1)$$

定义 3： 置信度 (confident) 表示数据库中包含一个事务 X 的前提下也包含事务 Y 所占的百分比，即值的信赖性。置信度 $(X \rightarrow Y)$ 可以定义为：置信度 $(X \rightarrow Y)$ 表示包含 X 和 Y 的元组数/包含 X 的元组数，如公式(2)所示：

$$\text{confident}(X \rightarrow Y) = P(Y/X) = \frac{\text{support}(X \cap Y)}{\text{support}(X)} \times 100\% \quad (2)$$

定义 4： 设 min_sup 、 min_conf 分别为最小支持度阈值、最小置信度阈值，假如事务集合 D 中 $X \rightarrow Y$ 关联规则符合公式(3)和(4)要求：

$$\text{support}(X \rightarrow Y) \geq \text{min_sup} \quad (3)$$

$$\text{confident}(X \rightarrow Y) \geq \text{min_conf} \quad (4)$$

则 $X \rightarrow Y$ 称为事务集合 D 中的强关联规则。

定义 5： 频繁项集是指 $\text{support} \geq (\text{min_sup})$ 的集合，反之当支持度小于最小支持度 (min_sup) 的集合称为非频繁项目集。

(2) Apriori 算法

在关联规则数据挖掘算法中，Apriori 算法是应用最广泛的算法之一，该算法产生的关联规则在分类上属于单层、布尔型关联规则。为了找到事务数据库中的频繁项集，Apriori 算法采用了迭代逐层搜索的方法，反复从长度为 k 的频繁项集里产生长度为 $k + 1$ 的候选项集，再进一步生成长度为 $k + 1$ 的频繁项集两个阶段生成频繁项集的思想。

表 1 Apriori 符号定义

k -项集	含有 k 个项目的集合
L_k	最小支持度阈值项集
C_k	频繁项集

Apriori 算法在执行时，首先进行事务数据库的扫描得到频繁项集中每个元素的支持度。判断该元素是否加入 L_k ，若产生的频繁项集是包含 n 个项目的大项集，则扫描事务数据库的次数至少是 n 次，事务数据库越大，I/O 的负载就越大；在算法执行的第二个阶段，频繁项集 L_{k-1} 产生候选项集 C_k 是指数级增长的，若频繁模式的长度为 200，候选项集最少为 2^{100} 个，候选 2-项集 C_m^2 中的 m 越大，得到的最后结果就越可观。Apriori 算法产生频繁项集时，当事务数据库数量较大、最小支持度阈值和最小置信度阈值设置太低，都将产生大量的关联规则，这些规则很多都是冗余规则。针对以上问题，本文提出一种项目约束下的聚类矩阵和剪枝策略相结合的关联规则数据挖掘算法。

2 一种含有项目约束的 Apriori 关联规则算法

传统关联规则的缺陷需要利用最小置信度阈值、最小支持度阈值来发现信息，算法执行结束后会产生较多的规则，若是在挖掘的过程中用户能够进行指导和控制，那么就可以增加

产生有效关联规则的数目,提高挖掘的效率。在挖掘的过程中设置适当的项目约束可以达到较好的效果,从而获取有效的关联规则。通过研究,首先可以利用约束条件产生较少频繁项集数目,从而降低数据库扫描的次数;其次利用构件约束条件过滤掉不符合条件的关联规则。将上述思想应用到生成关联规则和Apriori算法的频繁项集的产生中,具体的挖掘模型如图1所示。

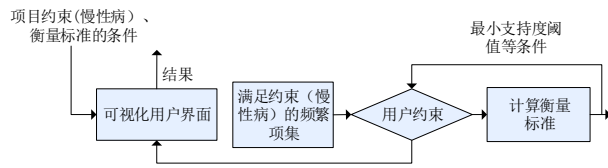


图1 添加用户约束的挖掘模型

通过上节分析已知Apriori算法存在的问题,本文将在原有算法基础上首先添加项目约束条件;其次利用压缩矩阵对事务数据库进行存储,通过矩阵向量的运算,计算候选项集的支持度;最后在自连接之前加入剪枝方法,从而减少频繁项集连接时产生的大量无用候选项集,进一步提高挖掘效率。

2.1 矩阵向量策略

在事务数据库 D 中进行以下定义:

定义1: 设 $I = \{I_1, I_2, I_3, \dots, I_n\}$ 是项目集合,集合中的每项 n_j 的向量记为 D_j :

$$D_j = \begin{bmatrix} d_{1j} \\ d_{2j} \\ \vdots \\ d_{nj} \end{bmatrix}, \text{ 其中 } d_{nj} = \begin{cases} 0, & n_j \notin D \\ 1, & n_j \in D \end{cases} \quad (5)$$

n_j 支持度频度为

$$\text{support_c}(n_j) = \sum_{n=1}^m d_{nj} \quad (6)$$

定义2: 2-项集 $\{i_i, i_j\}$ 的向量记为 D_{ij} :

$$D_{ij} = \begin{bmatrix} d_{1i} \wedge d_{1j} \\ d_{2i} \wedge d_{2j} \\ \vdots \\ d_{ni} \wedge d_{nj} \end{bmatrix} \quad (7)$$

其中,符号“ $\wedge$ ”表示“逻辑与”运算。

2-项集的支持度频度为

$$\text{support_c}(\{i_i, i_j\}) = \sum_{i=1}^m (d_{ii} \wedge d_{ij}) \quad (8)$$

定义3: k -项集 $\{i_1, i_2, \dots, i_k\}$ 的向量记为

$D_{1,2,\dots,k}$:

k -项集 $\{i_1, i_2, \dots, i_k\}$ 的支持度频度为

$$\text{support_c}(\{i_1, i_2, \dots, i_k\}) = \sum_{i=1}^n \{(d_{i1} \wedge d_{i2} \wedge \dots \wedge d_{i(k-1)}) \wedge d_{ik}\} \quad (9)$$

根据以上定义,矩阵向量策略执行过程如下:

(1) 将聚簇矩阵的策略实际应用应用到Apriori算法中。首先,扫描事务数据库 D ,并对其生成聚簇矩阵 A 。若该事务数据库中项目的个数是 k ,则将其聚簇到矩阵 $A(k)$ 中。聚簇矩阵 $A(k)$ 中的每一列代表集合中不同的项,每一行代表包含 k 个项集的详细事务记录,矩阵中的布尔量0和1分别表示该项在矩阵中是否有记录,有记录记为1,无记录记为0。对矩阵中的列向量利用逻辑运算的“与”进行计算,根据公式(5)计算出每个项目的支持度,最后确定频繁项集1-项集 L_1 。

(2) $L_{k-1} (k \geq 2)$ 自连接产生候选项集 C_k 的集合。

(3) 对聚簇矩阵中的列向量进行“与”运算并求和。若是聚簇矩阵“与”运算后,计算得到的支持度符合设定的阈值,则直接将该项集放入频繁 k -项集的集合 L_k 中,否则需要继续对聚簇矩阵 $A(k+1)$ 进行相应的列向量“与”运算,将运算所得的结果与上一个支持度频度进行累加,当累加的支持度小于设定的最小支持度阈值时,继续扫描后续矩阵,重复上述计算。

2.2 剪枝策略

Apriori算法利用 L_{k-1} 自连接生成候选项集的集合 C_k ,在自连接过程中,必然会生成大量的非频繁项集,若算法在生成 C_k 之前能够根据某种规则判断出某些项目是非频繁项集,便可以提前剪除这部分项集,这种情况下既节省了他们自连接的操作,又减轻了扫描事务数据库时计算其支持度频度的时间,从而在很大程度上降低了算法的时间复杂度。剪枝策略应用到频繁项集以下几个性质:

性质1: 若包含 k -项集 T 是频繁项集,则 T 的所有子集均是频繁项集。

性质2: 若 k -项集 T 中有一个子集不是频繁项集,则 T 不是频繁项集。

性质3: 若 k -项集 $T = \{t_1, t_2, \dots, t_k\}$ 中存在项目 $i \in T$, 且 $|L_{k-1}(i)| < i - 1$, 则 T 不是频繁项集, $|L_{k-1}(i)|$ 表示集合 L_{k-1} 中包含项目 i 的数目。

根据以上性质, 剪枝策略实施过程如下:

(1) 计算 $|L_{k-1}(i)|$, 即计算 L_{k-1} 中每个项目 i 出现的次数。

(2) 计算项集出现的频度, 对所有频度小于 $k - 1$ 的项目用 $A = \{i | L_{k-1}(i) < i - 1\}$ 方式记录。

(3) 删除 L_{k-1} 中所有含有 A 的任一元素的频繁项集, 并记录为 L'_{k-1} 。

(4) 集合 L'_{k-1} 进行自连接, 得到候选项集 C'_k 。

在剪枝策略的实施过程中, L'_{k-1} 与自连接后可以利用Apriori算法的剪枝策略再进一步缩小候选项集的生成数目。通过先剪枝策略减少了频繁项集连接时所需的时间, 同时也为后剪枝的实施减轻了工作量。在实施了前剪枝和后剪枝策略后, 候选项集的数量减少, 一定程度上降低了计算支持度时数据库扫描过程中的时间开销。

2.3 算法流程

该算法的执行过程是在原有的Apriori算法的思想上进行优化, 具体过程如下:

(1) 对事务数据库 D 进行扫描, 利用聚簇的方法对 D 中的项进行操作得到聚簇矩阵 A , 如果项集的个数是 k , 聚簇得到的矩阵为 $A(k)$ 。

(2) 通过求解每个列向量的元素之和, 利用公式(8)计算支持度频度。对计算结果进行筛选, 将满足条件($< \min_sup$)的项目加入到项集中, 频繁项集1-项集 L_1 。

(3) 对 L_{k-1} 执行前剪枝策略, 并计算每个项目出现的次数, 将所有支持度频度小于 $k - 1$ 项目的非频繁项集删除掉, 得到的集合记为 L'_{k-1} 。

(4) 集合 L'_{k-1} 进行自连接, 生成候选 k -项集 C'_k 。

(5) 利用原Apriori算法的性质对 k -项集 C'_k 进行后剪枝生成频繁项集 C_k , 减少候选项集中的无关项目。

(6) 生成频繁 k -项集, 频繁项集 C_k 中每个项集的长度为 k , 所以无需考虑长度小于 k 的聚簇矩阵。对聚簇矩阵 $A(k)$ 中的列向量进行“逻辑与”运算, 得到每个候选项集对应的列向量,

统计聚簇矩阵中列元素的和, 计算出对应候选项集支持度。若最小支持度阈值小于候选项的支持度, 则将该项集加入 L_k ; 若最小支持度阈值大于候选项的支持度, 则需要对 $A(k + 1)$ 中的列向量进行处理, 求出该项集的支持度频度, 并累加到计数中, 直到支持度不小于设置的最小支持度阈值时, 或者扫描完聚簇矩阵中的所有项目。

(7) 重复执行步骤(3)~(6), 直到 L_k 为空, 确定频繁项集 C_k 。

3 仿真实验

为了进一步验证算法的有效性和先进性, 本实验的环境使用Pentium(R)双核2.62 GHz/3.0 GB微型计算机(操作系统为Windows 7), 采用了MATLABR2012a仿真环境, 使用UCI标准测试数据集的Mushroom作为实验的数据集, 数据集中总共包含了9000条记录, 每条记录有27的属性, 每个属性有15个枚举值。

(1) 设置不同的最小支持度阈值, 相同的数据集。通过比较, Apriori算法和优化算法在不同支持度下生成候选项集的数量不同, 比较结果如图2所示。

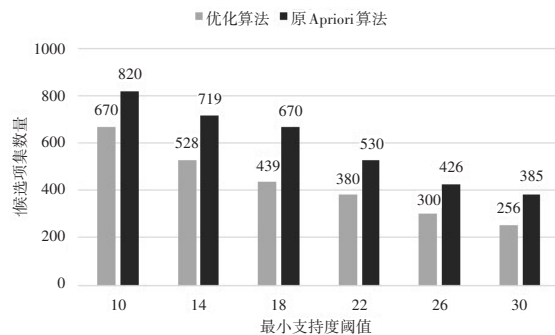


图2 支持度阈值设置不同时产生的候选项集

通过图2可知, 随着最小支持度阈值的逐渐增大, 优化算法和原Apriori算法的候选项集数量却都在减少, 但在相同的支持度下, 优化算法生成的候选项集的数量却小于原Apriori算法。所以, 优化算法在不同支持度阈值下生成候选项集的数量上具有明显优势。

(2) 设置不同的最小支持度阈值, 相同的数据集。通过对Apriori算法和优化算法在不同支持度下运行时间的比较, 得出如图3所示的结果。

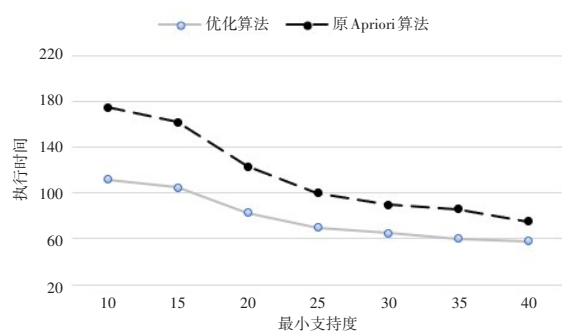


图3 不同支持度阈值下算法执行时间

通过图3可以发现,随着最小支持度阈值的逐渐增大,优化算法和原Apriori算法的执行时间都在减小,但在相同的支持度下,优化算法生成的候选项目集所用时间都小于原Apriori算法所用时间。因此可以得出,优化算法在不同支持度阈值下算法执行时间上具有明显优势。

(3) 假设 $\min_sup=0.3$, 对不同数量的数据集下优化算法和原Apriori算法的执行时间进行比较。所得结果如图4所示。

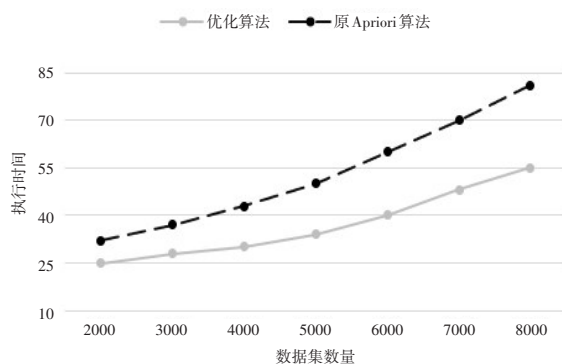


图4 不同数量数据集下的执行时间

从图4可以看到,随着数据集数量逐渐增大,优化算法和原Apriori算法的执行时间都在增大,但在相同的数据集数量下,优化算法所用的执行时间都小于原Apriori算法所用时间。所以,优化算法在不同数量数据集下的执行时间具有明显优势。

4 结语

本文提出一种项目约束条件下的聚类矩阵和剪枝策略相结合的关联规则数据挖掘算法,该算法在设定项目约束的前提下应用聚类矩阵和剪枝策略的方法,较大程度地降低了算法生

成候选项集的数量,避免了多次扫描数据库带来的I/O开销问题,提高了算法的执行效率。通过仿真实验从不同支持度阈值下生成候选项集的数量和算法执行时间、不同数量数据集下算法执行时间三方面进行了验证,实验结果表明提出的算法较好地提升了执行效率。鉴于本文中的算法生成关联规则是以形式化的公式展现,对于用户来说不容易理解,因此,以后研究工作的重心将是对数据挖掘过程可视化进行进一步的深入研究。

参考文献:

- [1] 韩存鸽. FP-growth算法的研究与改进[J]. 龙岩学院学报, 2020, 38(5): 8-13.
- [2] JAVIER L Z, TORRALBO J, CRISTOBAL R. Early prediction of student learning performance through data mining: a systematic review [J]. Psicothema, 2021, 33(3): 456-465.
- [3] 朱晴. 融合关联规则挖掘算法的信息化教学管理系统设计[J]. 现代电子技术, 2020, 43(23): 159-163.
- [4] 苗莉. 基于数据整合视角的高校教职工慢病健康管理的应用[J]. 北京联合大学学报, 2021, 35(2): 69-78.
- [5] 徐小兵, 李迪, 孙扬, 等. 基于关联规则的中国老年人慢性病共病分析[J]. 中国慢性病预防与控制, 2021, 29(11): 808-812.
- [6] SERBANATI L D. Health digital state and Smart EHR systems[J]. Informatics in Medicine Unlocked, 2020, 21: 29-36.
- [7] 颜建军, 陈松晔, 燕海霞, 等. 基于递归图和卷积神经网络的脉象分析识别[J]. 计算机工程与应用, 2020, 56(7): 170-175.
- [8] 周凯, 顾洪博, 李爱国. 基于关联规则挖掘Apriori算法的改进算法[J]. 陕西理工大学学报(自然科学版), 2018, 34(5): 40-44.
- [9] 文芳, 黄慧玲, 李腾达, 等. 基于FP-growth关联规则的图书馆数据快速挖掘算法研究[J]. 重庆理工大学学报(自然科学), 2020, 34(6): 189-194.
- [10] 郭倩, 殷丽凤. 基于散列技术的多层关联规则算法的改进[J]. 计算机工程与设计, 2021, 42(9): 2485-2491.
- [11] 刘彦戎, 杨云. 一种矩阵和排序索引关联规则数据挖掘算法[J]. 计算机技术与发展, 2021, 31(2): 54-59.

(下转第60页)

文章编号: 1007-1423(2023)24-0051-05

DOI: 10.3969/j.issn.1007-1423.2023.24.008

多维攻击下的网络入侵信号检测研究

岳 茹*

(山西省政法管理干部学院, 太原 030012)

摘要: 在多维攻击下网络运行环境中, 由于会遇到未知的入侵干扰信号, 且不同维度类型的入侵行为具有不同的未知特征, 导致入侵信号检测难度增大。为此, 提出多维攻击下的网络入侵信号检测方法。通过部署近场源均匀线列阵分布结构, 实现对网络中的源入侵信号进行捕获, 建立相应的入侵信号模型。针对不同的网络入侵行为, 提取相应的干扰信号特征。为了提高检测的准确性和可靠性, 对信号进行抗干扰滤波处理, 以排除或减少干扰信号的影响, 提高网络入侵信号检测的效果。对比实验结果表明, 该研究方法具有理想的去噪效果, 入侵信号检测仅耗时 0.40 ms, 应用效率较高。

关键词: 多维攻击; 网络入侵信号检测; 网络安全; 检测方法

0 引言

网络安全关系到用户隐私、军事信息安全、贸易保密传输安全和用户账户安全。一旦网络攻击和入侵发生, 将很容易造成信息泄露和网络扣押^[1]。网络具有自组织和多层节点的多维攻击, 容易受到网络入侵攻击。而网络攻击中的病毒移植过程主要是一个信号移植过程, 其中网络入侵数据是除固定时间差信号之外的多个组成部分。利用信号处理方法, 以实现对网络入侵的检测和主动防护。因此, 研究网络入侵信号检测方法具有重要意义。

目前, 网络入侵检测主要应用于防火墙。在复杂的网络环境中, 为实现网络安全, 国内学者展开大量研究, 并取得了一些研究成果。李响等^[2]针对这一问题, 采用哈里斯鹰方法, 对 CNN 网络结构进行自适应优化, 建立基于 CNN 网络的入侵检测模型, 针对传统卷积神经网络在全连接层中存在的过拟合问题, 提出了一种基于全局池化的神经网络模型, 并将其应用到卷积神经网络中。利用哈里斯鹰算法对

CNN 优化, 减少网络参数的选取, 以完成入侵检测。方熙宇等^[3]利用车载 CAN 网络, 并结合信号所属特征, 研究出一种基于信号的入侵检测, 通过构建通信矩阵, 将 CAN 网络分解成若干个信号, 以监测网络数据和状态, 从而实现入侵信号检测。

在以上研究方法的基础上, 提出多维攻击下的网络入侵信号检测研究。实验结果表明, 该方法有效地提高了网络入侵信号的检测能力, 检测效率高。

1 网络入侵信号检测方法设计

1.1 基于近场源均匀线列阵的入侵信号捕获与建模方法

为深入理解不同类型的入侵行为和攻击方式, 通过对已知入侵样本的分析, 构建网络入侵信号模型, 有助于提高对于入侵行为的认知和了解, 为网络安全防御提供基础。网络入侵数据是一种广泛而平稳的高斯线性时间信号模型^[4], 由相同的线性接近源划分为网络空间。

收稿日期: 2023-08-18 修稿日期: 2023-12-07

作者简介: *通信作者: 岳茹(1982—), 女, 山西运城人, 本科, 讲师, 主要研究方向为计算机科学与技术, E-mail: qi-yan3981@163.com

单个网络输入信号的近场来源的线性分布结构如图1所示。

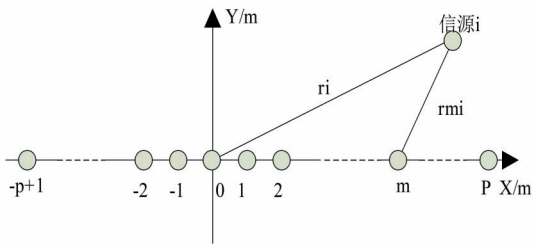


图1 近场源均匀线列阵分布结构图

图1显示了场源的最近分布结构中的网络入侵信号，且网络入侵信号间的交叉项具有非零峰度^[5]，通过计算不同信号频率分量的网络路由中心的空间分布向量的最大值，以捕捉到网络中的变化模式，通过比较不同时间段或不同网络状态下最大值的变化，可以检测到潜在的入侵信号，提高模型的检测能力。不同信号频率分量的网络路由中心的空间分布向量的最大值 r 表达式为

$$r \leq 2D^2/\lambda \quad (1)$$

在式(1)中， D 是入侵半径 λ 的恒定时间信号，为入侵信号的总频率。

时间序列可以提供关于入侵发生的时间和顺序信息，有助于确定入侵行为的时序模式和规律，能够尽早地发现和响应潜在的入侵威胁。假设网络攻击块中心线的协调为0，则 m 个块接收到的网络入侵时间序列为

$$T_m(t) = r \cdot s_i(t) \cdot \phi_{mi}/\alpha_m(t) \cdot n_m(t) \cdot \phi_{mi} \quad (2)$$

在式(2)中， $s_i(t)$ 是平面上的信号复包络， $\alpha_m(t)$ 表示网络入侵信号的时频相位， $n_m(t)$ 为序列中的高度非线性分布噪声， ϕ_{mi} 为网络入侵信号中的经验正交误差。采用傅立叶离散变换对入侵时间序列中的网络入侵信号进行解析，获取网络源入侵信号 $z(t)$ ：

$$z(t) = T_m(t) \cdot x(t) + i(t) \cdot a(t)/\theta(t) \quad (3)$$

在式(3)中， $x(t)$ 是通过傅立叶离散变换数据分析处理的实际入侵信号的一部分， $i(t)$ 为入侵信号分析模型的虚部， $a(t)$ 是入侵信号 $x(t)$ 的卷积， $\theta(t)$ 表示即时频率的入侵信号。

希尔伯特变换是一种时频分析方法，可以将信号从时域转换到方便处理的复频域，以提

供更全面的入侵信号模型。因此，利用希尔伯特变换方法获得入侵信号模型：

$$A = \sqrt{z^2(t) \cdot T_m(t) + \theta^2(t)}/\phi_{mi} \quad (4)$$

1.2 优化入侵检测模型中的干扰信号特征提取

入侵检测模型通常基于特定的入侵行为构建，由于在实际运行环境中可能遇到未知的入侵干扰信号，通过对入侵干扰信号特征提取，可以使模型更具适应性和泛化能力，能够在面对各种干扰信号时保持高效和准确的入侵检测能力。为了提高移动网络上通信干扰信号的识别质量，采用概率统计方法创建与随机过程中每个随机变化的频率相对应的新的频谱幅度序列，根据构建的入侵信号模型，提取干扰信号的特定频谱^[6]。具体措施是：假设 X 是干扰信号的 n 个意外变量集， y 是每组频率的范围， n 是恒定的噪声能量阈值，并且与加速过程中每个随机变量的频率相对应的频谱幅度构成新序列，表达式如下。

$$a_{afh} = \frac{\alpha_m(t) \cdot A}{n \cdot \{fgdhj\}} \cdot X \quad (5)$$

在上述方程中， $fgdhj$ 是分布的频谱码序列。

假设 X 是通信潜在干扰信号的有限可加性，由于在由通信干扰信号频谱序列的增加量 M 是随机的，因此所讨论的频谱幅度的潜力也增加了。根据频谱幅度，构建通信干扰信号的频谱矩阵：

$$X(f) = \frac{M}{n \cdot f_0} \cdot \frac{L \cdot a_{afh}}{2\pi} \quad (6)$$

在式(6)中， L 是线性调制频率， f_0 是初始调制频率， a_{afh} 是时域分段的数量， 2π 是干预信号采样周期。

在多维攻击中，基于经验状态特征的分布和能谱密度特征的提取，提出了一种网络入侵检测的动态方法。经验机制的特征共享应通过多维攻击中动态入侵信号的网络来执行^[7]，并且网络的动态信号入侵应划分为多个IMF组件。假设在多维攻击期间，作为多维攻击的一部分获得的网络入口特征的时间序列 $T_m(t)$ ，利用构建的通信干扰信号的频谱矩阵，提取网络入侵特征的时域特征为

$$W(t, v) = X(f)T_m(t) \exp(-2\pi a) d \cdot a \quad (7)$$

在式(7)中： a 表示多维攻击网络入侵信号域的频率， d 为由于网络入侵信号受到多维攻击的特定因素。

采用网络动态入侵信号检测器将经验机制分配给多维攻击中的动态网络入侵信号，以获得多维攻击中网络的动态入侵分量的数量。

$$B = c_i + r_n W(t, v) \quad (8)$$

其中， c_i 是每个经验机制划分的IMF分量的总能量频域， r_n 是将多维攻击的网络的动态入侵所得到的IMF分量的数量划分为一个经验系数，以获得动态网络入侵信号的经验特征：

$$E_x = P_x B \times W(t, v) \div X(f) \times T_m(t) \quad (9)$$

在式(9)中， P_x 表示多维攻击中网络入侵信号的频率。

在经验特征的基础上，对上述网络进行了入侵信号的动态分布^[8]，该设计是具有多维攻击的网络的入侵动态结构。其中， l_{fjl} 是源代码元素的速率， g_{djl} 是伪码直接分布的速率， b_{vm} 反映了在恒定频率信号下频谱分布的快速干预规律，入侵信号的动态分布特性 s_{rff} 反映了每个频率下所有振幅的总似然性，计算 s_{rff} 如下：

$$s_{rff} = \frac{b_{vm} \times E_x}{Bg_{djl}} \times [l_{fjl} \times j_{gkp}] \quad (10)$$

在式(10)中， j_{gkp} 是傅立叶变换分析窗口的长度。将入侵干扰信号的总体特征表示为 v_{cnkk} ， b_{xnpp} 是信号位置的变化，则 v_{cnkk} 表示为

$$v_{cnkk} = \frac{a_{afh} \cdot X(f)}{s_{rff} \cdot r_{dijk} \cdot E_x} \cdot h_{gkl} b_{xnpp} W(t, v) \quad (11)$$

在式(11)中， h_{gkl} 和 r_{dijk} 是用于时域和频域信号的转换的参数。通过提取入侵干扰信号特征，为实现通信入侵信号检测奠定了基础。

1.3 优化网络入侵信号检测的抗干扰滤波方案

由于入侵信号的特征提取过程中，存在来自非入侵源的干扰信号，对网络入侵信号检测的信噪比较低。网络空间中含有大量的合法和非合法数据交织干扰^[9]，采用一种级联滤波器进行降噪滤波需要进行抗干扰滤波处理，可以有效地消除这些干扰信号，提高入侵信号的信噪比，减少其对入侵信号特征的影响，有助于提高入侵检测模型的准确性和鲁棒性，从而更容易、更准确地识别入侵信号。

级联滤波器系统对网络入侵信号进行干扰

滤波的传递函数可表述为

$$G(z) = \frac{(1 - \delta^{-2}) \cdot r_{dijk}}{v_{cnkk} \delta \cdot h_{gkl}} \quad (12)$$

式(12)中， δ 表示级联滤波器系统传递系数。

针对将网络入侵信号集中到网络入侵信号上的过程，需要对网络入侵检测方法进行改进。为了重建网络入侵机制的特征，假设网络入侵信号是线性且呈正态分布的，并通过降低对不同机制特征范围的控制限制来实现这一目标。同时，假设网络信号的中值较高，并通过自适应波束技术来^[10]确定网络入侵信号的特定时间尺度，该时间尺度是通过极值之间的时间间隔来确定的。通过计算入侵信号的振幅包络值，可以得出以下结论：

$$f = \langle f, d_{r0} \rangle d_{r0} / R_f G(z) \quad (13)$$

在式(13)中， $\langle f, d_{r0} \rangle d_{r0}$ 表示网络输入信号 f 在 d_{r0} 的上下包络线中的平均值， R_f 为内模函数。 Q 是使用调频来检测入侵信号，其检测结果表示为

$$Q = R_f G(z) \cdot Af \cdot a_{afh} \cdot X(f) \quad (14)$$

在综合分析的基础上，实现多维攻击下的网络入侵信号检测。

2 实验论证

为了验证本文设计的多维攻击下的网络入侵信号检测研究的有效性，进行了对比实验，设置利用哈里斯鹰算法优化卷积神经网络的入侵检测研究方法为方法1、基于信号的车载CAN入侵检测与防御框架设计方法为方法2，将其与本文方法对比，对三种方法的网络入侵信号检测结果进行分析。

2.1 实验准备

实验用Matlab7编程实现多维攻击中的开放式检测网络入侵算法。在多维攻击中，网络动态入侵的采样时间间隔为0.24 ms，网络传输频带为1~5 kHz，采样率为-15~5 dB，网络的动态特征被划分为500~3000个采样点之间的多维攻击。在上述环境和参数的基础上，多维攻击进行了一个网络动态项目，得到两个采样通道中具有多维攻击的网络传输信号模型。在采样测试之后，在多维攻击中使用网络传输信号来

检测入侵类型。对网络动态入侵信号进行经验模式特征分解,将网络入侵的动态信号分解为几个IMF分量的总和,获得多维攻击下的网络动态入侵检测结果,并将结果进行记录,得到三种方法对比结果,见表1。

2.2 对比实验

为了模拟真实环境下的信号,获取一段来自非入侵源的干扰信号,以验证本文方法采用级联滤波器对信号降噪滤波处理的效果。将噪声信号输入级联滤波器中,获取滤波后的信号结果,对比0.40 ms滤波前后的信号结果,选取信号幅度为实验指标来评估降噪效果,以观察信号的波形和幅值的变化。滤波前后的信号结果如图2所示。

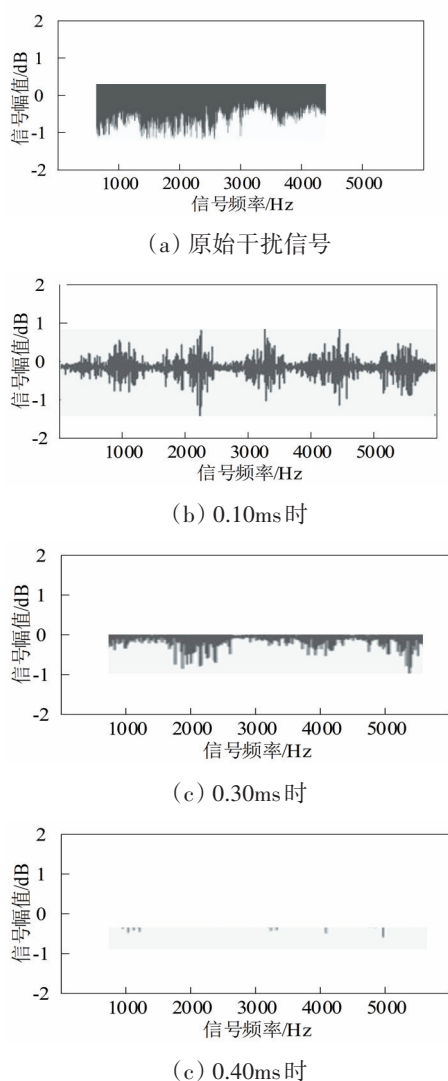


图2 不同时间下滤波前后的信号结果图

根据图2结果显示,通过滤波处理,信号的波形得到了改善并且幅值有所变化。通过时间的不断增加,经过时长0.40 ms的滤波处理后的信号显示更平滑的波形,成功地降低了噪声的干扰并改善了信号的质量。

为验证方法的有效性,选取入侵信号检测时间为实验指标,耗时越少,说明方法的效率越高。通过对比三种方法的实验结果,其结果见表1。

表1 三种方法结果对比

网络入侵信号/个	方法1网络入侵信号检测所用时间/ms	方法2网络入侵信号检测所用时间/ms	本文方法网络入侵信号检测所用时间/ms
500	1.43	2.85	0.31
1000	1.64	2.74	0.35
1500	1.65	2.94	0.53
2000	1.43	2.89	0.31
2500	1.85	2.85	0.30
3000	1.67	2.67	0.40

由表1可知,采用本文方法网络入侵信号检测效率最高,检测3000个网络入侵信号时,仅耗时0.40 ms,比方法1和方法2入侵检测时间分别减少1.27 ms和2.27 ms。实验结果表明,采用本文方法能够提高网络入侵信号检测效率。

3 结语

本文提出一种多维攻击下的网络入侵信号检测研究。首先进行了通信入侵干扰信号特征的提取,在此基础上构建网络入侵信号模型实现多维攻击下的网络入侵信号检测,最后通过对比实验进行验证。结果表明,采用本文方法能够实现信号的去噪,检测效率较高。

参考文献:

- [1] 曲洪权,王征一,盛智勇,等.基于梯度提升决策树算法的光纤入侵信号分类[J].激光与光电子学进展,2022,59(23):106-113.
- [2] 李响,缪祥华,张如雪,等.利用哈里斯鹰算法优化卷积神经网络的入侵检测研究[J].化工自动化及仪表,2023,50(4):513-520.
- [3] 方熙宇,韩培欣,刘霞,等.基于信号的车载CAN入侵检测与防御框架设计[J].信息与电脑(理论版),

- 2022, 34(17): 229-232.
- [4] 王俊彦, 衣然, 张斌, 等. 新型入侵检测技术在城市轨道交通信号系统中的应用研究[J]. 城市轨道交通研究, 2022, 25(7): 43-46.
- [5] 史诗洁, 刘正荣, 赵晖. 一类描述肿瘤入侵与具有信号依赖机制的趋化模型有界性与稳定性分析[J]. 数学物理学报, 2022, 42(2): 502-519.
- [6] 蒋岑, 吴迪. 隐蔽无线通信网络传输信息云存储密文检索[J]. 计算机仿真, 2021, 38(6): 125-128, 137.
- [7] 马海洲, 丁爱萍. 大数据时代无人船通信网络入侵信号攻击源识别模型[J]. 舰船科学技术, 2021, 43(24): 142-144.
- [8] 李睿, 向媛媛, 刘静. 异构无线网络环境下大规模入侵动态取证仿真[J]. 网络安全技术与应用, 2021, 43(10): 159-161.
- [9] 秦永俊, 唐增明. 改进的NetLinX开放网络动态入侵检测方法[J]. 西安工程大学学报, 2023, 31(4): 576-581.
- [10] 陈琳, 李勇, 王磊. 面向物联网的Sybil入侵防御系统设计与实现[J]. 计算机测量与控制, 2022, 25(3): 180-183.

Study on network intrusion signal detection under multidimensional attack

Yue Ru*

(Shanxi Politics and Law Institute for Administrators, Taiyuan 030012, China)

Abstract: In a multi-dimensional attack network operating environment, unknown intrusion interference signals may be encountered, and different types of intrusion behaviors have different characteristics, which increases the difficulty of intrusion signal detection. Therefore, a network intrusion signal detection method under multidimensional attacks is proposed. By deploying a uniform linear array distribution structure of near-field sources, the source intrusion signals in the network are captured and corresponding intrusion signal models are established. Extract corresponding interference signal features for different network intrusion behaviors. In order to improve the accuracy and reliability of detection, anti-interference filtering is applied to the signal to eliminate or reduce the impact of interference signals and improve the effectiveness of network intrusion signal detection. The comparative experimental results show that the research method has ideal denoising effect, and intrusion signal detection only takes 0.40ms, with high application efficiency.

Keywords: multidimensional attack; network intrusion signal detection; network security; detection method

文章编号: 1007-1423(2023)24-0056-05

DOI: 10.3969/j.issn.1007-1423.2023.24.009

基于 K-means 和 Dlib 的脸型分类算法设计与实现

李东阳^{1*}, 项辉宇¹, 冷崇杰¹, 张 勇²

(1. 北京工商大学人工智能学院, 北京 100048; 2. 中国特种设备检测研究院, 北京 100029)

摘要: 脸型分类是人脸分析领域中的重要任务, 对于人脸识别和美学评估等应用具有重要意义。针对传统脸型分类算法准确率低、分类速度慢的问题, 提出一种新的基于 Dlib 和 K-means 算法的人脸脸型分类算法。首先创建级联分类器和面部特征点预测器进行面部检测和关键点定位, 使用高斯滤波器对灰度图像进行平滑处理, 使用 K-means 算法对额头区域进行聚类, 通过测量额头线、颧骨线、下颚线、脸长线的长度和脸颊角度根据算法确定面部形状。根据实验集测试结果, 相对于现有算法而言, 此算法在脸型分类的准确性显著提升。

关键词: 脸型分类算法; K-means; Dlib

0 引言

近年来, 人脸脸型分类技术是计算机视觉领域中的一个重要问题, 在人脸识别、人脸美学^[1]、虚拟形象设计以及个性化推荐^[2]等领域中具有重要作用。例如, 在人脸识别前对数据库中图像进行脸型预分类, 可提高人脸识别的准确率和识别速度^[3]; 在个性化推荐领域, 通过对人脸形状进行分类, 可以帮助设计师和美学专家了解不同形状的人脸特征, 为顾客提供不同的发型、镜框以及配饰指导。因此, 脸型分类在多个领域中具有广泛的应用和重要作用。

通常, 脸型分类系统的主要组成部分为人脸图像输入、面部特征点的识别与定位、脸型特征分类模块, 而脸型分类系统的核心部分为面部特征点的识别与定位和脸型特征分类。在以往的研究中, 凌旭峰等^[4]提出一种基于轮廓线曲率特征的人脸分类及识别方法, 对人脸图像进行预处理之后得到人脸的轮廓线, 利用轮廓的曲率线对人脸进行分类和识别。王俊艳

等^[5]提出基于下颌轮廓线的人脸分类方法, 提取下颌轮廓特征对脸型进行分类。但是由于这两种方法轮廓曲率线提取精度不高, 导致脸型分类的准确率不高。张倩等^[6]采用主动形状模型算法提取人脸特征点和 K 近邻分类算法确定人脸脸型。刘双成等^[7]采用 ASM 算法得到人脸轮廓特征点, 计算特征点所围成区域面积和面型指数, 使用 K 近邻分类算法确定人脸脸型。王世洋等^[8]通过 ASM 算法定位得到的轮廓点数值化为一维向量, 并将其输入 DenseNet 网络进行训练, 实现脸型分类。以上三种方法都采用 ASM 算法, 但在实际应用中计算量大, 且不能解决 ASM 算法关键点定位不准造成的误差。

基于此, 本文提出一种新的基于 Dlib 和 K-means 算法的人脸脸型分类算法。预先将人脸脸型分为圆脸、正方脸、菱形脸、三角脸、长方形脸、椭圆脸, 使用集成回归树算法进行人脸关键点定位, 通过 K 均值聚类提取前额特征, 通过新的脸型分类算法确定人脸形状, 实验结

收稿日期: 2023-08-14 修稿日期: 2023-09-12

作者简介: *通信作者: 李东阳(1998—), 男, 山东威海人, 硕士研究生, 主要研究方向为智能设计与检测技术, E-mail: lidongyang1225@163.com; 项辉宇(1966—), 男, 山东青岛人, 博士, 教授, 主要研究方向为智能设计与检测技术; 冷崇杰(1965—), 女, 山东烟台人, 硕士, 副教授, 研究方向为智能检测技术; 张勇(1978—), 男, 山东济宁人, 博士, 高级工程师, 主要研究方向为游乐设施检测技术

果表明，此算法能够有效提取人脸面部的形状特征，有效地提高人脸脸型分类的准确率。

1 算法整体框架

本文提出的算法框架如图1所示，首先读取图像并进行预处理，将输入图像转换为灰度图像，使用高斯滤波器对灰度图像进行平滑处理，使用级联分类器检测人脸，使用Dlib算法定位人脸关键点，提取额头区域并对该区域进行K-means聚类，提取面部特征点中的额头线、颧骨线、下颌线、脸长线，根据其长度及角度关系设计算法确定人脸脸型。

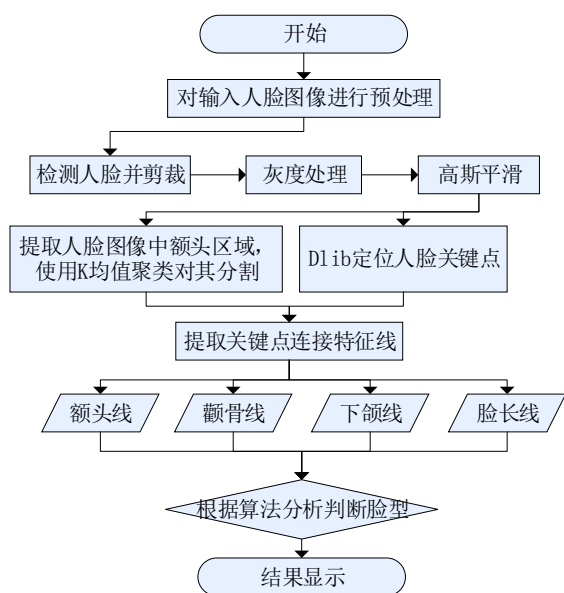


图1 整体框架图

2 算法实现

2.1 K-means 算法

为获取人脸额头区域特征信息，使用K-means算法对人脸图像的额头区域进行分割，分割为头发和皮肤两个区域，并且用不同的颜色表示。K-means算法是一种经典的聚类算法，用于将一个数据集划分为 K 个不同的簇(clusters)。每个簇由其内部的数据点组成，这些数据点在某种度量上彼此相似，而不同簇之间的数据点相似度较低。其优化目标是 minimized 簇内数据点与其所属簇中心之间的平方距离的总和，即最小化簇内平方误差和(SSE)。

具体实现方法如下：首先从原始图像中提取出前额区域，并将其转换为二维数组。将前额区域的像素作为输入数据，在聚类过程中选择 $K = 2$ ，即将前额区域的像素划分为两个簇。这样可以将头发像素和皮肤像素分别归为不同的簇。通过计算每个像素与其所属簇中心之间的距离，可以确定每个像素所属的簇，数据点到聚类中心点的马氏距离可表示为

$$D_M(X, U) = \sqrt{(X - U)^T \sum^{-1} (X - U)} \quad (1)$$

对于每个簇 C_k ，更新其聚类中心 μ_k 为该簇中所有数据点 x_i 的均值，表示为

$$\mu_k = \frac{1}{|C_k|} \sum_{x_i \in C_k} x_i \quad (2)$$

其中， $|C_k|$ 表示簇 C_k 中的数据点数量， $\sum_{x_i \in C_k} x_i$ 表示对簇 C_k 中所有数据点求和。

为了展示分割结果，使用不同的颜色来表示头发和皮肤区域。例如，使用黑色表示头发区域，使用白色表示皮肤区域。通过将分割结果以不同的颜色展示出来，可以直观地观察到头发和皮肤区域的分离效果。K-means算法处理前后对比图如图2所示。



图2 K-means处理对比图

当计算额头长度时，代码的主要思路是通过检测额头边缘(即头发的边缘)来确定额头的范围。使用K-means算法处理后的前额区域图像可以直观地看出头发和皮肤的分离效果，接下来通过计算图像的宽度和高度，获取额头的中点坐标，用于定位额头的大致位置，然后从额头的中点向左右两边遍历，代码使用一个循环来逐个检查额头区域的像素值，在每次循环中，代码检查当前像素点的颜色是否与额头中点像素点的颜色相同。在遍历过程中，如果检测到像素值发生变化，说明找到了额头的边

缘。记录下变化点的横坐标作为额头的边缘点,根据左侧点和右侧点的坐标,可以计算出额头的长度。

2.2 Dlib人脸关键点检测

Dlib是一个功能强大的C++开源库,提供了许多用于机器学习和计算机视觉的算法和工具。其中Dlib库中的人脸检测算法基于HOG(方向梯度直方图)特征和级联分类器的方法,具有高效和准确的特点。HOG是一种用于描述图像局部纹理和梯度信息的特征表示方法。在人脸检测中,使用HOG特征来描述人脸区域的纹理和边缘信息,能够有效地捕捉到人脸的形状和轮廓。级联分类器是由多个弱分类器组成的分类器,每个弱分类器只能对某个特定类型的人脸区域进行简单的判断。级联分类器通过级联的方式,将多个弱分类器组合起来,以实现高效的人脸检测,其总体公式可以表示为

$$y = C_1(x) * C_2(x) * \dots * C_n(x) \quad (3)$$

其中, x 是输入样本的特征向量, $C_1, C_2, \dots, C_n$ 表示级联分类器的各个阶段(级)。

shape_predictor_68_face_landmarks是Dlib库中用于人脸关键点检测的模型,人脸关键点是人脸具有特殊意义的位置,其中脸颊等部分的关键点的准确定位对于人脸脸型分析的应用非常重要,此模型旨在准确地预测人脸图像中关键点的位置,以达到对人脸脸型特征提取的目的,使用该模型对人脸图像进行关键点定位,如图3所示。

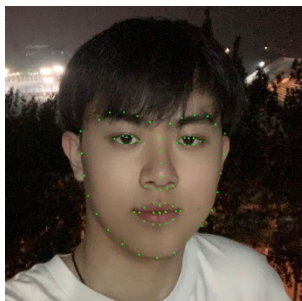


图3 人脸关键点定位示意图

2.3 脸型分类算法设计

根据以上工作,已经通过K-means算法提取

额头区域关键特征,包括额头中心点以及额头左右边缘的坐标点,连接额头左右边缘坐标点,并命名该直线Line1为额头线;接着,将Dlib中的68个关键点坐标定位出的左右脸颧骨处的关键点,即将第二个关键点和第十六个关键点连接,并命名该直线Line2为颧骨线;然后将左右脸颊处的关键点连接,即第四个关键点和第十四个关键点,并命名该直线Line3为下颚线;最后将额头中点与下巴中心关键点即第九个关键点连接,并命名该直线Line4为脸长线,如图4所示。

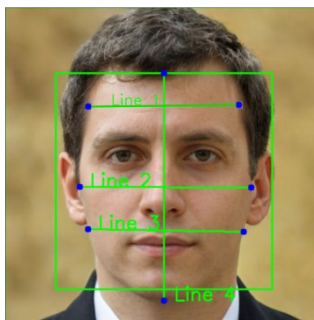


图4 关键点连接线示意图

根据上述四条连接线的长度以及脸颊线角度关系设计脸型分类算法,首先通过定位出的关键点坐标计算各连线的长度,为保证实际运用中的准确性,使用欧式距离计算线段长度,其公式如下:

$$D = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (4)$$

计算两组线段的标准差,第一组计算的是Line1、Line2、Line3三条线段的标准差SD1,第二组计算Line2、Line4的标准差SD2,标准差计算公式如下:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} \quad (5)$$

取人脸图像中第三到第六个关键点坐标,分别计算第三个关键点到第五个关键点和第四个关键点到第六个关键点这两个向量与x轴的夹角,取两者之差,使用abs()函数得到绝对值,并使用degrees()函数将角度转换为度数,得到两线段所形成夹角,并将其转换为角度值Angle,此角度值与脸颊角度密切相关,后续通过此角

度值、各连线长度、线段间标准差设计脸型分类模型，将脸型预先分为圆脸、正方脸、菱形脸、三角脸、长方形脸、椭圆脸，具体模型图如图5所示。

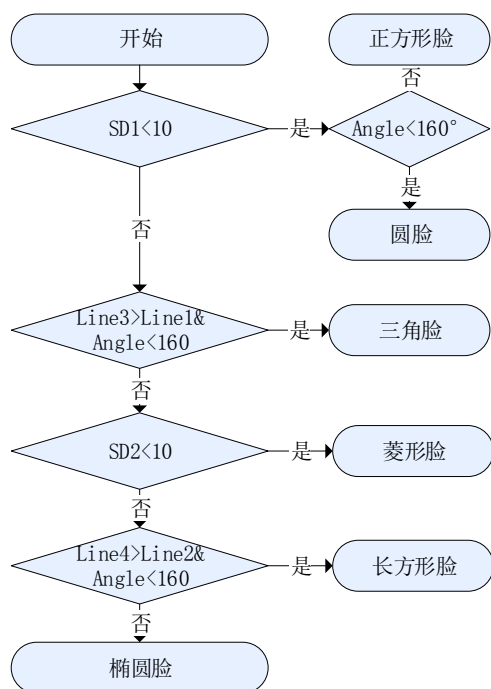


图5 脸型分类模型图

3 实验结果

本文采用LFW(labeled faces in the wild)人脸库中600张中性表情，温和光照的正脸人像照片作为实验集，采用20人的投票机制将图片分为圆脸、正方脸、菱形脸、三角脸、椭圆脸、长方形脸六类，分类准确性统计结果见表1。

表1 实验结果

脸型	样本数	正确分类数	错误分类数	正确率/%
圆脸	136	130	6	95.59
正方脸	65	64	1	98.46
菱形脸	89	85	4	95.51
三角脸	45	43	2	95.56
椭圆脸	145	143	2	98.62
长方形脸	120	115	5	95.83

从表1可以看出，此脸型分类算法具有较高的正确率，可以有效地解决实际应用中的脸型分类问题。

4 结语

针对目前脸型分类技术准确率较低，检测速度较慢的问题，本文以计算机视觉技术为基础，基于K-means和Dlib设计了一种新的脸型分类算法，该算法充分利用脸部特征，使用面部特征点预测器进行面部检测和关键点定位，使用K-means算法对额头区域进行聚类提取特征，测量额头线、颧骨线、下颚线、脸长线的长度和脸颊角度，根据特征点连线长度以及角度，依据脸型分类算法确定面部形状，分类方法科学高效。实验结果表明，该算法具有较高的正确率，具有实际应用意义。

参考文献：

- [1] 揭萍,彭佳林.基于脸型识别的形象辅助系统设计与实现[J].信息与电脑(理论版),2017(7):140-142.
- [2] 郭佳玥,韩豪杰,石金泽,等.基于脸型及肤色特征的饰品推荐系统[J].电子商务,2019(12):55-57.
- [3] 赵薇,汪增福.用于大库人脸识别的脸型分类研究[J].电子技术,2009,36(11):77-79,68.
- [4] 凌旭峰,杨杰,杨勇.基于轮廓线曲率特征的人脸分类及识别[J].红外与激光工程,1999(4):37-39.
- [5] 王俊艳,苏光大.基于下颌轮廓线的人脸分类方法[J].红外与激光工程,2004(2):159-163.
- [6] 张倩,丁友东,蓝建梁,等.基于ASM和K近邻算法的人脸脸型分类[J].计算机工程,2011,37(11):212-214,217.
- [7] 刘双成,蔡晓东,张力,等.基于主动形状模型和K近邻算法的人脸脸型分类[J].桂林电子科技大学学报,2014,34(6):479-483.
- [8] 王世洋,陈旺,武豪放.基于DenseNet的人脸脸型分类研究[J].软件导刊,2020,19(9):210-214.

Design and implementation of face shape classification algorithm based on K-means and Dlib

Li Dongyang^{1*}, Xiang Huiyu¹, Leng Chongjie¹, Zhang Yong²

(1. School of Artificial Intelligence, Beijing Technology and Business university, Beijing 100048, China;

2. China Special Equipment Inspection and Research Institute, Beijing 100029, China)

Abstract: Face shape classification is an important task in the field of facial analysis, with significant implications for applications such as face recognition and aesthetic assessment. In response to the low accuracy and slow classification speed of traditional face shape classification algorithms, a novel face shape classification algorithm based on Dlib and K-means is proposed. Firstly, a cascade classifier and facial landmark predictor are created for face detection and key point localization. The grayscale image is smoothed using a Gaussian filter, and the forehead region is clustered using the K-means algorithm. Facial shape is determined based on the lengths and angles of the forehead line, cheekbone line, jawline, and face length line measured by the algorithm. Experimental results on the test dataset demonstrate that this algorithm significantly improves the accuracy of face shape classification compared to existing algorithms.

Keywords: face shape classification algorithm; K-means; Dlib

(上接第 50 页)

A data mining algorithm based on clustering matrix and pruning strategy

Xu Kele^{*}

(School of Information Engineering, Xianyang Vocational Technical College, Xianyang 712000, China)

Abstract: In the field of medicine, the Apriori association rules algorithm has wide applications in areas such as chronic diseases. However, this algorithm has some issues, such as requiring multiple scans of the medical case database, generating large amounts of data sets, and producing redundant rules. Aiming at the related problems, an association rule data mining algorithm combining clustering matrix and pruning strategy is proposed. This algorithm adds constraint conditions and adopts the clustering matrix approach to compress the storage of transaction databases. Simultaneously, it introduces pre-pruning and post-pruning strategies for algorithm optimization. Experimental results have shown that this optimization algorithm significantly reduces the number of scans of the database, decreases the number of candidate item sets, shortens the execution time of the algorithm, and greatly improves operational efficiency.

Keywords: data mining; association rules; Apriori association rule algorithm; candidate item set; clustering matrix; pruning strategy

文章编号: 1007-1423(2023)24-0061-03

DOI: 10.3969/j.issn.1007-1423.2023.24.010

基于多尺度特征的遥感图像检测方法研究

陈 芳*

(湖南信息学院计算机科学与工程学院, 长沙 410151)

摘要: 遥感图像中的显著性目标检测在遥感信息处理领域具有重要研究价值。当前, 自然场景图像的显著性目标检测技术已经取得了显著的进展。可是, 遥感图像的目标尺度多样、目标种类丰富以及复杂的环境背景等特点, 使得直接在遥感图像中应用自然场景图像的显著性目标检测方法变得具有挑战性。针对这些局限性, 提出了一种密集连接多尺度特征网络模型, 该模型充分利用神经网络的强大的特征提取能力, 可以实现高效且准确的检测, 并为遥感图像显著性目标检测的实际应用提供技术支持。

关键词: 遥感图像; 多尺度; DCMM

0 引言

遥感图像是在宏观层面上反映各种地表地物特征的综合图像。遥感图像是成像设备在远距离条件下对目标或区域进行勘测所形成的图像, 它具有成像范围大、受环境限制少等特点。近些年, 随着遥感勘测技术的快速发展, 高分辨率遥感卫星发射数量的不断增加, 极大便利了遥感技术应用于地面勘察, 并为遥感图像处理应用提供数据支持。

与定位和框定特定对象的传统对象检测任务不同, 显著性目标检测旨在检测出一幅图像中最显著的目标或区域, 并将其分割为前景图像和背景图像。传统方法使用特征修改傅里叶变换、支持向量机、频谱变换等技术来提取遥感图像中的显著特征。但是, 不同的遥感场景对应的背景特征差异很大, 比如厚云遮挡, 复杂背景等干扰场景时, 使用遥感图像的传统处理技术很难满足显著性目标检测的性能要求。如何从大量遥感图像数据中快速而精确地筛选出显著目标, 是一个待解决的问题。

在真实场景中, 直接处理遥感图像时, 由于缺乏先验知识, 而导致检测性能下降。而显

著性目标检测方法可以快速提取出遥感图像中最显眼的区域或物体。图像可以作为预测因子, 降低后续图像处理的复杂度, 可以为遥感图像应用提供重要的初始信息, 在遥感图像处理图像融合、图像/视频压缩、图像质量评估、场景分类和变化检测等许多研究领域得到了成功的应用, 因此受到了广泛关注。

综上所述, 遥感图像中显著性目标检测不但可以促进多场景遥感数据的应用, 而且可以推动计算机视觉技术在遥感领域的发展, 具有重要的现实意义和实用价值。

1 显著性目标传统检测方法

1.1 基于区域对比的方法

由于图像中特征区域的差异, 基于区域比较的方法提取并比较图像不同区域的特征信息, 以计算显著性。Scharfenberger 等^[1]提出了一种基于纹理特征的局部对比度检测方法, 可以从不同角度提取旋转纹理不变特征来表示每个区域; Tong 等^[2]提出了一种基于组合多种特征信息的显著性目标检测方法, 该方法首先通过提取低级特征获得初始显著性特征, 然后利用重

收稿日期: 2023-07-26 修稿日期: 2023-12-08

基金项目: 2022 年度湖南省教育厅科学研究项目(22C1162)

作者简介: *通信作者: 陈芳(1980—), 女, 湖南长沙人, 副教授, 硕士, 研究方向为系统集成、网络安全与维护、网页图形图像, E-mail: 419695976@qq.com

建误差计算显著性信息附加显著性,最后得到最终显著性图像。基于区域比较的方法可以有效地从图像中提取不同的区域特征,并且可以在常规图像场景中得到更好的效果,但很难应用于云遮挡等干扰较大的场景。

1.2 基于图像块的内在线索的方法

因为基于单位像素构建内在线索的方法性能较低,部分工作将对比度计算扩展到了块级别。因此,Achanta等^[3]提出了一种用于显著性检测的频率调谐模型。Li等^[4]将中心-周围对比度表示为对成本敏感的最大边距分类问题。中心图像块被标记为阳性样本,而周围的图像块都被用作阴性样本。然后,根据中心图像块与周围图像块的可分离性来确定中心图像块的显著性。Lin等^[5]提出了一种通过测量图像块特征与平均图像块的距离来测量图像块的显著性的方法。一般来说,该算法基于使用内在线索的图像像素或块,但其存在如下缺点:首先,高对比度通常出现在对象的边缘,而不是整个对象区域;其次,当使用图像块作为计算单位时,它们对物体突出边缘的预测有更大的影响。

1.3 基于频域变化的方法

基于频域变化的检测方法通过捕获背景和突出物体之间的特征差异来进行显著性检测。Achanta等^[3]将图像分为低频域和高频域,然后对图像进行高斯平滑和均值平滑检测。Arya等^[6]提出了一种基于全局和局部对比的检测方法,首先从图像中提取局部特征,然后进行频域变化并计算全局显著性特征,最终获得显著性基于局部和全局特征加权的区域。虽然基于变分频域的方法取得了一定的性能提升,但对于处理背景环境干扰较大的图像仍然存在困难。

2 基于多尺度特征的显著目标检测方法

在遥感场景中,显著目标在形状、尺度和位置上有极大的变化。如果使用简单的解码器提取其特征,有可能编码器上采样的分辨率策略无法处理复杂的变化。密集级联多尺度模块(densely cascaded multiscale modules, DCMM)通过多个不同空洞率的空洞卷积来嵌入多尺度周边相关信息,捕获其中的目标特征,然后增强

这些特征并混合特征提取器的低级细节特征以产生最终的分割结果。

2.1 网络结构设计

给出一张遥感图像,DCMFNet的目的是检测图像中的突出物体,并将它们分组为二进制显著图。如图1所示,DCMFNet由特征编码和特征解码两部分组成,有三个主要结构:主干模型、密集级联多尺度模块(DCMM)和多级跳跃连接结构(MSCS)。在编码器阶段,主干模型是用来从输入遥感图像中提取基本语义特征;DCMM是用于从显著目标中提取多尺度融合信息,以增强模型的尺度感知能力。在解码器阶段,MSCS用于结合不同层次的多尺度特征,并结合编码和解码特征,以获得更好的显著特征。

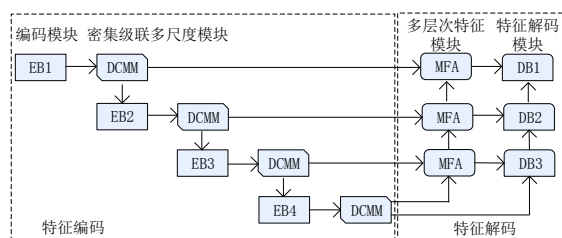


图1 DCMFNet结构图

2.2 密集级联多尺度模块

在遥感图像中,突出性物体的尺度变化很大,这在很大程度上影响了检测模型的性能。当前的基于多尺度的方法一般是直接整合卷积层和池化层所生成的特征图。虽然提取了不同尺度的特征,但特征的语义信息仍然一致,难以更好地捕获多尺度信息。针对以上问题,提出了如图2所示的DCMM。

DCMM不同于现有的多尺度方法,该模块通过渐进连接的方式提取各种不同尺度的融合特征信息,可以有效提高各种尺度之间的信息交互。输入特征经过不同池化率的池化层后,得到不同尺度的特征信息。接着,将三层池化后的特征进一步用卷积模块进行增强,提取语义特征,然后进行上采样,与两层池化后的特征进行拼接组合,得到最终的输出特征。这种级联拼接方法可以将多尺度信息相互融合,更好地显示出显著特征。

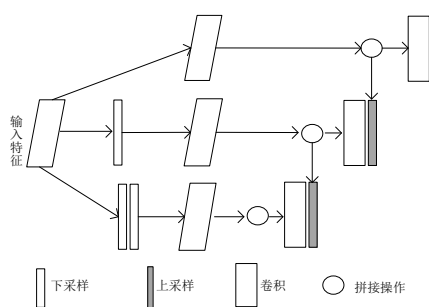


图2 DCMM结构图

2.3 多层次跳跃级联结构

在CNN中，每一层都包括显著的特征信息。所以，MSCS旨在聚合不同层之间的信息以改进特征提取。MSCS内置于DCMFNet的解码结构中。经过融合DCMM输出的多尺度信息，可以加强解码器中的特征提取。

如图1所示，DCMFNet中的MSCS从下至上依次融合了不同层的多尺度特征。当DCMM输出多尺度特征 M_4^8 时，MSCS融合多尺度特征如式(1)所示：

$$F^8 = M_4^8 \& \mu \left(\text{conv} \left(\text{up} \left(I_4^{8+1} \right) \right) \right) \quad (1)$$

式中， F^8 为融合后的输出特征， $\text{up}()$ 为上采样计算， $\text{conv}()$ 代表一个卷积层， $\&$ 为矩阵元素的相乘计算， μ 为Sigmoid函数。

3 结语

遥感图像中显著性目标对象尺度变化较大，尤其在背景环境复杂的场景下，严重影响显著性目标检测方法性能。针对这一问题，提出一种密集级联多尺度特征网络。在该方法中，采

用密集级联多尺度模块有选择地提取显著性目标的多尺度信息，进而实现不同尺度的信息相互融合，可以加强多尺度特征提取。此外，多层次跳跃连接结构能够优化不同尺度的多层次特征，有利于加强多尺度信息与不同层次信息的互相融合。DCMFNet可以较好地处理复杂场景图像，且适时地检测出目标对象的显著性特征，提升了检测结构完整性。

参考文献：

- [1] SCHARFENBERGER C, WONG A, CLAUSI D A. Structure-guided statistical textural distinctiveness for salient region detection in natural images [J]. IEEE Transactions on Image Processing, 2018, 24(1): 457-470.
- [2] TONG N, LU H, ZHANG Y, et al. Salient object detection via global and local cues [J]. Pattern Recognition, 2019, 48(10): 3258-3267.
- [3] ACHANTA R, HEMAMI S, ESTRADA F, et al. Frequency-tuned salient region detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019: 1597-1604.
- [4] LI F, GLEICHER M. Region enhanced scale-invariant saliency detection [C] // Proceedings of the IEEE International Conference on Multimedia and Photonics, 2019: 1477-1480.
- [5] LIN M, CHEN Q, YAN S. Network in network [C] // Proceedings of the International Conference on Learning Representations (ICLR), 2014.
- [6] ARYA R, SINGH N, AGRAWAL R K. A novel hybrid approach for salient object detection using local and global saliency in frequency domain [J]. Multimedia Tools and Applications, 2018, 75(14): 8267-8287.

Research on remote sensing image detection method based on multiscale features

Chen Fang*

(School of Computer Science and Engineering, Hunan Institute of Information Technology, Changsha 410151, China)

Abstract: Significant object detection in remote sensing images has important research value in the field of remote sensing information processing. At present, the technology of salient object detection in natural scene images has made significant progress. However, the characteristics of remote sensing images, such as diverse target scales, rich target types and complex environmental background, make it challenging to directly apply the salient target detection method of natural scene images in remote sensing images. In response to these limitations, this article proposes a densely connected multi-scale feature network model that fully utilizes the powerful feature extraction ability of neural networks, enabling efficient and accurate detection, and providing technical support for the practical application of remote sensing image saliency target detection.

Keywords: remote sensing images; multi scale; DCMM

实践与经验

文章编号: 1007-1423(2023)24-0064-06

DOI: 10.3969/j.issn.1007-1423.2023.24.011

基于藏成药数据的知识图谱可视化构建

樊 淼^{1*}, 高 屹^{1,2}

(1. 西藏民族大学信息工程学院, 咸阳 712082; 2. 西藏光信息处理与可视化技术重点实验室, 咸阳 712082)

摘要: 藏药是中华民族传统医药学的重要组成部分, 拥有悠久的历史, 它汇集了藏族人民多年抗击传染病的经验, 为青藏高原及周围区域民族群众的健康发展和繁衍作出了无可消灭的努力, 为中华民族的发展作出了巨大的奉献。为了有效地存储、管理和使用各种藏药数据, 提出了一种基于知识图谱的存储结构和可视化技术, 以实现藏药相关属性的可视化, 从而提升藏药知识的整合度和数据的关联性。通过使用知识图谱技术, 可以更好地发掘藏成药数据中的潜在价值。

关键词: 藏成药; 知识图谱; 数据可视化

0 引言

藏药是中华民族文化的组成部分, 拥有悠久的历史, 其独特的制剂工艺和卓越的疗效深受藏族人民和全国各族人民的青睐, 在国内外享有极高的声誉。在新冠肺炎的预防和治疗中, 藏成药所起到的作用与研究潜力, 受到了高度评价。

藏成药是藏药的重要流通方式。藏成药的相关数据大多数以半结构化或者非结构化的形式存在。这影响了藏成药的推广和相关应用。知识图谱作为大数据时代在大规模数据管理的新兴技术, 可以有效地将藏成药的相关规律以及隐藏价值挖掘出来, 以更直观的方式展示, 有效地实现了知识资源的整合, 使数据的服务能力大大提升。因此, 本文就藏成药的知识图谱可视化进行研究, 以便更好地利用藏成药现有的数据。

1 研究现状

中华民族优秀的历史文化源远流长, 影响

深远。藏医药是一种拥有悠久历史的文化遗产, 它是我国医疗体系中不可或缺的一部分, 具有深厚的历史底蕴和光荣传统。经过数千年的积累, 藏药拥有丰富的古籍文献。

藏药的相关数据较为丰富, 但与目前信息技术的数据要求有很大不同。有很大一部分的藏药数据以非结构化的文本形式存在, 数据噪声较大。一定程度上, 数据形式阻碍了藏药的推广和应用。如何从大量数据中发现研究思路, 根据藏药信息的总结规律, 深入探索藏药知识的隐藏价值, 以直观简洁的形式呈现, 让藏药资源得到更加充分的利用, 为人类健康做出更大贡献, 是值得深入研究的方向。

根据相关文献记载, 中医学中的知识图谱和可视化技术已经被广泛应用, 也有学者在研究基于汉语文本的知识图谱, 但是在藏药方面, 这些研究仍然相对较少。本文就知识图谱技术, 基于目前已有的互联网藏药数据, 通过对网站数据进行爬取, 形成藏药的数据集, 并通过该数据集达到实现藏成药知识图谱的效果。这为

收稿日期: 2023-08-17 修稿日期: 2023-12-06

基金项目: 西藏自治区自然科学基金项目(XZ202001ZR0043G)

作者简介: *通信作者: 樊淼(2000—), 男, 山西运城人, 在读研究生, 研究方向为大数据与知识图谱可视化, E-mail: 1608765146@qq.com; 高屹(1981—), 男, 山东烟台人, 硕士, 副教授, 研究方向为数据挖掘、分析及可视化应用、西藏文化产业与互联网融合发展研究等

藏药领域的潜在价值挖掘提供了帮助,也很好地呈现了藏成药的相关属性。

周雪阳等^[1]利用知识图谱技术和可视化技术,对多个开放网站的中成药数据进行了深入的分析,从而发掘出了其中蕴藏的巨大价值。李贡宇等^[2]通过采用MDL ISIS/Base数据库管理系统搭建了常用藏药数据库,收集了部分藏药数据,为藏药的数据挖掘提供了较好的数据基础。于彤等^[3]深入研究中药知识资源,并结合这一面向中医药领域的知识图谱,为相关研究人员提供更加全面、更有效的知识服务。

2 相关技术

知识图谱是一种新兴的技术,它将图像学、应用数学和信息技术可视化研究等学科专业的理论知识和办法融合在一起,通过计算学的引文解析和共现分析等方式,以图像可视化的形式展现学科专业的核心理论、发展过程和总体认识框架,从而帮助人们更好地理解 and 掌握专业知识。利用数据分析发现、数据处理、认识计算和图像绘画等技术,它可以将复杂的知识表达出来,深入探究知识领域的发展规律,为学科的发展提供有力的支撑和决策服务。

知识图谱以规模庞大的知识存储库为支撑,它的构造涉及到数据库、人工智能、自然语言处理等多个领域的计算机技术。知识图谱目前的相关应用通常是对某领域的发展趋势进行可视化预测,对前沿技术的动态发展进行预测,还与语义技术结合进行近似推理,从而进一步提升搜索引擎的智能化服务^[4]。

3 数据收集与数据处理

传统的医药体系有较多的实体概念,如药方、疾病等,它们相互之间都存在密不可分的联系,而研究这些实体之间的关系是医学领域的重点内容。伴随着现代信息技术的应用发展,有部分藏成药的基本属性、组成成分、治疗功能等已通过互联网的方式将数据进行存储、应用,这为藏药的数字化提供了技术支持,而数据挖掘、人工智能等技术的概念也被应用在藏药的数字化实现,并提供了强有力的理论支持和技术支撑。

(1) 藏医药网(www.tibetmd.com):藏医药网(如图1所示)收集了大量的藏药材、藏成药信息,包括药物的性状、主治功能等。该网站还建立了关于藏医的论坛,为用户提供了很好的借鉴和交互体验。

藏药药酒	
策划: 普智 编辑: 扎巴 日期: 2015-05-07 16:01:49 人气: 911 加入收藏 评论: 0	
红块糖药酒	
【藏文】	བྱ་མཚན་
【拉丁】	bur chang
【别名】	
【组成】	青稞、蕨蓼、小麦、红块糖。
【制法】	青稞、小麦、蕨蓼三味药,加水煎煮后,放置稍温,下曲发酵;红块糖粉碎用水溶解,慢慢兑入以上发酵液中,放置热处,密封贮存6-8天,即得。
【性味】	味甘、酸。
【主治】	祛风活血,滋补强壮。用于心隆、肺隆、肾隆、骨隆等一切隆病,老年滋补壮体。
【用法】	温服。
【贮藏】	密闭,置阴凉干燥处。

图1 藏医药网的相关数据

(2) 中国藏药网(<http://w.zyao.cn>):中国藏药网(如图2所示)为用户提供了药物的购买服务,包含了药物的成分、注意事项、用法用量等信息。有效地推动了藏药的线上销售,为大众提供了便捷的服务平台,为藏药的数据化提供了数据支持。

商品详情	用户评论 (0)
【商品名称】十八味杜鹃丸 【成 份】烈香杜鹃、草果、诃子、檀香、毛诃子、降香、余甘子、山胡叶、石灰华、藏西草、红花、紫草茸、肉豆蔻、秦艽花、丁香、甘草膏、豆蔻、沉香。 【性 状】本品为棕红色水丸,气香,味微苦、微甘。 【功能主治】祛风通络,活血。用于“白脉病”引起的四肢麻木、震颤、肌肉萎缩、筋脉拘挛、口眼歪斜等症。 【规 格】每丸重0.3克。 【用法用量】口服。一次2丸。一日3次,研碎后服用。 【注意事项】详见说明书。 【贮 藏】密闭,置阴凉干燥处。 【包 装】PVC板,12丸/板×2板/盒。 【有效期】42个月。 【批准文号】国药准字Z54020023。 【生产厂家】西藏甘露藏药股份有限公司。 本药来源于我国藏医古验方,现收录于《藏药标准》,经千百年的临床实践证明:其具有祛风通络、活血等功效,对现代医学中的脑中风后遗症及气血不利引起的四肢麻木、震颤、肌肉萎缩、筋脉拘挛、口眼歪斜等症等有特殊的治疗作用。本品在总结历代藏医经典验方的基础上,采用纯天然、稀有名贵藏药材精制而成。	

图2 中国藏药网相关数据

3.1 数据收集

本文主要是对藏成药的数据进行研究,首先搜索藏成药的相关词条,然后利用Python爬虫技术从以上开源网站对相关词条信息进行爬取,网站内容的解析由Python中的urllib分析包完成,具体爬取过程如图3所示。从以上开源网站所获取的数据有:名称、组成成分、功能和主治、性状、用法等。以十八味杜鹃丸为例,需要爬取的信息见表1。

ཐུ་གྲིག་ལོགས་ཁྱེད་དཀོན་པའི་མེན་, 我们可以通过分词更好地理解它们之间的语义关系, 从而更好地表达它们之间的内涵和意义。通过比较藏医实体的词库, 我们发现了一些有规律的实体, 例如འཛམ་གླིང་ཆེན་པོའ་བླ་, 它们是藏医药的重要组成部分。实体识别示例如图7所示。



图7 实体识别示例

3.2.3 关系抽取

实体关系抽取是一种自然语言处理技术, 它可以帮助我们从句子中抽取两个实物相互之间的关系类别。这种技术在数据搜索、认知图像和自动问答控制系统等方面都有广阔的应用前景。本文将重点介绍实体关系抽取的方法, 它利用语言学和自然语言处理学的基本知识, 通过人工构造特征关系词或规律, 将这些信息存储起来, 以便更好地理解句子中的实体关系。通过将规律与预处理后的非结构文本进行匹配, 我们可以从实体关系中抽取出三元关系组。

通过远程监督方法, 藏药实体关系抽取可以实现与实体识别类似的步骤, 即利用藏语中格助词的规律, 如属格、施格等, 提取出一定数量的实物关联, 并将其保存在数据库系统中。实验证明, 当两个实物同时出现在数据库中时, 它们之间的关系就会被认定为这类关联, 而且这种方法在藏医药等固定领域的文本上表现出色^[6]。

3.3 数据存储

知识图像的存储方式多种多样, 从 disk 或 mainmemory 的原始信息到关系型信息, 再到非关系型信息, 都可以提供有效的存储结构。本文采用非关系型数据库中的图形数据库, 以更好地发现实物彼此间的关联性, 从而更有效地保存知识图谱。将知识图谱保存在图数据库中, 就是将它的实物反映到图数据库中的结点, 而结点彼此间的关联性则可以通过链接结点的边来表现出来。当前, 许多主流的图形数据库系统都已经问世, 其中包括 Neo4j、GraphEngine、

gstore等, 这些数据库系统均由国家neo technology 机构开发, 并由北京大学提供。Neo4j 图数据库系统是一款高性能的图引擎, 它支持分布式存储, 并且与 Java 平台兼容性良好, 为后续工程开发提供了极大的便利。因此, 我们采用 Neo4j 图数据库系统来储存知识图谱, 以期获得更好的效果^[7]。

4 藏成药知识图谱的构建

本小节主要是对藏成药知识图谱的构建进行研究, 主要从两部分来构建藏成药知识图谱: ①从藏成药知识图谱的概念模式层设计模块。根据自底向上的方法, 构建藏成药知识图谱的概念模式。首先, 可以通过从藏药的文献和藏医们的经验总结中分析, 确定藏成药知识图谱的核心部分, 逐层扩展, 从而完成藏成药知识图谱的概念层设计。②藏成药知识图谱的生成模块^[8]。根据所确定的藏成药知识图谱的概念模式, 从数据集中实体进行填充, 完成藏成药的数据提取、数据融合和数据存储相关处理。

以藏成药的商品名为实体结点, 将其基础特征、安全性和经济性作为知识图谱的一级特征(如图8), 其中, 实体的主要规格、功能和主治、用法用量、贮藏等特性是从属基础特征的两级结点(如图9); 用药禁忌、疾病、注意事项、有效期等特性是从属安全的两级结点; 商品价格和成分来源则是从属经济性的两级结点。使用 Python 中的 Flask 架构, 我们能够轻松搭建一个可视化的藏成药知识图谱平台。与其他同类别的架构比较, Flask 架构更加简单易用, 安全性更高, 更容易实现。Ajax (Asynchronous JavaScript and XML) 是一项高效的互动式网页开发技能, 它能够让用户在不重复添加一个页面的前提下, 快速、动态地访问和更新网页内容, 从而提升网站的交互性能^[9]。Ajax 技术的异步交互功能有助于实现前端页面的实时更新, 大大提升了用户体验, 极大地改善了人机交互的效果。

平台使用 Neo4j 数据库对藏成药的相关数据进行存储, 将药物的三大类属性以不同颜色展示, 不同药物之间可以以相同属性为标准相互联系起来, 形成藏成药知识图谱体系, 再通过 Echarts 可视化技术对图谱体系进行展示。通过平台的交互功能, 用户可以查看更多相关数据信息, 并可以通过点击知识图谱模块中的实

体对象、对象属性和关联关系, 来实现对隐藏信息的有效控制。

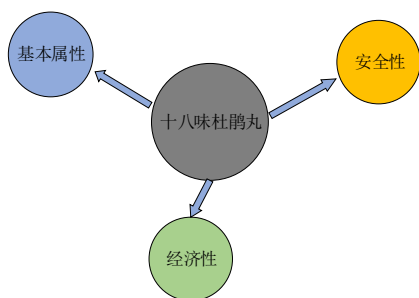


图8 知识图谱的一层展示

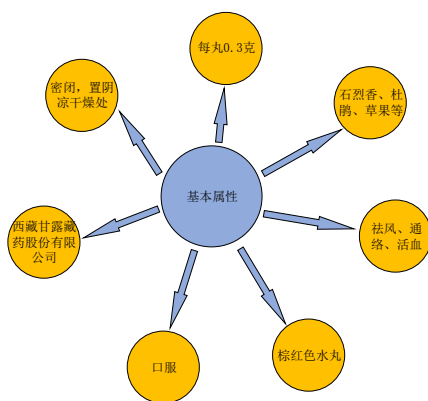


图9 知识图谱二层展示

5 总结和展望

5.1 总结

藏成药的相关数据较为复杂多样, 它的形式与当前信息技术对数据的要求有很大不同。有很多的藏成药数据仍未全面数据化, 致使在一定程度上阻碍了藏药的推广和有效应用。本文就知识图谱技术, 对现有的多个站内开放的藏成药数据进行整合, 将藏成药数据以新的存储方式和应用方式反馈给用户, 提出了藏成药知识图谱的平台构建, 本研究是藏成药与互联网技术相结合的一次新的实践, 可以促进藏成药的传播发展, 也可以为用户提供数据信息服务, 为后续藏成药的智能化提供了思路。

本文围绕藏成药的相关数据, 对藏成药的知识体系进行了深入研究, 介绍了藏成药知识图谱和平台搭建的思路, 描述了藏成药知识图谱的概念模型。通过Python爬虫技术, 我们可以收集藏成药开源资料, 并利用实体抽取和属

性抽取等方法构建基于Neo4j图形数据库系统的藏成药相关知识图谱体系。此外, 我们还利用Flask框架和Echarts可视化技术, 为未来藏成药数据可视化提出了一条新的思路。

5.2 展望

虽然本文对基于知识图谱技术的藏成药数据进行了分析研究, 但由于自身能力不够, 平台的搭建还未完全实现, 所以只提出了藏成药知识图谱体系的思路。本文尚存在以下几点不足:

(1) 本文对藏成药数据的收集均来自站内开源网站, 所采集的数据量和数据范围有限, 仍有大量的藏成药数据未被收录, 致使知识图谱的数据集不够全面。

(2) 由于数据来自不同的数据库, 所以数据的组织结构大不相同。致使一些信息在处理过程中有一些损失, 这会在一定程度上影响成果。

(3) 药物的数据属性不够全面, 对药物的成分来源、药物的发展脉络等研究都有待进一步扩展。

(4) 在对采集到的药物数据进行处理时, 实体抽取、关系抽取的人为规则指定过于局限, 可能会使部分实体数据未有效抽取, 有待优化。

参考文献:

- [1] 周雪阳, 廖诗雨, 董泽华, 等. 中成药数据图谱可视化与知识问答平台研究[J]. 软件导刊, 2022, 21(5): 158-162.
- [2] 李贡宇, 张燕玲. 常用藏药数据库的构建[J]. 世界科学技术(中医药现代化), 2012, 14(5): 2061-2062, 2064.
- [3] 于彤, 刘静, 贾李蓉, 等. 大型中医药知识图谱构建研究[J]. 中国数字医学, 2015, 10(3): 80-82.
- [4] 崔帅帅. 基于深度学习的中草药知识图谱构建及应用[D]. 阿拉尔: 塔里木大学, 2022.
- [5] 黄魏龙. 基于深度学习的医药知识图谱问答系统构建研究[D]. 武汉: 华中科技大学, 2019.
- [6] 尼玛顿珠, 洛桑嘎登, 景丽莎. 藏药知识图谱展示系统研发[J]. 电子元器件与信息技术, 2021, 5(6): 180-181, 185.
- [7] 曹皓伟, 徐建良, 窦方坤. 基于Neo4j生物医药知识图谱的构建[J]. 计算机时代, 2020(6): 35-38.
- [8] 李鹏飞. 医学关系抽取和知识图谱分布式表示研究[D]. 哈尔滨: 哈尔滨工业大学, 2020.
- [9] 王赫楠, 孙艳秋, 张柯欣. 知识图谱在中医药领域应用研究[J]. 辽宁中医药大学学报, 2022, 24(8): 182-185.

(下转第102页)

文章编号: 1007-1423(2023)24-0069-05

DOI: 10.3969/j.issn.1007-1423.2023.24.012

基于 AutoRec 的推荐系统模型研究

严武军, 刘守业*, 贺娇娇

(太原师范学院计算机科学与技术学院, 晋中 030619)

摘要: 传统推荐算法对用户的历史记录有严重的依赖, 同时面临工程复杂与表达能力差, 挖掘数据潜藏模式难的问题。为增强模型表达能力, 满足更多数据类型和业务场景, 以及解决推荐系统冷启动问题, 实验采用单隐层神经网络推荐模型 AutoRec 进行改进, 用于计算用户 User 与物品 Item 之间隐藏特征。同时, 采用概率统计学的方法, 解决系统冷启动问题。实验在 PyTorch 框架上进行数据集的训练, 与传统推荐算法相比, 该模型具有更强的表达能力和更精准的推荐结果。

关键词: 推荐系统; AutoRec; 神经网络; 冷启动

0 引言

信息技术的更新迭代, 已经改变人们获取信息的方式; 信息载体也在不断更新, 从传统的书籍、报纸和无线电广播等, 到现在的手机、电脑等, 信息载体的变化, 使得信息传播方式多种多样, 并且造成信息量爆炸式增长。在进入大数据时代, 人们每天在使用各种平台获取信息, 同时也会制造大量的垃圾信息。许多用户接触的信息, 并非自我感兴趣的或者是毫不相关的。为解决这一问题, 目前主流的平台都整合了自己的推荐机制, 用于粘合用户和给平台创造商业价值。平台需要在大量的资源中找到用户感兴趣的内容进行展示, 因此, 推荐显得尤为重要。

“推荐系统”这一核心概念, 从被提出, 到论证, 再到应用, 经过二十多年发展, 推荐系统逐渐被使用在各种场景, 尤其是工业界将其推到了顶峰。推荐系统的主流已经从经典框架到深度学习推荐模型过渡。

在传统的推荐算法中, 协同过滤算法从物品相似度和用户相似度出发, 衍生出用户协同

过滤和物品协同过滤两种算法, 开辟了推荐算法的先河, 但模型存在泛化能力差和不能良好地处理稀疏共现矩阵问题^[1]。Netflix 公司在 2006 年提出矩阵分解模型, 可以很好地解决上述两个问题; 矩阵分解在共现矩阵的基础上加入隐向量, 增强模型处理稀疏矩阵的能力, 但其无法加入用户、物品和上下文三者的相关特征, 这使得矩阵分解无法利用这些隐藏信息。逻辑回归模型是在一个线性模型的基础上添加 Sigmoid 函数, 使之能够融合不同的特征; 但其特征交叉与筛选能力弱, 不可避免地造成信息的损失^[2]。Facebook 公司首次提出基于 GBDT+LR 混合模型, 可以很好地解决高维特征组合和筛选问题。

与机器学习模型相比, 深度学习推荐模型具备更强的泛化能力, 能够根据不同的业务场景和数据特点进行泛化, 使模型与场景进行契合。澳大利亚学者首次提出一种单隐层神经网络推荐模型 (AutoRec 模型), 为后续采用深度学习模型解决推荐问题提供了思路。微软研究人员在 2016 年提出 Deep Crossing 模型, 该模型分

收稿日期: 2023-08-04 修稿日期: 2023-09-13

作者简介: 严武军 (1973—), 男, 山西太原人, 硕士, 副教授, 从事 Java EE 企业级开发、大数据方向研究; *通信作者: 刘守业 (1996—), 男, 山西朔州人, 在读硕士研究生, 研究方向为大数据、软件开发等, E-mail: lsytysf@163.com; 贺娇娇 (1996—), 女, 山西原平人, 在读硕士研究生, 从事大数据方向研究

别使用 Embedding 层、Stacking 层、Multiple Residual Units 层和 Scoring 层解决特征工程、稀疏向量稠密化和多层神经网络优化等问题。同年,上海交通大学研究员提出 PNN 模型,使用乘积层将不同特征的 Embedding 向量进行两两交叉后连接,更好地获取特征之间的交叉信息;但所有特征向量进行交叉,会丢弃原始特征向量中隐藏的价值信息。伦敦大学的学者提出 FNN 模型,采用 FM 模型训练各特征隐向量,完成初始化 Embedding 层的参数,可以加速神经网络的收敛速度。

冷启动是所有推荐系统都需要面临的问题。系统上线运营初期或新用户注册使用时,由于可用数据量小,无法精准进行用户标签化分类推荐。这里提出采用概率统计学的方法,即统计数据库中点击次数较多的信息进行展示,进而解决冷启动问题。

1 推荐系统模型介绍

1.1 逻辑回归模型

逻辑回归模型的实质是把线性回归模型通过 Sigmoid 函数进行一个非线性转换,获得一个介于 0 到 1 之间的概率值。因此,逻辑回归模型是把推荐问题转化为一个分类问题,通过计算样本的概率值对物品进行排序^[2]。基于逻辑回归的推荐模型重点在于,选取样本的特征向量和模型训练。逻辑回归模型的计算流程如图 1 所示。

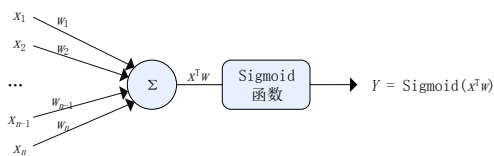


图 1 逻辑回归模型图

对逻辑回归模型进行分析可知,其将样本的特征向量进行加权求和,再经过 Sigmoid 函数进行非线性转换得到结果。采用逻辑回归模型详细推理流程如下:

(1) 将用户、物品和上下文等特征转换为数值型特征向量 $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ 作为模型的输入。

(2) 通过给各个特征赋予相应的权重 $\mathbf{w} = (w_1, w_2, \dots, w_n)$, 并将特征进行加权求和, 得到 $\mathbf{x}^T \mathbf{w}$ 。

(3) 将 $\mathbf{x}^T \mathbf{w}$ 输入 Sigmoid 函数, 使之映射在 0~1 区间, 得到最终的概率值。

(4) 根据最终的概率值大小进行排序。

由上可知, 逻辑回归模型整体推断过程的数学公式如下:

$$f(\mathbf{x}) = \frac{1}{1 + e^{-(\mathbf{w}\mathbf{x} + b)}} \quad (1)$$

式(1)中: $\mathbf{x}$ 表示特征向量, $\mathbf{w}$ 表示特征向量的权重向量, b 表示常数偏量; 说明: 对于逻辑回归模型来说, 权重向量 $\mathbf{w}$ 直接决定模型推荐效果。

1.2 AutoRec 模型

AutoRec 模型是由澳大利亚学者提出, 他把协同过滤思想和自编码器相结合, 利用协同过滤中的共现矩阵, 完成物品向量或者用户向量的自编码, 再利用自编码的结果得到用户对物品的预估评分进行推荐排序^[3]。AutoRec 模型从神经网络的角度出发, 使用一个单隐藏层的自编码器泛化用户或物品评分, 增强模型的泛化与表达能力^[4]。AutoRec 模型的提出, 找到使用深度学习解决推荐问题的思路; 在此基础上, 可以构建复杂的神经网络优化 AutoRec 模型。

自编码器在图像、音频和其他领域内, 使用极为广泛。自编码器主要是由编码器和解码器两部分组成, 其主要作用是将输入 x 转换成中间变量 z , 完成映射关系: $f: x \rightarrow z$, 将高维度的输入编码降低形成隐藏层, 将此过程看成数据编码阶段; 再将经过编码的隐藏层低维度变量 z , 完成映射关系: $g: z \rightarrow y$, 将低维度变量转换为高维度变量 y , 此过程看成数据的解码阶段^[5]。在其编码与解码阶段, 尽可能使输入变量 x 和最终结果 y 接近, 其过程如图 2 所示。

神经网络是由多个神经元组成。单个神经元是由输入、权重、偏移项、激活函数和输出五部分组成; 我们可以对一个输入进行各种不同的操作, 通过激活函数可以将正样本激活变大, 也可以让负样本激活变小^[6]。单神经元由于自身的局限性, 拟合能力差, 表达能力不足; 因此, 用多个神经元组成一个神经网络, 提升

模型的泛化能力,增强其拟合能力。通过对神经元的不同连接方式,演化出不同的深度学习神经网络模型,已经被很多领域使用,尤其在无人驾驶、自然语言处理和图像识别等领域。

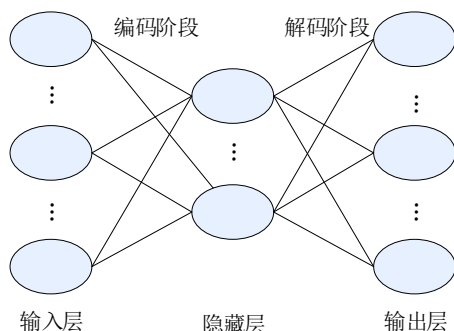


图2 自编码器编码与解码过程

AutoRec 模型是利用协同过滤中的共现矩阵,完成物品向量或者用户向量的自编码,再利用自编码的结果得到用户对物品的预估评分,再进行推荐排序^[7]。对一个物品 i 来说,所有 m 个用户对它的评分可形成 m 维的向量 $\mathbf{r}^{(i)} = (R_{i1}, R_{i2}, \dots, R_{mi})^T$, AutoRec 要解决的问题是构建一个重建函数 $h(\mathbf{r}; \theta)$,使所有该重建函数生成的评分向量与原评分向量的平方残差和最小^[8]。AutoRec 模型如图3所示, $\mathbf{V}$ 表示编码器的参数矩阵, $\mathbf{W}$ 表示解码器的参数矩阵。

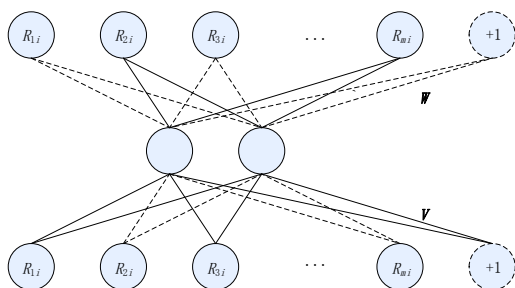


图3 AutoRec 模型

对 AutoRec 模型进行分析可知,其对物品或用户向量依次进行编码与解码,得到预估评分进行推荐。采用 AutoRec 算法详细推理流程如下:

(1) 确定自编码器的重构函数 $h(\mathbf{r}; \theta)$ 。

$$h(\mathbf{r}; \theta) = f(\mathbf{W} * g(\mathbf{V}\mathbf{r} + \mu) + b) \quad (2)$$

式(2)中,函数 g 隐藏层神经元的激活函数,函数 f 为输出层神经元的激活函数, $\mathbf{V}$ 表示编码器

的参数矩阵, $\mathbf{W}$ 表示解码器的参数矩阵, μ 、 b 表示偏差。

(2) 确定目标函数,即重建函数生成的评分向量与原评分向量的平方残差和最小。

$$\min_{\theta} \sum_{i=1}^n \|\mathbf{r}^{(i)} - h(\mathbf{r}^{(i)}; \theta)\|_0^2 \quad (3)$$

(3) 为防止重构函数的过拟合,加入 L2 正则化。

$$\min_{\theta} \sum_{i=1}^n \|\mathbf{r}^{(i)} - h(\mathbf{r}^{(i)}; \theta)\|_0^2 + \frac{\lambda}{2} (\|\mathbf{w}\|_F^2 + \|\mathbf{V}\|_F^2) \quad (4)$$

该模型是一个非常标准的三层神经网络,模型的训练利用梯度反向传播即可完成。

(4) 模型推荐

当输入物品 i 的评分向量为 $\mathbf{r}^{(i)}$ 时,模型的输入向量 $h(\mathbf{r}; \theta)$ 就是所有用户对物品 i 的评分预测。那么,其中的第 u 维就是用户 u 对物品 i 的预测 R_{ui} 。

$$R_{ui} = (h(\mathbf{r}^{(i)}; \theta))_u \quad (5)$$

通过遍历输入物品向量就可以得到用户 u 对所有物品的评分预测,进而根据评分预测排序得到推荐列表。

1.3 AutoRec 改进模型

上面已经详细介绍逻辑回归模型和单隐层神经网络推荐模型,其中逻辑回归模型泛化能力差,并且不具备特征组合与表达能力^[9];而单隐层神经网络推荐模型是利用自编码的泛化能力进行推荐,推荐效果不是特别突出,但是具有良好的特征组合能力;因此,结合两种算法模型的各自特点提出 AutoRec 改进模型,如图4所示。

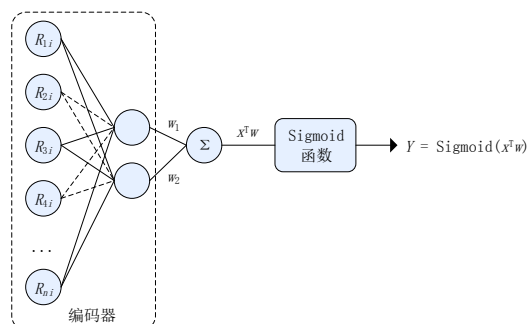


图4 AutoRec 改进模型

AutoRec 改进模型是使用 AutoRec 模型里训

练好的编码器对数据进行特征提取, 然后使用逻辑回归模型训练并评估的模型。AutoRec 改进模型的实验步骤具体如下:

(1) 使用 1.2 中的 AutoRec 模型, 完成自编码器训练。

(2) 将训练好的编码器对数据进行特征提取。

(3) 不使用解码器进行数据拟合, 改用逻辑回归模型拟合到训练数据集。

综上所述, 逻辑回归模型能够很好地融合特征, 但其存在特征筛选与交叉能力不足的问题。AutoRec 模型利用自编码器的泛化能力进行推荐, 但存在一定的表达局限性。AutoRec 改进模型结合以上两种模型的优势, 采用编码器进行特征融合处理, 并使用逻辑回归模型完成数据拟合; 这样使模型具有一定特征交叉能力和泛化能力。

2 实验过程与分析

本实验选择逻辑回归模型、单隐层神经网络模型与单隐层神经网络改进模型在相同数据集上进行对照实验, 获取实验结果。

2.1 实验环境设置

实验相关软件环境配置: Windows 10 操作系统, Python 3.10, Torch 2.0.0, CUDA 11.3; 实验相关硬件配置: Intel(R) Core(TM) i5-6300HQ 处理器, 32 GB 内存, NVIDIA GTX 1650 显卡。

2.2 实验数据集

目前存在的公开数据集有 MovieLens 电影评分推荐数据集、Last.fm 音乐推荐数据集和电影租赁网址 Netflix 数据库等, 这些数据集完整性与可用性参差不齐。因此, 本实验是依托于山西优逸客科技有限公司学员就业信息, 通过人工收集整理数据集进行模型训练, 共收集学员就业信息 2524 条。在学员就业信息中, 对非数值类型数据进行处理, 方便模型进行计算处理。本实验将处理好的就业信息数据集划分为训练集和测试集两部分, 80% 的样本数据用于模型训练, 20% 的样本数据用于模型评估。

2.3 评价指标

对于机器学习模型评价指标, 可以采用准

确率(*Accuracy*)、AUC 和 ROC 曲线等; 本实验使用准确率、精确率(*Precision*)、召回率(*Recall*)和 *F-score* 作为模型评估标准, 简洁清晰地展示实验结果^[10]。准确率是指在给定的分类数据中, 分类正确的样本数占总样本数的比例; 精确率表示分类的正样本数占分类器判定为正样本数的比例; 召回率表示在预测为正样本数中占有所有为正样本数的比例; *F-score* 是对精确率和召回率评价指标的调和值^[11]。

$$Accuracy = \frac{n_{correct}}{n_{total}} \quad (6)$$

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

式(6)中, $n_{correct}$ 为被正确分类的样本个数, n_{total} 为总样本个数; 式(7)、(8)中, *TP* 为模型正确地将正样本预测为正例的个数, *FP* 为模型错误地将负样本预测为正例的个数, *FN* 为模型错误地将正样本预测为负例的个数。

2.4 实验效果

实验分别使用逻辑回归模型、AutoRec 模型和 AutoRec 改进模型, 在上述 2524 条数据集上进行模型训练与预测, 实验结果见表 1。

表 1 实验结果

模型	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F-score</i>
逻辑回归模型	0.6476	0.6841	0.5956	0.6256
AutoRec 模型	0.7333	0.7023	0.6287	0.6421
AutoRec 改进模型	0.7561	0.7484	0.6462	0.6523

如表 1 所示, 相比于逻辑回归模型、AutoRec 模型, 在所用数据集上, 本文中提出的 AutoRec 改进模型在四个评价指标上均有提升。AutoRec 改进模型与逻辑回归模型、AutoRec 模型相比, 准确率分别提高 10.85 和 2.28 个百分点; 同时, AutoRec 改进模型比 AutoRec 模型在 *F-score* 评价指标系数上提高 1.02 个百分点, 这验证了采用训练好的编码器对数据进行特征提取, 用于辅助逻辑回归模型训练的有效性。

3 结语

自编码器作为一种神经网络, 不仅可以处

理原始数据进行特征压缩降维, 还可以通过编码器与解码器得到预估评分, 从而进行推荐排序。本文采用 AutoRec 模型进行特征融合与筛选, 数据集在该模型训练之后, 保存编码器模型, 使用逻辑回归模型替换解码器模型, 得到最终的推荐模型。实验结果表明, 本文的方法有很好的推荐效果。

参考文献:

- [1] 高长元, 黄凯, 王京, 等. 基于商品属性值和用户特征的协同过滤推荐算法[J]. 计算机工程与科学, 2017, 39(12): 2333-2339.
- [2] 常诗卉. 基于逻辑回归精排序模型的个性化推荐系统研究[D]. 广州: 东技术师范大学, 2020.
- [3] SEDHAIN S, MENON A K, SANNER S, et al. AutoRec: autoencoders meet collaborative filtering [C]//Proceedings of the 24th International Conference on World Wide Web, 2015: 111-112.
- [4] ZHU S Y, SHEN W N, QU A N. Weighted Auto-Encoding recommender system[J]. Statistical Analysis and Data Mining, 2022, 15(5): 570-585.
- [5] 刘金鑫. 自编码器在推荐系统中的应用研究[D]. 天津: 天津理工大学, 2022.
- [6] 李填填. 基于神经网络的推荐系统研究与实现[D]. 北京: 北京邮电大学, 2020.
- [7] 王雅思, 姚鸿勋, 孙晓帅, 等. 深度学习中的自编码器的表达能力研究[J]. 计算机科学, 2015, 42(9): 56-60, 65.
- [8] 余涛. 基于稀疏自编码器的手写体数字识别[J]. 数字技术与应用, 2017(1): 66, 68.
- [9] 杨帅, 王鹏. 基于堆栈降噪自编码器改进的混合推荐算法[J]. 计算机应用, 2018, 38(7): 1866-1871.
- [10] 张杰, 付立军, 刘俊明. 基于混合自编码器的协同过滤推荐算法优化[J]. 计算机系统应用, 2019, 28(5): 161-166.
- [11] 朱郁筱, 吕琳媛. 推荐系统评价指标综述[J]. 电子科技大学学报, 2012, 41(2): 163-175.

Research on recommendation system model based on AutoRec

Yan Wujun, Liu Shouye*, He Jiaojiao

(College of Computer Science and Technology, Taiyuan Normal University, Jinzhong 030619, China)

Abstract: The traditional recommendation algorithm relies heavily on the historical records of users, and at the same time, it is difficult to mine the hidden patterns of data due to the complexity of engineering and poor expression ability. To enhance the ability of model expression, meet more data types and business scenarios, and solve the cold startup problem of the recommendation system. In this experiment, AutoRec, a recommendation model of single hidden layer neural network, was used to calculate hidden features between User and Item. At the same time, the method of probability statistics is used to solve the cold start problem of the system. In this experiment, data sets are trained on PyTorch framework. Compared with traditional recommendation algorithms, this model has stronger expression ability and more accurate recommendation results.

Keywords: recommendation system; AutoRec; neural network; cold start

文章编号: 1007-1423(2023)24-0074-05

DOI: 10.3969/j.issn.1007-1423.2023.24.013

基于 BERT+BILSTM+CRF 的中国非物质文化遗产知识图谱构建

黄薇¹, 周卫^{2*}, 李宏杰², 张涛¹

(1. 广西民族大学电子信息学院, 南宁 530006; 2. 广西民族大学人工智能学院, 南宁 530006)

摘要: 中国非物质文化遗产是中国悠久历史和多民族文化遗产的重要组成部分, 当前数字化技术为非物质文化遗产的保护、传承和研究提供了新的途径和工具, 然而传统的数据库存储已经不能满足对节点关系深度查询的需求, 同时, 中国非物质文化遗产中存在着大量非结构化数据需要整理和清洗。为了提高非物质文化遗产的深度查询效率, 构建结构化数据的可视化图型, 引入 BERT+BILSTM+CRF 模型, 首先对数据进行清理和整理, 然后对数据进行标注, 对中国非物质文化遗产领域数据进行命名实体识别, 最后构建出了一个层次清晰的中国非物质文化遗产知识图谱。

关键词: 知识图谱; 中国非物质文化遗产; 命名实体识别; BERT; Neo4j

0 引言

中国的非物质文化遗产是一个了解中国文化、历史和人民智慧的窗口, 它为现代社会提供了联系过去与未来的桥梁, 代表着中华民族深厚的文化底蕴和独特的文化认同。

随着社会的不断发展和信息技术的快速进步, 数字化技术为非物质文化遗产的保护、传承和研究提供了新的途径和工具, 传统的数据库存储已经不能满足对节点关系深度查询的需求, 在此背景下, 知识图谱作为一种重要的知识表示和语义关联的数据模型, 已经引起了广泛关注^[1]。近年来, 知识图谱(knowledge graphs, KGs)已成为许多需要获取结构化知识的信息系统的基础^[2]。知识图谱以 RDF 风格三元组的形式存储节点与实体之间关系的客观信息, 三元组由两个实体和一个关系组成, 表现形式为<实体 1, 关系, 实体 2>或<实体, 属性, 属性值>^[3], 每个三元组可以看成 subject(主语)、

predicate(谓语)和 object(宾语)组成, 通过三元组的形式, 实体间的关系一目了然。

中国非物质文化遗产经过前人前仆后继的努力, 已经有了庞大的积淀, 虽然目前已有结构化的非物质文化遗产数据, 但仍存在许多非结构化资源亟待发掘与利用, 现代计算机视觉、虚拟现实技术、数字图像处理技术等数字复原和再现技术及设备日趋成熟, 为非物质文化遗产的传承和保护提供了更先进的手段与方法^[4], 在数据资源呈指数型增长的情况下, 数据的规模和复杂性使得数据管理、存储、处理和分析变得更加复杂和耗时。如何从非结构化的中国非物质文化遗产数据资源中抽取出信息是一个巨大的挑战, 由于语言的复杂性和多义性, 传统自然语言处理中手动定义的规则和模式已不适合处理非结构化数据。因此, 本文引入 BERT+BILSTM+CRF 模型, 对非结构化数据中的实体和关系进行抽取, 并构建了一个层次清晰的中国非物质文化遗产知识图谱。

收稿日期: 2023-08-15 修稿日期: 2023-09-12

作者简介: 黄薇(1998—), 女, 广西河池人, 在读硕士研究生, 研究方向为自然语言处理、文本生成; *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; 李宏杰(1998—), 男, 广东广州人, 在读硕士研究生, 研究方向为自然语言处理、知识图谱; 张涛(1998—), 男, 湖南长沙人, 在读硕士研究生, 研究方向为自然语言处理、问答系统

1 中国非物质文化遗产知识图谱的构建

1.1 数据来源及构建思路

非物质文化遗产领域因其独特的文化因素，暂时没有公开的数据集，因此，本文采用爬虫技术从多个网站，包括中国知网、中国民俗文化网、中国国家非物质文化遗产网爬取信息，数据涵盖3609个已被收入国家级非物质文化遗产的名录。

知识图谱分为自底向上和自顶向下的构建方式，自底向上的构建方式将互联网环境中的多模态数据进行汇总，提取出三元组的拆分最小单元（也即实体、属性与关系），依次按照原逻辑关系归并至数据层知识库，待知识库完善后按照一定的推理规则，将底层数据（如文本、标签、表格、图谱）依次加工构建顶层模式层中的结构、本体、规则，如Freebase。自顶向下与自底向上相反，先寻求结构化知识库去完成顶层模式层的本体学习构建，而后进行底层非结构化数据的填充与匹配，如Knowledge Vault^[5]。

本研究构建的中国非物质文化遗产知识图谱采用自底向上的方式，其具体构建流程主要分为数据获取、知识抽取、知识存储和可视化展示。

1.2 知识图谱的知识抽取

1.2.1 分词

在自然语言处理技术中，对中文文本进行处理的一大难点在于分词^[6]，非物质文化遗产的数据大部分如表1原文所示，使用当前热门的分词工具如jieba、LTP^[7]、Pkuseg^[8]，进行分词。

由于中国非物质文化遗产数据大部分都非通用词汇，并且夹杂着日常使用较少的文字，虽然可以通过用户自定义词典^[9]进行辅助分词，但仅靠人工标注和添加并不容易，分词效果会对知识图谱的构建产生直接影响，从分词效果来看，当前的分词工具已无法满足非物质文化遗产数据的分词处理。

表1 非物质文化遗产数据示例

分词工具	文本
原文	阿昌族户撒刀锻制技艺流传于云南省德宏傣族景颇族自治州陇川县西北部的户撒乡，是我国的传统技艺，于2006年收录，是第一批被收录的非物质文化遗产。
Jieba	阿昌族/户撒刀/锻制/技艺/流传/于/云南省德宏傣族景颇族自治州/陇川县/西北部的/户撒乡/，/是/我国/的/传统/技艺/，/于/2006/年/收录/，/是/第一批/被/收录/的/非/物质/文化/遗产/。/
LTP	阿昌族/户撒刀/锻制/技艺/流传/于/云南省/德宏/傣族/景颇族/自治州/陇川县/西北部的/户撒乡/，/是/我国/的/传统/技艺/，/于/2006/年/收录/，/是/第一批/被/收录/的/非/物质/文化/遗产/。/
Pkuseg	阿昌族户撒刀锻制技艺流传于云南省德宏傣族景颇族自治州陇川县西北部的户撒乡，是我国的传统技艺，于2006年/收录/，/是/第一批/被/收录/的/非/物质/文化/遗产/。/

1.2.2 词性标注

词性标注是将文本中的每个单词或词汇赋予相应的词性类别标签的过程。能够帮助计算机进行更准确的语言理解和处理。传统的词性标注方法包括基于规则的方法、基于词典的方法、基于统计的方法、基于规则和词典的混合方法以及基于转换的方法^[10]。目前传统的词性标注方法在一定程度上已经被深度学习的方法所取代，词性标注的结果同样会影响知识图谱的构建，因此，本文使用BERT预训练模型^[11]、双向长短期记忆网络(BiLSTM)以及条件随机场(CRF)，对非结构化数据中的实体和关系进行抽取，并构建中国非物质文化遗产知识图谱。

1.2.3 命名实体识别

命名实体识别是自然语言处理领域中的一项重要任务，旨在从文本中识别出特定的实体或命名实体，并将它们分类为预定的类别，由于分词和词性标注的方法都无法将非物质文化遗产的实体识别出来，因此，本研究使用BERT+BiLSTM+CRF模型，根据中国非物质文化遗产的数据特征，识别出了非物质文化遗产名称、分类、保护单位、批次、所属行政区，其中所属行政区精准划分出特别行政区、自治区、省、市、县等地理结构。

BERT+BiLSTM+CRF模型如图1所示，首先

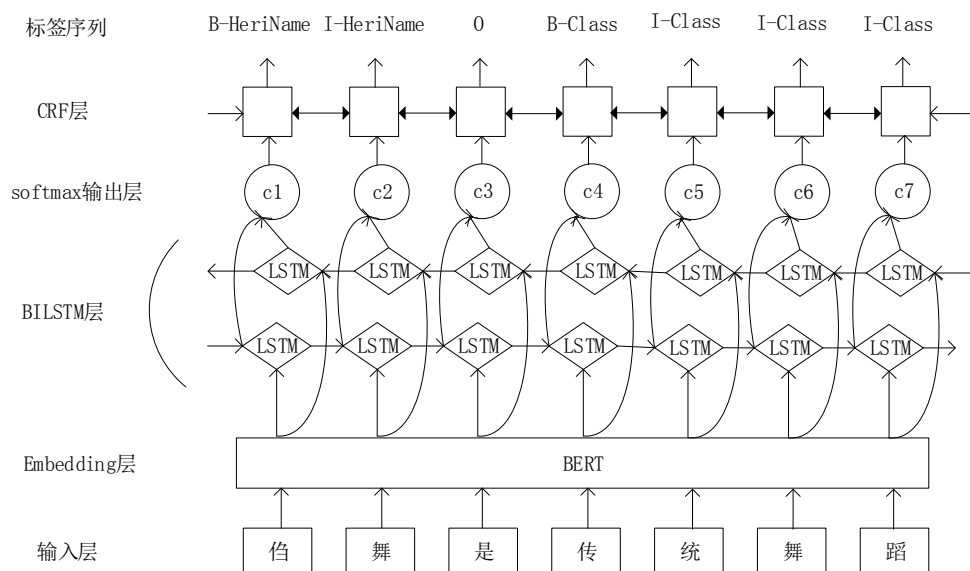


图1 BERT+BILSTM+CRF模型

将数据输入BERT预训练模型训练文本字向量，提取文本的丰富特征，并将训练好的字向量输入到双向LSTM层，从不同的方向捕获上下文信息，提取出双向文本信息，得到双向状态序列，输出层使用softmax预测各个标签的概率，最后结合CRF的状态转移矩阵，对输出序列进行约束处理，使得输出具有逻辑性，最终得到一个全局最优的标签序列。

本研究使用BIO标注形式进行数据标注，模型的识别效果示例见表2。

表2 非物质文化遗产实体抽取结果示例

原文	识别效果
伯舞发源于四川省九寨沟县，受九寨沟县文化馆保护，是第一批收录的传统舞蹈类非物质文化遗产。	'B-HerName', 'I-HerName', 'O', 'O', 'O', 'B-Province', 'I-Province', 'I-Province', 'B-County', 'I-County', 'I-County', 'I-County', 'O', 'O', 'B-ProtectUnit', 'I-ProtectUnit', 'I-ProtectUnit', 'I-ProtectUnit', 'I-ProtectUnit', 'I-ProtectUnit', 'I-ProtectUnit', 'O', 'O', 'O', 'O', 'B-Batch', 'I-Batch', 'I-Batch', 'O', 'O', 'O', 'B-Class', 'I-Class', 'I-Class', 'I-Class', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O'

1.2.4 关系抽取

本研究将识别出的中国非物质文化遗产名称与其命名实体按照先后顺序进行排序，以中

国非物质文化遗产该实体为本体，将其他命名实体与其连接，构成<实体1,关系,实体2>或<实体,属性,属性值>的三元组关系，主要工作为从构建好的知识图谱中取出所有关系R构成关系词表，将抽取出的关系在此表中进行对比，若此表中未出现过该关系，则进行相似度计算，并将关系词表中最接近该关系的词作为替换词，作为三元组的关系表示。

2 基于Neo4j的知识存储及可视化展示

本研究根据<实体1, 关系, 实体2>或<实体, 属性, 属性值>的三元组关系设计了RDF模型，图2为中国非物质文化遗产知识图谱的RDF模型。

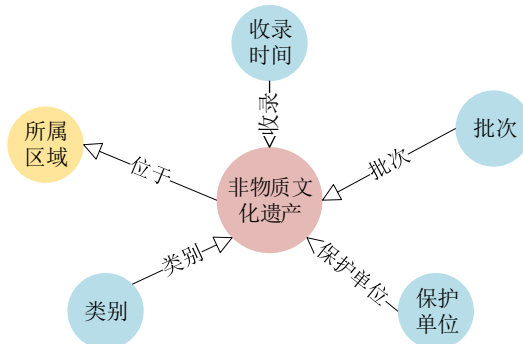


图2 中国非物质文化遗产RDF模型

使用py2neo的Python库进行知识图谱构建,首先使用CREATE_NODE命令创建中国非物质文化遗产名称、批次、收录时间、保护单位、类别、所属地域的节点,然后使用MATCH命令匹配具有相互关系的实体,最后使用CREATE_RELATIONSHIP命令进行两个节点间关系的创建,创建好的三元组关系如图3所示。

本文成功抽取出了中国非物质文化遗产知识图谱,包含8150个实体,19719个三元组,本研究构建的中国非物质文化遗产知识图谱如图4所示。

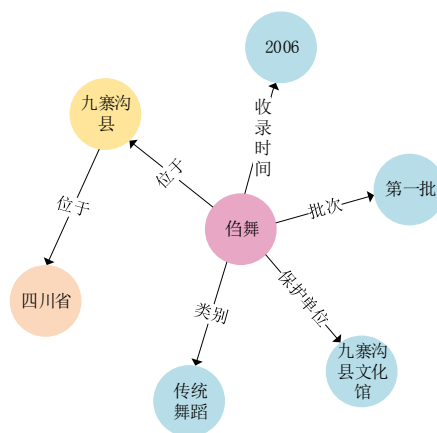


图3 三元组关系

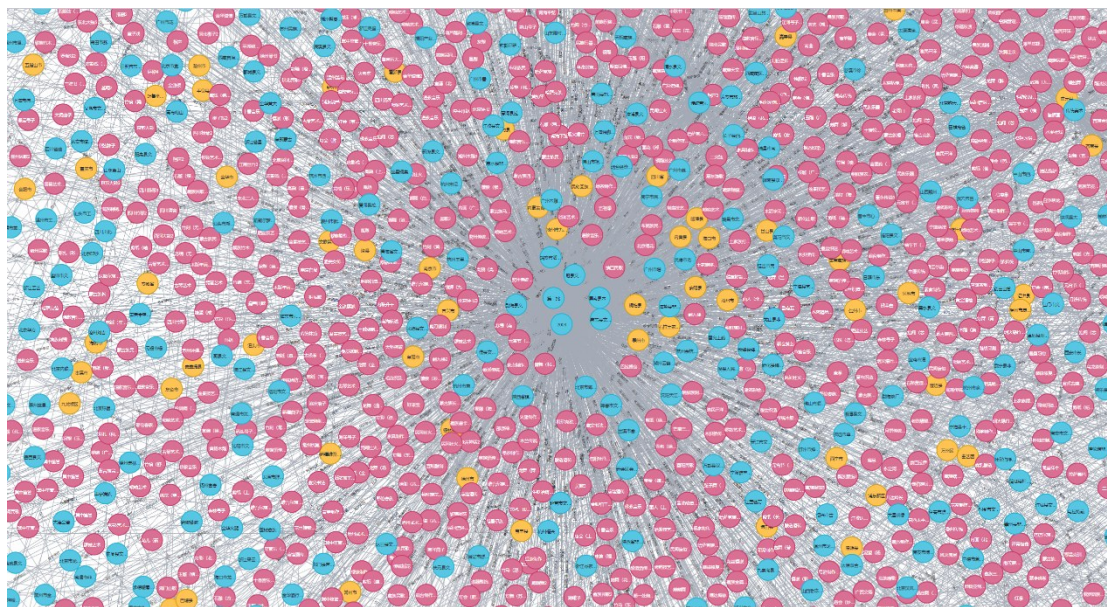


图4 中国非物质文化遗产知识图谱

3 结语

本文对中国非物质文化遗产数据进行了清洗和整理,制定了一个中国行政区域层级划分规则,将文字材料中非物质文化遗产所属区域精准分割至单元级别,提高了非物质文化遗产地理位置命名实体的细粒度;对非物质文化遗产进行命名实体标注,采用BERT+BILSTM+CRF模型对中国非物质文化遗产数据进行命名

实体识别,抽取出了非物质文化遗产名称、分类、保护单位、批次、所属行政区;使用Neo4j构建了一个庞大的中国非物质文化遗产知识图谱,为非物质文化遗产领域的图形可视化奠定了基础。由于采集的非物质文化遗产领域信息数量有限,仍有大量的信息未被充分使用,未来将对非物质文化遗产数据进行知识增强,丰富非物质文化遗产语义信息,为非物质文化遗产的保护与传承工作添砖加瓦。

参考文献:

- [1] ZOU X. A survey on application of knowledge graph [C]//Processing of the 2020 4th International Conference on Control Engineering and Artificial Intelligence (CCEAI 2020), Singapore, 2020, 1487 (1): 012016.
- [2] PU F, ZHANG Z W, FENG Y, et al. Knowledge graphs[J]. Journal of Data and Information Science, 2022, 7(2): 84-106.
- [3] MOHAMAD Z, HASSAN R, MINOU R. A review of knowledge graph completion [J]. Information, 2022, 13(8): 396.
- [4] 黄永林, 谈国新. 中国非物质文化遗产数字化保护与开发研究[J]. 华中师范大学学报(人文社会科学版), 2012, 51(2): 49-55.
- [5] 郭琳, 陈晓慧, 肖梅. 知识图谱研究综述[J]. 信息记录材料, 2023, 24(6): 17-19, 23.
- [6] DUAN S, ZHAO H. Attention is all you need for Chinese word segmentation [C]//Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), Online. Association for Computational Linguistics, 2020, 3862-3872.
- [7] CHE W, FENG Y, QIN L, et al. N-LTP: an open-source neural language technology platform for Chinese [C]//Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, Online and Punta Cana, Dominican Republic, 2021, 42-49.
- [8] LUO R, XU J, ZHANG Y, et al. PKUSEG: a toolkit for multi-domain Chinese word segmentation [EB/OL]. arXiv:1906.11455, 2019.
- [9] 赵继贵, 钱育蓉, 王魁, 等. 中文命名实体识别研究综述 [J/OL]. 计算机工程与应用: 1-14 [2023-08-05]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20230717.1232.002.html>.
- [10] 唐晓波, 郑杜, 谭明亮. 融合情报方法论与人工智能技术的企业竞争情报系统模型构建[J]. 情报科学, 2019, 37(7): 118-124, 162.
- [11] DEVLIN J, CHANG M W, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding [EB/OL]. arXiv: 1810.04805, 2018.

Construction of a knowledge map of China's intangible cultural heritage based on BERT+BILSTM+CRF

Huang Wei¹, Zhou Wei^{2*}, Li Hongjie², Zhang Tao¹

(1. School of Electronic Information, Guangxi University for Nationalities, Nanning 530006, China;

2. School of Artificial Intelligence, Guangxi University for Nationalities, Nanning 530006, China)

Abstract: China's intangible cultural heritage is an important component of China's long history and multi-ethnic cultural heritage. Currently, digital technology provides new ways and tools for the protection, inheritance, and research of intangible cultural heritage. However, traditional database storage can no longer meet the demand for deep query of node relationships. At the same time, there is a large amount of unstructured data in China's intangible cultural heritage that needs to be sorted and cleaned. In order to improve the efficiency of deep query of intangible cultural heritage and construct a visual graph of structured data, the BERT+BILSTM+CRF model is introduced. The data is first cleaned and organized, and then annotated to identify the names and entities of data in the field of China's intangible cultural heritage. Finally, a clear hierarchical knowledge graph of China's intangible cultural heritage is constructed.

Keywords: knowledge graph; China's intangible cultural heritage; named entity recognition; BERT; Neo4j

文章编号: 1007-1423(2023)24-0079-04

DOI: 10.3969/j.issn.1007-1423.2023.24.014

基于多线程对 TCP 端口扫描技术的研究

石志强*, 石美静

(南华大学计算机学院, 衡阳 421001)

摘要: 对计算机的 TCP 端口扫描问题进行了研究, 对比分析了常用端口扫描技术的原理与各自的优缺点, 采用多线程技术与端口扫描相结合的方法, 通过任务分配和优化线程数量与分配策略来提高性能。

关键词: TCP; 端口扫描; 多线程

0 引言

由于互联网的迅猛发展, 以及它在各行各业的普及, 网络安全已经成为一个越来越受到关注的话题。端口对于计算机和外部环境的交流至关重要, 它们可以帮助我们辨认出不同的网络服务和应用程序, 其中最常见的端口类型有 HTTP 端口和 FTP 端口, HTTP 端口可以提供 Web 服务, 默认端口号为 80; FTP 端口可以提供文件传输的功能, 默认端口号为 21。客户端和服务端通信时通常是客户端使用不固定的端口, 服务端使用固定端口, 客户端与服务器端的端口号不需要对应。

端口扫描技术的研究和分析对于网络安全具有重要意义。通过对常用端口扫描技术的原理和优缺点进行对比分析, 可以更好地了解攻击者的扫描方法和目标系统的脆弱性。同时, 使用多线程技术与端口扫描相结合的方法可以加速扫描速率, 提高效率。通过研究防御端口扫描的方法, 如限制 IP 访问频率、关闭不必要的端口和使用防火墙等, 可以有效减少端口扫描的威胁。总的来说, 这项研究对于提高计算机网络安全性和防御能力具有重要意义, 可以帮助安全管理员更好地评估和保护网络安全。

1 常见端口扫描技术

基于 TCP 的端口扫描^[1]分为以下几种类型。

1.1 TCP connect()扫描

TCP connect()扫描是端口扫描中最基本的一种技术, 利用操作系统提供的 connect()函数发送连接请求, 与目标计算机的端口尝试建立 TCP 连接, 以确定该端口是否处于开放状态。其主要优点在于无需特殊权限, 任何系统用户都可以使用该调用。此外, 扫描速度较快, 可以通过同时打开多个套接字并使用非阻塞 I/O 来加速扫描过程。然而, TCP connect()扫描也存在一些缺点和潜在风险。首先, 由于实际建立了 TCP 连接, 目标计算机的网络堆栈会记录连接尝试, 容易被防火墙或入侵检测系统察觉, 从而引起警报或阻止扫描。其次, 目标计算机的日志文件可能会显示大量连接和连接错误消息, 这可能迅速吸引管理员的注意并采取相应的防御措施。为了降低被察觉的风险, 黑客和安全专业人员通常采用更隐蔽的扫描技术, 例如 SYN 扫描、FIN 扫描、ACK 扫描等, 避免完全建立 TCP 连接, 降低被发现概率。

收稿日期: 2023-09-01 修稿日期: 2023-10-23

作者简介: *通信作者: 石志强(2002—), 男, 湖南怀化人, 在读本科, 研究方向为网络工程, E-mail: zqshi2002@qq.com; 石美静(2002—), 女, 安徽阜阳人, 在读本科, 研究方向为网络工程

1.2 TCP SYN 扫描^[2]

TCP SYN 扫描是一种常被称为“半开放”扫描的技术,广泛应用于计算机网络的各个方向。原理是扫描程序发送一个SYN数据包的连接请求,等待目标计算机的响应。如果目标主机的端口处于监听状态,它将返回一个SYN/ACK的响应;如果端口未处于监听状态,它将返回一个RST响应。如果收到SYN/ACK响应,扫描程序会发送一个RST数据包,则可以关闭连接。如图1所示,过程与TCP三次握手建立连接的过程类似。TCP SYN扫描的优点是可以隐蔽扫描的行为,因此一般不会对目标主机留下记录,但该方法的缺点是必须构造SYN数据包,而构造数据包具有root权限,这限制了非特权用户的使用。

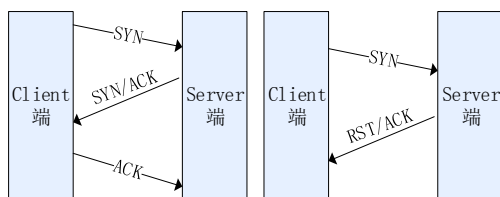


图1 TCP SYN扫描

1.3 TCP FIN 扫描^[3]

原理是向目标主机的特定端口发送一个FIN数据包。如果端口处于关闭状态,则目标主机丢弃该数据包并响应一个RST数据包。如果该端口处于开启状态,那么目标主机只会丢弃该数据包而不返回RST数据包。需要注意的是,这种扫描方法的有效性与系统实现有关,因为有些系统无论端口是否开放,都会回复RST数据包,那么TCP FIN扫描就失效了,例如Windows 95/NT操作系统。优点是TCP FIN扫描不涉及TCP三次握手协议,所以不会被目标主机发现。相比TCP SYN扫描,TCP FIN扫描更为隐蔽。缺点则是TCP FIN扫描也需要构造自己的数据包。TCP FIN扫描通常适用于目标主机是类UNIX操作系统。

本次研究采用的是TCP SYN扫描。

2 原始套接字编程原理

原始套接字^[4](Raw Socket)是一种可以直接

访问网络协议栈的套接字。它允许用户应用程序在数据链路层(如以太网)、网络层(如IP协议)和传输层(如TCP或UDP协议)之间发送和接收原始数据包。

下面是原始套接字编程的基本原理:对于链路层原始套接字的发送,套接字层会直接调用packet_sendmsg()函数,然后传递给网卡驱动的发送函数进行实际发送。这意味着数据包可以直接通过链路层发送,绕过了网络层的处理。而对于网络层原始套接字的发送,实现相对复杂一些。套接字层首先调用inet_sendmsg()函数,然后进一步调用raw_sendmsg()函数。接下来,数据包经过路由和邻居子系统的处理,最终传递给网卡驱动的发送函数进行实际发送。流程如图2所示。

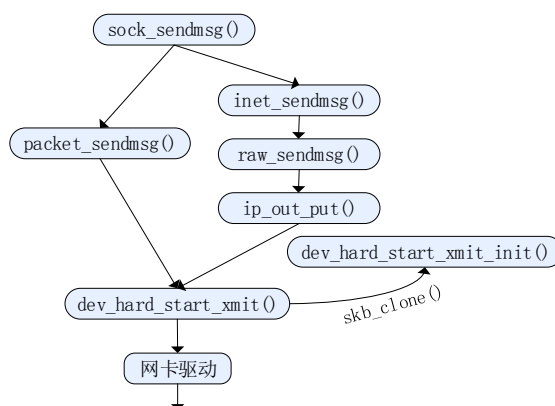


图2 原始套接字流程图

3 端口扫描的设计与实现

通常,每个电脑都拥有65535个端口,因此,要想全面检测出该IP地址范围内的所有主机的开放端口^[5],需要耗费大量的时间和精力。为提高对主机的端口扫描速率,实现同时对一个IP地址段内的主机进行高速的端口扫描,采用多线程技术,同时启用多个线程并行执行任务。通过各个线程并行执行sendto()函数与目标主机的端口进行连接^[6],如果目的端口处于开启状态,那么就会返回对应的响应报文,经过接收报文并分析判断即可知道该端口的状态;若目标主机未开放该端口,则操作失败,产生异常。具体流程如图3所示。

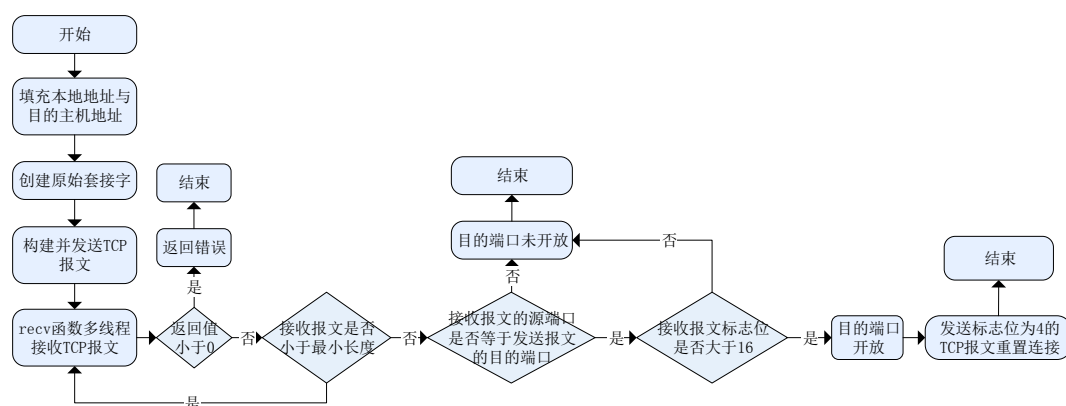


图3 端口扫描流程

3.1 发送TCP报文

在发送TCP报文前需要构建TCP随后发送。以下是构建TCP报文及发送TCP报文的关键代码：

```

typedef struct psd_hdr
{
    unsigned long saddr;           // 源地址
    unsigned long daddr;           // 目的地址
    char mbz;
    char ptcl;                      // 协议类型
    unsigned short tcpl;           // TCP长度
} PSD_HEADER;

typedef struct tcp_hdr
{
    unsigned short src_port;        // 16位源端口
    unsigned short dst_port;        // 16位目的端口
    unsigned int seq;               // 32位序列号
    unsigned int ack;               // 32位确认号
    unsigned char lenres;           // 4位首部长度/4位保留字
    unsigned char flag;             // 6位标志位
    unsigned short win;             // 16位窗口大小
    unsigned short sum;             // 16位校验和
    unsigned short urp;             // 16位紧急数据偏移量
} TCP_HEADER;

memcpy(buffer, &psdHeader, sizeof(psdHeader));
memcpy(buffer + sizeof(psdHeader), &tcpHeader,
        sizeof(tcpHeader));
tcpHeader.sum = checksum((unsigned short *)
    buffer, sizeof(psdHeader) + sizeof(tcpHeader));
memcpy(buffer, &tcpHeader, sizeof(tcpHeader));
在构建好TCP报文之后使用 sendto()函数发送
  
```

```

int ret = sendto(sockfd, buffer, sizeof(tcpHeader), 0,
    (struct sockaddr *)&target, sizeof(target));
if (ret == -1) {
    perror("sendro error");
    exit(-1);
}
  
```

3.2 接收响应报文

如果目的主机的端口开放，则会返回一段响应报文，此时需要对响应报文进行分析，以判断目的端口是否开启。以下是关键代码：

```

size = recvfrom(sockfd, msg, sizeof(msg), 0,
    (struct sockaddr *)&target, &len);
if (size < 0) {
    perror("recvfrom error");
    return 0;
}

recv_tcp = (struct tcp_hdr *) (msg + sizeof
    (struct ip_hdr));
if (size < (sizeof(struct ip_hdr) + sizeof
    (struct tcp_hdr))) {
    continue;
}

printf("%s ", timeStr);
printf("目标主机%s ", inet_ntoa(target.sin_addr));
if (ntohs(recv_tcp->src_port) == ntohs(target.sin_port)) {
    // 接收的相应报文的源端口与扫描的端口相等
    if (recv_tcp->flag > 16) {
        // 需要ack=1,此时ack=16
        printf("%d 端口开放! \n", ntohs(recv_tcp->
            src_port));
        sendTCP(4, target); // rst=1 则flag=4,需要断开连接
    }
}
  
```

```
break;
} else {
    printf("%d 端口未开放\n", ntohs(recv_tcp->
        src_port));
    break;
}
} else {
    printf("%d 端口未开放\n", ntohs(tar.sin_port));
    break;
}
}
```

4 结语

随着计算机网络的飞速发展, 端口扫描对于计算机安全与网络安全有着重要的意义, 本文对于常见的端口扫描技术进行了对比与研究, 指出了各自的优势与不足。为了实现端口扫描的高效性与准确性, 选取一项技术结合多线程

编写了端口扫描工具, 实现了对一段IP地址内多个主机端口的扫描功能, 得到了主机的开放端口序列。

参考文献:

- [1] 龚小刚. 网络漏洞扫描技术研究[C]//中国电子学会信息论分会. 中国电子学会第十七届信息论学术年会论文集, 2010: 180-184.
- [2] 侯美静. 基于智能爬行算法的网络扫描技术研究及实现[D]. 西安: 西安电子科技大学, 2018.
- [3] 肖微. 端口扫描技术的原理及应用[J]. 网络安全技术与应用, 2006(10): 18-19.
- [4] 徐建军, 丁常福. 网络安全端口扫描技术及防御方法[J]. 陕西电力, 2007(12): 41-44.
- [5] 陈珊珊. 网络安全扫描系统的设计与实现[D]. 大连: 大连理工大学, 2014.
- [6] 赵鹏, 薛冰冰. 基于加权关联规则的端口扫描检测[J]. 现代计算机(专业版), 2006(8): 23-25.

Research on TCP port scanning technology based on multi-threading

Shi Zhiqiang*, Shi Meijing

(School of Computer Science, University of South China, Hengyang 421001, China)

Abstract: The computer's TCP port scanning problem is studied, the principles of commonly used port scanning techniques and their respective advantages and disadvantages are compared and analyzed, and the combination of multi-threading technology and port scanning is used to improve the performance through task allocation and optimization of the number of threads and allocation strategies.

Keywords: TCP; port scanning; multi-threading

开发案例

文章编号: 1007-1423(2023)24-0083-07

DOI: 10.3969/j.issn.1007-1423.2023.24.015

蔬菜大棚温度可视化控制系统的设计与实现

郭金文*

(闽西职业技术学院信息工程学院, 龙岩 364021)

摘要: 温度控制是蔬菜大棚种植非常重要的环节, 对大棚温度的精准控制有助于提高棚内作物的生产效率。通过研究基于SpringBoot的蔬菜大棚温度可视化控制系统的设计和实现, 创建大棚温度监控任务, 将大棚的温度数据实时地展示在用户界面上, 并根据移动均线计算近期均温, 温度异常时发出告警, 并自动调节大棚内部的温度。在实现过程中, 结合Nginx负载均衡技术和Kafka消息中间件, 解决收集应用场景中存在的数据高并发问题, 使用了MySQL数据库存储温度数据历史记录, 通过使用MyBatisPlus框架来进行数据库操作。经过测试, 该系统能够稳定地运行, 并且在大棚温度控制管理方面具有良好的效果, 相比传统单纯依赖硬件系统进行温度控制, 该系统具有可视化控制和数据分析、远程控制等优点。

关键词: SpringBoot; 蔬菜大棚; 可视化管理; 自动化控制

0 引言

我国是一个农业大国, 农业在国民经济中占据着重要的地位, “三农”问题一直受到国家的高度重视^[1]。随着城市化的不断发展和人口的增加, 农村人口不断减少, 农业生产面临着许多困难和挑战, 尤其是农业现代化和可持续发展方面存在着较大的差距, 这对解决“三农”问题带来了不小的挑战^[2]。因此, 通过引入先进的科学和技术手段, 提高农业生产的智能化和自动化水平, 是解决“三农问题”的重要途径之一。

在农业生产中, 温度是一个非常重要的环境因素, 尤其对于蔬菜的生长, 温度的控制直接关系到蔬菜的品质和产量^[3]。采用蔬菜大棚种植方式, 使用供暖设备、通风设施, 通过热辐射、热对流等换热方式, 使棚内空气温度稳定在适宜植物生长的范围内^[4]。近年来, 也有一些蔬菜大棚温度控制系统方面的研究, 大部分研究更多的是从硬件设施系统角度展开^[5-6]。只靠硬件进行温度的调节控制, 存在一定不足之处,

如无法保存历史的大棚室内温度, 不便于分析, 也无法对大棚温度进行可视化控制、远程控制。因此, 设计一个蔬菜大棚温度可视化控制的应用系统, 对于提高蔬菜的生产效率、提高蔬菜的品质和产量具有非常重要的意义。

SpringBoot是一个基于Spring框架的开源框架, 它简化了Spring应用的构建和部署过程。它采用约定大于配置的方式, 提供了自动化配置、快速开发等功能, 使得开发者能够更加专注于业务逻辑的实现。因此, SpringBoot也在高校管理系统^[7]、政府系统^[8]、工业生产系统^[9]中得到广泛应用。

本文以SpringBoot框架为基础, 实现蔬菜大棚温度可视化控制系统的设计开发的目标。该系统可以实时监测大棚内的温度, 并根据预设的温度范围进行自动控制, 同时通过可视化界面, 让农业工作者可以直观地了解大棚内的温度变化和控制状态。通过该系统的实现, 可提高蔬菜生产的效率和品质, 帮助农民增加收入, 推动农业的现代化和可持续发展。

收稿日期: 2023-08-07

修稿日期: 2023-09-09

作者简介: *通信作者: 郭金文(1989—), 男, 福建长泰人, 硕士研究生, 助教, 研究方向为人工智能、数据驱动的软件工程、大数据分析等, E-mail: 3117845746@qq.com

1 应用系统设计

1.1 系统总体架构设计

本系统作为蔬菜大棚温度可视化控制系统平台,最终目标是利用温度数据的采集、计算,实现蔬菜大棚温度的可视化管理和自动调节,对蔬菜大棚温度进行全过程、全自动的有效管理。管理人员可以登录管理后台,创建大棚温度监控任务,根据不同种类的蔬菜大棚,设置适宜温度范围,实时查看当前大棚温度情况,开启温度异常告警,开启温度自动调节。

系统基于 SpringBoot 框架开发,采用 MVC (模型-视图-控制器)架构模式,其整体系统架构设计如图 1 所示,系统由视图层、控制器层、服务层、模型层以及外部接口层等五部分组成。

系统架构各部分详细设计如下。

(1) 模型层。模型层主要负责实现数据的存储和处理,包括数据库设计、数据访问对象(DAO)、数据实体类等。系统采用 MySQL 数据库进行数据的持久化存储,使用 MybatisPlus 作为 ORM(对象关系映射)框架,通过实体类映射数据库表结构,实现对数据库的访问操作。

(2) 视图层。视图层主要负责系统用户交互

的展示,包括用户登录、操作界面、数据可视化展示等。系统采用 HTML、CSS、JavaScript 等 Web 前端技术实现视图层,并通过 freemaker 模板渲染引擎实现前后端数据交互。

(3) 控制器层。控制器层主要负责处理用户请求和控制业务流程的执行,包括请求路由、请求参数解析、业务逻辑处理等。系统采用 Spring MVC 框架实现控制器层,通过 @RequestMapping 注解实现请求映射和参数绑定,使用 @RequestBody 注解实现请求体的解析。

(4) 服务层。服务层主要负责业务逻辑的实现和调用,包括用户登录认证、温度计算、数据查询、任务调度、数据管理、控制命令发送等。系统采用 Spring IOC(控制反转)和 AOP(面向切面编程)实现服务层的依赖注入和事务管理,使用 Java 多线程技术实现数据计算处理和命令发送的异步处理。

(5) 外部接口层。外部接口层主要负责与其他的硬件控制系统,如温度数据采集器和温度控制调节器进行交互。此外,还可以和第三方服务进行交互,包括短信通知、邮件发送、天气查询等。系统通过 HTTPS 协议和 RESTful API 实现与外部接口的交互。

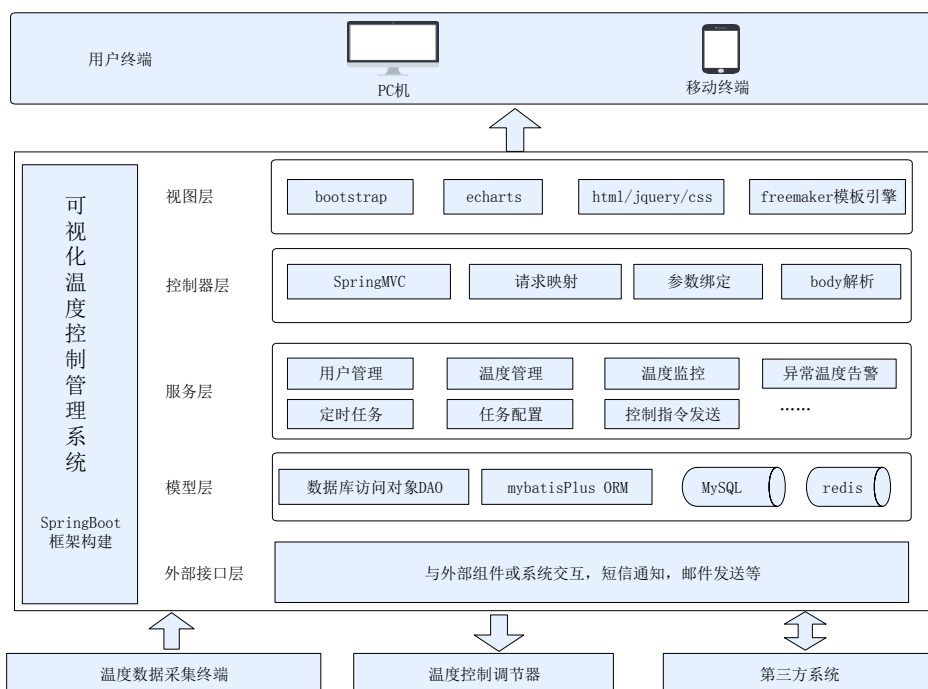


图 1 系统总体框架设计

总体来说, 本系统架构基于SpringBoot框架构建, 遵循了分层、松耦合、可拓展、易维护的设计原则, 符合现代软件工程的设计理念^[10]。

1.2 核心模块设计

1.2.1 温度数据采集存储模块

温度数据采集存储模块是本系统的一个核心模块。它负责接收多点温度采集器实时上报的温度数据, 并将数据存储在数据库中, 以供其他模块使用。该模块包含的主要部分包括:

(1) 数据采集。负责与传感器通信, 接收采集温度数据。

(2) 数据处理。该部分负责处理采集到的温度数据。它将多点温度采集器采集到的原始数据进行处理, 去除噪声, 并进行数据的平滑处理, 得到更加精确的温度数据。

(3) 数据存储。该模块负责将处理后的温度数据存储到数据库中。

1.2.2 温度管理模块

温度管理模块是系统中非常重要的部分。管理人员通过创建任务的方式, 配置各个蔬菜大棚不同的适宜温度范围, 开启任务并实时监控大棚温度。该模块主要部分包括:

(1) 创建温度监控任务。通过创建任务的方式, 针对不同的蔬菜大棚种类, 配置适宜的温度范围, 设置温度预警阈值, 自动调温的阈值。

(2) 监控并计算大棚温度。通过定时调度任务, 设置温度扫描的时间间隔, 定时从MySQL数据库或redis缓存数据库获取指定范围的温度数据进行计算。

1.2.3 温度调节控制模块

该模块主要是基于温度计算后, 根据异常的温度, 发送相关指令到相应的大棚温度调节器, 相应的降温或升温措施。主要部分如下:

(1) 异常温度获取。获取定时调度任务计算获得当前蔬菜大棚的温度, 如果是异常温度, 则进行温度预警。

(2) 温度预警。当温度超出预警阈值时, 系统可以通过短信或邮件等方式向相关人员发送预警信息。

(3) 调温指令发送。当温度达到自动调温阈

值时, 通过手动或自动发送调温指令到温度调节器, 进行相应的调温措施。

1.2.4 可视化操作模块

可视化模块部分, 主要是方便管理人员直接通过应用界面进行交互操作, 并且实时查看当前数据情况。主要包括:

(1) 用户交互。用户可以通过界面上的输入框、按钮、滑动条等交互元素来进行操作, 如设置温度范围、告警阈值、监控阈值、选择显示蔬菜大棚种类等。交互设计需要尽可能地符合用户的使用习惯和需求, 提供更好的用户体验。

(2) 数据展示。数据展示是数据可视化的核心部分。在浏览器或终端界面上, 通过条形图、柱状图、折线图等图表类型, 直观显示各大棚的温度数据, 提供给管理人员当前各蔬菜大棚的温度状态, 以便管理人员做出相应的决策。

1.3 数据库设计

在本系统中, 数据库设计是非常重要的的一部分。我们采用MySQL数据库进行数据存储, 其中核心表有6张, 包括大棚信息表、大棚温度记录表、管理员信息表、监控任务表、任务告警表、控温命令发送表。如图2所示, 是本系统的数据库模型。

2 核心功能实现

2.1 温度数据采集存储模块实现

系统的温度数据, 来自大棚内多点温度采集器的定时上报, 系统通过Nginx提供数据请求API接口服务, 开放https接口, 供温度采集器调用进行数据上报。Nginx配置负载均衡, 将数据请求分发到后端服务器进行处理。Nginx负载均衡策略采用ip_hash, Nginx具体负载配置如下所示:

```
upstream dataservers {  
    ip_hash;  
    server 172.18.45.11: 8081;  
    server 172.18.45.12: 8081;  
    server 172.18.45.13: 8081;  
}  
server {  
    listen 443;
```

```
ssl_certificate/cert/nginx/localhost.pem;  
ssl_certificate_key /cert/nginx/localhost.key;  
server_name localhost;  
location / {  
    #转发到负载服务上  
    proxy_pass http://dataservers;  
}  
.....  
}
```

考虑到大型的蔬菜大棚场地存在大棚数量较多，温度采集上报的温度数据量会比较大，存在瞬时高并发的问题。本系统在接收数据后，通过消息队列(MQ)进行数据的缓冲，我们采用 Kafka 这种高吞吐量的分布式发布订阅消息系统作为消息中间件。数据处理服务系统消费 Kafka 中的数据，进行数据清洗后存入到 MySQL 数据库中^[11]。数据采集处理存储实现原理图如图 3 所示。

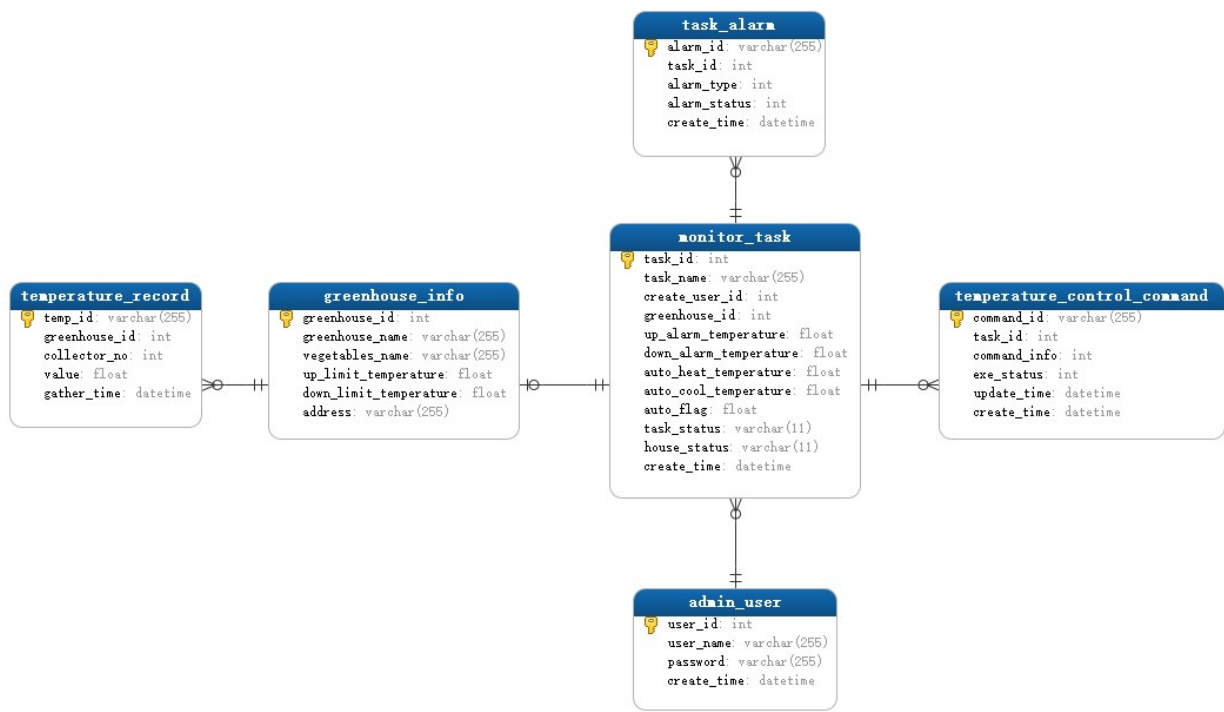


图 2 数据库模型图



图 3 数据采集处理实现原理图

2.2 温度管理模块实现

本系统的温度管理功能,通过监控任务进行,针对需要进行温度管理的蔬菜大棚,创建一个监控任务。创建任务前,管理人员需要录入蔬菜大棚的信息,包括名称、品种、地址、适宜温度上限、适宜温度下限等信息,存入表 `greenhouse_info`。

创建任务时,管理人员选择要进行温度管理的大棚,创建大棚温度监控任务,同时填写监控的相关信息,包括任务名称、温度上限告警阈值、温度下限告警阈值、温度自动调节阈值、是否自动调温等,创建任务信息存入 `monitor_task` 表中。后台任务创建可视化管理界面,如图4所示。

新建管理任务

请选择需要监控的蔬菜大棚

📌 新增任务

请命名任务名称

东区-一号监控任务

请输入温度上限告警阈值

请输入温度下限告警阈值

请输入自动加热阈值

请输入自动降温阈值

是否开启自动调温

☒ 开启 ☐ 关闭

新增任务

图4 温度管理任务创建界面

2.3 温度控制模块实现

创建并开启温度监控任务后,系统开始接收温度数据并入库,同时系统会启动定时调度任务,根据指定的时间间隔,从数据库获取温度数据进行计算,判断当前大棚的温度状态,系统一般配置的频率间隔为1分钟。为了避免一些瞬时温度异常导致的温度误判,我们采用线性加权移动均线的计算方法,取当前(包含)往前 N 个温度数据 $C_1, C_2, C_3, \dots, C_n$, 进行加权累加求均值计算,结果作为近期均温 D_v , 具体计算公式如式(1)所示,为了避免近期均温 D_v 和当前温度 C_1 相差过大,使判断温度能更加准确反映近期阶段的温度,我们需要限定 n 的取值范围。

$$D_v = \frac{\sum_{i=1}^n (n-i)C_i}{\sum_{i=1}^n i}, \text{其中 } n \leq 10 \quad (1)$$

当近期均温 D_v 计算结果大于任务中配置的温度告警上限阈值,或小于温度告警下限阈值,则系统将会触发告警,调用第三方组件,如邮箱或短信,发出告警通知。如果任务配置了开启自动调温机制,当近期均温 D_v 大于自动降温阈值,或小于自动加热阈值,则系统发出自动升降温的指令,指令通过 `https` 接口发送给温度控制调节器,调节器根据收到的指令进行相应的升降温处理,如打开通风阀门、打开遮阳网等。根据近期均温 D_v 的结果、自动升降温指令的发送结果,同时需要更新任务的状态,以便管理员从前端界面可以直观获取当前的任务情况。

2.4 可视化模块实现

本系统后台可视化管理,主要是基于 B/S (Browser/Server)交互模式设计,通过 Web 浏览器进行交互,通过浏览器直接登录后台管理。管理人员可直接使用 PC 机或移动设备进行监控任务配置和温度的管理,无需到大棚现场。

技术上,系统前端使用 `jQuery` 函数库进行 Dom 操作,采用 `Bootstrap` 响应式框架开发,一套代码就能适配 PC 大屏和 `mobile` 小屏,大大提高开发效率。数据可视化部分,系统引入了 `Echarts` 可视化库,通过折线图、柱状图、散点图等展示温度等数据,使数据呈现更加美观大气。数据管理表格,我们引入了 `BootstrapTable` 组件,通过该组件实现表格的功能,包括排序、筛选、分页等交互特性,使监控任务记录、温度数据呈现更加直观、易于管理。可视化界面开发设计,如图5所示。

3 系统测试

管理人员登录后台,配置好监控任务并启动,系统开始针对配置的大棚温度进行监控管理。系统经过一段时间的运行,我们在可视化界面得到如图6所示的温度走势图。



图 5 后台管理可视化界面



图 6 系统执行监控温度走势图

从图6可以看到,当近期温度 D_t 突破温度上限阈值时(图中蓝色虚线标线,交叉点A),系统发出告警。当温度突破自动降温阈值时(图中红色实线标线,交叉点B),系统发出自动降温指令到温度控制调节器,温度控制调节器接收到降温指令后,开启降温模式,例如打开大棚通风阀门执行降温。系统持续监控温度状况,当大棚开启降温后,温度回到预警线下(交叉点C)时,系统发出停止降温指令,温度控制调节器关闭降温模式。当近期温度 D_t 向下突破预警和降温阈值时,原理一致。

4 结语

本文介绍了一个基于SpringBoot的蔬菜大棚温度可视化控制系统的设计和实现。系统通过接收温度采集器发送的温度数据,将温度实时地展示在用户界面上,便于管理人员实时查看当前大棚温度状态,同时系统可根据用户设置的温度范围自动调节大棚内部的温度。在系统设计开发过程中,使用了MySQL数据库存储温度数据、用户信息、监控任务信息等,通过使用MyBatisPlus来进行数据库操作,使用线性加权移动均线计算近期均温,使得温度监控结果更加准确,更精准地进行温度自动化控制。除此之外,系统还集成了Kafka消息中间件,结合Nginx负载均衡技术,增强了系统的可扩展性和性能。

参考文献:

- [1] 郑有贵. 着力解决“三农”问题 推动乡村振兴[J]. 红旗文稿, 2022(10): 39-42.
- [2] 张新美. 论绿色发展理念与农业现代化[J]. 农业经济, 2019(6): 26-27.
- [3] 马振玉, 姜会飞. 最热(冷)月时段温度作为农业种植区划指标的探讨[J]. 中国农业大学学报, 2021, 26(9): 63-73.
- [4] 郭智勇, 王冬计, 王晓雪, 等. 地埋管供热对蔬菜大棚土壤温度分布影响的模拟研究[J]. 沈阳农业大学学报, 2022, 53(1): 73-82.
- [5] 贾宇琛, 李岩, 肖珂. 蔬菜大棚温度调节控制器[J]. 安徽农业科学, 2010, 38(24): 13393-13394, 13397.
- [6] 秦君琴, 李兴财, 杨有贞. 分数阶PID控制器在蔬菜大棚温度控制中的应用研究[J]. 西南大学学报(自然科学版), 2016, 38(1): 179-182.
- [7] 朱志慧, 蔡洁. 基于SpringBoot+Vue+Uni-app框架的校园失物招领系统[J]. 电子技术与软件工程, 2022(17): 62-65.
- [8] 韩佳彤, 杜芳, 甄必鑫, 等. 基于SpringBoot的分布式市政运维分析平台[J]. 内蒙古大学学报(自然科学版), 2022, 53(3): 309-316.
- [9] 徐小辉, 刘江涛, 高涵, 等. 基于SpringBoot+Vue框架的采气方案系统开发[J]. 计算机仿真, 2021, 38(6): 248-250, 382.
- [10] 吴化尧, 邓文俊. 面向微服务软件开发方法研究进展[J]. 计算机研究与发展, 2020, 57(3): 525-541.
- [11] 孙海勇, 金鸿宇, 富宏亚, 等. 面向实时智能控制的数据采集平台及其应用研究[J]. 航空制造技术, 2022, 65(22): 87-93.

Design and implementation of a visualized temperature control system for vegetable greenhouses

Guo Jinwen*

(College of Information Engineering, Minxi Vocational & Technical College, Longyan 364021, China)

Abstract: Temperature control is a crucial aspect of vegetable greenhouse cultivation, and precise control of greenhouse temperature can improve crop production efficiency. By studying the design and implementation of a visualized temperature control system for vegetable greenhouses based on SpringBoot, this paper creates a greenhouse temperature monitoring task that displays real-time temperature data on the user interface and calculates the recent average temperature based on the moving average. The system sends alerts when there are temperature abnormalities and automatically adjusts the temperature inside the greenhouse. In the implementation process, the system uses Nginx load balancing technology and Kafka message middleware to solve the problem of high concurrency in data collection scenarios. The system stores temperature data in the MySQL database and uses the MyBatisPlus framework for database operations. Testing shows that the system can run stably and has good effects on greenhouse temperature control management. Compared to traditional hardware-based temperature control systems, this system has the advantages of visualized control, data analysis, and remote control.

Keywords: SpringBoot; vegetable greenhouse; visualized management; automated control

文章编号: 1007-1423(2023)24-0090-07

DOI: 10.3969/j.issn.1007-1423.2023.24.016

基于 JavaWeb 的线上考试系统的设计与实现

操心慧*, 温智聪

(广州华商学院数据科学学院, 广州 511300)

摘要: 随着时代的发展, 考试、考核在日常工作学习中愈发频繁, 尤其是在学校中, 考试对于学生来说是必不可少的。而线下考试则面临着纸张浪费、教师批改繁琐、特殊情况无法开展等问题。线上考试系统可以通过线上发布试题、线上答题、系统自动批改等流程帮助师生完成相应考试, 并且提供学生管理、试题管理、试卷管理、成绩查询、自动寻题组成试卷等功能。系统基于 JavaWeb, 采用 B/S 模式, 数据库采用 MySQL 技术。

关键词: 线上考试系统; JavaWeb; B/S 模式; MySQL

0 引言

一场考试的组织过程包括: 设计试卷、打印试卷、发放试卷、学生答卷、收卷、教师阅卷、得分统计等, 在实际生活中是无法将这些事务的顺序打乱的, 其中设计卷、教师阅卷和得分统计这三件事最繁琐。当考试参与的人数过多, 就要花费大量的人力投入到阅卷工作中, 人的精神和批改有限, 效率底下, 并且阅卷之后, 还需进行成绩登记, 从经济方面思考人力不是最好的办法。此外, 站在学生的视角, 线上考试一方面省去了忘记携带纸笔的麻烦, 也能随时随地参与考试, 并且不用进行考前等候, 立即出成绩。站在教师的视角, 省去了打印试卷、批改试卷和统计分数等工作, 节省大量时间。因此本文希望能使用 JavaWeb 技术, 开发出一套能解决设计试卷、教师阅卷、得分统计等繁琐任务的软件。

随着高等院校课程内容不断增加, 学生需要学习的知识也变多了, 如果每一个科目都要设置一场线下考试, 不仅老师的负担增加, 也会牺牲学生更多的时间, 无法高效地学习。所以考试也要顺应时代发展, 充分利用网络分布

性的特点, 实现在线考试、在线阅卷、自动统计等功能, 有效减低成本, 节约资源。计算机快速响应和海量的存储力, 能准确抓取能力考核过程, 提供详细的分析数据, 为教育工作者提供帮助。

本文的研究目的在于研究如何实现线上考试系统, 需要实现什么功能, 还有研究设计时需要注意什么问题, 实现本设计需要做到以下几个方面:

(1) 查阅资料, 确认在线考试系统的基础功能、发展进程、工作流程。

(2) 根据考试执行步骤, 绘制结构图。

(3) 设计功能模块, 如: 后台控制模块。考生信息的增删改, 新试卷的发布和考试时长设置, 题目集合库管理, 还有学生考试分数统筹计算等功能。前台学生考试模块: 查看已发布的考试, 参与考试, 查询历史考试记录等功能。

(4) 运用 JavaWeb 技术和 MySQL 数据库, 并设计列表的主外键, 精简的程序代码编辑, 调试实现的系统功能, 大量的性能测试。

(5) 了解程序的运行原理, 了解 MySQL 表结构式的数据关联原理, 以及数据库与程序链接, 并交互的原理。

收稿日期: 2023-08-14 修稿日期: 2023-09-14

基金项目: 广州华商学院校级青年学术科研项目(2022HSXS084)

作者简介: *通信作者: 操心慧(1995—), 女, 安徽安庆人, 硕士, 助教, 研究方向为数据可视化、应用开发、多目标优化, E-mail: 945152118@qq.com; 温智聪(2001—), 男, 广东陆丰人, 学士, 研究方向为应用开发

1 开发技术及环境

本系统是基于JavaWeb的线上考试系统，所谓JavaWeb是使用Java语言开发Web应用程序的一种技术方案。JavaWeb应用程序一般由动态生成的网页组成，可用于各种类型的应用程序，如电子商务、企业管理系统等。采用JavaWeb具有跨平台性、安全性、灵活性等优点。

本系统以Java语言为基础进行考试系统开发，后台的代码编写中主要使用了Java、J2EE。Java是目前使用最为广泛的网络编程语言之一，它具有语法简单、面向对象、稳定、与平台无关、多线程、动态等特点。Java是一种编程语言，被特意设计用于互联网的分布式环境。J2EE的全称是Java 2 Platform Enterprise Edition，它是由SUN公司领导、各厂家共同制定的工业标准，也是Java的一种企业版，用于企业级的应用服务开发^[1]。前台使用了JavaScript、CSS等语言，JS是一种脚本，一门编程语言，C++展现的是静态信息，相比之下，前者可以实现更为多样化的功能，呈现的效果更加流畅。譬如：可跟随滚动屏移动视频播放位置的效果和交互式地图。虽然以开发Web页面的脚本语言闻名世界，但很多的非浏览器环境中，也有应用到JS。通过eclipse来编写这些程序，用MySQL进行数据库的设计与分析。MySQL是一个关系型数据库管理系统，在万维网应用方面，是非常适用的系统，采用表的形式来存储数据，增加了搜寻速率，让检索方式简易化、多样化^[2]。另外该系统的内置容器为spring Application，配置该容器，就可以进行考试。

2 系统需求分析

2.1 可行性分析

2.1.1 经济可行性

设计系统第一需要考虑的就是设计成本，成本需要控制在小于项目前期带来的预计利润。综合需求分析，本文设计所消耗的成本几乎可以忽略不计。首先该项目的前期开发几乎没有资金投入，只需自己翻阅资料，请教老师、同学，此外只需配备一台电脑，成本有限。

2.1.2 技术可行性

本系统以Java技术为基础，开发平台选用有着完整配备功能的Eclipse，新手也能轻松操作掌握，还有代码补全提示，精准报错的控制台，使用时更加得心应手。Java语言编码的优点：语言简洁，降低了出错的概率；兼容性强，无论是Windows平台，还是Linux、MacOs平台，编译后的程序都能运行；也能适应于变化的环境，是一个动态语言。同时MySQL具有可投射视图和不易受破坏的特点，安全性高，还有存储数据以及检索数据的能力。

2.1.3 操作可行性

该系统的部署容器是IDEA开发软件内置的springApplication容器，拥有该容器就可利用网络访问本系统，没有该容器，配置tomcat亦可。本系统使用jsp技术，利用网络可操作，且界面简洁。本系统具有易操作、易管理、交互性好的优点。

2.2 数据流程分析

数据流程分析就是舍去物理构建，只分析数据流动的一种数据分析模式。一般使用分层DFD进行。

2.2.1 顶层数据流程图

经由分析，顶层数据流程图如图1所示。

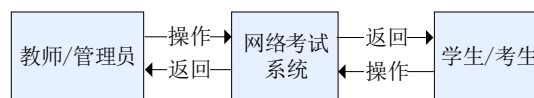


图1 顶层数据流程图

2.2.2 一层数据流程图

将顶层数据流程图每个进程再详细分析之后，可再绘制出如图2所示的数据流程图。

2.2.3 二层数据流程图

得到一层数据流程图，再对其深入解析数据流程图，如图3所示。此处以考试管理为例。

2.3 功能需求分析

本系统分为三大模块，分别为学生使用的考试功能模块、教师使用的考试管理模块以及管理员模块。

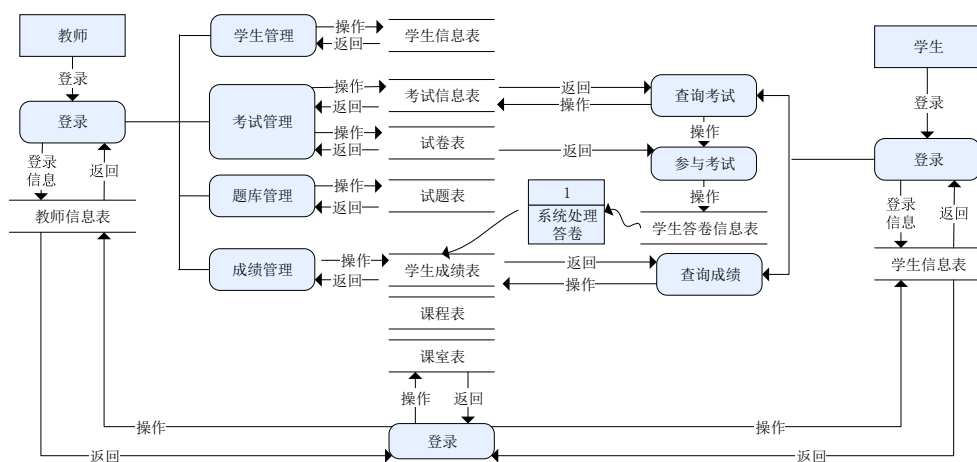


图2 一层数据流图

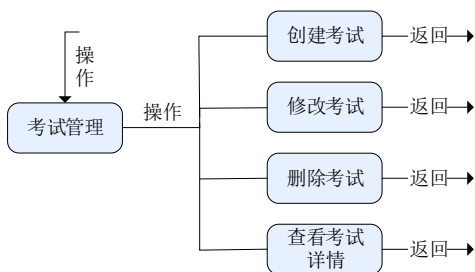


图3 二层数据流图

2.3.1 学生模块

学生使用的考试功能模块的功能分别为注册登录账号、考试中心、考试历史、修改密码等^[3]。

(1) 注册登录账号：学生根据身份证号以及学号、手机号等相关信息可进行验证注册，注册完成后可根据注册的账号、密码进行登录。

(2) 考试中心：学生在中心能看见发布过的所有考试，能参加在有效时间内的考试。可看见考试名，所属科目，考试时长和总题目量和满分分数，并进行答题。

(3) 考试历史：学生能看见曾经参加过的考试，完成试卷的时长和得分，还能看见自己历史答题所写的答案还有正确答案以及解析。

(4) 修改密码：学生可通过对旧密码的验证后修改自己的密码，防止被别人恶意操作。

2.3.2 教师模块

教师使用的考试管理模块的功能分别为注册登录账号、学生管理、考试管理、题集管理

和分数管理等。

(1) 教师同样需要根据身份证号以及工号、手机号等相关信息进行验证注册，注册完成后可根据注册的账号、密码进行登录。

(2) 学生管理：学生可以通过二维码、序列号等方式加入班级，同时教师可根据具体情况添加、删除相关学生。

(3) 考试管理：一个是查看考试信息，该功能可查看发布过的考试信息，能修改试卷，删除试卷，也能看见考卷的详细信息。另一个是创建考试，编辑试卷名称，编辑所属的课程和需要参与该考试的班级，设置题目数量，设置分值，规定考试时间范围和考试时长。生成试卷时，会自动组题组成试卷^[4]。

(4) 题集管理：教师可以编写题目，编辑选项和答案，并规定所属课程。

(5) 分数管理：该功能可查询到所有参与过考试的人员信息和得分。

2.3.3 管理员模块

管理员负责后台的系统管理。管理员可实现用户管理、考试管理、班级管理等功能，可对各个模块进行增、删、改、查操作。同时管理网站的配置，管理前端布局以及相应资讯^[5]。

3 系统总体设计

本文依据拥有的权限的不同进行划分，可以划分成学生、教师、管理员。教师权限包括：登录功能、成绩表、用户管理、题库管理、试卷管

理。学生权限包括：登录功能、参与考试、查询历史得分以及查询历史答卷信息与答案解析。管理员可进行用户管理、考试管理、班级管理等功能，可对各个模块进行增、删、改、查操作。

通过分析，该系统设计的几种重要逻辑表单及E-R图如下：学生用户表见表1，ER图如图4所示。用户id：每一个登陆的用户的系统生成的编号；用户名：兼账号；密码：登录密钥；真实姓名：用户真名；班级id：用户所在班级号。

表 1 users表

字段名	数据类型	长度	字段含义	是否主键	是否外键
userid	int	11	用户 id	是	
username	varchar	10	用户名		
userpwd	varchar	25	用户密码		
truename	varchar	15	真实姓名		
classid	int	11	班级 id		

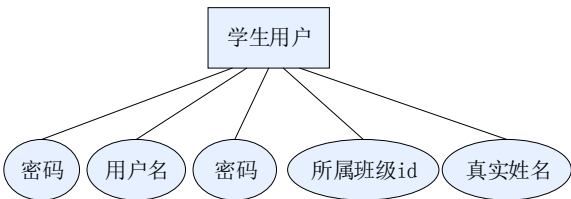


图 4 学生用户表ER图

教师用户表见表2，ER图如图5所示。

表 2 teouser表

字段名	数据类型	长度	字段含义	是否主键	是否外键
userid	int	11	教师 id	是	
username	varchar	15	教师用户名		
userpwd	varchar	20	密码		
truename	varchar	15	真实姓名		
classid	int	11	班级 id		

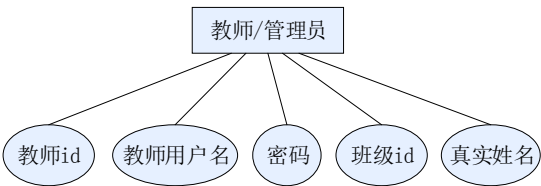


图 5 教师表ER图

教师id：每一个登陆的教师的系统生成的编号；教师用户名：教师登录时作为账号使用；密码：密码对应相应的账号；真实姓名：教师的真实姓名；班级id：显示教师所属的班级id。

考试信息表见表3，其ER图如图6所示。课程id：这里指代的是该考试属于哪个科目；教师id：出卷老师的教师id，生成试卷是绑定登录的教师账号；试卷id：系统自动递增生成的试卷编号；试卷名称：教师给予的试卷名称；班级id：在生成试卷的功能中限定可参与考试的班级；选择题数量：试卷题目数量；每道题分数：同类型题目分数；开始时间：开启时间；结束时间：关闭时间；考试时长：允许的答题时长。

表 3 exam表

字段名	数据类型	长度	字段含义	是否主键	是否外键
eid	int	11	试卷 id	是	
pname	varchar	25	试卷名称		
cno	int	11	课程 id		是
userid	int	11	教师 id		是
classid	int	11	班级 id		是
singlenunber	int	11	选择题数量		
singlecore	varchar	50	每道题分数		
examdate	date	0	开始时间		
examtime	date	0	结束时间		
testtime	int	11	考试时长		

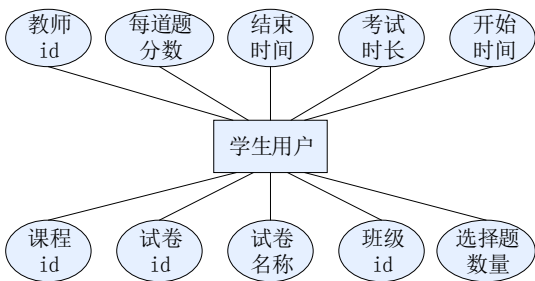


图 6 考试信息表ER图

考卷表见表4，其ER图如图7所示。考卷id：试卷编号；考卷名：考卷名称；题目：题目；选项：可供选择的答案；试题序号：题目的序号；考卷所属科目：试卷所属学科。

表 4 paper 表

字段名	数据类型	长度	字段含义	是否主键	是否外键
pid	int	11	考卷 id	是	是
eid	int	11	试卷 id		
sid	int	11	试题 id		是
cno	int	11	课程号		
scontent	varchar	255	题目		
sa	varchar	255	选项 A		
sb	varchar	255	选项 B		
sc	varchar	255	选项 C		
sd	varchar	255	选项 D		
skey	varchar	255	正确答案		

表 5 student exam 表

字段名	数据类型	长度	字段含义	是否主键	是否外键
seid	int	11	id	是	
userid	int	11	学生 id		是
classid	int	11	班级 id		是
eid	int	11	试卷 id		是
pname	varchar	255	试卷名称		
zscore	int	11	试卷总分		
score	int	11	学生分数		
titime	date	0	考试时间		

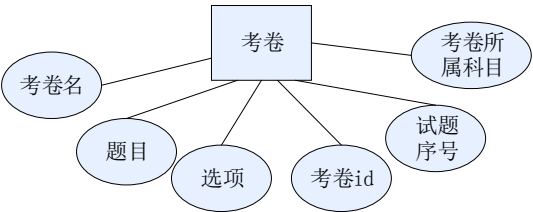


图 7 考卷表 ER 图

学生考试成绩表见表 5，其 ER 图如图 8 所示。成绩表 id：该数据库表格的 id；班级 id：班级号；学生 id：学生特有编号；试卷 id：进行答卷的试卷 id；试卷名称：学生进行答卷的考卷名称；试卷总分：学生进行答卷的考卷满分分数；学生分数：学生答卷所得分数；考试时间：考试时长。

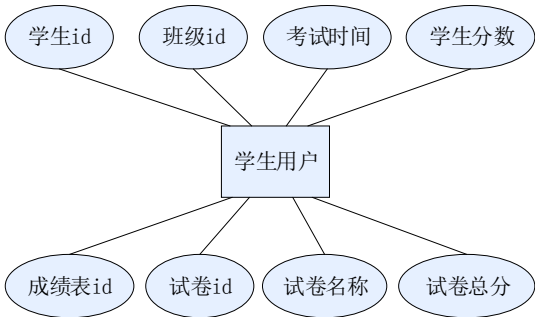


图 8 学生考试成绩表 ER 图

4 系统实现

用户可以在登陆页面选择学生身份或教师身份进行登录，如图 9 所示。

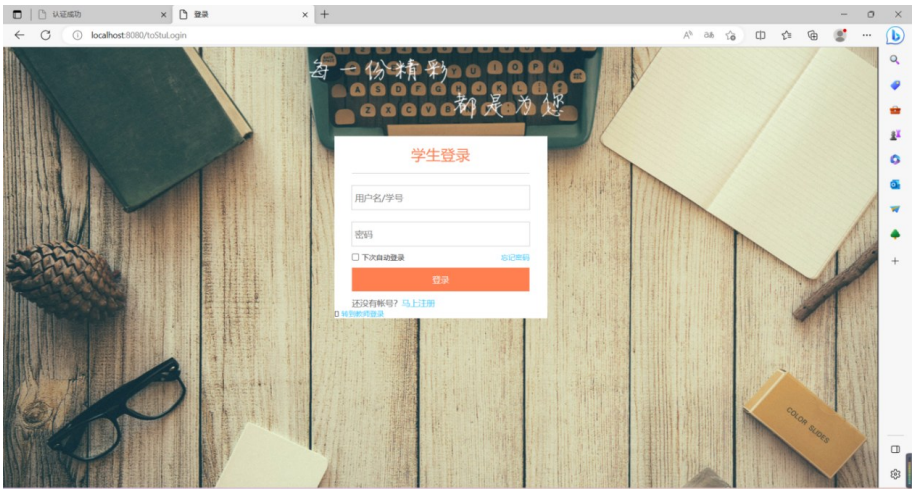


图 9 登录界面图

学生登录之后可以实现查看首页、管理个人信息、参加考试、查询成绩、修改密码等功能，如图 10 为学生参加考试的考试中心页面；图 11 为学生修改密码界面图。教师登录之后可以实现查看首页、管理学生、管理考试、题库管理、成

绩管理等功能，如图 12 为教师对学生进行管理的效果图；图 13 为教师创建新的考试的页面。管理员后端登录即可进入系统后台管理页面，可实现用户管理、考试管理、班级管理等功能，可对各个模块进行增、删、改、查操作。

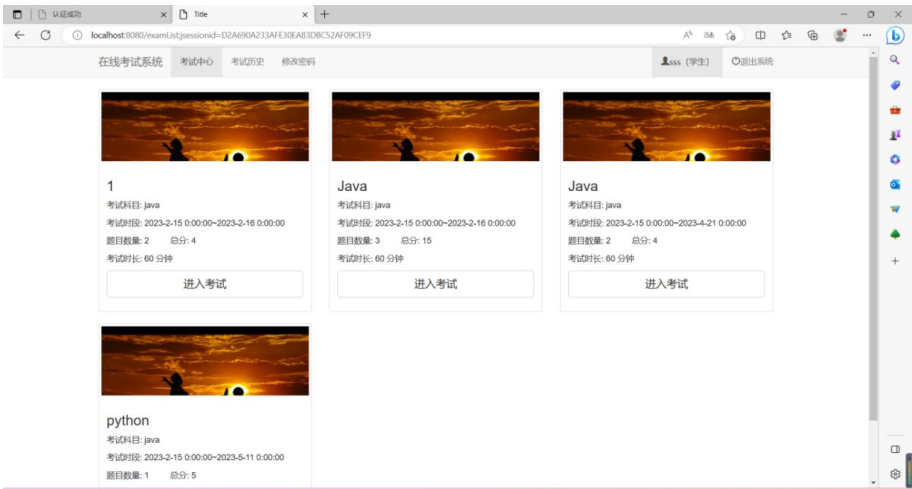


图 10 考试中心界面

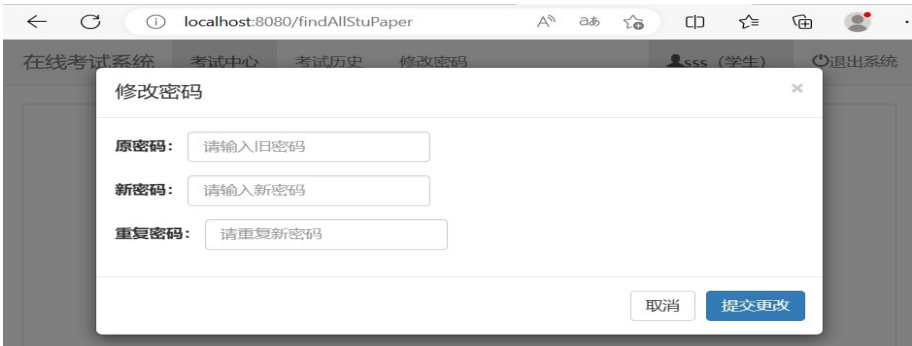


图 11 修改密码界面图



图 12 学生管理界面图

后台管理中心

Welcome to the online exam system !

试卷名称

命题人

所属课程

所属班级

单选题数量

每小題分数

考试日期

考试时长

生成试卷

图 13 创建新考试图

5 结语

本文运用Java、JavaEE、MySQL、JavaScript等技术设计并实现了基于JavaWeb的线上考试系统。与类似系统相比,本系统具有跨平台性,可以在不同平台上运行,减少了开发与维护的难度;具有良好的安全性,提供了多层防护机制,可以有效保护系统安全。

参考文献:

[1] 蔡兴壮. 基于C/S的无纸化考试系统的设计与实现

[D]. 青岛:中国石油大学(华东),2017.

[2] 刘洋. 无纸化自动评分考试系统的设计与实现途径探索[J]. 电脑知识与技术,2019,15(31):79-80.

[3] 齐琳. 职业技能培训系统模拟考核平台设计与实现[J]. 软件导刊,2014,13(11):90-92.

[4] 宋颖. 无纸化考试系统的设计与实现[D]. 大连:大连理工大学,2019.

[5] 刘崇巍. 移动实习就业跟踪系统的设计与实现[D]. 沈阳:东北大学,2015.

Design and implementation of an online examination system based on JavaWeb

Cao Xinhui*, Wen Zhicong

(School of Data Science, Guangzhou Huashang University, Guangzhou 511300, China)

Abstract: With the development of the times, exams and assessments are also widely present in daily life, especially in schools, where exams are essential for students. Offline exams, on the other hand, face problems such as paper waste, cumbersome teacher grading, and inability to carry out special situations. The online exam system can assist teachers and students in completing corresponding exams through online publishing of exam questions, online answering, and system automatic grading processes. It also provides functions such as student management, exam question management, exam paper management, score query, and automatic question finding to form exam papers. The system is based on Java web, adopts B/S mode, and the database adopts MySQL technology.

Keywords: online examination system; JavaWeb; B/S mode; MySQL

文章编号: 1007-1423(2023)24-0097-06

DOI: 10.3969/j.issn.1007-1423.2023.24.017

基于 GPR 模型的草地地上生物量估算软件设计

王志飞*

(成都理工大学旅游与城乡规划学院, 成都 610000)

摘要: 草地生物量的研究对植物生态学和植物生产都具有重要的意义。草地生物量研究有助于更好地了解草地生态系统, 草地生物量是草地生态系统健康评价、草地资源可持续开发利用的标准, 可以帮助我们更好地管理草地。近年来, 机器学习在估算草地生物量方面的研究取得了一定的进展。而高斯回归模型是一种有效的机器学习方法, 可以用来估算草地生物量。它的优点在于可以利用草地的多种信息(如气候、土壤、植物特征等), 从而更准确地预测草地的生物量, 对正确识别草地特征和提高草地生物量估算的准确性有重要的意义。然而高斯回归模型具有一定的复杂性, 对无程序基础的研究者来说实现具有一定的难度, 由于它需要调整大量的参数, 使模型能够很好地拟合数据, 这需要大量的计算时间和计算资源, 更需要不断尝试超参数的优化。如果不断更改代码的话, 耗时耗力, 并且需要一定程度的代码基础, 因此设计基于 GPR 模型的草地生物量估算软件就十分具有必要性。

关键词: 草地; 地上生物量; 高斯回归; 软件设计

0 引言

草地地上生物量(aboveground biomass, AGB)集中体现植被光能利用能力, 可反映草地生长状态和植被种类的动态变化, 是草地生态系统健康评价、草地资源可持续开发利用的关键指标^[1]。草地生态系统较为脆弱, 受扰动后变化过程的时空差异显著、不确定性突出, 因此以草地生物量为关键指标, 探讨草地生态系统碳汇特征及其对气候变化、人类活动的影响和响应, 已成为全球变化研究的重要方面^[2-4]。

目前, 草地生物量估算以地面实测与遥感估算建模最为常见^[5]。草地 AGB 地面采样实测方法适用于局部准确估算, 但空间不连续, 受到人力、物力和时间等客观条件的限制, 不适合在大范围或自然环境恶劣的地区开展^[6]。而 GPR 模型能够充分利用多光谱遥感数据, 在生物量估算中具有比函数回归更好的精度^[7]; 此外, GPR 根据高斯概率提供不确定性估算^[8], 它能对未知输入进行输出反演, 同时也能估算出该反演的精度, 因此建立一种基于 GPR 模型

的草地生物量估算系统十分有必要。

GPR 是近些年发展的一种机器学习方法, 它属于非参数回归模型, GPR 模型对草地生物量的估算是一个复杂的物理过程, 包含多种输入要素, 可以根据输入变量对目标变量进行估算^[9]。它使用参数估计来估计, 可以最好地拟合训练数据的函数, 从而可以将该函数应用于不同的输入变量, 以获得特定的输出结果。对草地生物量的估算, 受内部因素和外部因素的综合影响。内部因素包括 GPR 模型的内在影响因素, 均值函数、核函数、最大似然函数和推断函数的选择等因素, 以及对各种函数相对应的超参数结构的选取等等, 外部因素包括生成训练集的各种参数, 如叶面积指数、叶绿素含量、干物质含量、水含量、叶片结构参数、土壤湿度等参数的敏感性分析, 以及率定, 以生成最适配的训练集对训练集的训练和数据预处理也是重要影响因素。

近年来, MATLAB 高效的数值计算及符号计算功能, 能使用户从繁杂的数学运算分析中

收稿日期: 2023-05-06 修稿日期: 2023-09-02

作者简介: *通信作者: 王志飞(1997—), 男, 河南邓州人, 硕士, 研究方向为自然地理学, E-mail: 1179674226@qq.com

解脱出来；并且其具有完备的图形处理功能，实现计算结果和编程的可视化，以 MATLAB 软件为技术平台的模型估算一直是该领域的热门。本系统在 MATLAB 的支持下，对草地生物量进行估算，建立了一个独立的草地生物量估算系统，实现数据输入、参数调整、图形估算一体化，对没有计算机类基础的研究人员可以实现自主化进行草地地上生物量的估算。

该软件主要分为十大功能模块，由“估算的数据输入”“高斯回归原理”“均值函数、核函数、最大似然函数方法的选取”“对所选取参数的超参数输入”“训练数据的步长值选取”“估算图像显示”“运行结果 R^2 、RMSE、RRMSE 结果的显示”和“报错信息提示”“使用教程”“联系我们”等组成，每个大模块下又划分功能区，以便逐一实现功能，不失统一性与逻辑性，各个功能在此得以精化完善，以求用户得到最好的体验，软件总体结构如图 1 所示。

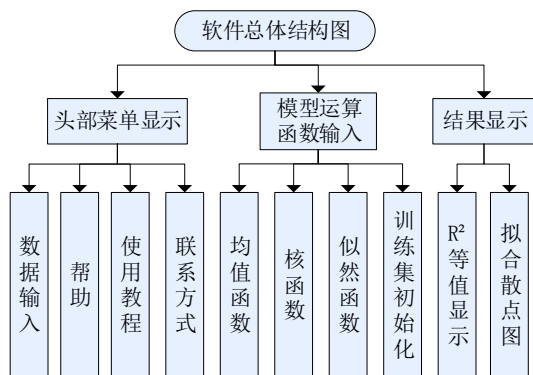


图 1 软件总体结构图

本软件主要特点如下：

(1) 丰富化。充分利用 App 形式的优势，将基于 MATLAB App 的 GPR 模型关于生物量的估算应用在表现层上，使得视觉效果更好、交互性更强、提供良好的用户视觉效果和兼容性。

(2) 实用性。基于 MATLAB App 的设计和基于 GPR 模型估算生物量的各种参数的选取，以及各种结果展示综合于一体的开发使得对相关专业的，以及专业部分对生物量的估算更加直观，同时系统结构简洁，功能方便、灵活，用户界面友好，且独立于界面化运行，便于系统操作人员的管理和使用，即使是不懂代码的人员也可以毫

无压力地进行相关区域的草地生物量估算。

(3) 统一性与规范性。软件设计遵循统一、规范的信息编码和框架系统、规范的数据输入精度与符号系统等。

(4) 可延展性与开放性。软件的设计考虑到软件的扩展和与其它系统的兼容，在整体框架的设计、App 的便捷性以及软件功能等方面，尽可能留有余地，方便该软件的扩充，更加入大量的数据输入，增加了对该软件操作的自由性。

1 软件概述

软件整体的主界面包括：导航菜单、主窗口、信息输入栏和关闭提示框。

(1) 导航菜单：导航菜单从左到右依次为数据、帮助、使用教程、联系我们。用户点击导航菜单上相应的按钮，就可以展开内容按钮，点击需求的内容后就可以实现窗口的跳转。

(2) 信息输入栏：数据输入、函数方法选取、函数超参数输入提示、超参数输入、训练数据集步长值输入、超参数关键参数 Length 值输入等，选择需要的函数类型，并选择其不同的算法后，点击提示按钮即可提示输入超参数初始化参数值的相关格式和数量。并可以根据下拉框内不同的选择实现实时地、动态地刷新提示，用户可以自由选择并且输入。

(3) 关闭提示框：为防止用户误触关闭按钮导致估算运行停止，增加了关闭软件时的警示框，避免前功尽弃。

基于 GPR 模型的草地地上生物量估算 App 软件的主界面如图 2 所示。

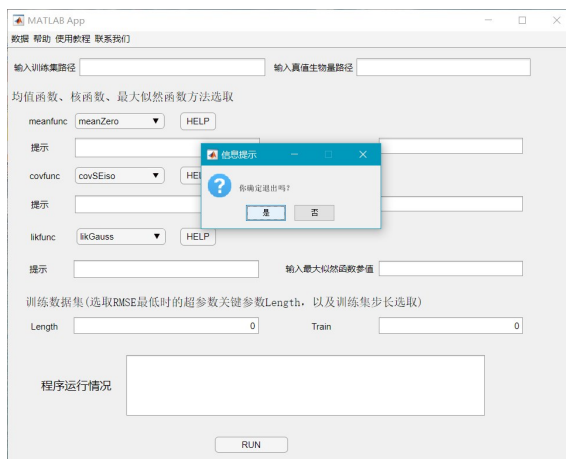


图 2 基于 GPR 模型的草地地上生物量估算软件的主界面

2 软件主要功能描述

软件整体的主界面包括：GPR模型相关原理介绍、软件使用视频教程、联系方式获取、数据输入、函数方法选取、函数超参数输入提示、超参数输入、训练数据集步长值输入、超参数关键参数 Length 值输入、程序运行状态提示、 R^2 、RMSE、RRMSE 值显示，错误信息提示、关闭弹窗提示等。

2.1 数据输入

打开数据按钮后显示所输入数据类型的下拉框，包括训练集和真值生物量，用户点击相应的按钮后可以调用电脑文件夹显示，因为设置了xlsx属性，用户只可以选择Excel文件，其他文件不会显示。用户选择文件后，其路径会显示在相应的文本框中。如图3所示。



图3 GPR模型估算软件数据输入

2.2 帮助

为使用户在使用GPR模型时更加方便、准确，在此设置了帮助功能，为用户全面化的地讲解了高斯回归估算的原理，更加便捷地为用户提供帮助。点击帮助时会出现GPR模型原理，点击后会出现讲解弹窗。用户可以实时阅读，如图4所示。

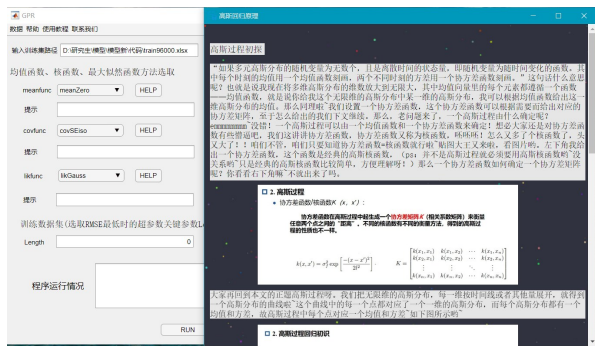


图4 软件头部菜单帮助功能

2.3 使用教程

更方便、更准确地使用GPR模型估算，增加了使用教程功能，包括图文教程和视频教程。可以指导用户快速实现GPR模型估算草地生物量软件App的使用。如图5所示。

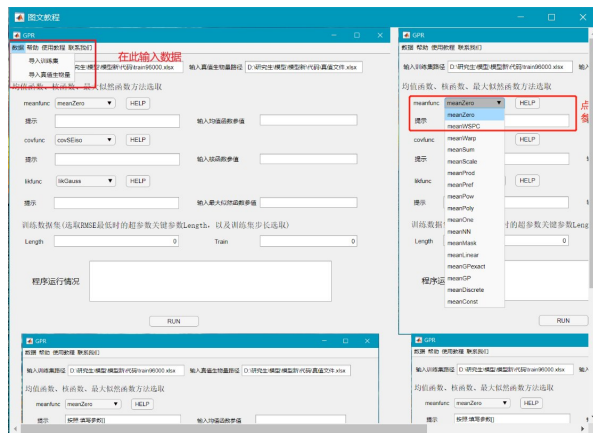


图5 软件头部菜单图文教程功能

2.4 函数方法选取

函数的方法选取从上至下依次是均值函数、核函数、似然函数等，通过将每种函数的各种方法封装在一个下拉框中以供用户选择相应函数不同的算术方法，选取后会显示在下拉框中，可以实现多种类化的应用。如图6所示。

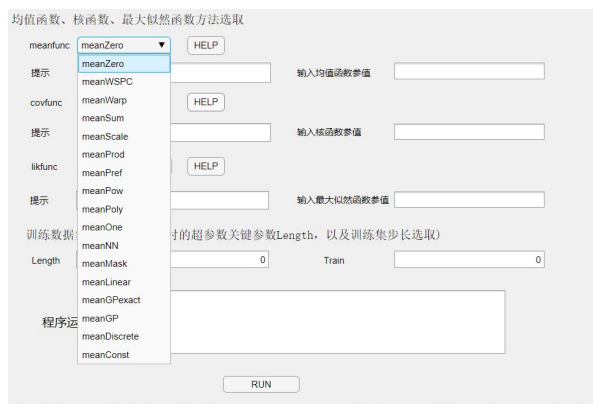


图6 软件函数方法选取

2.5 函数超参数初始化参值输入提示

每当选取一种相应的均值函数、核函数、似然函数后，为方便用户更好地对数据进行估

算, 要对超参数进行初始化来更好地优化 GPR 模型。而为了方便用户可以自主方便地选择不同的函数类型, 在此为用户增加参数值输入提示项, 点击函数选择后面的函数提示按钮, 会在提示的文本框中显示用户应该输入的参值类型以及数目。如图 7 所示。



图 7 超参数初始化参数值选取

2.6 训练数据集

选择核函数类型之后, 模型超参数是决定 GPR 模型的关键。超参数是在开始学习过程之前设置的参数, 并非通过训练得到。通常情况下, 需要对超参数进行优化, 给学习机选择一组最优超参数, 以提高学习的性能和效果。Length 是超参数的关键参数, 它的绝对值给出函数计算的最大允许次数, 控制分布的延展、平滑性。而为了测试 GPR 模型精度, 验证模型的泛化能力和鲁棒性, 采用 10 折交叉验证方法对模型进行评估, 交叉验证的目的是得到可靠稳定模型, 当训练集的步长值为 30 时 RMSE 总体较小。所以在进行模型估算时要对 Length 的值和训练数据集步长值 Train 进行输入。如图 8 所示。



图 8 Length 值和 Train 值输入

2.7 程序运行情况提示

当 GPR 模型所有的输入数据、步长值以及

所有的函数和参数选择完成后, 点击 Run 运行按钮, 会在程序运行框中显示程序运行状况, 可以让用户清楚地知道软件的运行情况, 如图 9 所示。



图 9 软件运行过程提示和结果显示

2.8 运行结果数据显示

模型运行完毕后, 该模型的估算 R^2 、RMSE、RRMSE 的结果将显示在运行框内, 可以使用户清楚地知道该模型的数据拟合精度, 方便用户根据结果对核函数、均值函数、似然函数以及相应参数做出改变。如图 10 所示。



图 10 软件运行结果数据显示

2.9 基于 GPR 模型的生物量估算 App 运行图像显示

基于 GPR 模型的草地生物量估算软件是为

了得出最终的生物量估算值和真值的拟合程度,以图像弹窗的形式显示最终结果,根据图像可以清楚地看到估算值和真值的差异性对比,以及拟合程度。以弹窗的形式展示还可以实现保存、打印、编辑等各项工作。如图11所示。

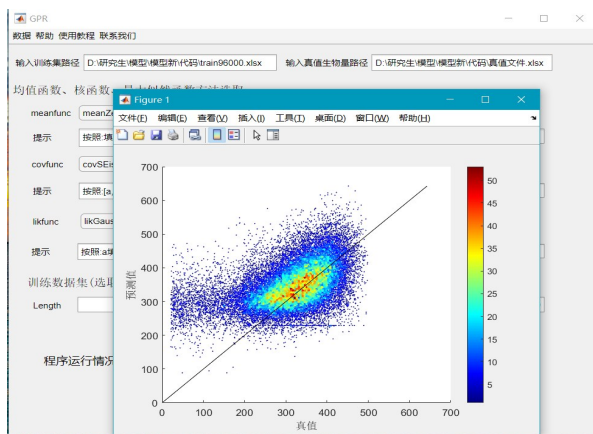


图11 基于GPR模型的生物量估算APP运行拟合图像

3 结语

近年来,随着环境保护意识的提高,草地上生物量的研究成为了热门研究,而基于GPR模型估算草地地上生物量是一种有效的方法,它可以从遥感影像数据中提取出关键特征,从而更准确地估算草地地上生物量。本文基于此开发的GPR模型估算草地地上生物量的软件可以很好地解决GPR模型代码复杂性的问题,在很大程度上减少了调整大量参数时耗时耗力的缺点,可以更加直观地更改参数,以及选择最优的超参数优化值,并且能够更加直观地观察到调整参数以及超参数优化得到的最终拟合散点图和 R^2 以及RMSE。很好地为缺少代码研究基础的研究者提供了一个大范围估算草地地上生物量的平台,能够更加直观、简洁地进行草地地上生物量的估算,更好地对草地进行检测以及保护。

参考文献:

- [1] 何丽. 基于PROSAIL和GPR模型的若尔盖草地上生物量遥感反演研究[D]. 成都: 中国科学院大学(中国科学院水利部成都山地灾害与环境研究所), 2020.
- [2] XIU X M, ZHOU S F, CHEN Q, et al. Above-ground biomass estimation of provincial scattered grassland based on Google Earth Engine and machine learning[J]. Bulletin of Surveying and Mapping, 2019 (3): 46-52, 75.
- [3] FU X, TANG C, ZHANG X, et al. An improved indicator of simulated grassland production based on MODIS NDVI and GPP data: a case study in the Sichuan province, China [J]. Ecological Indicators, 2014, 40(1): 102-108.
- [4] KUMAR L, MUTANGA O. Remote sensing of above-ground biomass [J]. Remote Sensing, 2017, 9 (9): 935.
- [5] 宋柯馨, 蒋馥根, 胡宗达, 等. 西藏自治区草地地上生物量遥感反演研究[J]. 生态学报, 2023(13): 1-14.
- [6] 赵天舸, 于瑞宏, 张志磊, 等. 湿地植被地上生物量遥感估算方法研究进展[J]. 生态学杂志, 2016, 35 (7): 1936-1946.
- [7] 秦格霞, 吴静, 李纯斌, 等. 基于机器学习算法的天祝藏族自治县草地地上生物量反演[J]. 草业学报, 2022, 31(4): 177-188.
- [8] VERRELST J, RIVERA J-P, MORENO J, et al. Gaussian processes uncertainty estimates in experimental Sentinel-2 LAI and leaf chlorophyll content retrieval [J]. ISPRS Journal of Photogrammetry & Remote Sensing, 2013, 86(1): 157-167.
- [9] 何志昆, 刘光斌, 赵曦晶, 等. 高斯过程回归方法综述[J]. 控制与决策, 2013, 28(8): 10.

Design of grassland aboveground biomass estimation software based on GPR model

Wang Zhifei*

(School of Tourism and Urban-Rural Planning, Chengdu University of Technology, Chengdu 610000, China)

Abstract: The study of grassland biomass is of great significance to plant ecology and plant production. Grassland biomass research helps to better understand the grassland ecosystem. Grassland biomass is the standard for grassland ecosystem health assessment and sustainable development and utilization of grassland resources, which helps us to better manage grassland. In recent years, machine learning has made some progress in estimating grassland biomass. The Gaussian regression model is an effective machine learning method that can be used to estimate grassland biomass. Its advantage is that it can use a variety of information of grassland (such as climate, soil, plant characteristics, etc.) to predict grassland biomass more accurately, which is of great significance for correctly identifying grassland characteristics and improving the accuracy of grassland biomass estimation. However, the Gaussian regression model has a certain complexity, and it is difficult for researchers without program basis to implement it. Because it needs to adjust a large number of parameters, so that the model can fit the data well, which requires a lot of computing time and computing resources. It is necessary to constantly try to optimize the hyper parameters. If the code is constantly changed, it is time-consuming and requires a certain degree of code base. Therefore, it is necessary to design a grassland biomass estimation software based on GPR model.

Keywords: grassland; aboveground biomass; GPR; software design

(上接第 68 页)

Visualization construction of knowledge graph based on Tibetan medicine data

Fan Miao^{1*}, Gao Yi^{1,2}

(1. School of Information Engineering, Xizang Minzu University, Xianyang 712082, China;

2. Xizang Key Laboratory of Optical Information Processing and Visualization Technology, Xianyang 712082, China)

Abstract: Tibetan medicine is an important component of traditional Chinese medicine, with a long history. It gathers the experience of Tibetan people in fighting against infectious diseases for many years, and has made indelible efforts for the healthy development and reproduction of ethnic groups in the Qinghai Tibet Plateau and surrounding areas. It has made tremendous contributions to the development of the Chinese nation. In order to effectively store, manage, and use various Tibetan medicine data, a knowledge graph based storage structure and visualization technology are proposed to achieve visualization of Tibetan medicine related attributes, thereby improving the integration of Tibetan medicine knowledge and the correlation of data. By using knowledge graph technology, the potential value of Tibetan medicine data can be better explored.

Keywords: Tibetan medicine; knowledge graph; data visualization

文章编号: 1007-1423(2023)24-0103-05

DOI: 10.3969/j.issn.1007-1423.2023.24.018

基于无纸化策略的胎儿医学中心信息平台改造

莫远明*, 乔梦薇, 林思阳, 黎鹏安

(中山大学附属第一医院信息数据中心, 广州 510080)

摘要: 阐述基于无纸化策略胎儿医学中心信息平台改造的必要性, 通过利用 HIE 数据平台、电子签名等技术, 实现信息平台的无纸化改造, 并介绍改造后信息平台主要业务流程、架构、功能以及应用情况。通过无纸化改造, 信息平台实现了数据闭环与互联互通, 解决了原来纸质化单据多、容易漏收错收费用等问题, 提高了医护人员工作效率, 节省了科室运营成本, 提高了患者服务水平, 助力智慧医疗的实现。

关键词: 胎儿医学中心; 无纸化; HIE 数据平台; 电子签名; 智慧医疗

0 引言

胎儿医学是当前围产医学最为引人注目的研究和发展领域, 涉及基因组学、蛋白组学、代谢组学、微创手术、基因治疗等最前沿的生物学技术, 是现代生物医学的集大成者。胎儿医学中心主要任务包括胎儿疾病的产前筛查和诊断、疾病的治疗与规范管理等^[1], 其中产前诊断是胎儿医学的重要内容^[2], 非侵入性产前诊断技术包括染色体微阵列分析、荧光原位杂交技术与细胞核型分析等, 都是目前产前诊断的主要手段^[3]。因此, 先进生物技术平台的建立, 为临床医护人员诊断治疗胎儿疾病提供了良好的工具。由于产前诊断和治疗涉及大量的实验室检测数据, 因此利用信息平台快速获取检查结果信息, 形成业务全流程的闭环, 提高胎儿医学中心运营效率, 为患者提供优质医疗服务, 有利于促进胎儿医学中心的发展。

1 无纸化概述

无纸化策略是解决医疗系统数字化程度低、信息互联互通程度差的重要抓手, 是现代信息

与网络发展的产物, 其优点在于不仅能节约资源、降低成本, 而且便于医学资料存档和检索, 提高工作效率, 从而有利于业务流程的规范化、程序化、便捷化^[4], 助力医院高质量发展。医疗系统的无纸化需要依靠医疗信息交互(health information exchange, HIE)数据平台、电子签名、物联网等技术, 来实现异构系统数据交互、业务的全流程闭环、电子签名的安全性及便捷性^[5]以及医疗设备数据的自动获取, 为数据的合法、可信、安全提供技术保障。无纸化改造并不代表完全无纸或者脱纸, 需要通过系统改造、流程再造和设备淘汰等措施逐步推行^[6], 是一个长期的过程, 同时对于系统用户的专业性、信息化操作水平也提出了更高的要求。

2 现状与问题

医院胎儿医学中心产前筛查与诊断检测主要包括 CMA 染色体微阵列分析、荧光原位杂交、外周血细胞染色体检查等项目, 标本有静脉血、羊水、脐血、绒毛等类型。随着医疗水平的提高和业务的发展, 原有系统逐渐不能适

收稿日期: 2023-08-16 修稿日期: 2023-09-12

作者简介: *通信作者: 莫远明(1989—), 男, 广东高州人, 本科, 工程师, 研究方向为医疗信息化, E-mail: 739703923@qq.com; 乔梦薇(1993—), 女, 河南安阳人, 硕士, 助理工程师, 研究方向为医院信息系统建设与运营管理; 林思阳(1988—), 男, 广东广州人, 硕士, 助理工程师, 研究方向为医疗信息化; 黎鹏安(1994—), 男, 广东广州人, 硕士, 助理工程师, 研究方向为信息技术在医疗领域的应用

应临床的需求,包括:①电子化程度低,有大量的纸质单据,例如项目申请单、标本运送单、结果报告单等,需要耗费大量的人力物力去管理和维护,造成资源浪费;②各个模块互联互通能力差,数据闭环程度低,标本登记、运送、报告发放等需要人工处理,临床不能在系统上直接获取患者检查报告,影响诊疗效率,而且容易漏收错收费用,给医院带来不必要的损失;③科室的质控、绩效、科研数据需要人工二次收集和记录,无法直接提供电子化数据支撑,给科室管理和科研工作带来不便;④纸质诊疗报告、检查结果发放效率低且容易丢失,给医疗工作带来风险,影响患者就医体验。因此,以无纸化为抓手,对原有系统进行信息化改造,实现数字化与互联互通,提高诊疗效率和医疗服务水平,具有重要的意义。

3 信息平台整体设计

3.1 主要业务流程

无纸化改造需要对原有业务流程进行重新梳理,实现流程再造。改造后的业务流程包括:①系统管理员维护项目基本信息、标本信息、人员信息等基础字典,并推送到收费系统、电子医嘱系统、护理系统等医疗业务终端;②医生站开立检查医嘱并自动生成检查申请单,收费处收费或护士站进行审核批费;③护士采集标本,工人运送标本到胎儿医学中心前台并与前台护士交接;④医护人员将标本上机,待采集平台获取检查结果后进行报告审核、发布;⑤信息平台将报告推送至患者App、自助机、电子病历系统、护理系统等应用端,实现患者自助服务以及医疗数据共享。信息平台业务流程如图1所示。

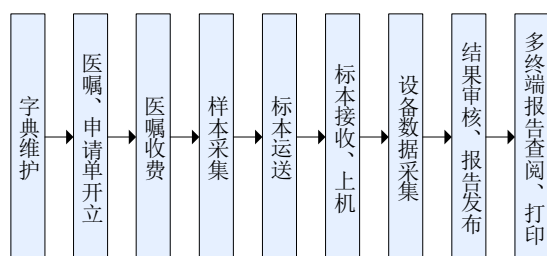


图1 信息平台主要业务流程

3.2 信息平台架构

信息平台架构主要包含三个层面。第一是基础设备和业务系统层,主要包含产前诊断实验室设备,以及挂号、住院、医嘱、护理、电子病历等基础医疗系统。第二是数据交互层,包含数据采集平台和HIE数据平台。数据采集平台基于物联网技术和接口通讯协议,从设备获取检查数据和结果,通过Web Service技术和XML标准将数据推送至信息平台。HIE数据平台则基于HL7、IHE等国际标准与规范,解决了不同结构系统间的信息交互和数据集成问题^[7],通过ESB技术实现消息传输和数据交换,把患者基本信息、费用信息、检查项目信息等数据传至胎儿中心信息平台,同时从信息平台获取检查报告、标本状态等数据并回传至其它医疗业务系统,实现业务数据的互联互通与数据共享。第三是胎儿中心信息平台业务逻辑和应用功能,通过单据电子化、数据闭环和共享、设备数据自动化采集等改造,达到全流程业务闭环和数据互联互通,实现字典、申请单、收费、标本、报告管理以及辅助决策功能,并通过HIE平台支撑全景电子病历、统一门户、患者App、自助机、医护工作站等上层医疗应用和患者服务,提高医院信息化和患者服务水平。信息平台架构如图2所示。

4 信息平台功能改造

为实现无纸化目标,信息平台通过现有问题分析、流程再造、架构设计,进行多项功能改造,包括单据电子化、数据闭环、数据互通及共享、决策支持等方面,消除信息孤岛,提高平台使用便利性,减少错收、漏收费用情况,便于追溯标本运送状态,监控报告发布情况,提高科室管理水平和减少医疗差错,构建智能胎儿医学中心信息平台。平台功能改造内容如图3所示。

4.1 单据电子化

单据电子化改造主要包括项目申请单、标本运送接单、检查报告单电子化改造。医生站在开立检查医嘱时,录入项目名称、标本类型、执行科室等信息后,系统自动调用并管规则,获取试管信息和采血次数,自动生成检查申请

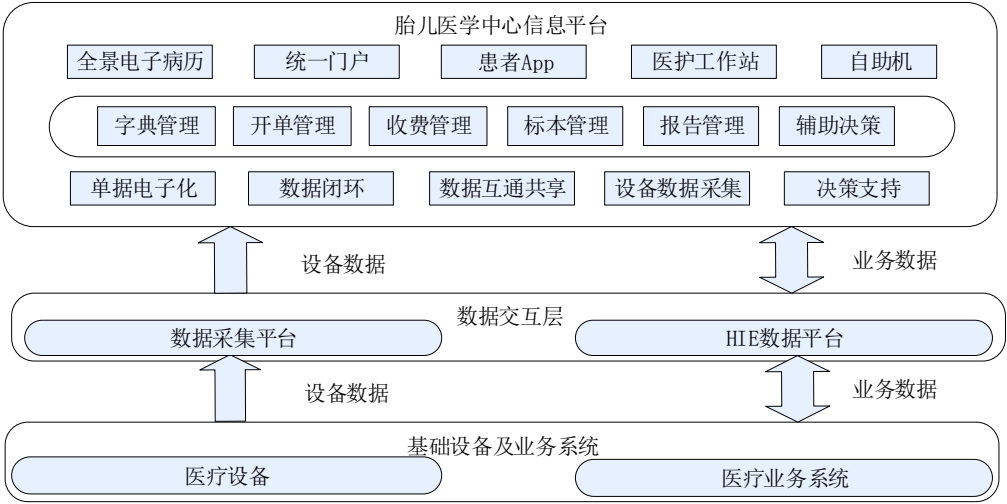


图 2 信息平台架构

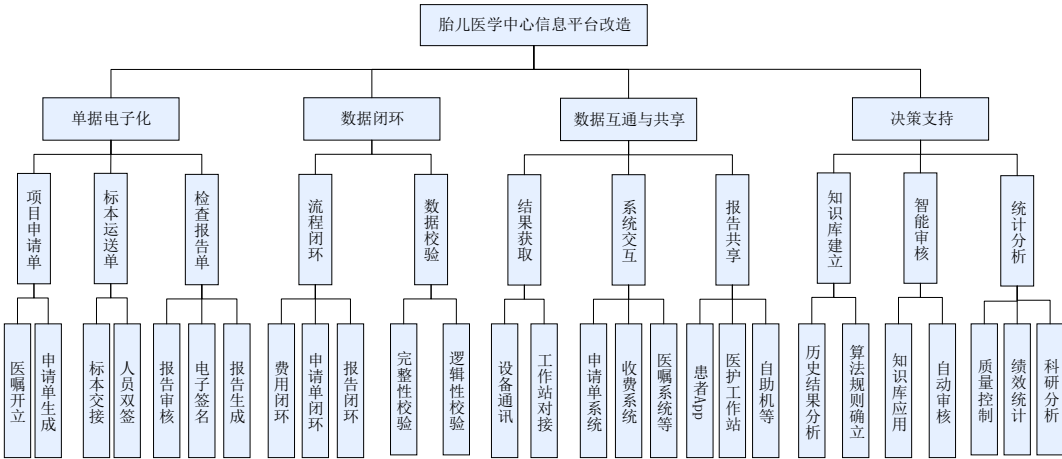


图 3 信息平台功能改造

单，并将项目费、试管费和采血费等发送到收费系统。护士站护士在采集标本后扫描试管条码并在系统打包，运送工人使用PDA扫描标本包后录入人员工号与护士进行签名确认，在送达胎儿医学中心前台后与前台人员进行标本交接双签，系统记录标本交接人员信息及标本状态。系统获取设备结果后形成检查报告，医生审核、批准后系统自动调用电子签名形成检查报告，生成报告的PDF文件。

4.2 数据闭环

数据闭环改造包括流程闭环和数据校验改

造。流程闭环由费用闭环、申请单闭环、报告闭环组成，医生开立医嘱后生成申请单，门诊患者凭申请单去收费处缴费，住院患者则由护士进行医嘱审核批费^[8]，收费后护士进行标本采集，采集完成后由工人运送到胎儿医学中心进行上机，系统生成报告并推送至临床医护工作站及患者 App，形成业务闭环。数据校验包括完整性校验和逻辑性校验，例如在标本运送时记录双方工号信息、交接时间、标本状态，在发布报告时记录审核批准人员电子签名、结果数据、异常信息等，形成全流程的数据轴，方便进行追溯和管理。逻辑性校验主要检查数

据是否符合业务和管理逻辑,对于不符合要求的操作进行提醒或者拦截,如未缴费或退费患者不能进行标本采集、上机、发报告等操作,已上机标本不能进行退费,能有效减少漏收、错收费用的情况。

4.3 数据互通与共享

数据互通与共享包括自动获取检查结果、系统交互以及报告共享等改造。系统获取设备检查数据有三种模式,第一是通过数据线和采集盒子接入医院内网,通过统一采集平台直接获取设备数据;第二是跟设备自带工作站交互,读取设备自带软件数据库或者共享文件内容得到检查结果;最后是医生利用设备自带软件审核生成报告 PDF,采集平台将 PDF 推送给信息平台,实现报告共享。系统交互主要是利用医院 HIE 数据平台与申请单系统、医嘱系统、收费系统进行交互,实现医疗数据的实时交换与共享。数据共享主要实现将医嘱信息、标本状态、检查报告等信息推送至医护工作站、自助机、患者 App 等终端,并形成临床数据仓库 CDR,为医疗、科研、管理提供数据支撑。

4.4 决策支持

决策支持主要包括知识库建立与应用、报告智能审核、统计分析等构建。医护人员根据项目信息、历史结果等内容,通过数据选择、分析、比较,以及差值比较、逻辑关系与关联性分析、一致性分析等规则,制定结果审核算法、标准和逻辑,并维护到知识库中,系统根据知识库定义自动对检查结果进行审核^[9]。在此过程中,当检查结果符合知识库预定义的标准时,系统自动将相符的结果输入相应的报告模板中,根据预设的模板内容自动生成报告,无需人工干预。当智能审核功能判断检查结果不满足预定要求后,对该数据加以记录,该信息将被保存,由人工完成相应的数据审核。若出现仪器设备更换、自动审核参数更改、信息系统更新等可能影响自动审核功能的变化时,由系统提示并对知识库进行校验,确定合格后方可继续使用^[10]。另外平台支持为科室的质量控制、绩效管理、科研分析提供数据报表,实现数据快速

检索、导出、打印,为科室提供数据支持。

5 改造效果及应用情况

信息平台改造基于 C# 编程语言、Microsoft SQL-Sever 数据库进行开发,开发环境为 Visual Studio^[11],对外服务接口采用 Web Service 技术。信息平台改造完成后无纸化程度有了较大提高,实现了数据闭环与互联互通,构建了智能胎儿医学中心信息平台,其功能包括:①字典管理:实现检查项目、标本类型、采集容器等信息的维护以及推送;②申请单及费用管理:实现申请单的自动生成、作废、执行,以及收费和退费的闭环管理以及费用校验;③标本管理:实现标本采集、打包、运送、上机等业务流程记录、状态上传,可在全景电子病历上查看其闭环时间轴;④报告管理:实现报告的产生、审核、批准、上传,以及多终端共享;⑤辅助决策:实现知识库的管理以及报告智能审核,并提供数据校验、结果验证、统计分析功能。

信息平台于 2022 年 6 月完成改造,截至 2023 年 7 月,共开立申请单 10766 张,产生标本包 1927 个,发布报告 8994 份,开单科室包括生殖中心、胎儿医学中心、儿科、妇产科等科室,节省纸张超过两万张,有效降低了科室运营成本。在使用效率方面,比起手工申请单,电子申请单效率提高 50%,标本运送对接时间减少 30%,报告发放时间减少 70%,有效提高了临床工作效率和患者服务水平。

6 结语

基于无纸化策略的胎儿医学中心信息平台改造,解决了原来信息孤岛多、纸质化单据多、容易漏收错收费用等问题,提高了医护人员工作效率,节省了科室运营成本,实现了报告的医护患共享,促进了智慧医院建设。但是在使用过程中发现仍有需要优化的地方,如部分设备数据无法直接获取,只能在设备自带工作站审核结果,然后再上传至医院信息平台;报告智能审核功能有待完善,辅助决策能力有待加强等,因此仍需不断优化平台功能,为临床提供更好的信息化服务。

参考文献：

- [1] 周祎,王子莲. 中国胎儿医学现况与地域发展差异[J]. 中国实用妇科与产科杂志, 2015, 31(12): 1102-1106.
- [2] 段涛. 中国胎儿医学的现状与未来发展方向[J]. 中国产前诊断杂志(电子版), 2011, 3(3): 1-3.
- [3] 付爱红,邱萍,潘芹. 胎儿低比率嵌合体产前诊断方法的比较[J]. 巴楚医学, 2023, 6(1): 59-63.
- [4] 钱云凤. 绿色低碳背景下高职院校无纸化办公的应用研究[J]. 办公室业务, 2023(12): 183-184.
- [5] 王浩,胡艳珂,郜勇. 医院移动电子签名的应用研究[J]. 中国卫生信息管理杂志, 2023, 20(3): 471-474, 481.
- [6] 张荣民,程庆林. 病案无纸化对公立医院精细化管理的作用[J]. 中国卫生信息管理杂志, 2022, 19(2): 256-259.
- [7] 莫远明,王毅,刘小洲,等. 基于五操作信息平台的多院区检验信息平台的构建[J]. 现代医院, 2020, 20(3): 407-410.
- [8] 卢瑞鸿,莫远明,陈康,等. 医院非常规检查全流程闭环管理信息化构建与应用[J]. 现代医院, 2023, 23(4): 614-618.
- [9] 俞华,陶正玄,赵英英. 围手术期智能临床辅助决策系统的构建与应用[J]. 中国卫生信息管理杂志, 2022, 19(6): 911-917.
- [10] 国家卫生健康委员会. 国家卫生健康委员会通告国卫通[2018]16号[J]. 中华人民共和国国家卫生健康委员会公报, 2018(8): 23.
- [11] 莫远明,卢瑞鸿,陈康. 临床分子诊断与基因检测信息平台的构建[J]. 现代计算机, 2022, 28(19): 106-110.

Reconstruction of fetal medical center information platform based on paperless office strategy

Mo Yuanming*, Qiao Mengwei, Lin Siyang, Li Peng'an

(Information Data Center of the First Affiliated Hospital of Sun Yat-sen University, Guangzhou 510080, China)

Abstract: This paper expounded the necessity of the transformation of the information platform of the fetal medical center based on the paperless office strategy. The paperless office transformation of the information platform was realized by using HIE data platform, electronic signature and other technologies. It introduced the main business processes, architecture, functions and applications of the information platform after the transformation. And the information platform realized the data closed-loop and interconnection by the paperless office transformation, which solved the problem of the original paper-based documents. The solution of easy omission and incorrect collection of fees has improved the efficiency of medical staff, reduced the department operating costs, improved patient service levels, and assisted in the implementation of smart healthcare.

Keywords: fetal medical center; paperless; HIE data platform; electronic signature; smart healthcare

文章编号: 1007-1423(2023)24-0108-05

DOI: 10.3969/j.issn.1007-1423.2023.24.019

基于 Arduino 的智能垃圾分类系统的设计与开发

张银霞*, 张新雨, 王冀伟, 聂鑫, 王鸣政, 王硕

(北京联合大学师范学院, 北京 100101)

摘要: 垃圾分类可以体现一个国家对经济循环发展、生态文明建设的谋篇布局。城市生活垃圾每年都在按一定的速度增长, 要解决“垃圾围城”问题, 就要进行生活垃圾分类, 变废为宝。随着人工智能科技的发展, 人们的智慧化生活水平也在不断提高, 可以设计开发一款智能垃圾分类系统, 使人们更加便捷准确地对生活垃圾进行分类。详细介绍了此款智能垃圾分类系统的结构设计、硬件平台、开发技术及核心功能实现的关键代码, 为此类开发设计提供参考模型。

关键词: Arduino; 垃圾分类; 语音识别; Arduino IDE; 串口通信

0 引言

我国城市垃圾的产量以每年约 8.98% 的速度在增长^[1], “垃圾围城”问题日渐突出。实际上城市生活垃圾中有很多资源是可以再回收利用的, 只要放对了位置, 垃圾就可以变废为宝。为了解决这一问题, 垃圾分类政策、垃圾管理条例和实施方案被相继发布, 各大重点城市开始了轰轰烈烈的垃圾分类宣传和活动, 号召人们进行垃圾分类。

但生活垃圾的品目较多, 人们记忆起来也不方便, 经常分错或者分不清楚, 因此希望设计一个智能垃圾分类系统, 利用智能科技帮助人们准确分类垃圾, 保障倡导垃圾分类活动的有效性。

1 系统分析与设计

1.1 系统功能结构设计

本系统要面向的使用场景是家庭生活垃圾

分类, 因此在进行系统的设计与开发之前, 对居民用户进行了实际的问卷调查, 参与调研的人群年龄范围在 18~50 岁之间, 其中约 69.23% 的用户希望此类系统具有自动翻盖功能, 71.79% 的用户表示要有垃圾识别的功能, 54.7% 的用户表示要进行语音交互, 此外还有 52.99% 的用户要智能可操作性的设置。

本系统在功能上包括了自动翻盖、垃圾识别、语音交互、并且可进行操作设置等功能。其 UML 用例图如图 1 所示, 用例规格说明见表 1。

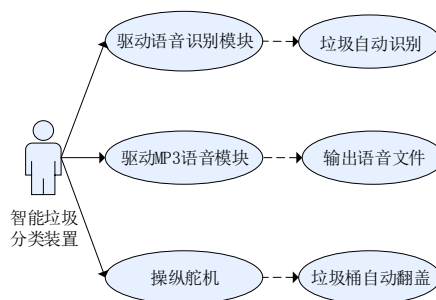


图 1 系统 UML 用例图

收稿日期: 2023-08-15 修稿日期: 2023-12-12

作者简介: *通信作者: 张银霞(1979—), 女, 河南郑州人, 硕士, 讲师, 研究方向为计算机网络、计算机软件设计与开发、Arduino 硬件开发, E-mail: zhangyinxia79@126.com; 张新雨(2000—), 男, 北京人, 本科生, 专业方向为计算机科学与技术(师范); 王冀伟(2003—), 男, 北京人, 本科生, 专业方向为计算机科学与技术(师范); 聂鑫(2002—), 男, 北京人, 本科生, 专业方向为计算机科学与技术(师范); 王鸣政(2002—), 男, 北京人, 本科生, 专业方向为计算机科学与技术(师范); 王硕(2003—), 男, 北京人, 本科生, 专业方向为计算机科学与技术(师范)

表 1 系统用例说明

参数	说明
执行者	智能垃圾分类系统
说明	通过此系统中的语音识别模块，此传感器模块可自动识别相应的语音，调用 MP3 语音播放传感器进行语音实时播放，并通过结合数据调用舵机角度，进行垃圾桶的自动翻盖
前置条件	Arduino 主板连接电源，进行数据导入输出
后置条件	智能垃圾分类系统完成语音识别，实现基础功能
主事件流	1) 通过语音识别模块进行语音的识别功能 2) 通过调用 MP3 语音播放传感器对语音文件进行输出 3) 通过舵机输出，对垃圾箱进行操作
备选事件流	E-1 线路连接中断，无识别内容

1.2 系统开发使用的设备和技术

本系统采用的开发板是 Arduino UNO，开发环境是 Arduino IDE 平台，开发语言为 C。还需要搭配 UNO R3 扩展板、ASR M08-B 语音识别模块、DFPlayer Mini MP3 模块、小喇叭等硬件设备。

1.2.1 Arduino 硬件开发板

Arduino 属于开放源码的电子原型平台，由各种型号的 Arduino 硬件板和配套的软件开发环境 Arduino IDE 两部分组成^[2]，可以接收来自各种传感器的输入信号，从而做到监测环境的效果，并通过控制光源、电机、马达以及其他执行器来反馈、影响环境。Arduino 硬件板有微控制器，控制功能是通过将编写的程序翻译成二进制文件烧录进微控制器中实现^[3]。

本系统采用 Arduino UNO 型号开发板。

1.2.2 Arduino IDE 开发环境

Arduino IDE 是基于 Processing IDE 开发的一个开放源码编程开发工具，兼容性强，可在 Windows、Linux 等操作系统上运行，以及 Arduino IDE 支持 Nano、Mega、UNO 等开发板，支持 C/C++、Python、Java 等主流编程语言。

1.2.3 Arduino UNO R3 扩展板

R3 扩展板可以引出大量扩展口、IIC 接口、串口，集成电机驱动和大电流电源管理，可以

更加方便地外接设备。

1.2.4 ASR M08-B 语音识别模块

ASR-08B 模块可以在 3 米以内进行非特定语音识别，识别率可以高达 85%~95%，TTL 串口通讯模式，接收语音信号的是其内置的高清麦克风。该模块有三种检测识别模式：循环检测、口令检测和按钮检测。本系统采用口令检测模式，实现智能垃圾分类的语音交互。

1.2.5 DFPlayer Mini MP3 模块

该模块与扬声器直接连接，与主控面板相结合，可以独立或串行控制。该模块将 MP3、WAV、WMA 等硬译码功能进行了很好的整合，用串行指令来实现音频文件的播放。本系统中，语音识别模块接收语音信息，一旦识别成功，就会进行语音文件的调用。

1.3 系统框架结构设计与搭建

智能垃圾分类的系统架构图如图 2 所示，主要包括 Arduino UNO 主控板、DFPlayer MP3 语音播放器、语音识别模块和小喇叭。

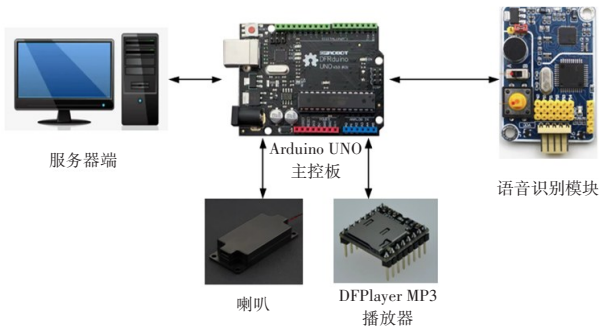


图 2 系统架构图

2 系统核心功能模块的设计与实现

2.1 业务流程设计

智能垃圾分类系统采用语音识别技术，其业务流程设计如图 3 所示，当连通电源后，唤醒硬件状态，通过关键词唤醒语音模块，若成功唤醒，语音识别模块则可说出垃圾名称；若唤醒不成功，需等待再次唤醒。语音识别成功后，系统调用 MP3 语音播放模块对相应垃圾名称进行语音文件输出，并操纵舵机转动，从而达到垃圾分类效果。

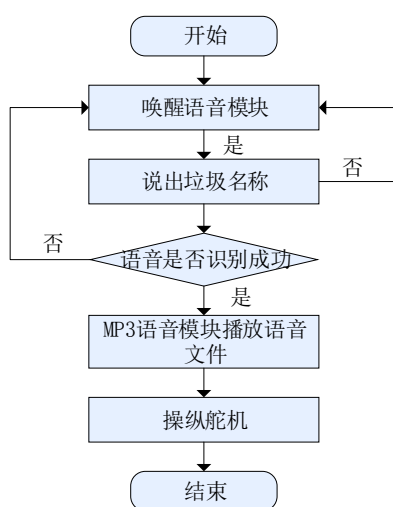


图3 系统业务流程图

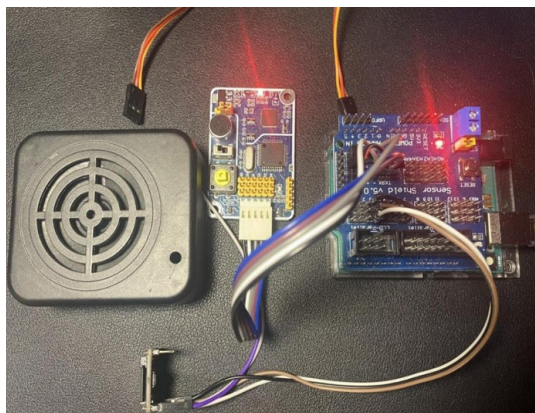


图4 系统实物连接图

2.2 智能垃圾分类系统功能的实现

2.2.1 实物连线

智能垃圾分类系统的实物连接效果如图4所示，语音识别模块对输入的声音进行识别，并将函数值通过主控板调用MP3语音播放传感器，输出语音文件，实现语音交互。根据返回值，再调用舵机实现垃圾分类和自动翻盖功能。

语音识别模块：内置5个引脚，GND、3.3 V电源引脚、5 V电源引脚、RX引脚和TX引脚。本系统中采用GND、5 V电源、RX引脚和TX引脚。该串口使用的UART通讯协议，其最大的优点是不需要时钟信号，只需要通过数据通讯线路就能进行数据传输。

MP3模块：引脚采用VCC引脚、RX引脚、TX引脚、SPK1引脚（驱动喇叭）、GND。通过串口连接扬声器，该模块支持异步串行通信方式，用串口接收控制指令。

2.2.2 设置编写语音识别口令

本系统的语音识别指令如图5所示。

```

@01,su liao ping,001,$
@02,fei tie,001,$
@03,bo li ping,001,$
@04,dian shi ji,001,$
@05,fei shui,002,$
@06,wei sheng zhi,002,$
@07,yan tou,002,$
@08,ya qian,002,$
@09,sheng cai,003,$
@10,guo pi,003,$
@11,ji dan ke,003,$
@12,ping guo he,003,$
@13,dian chi,004,$
@14,you qi tong,004,$
@15,lao shu yao,004,$
@16,nong yao,004,$

@49,xiao ding dang,254,$

@41,chu yu la ji,255,$
@42,you hai la ji,255,$
@43,ke hui shou la ji,255,$
@44,qi ta la ji,255,$
  
```

图5 垃圾分类语音识别指令

2.2.3 软件发送语音识别口令

通过ASR M08-B设置软件对模块进行关键词设置，并发送至语音识别模块的实物中，发送垃圾名称口令的软件窗口如图6所示，将图5所示的口令代码输入在图6所示的“数据发送区”中，点击“发送”按钮，即可发送。



图6 语音识别模块设置指令图

2.2.4 在 Arduino IDE 中输入智能垃圾分类装置核心代码

设置串口通讯的核心代码如图7所示。

```
void setup () {
  //设置4路舵机的引脚
  servo_4.attach(4);
  servo_7.attach(7);
  servo_10.attach(10);
  servo_12.attach(12);
  //设置舵机的初始角度
  servo_4.write(0);
  servo_7.write(0);
  servo_10.write(0);
  servo_12.write(0);
  //设置串口通讯
  Serial.begin(9600);
  mySerial.begin(9600);
  mp3_set_serial(mySerial);
  delay(1);
  mp3_set_volume(50); //设置mp3音量
}
```

图7 设置串口通讯图

编写变量对返回值进行判断,操纵舵机与MP3语音播放模块,如图8所示。

```
void loop () {
  int val = Serial.read(); //定义变量,并把串口值赋给这个变量
  if (val == 1) {
    mp3_play_physical(1); //播放第1个声音文件
    digitalWrite(3, LOW);
    servo_4.write(60); //设置舵机转动角度
    delay(2000);
    servo_4.write(0); //设置舵机回到初始角度
    digitalWrite(3, HIGH);
  }
  if (val == 2) {
    mp3_play_physical(2);
    digitalWrite(3, LOW);
    servo_7.write(60);
    delay(2000);
    servo_7.write(0);
    digitalWrite(3, HIGH);
  }
}
```

图8 识别操纵舵机与MP3模块代码图

3 系统运行效果展示

通过口令语音唤醒语音模块,语音识别模块被唤醒后会亮起蓝灯,根据输入不同种类的垃圾语音信息,识别并自动打开不同种类的垃圾桶盖子,实现结果如图9所示。

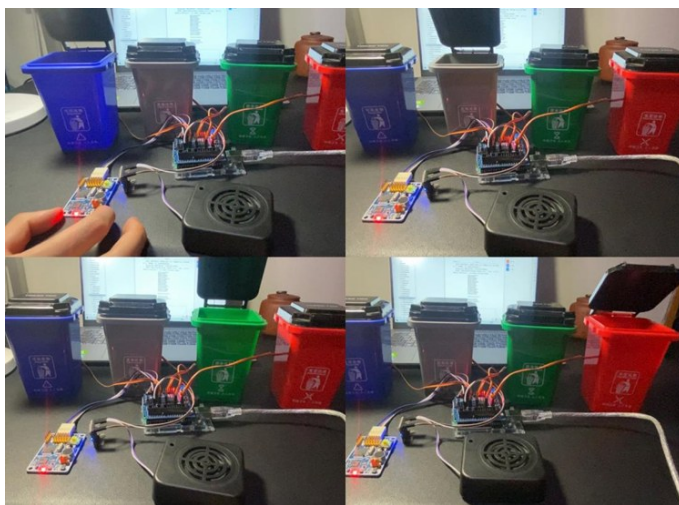


图9 不同垃圾种类的语音信息控制不同的垃圾桶自动打开

4 结语

本系统基于对居民生活垃圾分类的实际需求调研,利用语音识别技术,实现人和智能垃圾桶之间的语音交互,帮助进行垃圾的自动分类、垃圾桶的自动翻盖功能,解决居民垃圾分类中的识别困难和投放错误等实际问题,使居民生活更加有序化,促进环境资源优化。选用的 Arduino 开源硬件主板,可扩展功能强,后续可以

根据需要进行其他功能。

参考文献:

- [1] 杨海东. 城市垃圾处理及资源化应用[J]. 科技风, 2014(6):274.
- [2] 孟宪宁. 在 Arduino 平台上快速实现创意作品的原型设计[J]. 单片机与嵌入式系统应用, 2020(7): 40-42.
- [3] 胡宸晗. 基于 Arduino 和压电传感器的种子自动计数系统设计[J]. 电子制作, 2019(1):29-31.

(下转第116页)

文章编号: 1007-1423(2023)24-0112-05

DOI: 10.3969/j.issn.1007-1423.2023.24.020

一种微服务框架的浙江气象应用开放服务技术研究与实现

曾 悠, 封秀燕*, 杨 明

(浙江省气象信息网络中心, 杭州 310051)

摘要: 为了解决传统气象应用开发面临的开发周期长、技术多样化、移植难等问题, 采用微服务框架技术对浙江气象应用服务开放平台进行研究与实现, 阐述了平台开发中总体框架、系统架构及核心部件的设计, 通过对关键性微服务业务模块的研究, 形成了符合业务需求的微服务业务架构, 提高了气象应用服务的信息化支撑能力。

关键词: 微服务框架; 智慧气象; 气象应用; Spring Cloud; 开放平台

0 引言

伴随智能终端和互联网的高速发展, 智慧气象的发展所带来的智能观测、精准预报、开放服务为人们带来了便利^[1]。同时, 气象海量大数据以及社会对气象服务的需求日益增长, 使各类气象智能服务软件和应用的复杂度、对高并发性的要求都大幅提升, 逐渐呈现出“大平台、微开发”的特征^[2], 而传统气象应用的开发强依赖于开发人员的技术, 整个开发过程需经历数据存储、数据处理、数据信息展示等一系列环节。除开发周期长, 还面临着技术多样化、封闭性强、移植难等问题, 已无法满足当前气象发展对跨领域融合、开发效率、产品迭代频繁的高要求。因此, 为了提高气象数据服务的使用效率、降低业务人员软件开发门槛, 本文基于 Spring Cloud 微服务框架, 研究实现了一套集计算、组件、数据、地图和应用服务为一体的浙江气象应用开放服务平台。平台面向省、市、县气象行业业务人员或外部门的专业人士, 将各类气象功能模块以微服务的形式作为可复用的单元开放调用, 用户也可以上传共享本地化功能或服务模块, 实现气象应用的共享、共创、共建。

1 微服务架构

微服务(Micro Service)伴随着云计算、虚拟化技术的兴起和发展应运而生, 最初在 2011 年的一次软件架构研讨会上被讨论, 之后以 Martin Fowler 为代表的软件开发专家于 2014 年正式提出微服务架构^[3]。微服务架构^[4]是相对于单体架构^[5]而言的。基于单体架构的软件一般作为一个整体构建, 处理请求的所有逻辑都是单程运行的, 这意味着后期程序所需的任何更改都可能影响着整体的构建和部署, 造成系统冗余、难以维护等问题^[6]。而基于 SOA 架构思想^[7], 微服务架构是围绕业务功能组织服务, 将复杂庞大的系统拆分成多个体量更小的独立服务单元, 支持服务单元的独立开发、单独部署, 实现去中心化管理, 可以看做是对单体结构的一种解耦。因此, 对于拥有海量数据、业务逻辑复杂度高、实时性和伸缩性运行要求都高标准的气象服务系统应用来说, 微服务架构具有多项优势, 主要体现在^[8-11]:

(1) 开发维护更便捷: 微服务架构能够实现功能模块间的相互独立, 可对功能模块进行

收稿日期: 2023-08-14 修稿日期: 2023-11-30

作者简介: 曾悠(1989—), 女, 湖北宜昌人, 硕士研究生, 工程师, 研究方向为气象数据可视化研发; *通信作者: 封秀燕(1966—), 女, 浙江丽水人, 本科, 正研, 研究方向为地面气象资料质量控制技术和成果业务化应用, E-mail: fxy945@sohu.com; 杨明(1983—), 男, 江苏宿迁人, 硕士研究生, 高级工程师, 研究方向为气象信息技术与大数据系统

单独开发或维护，便于统筹分配时间和人员。

(2) 技术方案灵活性更好：不再只支持单一技术或平台，不同的服务单元可以根据功能模块需求选用合适的技术，有利于技术协同。

(3) 独立部署，可持续性更强：每个服务单元可作为独立模块开发、迭代或部署，无需对应用系统进行整体部署，缩短总体开发周期，易于控制成本，实现可持续性集成。

(4) 扩展性更强：每个服务单元可根据业务需求局部、有针对性地扩展，不影响应用系统总体的运行逻辑与功能。

(5) 容错能力更强：出现故障不需要对整体系统进行更改，只需更新相对应功能模块的微服务，可通过超时重试、多副本策略等机制实现应用层面的容错，避免全局不可用^[12]。

2 微服务架构的浙江气象应用开放服务平台

基于气象大数据云平台，浙江气象应用开放服务平台为用户提供各类服务和功能，在逻辑上通过数据收集微服务集、处理与计算微服务集、服务发布微服务集三部分实现，采用微服务架构，每个微服务集又分别由多个功能相近的微服务组合而成，以实现一系列功能。按照平台制定的运行标准规范，使用服务接口实现微服务之间的相互调用。浙江气象应用开放服务平台服务架构如图1所示。

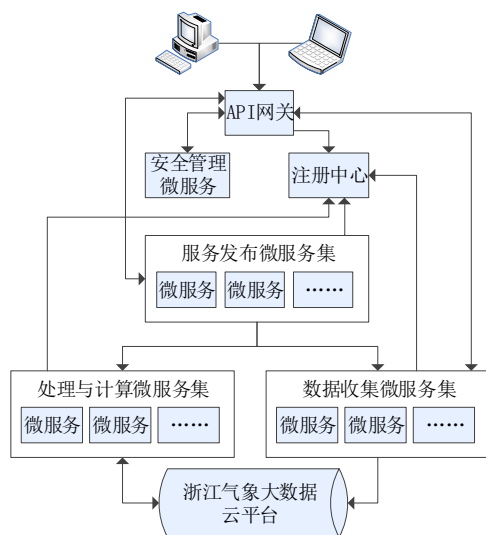


图1 浙江气象应用开放服务平台服务框架图

首先，注册中心完成对数据收集微服务集、处理与计算微服务集和服务发布微服务集三大模块的统一注册；然后根据用户的指令要求，经作为所有请求入口的API网关和安全管理微服务进行所需微服务的重组；最后，按照平台规定的规范，完成数据的获取、处理、计算和加工等，最终结果输出到用户端。

2.1 系统架构设计及核心部件

由于每个微服务在独立运行的同时又需相互配合，则微服务之间的通讯机制、服务发现机制以及统一的管理等都需要一套完整的框架实现。本文引入目前主流、综合性较强的Spring Cloud开源框架，它采用分布式服务思想，提供了构建基于微服务架构系统所需的各类功能和模块，主要通过其核心组件来实现，包括配置中心、服务调用、服务网关、服务熔断、安全控制、注册中心、负载均衡等等，具有跨平台、数据重用、易部署、易维护等特点。

基于Spring Cloud框架，本文为浙江气象应用开放服务平台设计了一套数据调用和接口的标准，定义了各个微服务之间的数据交换、查询获取等交互标准，为平台提供数据支撑。开放平台依靠Spring Cloud框架的核心组件实现了多项主要功能^[13]，包括：

(1) 利用SpringCloud Config对每个微服务集进行动态、集中的配置管理。

(2) 通过Zuul组件实现外部服务请求的统一入口，提供路由、过滤(权限验证、API监控等)和智能路由对具体服务等调用等功能。

(3) 利用服务中心Eureka组件，实现微服务调用发现和微服务注册，并提供服务查询获取、心跳检查机制及管理功能。

(4) 利用Hystrix熔断器，实现服务超时重试、限流操作、服务执行控制以及故障隔离等机制和功能。

(5) 利用Ribbon负载均衡组件，与Eureka组件共同部署，结合服务发现达到动态调用、负载均衡的目的。

在开放平台微服务架构下，面向用户的是

服务发布微服务集，其他微服务集和微服务在内部调用，无需展示在用户面前。

2.2 微服务业务架构

浙江气象应用开放服务平台面临着海量历史或实时数据、计算或算法、组件或模块等高并发访问和调用等问题，对数据存储和处理能力较高。本文基于浙江气象大数据云平台的分布式云存储，对需提供的服务进行分析、分解和再组合，根据数据收集、处理计算和服务发布三大模块，得到一系列独立的微服务和微服务集，形成符合业务需求的微服务业务架构。

2.2.1 数据收集微服务集

应用开放平台的数据来源主要包括浙江省自动站观测、数值预报、卫星、雷达、地理信息数据等，具有种类全面多元、精度高、信息量大、异构性强等特点，传统的关系型数据库在数据存储、读取方面已满足不了平台对访问速度和使用性能方面的要求。本文采用云关系型数据存储、NoSQL对象库存储和气象数据处理技术^[14]，实现对结构化和半/非结构化气象数据的分类存储。数据收集微服务集由基础数据、地图数据和数据存储三个微服务组成，如图2所示。

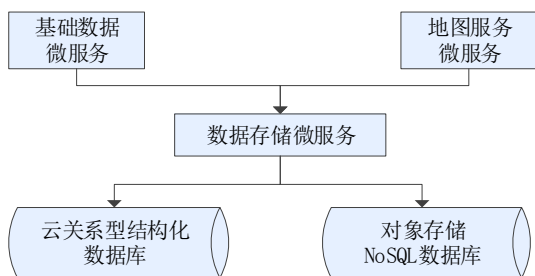


图2 数据收集微服务集结构图

2.2.2 处理与计算微服务集

根据不同的业务场景和业务需求，应用开放平台需满足用户对数据、计算、地图等各类气象服务的获取调用。对处理与计算微服务集进行微服务的拆分，可分为数据微服务、地图

微服务、可视化微服务和加工微服务。

其中，数据微服务为平台提供结构化和非/半结构化数据逾200多类，包括地面、高空、农气、雷达、卫星数据等等；地图微服务提供气象地图、卫星影像地图和地形图等基础地图，以及部分流域、水库等地理信息服务；可视化微服务提供气象等值线和等值面、流场以及各类显示空间等；加工服务提供气象等值线和等值面的制作、卫星、雷达产品制作等服务。

2.2.3 服务发布微服务集

服务发布微服务集直接面向用户。用户向平台提出服务使用申请后，经API网关合法性和使用权限审核通过，根据用户需求和业务逻辑，调用相匹配的数据服务进行处理和计算，以前端的可视化展示为反馈结果，并以API接口的方式提供给用户调用。同时，用户也可通过平台经合法性审核后发布自主研发的本地化气象服务，并开放相应权限供平台其他用户使用。以此实现浙江气象应用开放服务平台各类气象服务和应用的共享、共创、共建。

平台提供的功能有：数据服务，包括地面数据、高空数据、大气成分、卫星、雷达数据等；计算服务，包括面雨量计算、制图计算、科学计算、气象分析等；应用服务，包括卫星应用、雷达应用、Web应用、决策应用服务等；组件服务，包括地图组件、接口组件、算法组件、数据组件等；模块服务，包括专题服务、气象预警、天气预报等等。

3 应用实现

本文基于上述技术路线和方法，利用Spring Cloud微服务框架，实现了集气象数据收集、处理与计算、服务发布为一体的浙江气象应用开放服务平台。应用结果表明，该技术方案具有良好的使用体验、应用效果和较高的应用价值。应用开放平台前端建设首页效果如图3所示。



图3 浙江气象应用开放服务平台首页前端页面效果图

4 结语

“大平台、微服务”是科技高速发展的背景下，气象行业面对海量数据、跨行业融合以及数据安全高要求所提出的重要技术途径之一。本文针对传统气象应用软件开发所面临的复用率低、开发周期长、维护难等问题，提出了一种基于微服务架构的浙江气象应用开放服务平台的研究与实现。利用微服务架构的优势，采用Spring Cloud微服务框架解决方案，将气象应用开发所涉及的气象数据存储和获取、数据处理和计算、应用部署和发布等流程设计为三大微服务集，介绍了三大主要模块的功能、作用以及平台整体架构，最后介绍了气象应用开放平台首页的前端展示效果。微服务架构思想为气象应用的快速开发以及共享、共创提供了一种解决思路和实现途径，提高了气象应用服务的信息化支撑能力。

参考文献：

- [1] 刘莎. 智慧气象内涵及发展思路[J]. 陕西气象, 2018(5):37-39.
- [2] 刘魁, 舒红平, 何文春, 等. 基于领域知识的气象智能服务软件微开发技术及应用[J]. 气象科技进展, 2018, 8(1):92-96.
- [3] LEWIS J, FOWLER M. Microservices [EB/OL]. (2014-03-25) [2017-11-07]. <http://martinfowler.com/articles/microservices.html>.
- [4] RICHARDSON C. Monolithic architecture [EB/OL]. [2021-04-02]. <https://microservices.io/patterns/monolithic.html>.
- [5] 张晶, 王琰洁, 黄小锋. 一种微服务框架的实现[J]. 计算机系统应用, 2017, 26(4):82-86.
- [6] 江郑, 王俊丽, 曹芮浩, 等. 一种基于微服务架构的服务划分方法[J]. 计算机科学, 2021(12):17-23.
- [7] DRAGONI N, GIALLORRINZO S, LAFUENTE A, et al. Microservices: yesterday, today, and tomorrow [J]. Present and Ulterior Software Engineering, 2017, 4:195-216.
- [8] 吴化尧, 邓文俊. 面向微服务软件开发方法研究进展[J]. 计算机研究与发展, 2020, 57(3):525-541.
- [9] THÖNES J. Microservices[J]. IEEE Software, 2015, 32(1):116.
- [10] 辛园园, 钮俊, 谢志军, 等. 微服务体系结构实现框架综述[J]. 计算机工程与应用, 2018, 54(19):10-17.
- [11] 刘樱, 杨明, 徐集云, 等. 一种微服务框架的暴雨强度服务平台设计与实现[J]. 现代计算机, 2019, 25(18):91-94.
- [12] 黄小锋, 张晶. 微服务框架介绍与实现[J]. 电脑与信息技术, 2016, 24(6):14-16.
- [13] 龙新征, 彭一明, 李若森. 基于微服务框架的信息服务平台[J]. 东南大学学报(自然科学版), 2017, 47(A01):48-52.
- [14] 杨明, 陈晔峰, 陈晴, 等. 气象数据云数据存储技术及应用[J]. 气象科技, 2017, 45(6):1017-1021.

Research and implementation of Zhejiang open service technology for meteorological application based on microservice framework

Zeng You, Feng Xiuyan*, Yang Ming

(Zhejiang Meteorological Information Network Center, Hangzhou 310051, China)

Abstract: The development of traditional meteorological applications often depends on the technical ability of developers, which is at great expense. In order to solve the problems faced by traditional meteorological application development, such as long development period, diversified technology and difficult transplantation, the microservice framework technology is used to study and implement Zhejiang open platform of meteorological application service. Firstly, the overall framework, system architecture and core components of the platform are instructed. Through the study of key microservice business modules, a microservice business architecture that meets business needs has been formed, and the information support capability of meteorological application services has been improved.

Keywords: Microservices framework; smart weather; meteorological application; Spring Cloud; open platform

(上接第 111 页)

Design and development of an intelligent garbage classification system based on Arduino

Zhang Yinxia*, Zhang Xinyu, Wang Jiwei, Nie Xin, Wang Mingzheng, Wang Shuo

(Teacher's College of Beijing Union University, Beijing 100101, China)

Abstract: Garbage classification can reflect a country's strategic layout for the economic circular development and the ecological civilization construction. Urban household waste is increasing at a certain rate every year. To solve the problem of "garbage siege", it is necessary to classify household waste and turn it into treasure. With the development of the artificial intelligence technology, people's intelligent living standards are also constantly improving. It is possible to design and develop an intelligent garbage classification system, making it more convenient and accurate for people to classify the household waste. The article provides a detailed introduction to the system architecture design, the hardware platform, the development technology, and the key code for implementing core functions of this intelligent garbage classification system, and provides a reference model for such developments and designs.

Keywords: Arduino; garbage classification; speech recognition; Arduino IDE; serial communication

文章编号: 1007-1423(2023)24-0117-04

DOI: 10.3969/j.issn.1007-1423.2023.24.021

基于 SeetaFace6 的人脸识别信息管理系统

李 娟*, 杨溢超

(福州理工学院计算与信息科学学院, 福州 350000)

摘要: 基于人脸识别的信息管理系统能实现身份信息的高效化管理, 但传统的人脸识别信息管理系统仍存在诸多问题, 如人脸图像质量要求高、系统运行速度慢等。基于 SeetaFace6 设计与实现了人脸识别信息管理系统, 系统采用轻量级、易部署的 SeetaFace6 人脸识别引擎实现人脸检测、跟踪、注册与识别等信息管理与可视化分析, 具有较高的准确率和快速响应能力, 可应用于工地人员进出通行、考勤打卡、安全监控等方面。

关键词: 人脸识别; 信息管理系统; SeetaFace6; 深度学习

0 引言

人脸识别信息管理系统是企业单位日常人员管理的重要组成部分, 主要包括人员考勤、人员信息管理、人员数据分析等。传统的人脸识别信息管理系统广义只包括人员考勤模块, 且存在对人脸图像质量要求高、要求被识别人员保持固定姿势、系统运行速度慢、不能跨平台使用、功能单一等诸多问题^[1-3]。本文基于 SeetaFace6^[4]设计开发了一款人脸识别信息管理系统, 该系统使用 OpenCV 对视频流数据进行处理, SeetaFace6 对视频流中帧图像进行实时人脸检测、人脸跟踪、人脸识别, 对人脸识别数据进行信息管理以及数据可视化, 特别地, 系统还能处理戴口罩人脸检测与戴口罩人脸识别, 实现了对单位部门员工的信息管理, 能够方便地进行人脸注册、人脸识别、数据汇总、数据可视化等工作, 满足诸如智慧工地的人脸信息管理系统的需求, 提升信息管理水平, 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

设计开发的人脸识别信息管理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。

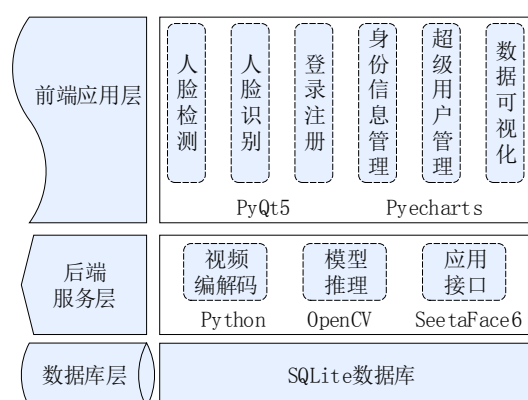


图 1 系统整体结构

前端应用层使用 PyQt5^[5]开发页面, 方便用户与系统的交互, 包括人脸注册、识别和检索等; 使用 Pyecharts^[6]的图形报表展示不同岗位的

收稿日期: 2023-12-04 修稿日期: 2023-12-15

基金项目: 2021 年福州理工学院校级科研基金项目 (FTKY21047); 2021 年度福建省中青年教育科研项目 (JAT210607); 2022 年移动通信和物联网福建省高校工程研究中心课题 (Kfkt202202)

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; 杨溢超 (2001—), 男, 河南许昌人, 在读本科, 研究方向为智能系统

考勤信息,方便用户了解系统运行状态和人员信息情况。后端服务层使用Python语言进行开发,主要聚焦人脸处理的各种任务,包括OpenCV获取帧图像、SeetaFace6的人脸检测、人脸跟踪、人脸关键点定位与遮挡识别、人脸特征提取、人脸比对等,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SQLite存储用户的系统信息,实现对用户信息、人脸图片存储路径、人脸特征等数据的管理和维护。

1.2 系统功能结构

设计开发的人脸识别信息管理系统包括人脸识别、登录注册和信息管理三个模块。其中,人脸识别模块涵盖了人脸检测、人脸跟踪、人脸注册、人脸特征提取、人脸比对以及陌生人人脸检测等多个功能,登录注册模块包括权限分配,面对超级管理员用户开放管理系统中用户的权限设定与修改,信息管理模块主要针对管理员用户开放,包括数据查询与修改、陌生访客记录、考勤信息统计、历史记录查阅、数据分析大屏展示。

2 主体功能的实现

系统工作流程上,人脸识别信息管理系统的管理员需要先注册本单位各部门员工的基本信息,可直接通过系统上的人脸注册模块的摄像头现场采集人脸信息和填写个人信息进行注册。人脸注册的人脸特征信息和个人信息将持久化保存到数据库中,作为后续人脸识别1:N的

人脸库。

2.1 人脸识别

设计开发的系统可实时从视频中进行多人人脸检测,并对最大人脸进行人脸跟踪与对齐,并提取人脸区域的特征向量,将提取的特征向量与系统中存储的人脸特征进行匹配,通过预设的阈值与相似度比对判断是否匹配成功,以确认人物身份。特别地,在提取人脸区域的特征向量前,系统对对齐的人脸进行有无遮挡识别,无遮挡的情况采用全脸区域的特征提取;有遮挡的情况采用口罩遮挡模型提取未佩戴口罩部分的特征,将其与已存储的未佩戴口罩部分特征进行匹配,具体流程如图2所示。

2.1.1 人脸检测与人脸跟踪

人脸检测是从OpenCV中解码获取的彩色图像或者灰度图像定位人脸信息,返回所有检测到的人脸边界框(Bounding Box)及该人脸的置信分数。人脸跟踪器则从连续传入的图像中进行人脸跟踪,返回所有跟踪到的人脸信息,包括人脸位置、人脸置信分数、视频帧的索引以及跟踪的人脸标识ID。实际采用的SeetaFace6的人脸检测器使用RetinaFace算法,基于MobileNet0.25骨干网络,使用掺入50%带口罩人脸的检测数据进行训练,使得人脸检测器可以实时从视频帧中同时检测戴口罩和未戴口罩人脸。通过提取人脸边界框坐标计算检出每个人脸的面积,并标记为关注人脸,记录人脸跟踪器对该人脸的标识ID并持续跟踪该人脸。

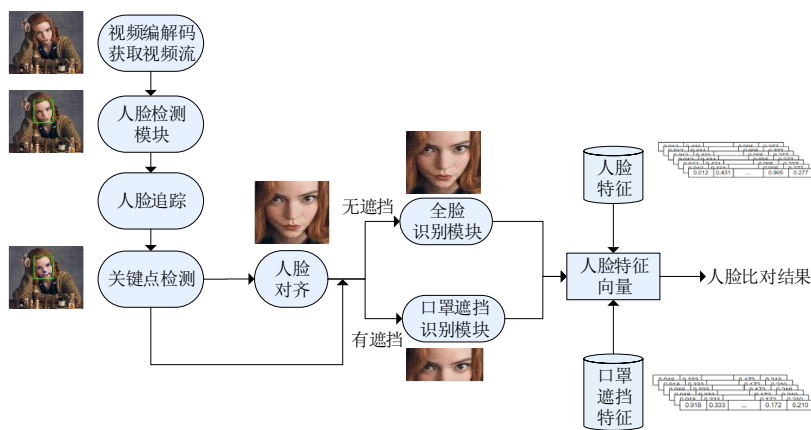


图2 人脸识别流程

2.1.2 人脸关键点检测与人脸对齐

采用SeetaFace6的人脸关键点检测和口罩遮挡是一个多任务模型，在传统的五个人脸关键点算法基础上，增加一个新任务用于人脸遮挡物分类(未遮挡，口罩遮挡)的识别，训练数据组成上增加了20%比例的戴口罩数据，最终得到可以同时进行关键点定位和口罩遮挡识别的模型。人脸关键点检测模型在最大人脸提取左眼、右眼、鼻、左嘴角、右嘴角五个关键点检测并经过仿射变换得到对齐的人脸图像，分辨率为256×256。

2.1.3 人脸特征提取与人脸比对

采用SeetaFace6的戴口罩人脸特征提取模型基于ResNet50骨干网络，其在新增了20%以上戴口罩的人脸识别数据集，将分辨率为256×256的对齐人脸裁剪上半部分的256×128的新人脸图像数据集上进行训练得到。识别过程则根据人脸关键点定位与口罩遮挡识别模型判断是否有遮挡，如果判断为有遮挡则使用戴口罩的识别模型进行特征提取，否则使用常规识别模型进行特征提取，两者提取到的人脸特征向量均为1024维。由于系统中已经保存了所有人员的人脸图像，根据提取的待识别人员人脸特征与

所有人员人脸特征进行1:N余弦相似度比较，以确认人物身份。

2.2 信息管理

2.2.1 数据管理

为了使系统更稳定、健壮地运行，分别设计了：①用户管理模块，允许超级管理员账户来管理普通用户和管理员用户的权限，能够进行相应权限的修改与删除操作；②数据管理模块，允许用户对数据进行检索、修改、删除、导出等操作。

2.2.2 数据分析大屏展示

为了让管理员用户实时了解系统运行状态和数据情况，以便为管理者提供决策分析，通过图表、报表等形式展示特定岗位的信息考勤记录和记录等信息，具体包括：到岗情况、重点岗位、访客情况、天气状态预报、人员流动量情况、考勤通行记录等。

3 系统测试运行

系统的核心模块为视频帧中的人脸检测与识别，其余模块为基于人脸识别结果的数据管理和数据可视化功能。系统能在CPU或嵌入式等轻量级的算力平台上实现实时检测识别，部分系统运行界面如图3所示。



图3 系统部分运行界面

4 结语

基于 SeetaFace6 设计实现了一款人脸识别信息管理系统, 该系统通过对人脸特征的准确提取和匹配识别实现了对人物身份的快速确认; 人脸识别引擎对硬件配置要求不高, 可在小算力移动终端运行; 系统还支持对身份信息的高效化管理, 可以为各种数据分析提供便利, 可应用于智慧工地实现自动化管理, 保障工地安全。

参考文献:

- [1] 李雄, 文开福, 钟小明, 等. 基于深度学习的人脸识别考勤管理系统开发[J]. 实验室研究与探索, 2019, 38(7): 115-118, 123.
- [2] 王琴, 米丹, 周丽. 基于 CNN 的人脸识别课堂考勤系统[J]. 现代计算机, 2019(25): 88-91.
- [3] 徐建峰, 孙浩, 陆萍, 等. 基于深度学习的人脸识别会议考勤系统开发[J]. 现代计算机, 2020(13): 86-89.
- [4] Seetatech. SeetaFace6 [Source code], 2021. <https://github.com/seetafaceengine/SeetaFace6>.
- [5] WILLMAN J. Modern PyQt: create GUI applications for project management, computer vision, and data analysis[M]. 1st ed. Berkeley, CA, Apress, 2021: 1-42.
- [6] CAO S, ZENG Y, YANG S, et al. Research on python data visualization technology [J]. Journal of Physics: Conference Series, 2021, 1757(1): 012122.

Face recognition information management system based on SeetaFace6

Li Juan*, Yang Yichao

(School of Computing and Information Science, Fuzhou Institute of Technology, Fuzhou 350000, China)

Abstract: Information management system based on face recognition can realize efficient management of identity information, but the traditional face recognition information management system still has many problems, such as high requirements for face image quality, slow system operation speed, etc.. The face recognition information management system is designed and implemented based on SeetaFace6, which adopts the light and easy-to-deploy SeetaFace6 face recognition engine to realize face detection, tracking, registration and identification information management and visualization analysis, with high accuracy and fast response capability, it can be applied to the in-and-out access of construction sites, attendance check-in, security monitoring.

Keywords: face recognition; information management system; SeetaFace6; deep learning