

M 现代计算机

XIANDAI JISUANJI

第28卷第19期 (总第763期)

半月刊 (1984年创刊)

2022年10月10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

ISSN 1007-1423



9 771007 142222 19

M 现代计算机

第28卷 第19期 (总第763期)

2022年10月

2022年10月 第28卷

第19期 (总第763期)

M

ISSN 1007-1423
CN 44-1415/TP

MODERN COMPUTER

现代计算机



中山大学出版社 主办

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

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

版权声明

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

目次

研究与开发

基于迁移学习与标签平滑策略的宫颈细胞分类方法..... 刘 美, 周 龙 (1)

中国养老服务床位数的预测与分析..... 唐鹏翔, 张晓美 (10)

基于迁移学习的食道癌影像识别研究..... 凌其颖, 刘小芳, 张 阳 (19)

一种基于单形空间缺失数据的补全方法..... 刘 冰, 李玮琦 (26)

基于YOLO框架的农田火源自动检测系统
..... 白 玉, 马广焜, 彭新茗, 王楠翔, 白清扬, 孟峻锋, 刘鑫蕊, 朱硕闻 (33)

角钢印字符识别方法的研究综述..... 何 迅, 刘小芳 (39)

基于SIFT算法的人群密集检测应用研究 郭 妍, 谢光敏, 向蓓蓓, 古荣龙, 王 磊 (45)

基于四足仿生机器人快递配送的研究
..... 靳 晨, 刘 刚, 吕兴东, 黎航宇, 杨雅子琪, 邓艺霖 (51)

面向威胁信息的网络安全态势感知研究..... 谷晓鹏 (57)

基于支持向量机的微博情感分析方法研究..... 李首政, 王 琪, 王 力 (63)

基于U-net算法和FCOS目标检测算法的洗衣机外壳钣金孔洞缺陷检测系统
..... 邓秋菊, 张振龙 (67)

基于AI深度学习的病害识别方法研究 刘小红 (71)

实践与经验

利用遗传算法进行智能家居设备识别特征选择
..... 王靖翔, 刘翔宇, 尹钟舒, 刘丹丹, 谢镛镛, 杜彦辉 (75)

轻量化卷积神经网络在药片检测分类中的应用 黄开坤, 徐 兴 (81)

轮播图导航制作在Web前端中的实践与应用 赵素萍, 杨 璐, 康晶晶 (87)

基于“碳中和”的高性能计算集群组网建设方法探析..... 陈 阳, 陈坚泽 (91)

开发案例

基于设备规则的变电站自动成票推理系统 顾明煜, 丁雪峰, 江红成, 花 冉, 赵 军 (95)

机载气吸式抓取一体装置设计与试验
..... 吴子彬, 吴志坤, 许绍聪, 余乃浩, 林嘉乐, 殷惠莉 (101)

临床分子诊断与基因检测信息平台的构建..... 莫远明, 卢瑞鸿, 陈 康 (106)

基于Unity 3D酷跑游戏的设计与实现 李爱军 (111)

地方高校班级管理系统的设计与实现 刘思云, 陈贵松 (116)

Modern Computer

(Vol.28, No.19; Oct.10, 2022)

CONTENTS

Research and Development

Classification of Cervical Cells Based on Transfer Learning and Label Smoothing Strategy	(1)
Prediction and Analysis of the Number of Pension Service Beds in China	(10)
Research on Recognition of Esophageal Cancer Images Based on Transfer Learning	(19)
Missing Data Completion Method Based on Simplex Space	(26)
Automatic Detection System of Farmland Fire Source	(33)
A Review of Research Methods for Character Recognition of Angle Steel Stamps	(39)
Research on Population Aggregation Features Detection Based on SIFT Algorithm	(45)
Research on Express Delivery Project Based on Quadruped Bionic Robot	(51)
Research on Network Security Situational Awareness Oriented to Threat Information	(57)
Research on Weibo Sentiment Analysis Method Based on Support Vector Machine	(63)
Detection System for Sheet Metal Hole Defect of Washing Machine Shell Based on U-net Algorithm and Fcos Target Detection Algorithm	(67)
Research on Disease Identification Method Based on AI Deep Learning	(71)

Practice and Experience

Genetic Algorithm is Used to Select the Features of Smart Home Devices	(75)
Application of Lightweight Convolution Neural Network in Tablet Detection Classification	(81)
The Practice and Application of the Navigation Production of the Carousel Map in the Web Front-End	(87)
Analysis of High Performance Computing Server Cluster Construction and Management Methods	(91)

Development Solution

Reasoning Algorithm of Substation Automatic Ticket Generation Based on Equipment Rules	(95)
The Design of an Airborne Air Suction Grab Integrated Device	(101)
Construction of Clinical Molecular Diagnosis and Gene Detection Information Platform	(106)
Design and Implementation of Virtual Parking Lot Based on Unity 3D	(111)
Design and Implementation of Class Management System in Local Universities	(116)

研究与开发

文章编号: 1007-1423(2022)19-0001-10

DOI: 10.3969/j.issn.1007-1423.2022.19.001

基于迁移学习与标签平滑策略的宫颈细胞分类方法

刘 美, 周 龙

(武汉轻工大学电气与电子工程学院, 武汉 430023)

摘要: 为了避免宫颈病理切片中误诊漏诊和细胞识别效率低下的问题, 提高细胞识别效率和精度, 针对宫颈细胞数据集少的特点, 建立以宫颈细胞识别准确率和特异性为目标的深度学习模型 TLS-Net50 (Transfer Label Smooth Net 50)。该算法引入迁移学习, 增强了模型的特征提取能力, 减少了模型的训练时间和训练成本。为了提高模型的泛化性和鲁棒性, 在此基础上使用标签平滑正则化策略。改进后的算法在公开数据上的平均分类准确率达到 98%, 对比其他经典深度学习算法表现优异, 这证明了 TLS-Net50 算法在计算机辅助宫颈细胞分类识别领域是具有优越性的, 具有一定的应用前景。

关键词: 宫颈细胞; 卷积神经网络; 残差结构; 迁移学习; 泛化性

基金项目: 国家自然科学基金资助项目(61075015)

0 引言

宫颈癌是影响妇女健康的重大疾病, 在中国, 宫颈癌是导致女性死亡的第二大杀手, 2020 年我国新发宫颈癌数量达到了 11 万, 占全球总量的 18%^[1]。然而, 宫颈癌是一种早期可以发现的肿瘤, 开展人群筛查是促进宫颈癌早诊早治的有效措施。宫颈癌筛查, 包括妇科检查、宫颈癌初筛、阴道镜检查、病理组织学检查等流程, 其中宫颈细胞学检查是早期发现宫颈癌最简单、有效、方便的方式, 且其准确性较高, 适用于宫颈癌的早期诊断、普查和预防^[2]。宫颈细胞学诊断是基于细胞形态学对宫颈细胞进行分类和统计, 是临床上对宫颈癌诊断的依据, 医学界称之为巴氏涂片检查。

传统的巴氏涂片阅片方法是由病理医师人工阅片, 这种方法不仅效率低下, 而且受限于病理医生的主观性和专业性, 时不时会发生误判和漏判。人工阅片的方法一方面需要消耗大量的人力物力, 另一方面准确率得不到保证, 这就导致了我国基层宫颈癌筛查工作进展缓慢, 筛查质量不高, 因此应用计算机辅助宫颈癌筛查是非常有必要的。

过去, 计算机辅助系统中的宫颈细胞分类任务都是根据细胞核和细胞质的形态学特征来分类。而这些特征的选取是依靠人工来筛选的, 不同的方法筛选的特征不同。Marinakos 等^[3]使用了基于遗传算法的特征选择和最近邻分类来进行巴氏涂片诊断; Kumar 等^[4]设计了一个基于生物学可解释特征的框架来分类癌症; William 等^[5]设计了一个模糊 C 均值算法的宫颈癌辅助诊断工具。然而, 这些手工设计的特征都有一些缺点, 因为它们或多或少会受到特征或分类器选择过程的影响。

人工智能(AI)技术正在改变当今的医疗实践现状。人工智能可以应用在各种各样的任务中, George 等^[6]将深度学习用于乳腺癌诊断; Böhländ 等^[7]用机器学习对甲状腺癌进行自动分类; Song 等^[8]对骨髓组织学的图像进行检测和分类, 许多人工智能应用已经达到了专家级的诊断准确度。

卷积神经网络(CNN)已被证明对医学图像分析处理中人工智能的成功具有重要意义。在宫颈细胞分类领域, 近年来基于深度学习和 CNN 的研究取得了很大进展。研究人员发现, CNN 可以作为特征提取工具, 自动提取深层特

征，对宫颈细胞进行分类。Zhang等^[9]将深度学习应用在宫颈细胞图片分类中，使用CNN模型自动提取宫颈细胞的深层特征，且无需事先分割；Plissiti等^[10]从细胞特征、图片特征、深层特征三个方面出发探讨了CNN在辅助计算机检测方面的可行性，并且提出了全新的宫颈细胞数据库。在此基础上，Shanthi等^[11]用卷积神经网络构建了预测模型，对宫颈癌进行不同等级的分类；Shi等^[12]用图神经网络对宫颈细胞进行分割和分类，深刻地探索了各类宫颈细胞图像之间的潜在关系，进一步提高了分类性能。虽然上述方法在一定程度上提升了准确率，但在宫颈癌诊断上，特异性和召回率还存在不足，用于医疗辅助诊断还需要进一步提高指标。

为了进一步提高宫颈细胞诊断的准确率和特异性，本文提出了基于迁移学习和标签平滑正则化损失函数的TLS-Net50算法，在特征提取阶段使用预训练参数，使网络能更好地学习到细胞特征，同时改进了交叉熵损失函数，使模型具有更好的泛化性和鲁棒性。

1 基于迁移学习和标签平滑策略的算法分析

TLS-Net50总体框架如图1所示，首先对输入的宫颈细胞图像进行预处理。预处理包括对图像的裁剪缩放。之后输入到模型中进行特征提取，特征提取基于ResNet50框架，其中卷积层的参数是通过迁移学习得到的，用卷积层提取图像的深层特征。全连接层的参数则由SIPaKMeD数据集进行训练得到。在此基础之上将原来的交叉熵损失函数进行改变，在原来函数的基础上加上一点噪声，最后经过softmax函数进行输出，得到模型的预测结果。

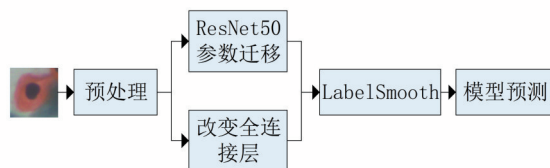


图1 TLS-Net50模型总体框架

1.1 迁移学习

迁移学习是指在其他大规模图像数据集上预训练的深度学习模型的微调。深度学习需要大量有标注的数据，然而现实条件下无法获得大量已标注的数据，因此考虑通过迁移学习的方法，将已有的知识迁移到目标领域中，解决仅有少量已标注样本的学习问题。ImageNet是广泛应用于计算机视觉领域的大型数据库，其中包含了100万幅图，共1000个类别。因此，在ImageNet数据库预训练过的ResNet50模型具有提取宫颈细胞深度特征的能力，尤其是卷积层具有充分的学习图片边缘和轮廓特征的能力。

经过预处理后，我们用在ImageNet数据库上预训练过的ResNet50的前几个卷积层和池化层用作网络的初始参数，在此基础之上对全连接层进行随机权重初始化。利用Sipakmed数据集对网络参数进行微调，就可以获得融合了迁移学习的ResNet50模型。迁移学习的过程如图2所示。

1.2 标签平滑

标签平滑是一种防止模型过拟合的正则化手段，在训练深度学习网络模型时使用标签平滑的方法是很有效的^[14]。标签平滑可以应用在很多任务中，例如图像分类、图片分割、机器翻译和语音识别，使用该方法可以提高深度学习模型的泛化性能和鲁棒性。

在分类问题中，常用one-hot(独热分布)对向量进行编码，编码的标签向量 y_i 为：

$$y_i = \begin{cases} 1, & i = \text{target} \\ 0, & i \neq \text{target} \end{cases} \quad (1)$$

其中正确的类概率是1，不正确的类概率是0。在训练时，模型最小化损失函数 $H(y, p)$ ：

$$H(y, p) = \sum_j y_j \log p_j \quad (2)$$

其中， p_i 由模型前一个全连接层输出的logits向量 z 应用Softmax函数得到：

$$P_i = \frac{\exp(z_i)}{\sum_i \exp(z_i)}, i \in \{0, \dots, K - 1\} \quad (3)$$

在正常的独热向量编码标签网络学习过程中，鼓励模型估计目标类别的概率接近1，非目标类接近0。但是，如果训练数据不足以覆盖所

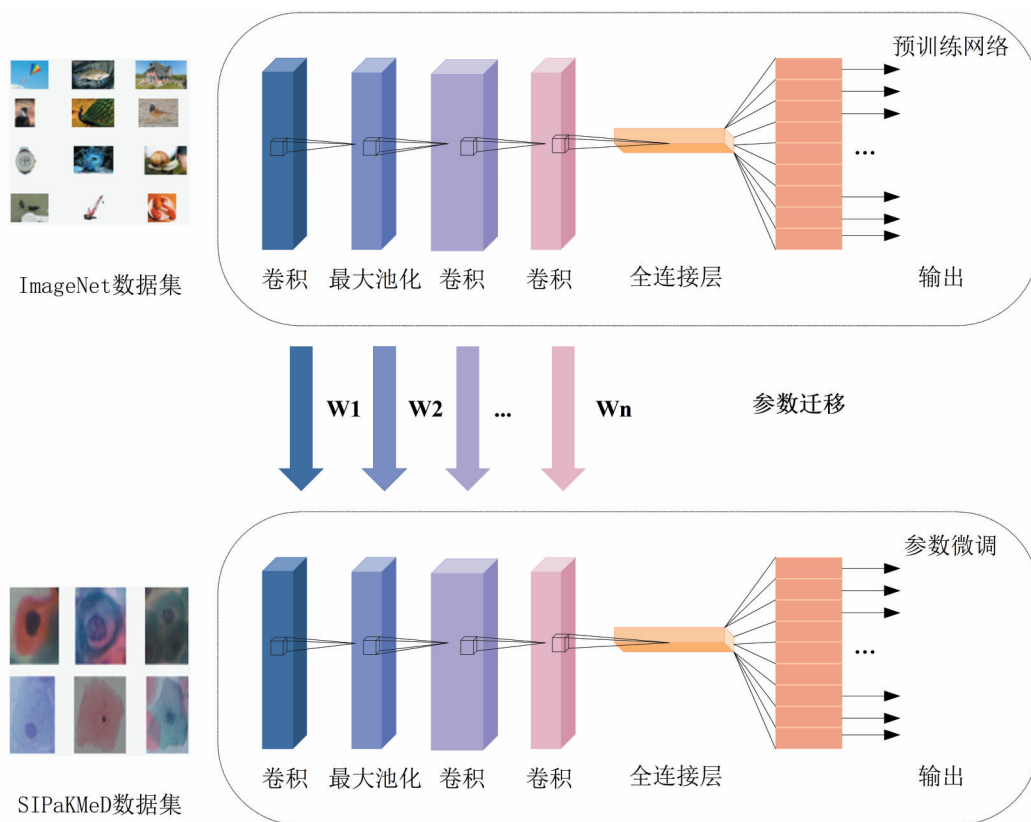


图2 迁移学习

有情况，这不仅会导致网络的过拟合还会降低模型的泛化能力，导致模型在预测集上表现不好。在数据集较少的情况下，使用交叉熵损失函数作为目标函数不一定是最优的。

因此考虑通过标签平滑正则化 (Label Smooth Regularization, LSR) 来优化模型，LSR 是通过在输出 y 添加噪声，实现对模型的预测概率的约束，降低模型过拟合程度的一种约束方法。运用了标签平滑正则化策略后，我们用更新的标签向量 \hat{y}_i 来替换独热编码的标签向量 y_{hot} ：

$$\hat{y}_i = y_{hot(1-\varepsilon)} + \frac{\varepsilon}{K} \quad (4)$$

其中 K 为多分类的类别总个数， ε 一般是一个较小的超参数(一般取0.1)，即：

$$\hat{y}_i = \begin{cases} 1 - \varepsilon, & i = \text{target} \\ \frac{\varepsilon}{K}, & i \neq \text{target} \end{cases} \quad (5)$$

带有标签平滑的交叉熵损失函数转化为下面的公式：

$$H(y, p) = \sum_j^K \hat{y}_i \log p_i \quad (6)$$

这样，标签平滑后的分布就相当于往真实分布中加入了噪声 ε ，避免模型对于正确标签过于“自信”，减少正确目标和错误目标的概率的差距，从而避免过拟合，提高模型的泛化能力。

标签平滑一般用在以下三个场景中：

(1) 数据集比较大，且图片中包含噪音，加入标签平滑之后可以避免模型错误地学习到图片中的噪音。

(2) 分类任务中不同种类图片具有很高的相似度，比如宫颈细胞图片分类中，有些图片又像空泡细胞又像化生细胞，利用此策略可以给两类都提供监督效果。

(3) 在模型比较小时，为了避免模型被训练得太自信，这会导致过拟合，我们希望提高模型学习的难度，也会引入标签平滑。

本课题研究的情况符合标签平滑的使用场景，因此考虑引入标签平滑策略。

1.3 TLS-Net50 总体框架

宫颈细胞识别模型框架主要利用迁移学习将 ResNet50 经过预训练的参数迁移到卷积层、池化层,然后将原来 1000 分类的 softmax 函数改成所需要的目标分类函数(在此文中即五分类)。宫颈细胞识别模型的训练和测试流程如下:

(1) 输入样本。从宫颈细胞数据库中每种细胞图片按 9:1 比例抽取训练集和验证集,将训练集样本作为模型训练输入。

(2) 对输入图片进行预处理。为了保证所有图片输入大小相同并且去除图片中可能存在的噪音,对输入的图像统一进行缩放和裁剪操作。在训练阶段,将单个细胞图像裁剪为 224 像素×224 像素大小,并且进行归一化处理,使特征图像趋近于正态分布,加快模型的收敛和训练。

(3) 构建 TLS-Net50 宫颈细胞识别模型。基于 ResNet50^[13]模型,优化全连接层,将原来的分类器换成五分类输出的分类器,这样不仅可以减少模型学习和训练的参数,还可以提升模型的识别效率。

(4) 微调卷积层的参数。用 Resnet50 经过 ImageNet 数据集预训练得到的参数通过迁移学习方式优化 TLS-Net50 模型参数,主要是确定卷积层的权重、偏置等参数。

(5) 训练模型。随机初始化模型参数,设定优化器的学习率和训练轮数,通过损失函数的反向传播,冻结卷积层的参数,训练全连接层的权重。

(6) 加入标签平滑策略进行训练。通过对损失函数进行改正,使模型具有更好的泛化性,避免了模型的过拟合。

(7) 模型测试。抽取训练集的图片,作为测试样本进行模型测试,以验证模型的精度。

ResNet50 预训练模型的权重参数数量为 25.5 M,计算量为 4.1 G,其中全连接层参数集中度较高,ResNet50 的预训练参数是为 1000 分类而设计的,而本模型只针对五个不同的类。因此本文将原来的全连接层改成五分类的输出,这样虽然稍微增加了模型的计算量,但是减少了模型的参数,还提高了模型的效率和精度。

改进后的模型参数(params)与原来的模型参数和计算量(FLOPS)对比结果如表 1 所示。

表 1 模型参数和计算量对比

模型	Params(M)	FLOPS(G)
ResNet50	25.5	4.10
TLS-Net50	22.99	4.26

TLS-Net50 模型的各层参数大体上与 ResNet50 一致。表 2 给出了模型每层的名称,每层的输入和输出特征图大小,以及每层的层结构。可以看到本模型的一个标志性特点就是特征图的通道增加一倍时,特征图的宽高会缩小一倍,这是为了保证在堆叠网络的过程中,网络不会因为继续堆叠而产生退化。

表 2 TLS-Net50 模型各层设计

层名称	层类型	输入特征图	输出特征图维度
Layer0	卷积+最大池化	3×224×224	65×56×56
Layer1	卷积+下采样模块	65×56×56	256×56×56
Layer2	卷积+下采样模块	256×56×56	512×28×28
Layer3	卷积+下采样模块	512×28×28	1024×14×14
Layer4	卷积+下采样模块	1024×14×14	2048×7×7
Layer5	全连接层	2048×7×7	1×5

2 在宫颈细胞数据集上的实验与分析

2.1 数据集的获取

本实验所用的宫颈细胞数据集是来自希腊约阿尼纳大学在 2018 年发表的公开数据集 SIPaKMeD^[10],该数据集包含了五种宫颈细胞,共 4049 张从 966 个 Pap 涂片的细胞簇图像中手工切分的孤立细胞图像,每类细胞的示例如图 3 所示。

可以看出五类宫颈细胞图像形态各异,且它们在成熟、衰老、癌变的过程中,其形态、大小都在发生着变化,因此可以根据每种细胞特异性的形态学特征进行辨别。

该数据集包含 4049 张孤立细胞图像,其中有 831 张浅中层细胞,787 张副基底细胞,825 张空泡细胞,813 张角化不良细胞,793 张化生细胞。表 3 所示为五类细胞的训练集、验证集分布。

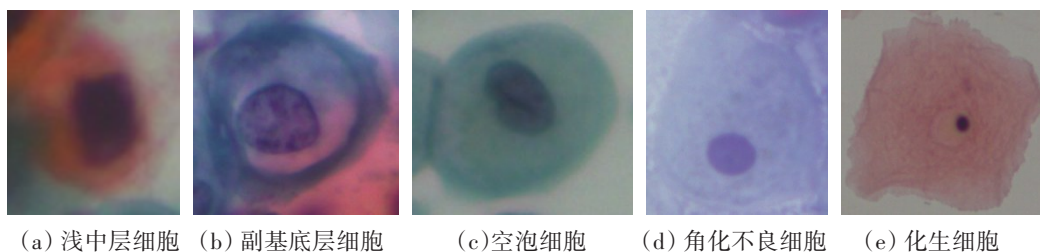


图3 SIPaKMeD中的五种宫颈细胞图像

表3 SIPaKMeD宫颈细胞数据集

细胞类别	细胞数量			备注
	训练集	验证集	总计	
superficial-intermediate	784	83	831	正常
parabasal	709	78	787	
koilocytes	743	80	825	异常
dyskeratotic	732	81	813	
metaplastic	714	79	793	良性
总计	3646	403	4049	

2.2 实验环境及评价指标

2.2.1 实验环境

本文所有实验结果均在同一实验环境下得出。实验环境如下:采用Intel(R) Core(TM) i5-6300HQ CPU @ 2.30GH 处理器, NVIDIA GeForce GTX 960M 显卡, windows10 操作系统, PyTorch=1.7.0+cu110 深度学习框架。编程语言为 python3.6, 相关程序依赖包有: numpy、skimage、matplotlib、tqdm 等, 模型均训练 30 轮, 采用 Adam 优化器, 学习率为 0.0001。

2.2.2 评价指标

以混淆矩阵为依据选取评价指标, 混淆矩阵可以理解为下面一张表格, 如表4所示, 表中有四个基础指标, 分别是 TP 、 TN 、 FP 、 FN , 这四个指标是评价模型的一级指标。

表4 混淆矩阵

真实标签	预测结果	
	正类	负类
正类	TP (True Positive) 正类分类为正类	FP (False Positive) 正类分类为负类
负类	FN (False Negative) 负类分类为正类	TN (True Negative) 负类分类为负类

在分类任务中一般用 ACC (准确率)、 $Precision$ (查准率)、 $Recall$ (查全率)、 $Specificity$ (特异度)、 $F1$ (综合评价指标)来评价模型。查准率和查全率是一对矛盾的度量, 一般来说, 查准率高时, 查全率往往偏低; 而查全率高时, 查准率往往偏低。而 $F1$ 就是综合了查准率与查全率的加权调和平均值, $F1$ 越高说明模型查准率查全率越好。计算公式如表5所示。

表5 评估指标标准公式

评价指标	公式
ACC	$\frac{TP}{TP + FP + FN + TN}$
P	$\frac{TP}{TP + FP}$
R	$\frac{TP}{TP + FN}$
S	$\frac{TN}{FP + TN}$
$F1$	$\frac{2TP}{2TP + FP + FN}$

2.3 实验过程及分析

为了证明本文算法的有效性, 本文基于 ResNet50 模型在测试集上做了相关的消融实验, 如表6所示。下面的算法均以 ResNet50 为基础, 评价指标为上文所提到的五类指标, 其中 TL 代表迁移学习(Transfer Learning), LS 代表标签平滑(Label Smooth)。可以看到, $N2$ 模型在 $N1$ 的基础上 $F1$ 值提升了 4.6% 左右; $N3$ 相比 $N1$, 其准确率、查准率、查全率、特异度、综合评价指标均有大幅度提升, 查准率提升了 8%, 查全率提升了 8.7%, $F1$ 提升了 8.5%。NA 为本文的 TLS-Net50 算法, 对比 $N3$ 又提高了 1.8%、0.4%、0.5%、0.2% 和 0.6%。

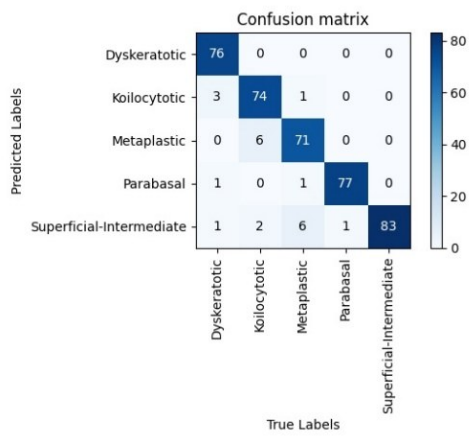
通过消融实验可以知道，融合迁移学习和标签平滑策略可以在单独使用的基础上进一步提升模型效果。因此可以看出，迁移学习和标签平滑策略是非常有用的。这也验证了本文算法在宫颈细胞识别领域的有效性。

表 6 消融实验结果

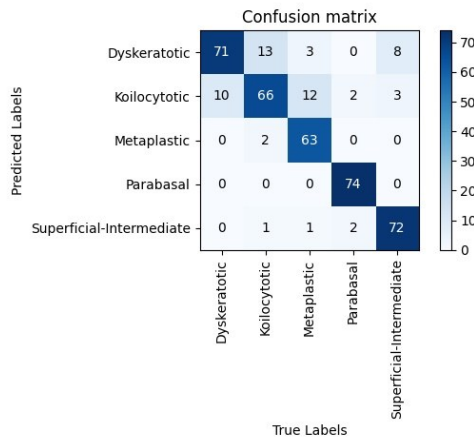
种类	算法			评价指标				
	ResNet50	TL	LS	ACC/%	P/%	R/%	S/%	F1/%
N1	√			94.3	86.8	85.8	96.5	86.0
N2	√		√	94.3	87.5	85.9	96.5	90.6
N3	√	√		97.8	94.8	94.5	98.6	94.5
N4	√	√	√	98.0	95.2	95.0	98.8	95.1

为了进一步分析不同模型对每一类宫颈细胞的识别结果，图 4 给出了 N1、N2、N3、N4 的混淆矩阵，混淆矩阵对角线代表着每类细胞被正确识别的数量。可以看到在四个模型中，副基层细胞最好辨认，空泡细胞最难分类正确，且它和化生细胞有互被错分类的现象，化生细胞有一定的概率被错认为浅中层细胞，角化不良细胞有时候被认为是空泡细胞，浅中层细胞则有时被认为是角化不良细胞。

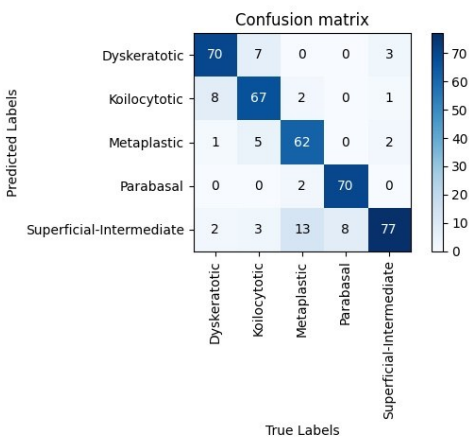
表 7 是 N4 模型在测试集上的表现，对于五类不同的细胞，TLS-Net50 平均准确率、查准率、查全率、特异度、综合评价指标分别达到了 98.0%、95.2%、95.0%、98.8 和 95.1%，副基



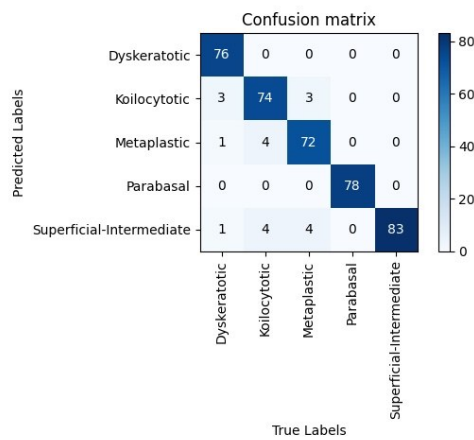
(a) N1



(b) N2



(c) N3



(d) N4

图 4 N1、N2、N3、N4 模型混淆矩阵

底层细胞和浅中层细胞被完全分类正确,而空泡细胞和化生细胞的准确率稍低,说明这两类细胞具有较高的相似度,区分较为困难。

表7 测试集评估指标

细胞种类	ACC/%	P/%	R/%	S/%	F1/%
superficial-intermediate	97.8	90.2	100.0	97.2	94.9
parabasal	100.0	100.0	100.0	100.0	100.0
koilocytes	96.5	92.5	90.2	98.1	91.4
dyskeratotic	98.8	100.0	93.8	100.0	96.8
metaplastic	97.0	93.5	91.1	98.5	92.3
平均识别准确率	98.0	95.2	95.0	98.8	95.1

为了进一步说明本文算法对宫颈细胞分类的有效性,本文与其他的深度学习经典算法进行了对比,其中实验参数均设置一样,实验结果如表8所示。实验说明了TLS-Net50算法的良好性能,也说明了TLS-Net50在计算机辅助宫颈细胞检测中的有效性。

表8 与其他算法的性能对比

算法	ACC/%	P/%	R/%	F1/%
LeNet ^[15]	95.0	88.0	87.7	87.6
AlexNet ^[16]	95.6	89.7	89.0	89.1
VGG16 ^[17]	96.9	92.6	98.1	92.4
GoogLeNet ^[18]	96.3	91.1	90.9	90.8
TLMResNet50	98.0	95.2	95.0	95.1

2.4 特征图可视化

为了更好地理解深度网络,可以将特征图进行可视化。卷积核在训练的过程中没有对图像的空间结构进行改变,因此直接将特征图像素的值映射到0~255的范围,就可以变成图像。

本文对TLS-Net50模型的特征图进行可视化,结果如图5所示。图5(a)是第一个卷积层的前12张特征图,从图片可以看到,每个卷积核学习到的内容都不一样。从不同层可视化出来的特征图可以看出,模型前面几层网络可以提取到图片纹理、细节特征,模型后面的深层网络提取的是细胞的轮廓、形状等特征。一般来说,浅层网络因为卷积的程度不高,可以学习到更多的特征,也包含有图片的关键特征。例如第一组特征图的第6张特征图,提取出的是细胞核边缘。相对而言,层数越深,模型提取到的特征就越抽象。图像的分辨率也会随着卷积的操作而越变越小。

这些经过训练的卷积核包括各种频率和方向的梯度以及不同的颜色斑点,这是宫颈细胞分类任务所必需的。除了卷积核之外,还提供了池化、激活函数的示例单元的特征图,可以观察到池化操作总结了先前的特征,并且特征变得越来越抽象。

由图片所示的各层显示结果可以知道,卷积神经网络能够有效地提取细胞的特征,说明本算法通过迁移学习和标签平滑正则化能够降低背景干扰、增强目标特征。

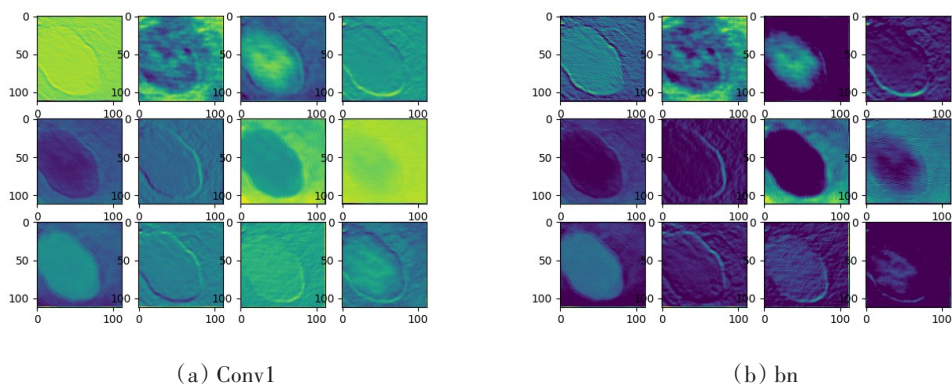


图5 TLS-Net50模型特征图可视化

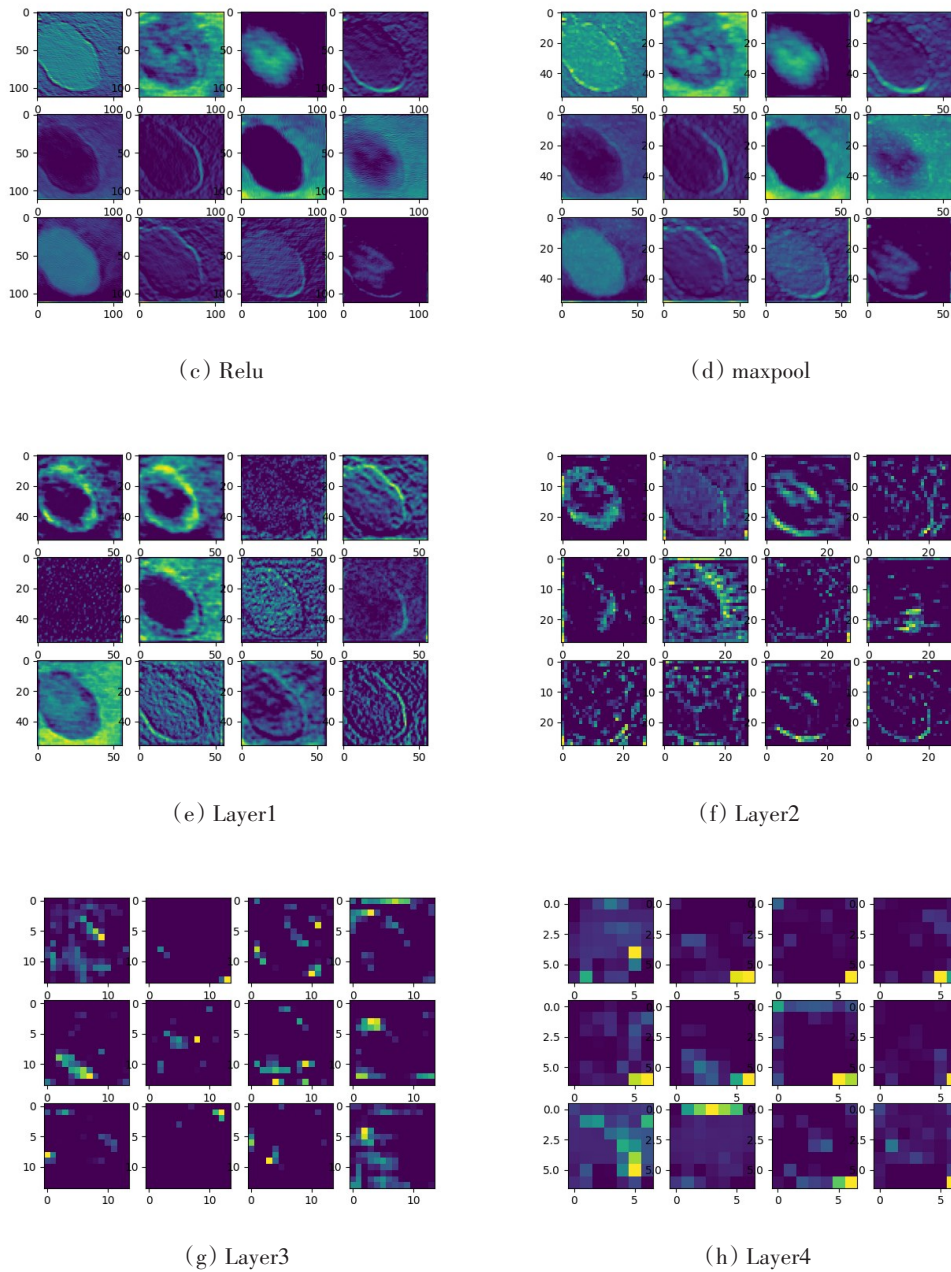


图5 TLS-Net50模型特征图可视化 (续)

3 结语

针对计算机辅助诊断宫颈癌的需求, 本文提出了一种基于卷积神经网络的宫颈细胞分类方法 TLS-Net50。本文对 ResNet50 模型进行参数优化和标签正则化, 并通过实验验证, 表明改进后的模型可以自动从图像中学习到细胞特征, 这样就避免了常规方法中人工选取特征的

不足。同时, TLS-Net50 模型相对于 ResNet50 平均准确率、查准率、查全率、特异性、F1 值依次提高了 3.7%、8.4%、9.2%、2.3% 和 9.1%, 可以满足计算机辅助诊断宫颈癌的需求。

此外, 本文所用模型可以准确地识别宫颈细胞的类型, 并且减少了模型的参数量, 可以在计算性能一般的硬件上进行应用, 有利于基层宫颈癌筛查工作的展开。

参考文献:

- [1] 刘萍. 中国大陆13年宫颈癌临床流行病学大数据评价[J]. 中国实用妇科与产科杂志, 2018, 34(1): 41-45.
- [2] 赵方辉, 戎寿德, 乔友林. 宫颈癌及其癌前病变筛查方法现状[J]. 中国医学科学院学报, 2001(6): 638-641.
- [3] MARINAKIS Y, DOUNIAS G, JANTZEN J. Pap smear diagnosis using a hybrid intelligent scheme focusing on genetic algorithm based feature selection and nearest neighbor classification[J]. Computers in Biology and Medicine, 2009, 39(1): 69-78.
- [4] KUMAR R, SRIVASTAVA R, SRIVASTAVA S. Detection and classification of cancer from microscopic biopsy images using clinically significant and biologically interpretable features[J]. Journal of Medical Engineering, 2015: 1-14.
- [5] WILLIAM W, WARE A, BASAZA-EJIRI A H, et al. A pap-smear analysis tool (PAT) for detection of cervical cancer from pap-smear images[J]. BioMedical Engineering OnLine, 2019, 18(1): 16.
- [6] GEORGE K, FAZILUDEEN S, SANKARAN P, et al. Breast cancer detection from biopsy images using nucleus guided transfer learning and belief based fusion[J]. Computers in Biology and Medicine, 2020, 124: 103954.
- [7] BÖHLAND M, THARUN L, SCHERR T, et al. Machine learning methods for automated classification of tumors with papillary thyroid carcinoma-like nuclei: a quantitative analysis[J]. PLOS ONE, Public Library of Science, 2021, 16(9): e0257635.
- [8] SONG T-H, SANCHEZ V, EI DALY H, et al. Simultaneous cell detection and classification in bone marrow histology images[J]. IEEE Journal of Biomedical and Health Informatics, 2019, 23(4): 1469-1476.
- [9] ZHANG L, LE LU, NOGUES I, et al. DeepPap: deep convolutional networks for cervical cell classification[J]. IEEE Journal of Biomedical and Health Informatics, 2017, 21(6): 1633-1643.
- [10] PLISSITI M E, DIMITRAKOPOULOS P, SFIKAS G, et al. Sipakmed: a new dataset for feature and image based classification of normal and pathological cervical cells in pap smear images[C]//2018 25th IEEE International Conference on Image Processing (ICIP). Athens: IEEE, 2018: 3144-3148.
- [11] B S P, FARUQI F, S H K, et al. Deep convolution neural network for malignancy detection and classification in microscopic uterine cervix cell images[J]. Asian Pacific Journal of Cancer Prevention : APJCP, 2019, 20(11): 3447-3456.
- [12] SHI J, WANG R, ZHENG Y, et al. Cervical cell classification with graph convolutional network[J]. Computer Methods and Programs in Biomedicine, 2021, 198: 105807.
- [13] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [EB/OL]. arXiv: 1512.03385[cs], 2015.
- [14] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision[EB/OL]. arXiv: 1512.00567[cs], 2015.
- [15] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [16] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [17] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. arXiv: 1409.1556[cs], 2015.
- [18] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[EB/OL]. arXiv: 1409.4842[cs], 2014.

作者简介:

刘美(1998—),女,湖北荆州人,硕士研究生,主要研究方向为深度学习、图像处理

通信作者:周龙(1970—),男,湖北武汉人,教授,硕士生导师,博士,主要研究方向为电气控制、图像处理与分析、农业工程与信息技术,E-mail:profzli@126.com

收稿日期: 2022-06-19 修稿日期: 2022-08-04

(下转第32页)

文章编号: 1007-1423(2022)19-0010-10

DOI: 10.3969/j.issn.1007-1423.2022.19.002

中国养老服务床位数的预测与分析

唐鹏翔, 张晓美

(吉林医药学院药学院, 吉林 132013)

摘要: 随着我国人口老龄化进程加速, 养老服务床位数和 65 岁以上老年人口数的预测成为亟待解决的问题。根据中国统计年鉴和民政局年度公报提供的数据建立了养老服务床位 logistic 模型和 ARMA(1,2,1) 模型, 为了研究床位与养老人口数比例之间的关系, 又建立了 logistic 模型, 对 65 岁以上的老年人口数建立指数模型。对上述四个模型使用 MATLAB 和 Eviews 进行求解, 给出相关预测结果, 并对不同模型的预测结果进行比较和分析, 为未来我国养老事业政策的制定和抉择提供理论依据。

关键词: logistic 模型; ARMA 模型; 指数模型; 预测

基金项目: 吉林省高教科研一般课题(JGJX2021D675); 吉林医药学院博士科研启动项目(JYBS2021010LK); 吉林省职业教育与成人教育教学改革研究一般课题(2021ZCY290)

0 引言

中国是世界上人口最多的发展中国家, 随着老龄化进程加速, 养老服务问题日益突出, 增加养老服务床位是一个亟待解决的现实问题, 参考中国统计局官网^[1]、中国民政部公报^[2]等相关数据, 利用数学模型, 对未来养老服务床位数量进行合理预测, 有较强的现实意义。通过预测结果进行综合研判, 可以对未来养老工作的开展提供便利, 更好地服务国家战略, 为将来中国老龄化问题提供一条解决思路。

谭英花在《上海机构养老资源配置研究》^[3]中利用曲线参数估计模型和年龄模型对养老资源进行预测, 从微观和宏观两个层面构建资源的评价体系。徐宗煌^[4]在《基于多元回归模型的养老床位需求预测》中利用 Leslie 矩阵模型、BP 神经网络预测养老床位数量情况。王子鑫等^[5]在《养老服务床位需求预测与运营模式模型分析》中使用多元线性回归模型和 Dijkstra 算法解决了养老床位需求问题。

本文通过中国统计局官网和民政部公报提供的数据, 建立养老服务床位数的 logistic 模型、ARMA 模型, 使用 MATLAB 和 Eviews 两种软件计算得到我国未来养老服务床位数; 另外建立养老服务床位与养老人口数比例的数学模型, 65 岁以上的老年人口的指数模型也可以得到未来我国养老服务床位数的变化情况, 并对这三种预测结果进行对比分析。

1 数据来源

本文所使用的养老服务床位数来自中华人民共和国民政部 2006 年后的年度统计公报。65 岁以上老年人口数来源于中国统计局官网统计数据。

2 模型的建立与计算

2.1 模型 1: 养老服务床位数的 logistic 模型

根据民政部公布的养老服务床位数统计情况(见表 1), 使用 MATLAB 对表 1 数据绘制散点

表 1 民政部公布的养老服务床位数统计情况

年份	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
床位数	153.5	212.8	234.5	299.3	314.9	353.2	416.5	493.7	577.8	672.7	730.2	744.8	727.1	775	821

万张

图(以2006年为初始时刻0),结果如图1所示。

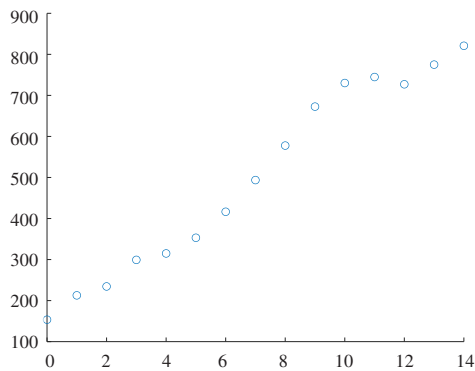


图1 养老服务床位数变化情况散点图

从已有数据的散点图分析可知,养老服务床位数在增加,但是增加幅度在逐渐减缓。从趋势上看,符合“S”型曲线趋势^[6]。从机理上分析老年人养老服务床位数量变化,由中国民政部每年年度公报中发布的养老服务床位数可知,养老服务床位数虽然在不断增加,但是由于建设资源、选择意愿等因素对老年人养老服务床位数增长起到了阻滞作用,并且随着时间的推移,阻滞作用越来越大。因此,对养老服务床位数设立阻滞增长模型(Logistic模型)。

设第 t 年的养老服务床位数为 $x(t)$,年增长率为 r ,初始养老服务床位数为 x_0 ,则 $x(t)$ 满足的微分方程是:

$$\frac{dx}{dt} = rx(1 - \frac{x}{x_m}), x(0) = x_0 \quad (1)$$

其中 x_m 为最大养老服务床位数。其解为

$$x(t) = \frac{x_m}{1 + \left(\frac{x_m}{x_0} - 1\right)e^{-rt}} \quad (2)$$

使用MATLAB曲线拟合工具箱cftool进行参数拟合^[7],计算时床位数以百万为单位,得出图2。

计算得 $x_m=9.19$, $r = 0.2704$,拟合优度为0.9827,残差值为1.118,拟合效果较好,故养老服务床位数预测的logistic模型为

$$x(t) = \frac{9.19}{1 + 5.23 \times e^{0.2704t}} \quad (3)$$

得到养老服务床位数的预测模型后,利用模型计算相关已知年份的床位数据并与已知数据进行对比,发现模型计算结果与已知数据相差较小(见表2)。

表2 模型结果与实际比较

年份	实际床位数	预测床位数	预测值与实际值的差值
2006	153.5	147.5	-6.0
2007	212.8	148.1	-64.7
2008	234.5	227.1	-7.4
2009	299.3	276.5	-22.8
2010	314.9	331.4	+16.5
2011	353.2	390.5	+37.3
2012	416.5	452.2	+35.7
2013	493.7	514.0	+20.3
2014	577.8	573.9	-3.9
2015	672.7	630.0	-42.7
2016	730.2	680.7	-49.5
2017	744.8	725.3	-19.5
2018	727.1	763.4	+36.3
2019	755.0	795.3	+40.3
2020	821.0	821.5	+0.5

最后,使用养老服务床位数模型进行计算,获得未来30年(2020—2050)的数据如表3所示。

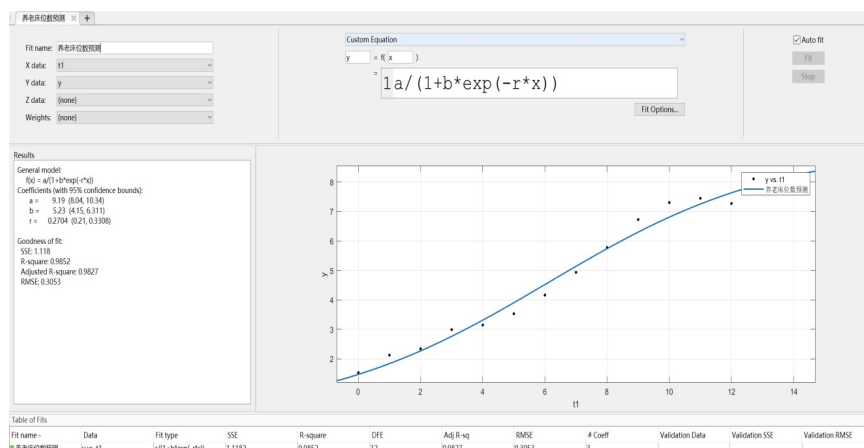


图2 基于MATLAB养老服务床位拟合结果

表 3 未来30年养老服务床位数据结果

															万张	
年份	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035
预测床位数	821.5	842.7	859.6	873.0	883.4	891.6	898.0	902.9	906.6	909.5	911.8	913.5	914.8	915.8	916.5	917.1
年份	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	
预测床位数	917.6	917.9	918.2	918.4	918.5	918.6	918.7	918.8	918.8	918.9	918.9	918.9	918.9	919.0	919.0	

利用已知和计算后的数据作图(见图3), 展示后可以看出, 养老服务床位数据图像呈“S”曲线, 较符合阻滞增长模型^[8]情况, 具有良好的预测意义。

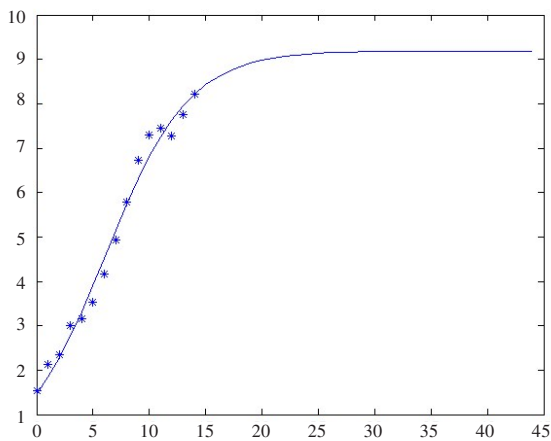


图 3 logistic模型预测结果图

通过所建立的 logistic 模型计算结果可得, 我国养老服务床位数将逐年稳定至919.0万张。

2.2 模型2:养老服务床位数的自回归移动平均模型

2.2.1 时间序列自回归移动平均模型

将养老服务床位数视为一个时间序列 $\{X_t\}$, 可以记为 $ARMA(p, q)$ 。其中 $ARMA(p, q)$ 模型^[9]的数学表达式可表示为

$$\begin{cases} X_t = c + \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \\ \varepsilon_t \sim WN(0, \sigma^2), \forall s < t, E(X_s \cdot \varepsilon_t) = 0 \end{cases} \quad (4)$$

其中, $\varphi_1, \dots, \varphi_p$ 为自回归系数, $\theta_1, \dots, \theta_q$ 为移动平均系数。

2.2.2 数据的平稳性检验

将2006—2020年养老服务床位数的数据输入到Eviews软件中, 对该数据进行时间序列的

预处理^[10], 首先得到图4的时序图, 该时序图中的数据呈递增趋势, 该图趋势性明显, 而对于平稳序列的要求是时序图应该要在一个值附近上下波动, 所以通过对该时序图的观察, 该序列不是平稳序列^[11]。

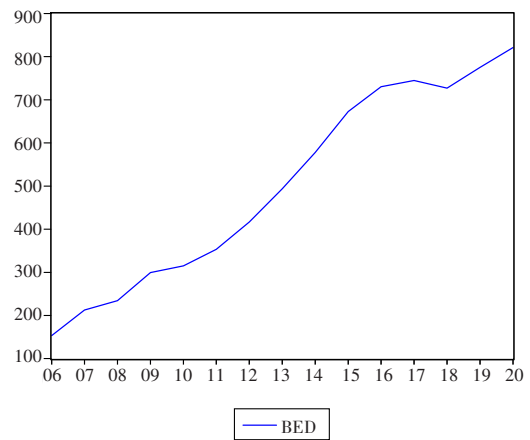


图 4 养老服务床位数的时序图

对于不平稳序列, 可以使用差分使其平稳化^[12], 首先采用一阶差分的方法, 将数据在Eviews软件下通过一阶差分得到新序列, 然后对其进行平稳性检验。

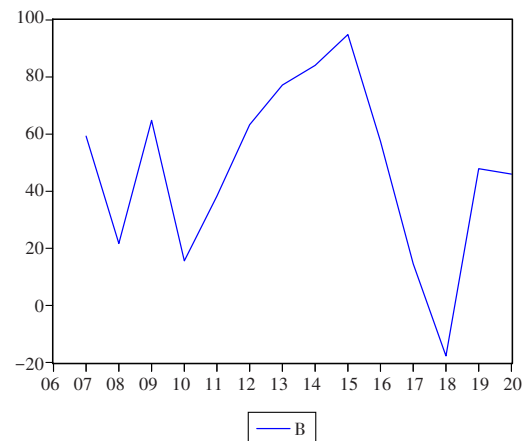


图 5 养老服务床位数据序列的时序图(一阶差分后)

进行单位根检验^[13],运用Eviews软件得到的结果见表4。

表4 一阶差分后的单位根检验结果

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.300073	0.1857
Test critical values:		
1% level	-4.057910	
5% level	-3.119910	
10% level	-2.701103	

通过对表4的结果分析,ADF检验中t统计量为-2.300073,与1%、5%、10%的t统计量比较,其更大,则应该接受该序列存在单位根的原假设,所以与上面时序图检验结果相同,认为该序列是不平稳的^[14],再将养老服务床位数进行二次差分处理,结果如图6所示。

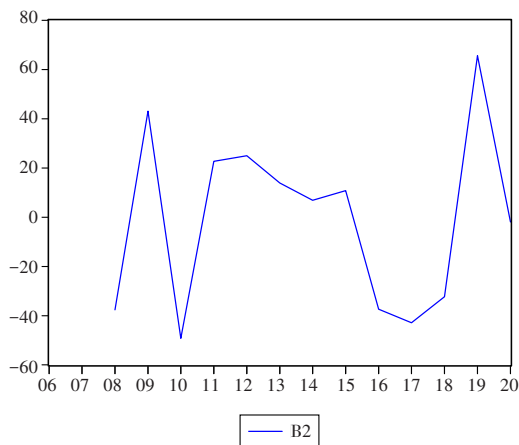


图6 养老服务床位数序列的时序图(二阶差分后)

再次进行单位根检验,结果见表5。

表5 二阶差分后的单位根检验结果

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-4.143441	0.0097
Test critical values:		
1% level	-4.121990	
5% level	-3.144920	
10% level	-2.713751	

t统计量相应的概率值 $P=0.0097$,非常小,则应该拒绝该序列存在单位根的原假设,认为该序列是平稳的^[15]。

2.2.3 ARMA模型的识别

养老服务床位数的ARMA模型识别即识别模型的阶数 (p, q) ,通常借助序列的相关图。

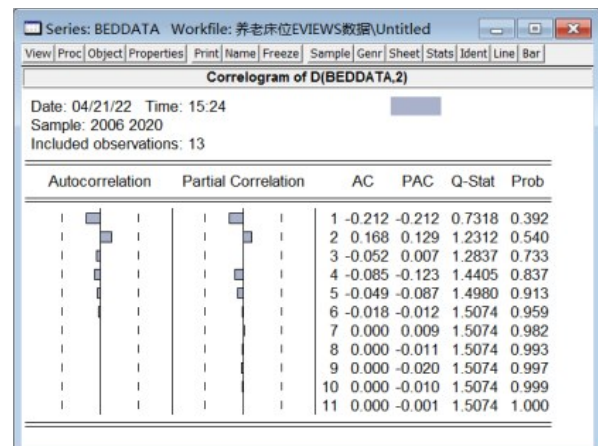


图7 二阶差分后的自相关和偏自相关图

根据得出的自相关和偏自相关图(见图7),目前预测该模型可为ARMA(1,1), ARMA(2,2), ARMA(4,4)等。但是具体选择哪个模型,还需要综合考虑不同模型的各项指标,通过对比选择最优的一个模型来进行拟合。

2.2.4 ARMA模型的选择

用Eviews软件建立ARMA模型后,需要对模型进行对比分析,然后选择一个最优模型,以下所给出的两个模型是经过初步对比后效果相对较好的模型,下面就对这两个模型进行对比分析选最优。

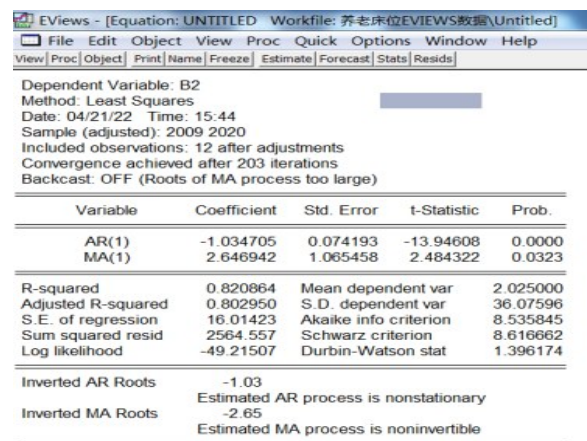


图8 建立ARMA(1,2,1)模型结果

通过图8可以看到,参数的t统计量p值都小于0.05,说明模型中的参数估计显著非零;决定系数和调整后的决定系数(即R-squared和Adjusted R-squared)分别约为0.820和0.803;

AIC 和 *SBC* 的值 (即 Akaike info criterion 和 Schwarz criterion) 分别约为 8.535 和 8.167; *DW* 值 (即 Durbin-Watson stat) 约为 1.396^[16]。模型 ARMA(1, 2, 1) 的拟合效果非常好, 但是我们要看有没有比它更好的模型。

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(2)	0.637376	0.052655	12.10468	0.0000
MA(2)	-1.326498	0.260877	-5.084758	0.0007

R-squared	0.470431	Mean dependent var	-1.709091
Adjusted R-squared	0.411589	S.D. dependent var	35.32094
S.E. of regression	27.09396	Akaike info criterion	9.599464
Sum squared resid	6506.744	Schwarz criterion	9.671809
Log likelihood	-50.79705	Durbin-Watson stat	1.436291

图9 建立 ARMA(2,2,2) 模型结果

通过图9可知, 参数的 *t* 统计量 *p* 值都小于 0.05, 说明模型中的参数估计显著非零; 决定系数和调整后的决定系数 (即 *R*-squared 和 Adjusted *R*-squared) 分别约为 0.470 和 0.412; *AIC* 和 *SBC* 的值分别约为 9.599, 9.672; *DW* 值 (即 Durbin-Watson stat) 约为 1.436。由此可以对比两个模型的各项指标, 结果见表6。

表6 模型指标对比结果

模型	决定系数	调整后的决定系数	<i>AIC</i> 值	<i>SBC</i> 值	<i>DW</i> 值
ARMA(1, 2, 1)	0.821	0.803	8.535	8.167	1.396
ARMA(2, 2, 2)	0.470	0.412	9.599	9.672	1.436

由以上五个指标的对比可以看出, 相比之下, ARMA(1, 2, 1) 模型的决定系数较大, *AIC* 值、*SBC* 值两模型差别不大^[17]。因此, 按照简单性原则, 选择 ARMA(1, 2, 1) 模型, 对拟合模型的残差序列做纯随机检验, 对 ARMA(1, 2, 1) 模型进行残差序列相关性检验结果如图10所示。

通过图10, 可以得到残差为纯随机序列, 残差值较小, 选用模型较为恰当。因此, 可以得到预测模型公式 (数据经过二阶微分处理):

$$X_t = -0.220720X_{t-1} + 3.169663\varepsilon_{t-1} \quad (5)$$

绘制预测图, 结果如图11和图12所示。

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-0.220720	0.026844	-8.222177	0.0000
MA(1)	3.169663	0.385502	8.222166	0.0000
RESID(-1)	-0.034752	0.119759	-0.290183	0.7791
RESID(-2)	-0.112103	0.125954	-0.890031	0.3994

R-squared	0.922632	Mean dependent var	0.688279
Adjusted R-squared	0.894994	S.D. dependent var	15.25204
S.E. of regression	4.942375	Akaike info criterion	6.294771
Sum squared resid	195.4165	Schwarz criterion	6.456406
Log likelihood	-33.76862	Durbin-Watson stat	2.105704

图10 ARIMA(1,2,1) 模型相关性检验

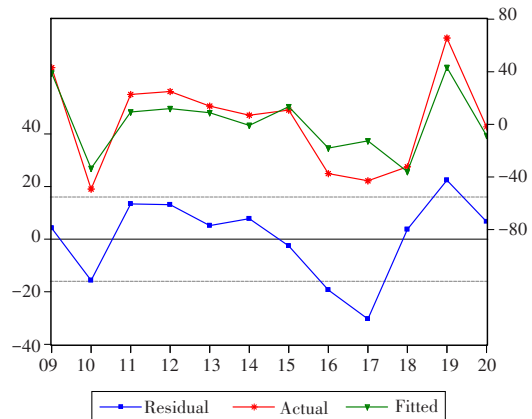


图11 模型拟合效果图

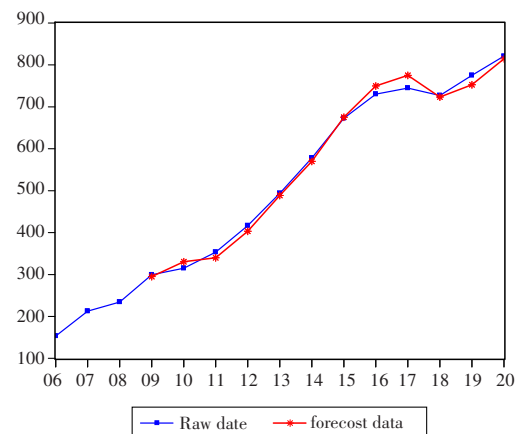


图12 模型预测图

从图11和图12可以看出, 模型预测效果较好^[18]。于是计算模型预测结果, 见表7。

表 7 ARIMA(1,2,1)模型预测结果

年份	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
预测床位数	NA	NA	NA	295.1049	315.4547	377.4567	396.3609	459.8586	477.2153	542.3142	558.014	624.8274	638.7533
年份	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	
预测床位数	707.4021	719.4288	790.0427	800.0362	872.7539	880.5705	955.5406	961.0267	1038.408	1041.399	1121.362	1121.682	

表 8 养老服务床位数与养老人口数比例

年份	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
床位数(万张)÷ 老龄人口数	1.48	1.99	2.13	2.64	2.64	2.88	3.26	3.72	4.16	4.65	4.86	4.67	4.35	4.37	4.31

2.3 模型 3:养老服务床位数与 65 岁以上老年人口数比例的 logistic 模型

利用已知的中国民政部官网每年公布的养老服务床位数和中国统计局 65 岁以上老年人口数据,计算床位数与养老人口数比例,结果见表 8。

以时间为横坐标,床位数与老龄人口的比值为纵坐标,绘制散点图,如图 13 所示。

图 13 中散点随着时间的推移,增长趋势受到阻滞且最终较平稳,与 logistics 模型曲线较为相似。利用 MATLAB 对时间和每年床位数与老龄人口的比值进行曲线拟合^[19]。利用已知的人口阻滞增长模型(logistics 模型)来预测养老服务床位数与养老人口数比例,结果如图 14 所示。

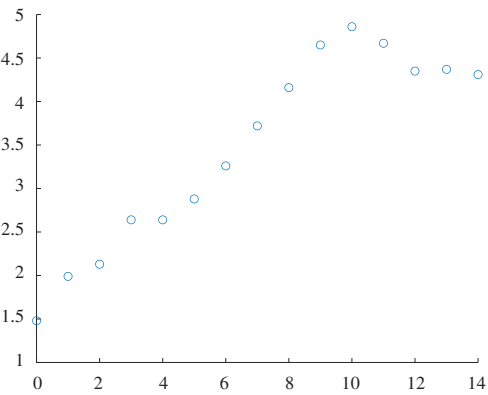


图 13 养老服务床位数与养老人口数比例散点图

通过曲线拟合得到:

$$x(t) = \frac{4.791}{1 + 2.328e^{-0.3115t}} \quad (6)$$

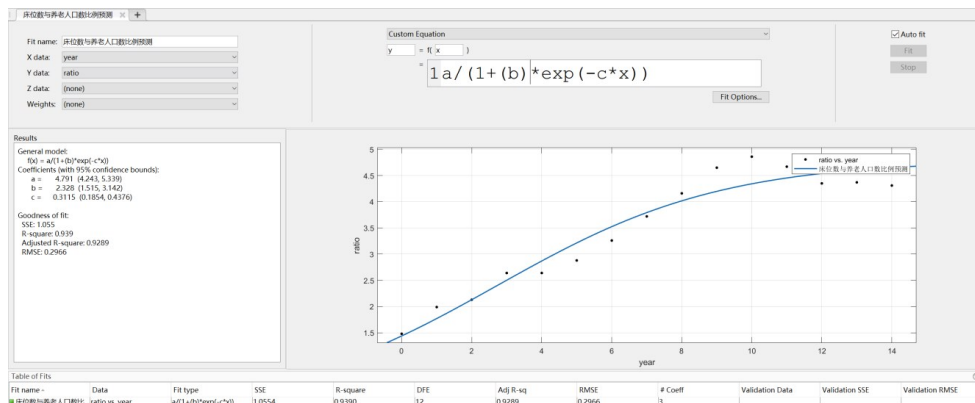


图 14 床位数与养老人口数比例拟合结果

模型显示残差值 $SSE=1.055$ ，拟合优度 $R=0.939$ ，拟合结果较好。

模型计算结果数据，见表9。

表9 未来养老服务床位数与养老人口数比例 %

年份	床位数(万张)/老龄人口数
2021	4.69
2022	4.72
2023	4.74
2024	4.75
2025	4.76
2026	4.78
2027	4.78
2028	4.78
2029	4.78
2030	4.78

通过拟合结果可以发现，未来随着时间推移，床位数与养老人口数比例逐渐趋于4.78%附近，但是65岁以上老年人口数还在增长，说明65岁以上老年人对于养老机构养老的选择只有微小增长，比例较稳定。

通过模型结果发现，我国65岁以上老年人对选择养老服务床位的意愿还处于较低水平。

造成这一现象可能有多种原因，如家庭状况等^[20]。在我国65岁以上老年人口持续增长的背景下，我国应加快养老机构的建设、完善相关政策等^[21]，养老机构应当加强对于养老的研究与探索，找到符合我国或某地区的合理养老规划方案，使更多65岁以上老年人到养老机构养老，减缓社会相关养老压力^[22]。

2.4 模型4:65岁以上老年人口的人口指数模型

二百多年前，英国人口学家马尔萨斯(Malthus, 1766—1834)调查了英国一百多年的人口统计资料，得出了人口增长率不变的假设，并据此建立了著名的人口指数增长模型^[23]，本文利用此模型进行未来65岁老年人口数的短期预测。

设 t 年后65岁以上人口数为 $x(t)$ ，年增长率为 r ，初始人口数为 x_0 ，则 $x(t)$ 满足的微分方程是：

$$\frac{dx}{dt} = rx, x(0) = x_0 \quad (7)$$

其解为：

$$x = x_0 e^{rt} \quad (8)$$

根据中国统计局官网数据，得到我国65岁以上人口数据，见表10。

表10 我国65岁以上人口数据

年份	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
65岁以上人口数	6368	6938	7218	7289	7622	7510	7845	8085	8359	8679	8821	9062	9377	9692	9879	10068
年份	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	
65岁以上人口数	10384	10702	11023	11343	11934	12277	12777	13262	13902	14476	15037	15961	16724	17725	19059	

相关数据绘制成散点图(图15)，发现我国65岁以上老年人口从1990年—2020年大概呈指数函数型增长。

使用MATLAB曲线拟合工具箱cftool，进行参数 x_0, r 的计算，以1990年为初始年份，其对应的 t 值为0，拟合结果如图16所示。

计算得 $x_0 = 61.61, r = 0.03499$ ，拟合优度为0.9805，拟合效果较好，故65岁以上人口的预测模型为

$$x = 61.61e^{0.03499t} \quad (9)$$

根据模型4，绘制出拟合效果图和至2030年65岁以上人口数的预测曲线，结果如图17所示。

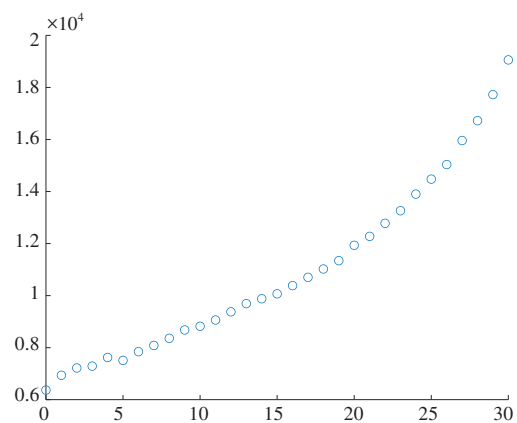


图15 我国65岁以上老年人口散点图

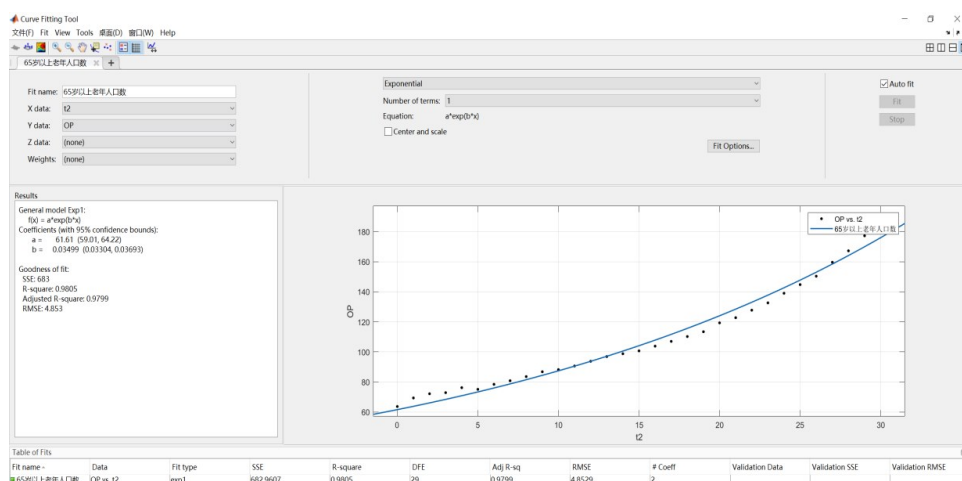


图 16 65 岁以上老年人口指数模型拟合图

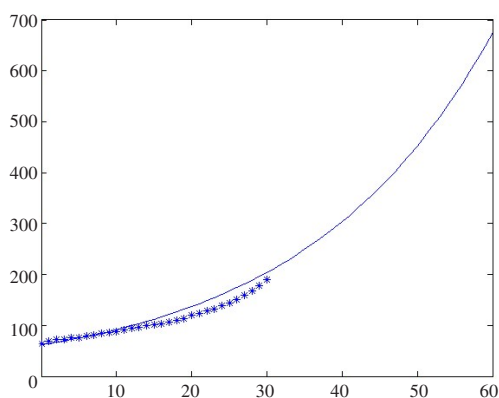


图 17 65 岁以上人口拟合图预测图

从预测值看，到 2050 年我国老年人口将达到 7 亿，这与实际情况有一定差距，主要是各种因素影响使人口增长率为常数的假设不成立^[24]。但是我国老年人口数实际上还在增加，此人口指数模型进行短期预测能起到一定参考作用。

3 模型比较与分析

根据上海养老网站整理得到民政部 2021 年公布养老服务床位数，用于不同模型间比较，结果见表 11。

表 11 模型预测比较

年份	实际床位数	万张		
		模型 1 预测值与实际值的差值	模型 2 预测值与实际值的差值	模型 3、模型 4 预测值与实际值的差值
2021	813.5	+29.2	-23.4573	+41.3679

通过表 11 可知，三个模型预测结果中 logistic 模型与 ARIMR(1, 2, 1) 模型预测效果相似且较好，指数模型预测得到 65 岁以上老年人口数，再乘以床位数与养老人口数比例得到养老服务床位数预测误差稍大。但是随着未来各种未知因素影响(如：生育政策等)，利用该方法进行预测能及时调整，更加动态灵活，也具备一定的预测价值。

4 结语

本文使用两种软件建立四种数学模型预测养老服务床位数、床位数与 65 岁以上老年人口比例、65 岁以上老年人口数。目前 2021 年养老服务床位数仅为 821 万张。模型 1 计算结果显示 2030 年我国养老服务床位为 911.8 万张；模型 2 计算结果显示为 1121.682 万张，还有较大的提升空间；模型 3 预测结果显示，未来养老服务床位与 65 岁以上老年人口数比例继续呈上升趋势，但上升空间较小，在 2026 年将逐年稳定至 4.78%；模型 4 预测结果显示未来 65 岁以上老年人口数还会继续增加。

从模型比较和应用来看，根据历史养老服务床位数建立模型 1 和模型 2，基于当前影响因素条件下进行的预测。如果未来出现新的影响因素(如新生育政策等)，则需要重新计算模型系数，从而保持模型预测的准确性。模型 3 计算比例显示，我国 65 岁以上老年选择养老服务床位的意愿处于较低水平，面对日益增加的老龄人口，可以改善养老服务模式，提升养老服务

机构建设等,从而增加老年人选择养老服务床位的比例^[25]。而模型3和模型4可以结合使用,当未来65岁以上老年人口数或者65岁以上老年人选择养老方式的意愿发生较大变化时,调整其中一个模型进行计算就能得到预测结果。更加简便快捷,能更加适应短期内的动态变化调整,短期内时效性更强。总的来说,本文核心研究内容是养老服务床位的需求预测,四个模型各有特点,能适应长期和短期养老服务床位数预测的需要,也提供了未来具有不确定性因素影响下的调整思路。为相关养老服务床位数的研究提供研究思路,为未来我国养老事业政策的制定和抉择提供理论依据。

参考文献:

- [1] 中华人民共和国国家统计局. 2020年数据地图 65岁以上老年人口[EB/OL], 2020. <http://www.stats.gov.cn/>.
- [2] 中华人民共和国民政部. 民政事业发展统计公报[R]. 2006~2020.
- [3] 谭英花. 上海机构养老资源配置研究[D]. 上海: 上海工程技术大学, 2014.
- [4] 徐宗煌. 基于多元回归模型的养老床位需求预测[J]. 沈阳大学学报(社会科学版), 2022, 24(1): 52-61.
- [5] 王子鑫, 王森, 卢悦冉, 等. 养老服务床位需求预测与运营模式模型分析[J]. 中国市场, 2021(11): 127-128.
- [6] 李哲瑜, 简宋全, 李青海. 基于Logistic回归的个人消费贷款预测[J]. 现代计算机(专业版), 2017(29): 9-12.
- [7] 唐家德. 基于MATLAB的非线性曲线拟合[J]. 计算机与现代化, 2008(6): 15-19.
- [8] 陈平平, 黄仲羽, 黄振, 等. 基于Logistic回归分析的“三高”移动健康管理系统设计[J]. 现代计算机, 2013(14): 58-61, 66.
- [9] 熊志斌. 基于ARIMA与神经网络集成的GDP时间序列预测研究[J]. 数理统计与管理, 2011, 30(2): 306-314.
- [10] 林在超. 基于滚动ARIMA的农作物价格预测方法应用[J]. 现代计算机, 2021, 27(35): 80-83.
- [11] 樊欢欢. EViews统计分析与应用[M]. 2版. 北京: 机械工业出版社, 2014: 222.

- [12] 王丽娜, 肖冬荣. 基于ARMA模型的经济非平稳时间序列的预测分析[J]. 武汉理工大学学报(交通科学与工程版), 2004(1): 133-136.
- [13] 陈双金. 时间序列单位根检验方法比较[D]. 成都: 电子科技大学, 2013.
- [14] 景倩. ARMA模型在我国农林牧渔业发展中的应用[J]. 消费导刊, 2008(7): 189-190.
- [15] 薛冬梅. ARIMA模型及其在时间序列分析中的应用[J]. 吉林化工学院学报, 2010, 27(3): 80-83.
- [16] 邱立国. 日度负荷预测算法的研究和应用[J]. 现代计算机(专业版), 2017(21): 21-25, 35.
- [17] 陈立新, 张磊. 一种面向金融数据的基于ARIMA的模糊时间序列预测模型[J]. 现代计算机(专业版), 2014(7): 3-8, 13.
- [18] 丁明, 张立军, 吴义纯. 基于时间序列分析的风电场风速预测模型[J]. 电力自动化设备, 2005(8): 32-34.
- [19] 方匡南, 范新妍, 马双鸽. 基于网络结构Logistic模型的企业信用风险预警[J]. 统计研究, 2016, 33(4): 50-55.
- [20] 张文娟, 魏蒙. 城市老年人的机构养老意愿及影响因素研究: 以北京市西城区为例[J]. 人口与经济, 2014(6): 22-34.
- [21] 穆光宗. 我国机构养老发展的困境与对策[J]. 华中师范大学学报(人文社会科学版), 2012, 51(2): 31-38.
- [22] 张洋. 我国社会养老服务体系完善研究[D]. 长春: 东北师范大学, 2016.
- [23] 何春. 马尔萨斯人口模型在广州市人口预测中的应用[J]. 广东工业大学学报, 2010, 27(3): 31-34.
- [24] 张金明, 李骞. 基于马尔萨斯模型的北京市人口预测[J]. 特区经济, 2013(7): 63-65.
- [25] 穆光宗, 张团. 我国人口老龄化的发展趋势及其战略应对[J]. 华中师范大学学报(人文社会科学版), 2011, 50(5): 29-36.

作者简介:

唐鹏翔(2000—), 男, 云南楚雄人, 在读本科, 研究方向为生物制药与数据处理

通信作者: 张晓美(1987—), 女, 河南漯河人, 博士, 研究方向为生物制药与数据处理, E-mail: zirenjlcc@163.com

收稿日期: 2022-06-12 修稿日期: 2022-07-09

(下转第44页)

文章编号: 1007-1423(2022)19-0019-08

DOI: 10.3969/j.issn.1007-1423.2022.19.003

基于迁移学习的食道癌影像识别研究

凌其颖, 刘小芳, 张 阳

(四川轻化工大学计算机科学与工程学院, 宜宾 644000)

摘要: 针对食道癌影像在深度学习中因为数据集小、数据不平衡而识别效果不佳等问题, 提出一种基于 AlexNet 神经网络模型的迁移学习识别方法。首先对下载好的食道癌数据集进行相应的数据增强和预处理作为目标样本备用, 接着挑选大量标签完整有关联的 ImageNet 图像对 AlexNet 网络进行训练, 将训练好的 AlexNet 网络有效的特征提取层进行冻结, 对未冻结的替换层进行符合目标样本分类任务的参数调整和结构优化, 然后将目标样本放入完成迁移学习的 AlexNet 网络中进行训练。最后将目标样本直接放入传统的 AlexNet、GooLeNet、ResNet 神经网络中直接训练, 并与进行迁移学习以后的 AlexNet 模型进行结果比较。结果表明基于迁移学习的 AlexNet 模型识别率明显高于传统的 AlexNet 神经网络模型, 验证了迁移学习的可行性。

关键词: 迁移学习; AlexNet; 食道癌; 食道内窥镜

基金项目: 四川省科技计划资助(2017GZ0303); 四川省院士(专家)工作站基金项目(2016YSGZZ01); 高层次新人才培养专项资助(B12402005); 四川轻化工大学人才引进项目(2021RC16)

0 引言

食道癌是最致命的癌症之一, 全球范围内死亡率排第六。其发病率呈世界分布, 有一定的地域性^[1], 占全球一半食道癌发病率的中国尤为明显, 河南、河北、广东、四川等省均为食道癌的高发区。由于食道癌早期不易察觉, 而且大部分依赖于专家的诊断, 我国目前食道癌的早期诊断率几乎为零^[2]。研究表明^[3], 早期确诊的食管癌^[4], 经合理有效治疗后五年生存率高达 70%^[5], 而对于中晚期患者, 出现邻近组织转移, 五年生存率下降到 24%, 尤其是远处转移的患者, 五年生存率仅为 5%。所以, 早发现、早治疗是降低食道癌患者死亡率的关键。

随着深度学习的兴起, 公开数据的增加, 以及计算机硬件的升级, 促使深度学习的应用领域急速扩展。在医学领域, 经过训练的深度学习模型可以快速整合大量医学图像信息并进行分析, 有效提高大多数疾病的筛查、诊断效率, 从而成为各科医生的有效诊断辅助工具。但是由于医学领域的特殊性, 样本类别不平衡

的小数据集成为研究人员望而却步的原因, 虽然这些有着许多难题的小数据隐含着丰富的各类信息^[6]。对医学数据集进行人工标注是一件费时费力的事, 原因是用于神经网络模型学习的数据集在制作以及处理等各方面的标准都极高^[7]。因此将在其他领域学习到的知识转移到目标领域的迁移学习, 成了解决以上问题的大趋势。

1 迁移学习

为了解决深度学习中数据集小而导致训练样本不足的问题而提出了迁移学习^[7]。迁移学习可以根据两类样本之间的联系以及最终学习的目的^[8], 划分为直推式迁移学习、归纳迁移学习, 以及无监督迁移学习等。迁移学习中可以使用类型截然不同的源领域样本和任务领域样本, 但并不能毫无干涉, 其基本要求和原则就是基于源域训练学习得到的预训练模型, 对于目标任务也有着不错的测试效果。迁移学习广泛存在于人们的生活中, 两个不同领域交集

越多，迁移学习的难度相对越好，反之迁移的难度和效果较不理想，甚至产生“负迁移”现象^[9]，与理想效果背向而驰。

近些年来迁移学习不仅在理论研究上到处开花结果，而且还被频繁地应用在类型不同但拥有关联的各行各业中。Yang等^[10]为了克服深度学习在乳腺癌前哨淋巴结转移CT图像数据不足的缺陷，通过迁移学习，将ImageNet数据集上训练好的CNN-Fast网络模型用于乳腺癌前哨淋巴结的分类，取得了当前基于CT研究的最优结果。Van等^[11]提出了一种使用带有卷积神经网络的迁移学习进行早期食道癌检测的新方法，在只给定少量带注释的数据情况下，可使用CNN进行食道癌分类。Lévy将预训练好的AlexNet^[12]深度学习模型用于乳腺癌良恶性鉴别，并取得了优于现有研究成果的诊断结果^[13]。

以上研究表明，迁移学习在诸多肿瘤类型中应用颇多且成效较好，现将其用于食管癌内窥镜识别，是一种可行方案。

2 神经网络模型

2.1 AlexNet神经网络模型

AlexNet在机器视觉领域学术竞赛ImageNet中的Top-5上的分类精度位居第一，达到了84.7%，且第二名仅为73.8%^[14]。AlexNet神经网络作为识别效果较好的图像识别模型，包含八个带权重的层^[15]，五个卷积层，三个全连接层。AlexNet神经网络模型图如图1所示，前五层的卷积层主要对图像进行卷积、池化或修正线性单元等特征检测操作，最后三层是用于分类的全连接层。

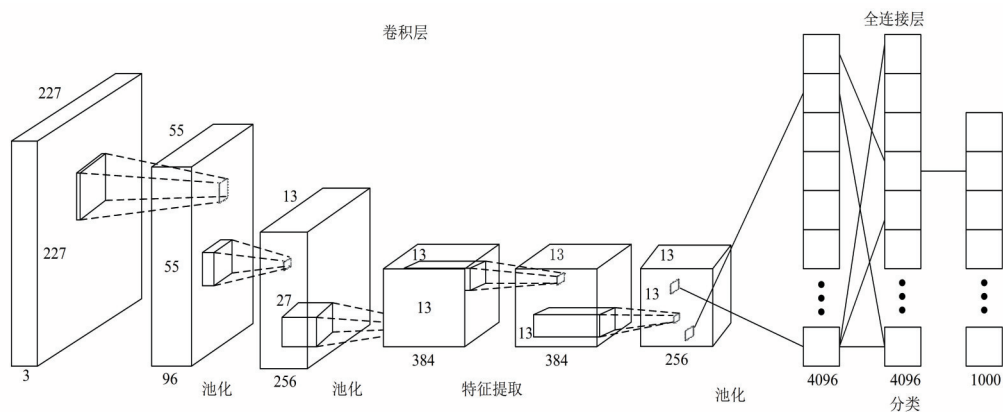


图1 AlexNet神经网络模型图

表1 AlexNet网络层级结构

网络层级	网络结构	特征图大小
Input Image	原始图片	227×227×3
Layer1	Conv+Pool+ReLU	55×55×96
Layer2	Conv+Pool+ReLU	27×27×256
Layer3	Conv+ReLU	13×13×384
Layer4	Conv+ReLU	13×13×384
Layer5	Conv+Pool+ReLU	13×13×256
Layer6	Full+ReLU+Dropout	4096
Layer7	Full+ReLU+Dropout	4096
Layer8	Softmax	1000

2.2 GoogLeNet神经网络模型

GoogLeNet是谷歌公司在2014年提出的一种深度学习结构^[16]。GoogLeNet引入的Inception模块核心思想就是将输入样本通过不同大小的卷积核进行特征提取操作，然后再将各个部分得到的结果进行组合，最后再输入到下一层的网络中，可使计算资源得到更加充分的使用，在同等条件下提取到更丰富的图像特征，得到更高的准确率。图2为Inception原始模块图。

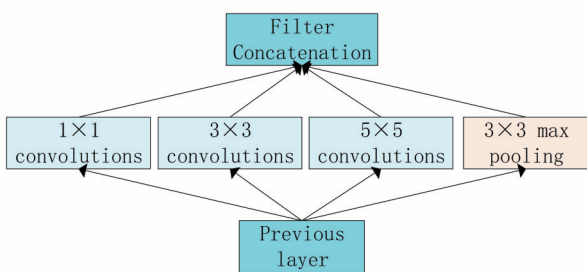


图 2 Inception 原始模块

但在实际应用中仍然存在许多问题, 因此将 Inception 模块进行改进, 具体改进模型如图 3 所示。改进后的 Inception 结构不仅大大减少了参数量, 并增加了 1×1 卷积进行特征降维, 同时随后跟着非线性激活函数, 使得神经网络的泛化能力得到了提升。GoogLeNet 采用 Inception Module, 使用辅助分类节点, 以及用多个小卷积核替代大卷积核, 提高了参数的利用率, 具有良好的分类性能。

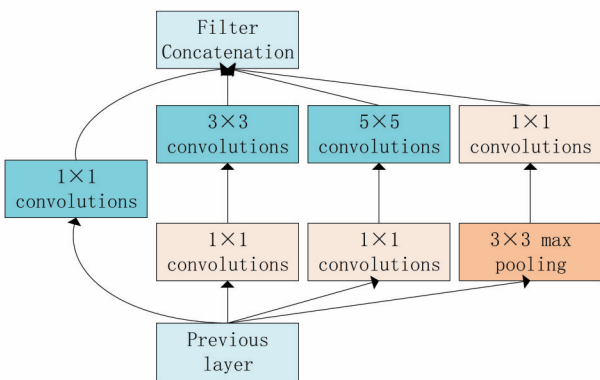


图 3 Inception 改进模型

2.3 ResNet 神经网络模型

从 VGGNet 到 GoogLeNet, 神经网络的深度被公认为影响网络性能最重要的因素之一, 一般认为神经网络模型的层次越多, 网络的准确率越高, 效果越好。但研究表明, 随着网络层次的增加, 在达到一定程度后, 会出现了网络性能退化现象, 效果也会随之变差。这是由于随着模型的加深, 网络模型的训练复杂程度也会随之加大, 深层的卷积神经网络可能会出现梯度向前传递衰减甚至消失为 0 的情况, 网络层数越多, 这种情况造成的影响就会越大。为了解决梯度消失所造成的影响, He 等^[17]等提出了

残差模块, 其点睛之笔就是在某些神经网络层之间添加一条回路, 这样就可以让传递的信息得到对比, 避免部分卷积操作使得梯度消失^[18], 图 4 中这种类似于电路中的短路连接就是残差模块中重要的创新点。当输入 X 时, 其学习到的特征为 $H(x)$, 希望学习到的残差 $F(x)=H(x)-x$, 那么原始的学习特征就应为 $H(x)=F(x)+x$ 。最理想的状态是将网络模型做了残差为零的恒等变换。

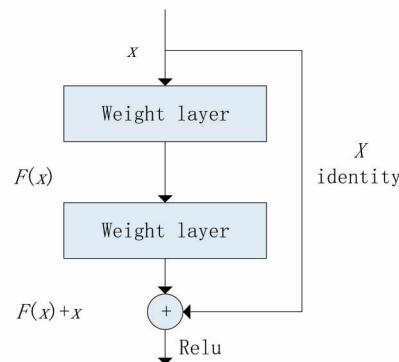


图 4 残差模块

3 迁移学习的实现

为了解决两个完全不相关领域之间的迁移学习而产生的神经网络参数不匹配问题, 参数的调整是非常重要的步骤。迁移学习主要可分为以下几个步骤^[19]:

(1) 选取大量标注完整且与目标领域相关联的数据集, 用其对神经网络模型特征提取层进行参数训练, 用训练好的模型对目标样本进行有效的特征提取;

(2) 贴合实际任务对神经网络剩下的全连接层以及分类层进行有效的自定义重构;

(3) 最后根据实际情况, 使用最终任务样本对最后重构好的网络模型进行训练和参数微调, 完成迁移学习。

因为实验数据集相对较小, 且选用的图像与源域的各类图片存在差异性, 所以采用冻结模型的方式对模型进行微调。AlexNet 的迁移过程如图 5 所示, 将 AlexNet 的最后三层替换为与食道癌类别数目相同的节点数的全连接层、softmax 层和分类输出层。

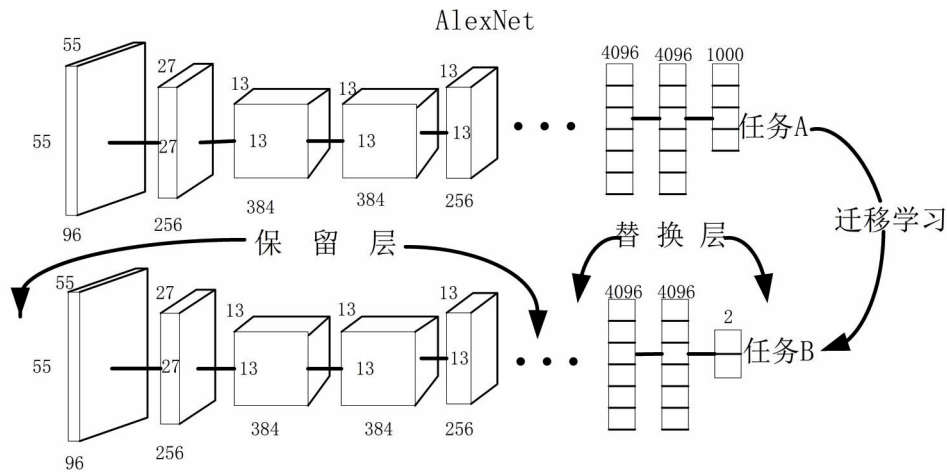


图5 AlexNet迁移过程

通过使用源域上万张图片对网络中的参数进行预训练，为了加快后期目标任务训练的收敛速度，只需要对这些参数进行微调即可。在整个网络中替换层参数占很小一部分，即可使用小数据集进行训练。

4 结果与分析

电子计算机断层扫描(Computed Tomography, CT)是食管癌诊断评估常用的影像手段，一般认为当食管壁的厚度大于5mm或者出现不对称增厚时应考虑可能为食管癌^[20]，但在前期病发中会考虑用食道内窥镜。数据来源于2020年Kaggle^[21]上发布的食管内窥镜图像数据集，该数据集包含非病理性患者的食管内窥镜图像1689张，病理性患者的食管内窥镜图像8973张。食道内窥镜图像如图6所示。



图6 食道内窥镜图像

首先考虑到数据集中食道内窥镜图像在获取过程中，因为操作设备、患病案例、现场情况的不同，对图像进行逐一的挑选检查。对挑选检查后的图像进行样本标记，食道癌患者的内窥镜图像设为标签为1的正样本，未患食道癌的健康内窥镜图像设为标签为0的负样本。将标记好的图像使用OpenCV统一裁剪到224×224的大小，然后对所有的图像进行归一化处理，最后进行简单的数据增强操作。研究表明在对图像进行几何变换以及缩放操作时，对图像的特征提取不会产生影响^[22]，所以不必担心数据增强会对后面的神经网络学习造成影响。最后将预处理好的图像进行简单的打乱，按照要求划分为需要的集合。

4.1 实验环境

本次实验深度学习的基本框架，使用国内的PaddlePaddle2.1.2以及python3.8作为编程工具。选择PaddlePaddle的主要原因是其对于国人友好的API，以及框架中可以加载结构优化以后的各类经典的神经网络模型，如LeNet、AlexNet、GoogLeNet等。AlexNet的正确率相较于VGG、GoogLeNet、ResNet略微偏低，但在验证集中也可达到70%以上^[21]。实验中的样本较少，所以选择一种基于迁移学习和AlexNet的一种微调模型，与从零开始训练的AlexNet、GoogLeNet、ResNet形成对比，验证迁移学习的可行性以及必要性。本实验网络用到的激活函数有ReLU函数，公式为

$$f(x) = \max(0, x) \quad (1)$$

其中: x 是输入变量。

与传统的激活函数 Sigmoid 函数、Tanh 函数相比, ReLU 在梯度下降和反向传播上具有更高的效率。ReLU 是非饱和激活函数, 因此当数值过大或过小时, 导数不会接近于 0, 可以避免梯度消失以及过拟合。

实验中的损失函数为交叉熵函数, 公式为

$$L = \frac{1}{N} \sum_i -[y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \quad (2)$$



图7 食道癌内窥镜图中间层可视化

对比发现, 输入类型差别很大的图像, 通过前面低卷积层运算后, 仍可通过轮廓区分图像类别; 但在高层卷积池化以后, 图像的轮廓变得非常抽象模糊, 肉眼已经难以区分。由此可以看出迁移学习利用的是用于源领域训练好的通用特征提取的底层卷积网络, 结合用于特定任务的定向特征提取的高层卷积网络, 来完成目标样本的指定任务, 侧面验证了迁移学习的可行性。

4.3 模型训练效果对比

根据前文将四个模型分别训练, 并对四个模型的训练集损失值、验证集损失值、测试集损失值进行简单的对比, 然后结合验证集以及

其中: y_i 表示样本的标签, 正类为 1, 负类为 0。
 p_i 表示样本预测为正类的概率。

交叉熵函数不仅在反向传播中的结果简洁, 而且配合输出层的激活函数还能提升深度学习的学习速度。

4.2 中间卷积层输出可视化

为了对卷积神经网络的训练过程有更详细的理解, 这里进行了神经网络模型中不同卷积层的输出。食道癌患者的食道内窥镜图在 AlexNet 模型中间层的输出结果如图 7 所示, 图 8 为其它图像在 ResNet 模型中间层的输出可视化。



图8 其它图像中间层可视化输出

测试集的准确率作为综合评价方法。如图 9 训练集损失值对比图所示, 可以看出传统的深度学习除了 AlexNet 以外, GoogLeNet 和 ResNet 损失值都相对较低且下降平缓, 迁移微调以后的 AlexNet 得到的效果明显优于从零开始训练的 AlexNet 网络。

由图 10 验证集损失值对比图可以看出, 在相同的迭代次数从零开始的 AlexNet 损失值相对较高, 收敛值不理想, 而迁移学习的 AlexNet 收敛迅速, 与结构更加优良的 GoogLeNet 和 ResNet 几乎不相上下。对于测试集损失值, 从图 11 四个模型测试集损失值对比图可以看出, 从零开始训练的 AlexNet 损失值开始于较高位置, 且损失值收敛的过于异常, 且波动较少, 而基于迁

移学习的 AlexNet 以及从零开始训练的 GoogLeNet 以及 ResNet 损失值开始较低, 收敛速度稳定。

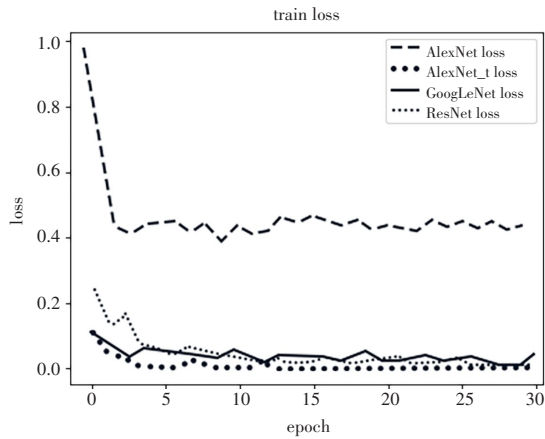


图 9 模型训练集损失值对比图

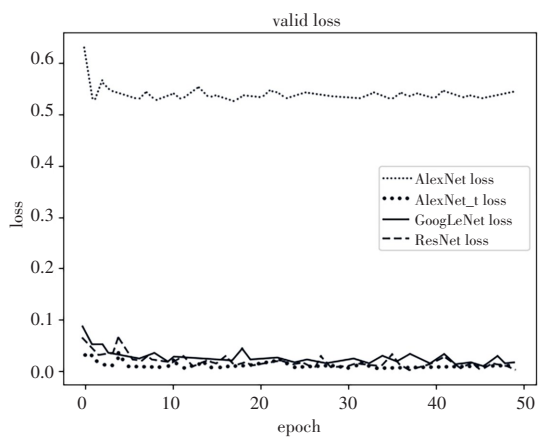


图 10 模型验证集损失值对比图

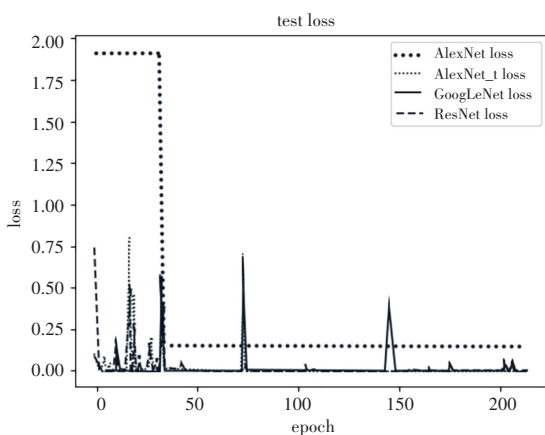


图 11 模型的测试集损失值对比图

四个模型的训练准确率对比如表2所示, 通过观察表2以及上图发现, 除了对于分类有着良好优势的 GoogLeNet 以及 ResNet 网络模型外, 迁移学习的 AlexNet 模型识别效果较好, 验证集和测试集的准确率高达 93.21% 和 91.26%, 明显优于从零开始训练的 AlexNet 模型 78% 的准确率。经过迁移学习以后的 AlexNet 神经网络模型对食道内窥图像有着更好的特征提取和泛化能力, 侧面验证了迁移学习应用于小数据集医学领域的可行性。

表 2 四个模型训练准确率对比

模型	验证集准确率	测试集准确率
AlexNet	78.1	84.21
AlexNet(迁移学习)	93.21	91.26
GoogLeNet	95.13	94.22
ResNet	96.73	95.84

5 结语

本文提出一种基于迁移学习和 AlexNet 对标注少、数据集小, 且数据相对不平衡的食道癌数据集进行识别的方法。与传统直接从零开始训练的 AlexNet、GoogLeNet、ResNet 网络模型相比较, 实验结果表明, 通过迁移学习模型验证集和测试集的准确率分别为 93.21% 和 91.26%, 与其它三个模型相比优势明显, 验证了迁移学习在数据量少、数据不平衡、标注信息少的数据集上的可行性和优良性。

实验使用的数据集较小、分类不够细致、健康的食道图像差异过少, 都是可能对模型的准确率产生影响的原因。为了能够提高神经网络模型在识别食道癌图像中的识别率, 在以上问题无法得到解决的情况下, 使用结构更为优良的神经网络对小数据集进行迁移学习成为了更好的选择, 能够更快、更方便地提高模型的训练速度和准确率。

参考文献:

- [1] 杨春霞. 生活方式和基因多态与食道癌的关系[D]. 成都: 四川大学, 2005.
- [2] 吴磊. 基于深度卷积神经网络的食管癌术前影像诊断关键问题研究[D]. 广州: 华南理工大学, 2020.

- [3] CHEN M F, CHEN P T, LU M S, et al. Survival benefit of surgery to patients with esophageal squamous cell carcinoma [J]. *Scientific Reports*, 2017, 7 (1):46-49.
- [4] NASSRI A, ZHU H, MUFTAH M, et al. Epidemiology and survival of esophageal cancer patients in an American cohort [J]. *Cureus*, 2018, 10 (4) : 2507-2508.
- [5] SUN Y, WANG J, LI Y, et al. Nomograms to predict survival rates for esophageal cancer patients with malignant behaviors based on ICD-0-3 [J]. *Future Oncology*, 2018, 15(2):121-132.
- [6] 王跃,王卫东,赵蕾,等. 基于迁移学习的胃镜图像自动识别多分类系统的研究[J]. *中国医疗设备*, 2021, 36(3):81-84.
- [7] JANSEN I, LUCAS M, SAVCI HEIJINK C D, et al. Histopathology: ditch the slides, because digital and 3D are on show [J]. *World J Urol*, 2018, 36(4):549-555.
- [8] PAN S J, YANG Q. A survey on transfer learning [J]. *IEEE Trans Knowl Data Eng*, 2010, 22(10):1345-1359.
- [9] DAI W Y, JIN O, XUE G R, YANG Q, YU Y. Eigen transfer: a unified framework for transfer learning [J]. *Morgan Kaufmann Publishers*, 2009, 14(6):193-200.
- [10] YANG X, WU L, YE W, et al. Deep Learning signature based on staging CT for preoperative prediction of sentinel lymph node metastasis in breast cancer [J]. *Academic Radiology*, 2020, 27(9):8-10.
- [11] VAN RIEL S, VAN DER SOMMEN F, ZINGER S, et al. Automatic detection of early esophageal cancer with CNNs using transfer learning [C] // *IEEE 2018 25th IEEE International Conference on Image Processing (ICIP)*, 2018:1383-1387.
- [12] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks [J]. *Advances in Neural Information Processing Systems*, 2012, 25(2):1097-1105.
- [13] D LÉVY, JAIN A. Breast mass classification from mammograms using deep convolutional neural networks [J]. *CoRR*, 2016, 16(12):542-544.
- [14] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [C] // *Communications of the ACM*. New York: Association for Computing Machinery, 2017:84-90.
- [15] 卢宏涛,张泰川. 深度卷积神经网络在计算机视觉中的应用研究综述 [J]. *数据采集与处理*, 2016, 31(1):1-17.
- [16] CAO F K, BAI T, XU X L. Vehicle detection and classification based on highway monitoring video [J]. *Computer Systems & Applications*, 2020, 29(10):267-273.
- [17] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [J]. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, 2016, 15(12):770-778.
- [18] 庄福振,罗平,何清,等. 迁移学习研究进展 [J]. *软件学报*, 2015, 26(1):26-39.
- [19] 教良忠,马瑞阳,杨学文. 基于DenseNet和ResNet融合的发动机孔探图像分类研究 [J]. *计算技术与自动化*, 2021, 40(3):105-110.
- [20] HONG S J, KIM T J, NAM K B, et al. New TNM staging system for esophageal cancer: what chest radiologists need to know [J]. *Radiographics*, 2014, 34(6):1722-1740.
- [21] 山巅之鹰. Esophageal endoscopy image [DB/OL]. [2022-01-07]. <https://www.kaggle.com/datasets/chopinforest/esophageal-endoscopy-images>.
- [22] ZEILER M D, FERGUS R. Visualizing and understanding convolutional networks [C] // *Computer Vision—ECCV 2014: Part I*. Zurich, 2014:818-833.

作者简介:

凌其颖(1998—),女,四川西昌人,硕士生,主要研究方向为深度学习

通信作者:刘小芳(1969—),女,重庆垫江人,教授,研究生导师,主要研究方向为模式识别与智能系统, E-mail: lxf1969@163.com

张阳(1995—),男,四川南充人,硕士生,主要研究方向为目标检测

收稿日期: 2022-05-21 修稿日期: 2022-07-06

(下转第56页)

文章编号: 1007-1423(2022)19-0026-07

DOI: 10.3969/j.issn.1007-1423.2022.19.004

一种基于单形空间缺失数据的补全方法

刘 冰, 李玮琦

(达州职业技术学院人工智能学院, 达州 635001)

摘要: 在对基于单形空间数据进行多元回归分析时, 回归系数值极不稳定, 导致现有的一般统计分析方法不能恰当地对其解释和处理。本文提出了一种基于单形空间缺失数据的补全方法: 首先定义了单形空间上完备的代数运算体系, 然后运用K均值法对缺失数据进行初始补全, 并对其进行等距对数比变换, 最后运用主成分法再对初始补全值进行修正补全。实例结果表明, 该补全法要比运用其他单一补全方法的效果更好。

关键词: 单形空间; 成分数据; 对数比变换; 主成分分析; 数据补全

基金项目: 四川省教育厅自然科学基金重点资助项目(18ZA0217)

0 引言

成分数据是一种具有比例结构的多维数据, 其数学形式定义为

$$\begin{aligned} \forall X = [x_1, x_2, \dots, x_D]^T \in S^D \\ \text{s.t. } \sum_{i=1}^D x_i = 1, 1 \geq x_i \geq 0 (i = 1, 2, \dots, D) \end{aligned} \quad (1)$$

则称向量 X 为 D 维成分数据^[1], 向量空间 S^D 为单形空间, 其中 X 中的每一个元素 x_i 表示其在整体中所占的比重。相较于普通数据而言, 成分数据除了用于分析整体中各部分数据间的相对关系外, 还有利于揭示普通数据所隐藏的相对信息, 在诸如社会学、经济学、气象学、地质学、医学等领域都有十分广泛的应用^[2-4]。

由公式(1)可以看出, 单形空间中的数据受到两个条件的限制, 一个是有界约束, 另一个是定和约束, 而现有一般的统计分析方法对被分析数据是没有约束要求的。显然, 这就导致现有一般的统计分析方法在单形空间中无效, 一个最主要的原因在于单形空间中数据的协方差矩阵通常是奇异矩阵, 其含义与普通的数据不同^[5]; 另一个原因在于单形空间中的数据总体一般不满足多元正态分布的假设, 这就会导致建立模型十分困难。为了能运用现有一般的统计分析方法去分析单形空间中的数据, 通常

的做法是先进进行预处理, 即经过一定的变换使之成为无约束数据。文献[1]首次提出通过对成分数据的对数比变换建立成分数据的逻辑正态分布模型, 有效地解决了现有一般的统计分析方法对单形空间中数据有界定和约束的限制, 但建立的模型对数据的解释性较差; 文献[6]提出了一种新变换, 即对称对数比变换, 该方法能很好地解释数据, 但在某些情况下容易造成变换后的数据间存在较高的冗余性, 从而损失部分数据信息; 文献[7-8]在此基础上又进行了改进, 即对称等距对数比变换 (isometric log-ratio transformations, ILR), 但该方法对数据的完整性要求较高, 若存在数据缺失, 当补全为0时会造成变换后的数据为无穷的情况, 显然失去了实际的意义。对于成分数据的缺失值补全, 目前一般采用单形空间均值(SM)补全法、极大似然补全法、期望最大填补法、 k 近邻补全法(KNN), 等等。这些方法对于回答信息而言其实现较为容易, 但稳健性差, 结果偏差较大, 补全后的数据冗余度高, 缺少解释性。

消除变换后的数据冗余性的最有效方法是实现单形空间到欧氏空间的正交变换, 为此, 本文首先给出单形空间的代数运算体系, 在给出文献[4,9-11]的相关变换过程的基础上, 着重研究对于成分数据存在缺失值时通过对数比

变换后存在的多重共线性数据的填补方法,并结合主成分分析法提出了一种较为有效的参考解决路径。

1 单形空间的代数体系

为了实现单形空间到欧氏空间的正交变换,本文给出单形空间对于向量的加法、数乘、内积以及距离的如下定义。

(1) 向量的加法:对于任意 $X, Y \in S^D$, 向量 X, Y 的加法运算 \oplus 定义为

$$X \oplus Y = A[x_1y_1, x_2y_2, \dots, x_Dy_D]^T \quad (2)$$

式中, $A[\cdot]$ 为封闭运算, 即

$$A(z_1, z_2, \dots, z_D) = \left[\frac{z_1}{\sum_{i=1}^D z_i}, \frac{z_2}{\sum_{i=1}^D z_i}, \dots, \frac{z_D}{\sum_{i=1}^D z_i} \right]$$

(2) 向量的数乘:对于任意 $X \in S^D$, 任意实数 $a \in R$, a 与向量 X 的数乘运算 \otimes 定义为

$$a \otimes X = A[x_1^a, x_2^a, \dots, x_D^a]^T \quad (3)$$

(3) 向量的内积:对于任意 $X, Y \in S^D$, 向量 X, Y 的内积定义为

$$(X, Y) = \sum_{i=1}^{D-1} \sum_{j=i+1}^D \left(\ln \frac{x_i}{x_j} \ln \frac{y_i}{y_j} \right) \quad (4)$$

(4) 向量的距离:对于任意 $X, Y \in S^D$, 向量 X, Y 的 Aitchison 距离^[10] 定义为

$$d_A(X, Y) = \sqrt{\sum_{i=1}^D \left(\ln \left(\frac{x_i}{\sqrt{\prod_{j=1}^D x_j}} \right) - \ln \left(\frac{y_i}{\sqrt{\prod_{j=1}^D y_j}} \right) \right)^2} \quad (5)$$

2 成分数据的缺失

在实际工作中,待分析处理数据集中某些数据或属性值缺失的原因是多方面的,或者是调查者基于主观的判断认为不重要而丢弃某些数据;或者是由于客观的问卷设计存在瑕疵、录入失误、受访者拒绝回答而没能采集到某些数据;或者是在原始数据的存储过程中,由于设备的故障造成存储数据的不全或失败而丢失某些数据等,使得没能满足设计预期获得详细而全面的资料数据,显然,如果缺失数据占比较大,对于后续的数据分析处理会造成难以估计的影响。

从式(1)易知,若一个成分向量中只有一个

元素值缺失,即可根据定和限制求出该缺失值,因此一般来说,成分数据的缺失值是指某个样本或属性值中至少有两个或两个以上的缺失值。其数学形式化定义如下:

$$Z = (x_{ij})_{n \times D} = [X_1, X_2, \dots, X_n]^T \quad (6)$$

若某个向量 $X_k (k = 1, 2, \dots, n)$ 中至少有两个元素值存在缺失,则称 Z 为缺失数据矩阵。

3 缺失数据的补全

由于主客观等因素的影响,经常会碰到待分析处理的数据集中某个数据或某些属性值出现为零或缺失的情况。对于前者,通常的做法是将其处理为缺失值;而对于后者,一般先要考虑缺失数据的占比情况,若某行(列)缺失数据比超过 90%,一般进行剔除处理,或重新进行该行(列)数据的采集。对于缺失数据比小于 90% 的情况,则对缺失数据进行某种策略的填补。一种经典的填补方法是基于 k 近邻(KNN)方法^[12],即用通过某缺失值的 k 个最近邻样本信息来估算该缺失值;另一种是把缺失值当作一类随机变量或者隐变量,建立概率隐变量模型,然后通过 EM、VI(Variational Inference)或者 MCI (Monte Carlo Inference)来估计缺失值的分布,具体做法是:

对于式(1),进行如下的处理:

$$X_i = [x_{oi}; x_{mj}]^T, O_i + M_j = D \quad (7)$$

其中, x_{oi} 为非缺失值数据, x_{mj} 为缺失值数据,则缺失值 x_{mj} 的分布估计为

$$p(x_{mj}|x_{oi}) = \sum p(x_{mj}|X_i, \theta) p(X_i|x_{oi}, \theta) \quad (8)$$

然后计算该分布的期望值,并将其置为缺失值的估计值。

对于简单的模型,其解析解可用 EM 算法求解;若模型复杂,则可借助 MCI 去进行逼近求解,但无法解决结构带来的不实用的问题。

此外,对于多元线性回归模型,若变量之间线性无关,还可采用回归估计法对缺失值进行填补。但变量之间完全线性无关仅仅具有理论上可能,在实际情况下,变量之间往往存在多重共线性,若直接采用回归估计法,其估计结果会与实际情况相去甚远。

对于多重补全法,文献[9-11,13]给出的方

法较有代表性，下面作简要介绍：

3.1 非对称对数比变换

对于式(1)，定义如下变换：

$$y_i = \ln \frac{x_i}{x_D}, i = 1, 2, \dots, D - 1 \quad (9)$$

其逆变换式为 $x_D = 1 / \left(1 + \sum_{j=1}^{D-1} e^{y_j} \right)$ ，进而有：

$$x_i = \frac{e^{y_i}}{\left(1 + \sum_{j=1}^{D-1} e^{y_j} \right)}, i = 1, 2, \dots, D - 1$$

从式(9)中可以看到，该变换是一个从单形空间 S^D 到欧氏空间 R^{D-1} 上的线性变换，而非正交变换，变换后的 y_i 与变换前的 x_i 不具有一一对应的关系，即存在非对称关系，这就会导致建立的模型不能合理准确地解释数据。

3.2 对称对数比变换

针对非对称对数比变换存在的缺陷，张尧庭^[6]在《成分数据统计分析引论》中提出了一种新变换，使得变换后的 y_i 与变换前的 x_i 具有了一一对应的关系，即存在对称关系，这就使得建立的模型具有了一定的可解释性。其具体变换式如下：

对于式(1)，定义如下变换：

$$y_i = \ln \frac{x_i}{\sqrt{\prod_{j=1}^D x_j}}, j = 1, 2, \dots, D \quad (10)$$

其逆变换式为 $x_D = 1 / \left(1 + \sum_{j=1}^{D-1} e^{(y_j - y_D)} \right)$ ，进而有：

$$x_j = \frac{e^{(y_j - y_D)}}{\left(1 + \sum_{j=1}^{D-1} e^{(y_j - y_D)} \right)}, j = 1, 2, \dots, D - 1$$

从式(10)看到，该变换是正交变换，但当 $0 < x_i < 1$ 时， $\forall [\alpha_1, \alpha_2, \dots, \alpha_D]^T \in S^D, \exists \sum_{i=1}^D \alpha_i y_i \neq 0$ ，即变换得到的数据存在一定的相关性，导致了变量间协方差矩阵不满秩，从而使得基于协方差结构的统计方法无效，在实际应用中，应当避免使用该变换对成分数据进行预处理。

3.3 对称等距对数比变换

文献[4, 11, 13]又在对称对数比变换基础上进行了改进，即对称等距对数比变换 (isometric log-ratio transformations, ILR)，具体如下：

对于式(1)，定义如下变换：

$$y_i = irl(x_i) = \left(\frac{D-i}{D-i+1} \right)^{\frac{1}{2}} \ln \frac{\left(\prod_{p=i+1}^D x_p \right)^{\frac{1}{D-i}}}{x_i} \quad i = 1, 2, \dots, D - 1 \quad (11)$$

容易得出：式(11)的逆变换式为

$$x_D = \exp \left(\sum_{i=1}^{D-1} \frac{1}{\sqrt{(D-i+1)(D-i)}} y_i \right)$$

进而有：

$$x_j = \exp \left(\sum_{i=1}^{j-1} \frac{1}{\sqrt{(D-i+1)(D-i)}} y_i - y_j \left(\frac{D-j}{D-j+1} \right)^{\frac{1}{2}} \right) \quad j = 1, 2, \dots, D - 1 \quad (12)$$

从式(11)可以看出，该变换实现了从单形空间 S^D 到欧氏空间 R^{D-1} 的正交变换，确保了在变换后的空间中运用传统的统计分析方法进行合理的模型建立。但在 $x_i = 0$ 时，对应的 y_i 的结果将为无穷，失去了实际的意义，对后续的进一步分析处理造成了障碍。

4 成分数据的补全

一般情况下，对于缺失数据不宜贸然进行删除处理，通常需要采用某种方法进行补全操作。常用的方法有：均值补全法、极大似然估计法、多重补全法等，其中多重补全法是通过估计出待补全的值加上不同的噪声来得到补全值。对于成分数据缺失值的补全，Hron等^[12]提出的k近邻法较有代表性，该方法是通过用Aitchison距离来寻找到含缺失值样本的k个近邻，并用该k个近邻的中位数来进行初始补全，然后用最小二乘法来进行迭代补全。本文在前面定义的单形空间的加法运算 \oplus 以及数乘运算 \otimes 的基础上结合文献[12]的方法，提出一种基于单形空间缺失成分数据的补全方法，同时运用主成分分析法，处理将成分数据变换为一般数据后可能存在的多重共线性的情况。

4.1 多重共线性

在进行多元回归分析时,若某些解释变量之间存在严格或近似的线性关系,其样本点或属性值的一个微小改变都会极大地扰动回归系数的估计值,使得回归系数极不稳定^[14]。因为某些解释变量之间存在的强相关关系将极大地降低 $Z^T Z$ 的可逆性,大多数情况变得不可逆,即使通过某种计算使其变得可逆,其逆矩阵的特征值也往往会较大,导致标准误差值也较大,进而降低了参数估计值的精度,无法得出稳定的回归模型,回归系数及符号也与实际情况相去甚远。

检测多重共线性的方法主要有:

(1) 通过计算自变量间的相关系数与显著性来进行判断,即若某些变量间的相关系数显著,则认为它们之间可能存在多重共线性问题。

(2) 使用回归分析中的方差膨胀系数(Variance inflation factor, VIF)值来进行判断, VIF的计算公式为 $VIF = 1/(1 - R_i^2)$ 。其中, R_i 为负相关系数。自变量之间共线性程度与VIF值存在较强的正相关关系。根据Hair(1995)标准,当 $VIF \leq 10$ 时,模型的多重共线性较弱;当 $10 < VIF \leq 100$ 时,模型的多重共线性较为严重;当 $VIF > 100$ 时,模型的多重共线性很严重。

(3) 容忍值(Tolerance)法,也是较为常用的方法。其计算公式为 $Tol = 1/VIF$ 。显然,其与方差膨胀系数法的判定标准相反,自变量之间共线性程度与 Tol 值存在较强的负相关关系。在实际中,通常为 Tol 指定一个阈值,确保小于阈值的变量间的相关系数矩阵可逆,使回归系数的估计值具有较强的稳定性。该方法的缺陷在于 Tol 阈值的确定存在随意性,没有一个统一的标准。

(4) 主成分回归法,对于矩阵(6),设 $Z^T Z$ 的特征值为 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n > 0$,称 $h = \lambda_1/\lambda_n$ 为 $Z^T Z$ 的条件数,一般地,若 $h < 100$,则认为模型的多重共线性程度较小;若 $100 < \lambda_1 < 1000$ 则认为模型的多重共线性程度较强;若 $h > 1000$,则认为模型的多重共线性程度严重。

需要说明的是,在现实工作中,获得的数据集一般都存在多重共线性,只是程度不同而已,对于共线性程度较小或一般的问题可以不

必采取措施。另外,如果学得模型的拟合度好,也可不需处理多重共线性问题。

4.2 单形空间上的均值补全

根据公式(5)易知,若两个样本各自成分数据子集相似,则它们之间的Aitchison距离可以用其对应子集的Aitchison距离大约表示。即 $d_A(x_i, x_j) \approx d_A(\mathbf{X}_{M_i}, \mathbf{X}_{M_j}) \approx d_A(\mathbf{X}_{O_i}, \mathbf{X}_{O_j})$,其中: $\mathbf{X}_{M_i}, \mathbf{X}_{M_j}$ 和 $\mathbf{X}_{O_i}, \mathbf{X}_{O_j}$ 分别是样本 x_i, x_j 各自所对应的缺失值和确定值成分数据子集。

下面根据第2节的相关定义及文献[12]的方法给出 x_i 的某一缺失成分 $x_{m_i} \in \mathbf{X}_{M_i}, m \in M$ 的补全步骤:

(1) 根据Aitchison距离找到含缺失值 x_i 相应子成分 \mathbf{X}_{M_i} 的 $k (k < n)$ 个最近邻,并记为 $x_{M_i}^{[1]}, x_{M_i}^{[2]}, \dots, x_{M_i}^{[k]}$,其对应的 k 个全样本依次为 $x_i^{[1]}, x_i^{[2]}, \dots, x_i^{[k]}$ 。

(2) 根据定义1和定义2,计算出 k 个全样本的均值:

$$\bar{x}_i = \frac{1}{k} \otimes (x_i^{[1]} \oplus x_i^{[2]} \oplus \dots \oplus x_i^{[k]}) \quad (13)$$

(3) 求出 x_{m_i} 的补全值 \tilde{x}_{m_i} :

$$\tilde{x}_{m_i} = \frac{I_j^T \bar{x}_i}{\sum_M (I_j^T \bar{x}_i)} \quad (14)$$

其中: $I_j = (0, \dots, 1, \dots, 0)^T \in R^n$ 的第 j 个元素为1, $j = 1, 2, \dots, n$ 。

4.3 基于主成分的补全

在大多数情况下,在单形空间上由子成分的Aitchison距离对缺失值进行均值补全后的成分数据存在多重共线性,基于4.1节所述,下面采用主成分回归分析法对上节初始补全后的成分数据再次进行修正补全,主要步骤如下:

(1) 将含有缺失成分的样本 x_i 和其 k 个最近邻样本 $x_i^{[1]}, x_i^{[2]}, \dots, x_i^{[k]}$ 组成一个单形空间矩阵,并将缺失值 x_{m_i} 初始补全后的 \tilde{x}_{m_i} 变换到第1行第1列,记为: $\mathbf{A}_{(k+1) \times D}$ 。

(2) 根据公式(11),将单形空间矩阵转换为欧式空间矩阵,如下所示:

$$\text{irl}(\mathbf{A}_{(k+1) \times D}) = \begin{bmatrix} \alpha & \mathbf{B}^T \\ \omega & \mathbf{A} \end{bmatrix} \quad (15)$$

其中, $\alpha = \text{irl}(\tilde{x}_{m_i})$, \mathbf{A} 为一 $k \times (D - 1)$ 阶矩阵,令:

$$A = \begin{bmatrix} z_{i2}^{[1]} & \cdots & z_{i(D-1)}^{[1]} \\ z_{i2}^{[2]} & \cdots & z_{i(D-1)}^{[2]} \\ \vdots & \ddots & \vdots \\ z_{i2}^{[k]} & \cdots & z_{i(D-1)}^{[k]} \end{bmatrix} \quad (16)$$

(3) 对矩阵 A 做主成分分析, 其协方差矩阵记为 $\Lambda = \sum_{u=1}^k \sum_{v=1}^k (z_u - \bar{z})(z_v - \bar{z})^T$, 其中, z_u, z_v 为 A 的行向量, \bar{z} 为 A 的行向量均值。

(4) 计算 Λ 的前 p 个主成分, 依次为 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$, 则响应变量 Y 与 $z_{ij}^{[m]}$ 间的关系为

$$Y = \sum_{j=2}^{D-1} \beta_j z_{ij}^{[m]} + \varepsilon \quad (17)$$

其中: $m = 1, 2, \dots, k, \varepsilon$ 为误差项。

(5) 通过上式得到 β_j 的估计值 $\hat{\beta}_j$, 计算到缺失值 x_{mi} 的补全值为

$$\hat{x}_{mi} = \sum_{j=2}^{D-1} \hat{\beta}_j z_{ij} + \varepsilon \quad (18)$$

(6) 运用公式(12)将数据 x_{mi} 还原为成分数据, 并通过第(1)步将其调回到原始位置。

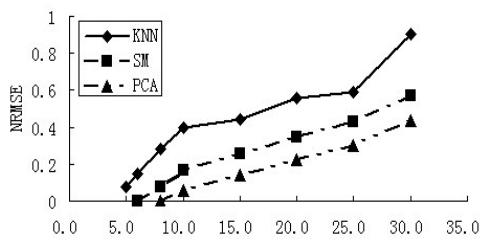
5 评价与比较

5.1 评价指标

为了有效地评价上述方法对数据集中缺失值的补全效果, 本文采用正规化方均根差(the normalized root mean squares error, NRMSE)作为判别准则, 即:

$$NRMSE = \mu \frac{[(v_g - v_r)^2]}{\sigma(v_r)} \quad (19)$$

其中, v_g 为补全值, v_r 为真实值, $\mu(\cdot)$ 为均值, $\sigma(\cdot)$ 为方差。NRMSE 值的大小反映了真实值与补全值之间差距, 若 NRMSE 值较大, 则说明补全值与真实值存在较大差距; 若 NRMSE 接近于 0, 则说明补全值非常接近真实值。



(a1) $MR(p=0.3)$

5.2 比较分析

为了验证前述方法的有效性, 选用文献[1]中 Hongite 数据, 该数据集包含 25 个样本, 每个样本包含 5 个特征: ablite, blandite, cornite, daubite, endite, 根据 4.1 节所述计算得到条件数 $h=2747.238 \gg 1000$, 即认为该数据集存在严重的多重共线性。下面假定该数据集的 ablite 和 cornite 特征数据缺失, 分别运用 k 近邻补全法 (KNN), 单形空间均值 (SM) 和主成分补全法 (PCA) 对缺失值进行补全, 得到比较结果见表 1。

表 1 KNN, SM 和 PCA 补全操作比较结果表

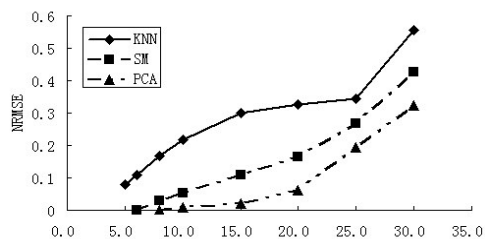
方法	Ablite	Cornite	评价
原始数据	48.8	3.8	NRMSE
KNN	45.7	3.5	0.0967
SM	47.4	4.7	0.0519
PCA	47.9	3.9	0.0270

从表 1 可以看出, 当条件数 $h \gg 1000$ 时, 用 PCA 方法进行补全的结果最好, KNN 的补全结果最差。

根据文献[10]的结论模拟 100 个 5 维的成分数据 $x \sim N_{\varphi}^5(\mu, \Sigma)$, 其中 $\mu = (0, 0, 0, 0, 0)^T$, Σ 是一个主对角线上全为 1, 其余全为 p 的 4 阶方阵。在假定 p 的取值分别为 0.3、0.7、0.995, 缺失率 (MR) 分别为 10%、20% 和 30% 情况下, 分别运用 KNN、SM 和 PCA 方法进行缺失值补全, 并用 NRMSE 进行评价比较, 结果如图 1 所示。

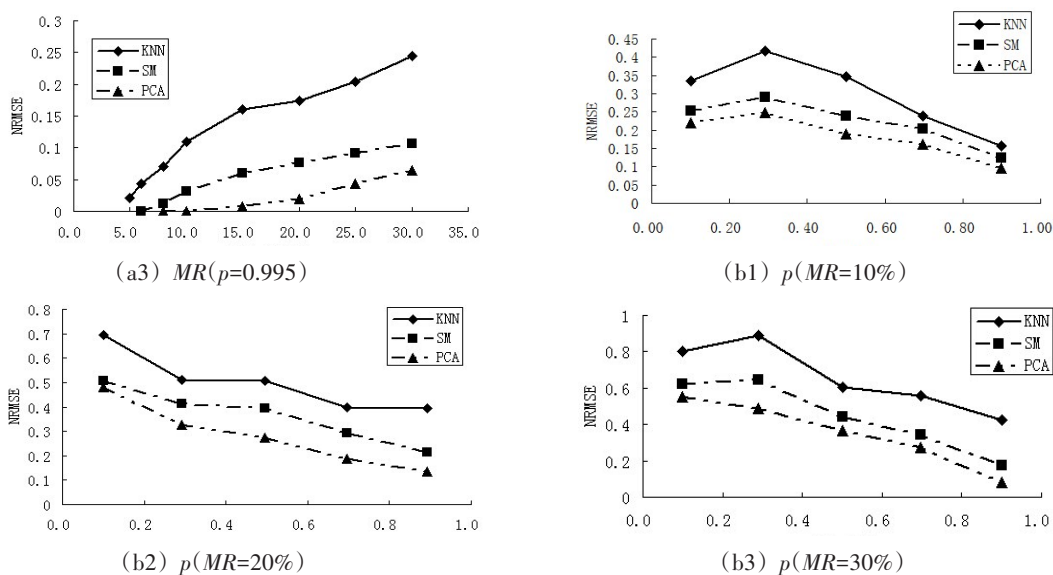
其中, 在图 1 中的图 a1~a3 是在 MR 不同 p 一定时三种方法的补全比较结果, 图 b1~b3 是在 p 不同 MR 一定时三种方法的补全比较结果。

从图 b1 与图 b2 中可以看出, PCA 比 KNN, SM 的结果都要好。这三种补全法在 MR 一定时, p 与 NRMSE 呈负相关关系, 也就是说若数据间



(a2) $MR(p=0.7)$

图 1 MR 与 P 分别取值时三种方法的补全比较结果

图1 MR 与 p 分别取值时三种方法的补全比较结果(续)

的多重共线性程度越大,无论哪种方法的补全效果都越好。而在 p 一定时,三种方法的 MR 与 $NRMSE$ 呈正相关关系。作为初始的补全法,KNN法明显比SM差,随着 MR 的增大,结果会更差。但在 MR 变大时,PCA法明显比SM效果好,随着 MR 的增大,结果会更明显。

6 结语

基于单形空间完备的代数体系提出的等距对数比变换是一个正交变换,该变换既克服了非对称对数比变换改变内积及距离等几何概念的缺陷,同时,又避免了对数比变换导致的多重共线性给多元分析方法带来的不利影响。对于含有缺失值的多元数据来说,无论是基于模型还是基于距离,多变量补全法比单变量补全法结果更为准确:在单形空间上先进行均值补全,然后运用等距对数比对补全后的数据进行变换,最后再对变换后的数据运用主成分法进行第二次补全,实例分析表明,再次运用主成分法进行二次补全要比其他方法的效果更好。

参考文献:

[1] AITCHISON J. The statistical analysis of compositional data[M]. New York:Chapman and Hall,1986.
 [2] BUCCIANTI A,PAWLOWSKY-GLAHN V. New perspectives on water chemistry and compositional data analysis [J]. Mathematical Geology, 2005, 37:

703-727.

- [3] JARAUTABRAGULAT E, HERVADASALA C, EGOZCUE J J, et al. Air Quality index revisited from a compositional point of view [J]. Mathematical Geosciences, 2015, 48:581-593.
 [4] SNYDER R D, ORD K, KOEHLER A B, et al. Forecasting compositional time series: a state space approach [J]. International Journal of Forecasting, 2017, 33(2):502-512.
 [5] BILLHEIMER D, GUTTORP P, FAGAN W F. Statistical interpretation of species composition [J]. Journal of the American Statistical Association, 2001, 96 (456):1205-1214
 [6] 张尧庭. 成分数据统计分析引论[M]. 北京:科学出版社,2000.
 [7] PAWLOWSKY-GLAHN V, EGOZCUE J J, TOLOSANA-DELGADO R. Modeling and analysis of compositional data [M]. New Jersey: John Wiley & Sons, Ltd, 2015.
 [8] KYNCLOVÁ P, FILZMOSER P, HRON K. Modeling compositional time series with vector autoregressive models [J]. Journal of Fore-Casting, 2015, 34:303-314.
 [9] 郭丽娟,王惠文,关蓉. 基于等距logratio变换的成分数据判别分析方法[J]. 系统工程, 2016, 34(2): 153-158.
 [10] AITCHISON J, BARCELÓ-VIDAL C, EGOZCUE J, et al. A concise guide to the algebraic-geometric structure of the simplex, the sample space for compositional data analysis [J]. Proceedings of IAMG, 2002, 2:387-392.

- [11] AITCHISON J. The statistical analysis of compositional data[M]. New York:Chapman and Hall, 1986.
- [12] HRON K, TEMPI M, FILZMOSER P. Imputation of missing values for compositional data using classical and robust methods [J]. Computational Statistics & Data Analysis, 2010, 54(12):3095-3107.
- [13] EGOZCUE J J, PAWLOWSKY-GLAHN V, MATEU-FIGUERAS G, et al. Isometric logratio transformations for compositional data analysis [J]. Mathematical Geology, 2003, 35(3):279-300.

- [14] 王松桂, 史建红, 伊素菊, 等. 线性模型引论[M]. 北京: 科学出版社, 2004.

作者简介:

刘冰(1970—), 男, 四川达州人, 学士, 教授, 研究方向为数据处理、图像技术

李玮琦(1982—), 女, 四川达州人, 硕士, 副教授, 研究方向为数据处理、统计建模及其应用

收稿日期: 2022-06-09 修稿日期: 2022-07-13

Missing Data Completion Method Based on Simplex Space

Liu Bing, Li Weiqi

(Artificial Intelligence College of Dazhou Vocational and Technical College, Dazhou 635001)

Abstract: In the multiple regression analysis of data in simplex space, the value of regression coefficient is very unstable, which leads to the existing general statistical analysis methods can not properly explain and deal with it. In this paper, a method based on missing data in simplex space is proposed: firstly, a complete algebraic operation system in simplex space is defined, then the k-means method is used to complete the missing data, and the equidistant logarithmic ratio transformation is carried out. Finally, the principal component method is used to correct the initial completion value. The example results show that the effect of this method is better than that of other single methods.

Keywords: simplex space; component data; log ratio transformation; principal component analysis; missing data completion

(上接第9页)

Classification of Cervical Cells Based on Transfer Learning and Label Smoothing Strategy

Liu Mei, Zhou Long

(School of Electrical and Electronic Engineering, Wuhan Polytechnic University, Wuhan 430023)

Abstract: In order to avoid the problems of misdiagnosis and leakage and low efficiency of cell recognition in cervical pathological sections, and to improve the efficiency and accuracy of cell recognition, a deep learning model TLS-Net50 (Transfer Label Smooth Net 50) targeting the accuracy and specificity of cervical cell recognition was developed for the characteristics of small cervical cell data set. The algorithm introduces migration learning, which enhances the feature extraction ability of the model and reduces the training time and training cost of the model. To improve the generalization and robustness of the model, a label smoothing regularization strategy is used on top of this. The improved algorithm achieves an average classification accuracy of 98% on publicly available data and outperforms in comparison with other classical deep learning algorithms, which proves that the TLS-Net50 algorithm is superior in the field of computer-aided cervical cell classification and recognition and has some application prospects.

Keywords: cervical cells; convolutional neural network; residual structure; transfer learning; generalization

文章编号: 1007-1423(2022)19-0033-06

DOI: 10.3969/j.issn.1007-1423.2022.19.005

基于 YOLO 框架的农田火源自动检测系统

白玉, 马广焜, 彭新茗, 王楠翔, 白清扬, 孟峻锋, 刘鑫蕊, 朱硕闻

(沈阳工业大学软件学院, 沈阳 110870)

摘要: 农田安全是社会普遍关注的热点话题, 农田火灾也是农田安全隐患中的重要因素。为了降低农田中火源所带来的危害, 提出了针对智能火源检测的农田火源自动检测系统。虽然针对火灾的检测, 已经存在温度、烟雾传感器等检测手段, 但是无法保证监测实时性。针对这一问题, 提出了基于 YOLOv5 框架的农田火源自动检测系统。通过小目标检测技术, 实时检测农田当中的火源隐患, 节约了人工对火源排查的时间。采用最新的 YOLOv5 算法, 此算法有较短的推理时间和训练时间, 大大缩短了从火源产生到人工对火源采取措施的间隔时间。此外, 本系统提供多种接口, 能够与智能农田系统结合, 使农田受到更好的保护。可以将数据整合到云端, 通过任意设备调用云端的接口, 实现系统的多样性, 扩大应用空间。

关键词: 深度学习; 小目标检测; YOLOv5; 火灾检测

基金项目: 辽宁省 2021 年大学生创新创业训练计划项目(S202110142021)

0 引言

农业产业是我国的根本产业, 农业的安全问题关系到国民生活。火灾是农业生产中最可怕的灾难, 火灾会给农作物造成毁灭性的打击, 同时也无法保障人们的人身安全和财产安全。在农作物种植周期中, 由于人为、农业机械故障或自然条件等原因, 有时会引起农田火灾, 但又由于农民对农田火灾发现得不够及时, 火灾快速扩散, 错过了黄金救援时间, 给农民造成极大的经济损失。

国内外, 智能农业系统已经开始广泛应用, 功能全面, 但对消防安全隐患的关注度还有待提高。如果突发意外引起农田火灾, 势必会带来巨大的损失, 而越早发现火源, 将会大大延长火灾的黄金救援时间。YOLOv5 框架在目标检测精度方面有很大的优势, 因此本研究设计的农田火源自动检测系统基于 YOLOv5 网络结构, 针对农田火源研究了自动检测系统, 相比于传统的人工排查, 该系统可以快速响应, 防止火灾进一步扩散。

现阶段国内外对火灾识别普遍采用传统的人工监控方法, 但传统监控在模糊环境如浓烟

下比较容易形成误判。于是我们需要采用目标检测方法, 而目标检测又是深度学习领域中重要的研究方向。由于火源图像的复杂性, 需用训练深层神经网络提高模型对事实的描述^[1-2]。

农业是我国的根本产业, 人们希望把目标检测应用到农业中, 研究农田火源检测的问题, 对于预防农田火灾有着重要意义。在国外, 以色列农业技术发达且具有先进的辅助设备, Gilad Axelrad 等^[1]开发了可实现农业污水污泥回收利用的农田模型, 可以很好地提高水利用率, 但对消防安全隐患的重视度还不够。国内相对智能的农田可以实现农田信息实时采集、灾害预防。王朝泽^[4]设计和实现的智能灌溉系统可以通过传感器精确检测土壤温湿度、环境温湿度、二氧化碳浓度、光照强度等环境参数, 但未考虑到火源检测。

目前 YOLO 框架是最先进的物体检测方法之一, 它能完成不同尺度大小的目标检测的能力。该方法已用于安全帽佩戴检测^[5]、白细胞检测^[6]等小目标检测。本文提出的农田火源自动检测系统, 以火源作为目标搭建检测系统, 基于 YOLOv5 框架通过图像识别小目标, 连接终

端实时报警，防止火灾进一步扩散。相比于传统的监控方法，本方法具有可以降低时间浪费率，同时避免产生大量浓烟污染环境，防止引起火灾造成生命危险等优点。

1 系统框架结构设计

本系统在设计上主要分成三部分：终端部分、服务器部分和图像识别接口部分。终端部分主要通过连接电脑、手机两个设备来进行实时监控；服务器部分可将采集到的视频传递给图像识别部分，上传图像给终端部分起到中转作用；图像识别接口部分获取视频进行模型训练，且可以通知报警。示意图如图1所示，通过网络摄像头将图像以视频流的方式实时传输给服务器，服务器调用YOLOv5的算法接口，对图像中的火焰进行标注，将所标注信息渲染到前端页面当中，当所标注的火焰图像置信度达到一定阈值时，发出指令使蜂鸣器报警。

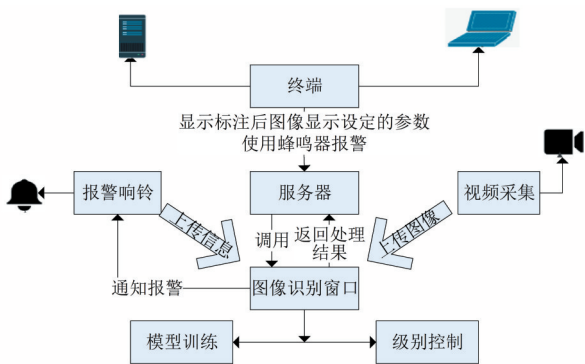


图1 系统示意图

1.1 服务器部分

服务器为租用的腾讯云服务器，视频采集部分将视频图像上传到服务器端，服务器将连接、上传图像，实时报警给终端，例如手机或电脑，服务器的后台程序接收到移动端数据后，将其存入后台供客户端调用，也可以将标注后的图像所设定的参数显示在终端供用户查询使用。如果发现火源，蜂鸣器将报警，同时调用本地服务器接口实时查看农田的录像，并可收到报警提醒。服务器可以调用图像识别接口，同时也可以收到处理结果的反馈。

1.2 图像识别接口部分

当图像识别模块识别到感兴趣目标时会通知报警模块传递信号，报警模块在收到信息后报警鸣铃，同时把信息上传至服务器，获取级别服务器。图像识别部分基于深度学习的目标检测算法来对火源进行识别和检测。

1.3 终端部分

终端连接手机摄像头和监控实时检测火源，监控识别到火源后会在客户端界面上报警，自动标记火源发生处，将级别调整信息传递给服务器，服务器针对不同级别选择不同的训练模型，进而调整报警强度。

2 系统功能设计

该火灾报警系统由四个模块组成，分别为视频采集模块、图像识别模块、报警响铃模块和服务器监控模块，如图2所示。

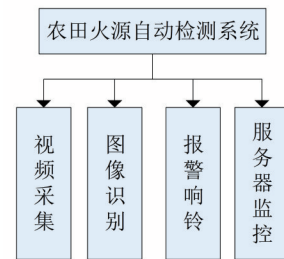


图2 系统功能图

2.1 视频采集模块

远程实时视频监控采集，搭建采集系统平台，收集视频信号源进行视频采集。设计好服务器端软件、客户端软件，确保此模块通过多个摄像头实时采集农田内场景，使整个农田都能被监控到，将采集到的视频上传到服务器。

2.2 图像识别模块

将采集到的视频实时传输到图像识别模块，即对图像数据内火源进行检测，是该系统的检测核心。使用的是YOLO框架，基于YOLOv5框架来对图像数据内火源进行检测，总共将其划分为四个通用的模块：输入端、Backbone、Neck网络与Head输出端。输入端，即为输入的图片；基准网络，主要用来提取通用的特征表示；Neck网络，可以增加特征的多样化和鲁棒

性; Head输出端,即目标最终检测结果的输出级别控制模块。本模块还包含训练模型,训练模型通过强化学习集、训练集的方式对火源进行分类训练。

2.3 报警响铃模块

此模块可以调用服务器接口,检测到火源后快速响应,及时将信息传达给服务器。然后将信号送至火灾报警控制器,发生报警,同时也能够通过控制器联动其他有关的设备动作,最终上传信息到终端,用户可以及时查看到火警信息。

2.4 服务器监控模块

负责开发多个接口,确保实时监控、上传图片、实时报警,起到总控作用,是系统化、直观化的管理火源检测系统。

3 关键技术的实现

YOLOv5的整体网络结构一共分为四个部分,输入端、Backbone(主干部分)、Neck和输出端(Prediction),图3所示为YOLOv5的整体网络结构图。

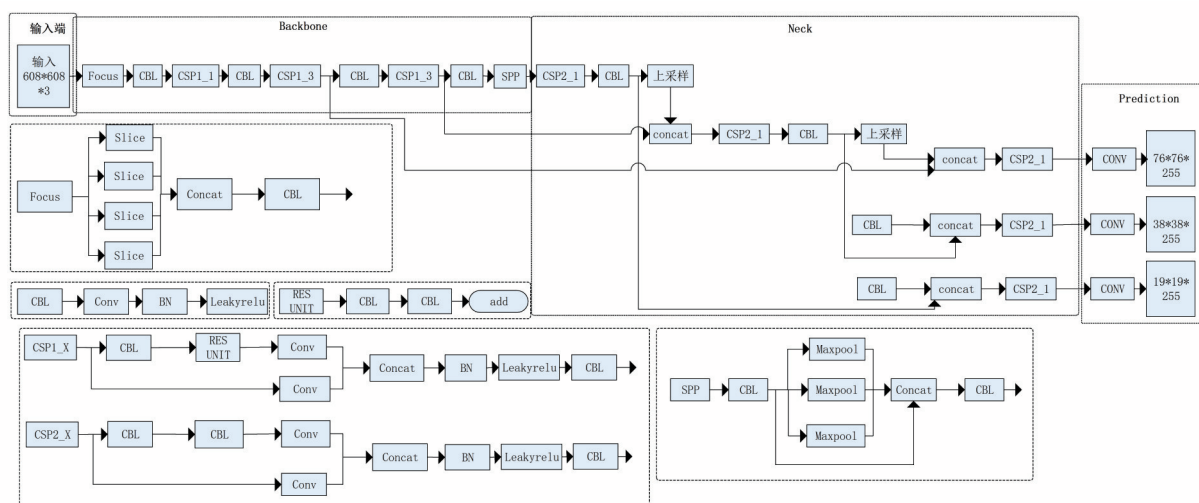


图3 YOLOv5整体网络结构图

3.1 输入端(Head)

首先输入图像,其大小为608x608,本阶段包括一个图片预处理阶段,将输入后的图片缩放到网络的输入大小,对其进行归一化等操作。

(1) Mosaic数据增强:使用随机裁剪、随机缩放、随机排布的方法来对图像进行拼接,然后达到对小目标的检测,数据增强能够丰富数据集,同时也增加网络鲁棒性,提高训练速度。

(2) 自适应图片缩放:首先将原始图片全部缩放到一个标准尺寸,根据输入图像的尺寸大小来进行自适应填充,然后再将其送入检测网络中,对其原始图像自适应地添加最少的黑边,可以减少图像高度上两端的黑边,同时也减少了其计算量,极大提升了目标的检测速度。

(3) 自适应锚框计算:采取anchor的计算方法,在初始锚框的基础上输出预测框,利用K均值聚类(K-means)方法聚类出多种尺寸大小的anchor,在每次训练时计算不同训练集中的最佳锚框值。在预测时采用自适应图片大小的缩放模式,在减小黑边的同时,也提升了预测速度。下面是本系统设定的anchor:

- [10, 13, 16, 30, 33, 23] # P3/8
- [30, 61, 62, 45, 59, 119] # P4/16
- [116, 90, 156, 198, 373, 326] # P5/32

3.2 Backbone

Backbone阶段主要特征提取图片,对图片进行卷积操作等,包括Focus和CSP等结构。

3.2.1 Focus 结构

首先将 608×608×3 大小的图像输入到 Focus 结构，然后采用 slice 操作对输入的图片进行裁剪，之后采用切片操作输出可以变成 304×304×12 大小的特征图，再进行一次卷积操作，最后形成 304×304×32 大小的特征图，Focus 结构提升了图片特征提取的质量。

3.2.2 CSP 结构

CSP 结构共分为两种，分别是应用于 Backbone 网络的 CSP1_X 结构和应用于 Neck 网络的 CSP2_X 结构，其过程是先将基础层的特征映射分成两部分，然后采用跨阶段层次结构将其合并在一起。CSP 结构能够降低计算量，并提高准确率，同时增强了网络对特征的融合能力。

3.3 Neck 网络

在卷积过程中，网络层数越深，该目标的特征信息就越强，模型对目标的预测效果会更好，但目标的位置信息也会因此越弱，而且容易使小目标的信息损失，所以需要进行不同尺度的预测来减少损失。

Neck 网络层采用 FPN 和 PAN 的结构对特征进行多尺度融合。FPN 结构是通过自顶向下进行上采样，让底层的特征图包含更强的火源语义信息；PAN 结构自底向上进行下采样，让顶层特征包含强火源位置信息，最后让两个特征进行融合，使不同尺寸的特征图都包含强火源语义信息和强火源特征信息，因此对不同尺寸的火源图片的准确预测有了保证。图 4 所示为 FPN 和 PAN 的结构示意图。

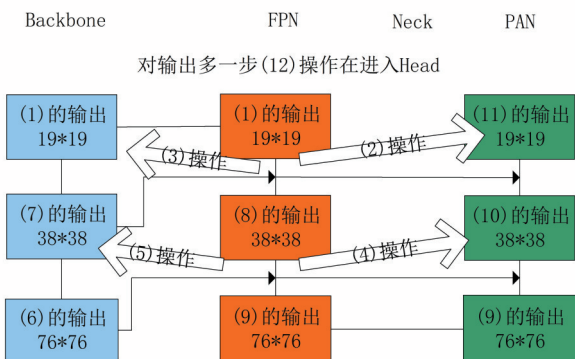


图 4 FPN 和 PAN 结构示意图

3.4 输出端(Prediction)

输出端分别对 Neck 的三个输出进行卷积操作，利用 CIOU_Loss 函数、nms 非极大值抑制和 FocalLoss 损失函数进行最终的预测输出。

3.4.1 CIOU_Loss

本系统采用 CIOU 来计算交并比，因为 CIOU 对于重叠目标的识别穿透性更好，而且优化了两识别框不重合和过度覆盖的情况。CIOU 的计算公式如下：

$$CIOU_Loss = 1 - CIOU$$

$$= 1 - \left(IOU - \frac{Distance_2^2}{Distance_C^2} - \frac{v^2}{(1 - IOU) + v} \right)$$

其中，IOU 是预测框和真实框的交并比，v 是衡量长宽比一致性的参数，C 是图片的最小外接矩形，Distance_C 是图片的对角线距离，Distance_2 是两个中心点的欧氏距离。

3.4.2 nms 非极大值抑制

为了得到最后的框序列信息，首先要将所有的矩形框按照不同的类别标签分组，组内按分数高低进行排序，将得分最高的矩形框先放入结果序列，然后通过遍历剩余矩形框计算与当前得分最高的矩形框的交并比，如果大于预设的阈值则剔除，对剩余的检测框重复上述操作，直到处理完图像内所有的候选框，该阶段对那些有重叠图像的目标检测效果更好。

3.4.3 FocalLoss 损失函数

使用 FocalLoss 进行评价目标框和预测框的类损失和置信度损失，其可以减少容易分类的样本对损失函数的影响，重点训练区分较难分类的样本。FocalLoss 的公式为

$$FL(p_i) = -(1 - p_i)\gamma \log(pt)$$

其中， $\gamma >= 0$ 。 $p_i = \begin{cases} p & \text{if } y = 1 \\ 1 - p & \text{otherwise} \end{cases}$ ，p 表示预测样本属于 1 时的概率，y 表示标签，y 的取值范围为{-1, +1}。

4 系统实验及输出结果

4.1 数据集

采用的是收集到的各种农田发生火灾的图片，然后进行分类存储，转化为数据集进行具体的训练集、验证集和测试集的划分，对图像

进行增广, 并采用Label软件图像标注工具对火源图片进行标注, 设置标签为: 烟雾(smoke)和火源(fire)。

4.2 测试结果

首先使用本系统对近 1000 张农田火源照片数据集进行测试, 如图 5 所示为混淆矩阵测试结果图像, 可以看出模型对烟雾和火焰的混淆矩阵fire的正例率为94%, smoke的正例率为85%, 两者的正利率并不低, 说明本模型较好, 系统是可靠性的。

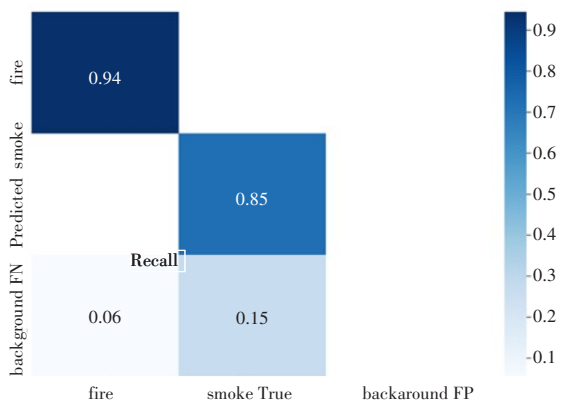


图 5 混淆矩阵图像

运用本系统进行测试的召回率图像如图 6 所示, 其横轴为召回率, 纵轴为精确率。从图 6 可以看出, 在召回率小于 0.7 时精确率较稳定; 之后随着召回率的不断增大, 精确率整体上呈现出下降趋势, 系统的最佳效果是在召回率小于 0.7 时, 其性能较好。

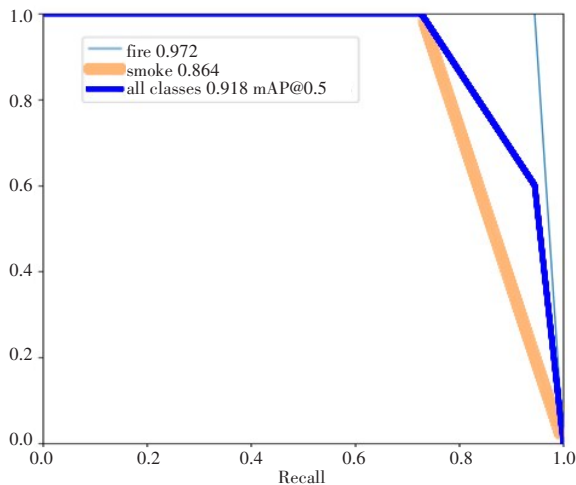


图 6 召回率图像

系统测试的置信度图像如图 7 所示, 可以看出当置信度小于 0.3 时, 精确率稳定, 系统的效果最好, 稳定性最高, 之后随着置信度的不断增大, 精确率迅速下降。

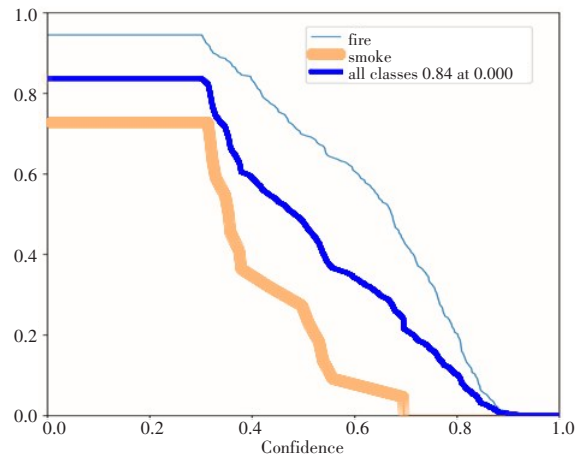


图 7 置信度图像

基于上述三次测试, 可以得出本系统的稳定性较好, 可靠性较高, 但系统模型训练还存在一些误差, 因影响因素过多, 训练效果并不理想, 最终的火源测试检测效果如图 8 所示, 图 9 为可视化大屏实时火源检测结果界面。



图 8 火源检测效果图

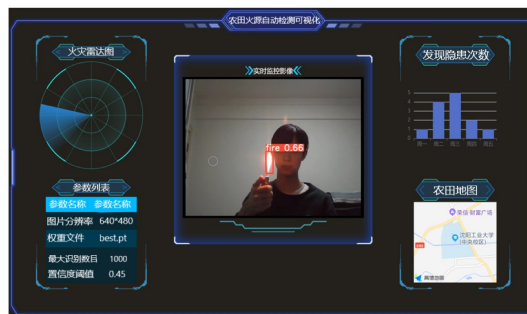


图 9 可视化大屏实时火源检测结果界面

5 结语

针对当前农田火灾频发,不能确保火灾检测实时性等问题,为避免造成较大威胁,本文提出了农田火源自动检测系统。该系统能快速从视频中获取火源图像并报警,有较短的推理时间和训练时间,大大缩短了火源产生和人工对火源产生措施的间隔时间,与深度学习相结合,通过强大而快速的计算,代替了人工对火源的排查,很好地预防了火灾的发生,尽可能地火灾发生后的人身和财产进行保护,在各模块的配合下,最终实现了农田火源自动检测功能研发。

参考文献:

- [1] WU S X, ZHANG L B. Using popular object detection methods for real time forest fire detection[C]// 2018 11th International Symposium on Computational Intelligence and Design (ISCID). Hangzhou, China:IEEE,2018.
- [2] 詹琪. 基于视频图像的森林火灾自动监测识别技术研究[D]. 成都:电子科技大学,2017.
- [3] AXELRAD G, GERSHFELD T, FEINERMAN E. Reclamation of sewage sludge for use in Israeli agriculture: economic, environmental and organizational aspects[J]. Journal of Environmental Planning and Management,2013,56(10):1419-1448.

- [4] 王朝泽. 基于物联网的农田智能灌溉系统研究和设计[D]. 天津:天津理工大学,2019.
- [5] 张锦,屈佩琪,孙程,等. 基于改进YOLOv5的安全帽佩戴检测算法[J]. 计算机应用,2022,42(4):1292-1300.
- [6] 王静,孙紫雲,郭革,等. 改进YOLOv5的白细胞检测算法[D]. 计算机工程与应用,2022,58(4):134-142.

作者简介:

白玉(2001—),男,辽宁铁岭人,本科生,研究方向为信息与计算科学

马广焜(1977—),女,辽宁沈阳人,博士,讲师,研究方向为计算机图形学和图像处理

彭新茗(1989—),女,辽宁沈阳人,硕士,工程师,研究方向为模式识别

王楠翔(2000—),男,辽宁辽阳人,本科生,研究方向为信息与计算科学

白清扬(2000—),男,辽宁沈阳人,本科生,研究方向为信息与计算科学

孟峻锋(2001—),男,辽宁抚顺人,本科生,研究方向为:信息与计算科学

刘鑫蕊(2001—),女,辽宁朝阳人,本科生,研究方向为信息与计算科学

朱硕闻(2001—),女,辽宁鞍山人,本科生,研究方向为信息与计算科学

收稿日期:2022-06-02 修稿日期:2022-07-16

Automatic Detection System of Farmland Fire Source

Bai Yu, Ma Guangkun, Peng Xinming, Wang Nanxiang, Bai Qingyang, Meng Junfeng,
Liu Xinrui, Zhu Shuowen

(College of Software, Shenyang University of Technology, Shenyang 110870)

Abstract: Farmland safety is a hot topic of general concern in the society, and farmland fire is also an important factor in the hidden dangers of farmland safety. In order to reduce the harm of fire source in farmland, an automatic fire source detection system for intelligent fire source detection is proposed. Although there are already temperature and smoke sensors for fire detection, the real-time monitoring cannot be guaranteed. To solve this problem, a farmland fire source automatic detection system based on Yolo V5 framework is proposed. Through the small target detection technology, real-time detection of potential fire sources in farmland, saving the time of manual fire source investigation. The latest yolov5 algorithm is adopted, which has shorter reasoning time and training time, and greatly shortens the interval between fire source generation and manual fire source generation measures. In addition, the system provides a variety of interfaces, which can be combined with the intelligent farmland system to better protect the farmland. The data can be integrated into the cloud, and the cloud interface can be called through any device, realizing the diversity of the system and wide application space.

Keywords: deep learning; small target detection; YOLOv5; fire detection

文章编号: 1007-1423(2022)19-0039-06

DOI: 10.3969/j.issn.1007-1423.2022.19.006

角钢钢印字符识别方法的研究综述

何 迅, 刘小芳

(四川轻化工大学计算机科学与工程学院, 自贡 643000)

摘要: 随着互联网在世界范围内的快速发展, 我国已经走上了快速发展的轨道, 工业互联网已经成为其中的重要组成部分。工业智能化已成为未来发展的必然趋势, 而角钢作为国家电网建设中的重要材料之一, 通过深度学习使用 OCR 技术对处于复杂背景下的角钢钢印字符进行检测与识别, 提取型号信息并进行存储, 为角钢信息智能化管理提供便利。通过分析研究背景及意义、分析研究现状, 为角钢钢印字符识别提供方法和依据。

关键词: 工业智能; 深度学习; OCR; 复杂背景

0 引言

随着信息技术的飞速发展和现代工业生产技术的不断进步, 人类已经进入了一个全新的工业时代。目前国家电网也在大力推广工业智能化, 而如何将智能化引入传统制造业是当今乃至今后一段时间内的重要课题。角钢是国家电网建设的重要材料之一, 角钢钢印字符便是角钢的身份标识, 也是人们用来识别角钢型号的依据。一个完整的高架铁网需要不同型号的角钢组成, 而不同型号的角钢所需的数量也不同, 人工对成堆角钢钢印字符进行识别记录时很容易造成失误从而降低效率, 人工造成的失误会耗费大量的人力物力成本去重新匹配正确的型号组合, 在如今信息化的时代这样的方式极为低效。受益于图像识别技术的高速发展, 字符识别已经应用于社会的许多领域, 角钢钢印字符识别便是将智能化带入传统制造业的重要途径之一, 为日后解决工业角钢仓储管理信息化、出入库自动化等场景问题提供重要的技术思路。

1 研究背景及意义

在图像处理领域中字符识别是一个重要的研究方向。字符识别的方法是对图像中字符信息进行定位、识别和提取, 然后输出结果让计

算机和人们理解。现如今字符识别技术已经应用于人们生活中的多个方面, 如快递单号^[1]、证件信息^[2]、纸质文档^[3]、打印发票^[4]、汽车车牌^[5]等。而工业环境下的角钢钢印字符与此不同, 角钢有更为复杂多变的字符背景。角钢字符是一种钢印字符, 相比于打印字符特征更加难以泛化, 并且字符表面存在锈迹、磨损和字迹不清等干扰; 字符图像存在相机设备与工厂加工设备工作振动导致字符抖动的影响。面对复杂的工业环境和角钢字符表面的干扰, 人工设计特征定位提取字符再识别的传统算法很难应对。

总的来说, 角钢字符背景环境复杂、干扰因素多、人工检测效率不高且无法满足工业需求, 传统算法角钢字符特征提取困难。所以, 使用如今高速发展的深度学习可以很好地解决这个问题。相机定点采集角钢字符数据集, 利用深度学习训练检测与识别模型, 开发角钢钢印字符识别系统, 从而提高角钢字符的识别, 促进工业智能化发展。

2 研究现状及存在的问题

光学字符识别 (Optical Character Recognition, OCR) 是一门科学, 能够将各种类型的文档或图像转换为可分析、可编辑和可搜索的数据^[6]。OCR 的设想最早起源于 20 世纪 20 年代末,

由奥地利的工程师Tauschek^[7]提出,并获得了技术专利,这是OCR概念首次出现在历史上。纵观OCR迅速发展的几十年,在深度学习流行于识别领域前,OCR多使用传统算法。

传统的字符识别算法是定位和分割单个字符,然后利用人工设计的特征训练分类器。分割算法决定了最终识别结果的质量,常见的分割算法一般基于连通域和垂直投影^[8-10]。图像二值化后,分割算法通过连通域或灰度直方图分割字符。背景简单、风格规整的字符对此有不错的效果,但对于像角钢字符一样背景复杂的字符很难达到好的效果。于是有的研究者开始探索另一条道路,如Girshick等^[11]提出R-CNN模型,便标志着深度学习^[12]与OCR技术结合道路的开始。随后Hinton课题组使用深度学习创建了AlexNet^[13],在ImageNet图像识别比赛上一举夺冠,并且碾压第二名采用SVM方法的分类性能,这一成果吸引了众多学者开始关注深度学习,此后深度学习的发展进入爆发性增长,更多的研究人员将OCR技术与深度学习结合在一起,一些具有优秀性能的网络结构被提出^[14-16],基于深度学习的各种文本检测和文本识别模型也逐渐涌现。

文本检测模型方面,Tian等^[17]提出CTPN模型可以检测水平和轻微倾斜的文本,水平长文本的检测效果也更好,CTPN模型是在Faster-RCNN的基础上移除ROI层,引入anchor机制,并加入RNN网络对检测目标的位置偏移量和置信度分数的计算;CTPN模型的检测效果很好,但是由于加入了RNN导致网络速度会比较慢。Zhou等^[18]提出EAST模型,采用端到端的训练方法,直接将OCR常用的多阶段简化为两个阶段,大大提高了网络的执行速度。此外,由于该模型设置的检测框可以是任意四边形,致使EAST能够检测倾斜文本,但扭曲文本EAST仍然存在不足。之后Liao等^[19]提出DBNet模型,解决了EAST模型不能识别扭曲文本的问题,该模型在基于分割的文本检测中,不同于之前需要指定二值化阈值,该模型可以自适应地对每一个像素点进行二值化,并通过从网络学习得到二值化阈值,直接将二值化步骤添加到网络中一起训练,大大提高了输出图对于二值化阈值的鲁

棒性,从而达到高性能和高准确度。近年来扭曲文本的检测仍在提高,Zhu等^[20]提出傅立叶轮廓嵌入(Fourier Contour Embedding)方法,无论什么形状的文本轮廓都可表示为Compact Signatures,FCENet先预测文本中的Compact Fourier Signatures,接着就在测试的过程中通过傅里叶变换和非最大抑制重新创建文本轮廓,以此解决高度弯曲的文本检测。

文本识别模型方面,由Shi等^[21]提出CRNN模型,结合了CNN与RNN,利用BLSTM和CTC^[22]组件学习字符图像中的上下文关系,可以有效提高文本识别的准确性,还解决了基于图像的序列识别问题,特别是很好地处理了在实际场景中的文本识别应用问题,能够识别较长的文本。可CRNN的缺点是对有较大形变的文本识别准确率较低,针对较大形变这个问题,Shi等^[23]又提出RARE模型,RARE支持一种称为TPS(Thin-Plate Splines)的空间变换,可以更准确地识别透视变换文本,以及弯曲文本。形变文字的问题得以解决后,发现文字识别时排列文字不规则也会影响模型效果,对此针对不规则文字,Shi等^[24]继续对RARE进行改进,提出ASTER模型,实行先把文本矫正成正常线性排列的文字,然后再进行识别,并把矫正网络和识别网络整合起来作为一个端对端网络来训练,进一步提高识别效率。但这种方法需要维护复杂的管道,不适合实时应用,为了解决这个问题,Liu等^[25]提出自适应贝塞尔曲线网络(Adaptive Bezier-Curve Network),通过将贝塞尔曲线参数化来自适应地处理任意形状的文本,显著提高了速度,并自主设计了新的Bezi-align层,在提取任意形状文本的卷积特征上,相比之前的方法精度更高,在实时识别应用上也具有竞争性的识别精度。

除了大多数研究者努力研发文本检测和文本识别的模型,也有研究者对其他方面进行开拓,Nguyen等^[26]提出一种新的语言感知方法,在训练和推理阶段运用该方法利用字典可以解决许多条件下的视觉模糊性问题。Hou等^[27]提出新型注意力机制(Coordinate Attention, CA)模块,通过嵌入位置信息到通道注意力,从而使轻量级网络获取更大区域的信息而避免引入大

的开销。为了避免2D全局池化引入位置信息损失,将通道注意力分解为两个并行的1D特征编码来高效地整合空间坐标信息到生成的Attention Maps。Squeeze-and-Excitation Networks(SE)模块^[28]、Convolutional Block Attention Module(CBAM)模块^[29]与新型CA注意力机制模块结构对比示意图如图1所示。

图1中三个模块都是被嵌入Residual结构中的例子,在结构上SE模块分为两个步骤,压缩(Squeeze)和激励(Excitation),分别设计用于全局信息嵌入和通道关系的自适应重新校准,具体结构如图1(a)所示,使用Global Avg Pool(GAP)作为Squeeze操作,紧接着一个Fully Connected(FC)层将特征维度下降 r 倍,再经过Non-linear函数激活后再通过一个FC层升回到原来的特征维度,然后通过Sigmoid函数转化为一个0~1的归一化权重,这一段为Excitation操作。最后便将新生成的特征图和原来的特征图进行相乘校准,即Re-weight操作;CBAM模块除了对通道域进行加权操作,还增添了空间域的处理,

具体结构如图1(b)所示,首先对通道域方面进行操作,使用全局平均池化(GAP)和全局最大池化(GMP)来提取高层次特征,之后与SE模块过程类似,只是将FC层换成了1维卷积,激活函数换成了ReLU,接下来便是空间域上的操作,基于Channel再次进行GAP和GMP处理,由于进行了两次池化,拼接后通道数变成了2,所以使用一个7维卷积核进行卷积,将通道数压缩成了1,然后进行BatchNorm(BN)和Sigmoid操作,最后完成Re-weight操作;CA模块通过空间域对通道关系和远程依赖进行编码,具体结构如图1(c)所示,首先沿 X 方向和 Y 方向分别进行平均池化(AvgPool)操作,得到两个分别基于 X 和 Y 的特征图,然后进行拼接(Concat)和卷积操作,建立起远程依赖关系,接着进行BN操作和Non-linear函数激活后,再将有着全局信息的该特征图重新分为基于 X 和 Y 的特征图,分别继续卷积和Sigmoid归一化操作,最后将沿 X 和 Y 的信息加权于原特征图。图1中“ C ”表示通道,“ H ”和“ W ”分别表示高和宽。

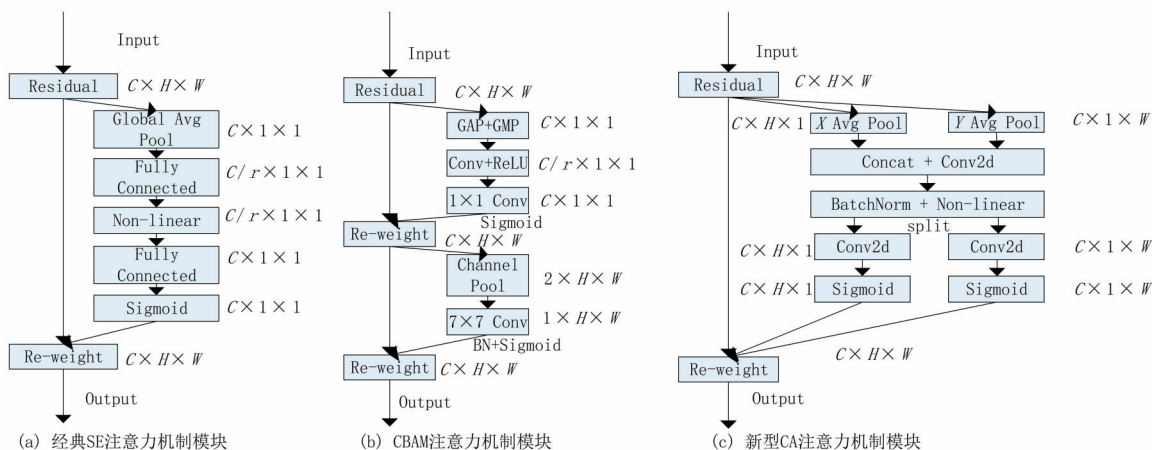


图1 各注意力机制结构对比示意图

目前与深度学习结合的OCR技术已经被广泛应用于金融、交通、物流等多个领域,但是就工业领域而言,该OCR技术仍然没有得到充分应用,所以近年来针对于工业领域的字符识别开始大力推进。工业上的图像数据和公共数据集上的图像数据特征差异过大,使用公共数据集训练的通用型模型对于工业字符来说效果并不好,所以需要实验者自己收集实验对象的

数据集。因此,工业OCR便存在两个问题,如图2所示。

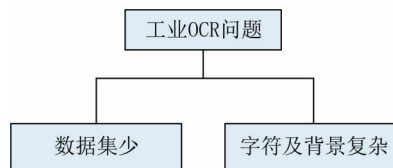


图2 工业OCR问题示意图

在这两方面的问题上有研究者也已经做出研究:

(1) 数据集少的问题, 实验者个人收集的实验对象数据一般来说是没有完善的公共数据集, 通常收集的数据图像不全面且数量不多, 这样的数据集, 其多样性就不足。在不增加实拍数据集的情况下, 林惠^[30]提出一种基于条件对抗网络的数据增强算法, 通过以字符类别标签作为约束条件, 可以实现字符识别数据集扩充, 解决芯片字符图像数据集不足、多样性差的问题。而陈妃奋^[31]采用YOLOv4算法, 通过迁移学习的训练方式也解决了工业样本量较少而引起的过拟合问题。两者的思路都是解决工业上数据集少的可参考方法。

(2) 在字符及其背景问题上, 与普通文本检测相比, 工业字符检测的难点最主要就是工业字符类型多样、字符特征不同、排版也不同, 并且工业环境下的图像字符背景也更复杂多样, 所以市面上的通用模型很难实现字符定位与分割。吴慧英^[32]提出以自然场景文本识别网络CRNN为基础加入CTPN网络, 同时引入注意力机制, 不仅能在复杂的图像中检测出指定的文本区域, 还能提高在图像背景复杂及噪声干扰下的稳定能力, 较好地复杂背景下完成字符识别。李建明^[33]提出改进的PSENet和CRNN算法进行字符检测, 检测算法方面采用ShuffleNetv2-FPN替换PSENet的骨干网络完成字符定位; 识别算法方面在CRNN的骨干网络ResNet18中引入多头自注意力机制模块, 且结合Center Loss和CTC Loss损失函数来完成字符识别。将改进的定位算法和识别算法进行串联封装, 很好地提高了工业环境下的字符检测能力。

3 结语

总结目前工业OCR的现状和问题, 人工检测与记录工业工件上的字符效率较低, 通过OCR技术可以极大地提高识别效率, 进而实现字符识别流程自动化。现在公开的ICDAR、ICPR等比赛数据集都是自然场景文本, 而工业字符公共数据集却鲜有公开, 且字符背景又复

杂, 导致通用模型效果都比较差, 结合一些研究者的成果, 总结出可以从以下两个方面改进角钢印字符识别性能:

(1) 最简单的便是从数据集方面入手, 在工业实地多角度、多型号地尽可能收集足量的图像, 从而训练出泛化性和鲁棒性不错的模型。

(2) 灵活地从网络结构方面进行改进, 对于不同类型的工业字符产品适当地修改网络结构。就提取特征而言, 如只有英文和数字, 可以使用轻量级网络结构, 这样既能减少训练所消耗的成本, 也能加快检测速度; 反之, 有中文或较多字符类型时, 可以选择较深、较复杂的网络结构。

当然这两个方面也可同时进行, 对于(1)很难有条件得到较多现场数据时, 也可以使用数据增强的方法辅助处理, 再基于此对(2)进行针对性的改进, 来解决识别角钢印字符的性能问题。在基础工业字符数据集的不断公开和技术的不断进步下, 如果能够得到一个具有泛化性和鲁棒性的工业字符通用模型, 将具有十足的现实意义, 这种意义会随着工业智能化的发展继续产生影响, 并将在其他领域引发类似的启发。

参考文献:

- [1] 张震. 基于深度学习的快递表单信息处理及应用[D]. 南京: 南京邮电大学, 2020.
- [2] 段金宝. 基于深度神经网络的证件图像文本识别方法[D]. 北京: 北京邮电大学, 2018.
- [3] 唐董琦. 文档图像中端到端文字识别及其应用研究[D]. 南京: 南京大学, 2020.
- [4] 谢阳. 基于深度学习的发票识别研究[D]. 南京: 南京邮电大学, 2020.
- [5] MONTAZZOLLI S, JUNG C R. Real-time Brazilian license plate detection and recognition using deep convolutional neural networks [C] // SIBGRAPI-Conference on Graphics, Patterns and Images. IEEE, 2017: 55-62.
- [6] MEMON J, SAMI M, KHAN R A, et al. Handwritten optical character recognition (OCR): a comprehensive systematic literature review (SLR) [J]. IEEE Access, 2020, 8: 142642-142668.
- [7] PLAMONDON R, SRIHARI S N. On-line and off-line handwriting recognition: a comprehensive

- survey[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2000, 22(1):63-84.
- [8] INUNGANBI S, CHOUDHARY P, SINGH K M. Local texture descriptors and projection histogram based handwritten Meitei Mayek character recognition[J]. Multimedia Tools and Applications, 2020, 79(3):2813-2836.
- [9] GRZEGORZEK M, LI C, RASKATOW J, et al. Texture-based text detection in digital images with wavelet features and support vector machines[C]// Proceedings of the 8th International Conference on Computer Recognition Systems CORES 2013. Switzerland: Springer, 2013:857-866.
- [10] TIAN S, BHATTACHARYA U, LU S, et al. Multilingual scene character recognition with co-occurrence of histogram of oriented gradients[J]. Pattern Recognition, 2016, 51(3):125-134.
- [11] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6):1137-1149.
- [12] HAO X, ZHANG G, MA S. Deep learning[J]. International Journal of Semantic Computing, 2016, 10(3):417-439.
- [13] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet classification with deep convolutional neural networks[J]. Advances in Neural Information Processing Systems, 2017, 60(6):84-90.
- [14] 吴帅,徐勇,赵东宁. 基于深度卷积网络的目标检测综述[J]. 模式识别与人工智能, 2018, 31(4):45-56.
- [15] HUANG G, LIU Z, DER MAATEN L V, et al. Densely connected convolutional networks [C] // IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017: 2261-2269.
- [16] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8):2011-2023.
- [17] TIAN Z, HUANG W, TONG H, et al. Detecting text in natural image with connectionist text proposal network [C] // European Conference on Computer Vision, Cham, Switzerland, 2016:56-72.
- [18] ZHOU X, YAO C, WEN H, et al. EAST: an efficient and accurate scene text detector [C] // IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 2017:2642-2651.
- [19] LIAO M, WAN Z, YAO C, et al. Real-time scene text detection with differentiable binarization[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020; 34(7):11474-11481.
- [20] ZHU Y, CHEN J, LIANG L, et al. Fourier contour embedding for arbitrary-shaped text detection [C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 2021:3122-3130.
- [21] SHI B, XIANG B, CONG Y. An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016, 39(11):2298-2304.
- [22] GRAVES A, SANTIAGO F, GOMEZ F. Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks [C] // 23rd International Conference on Machine Learning, Pittsburgh, Pennsylvania, USA, 2006:369-376.
- [23] SHI B, WANG X, LYU P, et al. Robust scene text recognition with automatic rectification [C] // IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 2016:4168-4176.
- [24] SHI B, YANG M, WANG X, et al. ASTER: an attentional scene text recognizer with flexible rectification [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41(9):2035-2048.
- [25] LIU Y, CHEN H, SHEN C, et al. ABCNet: real-time scene text spotting with adaptive bezier-curve network [C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 2020:9806-9815.
- [26] NGUYEN N, NGUYEN T, TRAN V, et al. Dictionary-guided scene text recognition [C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 2021:7379-7388.
- [27] HOU Q, ZHOU D AND FENG J. Coordinate attention for efficient mobile network design [C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 2021:13708-13717.
- [28] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8):2011-2023.
- [29] ZHAO Z, CHEN K, YAMANE K. CBAM-Unet++: easier to find the target with the attention

module "CBAM" [C] //IEEE 10th Global Conference on Consumer Electronics (GCCE), Kyoto, Japan, 2021:655-657.

- [30] 林惠. 基于深度学习的芯片字符实时识别方法研究[D]. 武汉:华中科技大学, 2019.
- [31] 陈妃奋. 基于深度学习的字轮式水表读数识别研究与应用[D]. 桂林:桂林电子科技大学, 2021.
- [32] 吴慧莹. 复杂背景下的工业字符分割与识别研究[D]. 桂林:广西师范大学, 2020.
- [33] 李建明. 基于卷积神经网络的工业字符检测方法研究[D]. 无锡:江南大学, 2021.

作者简介:

何迅(1998—),男,四川绵阳人,硕士研究生,研究方向为人工智能

通信作者:刘小芳(1969—),女,重庆垫江人,博士,教授,硕士研究生导师,从事智能信息处理、模式识别、数据挖掘等方面的教学及科研工作, E-mail: 150507076@qq.com

收稿日期: 2022-05-19 修稿日期: 2022-07-24

A Review of Research Methods for Character Recognition of Angle Steel Stamps

He Xun, Liu Xiaofang

(School of Computer Science and Engineering, Sichuan University of Science & Engineering, Zigong 643000)

Abstract: With the rapid development of the Internet in the world, China has embarked on the track of rapid development, and the industrial Internet has become an important part of it. Industrial intelligence has become an inevitable trend of future development, and angle steel, as one of the important materials in the construction of the State Grid, uses OCR technology to detect and identify angle steel stamp characters in a complex background through deep learning, extract model information and store it, providing convenience for intelligent management of angle steel information. Through the analysis of the research background and significance, and the analysis of the research status, it provides methods and basis for the identification of angle steel stamp characters.

Keywords: industrial intelligence; deep learning; OCR; complex background

(上接第 18 页)

Prediction and Analysis of the Number of Pension Service Beds in China

Tang Pengxiang, Zhang Xiaomei

(College of Pharmacy, Jilin Medical University, Jilin 132013)

Abstract: With the acceleration of China's population aging process, the prediction of the number of beds for pension services and the number of elderly population over 65 years old has become an urgent problem to be solved. According to the data provided by the China Statistical Yearbook and the annual bulletin of the Civil Affairs Bureau, a logistic model and an ARMA(1, 2, 1) model were established for the number of elderly care beds. In order to study the relationship between the number of elderly care beds and the proportion of the elderly population, a logistic model was established, and an exponential model was established for the elderly population over 65. MATLAB and Eviews were used to solve the above model, and the relevant prediction results were given. The prediction results of different models were compared and analyzed, which will provide a theoretical basis for the formulation and choice of China's pension policy in the future.

Keyword: logistic model; ARMA model; exponential model; prediction

文章编号: 1007-1423(2022)19-0045-06

DOI: 10.3969/j.issn.1007-1423.2022.19.007

基于 SIFT 算法的人群密集检测应用研究

郭 妍, 谢光敏, 向蓓蓓, 古荣龙, 王 磊

(四川水利职业技术学院, 成都 610039)

摘要: 为了检测视场开阔下的人群密集情况, 采用基于图像局部特征点提取的 SIFT 算法。首先, 对人群图像预处理, 其中包括灰度化、直方图均衡化、二值化和形态学处理等。其次采用背景差分法进行视频图像的前景提取。再次, 采取 SIFT 算法对图像中的人群部分进行特征点的检测, 采取高斯核对待检测的人群图像进行尺度变换, 获得图像的尺度空间表示序列。之后对极值点检测和精确定位, 然后指定其方向参数, 生成关键点描述子。最后, 统计人群目标生成的关键点的数目。根据关键点数目判断人群密集程度。多组实验数据表明, 该方案能较好地检测出视频图像中的人群密集情况。

关键词: 尺度变换; SIFT; 关键点数目

基金项目: 四川水利职业技术学院校级科研项目(KY2022-07)

0 引言

随着经济持续稳定的发展, 生活水平的日益提高, 城市公共场所作为经济文化的主要载体承担着各种各样的活动, 而随着人口密度的增大以及人们对生活质量追求的提高, 每一次的活动都会伴随着人群的聚集。近年来, 人口聚集的场所突发事件时有发生, 每一次的突发事件不仅造成了人们的恐慌, 也伴随着经济的损失和一系列的安全问题。其实大多时候人群聚集而引发事故发生进而导致重大的后果并不是事故自身的问题, 而是因为人群过于密集, 事发时慌乱拥挤造成的。从目前社会的整体情况来看, 我国的社会形势已经基本保持在一个稳定的状态。但在经济蓬勃发展的今天, 由于人们对物质文化需求的增多, 各种人群聚集的活动也逐渐增加。而公共场所作为承担着各种活动的重要载体, 突发事件发生的可能性也比较大。

2014 年的 12 月 31 日晚, 这本来是一个令人激动的跨年之夜, 很多游客市民聚集在上海外滩准备迎接新年的到来, 但因有人失衡摔倒接

连诱发多人摔倒、叠压, 致使拥挤踩踏事件发生, 最终导致 36 人遇难, 49 人负伤^[1]。因此, 为了防止类似事件的发生, 降低其概率, 对人群密集的检测是非常有必要的。

1 SIFT 算法简介

SIFT 算法, 即尺度不变特征变换算法, 它可以用于检测与描述影像局部特征, 找到空间尺度上的特征点, 并提取其位置、尺度、旋转不变量。辨识物体可以通过描述与检测局部影像特征起到作用^[2]。SIFT 特征与图像的尺寸和旋转无关, 它是基于物体上的一些局部外观兴趣点的, 对光线、噪声以及些微视角变化也有较高的容忍度。它们因为以上特征而呈现高度显著的特点并相对容易获得, 而且能够较容易且很少错误地识别基数庞大的特征数据库中的对象。

在不同的尺度空间上找到特征点, 并获得其方向是 SIFT 算法的实质。它通过高斯模糊来达到尺度空间的获取。因此, 特征点的提取筛选资源从高斯金字塔产生的差分高斯空间中的局部特征点获得。

2 方案介绍

2.1 提取人群特征点

提取人群特征点，需要先提取出能较好反映出人物事件的特征，可以成为影响后期人群密集检测的重要因素。即待检测图像读入内存并进行图像预处理之后，进行背景差分，提取图像的人群特征，对提取后的图像采用SIFT算法建立差分金字塔图像^[3]，进行极值检测。最后通过关键点精确定位生成人群特征点，根据视频图像中人群的特征点数目判断人群是否密集。算法流程如图1所示。

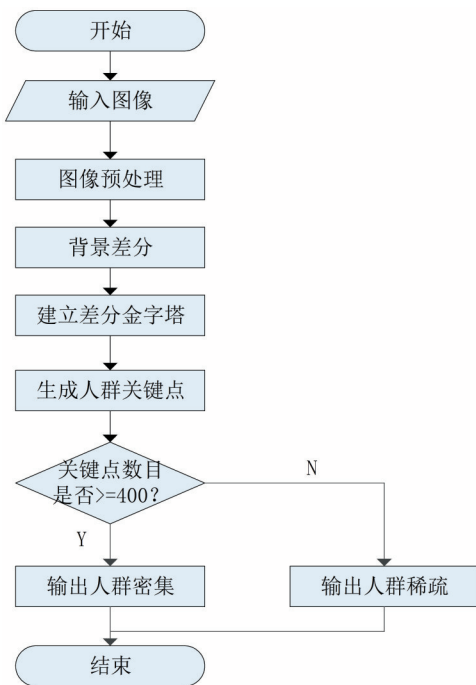


图1 算法流程图

2.2 图像预处理

通过图像预处理去掉原始图像所夹杂的一些不相干信息，可以提高我们所需的有用信息的可检测性且尽可能精简数据，以此达到特征提取的目的^[3]。目前常见的图像预处理技术包括彩色图像的灰度化、直方图均衡化、图像的二值化处理和形态学处理等。

其中图像灰度化处理^[3]是将彩色图像转化为灰度图像。彩色图像中的每个像素的颜色

是由红(R)、绿(G)、蓝(B)三个分量决定，而每个分量有255个值(0~255)，经计算，一个像素点约有1600多万(256*256*256)色。而灰度图像是红(R)、绿(G)、蓝(B)特殊的彩色图像，一个像素点的取值范围为0~255。因此，数字图像处理中，通常将图像进行灰度化处理以减少计算量。图像的灰度化处理过程可用下列公式表示：

$$Gray = (R*0.229 + G*0.587 + B*0.114) \quad (1)$$

其中 W_R 、 W_G 、 W_B 分别为 R、G、B 的权值。研究表明，人眼对颜色的敏感度由高到低依次为绿、红、蓝。因此若想取得较好的结果，取 $W_G > W_R > W_B$ 时的权值较合理。

2.3 背景差分

背景差分法^[4]又称背景减法，它的原理是将当前帧与背景图像进行差分从而得到运动目标区域。背景差分法将视频帧图像视为两个部分，即前景和背景。它最初需要选取一幅图像当作背景，而且这幅图像中不能含有运动前景。这种方法能够更好地识别和提取目标。由于摄像头和背景的固定性，采用背景差分法可以简单、快速地提取出前景图像，即用每一帧图像减去背景图像，便可得到我们所需的前景图像。它可以在背景平稳时达到良好的检测效果，并且视频中静止的目标也可以被检测到。

在不考虑噪声 $n(x, y, t)$ 的影响下，可以将运动目标 $m(x, y, t)$ 与背景图像 $b(x, y, t)$ 组成的图像称为视频帧图像 $I(x, y, t)$ ：

$$I(x, y, t) = b(x, y, t) + m(x, y, t) \quad (2)$$

由式(2)并不能够得到真正的运动目标，而是由运动目标区域与噪声所组成的差分图像 $d(x, y, t)$ ，即：

$$d(x, y, t) = I(x, y, t) - b(x, y, t) + n(x, y, t) \quad (3)$$

对其进行进一步处理后，可得到运动目标：

$$m(x, y, t) = \begin{cases} I(x, y, t) & d(x, y, t) \geq T \\ 0 & d(x, y, t) < T \end{cases} \quad (4)$$

其中， T 代表阈值。样本图像和经背景差分法处理后的图像分别如图2和图3所示。



图2 样本图像



图3 背景差法处理后的图像

2.4 建立差分金字塔

SIFT为获得连续变化的尺度,采用高斯金字塔来构建尺度空间。尺度空间上的极值点可以在计算获得金字塔后被检测^[5]。尺度规范的算子(Laplacion of Gaussian)被 Tony Lindeberg 说明具备真正的尺度不变性。高斯差分金字塔类似算子被 Lowe 用来对稳固的特征点在尺度空间中进行检测。对原始图像进行不停降阶采样后会获得相应的有差别的图像,将其由大到小、由下至上排列组成塔状模型,这就是图像的金字塔模型^[6]。通常情况下,金字塔的第一层是原始图像,金字塔层数为,每层(一张图像对应一层)是每次降采样后获得的新图像。图像的最初尺寸以及塔顶图像的尺寸一起决定了金字塔的层数,层数的计算公式如下:

$$n = \log_2\{\min(M, N)\} - t, t \in [0, \log_2\{\min(M, N)\}] \quad (5)$$

其中,原始图像的尺寸用 M, N 表示, t 为塔顶图

像的最小维数的对数值。

本方案在现实中进行极值检测^[7]时,需先建立高斯差分金字塔,其通过用高斯金字塔每组中相邻的上下两层图像作差而产生。经过同一组内 DOG 各相邻两层图像之间的对比,由 DOG 空间的局部极值点组成的关键点可以获得初步检测。本文中 DOG 函数的极值点的检测要求每一个像素点与它全部邻近的点作对比,观察结果是高于还是低于它的图像域和尺度域的邻近点。将中间的检测点与具有和其相同尺度的8个邻近点以及与其相邻尺度中对应的 9×2 个点进行对比,以确保尺度空间和二维图像空间中的极值点都可以被检测到。

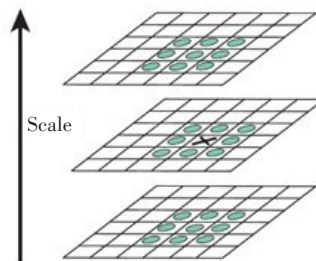


图4 DOG空间极值检测

2.5 生成人群关键点

本文结合特征点邻域像素的梯度方向分布特性,指定方向参数^[6]给每个特征点,以此让其在图像发生改变时不会发生变化。在对关键点进行运算后,邻域内像素的梯度和方向通过直方图来表示。该梯度直方图以每柱 10° 的区间将 $0 \sim 360^\circ$ 的方向范畴平均分成 36 个柱(bins)。如图5所示,关键点的主方向即直方图的峰值对应的方向。

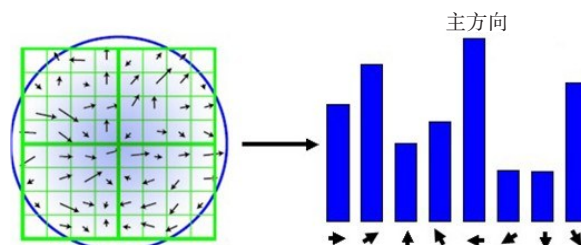


图5 关键点方向直方图

在获得关键点邻域的像素梯度分布特征信

息之后便可明确关键点主方向^[8]。一个特征点处邻域梯度的方向由方向直方图的峰值代表，该关键点的主方向即是图中最大值对应的方向。本文将该关键点的辅方向定义为高于主方向峰值80%的方向。为了向其提供稳定性，每个关键点都可以存在多个辅方向。本文在实际处理时，将其复制成多份。这些复制后的关键点将被赋予该点的多方向。而为了获得更加准确的方向角度值，要对离散的梯度直方图进行插值拟合处理^[9]。生成关键点后的检测效果如图6所示。



图6 检测效果图

2.6 输出检测结果

在生成关键点后，统计所生成的关键点数目*i*，将生成的关键点数目*i*与预先设定的中间值*c*做比较。如果*i*≥*c*，则待检测样本判断为密集人群，反之则为稀疏人群。

本文所采取的运行的待检测视频帧图像来源于PETS 2009视频数据库中的视频序列P2和P3，背景原图来自于该视频数据库中视频序列P1。这些相关图像序列均是分辨率为768×576的24位JPG图像。根据算法比对，相同场景下具有不同人群密集程度的视频序列图像对其进行相应的实验检测，检测结果如图7所示。

根据检测结果，分别对这两种情况下的人群图像进行讨论分析，检测结果如表1所示。

表1 对图7中不同人群图像的检测对比

	稀疏人群图像	密集人群图像
检测出的特征点数	221	720
图像预估的特征点数	231	810
置信度	95.67%	88.89%



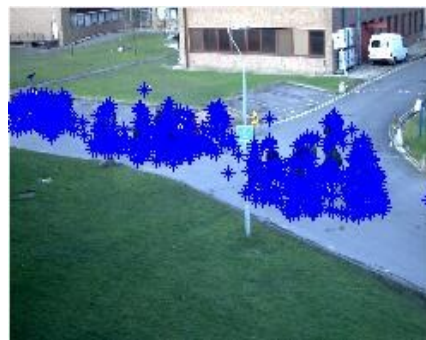
(a)稀疏人群原始图像



(b)生成特征点后的稀疏人群图像



(c)密集人群原始图像



(d)生成特征点后的密集人群图像

图7 相同场景下不同人群密集程度

针对不同场景下的不同人群图像,也分别对它们进行相应的SIFT提取特征点的检测,检

测结果如图8所示。



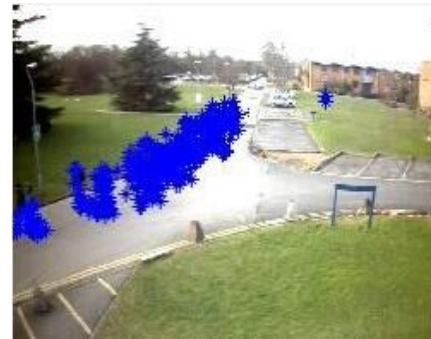
(a)稀疏人群原始图像



(b)生成特征点后的稀疏人群图像



(c)密集人群原始图像



(d)生成特征点后的密集人群图像

图8 不同场景下不同人群密集程度

根据检测结果,分别对这两种情况下的人群图像进行讨论分析,检测结果如表2所示。

表2 对图8中不同人群图像的检测对比

	稀疏人群图像	密集人群图像
检测出的特征点数	250	514
图像预估的特征点数	270	600
置信度	92.59%	85.57%

3 结语

随着城市和经济的全球化发展,人口密度的增加而带来的各种经济、安全问题已经成为当今社会所无法避免的现实问题。因此对人群是否密集的检测显得尤为重要。本文重点研究了SIFT提取特征算法。通过软件测试检验,

SIFT针对不同人群密集程度的图像进行特征点提取,都能够很好地将图像人群区域的特征点检测出来。

参考文献:

- [1] 陈耿. 防范公共场所人群拥挤踩踏事故的措施[J]. 城市与减灾, 2015(3): 42-43.
- [2] 屈有佳. 基于SIFT特征的关键帧提取算法研究[D]. 北京: 北京交通大学, 2015: 1-55.
- [3] 丁苏楠, 张秋菊. 基于改进SIFT算法的图像匹配方法[J]. 传感器与微系统, 2020, 39(10): 45-47, 50.
- [4] 乐英, 赵志成. 基于背景差分法的多运动目标检测与分割[J]. 中国工程机械学报, 2020, 18(4): 305-309. DOI: 10.15999/j.cnki.311926.2020.04.005.
- [5] 王祖光. 图像挖掘中的图像预处理技术浅析[J]. 科技风, 2011(19): 246. DOI: 10.19392/j.cnki.1671-7341.2011.19.206.

- [6] 于祥春. 一种基于SIFT的特征提取在人脸识别算法中的研究[D]. 长春: 吉林大学, 2015: 1-53.
- [7] LAPTEV I, LINDERBERG T. Local descriptors for spatio-temporal recognition [M]. Spatial Coherence for Visual Motion Analysis. Springer Berlin Heidelberg, 2006: 91-103.
- [8] 高晓玲, 王娟, 尹慧敏. 基于SIFT算法的图像匹配技术的研究[J]. 信息技术与信息化, 2020(12): 65-67.
- [9] 刘鹏南, 徐冬冬, 白春梦. 基于尺度不变特征变换的异源图像配准方法[J]. 激光与光电子学进展, 2021, 58(24): 172-181.

作者简介:

通信作者: 郭妍(1994—), 女, 吉林白山人, 硕士, 研究实习员, 研究方向为图像处理、模式识别, E-mail: 363934069@qq.com

谢光敏(1996—), 女, 四川资阳人, 硕士, 助教, 研究方向为大数据分析、图像篡改

向蓓蓓(1992—), 女, 重庆潼南人, 学士, 助教, 研究方向为电路与系统、信号与信息处理

古荣龙(1990—), 男, 四川成都人, 学士, 助教, 研究方向为软件开发及大数据应用

王磊(1995—), 男, 山东潍坊人, 学士, 助教, 研究方向为计算机应用技术、大数据分析

收稿日期: 2022-03-10 修稿日期: 2022-04-26

Research on Population Aggregation Features Detection Based on SIFT Algorithm

Guo Yan, Xie Guangmin, Xiang Beibei, Gu Ronglong, Wang Lei

(Sichuan Water Conservancy Vocational College, Chengdu 610039)

Abstract: In order to detect the dense crowd in a wide field of view, the SIFT algorithm based on the extraction of local feature points of the image is adopted. Firstly, the population image was pretreated, including gray-scale, histogram equalization, binarization and morphological processing, etc., and secondly, the background difference method is used to extract the foreground of the video image. Then, the SIFT algorithm is used to detect the feature points of the crowd in the image. The Gaussian kernel is used to scale the image of the population to be detected, and the scale space representation sequence of the image is obtained. Then the extreme point was detected and precise positioning, the specify is direction parameters, generate key descriptors. Finally, count the number of key points generated by the target population. According to the number of key points, the crowd is dense in Chengdu. Multiple sets of experimental data show that this scheme can better detect crowded situations in video images.

Keywords: scale transformation; SIFT; number of key

文章编号: 1007-1423(2022)19-0051-06

DOI: 10.3969/j.issn.1007-1423.2022.19.008

基于四足仿生机器人快递配送的研究

靳 晨, 刘 刚, 吕兴东, 黎航宇, 杨雅子琪, 邓艺霖

(北京信息科技大学仪器与光电工程学院, 北京 100192)

摘要: 简要概述了中国机器人大赛中四足仿生机器人的参赛机制, 详细介绍了四足仿生机器人的硬件平台结构以及快递配送项目的软件实现方案; 从实践结果来看, 仿生机器人具备过减速带、上台阶、过窄桥、下斜坡、过草地等地形障碍以及二维码识别、快递配送和快递投放任务的功能; 从参赛结果来看, 该机器人取得的优异成绩, 验证了本文硬件设计与控制策略的正确性。

关键词: 四足仿生机器人; 控制策略; 快递配送

基金项目: 2022 年大学生科技创新计划项目(202211232002)

0 引言

仿生机器人是机器人技术领域中的一个发展分支, 是当前机器人领域的研究热点。人们可以利用仿生机器人的灵活性和环境适应性从事复杂或危险的工作。近年来, 国内外足式仿生机器人的研发迅猛, 如 1992 年成立的美国波士顿动力(Boston Dynamics)公司在四足机器人方面的研发一直领跑世界^[1]; 国内对四足机器人的研究热度也在持续升温, 2007 年成立的杭州宇树(Unitree)科技有限公司, 在智能机器人领域的研发也日新月异。由于足式运动的特点, 此类型的机器人在崎岖不平的地面支撑、运动灵活, 环境适应能力强, 同时具有很强的平衡能力, 多用于物资运输、警备执勤、抢险救灾及娱乐表演等, 应用发展前景广阔。

1 四足仿生机器人快递配送比赛规则^[2]

如图 1 所示, 四足仿生机器人快递配送赛开始比赛时, 四足仿生机器人须放置在快递集散区域, 四脚落地并全部位于快递集散区框线内。由参赛队员按照配送快递序列顺序, 向参赛四足仿生机器人背筐中放置一个快递球, 并辅助四足机器人识别快递球颜色, 然后将快递球放到四足机器人背筐里。机器人从快递集散区出发进行配送, 沿途依次通过减速带、上台阶、

窄桥、下斜坡、草地等障碍地形, 并快速通过平整地面赛段, 完成快递配送任务, 并返回集散区。当机器人四足落脚点完全进入集散区域后, 参赛队员方可再次向背筐中放球, 机器人进行下一轮配送。每一轮仅配送一个快递球。共有 3 家住户, 快递共 9 个, 每家住户 3 个快递球, 快递球配送顺序在赛前打乱生成。比赛时间最长为 20 分钟。9 个快递球配送完毕、机器人返回集散区时比赛立即结束; 若 20 分钟时间到快递球未配送完毕, 比赛立即结束。

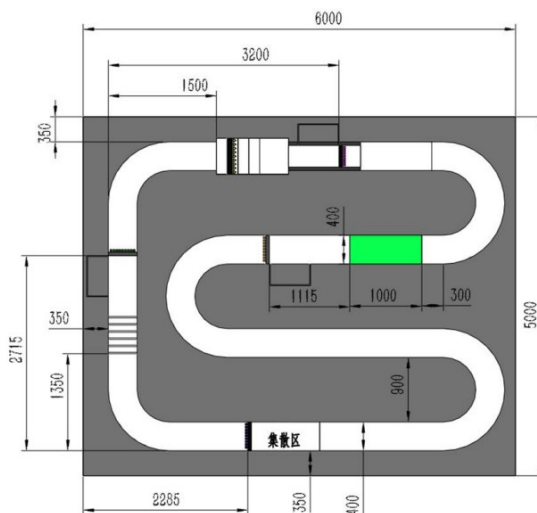


图 1 比赛场地整体尺寸(单位:mm)

比赛对四足仿生机器人的要求是四足仿生机器人重量不得超过 3.0 kg，四条腿直立时，其体积小于长 400 mm×宽 240 mm×高 260 mm，且腿长小于 200 mm，机器人正常行走时，躯干下表面离地高度不小于 100 mm(注：尺寸为去除背筐时的尺寸)。四足机器人为四足哺乳类动物仿生腿足结构，不得使用爬行类等动物仿生腿足结构。不得使用并联机构，各关节需通过连杆或支架串联形成腿。

图 1 为快递配送项目比赛的场地，大小为 6000 mm×5000 mm，在硬质平整地面搭建，表面为黑色。赛道宽度为 400 mm，由白色无纺布或者白色哑光喷绘纸铺设(视场地情况可直接喷

涂白色非反光漆)。赛道中的减速带、台阶、窄桥、斜坡为白色密度板材或者白色木板，草地为 5 mm 厚的人工草皮。场地周围 3 个住户区域，各住户的三面由长宽高为 570 mm×300 mm×150 mm 的挡板围成，一面开口，朝向赛道。挡板厚度为 10 mm，颜色为白色。

2 硬件平台设计^[3]

四足仿生机器人的硬件平台如图 2 所示，仿生机器人本体共有五大部件：舵机组、香蕉派开发板、运动控制板、CCD 摄像头及姿态传感器等。

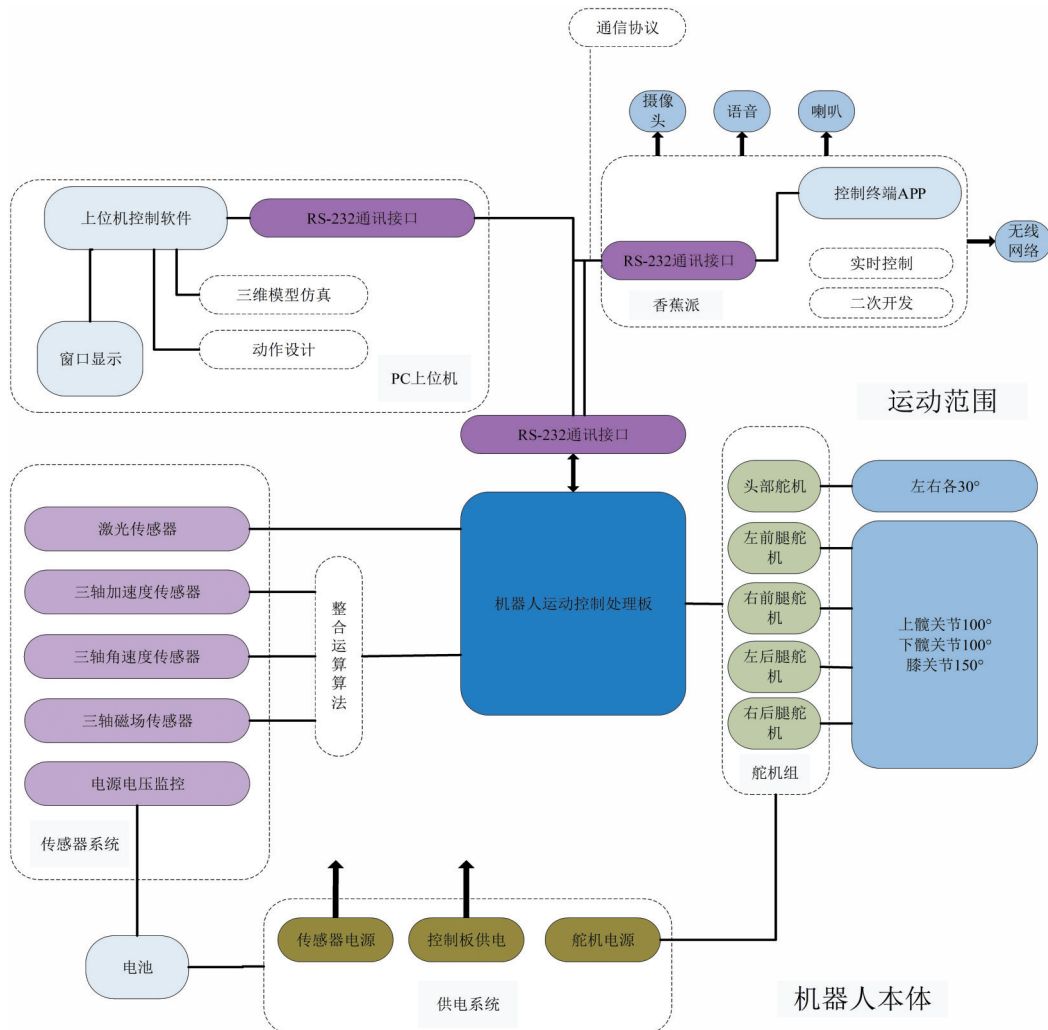


图 2 四足仿生机器人硬件系统平台框图

2.1 舵机组

头部设有一个水平方向的关节舵机,可以左右运动各 30° 角度(运动范围 60°);每条腿设有三个关节舵机,最上面为上髌关节舵机,运动范围 100° ;中间为下髌关节舵机,运动范围 180° ;最下面为膝关节舵机,运动范围 150° 。

四足仿生机器人的姿态是指机器人每条腿抬起、放下的规律,是确保步行机构能够稳定行进的非常关键的条件,姿态生成与控制的最终目标就是设计机器人运动时各条腿迈腿顺序以适应复杂环境^[4]。

如图3所示,用户可采用Action designer应用程序对四足仿生机器人的舵机进行调零设置。初次调试时,动作列表为空,用户可以点击添加按钮,添加动作,点击重命名来重命名动作。当动作列表中有一个动作存在时,需选定一个动作,然后点击添加按钮,这样就在被选动作前添加了一个动作。当选中一个动作,并点击设置页面的下载按钮时,就可以实时地下载动作到下位机(四足机器人)。



图3 Action Designer软件界面

接下来是姿势的添加,既是本机器人调试的难点又是其特色。在姿势列表中,用户可以通过鼠标右键拖动3D四足机器人,显示区域中的四足机器人四肢来设计每个姿势。在动作姿势列表内,当选定一个动作列表中的动作时,可以拖动姿势列表中的姿势到动作姿势列表中,这样就将姿势添加到了动作。当设计好一个动作时,可以拖动该动作到时间轴,点击播放,查看动作的连续播放演示。我们可以拖动动作列表中的一个动作到时间轴,添加到时间轴上,当点击播放时,下位机与3D四足机器人显示区

域会同步播放四足机器人的动作(因为下位机存在一定的通信延迟,所以下位机动作会比3D四足机器人显示区域的动作稍显延迟)。

2.2 香蕉派开发板

如图4所示,型号BananaPI M2,采用Allwinner A31S作为系统芯片,1Ghz ARM7四核处理器,Allwinner A31S Cortex-A7为CPU,PowerVR SGX544MP2为GPU,内存为1GB DR3(与GPU共享),需用5V/2A直流电供电,目前安装了安卓操作系统,可以下载应用,可以外接扩展模块,开发板具有4个USB2.0接口、一个microUSB接口、一个10/100/1000Mbps的有线以太网口、40个GPIO接口和一个HDMI接口,通过HDMI接口和USB接口可以外接控制设备及显示设备,从而对其进行二次开发。

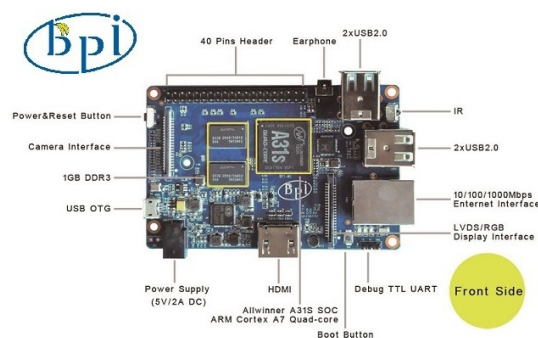


图4 BananaPI M2开发板

2.3 运动控制板

STM32运动控制处理板,通过TTL串口与BPI开发板相连,在外壳上留有RS-232串口通信接口,根据串口通信接口文档可以通过其他设备直接与机器人控制板通信以控制机器人执行相应指令,还可以使用上位机软件通过串口下载编辑的动作,并让机器人执行编辑的相应动作。

2.4 CCD摄像头

机器人头部有500万像素USB摄像头,从机器人内部连接到安卓开发板,用于获取机器人外部图像信息。机器人通过CCD摄像头获取的视觉信息进行二值化处理,加入改进PID算法,使四足仿生机器人可以沿着既定轨迹前进。

2.5 姿态传感器

能感受物体姿态（轴线对重力坐标系的空间位置）并转换成可用输出信号的传感器叫做姿态传感器^[5]。连接在运动控制板上，通过内嵌的低功耗ARM处理器输出校准过的角速度、加速度、磁数据等，通过基于传感器数据算法进行运动姿态测量，实时输出三维姿态数据。

3 软件设计

软件设计部分所完成的主要功能包括图像处理、循迹、颜色识别以及控制命令下发等。

主程序流程图如图5所示。

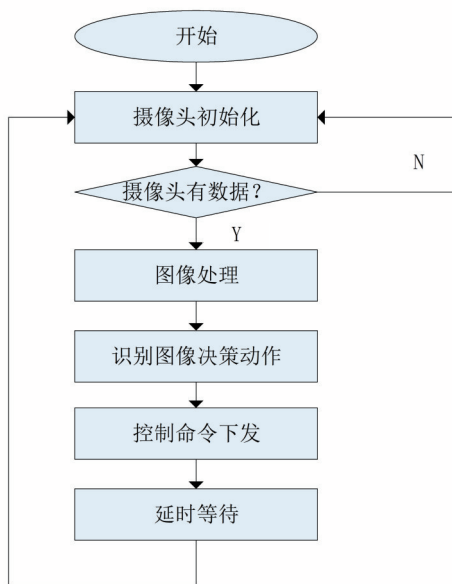


图5 主程序流程图

3.1 图像处理

图像处理部分的目标就是将摄像头拍摄的彩色画面变为二值化的黑白图像，以便后续图像处理和识别。具体操作包括图像分割、中值滤波和二值化。

3.1.1 图像分割

在将一帧图像读取到程序内部后，如图6所示，首先要对其进行分割，降低图像分辨率。对高分辨率的原始图像直接处理会浪费大量时间和算力，所以我们使用OpenCV中的resize函数将图像分辨率降低到80×60，在不影响精度的情况下可以提高程序运行速度和效率。

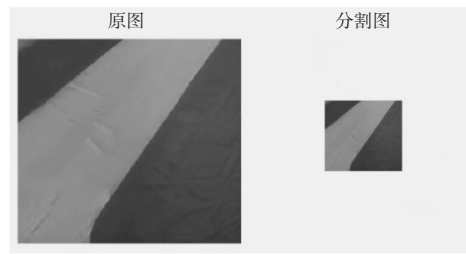


图6 图像分割前后对比

3.1.2 中值滤波

导入的图像通常会掺杂一些噪声，每一个噪点都可能影响边缘的判断，从而影响循迹效果，因此在图像分割后需要采用中值滤波减少噪声。本文采用OpenCV中的medianBlur函数对图像进行中值滤波处理。中值滤波是基于排序统计理论的一种能有效抑制噪声的非线性信号处理技术，中值滤波的基本原理是把数字图像或数字序列中一点的值用该点的一个邻域中各点值的中值代替，让周围的像素值接近真实值，从而消除孤立的噪声点，滤波前后效果如图7所示。

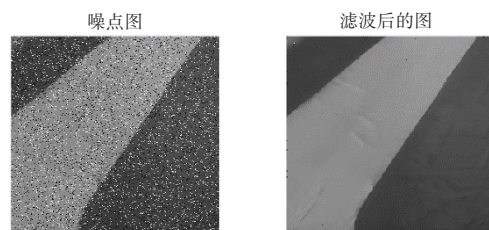


图7 中值滤波前后对比

3.1.3 二值化

如图8所示，滤波后的图像仍是彩色，通过二值化可以筛选出特定的颜色进行识别，OpenCV中的inRange函数可以将介于给定颜色范围内的像素值变为255，其他像素变为0。



图8 二值化前后对比

3.2 循迹策略

循迹算法受到摄像头高度限制,四足仿生机器人正常站立时摄像头距离地面高度约160 mm,赛道左右两侧不能同时扫描到,所以采取单边循迹的方式,沿着赛道内侧或者外侧循迹;由于赛道外侧经常有参赛队员(或观众)流动,会影响到仿生机器人的边界线识别;而进行内侧循迹可避免此类干扰,循迹效果更佳。

单边循迹策略具体做法是将二值化后的图像进行逐行、逐列扫描,判断并记录每行边界位置,对每行边缘位置取均值,将该均值作为机器人距边线的距离,通过控制其左右行走将该值控制在一定范围即可实现循迹。

3.3 颜色识别

在完成循迹的同时四足仿生机器人需要通过一系列障碍地形,为住户完成快递配送任务;在执行任务区域带有特定颜色的色块进行辅助提示,对这些颜色进行识别之后才能判定接下来将要执行的任务。

如图9所示,在图像处理阶段已经将图像对特定的颜色做了二值化处理,对特定颜色的二值化图像扫描,若该颜色像素数占总像素数百分之三十以上就可以认为识别到了该特定颜色,此时就可以执行相应的任务。

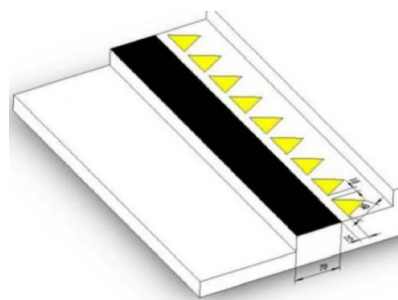


图9 特定颜色区域辅助提示图

3.4 控制命令下发

四足仿生机器人的控制驱动板采用串口通信,其内部的简单指令集已经可以实现直行、左转、右转、高速奔跑等动作。也可以利用其复杂指令集准确控制其各个关节的位置及运动速度快慢。根据上位机发送到下位机的指令包格式利用串口向驱动板发送指令包,即可实现对仿生机器人动作的控制。

4 结语

研究了四足仿生机器人的快递配送项目:包括规则解读、硬件平台搭建、相关机器视觉的软件设计等。着重论述了机器视觉中图像处理、循迹策略和颜色识别等主要内容。参加中国机器人大赛及人工智能挑战赛近三年的获奖成绩(见表1),验证了本文所设计的四足仿生机器人具有一定的实际应用价值。

表1 近三年四足仿生机器人参赛获奖表

参赛项目	本科生参赛人次	证书编号	获奖等级
2019中国机器人大赛	10	Y1906R039A0032	二等奖(国家级)
2020中国机器人大赛	11	Y2009R038A0005	一等奖(国家级)
2021年第二十三届中国机器人及人工智能大赛(全国初赛)	4	CRAIC2021-CS-Y4PB	二等奖(省级)
2021年第二十三届中国机器人及人工智能大赛(决赛)	4	CRAIC2021-NF-9MBR1B	三等奖(国家级)

参考文献:

- [1] 姚华. 浅析四足仿生机器人的研究现状及趋势[D]. 哈尔滨:哈尔滨远东理工学院,2019.
- [2] 中国自动化协会机器人竞赛工作委员会. 2021中国机器人大赛比赛规则[EB/OL].(2021-06-22). <http://robocup.drct-cao.org.cn/index.php/category/view?id=794>.

- [3] 山东优宝特智能机器人有限公司. 四足仿生机器人使用说明书[M],2018.
- [4] 邹明峻. 全肘式四足仿生机器人爬台阶步态研究[J]. 电气传动,2019(11):64-69.
- [5] 张琳. 一种基于MEMS运动姿态传感器的研究与设计[J]. 现代计算机,2017(6):12-13,24.

作者简介:

靳晨(2000—),男,河北张家口人,在读本科,研究方向为测控技术与仪器

刘刚(1963—),男,江苏泰兴人,本科,研究方向为计算机测控技术

吕兴东(2001—),男,山东济南人,在读本科,研究方向为光电信息科学与工程

黎航宇(2001—),男,四川成都人,在读本科,研究方向为测控技术与仪器

杨雅子琪(2002—),女,北京丰台人,在读本科,研究方向为测控技术与仪器

邓艺霖(2002—),男,北京西城人,在读本科,研究方向为智能感知工程

收稿日期: 2022-03-16 修稿日期: 2022-07-04

Research on Express Delivery Project Based on Quadruped Bionic Robot

Jin Chen, Liu Gang, Lyu Xingdong, Li Hangyu, Yang Yaziqi, Deng Yilin

(School of Instrument Science and Optoelectronic Engineering, Beijing Information Science and Technology University, Beijing 100192)

Abstract: The competition mechanism of the quadruped bionic robot in China Robot Competition is briefly summarized and the hardware platform structure of the quadruped bionic robot and the software scheme for express package delivery are presented in details in this paper. The practical results show that the bionic robot has the function of passing through terrain obstacles, such as speed bumps, steps, narrow bridges, slopes and grasslands, as well as QR codes recognition, express package distribution and express package delivery. From the perspective of the competition, the excellent results of the robot verified the correctness of the hardware design and control strategy.

Keywords: quadruped bionic robot; control strategy; express package delivery

(上接第 25 页)

Research on Recognition of Esophageal Cancer Images Based on Transfer Learning

Ling Qiyang, Liu Xiaofang, Zhang Yang

(School of Computer Science and Engineering, Sichuan University of Science & Engineering, Yibin 644005)

Abstract: Aiming at the poor recognition effect of esophageal cancer images in deep learning due to small data sets and unbalanced data, a recognition method based on transfer learning and AlexNet neural network model is proposed. First, perform corresponding data enhancement and preprocessing on the downloaded esophageal cancer data set as a target sample backup, and then select a large number of fully-labeled and associated ImageNet images to train the AlexNet network, and perform the effective feature extraction layer of the trained AlexNet network Freeze, perform parameter adjustment and structural optimization of the unfrozen replacement layer in line with the target sample classification task, and then put the target sample into the AlexNet network that completes the migration learning for training. Finally, the target samples are directly put into the traditional AlexNet, GooLeNet, and ResNet neural networks for direct training, and the results are compared with the AlexNet model after migration learning. The results show that the recognition rate of the AlexNet model based on transfer learning is significantly higher than that of the traditional AlexNet neural network model, which verifies the feasibility of transfer learning.

Keywords: transfer learning; AlexNet; esophageal cancer; esophageal endoscopy

文章编号: 1007-1423(2022)19-0057-06

DOI: 10.3969/j.issn.1007-1423.2022.19.009

面向威胁信息的网络安全态势感知研究

谷晓鹏

(91001 部队, 北京 100841)

摘要: 随着计算机的日益普及和通信技术的快速发展, 互联网在各领域发挥的作用越来越大, 计算机网络安全问题也越发受到关注。网络安全态势感知技术的目的是能够对安全要素进行获取、理解和预测, 是确保网络安全的重要手段之一。基于此, 首先参照国家相关标准, 面向网络威胁信息设计安全态势感知模型, 并在此基础上构建网络安全态势感知系统, 形成安全管控的智能决策。

关键词: 网络安全; 态势感知; 威胁信息

0 引言

计算机网络的日益普及和组网技术的快速发展, 网络在深入社会、政治、经济、军事等各个领域, 对社会治理、产业发展、个人生活、思想文化产生了巨大影响。网络在人类社会中的作用越来越大, 同时入侵攻击带来的损失也日益增大, 网络安全问题逐渐严重。网络攻防也日趋规模化、分布化、复杂化, 攻击手段向着态势变化迅速、破坏性越来越大的方向发展。日益严峻的安全威胁迫使政府、社会、企业不得不加强网络系统安全防护。目前网络攻击向着高级持续威胁等有组织的攻击发展, 传统的单点防御、各种独立的安全防御措施, 在新的攻击技术、新形势下, 已经完全无法满足需求。构建协同化、立体防御系统, 核心在于有效生成和发布网络安全态势信息, 使得防御方难于响应复杂网络攻击。在这种情况下, 迫切需要一种新的网络安全态势系统, 可协同各个网络防御单元实时掌握网络安全态势信息, 并及时消除网络威胁, 降低攻击造成的损失。网络安全态势信息是不同的安全厂商、网络运营单位之间进行安全情报共享、敌情我情评估的共性数据, 而且具备机器可读信息。因此, 开展网

络安全态势感知研究至关重要。

本文首先参照国家相关标准, 面向网络威胁信息设计安全态势感知模型, 并在此基础上构建网络安全态势感知系统, 为全面的网络安全防护体系构建提供前瞻且务实的指导思想。

1 国内外研究现状

态势感知(Situation Awareness, SA)最早用于研究飞行员对当前所处飞行状态的认识^[1]。Endsley等^[2]认为态势是对抗双方或多方, 在一定时空环境内的各方态势要素进行探测与信息收集, 并对获得的信息进行理解, 预测它们在不久将来的状态的方法。态势感知通常分为三个层次, 分别为察觉、理解和预测。

1999年, Bass^[3]提出了网络空间态势感知(Cyberspace Situation Awareness, CSA)和网络态势感知(Network Situation Awareness, NSA)的概念, 是首次将态势感知概念引入网络安全领域。本文基于Bass对CSA的定义, 认为网络安全态势感知是在大规模网络环境中, 对能够引起网络态势变化的安全要素进行获取、理解、可视化, 并能够对未来发展趋势进行一定的顺延性预测的技术方法, 其最终的目的是要获得网络安全防御的正确决策, 采取正确行动。所以,

网络空间安全态势既是一种状态，又是一种趋势，必须考虑整体、全局多个层级以及它们的关联性，用普遍联系的思维研究网络安全问题，因此任何单一的安全事件、局部的安全状态都不能称之为网络安全态势。

2007年，美国在爱因斯坦计划^[4]的基础上提出可信互联网连接(TIC)计划，提出网络出口管理。2010年又通过《联邦信息安全管理法》要求各机构的网络信息安全方案中，必须包含对各类信息系统进行持续监测^[5]。美国国家标准与技术研究院在2016年发布了网络威胁信息共享指南，该指南中所涉及的信息源的选择、威胁情报类型、数据源的选择、威胁指标等内容可以作为态势感知系统的有效参考^[6]。

国内对态势感知发展建设同样重视。2017年，中国移动通信集团公司业务支撑系统部制定了《中国移动业务支撑网升级安全威胁分析与预警平台技术规范》^[7]，规定了中国移动业务支撑网省级安全威胁分析与预警平台对业务支撑系统、管理信息系统各类安全数据的采集、存储、安全威胁分析、安全告警和安全预警的功能要求。2019年5月，发布的国家标准《信息安全技术网络安全威胁信息格式规范》(GB/T36643-2018)^[8]中，对网络安全威胁信息描述做出了标准规范，有助于整体的网络安全威胁态势感知能力的提升。目前态势感知已经逐步成为网络安全系统建设中越来越重要的组成部分。

2 安全态势感知模型

为确保构建的安全态势感知模型具备通用性、一致性和可移植性，本文在遵循GB/T36643-2018标准要求的基础上开展设计，首先介绍上述标准中的威胁信息模型。

2.1 威胁信息模型

为了能够实现不同组织间网络安全威胁信息的共享和利用，GB/T36643-2018标准定义了威胁信息模型，从对象、方法和事件三个维度对各类威胁信息进行了划分。经典的威胁信息模型包括可观测数据(Observation)、攻击指标

(Indicator)、安全事件(Incident)、攻击活动(Campaign)、威胁主体(Threat Actor)、攻击目标(Exploit Target)、攻击方法(TTP)、应对措施(Course Of Action)在内的八个威胁信息组件描述网络安全威胁信息。又可划分为对象域、方法域和事件域：

(1) 对象域：描述了网络安全行为的参与角色，核心就是攻防双方，包括“威胁主体”类角色(一般是攻击者)和“攻击目标”类角色(一般是被攻击方)。

(2) 方法域：描述网络安全威胁中的方法类元素，包括两个方面：一是“攻击方法”，二是“应对措施”，分别对应了攻击者的主要方法，也就是攻击和入侵所采用的方法、技术和过程，和防御方的主要技术方法，也就是针对攻击行为的防御措施，如预警、检测、防护、响应等动作。

(3) 事件域：描述网络安全威胁相关的事件，包括四个组件：攻击活动、安全事件、攻击指标和可观测数据，其中攻击活动通常以经济或政治为攻击目标，攻击事件是指对目标信息系统进行网络渗透的行为、安全时间主要指对终端、网络节点、设备实施的具体攻击，可观测数据主要是从网络或主机层面捕获流量和日志信息。

在对上述模型信息组件研究的基础上，经过各类安全时间和攻防模型的研究，本文在微观、中观和宏观三个层面构建了安全态势感知模型，并可划分为运行状态态势、攻击活动态势和攻击方法态势。

2.2 运行状态态势

运行状态态势是针对所要描述的网络环境，给出其运行状况的定性及定量评估，如图1所示。运行状态所基于的网络环境，应当是单个资产，或基于一定关系组织起来的局部网络或资产组合，从微观到中观到宏观，可基于应用、设备、信息系统、业务活动、域、子、整体网络等给出运行状态态势。基于时间序列给出的运行状态评估值序列称为运行状态变化趋势。

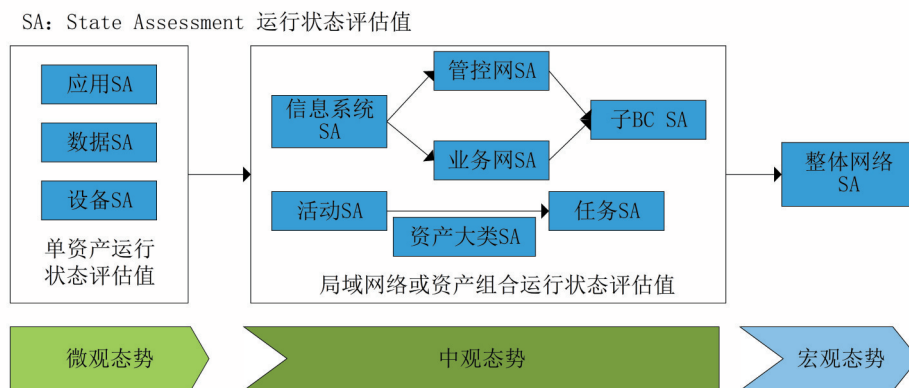


图1 运行状态态势

2.3 攻击活动态势

安全事件基于威胁主体及具体意图的融合理解，形成攻击活动。对于安全事件最直接的融合是基于过滤条件(如时间、威胁主体、攻击方法、攻击目标等)输出事件列表及事件次数统计信息。从微观、中观到宏观视角对安全事件-攻击活动的态势如图2所示。

同样，也可针对事件数量及威胁评估值在时间序列上的变化形成事件数量时间趋势及威胁评估值时间趋势，如图3所示。

从态势感知角度，基于威胁事件的特征把事件分配到攻击链各个环节，从IP关联、时间关联、大类关联、阶段关联等多种关联手段分析攻击者意图，建立起事件与事件之间的关联关系，还原攻击过程。

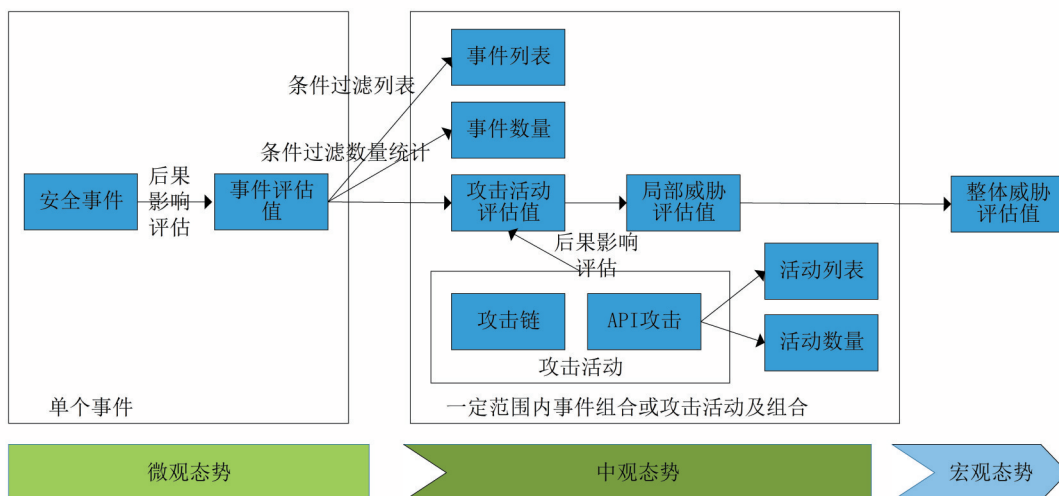


图2 安全事件-攻击活动态势



图3 攻击链模型

从攻击链模型的七大阶段看，前三阶段事实上是攻击的试探及准备阶段，第四个阶段“渗透”是个分界线，后面三个阶段都是渗透成功，获得部分权限后所执行的威胁活动。攻击链的具体实例，基于不同的威胁过程，可能不一定包含七个步骤的内容。

2.4 攻击方法态势

攻击方法的分类分为两种维度。其一是对攻击技术手段的维度，其二是针对攻击目标所利用的脆弱性进行分类。

对攻击技术手段的分类分为基于攻击机制及攻击领域两种分类方式。针对攻击目标的脆弱性的分类方式可按研究概念视图、开发概念视图、架构概念视图等不同视图进行分类。

不管采用哪种分类方法，攻击方法的分类都是多层次的结构，如图4所示。攻击方法在微观上的态势分析是针对单个事件具体使用的攻击方法的分析。对攻击方法的态势融合在分类层面的从微观层次到中观再到宏观的融合遵循从单个具体的方法到小类型到上级类型的规律。

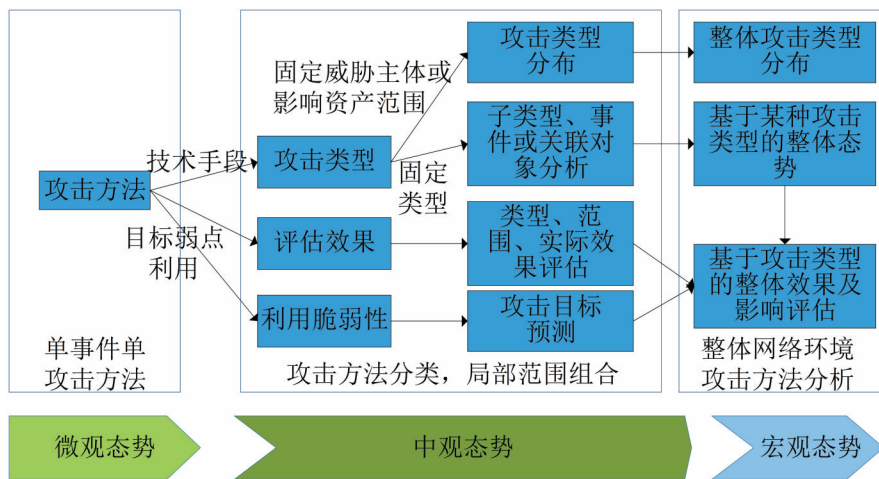


图4 攻击方法态势

3 网络安全态势感知系统框架

结合大数据、人工智能等技术，实现安全态势感知模型的应用，形成网络安全态势感知系统框架如图5所示。该系统分为资源层、采集管控层、大数据层、服务层和业务层，采集层对资源层进行安全数据采集，通过大数据层实现数据存储与分析，这三层为平台的服务层和业务层提供数据支撑。而平台的服务层和业务层，为用户提管理服务和技术服务。

3.1 数据支撑服务

网络安全态势感知系统在数据支撑上，支持对网络下所有节点日志的统一管理，以及多种日志采集方式收集云平台设备和系统日志，并通过对日志的分类、过滤、强化、分析和存储，为技术服务与管理服务提供数据支撑。

通过内置过滤器，系统对网络设备、安全设备、安全资源池等进行日志过滤、归并和规范化，过滤掉严重程度较低的原始日志信息。通过指定日志影响的设备、日志采用协议、日志类别、日志标题等日志属性进行过滤，对日志数据过滤的开启状态进行手工设定，多级过滤。日志采集引擎支持预处理安全日志，在采集引擎段即可生成标准化的日志格式。通过交互式界面，动态识别和提取日志关键信息，并自动生成正则表达式，通过所见即所得的交互式模式进行正则的验证。

3.2 管理服务功能

系统以安全管理体系为输入，进行管理流程设计。如结合安全管理体系中的规章制度、规范流程设计平台的工作流；参照管理机构与人

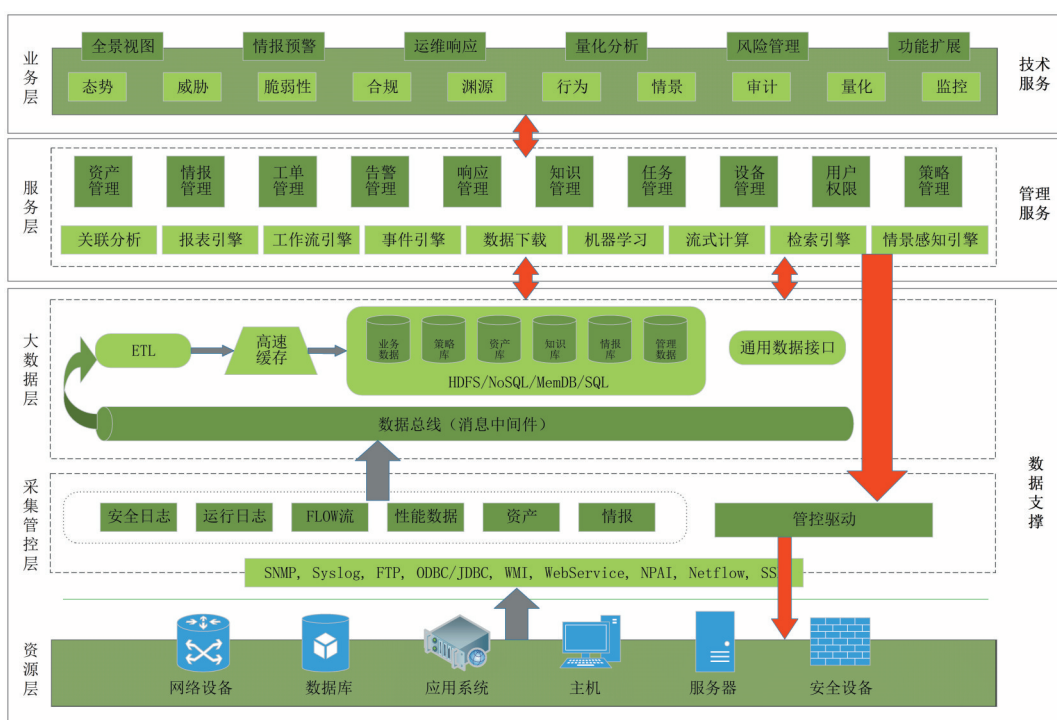


图5 网络安全态势感知系统框架

员，对应用户设置不同权限，制定审批链，并以此为参照，进行用户权限管理；结合运维过程中的记录、表单，设计工单管理中的流转页面。在平台运营过程中，结合平台资产库、漏洞库、配置基线库等，通过平台开展运维管理。

结合网络内的扫描类产品，系统支持网络资产自动发现功能，能够自动发现和识别资产。能够实现全局模式下通过搜索IP地址、资产名称、MAC、操作系统、资产编号、物理位置、资产类型等方式查询资产表信息，实现事件快速定位。

脆弱性管理包括漏洞脆弱性和配置脆弱性两个部分，是网络环境重要的风险评估项，在等保与分保标准中，也要求定期开展配置检查和漏洞扫描。脆弱性管理功能从脆弱性、漏洞情报和资产管理三维度出发，分析资产受影响情况，提供防护措施。通过脆弱性管理功能模块加强漏洞风险管理。

3.3 技术服务功能

系统以安全态势感知模型为基础，综合应用威胁情报系统，建立攻防场景模型，对态势

进行大数据分析，对各类相关数据设计了可视化展示方法，能够建立对安全态势的全面监控，在综合其他数据源的基础上，可以进行安全威胁的实时预警和安全事件的智能决策，具备安全事故联动响应的能力。系统能够高效地结合情境上下文分析，协助安全运维人员和安全分析工程师快速发现和分析安全问题。指导实际运维手段，提升网络安全防御水平。

平台在技术服务上包含三大核心：情报预警中心、态势感知中心、联动响应中心。

威胁情报管理根据来自内部预警信息和外部预警信息，分析获得对可能发生的威胁的提前通告。

安全中心安全态势分析的定位是可以针对整体范围或某一特定时间与环境，基于条件开展微观、中观和宏观的态势感知评估，最终形成历史的整体态势以及对未来短期的预测。

联动响应的安全架构自顶向下可分为安全应用、安全控制平台和安全设备三层。态势感知应用对整体攻防态势进行感知，决策引擎进行分析判断，最终生成安全决策，向安全控制平台下达处置策略。

4 结语

网络安全态势感知系统是安全体系中融合技术与管理,提供整体运维和安全管控的支撑平台。它一方面从网络中采集安全数据,进行安全状态实时监控,并通过大数据分析,结合威胁信息,形成多层面的态势感知呈现。通过对全网安全数据的采集,应用大数据、机器学习等技术,形成安全管控的智能决策。

参考文献:

- [1] DURSO F T, GRONLUND S D. Situation awareness[J]. Handbook of Applied Cognition, 1999, 283: 314.
- [2] ENDSLEY M R, HOLDER L D, LEIBRECHT B C, et al. Modeling and measuring situation awareness in the infantry operational environment[R]. Alexandria, VA: U.S Army Research Institute for the Behavioral and Social Sciences Infantry Forces Research Unit, 2000.
- [3] BASS T. Multisensor data fusion for next generation distributed intrusion detection systems[C]//Proceedings of the IRIS National Symposium on Sensor and Data Fusion. Citeseer, 1999, 24(28): 24-27.
- [4] 王兵, 武杰. 爱因斯坦-3 计划解读与分析[J]. 中国信息安全, 2011(2): 51-54.
- [5] 徐丽萍, 郝文江. 美国政企网络威胁情报现状及对我国的启示[J]. 信息安全, 2016(9): 278-284.
- [6] 李留英. 美国网络威胁情报共享实践研究[J]. 信息安全研究, 2020, 6(10): 41-46.
- [7] 赵康. 网络安全态势感知在企业应用中的研究[D]. 太原: 太原理工大学, 2021. DOI: 10.27352/d.cnki.gylgu.2021.001577.
- [8] 李书峰. 信息安全技术网络安全威胁信息格式规范正式发布[J]. 信息安全与通信保密, 2018(11): 7.

作者简介:

谷晓鹏(1976—),女,河北承德人,硕士研究生,高级工程师,研究方向为数据工程

收稿日期: 2022-05-13 修稿日期: 2022-06-24

Research on Network Security Situational Awareness Oriented to Threat Information

Gu Xiaopeng

(Army of 91001, Beijing 100841)

Abstract: With the increasing popularity of computers and the rapid development of communication technology, the Internet has played an increasingly important role in various fields, and the issue of computer network security has also attracted more and more attention. The purpose of network security situational awareness technology is to obtain, understand and predict security elements, which is one of the important means to ensure network security. Based on this, firstly refer to the relevant national standards, design a security situational awareness model for network threat information, and build a network security situational awareness system on this basis to form intelligent decision-making for security management and control.

Keywords: cybersecurity; situational awareness; threat information

文章编号: 1007-1423(2022)19-0063-05

DOI: 10.3969/j.issn.1007-1423.2022.19.010

基于支持向量机的微博情感分析方法研究

李首政¹, 王 琪¹, 王 力²

(1. 南阳理工学院信息工程学院, 南阳 473000; 2. 南阳理工学院土木工程学院, 南阳 473000)

摘要: 微博是当前国内最流行的社交平台之一, 微博文本的情感分析有助于进一步分析实现其媒体价值, 然而, 微博数据庞大且冗余性高, 使得文本特征具有较高的稀疏性和局限性, 在小样本数据分析上情感判断结果并不理想。因此, 提出一种基于支持向量机分类模型的微博数据情感分析方法, 首先通过 weibo Spider 爬取微博数据, 进行人工标注构建微博文本数据集, 然后联合优化 TF-IDF 算法和传统词袋, 提出一种基于关键词的词袋模型, 获取文本特征矩阵以解决微博文本高稀疏、高冗余的问题, 最后构建高斯核的支持向量机分类器实现对微博数据的情感分析。实验结果显示, 对比朴素贝叶斯、决策树等方法, 提出的方法可获得较高的准确率, 且在小样本数据上有明显优势。

关键词: 微博文本; 情感分析; 支持向量机; 机器学习

0 引言

微博作为国内最大的社交媒体之一, 是一种极其丰富的文本来源, 具有内容精炼、时效性强、简短、通俗等特点, 研究专门的情感分析技术有利于决策者分配业务及合理的决策^[1]。

过去人工监控和分析的方法不仅需要耗费大量的人力成本, 而且产生了很大的滞后性。目前对微博数据进行情感分析的方法主要包括两大类, 基于情感词典的方法和基于机器学习的方法。基于词典的方法虽然设计思想简单, 准确率较高, 但需依赖于专业人士的词典构造, 人工工作量巨大, 且受限于词典先验知识, 普遍性不强^[2]。基于 KNN 的分类方法虽然算法简单, 但准确率较低且内存占用量较大^[3]。Krishnaveni 等^[4]提出面向文本的朴素贝叶斯分类器的基本思想, 然而经典的回归思想和贝叶斯方法都是基于一个假设, 即概率分布及分布模型是先验的, 但实际数据集往往并不同分布, 因此准确率较低。基于深度学习的方法采用 RNN 的方法虽然准确率较高, 但需大量数据进行模型训练, 且参数较大, 实际应用中小样本数据存在过拟合问题^[5]。支持向量机(SVM)方法采用监督方式学习数据的分布模型, 从而有效避免了

先验模型与实际模型差距较大的问题^[6-7]。

本文分析微博文本数据, 提出了一种基于支持向量机的微博文本分析方法, 首先爬取微博文本数据并进行预处理, 然后采用 TF-IDF 算法进行文本向量化, 通过词袋模型获得文本特征向量, 设计高斯核支持向量机分类模型对文本特征向量进行分类器训练, 最终, 获得微博文本数据的情感分类结果, 通过实验分析本文支持向量机方法的准确性, 总体流程如图 1 所示。

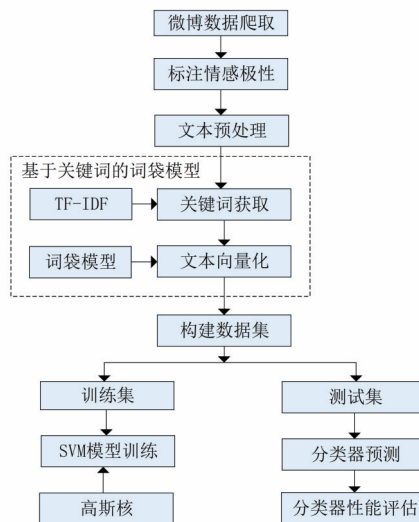


图 1 总体结构

1 文本预处理及特征获取

通过微博数据爬取工具 Weibo Spider 爬取微博平台个人发布数据, 获得微博文本原始数据。然而, 直接爬取的数据存在大量冗余和无效内容, 且计算机并不能直接处理文本信息, 因此需要对文本数据进行预处理, 使得文本特征向量更准确, 从而达到提高训练准确性的目的。文本预处理包含文本清洗、情感标注、文本分词、去除停用词四部分。预处理后的文本计算机并不能直接处理, 本文提出基于关键词的词袋模型, 获得文本特征矩阵降低文本特征矩阵稀疏性, 最终, 构建了实验数据为 10000 条的训练集和 500 条的测试集数据。

1.1 文本预处理

1.1.1 文本清洗和情感标注

由于爬取程序是直接复制博主发布的内容, 导致爬取的数据中包含了“#”开头和“#”结尾的 Tag 等较多无效内容, 因此需要对无效信息进行清洗。同时, 由于微博短文本的特点, 爬取信息中包含较多表情符号, 而这些表情符号可以很好地体现人的情感, 因此将表情符号转换为同义文本表示, 比如微博常见的允悲表情, 用文字“允悲”表示。文本分类是有监督学习, 因此清洗后的训练集文本需进行情感标注, 根据自然语言知识及人类普遍认知对爬取的文本逐条进行情感标注, 本文设计的情感分为积极和消极两种情感, 积极情感用标签“1”表示, 消极情感用标签“0”表示。

1.1.2 文本分词和去除停用词

文本由词语组成, 而与英文句子相比, 中文句子中若没有空格则无法简单地识别词语, 因此需要分词。本文采用 jieba 分词工具进行文本分词。停用词是指一些无意义词的中文词, 比如“你, 我, 他”, “如果, 那么”。去除停用词是常见的 NLP 预处理的一个步骤, 能帮助减少特征矩阵的大小, 从而减少特征矩阵的计算消耗, 实现减少计算时间和成本的目的。清洗后的文本标注“积极”或“消极”情感标签, 用于构建数据集及模型训练。

1.2 基于关键词的词袋模型

本文采用词袋模型将文本进行向量化。但传统词袋模型获得的文本向量稀疏性较强, 存在很大冗余且不能体现句子关键词, 因此本文提出基于关键词的词袋模型, 从而获得文本特征向量, 算法通过 IF-IDF 方法获得, 包括三个步骤。

第一步, 通过 TF-IDF 算法提取文本特征词, 由特征词构建新的文本词袋, 从而解决了传统词袋模型高冗余性的问题, 同时能够体现词的权重概率。

特征词提取主要是指提取出有利于情感分析分类的情感词, 本文采取 TF-IDF 算法计算每个词的 IF-IDF 值, 并采用 L2 范数进行标准化, 选取文档权值最高前 20 个词作为特征词。TF 是词出现的频率, TF-IDF 的基本思想是如果某个单词在一篇文章中出现的频率 TF 高, 并且在其他文章中很少出现, 则认为此词或者短语具有很好的类别区分能力, 适合作为特征词, 算法见式(1)~(4)。

$$T_{\omega} = \frac{\text{在某个词条}\omega\text{出现的次数}}{\text{该类中所有的词条数目} + 1} \quad (1)$$

$$IDF = \log\left(\frac{\text{语料库的文档总数}}{\text{包含词条}\omega\text{的文档数} + 1}\right) \quad (2)$$

$$TF - IDF = TF * IDF \quad (3)$$

$$TF - IDF_{\text{normal}} = \frac{TF - IDF}{\|TF - IDF\|_2} \quad (4)$$

第二步, 运用新建的词袋表计算句子的词袋向量, 并将每个词的 TF-IDF 值替代传统词袋向量中词出现的次数。

第三步, 数据标准化的目的是通过标准化得到均值为 0 和标准差为 1 的服从标准正态分布的数据, 此外, 标准化能提升分类器拟合的速度, 同时缩小可能的异常值的影响。本文采用数据标准化的方法是减去均值然后除以方差或标准差, 经过这种数据标准化方法处理过的数据符合标准正态分布。

2 分类器构建

2.1 线性分类器

支持向量机算法的基本思想是寻找一个分

类器,使得超平面和最近的数据点之间的分类边缘(超平面和最近的数据点之间的间隔)最大。对于SVM算法,通常认为分类边缘越大,平面越优,通常定义具有最大间隔的决策面就是SVM要寻找的最优解。因为支持向量到决策边界有一定的距离,因此支持向量机容许一定的误差,算法的鲁棒性得到一定的提高。对线性可分样本集 $T = \{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \dots, (x^{(m)}, y^{(m)})\}$, 其中 $x^{(m)} \in \mathbb{R}^n$, $y^{(m)} \in \{-1, 1\}$; 选择惩罚参数 $C > 0$, 构造并求解凸二次规划。

$$\min_a \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m a_i a_j y_i y_j (x_i \cdot x_j) - \sum_{i=1}^N a_i \text{ s.t. } \sum_{i=1}^m a_i y_i = 0, \quad 0 < a_i < C, i = 1, 2, \dots, N \quad (5)$$

其中 a_j 为拉格朗日乘子, 得到最优解 $a^* = (a_1^*, a_2^*, \dots, a_N^*)^T$ 。

计算 $\omega^* = \sum_{i=1}^m a_i^* y_i x_i$, $b^* = y_j - \sum_{i=1}^m a_i^* y_i (x_i \cdot x_j)$, 得到超平面 $\omega^* x + b^* = 0$, 则分类决策函数:

$$f(x) = \text{sign}(\omega^* x + b) \quad (6)$$

可见, SVM分类问题需求解样本的内积, 然而实际样本往往线性不可分, 因而需引入核函数, 核函数可将样本向量映射到高维核空间, 使得原本线性不可分的样本变为线性可分, 核函数一般包括线性核函数、多项式核函数和高斯核函数。

2.2 高斯核函数

高斯核函数可以将输入特征映射到无限多维, 不仅可以解决样本线性不可分的问题, 而且能够突出样本特征的相似性, 同时相比多项式核函数计算量更小, 因其仅需一个参数 σ , 所以调参较易。

$$K(x^{(i)}, x^{(j)}) = \exp\left(-\frac{(x^{(i)} - x^{(j)})^2}{2\sigma^2}\right) \quad (7)$$

引入高斯核函数后分类器优化问题变为

$$\min_a \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m a_i a_j y_i y_j K(x_i \cdot x_j) - \sum_{i=1}^N a_i \text{ s.t. } \sum_{i=1}^m a_i y_i = 0, \quad 0 < a_i < C, i = 1, 2, \dots, N \quad (8)$$

3 实验与分析

采用本文方法爬取数据并进行文本处理后构建了10000条数据作为训练集, 500条作为测试集。统计出训练集的文本长度分布如图2所示, 文本长度集中在31—151区间, 更接近于真实语料环境。训练集经过标注后得到1标签和0标签的文本数量分布如图3所示。图3横坐标的0代表着文本的情感是消极的; 横坐标1代表着文本的情感是积极的。由图3可见训练样本中积极情感占比54.96%, 消极情感占比45.04%, 积极情感略高于消极情感, 但整体数量相当, 合理的数据分布有利于提高分类器训练的准确性。

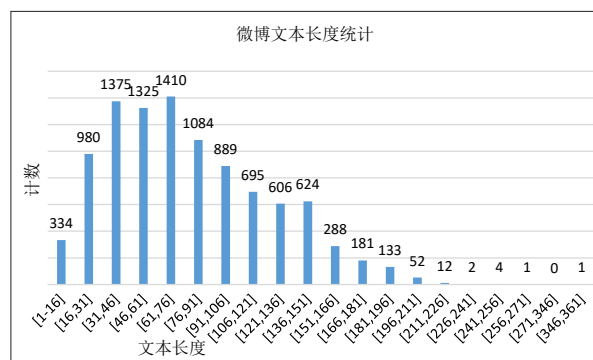


图2 微博训练集文本长度统计

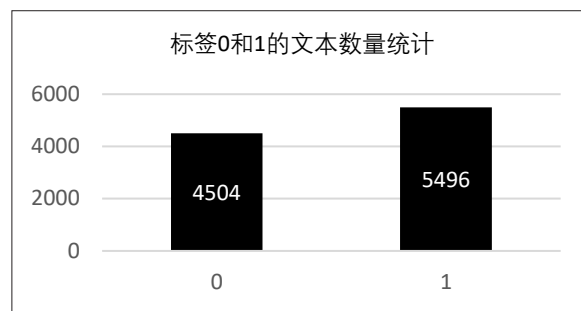


图3 微博训练集情感极性分布

通过基于关键词的词袋模型对10000条训练集文本进行特征词提取, 共提取出61949个特征词, 并获得(10000, 61949)大小的文本特征矩阵, 500条测试集特征矩阵大小为(500, 1949)。利用本文高斯核SVM的方法对训练集数据进行训练, 学习曲线如图4所示, 由图4中准确率的变化趋势可见, 训练集和验证集的准确率随着样本数的增大能够较好地拟合, 最终稳定在较

高准确率，说明本文方法获得的训练模型有较好的泛化能力。与通过传统词袋模型采用多项式SVM、线性SVM、朴素贝叶斯、KNN和决策树五种方法的训练集和测试集得分进行对比，结果如表1所示，SVM三种分类器表现出较理想的效果，其中本文高斯核的SVM方法在训练集和验证集上都取得了较高的准确率，且拟合度较好，KNN方法和朴素贝叶斯方法的准确率不高，决策树方法出现了较严重的过拟合现象。采用本文高斯核SVM方法和上述五种方法分别建立分类器对测试集进行验证，其预测的准确率如图5所示。

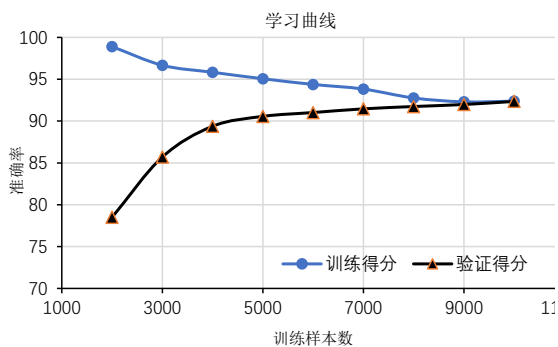


图4 高斯SVM分类器学习曲线

表1 各种方法的准确率对比结果

	线性核 SVM	高斯 SVM	多项式 SVM	KNN	朴素贝 叶斯	决策 树
训练集	91.65	95.23	94.38	93.54	87.63	95.73
验证集	88.91	92.35	91.83	90.57	85.84	71.38

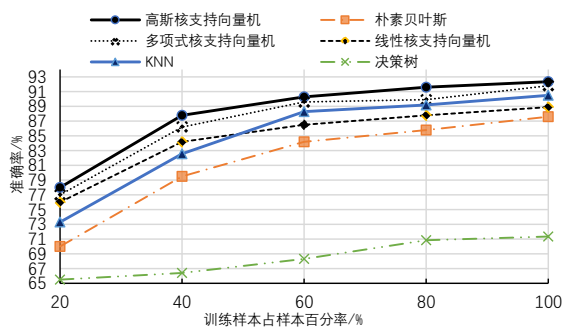


图5 分类器准确率对比图

图5可见，使用小样本训练支持向量机分类器已经取得不错的准确率，与同等样本数训练

的其他分类器相比有明显的优势，随着样本数的增加，六种分类器的准确率均有所提高，且KNN和朴素贝叶斯分类器与SVM分类器准确率差距有所缩小，其中本文高斯SVM的方法准确率最高。决策树方法由于存在过拟合现象导致总体效果不佳。由此可见，本文采用的高斯SVM方法在微博情感分析中能取得较高的准确率，且相比KNN、朴素贝叶斯等方法，本文方法在小样本数据集上有较明显优势。

4 结语

针对当前需要对内容精炼、时效性强、通常包含简短且非正式缩写词的微博短文本情感分析问题，本文提出一种更适合微博文本情感分析的方法，通过weibo Spider爬取微博数据并进行人工标注，构建数据集，TF-IDF算法和词袋模型构建基于关键词的词袋模型，获得文本特征矩阵，最后采用了高斯核的支持向量机方法进行分类器训练，通过测试集验证准确率，对比多项式支持向量机、KNN、朴素贝叶斯决策树等方法，实验结果显示本文方法准确率较高，在小样本数据的情况下可获得较高的准确率。

参考文献:

- [1] 朱琳琳,徐建. 网络评论情感分析关键技术及应用研究[J]. 情报理论与实践,2017,40(1):121-126.
- [2] 杨立月,王移芝. 微博情感分析的情感词典构造及分析方法研究[J]. 计算机技术与发展,2019,29(2):19-24.
- [3] ZHANG S, LI X, ZONG M, et al. Efficient kNN classification with different numbers of nearest neighbors[J]. IEEE Transactions on Neural Networks and Learning Systems,2018,29(5):1774-1785.
- [4] KRISHNAVENI G, SUDHA T. Nave Bayes text classification: a comparison of event models[J]. Imperial Journal of Interdisciplinary Research, 2017, 3(1):290-294.
- [5] 杨丽,吴雨茜,王俊丽,等. 循环神经网络研究综述[J]. 计算机应用,2018,38(S2):6-11,31.
- [6] 周杰,林琛,李弼程. 基于机器学习的网络新闻评论情感分类研究[J]. 计算机应用,2010,30(4):1011-1014.
- [7] YE Q, ZHANG Z Q, LAW R. Sentiment Classification of online reviews to travel destinations by supervised machine learning approaches [J]. Expert Systems with Applications,2009,36(3):6527-6535.

文章编号: 1007-1423(2022)19-0067-05

DOI: 10.3969/j.issn.1007-1423.2022.19.011

基于 U-net 算法和 FCOS 目标检测算法的洗衣机外壳钣金孔洞缺陷检测系统

邓秋菊, 张振龙

(重庆移通学院, 合川 401520)

摘要: 对外壳钣金冲孔过程的孔洞检测是洗衣机自动化生产线上不可缺少的环节。针对洗衣机外壳钣金的孔洞检测这一问题, 研发设计了一套基于 U-net 图像分割算法和 FCOS 目标检测算法的检测系统。针对钣金孔洞的图像分割需求, 引入 U-net 算法进行实现, 并根据洗衣机外壳钣金孔洞的检测特点, 利用 FCOS 算法来解决孔洞目标的检测问题。最终检测数据表明: 本系统在自动化洗衣机生产线上对于洗衣机外壳钣金孔洞缺陷的检测精度达到行业标准, 具有良好的检测精度, 完全能够满足正常生产时的检测需求。

关键词: 外壳钣金孔洞; 目标检测; 图像分割; 缺陷检测

基金项目: 重庆市教育委员会科学技术研究项目(KJQN202102401)

0 引言

随着制造业的迅速发展以及现代化车床的广泛应用, 工业制造生产效率得到了巨大的提升, 在生产效率提升的同时伴随出现了生产缺陷的问题, 例如洗衣机外壳钣金冲孔是否按标准模板进行冲压, 一旦冲孔出错, 将会造成批量问题产品。对于工厂来说, 缺陷检测是至关重要的一项工作, 然而人工进行工业钣金的缺陷检测费时费力, 且长时间高强度检测作业时误判率较高, 亟需一种自动化的检测系统来进行高效、准确的工业钣金缺陷检测^[1]。得益于人工智能技术的日益发展, 由机器视觉+深度学习的技术检测组合在现代化工业生产线上得以应用。

本文针对洗衣机钣金冲孔的缺陷检测问题, 基于 U-net 图像分割算法和 mmdetection 框架下的 FCOS 算法 (FCOS: Fully Convolutional One-Stage Object Detection), 该算法以逐像素预测的方式解决目标检测问题, 类似于语义分割, 最终实现了钣金缺陷检测的效率和准确性的提升。

1 U-net 图像分割

U-net 网络是一种图像语义分割网络, 计算机根据图像的语义对图像中的每一个像素点进行

分类, 再根据像素点的类别进行区域的划分, 从而可以实现在复杂环境的图片中对特定目标进行指定分割的功能。

U-net 网络结构如同一个字母 U, 如图 1 所示, U-net 网络的左侧对图像进行特征提取, 随着网络层的加深, 图像将会变小, 提取的图像特征也会越精确。在特征提取的过程中可能会出现部分图像信息的丢失, U-net 网络的右侧部分将会对图像特征进行还原, 加入对应的丢失部分的图像信息。

U-net 的工作原理如下, 针对一张待处理的图片, U-net 会将图像的识别区域转化为不同类别的标签, 每一个需要识别的物体都需要一个对应的通道, 所有物体通道的叠加就是整张图片最后得到的分割结果^[2]。

one-stage 和 two-stage 目标检测算法作为两种被广泛使用的经典目标检测算法, one-stage 目标检测算法以算法结构简单、运行速度快作为其特点, 具有代表性的 one-stage 目标检测算法有 RetinaNet、YOLOv3、CenterNet、SSD 等。这类算法能够直接对图像进行检测处理来实现实时图像检测, 例如钣金孔洞的缺陷检测就是 one-stage 目标检测算法的典型应用场景, 但

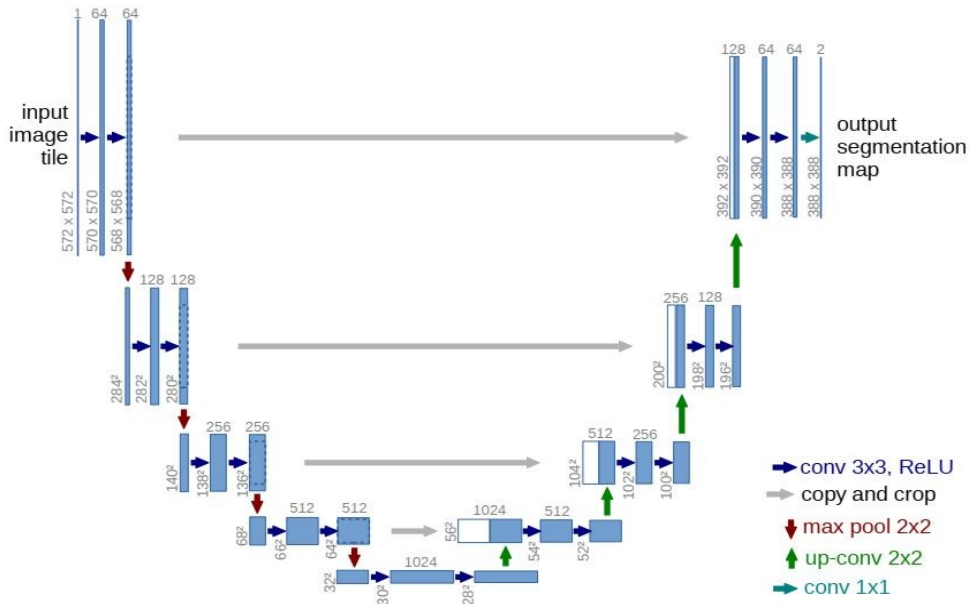


图 1 U-net网络结构

one-stage 算法对细小目标和复杂场景下的目标检测精度较低，这一问题在 two-stage 目标检测算法内得到解决。得益于 two-stage 目标检测算法先提取候选区，使用该类算法能够最大化地学习到目标的特征，其检测精度相较于 one-stage 检测算法有极大的提升，但其网络结构较为复杂，获取候选框时计算量较大^[2]，two-stage 目标检测算法的检测速度较慢，无法满足实时检测功能的需求，典型的 two-stage 目标检测算法有 Faster R-CNN、TridentNet 以及 Cascade R-CNN 等。

相较于经典的 one-stage 和 two-stage 检测算法，FCOS 的核心思想是预测输入图像中每个点所属的目标类别和目标框，由此 FCOS 摆脱了对锚框 (anchor box) 的依赖^[2]，对于不同的任务，FCOS 算法不再需要重新配置 anchor box 的参数，因此 FCOS 具有很好的普适性。此外，在训练模型时大量的 anchor box 会造成正负样本间的不平衡，并且计算这些 anchor box 与真实框的交互比 (IOU) 需要消耗大量的内存和时间^[3]。

FCOS 在网络结构上大体与 FCN 一致，其关键在于图 2 中的虚线框内的预测层。

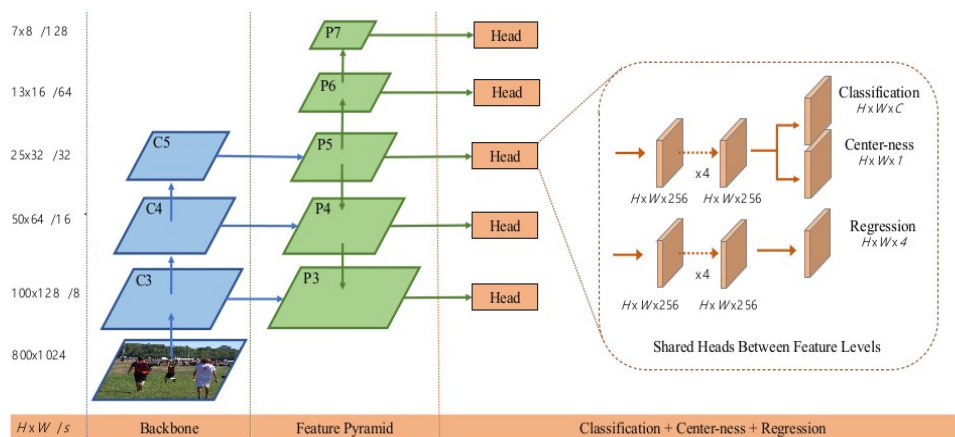


图 2 FCOS 网络结构

FCOS的预测层由分类支路、Centerness和回归支路三部分组成^[3]:

(1) 分类支路,主要功能是将特征上的位置坐标 (x, y) 转换成输入图像的位置,图2中 C 表示类别数, $H*W$ 则表示特征的大小。

(2) Centerness,负责计算每个点到目标中心点的距离,去除距离目标中心点较远的点。

(3) 回归支路,将会由 (l, t, r, b) 4个值分别表示目标框内一点距离上下左右边框的距离。

总结FCOS算法的实现步骤,可分为4步:

(1) 图片预处理操作。

(2) 在backbone网络结构中获取原图片的特征图,在特征图中的每一点执行回归操作,不断训练获取网络模型。

(3) 将测试集内的图片输入到训练出的网络模型进行测试,获得多个预测结果。

(4) 使用NMS等后处理操作获得最后的结果。

2 采取U-net图像分割算法+FCOS目标检测算法的优势

洗衣机外壳钣金孔洞缺陷检测系统能够在不同的生产线环境下保持较高精度完成钣金的缺陷检测任务,关键得益于采取了U-net图像分割算法+FCOS目标检测算法对待测目标进行处理。

洗衣机外壳钣金孔洞缺陷检测系统的其中一个难点问题在于,应用到不同环境的生产线时检测结果仍要保持高精度。在不同的生产线中,例如灯光照明强度、可支持拍摄角度、照明的稳定性、钣金位置的一致性环境因素,通过搭建暗箱能够有效解决外界自然光照影响。对不同钣金如何选择合适的模型且保证最终检测的高精度,对于此问题,通过引入U-net图像分割算法根据图像的语义对逐个像素点进行分区和区域划分,在进行钣金检测时对缺陷检测区域进行像素级的区域划分,从而达到对检测目标与无关背景区域的明确区分。后续在对目标特征的进一步提取过程中,目标特征的准确度将会进一步提高,增强提取到特征的鲁棒性,并且U-net算法对于图像特征的还原功能能够保障随着网络层加深,目标特征提取精确度不会出现下降的现象^[4]。

另外,我们还要面对的两个主要问题分别是:

(1) 针对不同的钣金生产线,其钣金大小和具体特征也不尽相同,洗衣机外壳钣金孔洞缺陷检测系统要保证应用于不同钣金生产线时具有普适性,而不是针对不同生产线每次都要对检测器进行重新调整。

(2) 部分钣金存在检测点相邻过密的问题,出现重叠区域,对目标检测造成了干扰。

为了解决以上两个问题,我们采用了FCOS目标检测算法。作为Anchor-Free的目标检测算法,在对不同生产线的钣金进行目标检测时FCOS算法会对锚框(anchor box)进行预测,无需对锚框进行重新设置,对应用于不同生产线的普适性问题得以解决。

在目标检测当中存在一种重叠现象,例如:第一张图片中是一个人手持一部手机,第二张图片中是一对恋人拥抱,此时算法对这两种图片进行预测时就会发生人和手机、人和人两两预测框重叠的现象,结果会造成第一张图片预测结果受到干扰,第二张图片出现重复标记的现象。

钣金孔洞检测区域的重叠现象是无法避免的,我们检测的钣金当中,如果出现一个较大的孔洞有一个或多个较小的螺丝时,就可能被认定为包含关系从而产生重叠,FCOS算法引入特征参数 P_i 为特征层 i 的最大距离,由此来对负样本进行过滤,减少重叠区域,判断是否满足为负样本的公式为

$$\max(l^*, t^*, r^*, b^*) > P_i$$

或

$$\max(l^*, t^*, r^*, b^*) < P_{i-1}$$

其中 P_i 设置为0, 64, 128, 256, $+\infty$,形成5个层来限制尺寸以减少重叠区域。

对于正样本的检测,FCOS采用了Centerness来进行优化,公式为

$$Centerness^* = \sqrt{\frac{\min(l^*, r^*)}{\max(l^*, r^*)}} * \sqrt{\frac{\min(t^*, b^*)}{\max(t^*, b^*)}}$$

当损失 l 越小时,Centerness越接近于1,说明该预测框的中心距离GT Box的中心越近,这样的样本则作为正样本加入到后续的回归工作中^[5]。运用了FCOS算法后的目标检测效果,如图3所示。

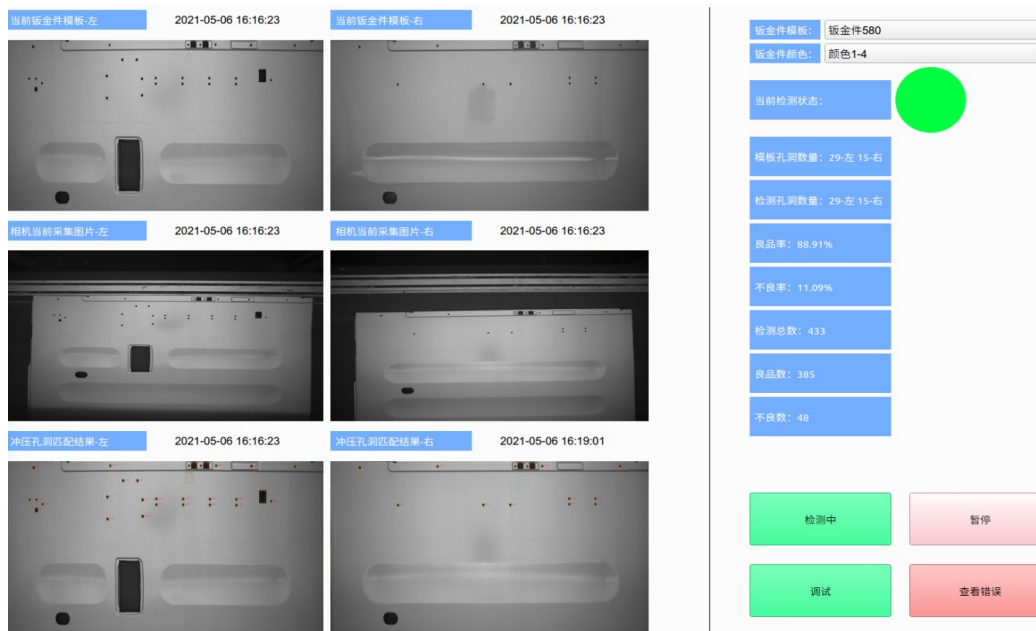


图3 引入FCOS目标检测算法后的检测结果

在检测系统中挑选了两组样本进行检测，引入了FCOS目标检测算法后的缺陷检测结果如表1所示。

表1 引入后检测结果

总检测数	漏检数	误检数	整体准确度/%
381	3	1	98.95
479	2	0	99.58

引入FCOS目标检测算法前同样的两组样本缺陷检测结果如表2所示。

表2 引入前检测结果

总检测数	漏检数	误检数	整体准确度/%
381	23	17	89.50
479	33	14	90.18

通过上述两组样本在引入FCOS目标检测算法前后的检测结果对比，引入后的漏检数和误检数显著降低，整体准确度也达到了98.5%以上。

3 结语

(1) 本文设计了基于U-net图像分割算法和FCOS目标检测算法的洗衣机外壳钣金孔洞缺陷检测系统，解决了应对不同生产线环境下对洗衣机外壳钣金孔洞缺陷检测的高精确度问题，使得本系统具有普适性和高精度的优点。

(2) 通过对检测结果的统计分析发现，对目

标检测影响较大的是因目标重叠而造成的包含或遮挡问题，设计时应引起注意，选择合适的算法来使目标重叠对检测结果影响最小化。

参考文献:

- [1] 王晓臣. 钣金件滚弯工艺成形及其缺陷检测[J]. 世界有色金属, 2020(1):143,145.
- [2] 李春, 李琳, 邹焱飏, 等. 基于机器视觉的钣金件缺陷在线检测算法[J]. 制造业自动化, 2016, 38(7): 56-59.
- [3] 单鹏, 李丹. 基于FCOS算法的盲人辅助障碍物检测[J]. 电子测试, 2021(21): 50-52, 15. DOI: 10.16520/j.cnki.1000-8519.2021.21.017.
- [4] 周涛, 董雅丽, 霍兵强, 等. U-Net网络医学图像分割应用综述[J]. 中国图象图形学报, 2021, 26(9): 2058-2077.
- [5] 范斌. 车辆辅助驾驶中基于FCOS算法的行人检测方法[J]. 电子测试, 2021(23): 49-51. DOI: 10.16520/j.cnki.1000-8519.2021.23.015.

作者简介:

邓秋菊(1987—),女,重庆铜梁人,硕士,副教授,研究方向为自然语言理解、机器视觉

通信作者:张振龙(2000—),男,河北保定人,本科,研究方向为机器视觉、图像处理,E-mail:583926577@qq.com

收稿日期:2022-06-05 修稿日期:2022-07-20

(下转第86页)

文章编号: 1007-1423(2022)19-0071-04

DOI: 10.3969/j.issn.1007-1423.2022.19.012

基于 AI 深度学习的病害识别方法研究

刘小红

(湖南信息学院软件工程系, 长沙 410151)

摘要: 快速准确地识别黄瓜病害类型是病害防治的前提, 针对黄瓜病害图像识别准确度不高的问题, 提出一种基于 AI 图像处理技术的卷积神经网络模型方法, 以提高病害图像识别率。首先, 利用 AI 图像处理技术对病害图像进行预处理; 其次, 选定卷积神经网络模型及参数, 利用样本数进行训练并加载到 Web 程序端; 最后, 在四种常见黄瓜病害图像构建的数据库上进行测试, 识别准确率均大于 90%。结果表明, 基于 AI 图像处理技术的卷积神经网络模型能够准确识别病害种类, 为田间开放环境下实现病害的快速识别提供了依据。

关键词: 病害识别; AI; 深度学习; 卷积神经网络

基金项目: 湖南省教育厅科学研究项目(20C1311)

0 引言

农作物由于易遭受病害的侵扰, 其产品质量和产量受到很大的影响, 为控制病害的扩散、减少经济损失和提高产品质量, 及时准确地识别病害具有重要意义^[1]。随着物联网、人工智能等信息技术的快速发展, 基于 AI 图像处理技术因具有操作简便、处理快、成本低等优点, 已逐渐成为农作物病害图像处理的主流技术。

深度学习技术使用学习程序从原始数据中自动学习特征, 具有突破人工局限的特点。近几年来, 作为深度学习核心的各种卷积神经网络被陆续提出, 并逐渐引入到植物病害检测与识别中^[2]。利用卷积神经网络(CNN)的深度学习技术对病害识别进行了大量探索研究, 如孟亮^[3]采用一种轻量级 CNN 农作物病害图像识别模型对辣椒、番茄和马铃薯的病害图像进行识别, 模型在训练集上的总识别准确率为 99.33%; 贾鹤鸣等^[4]在卷积神经网络中引入深度可分离卷积(DSC)和全局平均池化(GAP)方法, 对植物叶片进行病害识别, 病害识别精度达 99.42%;

廖经纬等^[5]在使用 Ostu 算法与基于 Sobel 算子将植物叶片的前景与背景有效分离后采用卷积神经网络对植物病害进行识别, 识别率达到 98.44%。虽然基于卷积神经网络(CNN)的深度学习技术在农作物病害识别方面已取得一定研究成果, 但由于受提取算法、环境条件和样本数量等因素的限制, 识别效果很难满足应用需求。

本研究以黄瓜四种病害(白粉病、斑点病、炭疽病、霜霉病)图像作为研究对象, 从病害识别需求出发, 采用 AI 图像处理技术结合卷积神经网络模型对病害图像进行病害识别, 并应用在 PC 端的病害识别系统上, 为农作物病害识别的精准性提供支持。

1 材料采集

利用高分辨率的相机采集黄瓜四种病斑(白粉病、斑点病、炭疽病、霜霉病)图像作为样本, 病斑图像各 100 张(总共 400 张)作为样本数, 要求图像清晰, 病斑区域和绿色区域的光照保持一致性, 黄瓜病斑图像样本如图 1 所示。

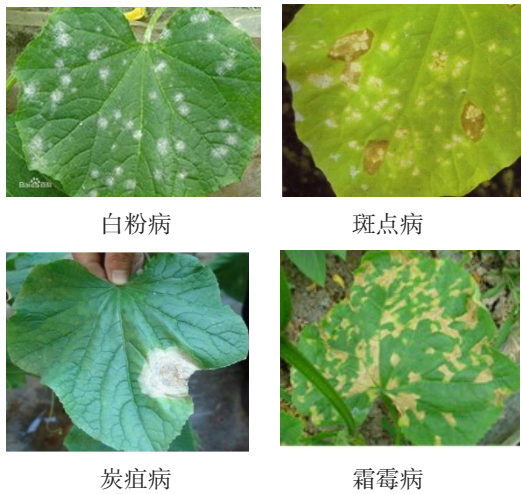


图1 黄瓜病斑图像样本

2 深度学习的识别模型构建

2.1 图像预处理

2.1.1 图像预处理过程

对病害图像先进行适当裁剪，去除多余部分，保留清晰的病斑区域；接着将原图RGB空间转成灰度化的目标图；最后利用最大类间方差法(即OTSU算法)将灰化图像进行分割，实现病斑区域的提取。将裁剪后的原图转换为灰度图像，采用转换公式： $Gray=0.299 \cdot R+0.587 \cdot G+0.114 \cdot B$ ，其中R，G，B分别为原图的红绿蓝分量，转换后得到灰度图，保留清晰的病斑特征，方便后期的病斑分割^[6]。OTSU算法是将灰度化图像归一化处理，利用阈值将图像分成背景和前景两部分，循环计算背景和前景之间的类内方差，求取类内方差为极大值时对应的阈值即为最佳分割阈值^[6]。

2.1.2 OTSU 阈值分割过程

(1) 计算灰度化病害图像的归一化直方图，

得到图像灰度级 $1 \sim M$ ，总像素数为 N ，则第 i 级灰度出现的概率为： $P_i = \frac{n_i}{N}$

(2) 计算图像灰度均值： $\mu = \sum_{i=1}^M i \cdot P_i$

(3) 计算直方图的零阶 $w[K]$ 和一阶矩阵 $u[K]$ ： $w_0 = w[k] = \sum_{i=1}^k P_i$ ， $\mu[k] = \sum_{i=1}^k i \cdot P_i$ ，对一阶矩作以下处理： $w_1 = 1 - w[k]$ ； $\mu_0 = \mu[k] / w[k]$ ； $\mu_1 = [\mu - \mu[k]] / [1 - w[k]]$

(4) 利用公式 $\sigma^2[k] = w_0(\mu - \mu_0)^2 + w_1(\mu - \mu_1)^2$ 找到最大的类间方差，对应此方差最大的灰度值即为要找的阈值， k 从 $1 \sim M$ 变化。

(5) 利用计算得到的阈值进行图像分割。图像预处理后的效果如图2所示。

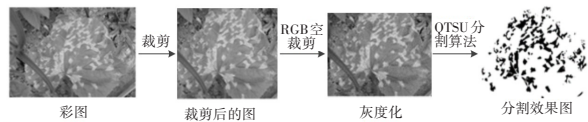


图2 图像预处理效果图

2.2 模型结构、训练和测试

病害图像在识别前先进行图像预处理，特别是利用OTSU算法将病害图像的绿色区域分割后只保留病斑区域，大大降低图像大小，利于后期的深度学习训练。本文根据黄瓜病害图像的特点，采用基于CNN深度学习网络结构—LeNet，构建具有多层的CNN网络识别模型，实现对黄瓜病害准确、快速识别。

2.2.1 模型结构

本文对LeNet网络模型加以改进，包含有输入层、卷积层、池化层、全连接层和输出层^[7]。病害识别模型网络如图3所示。

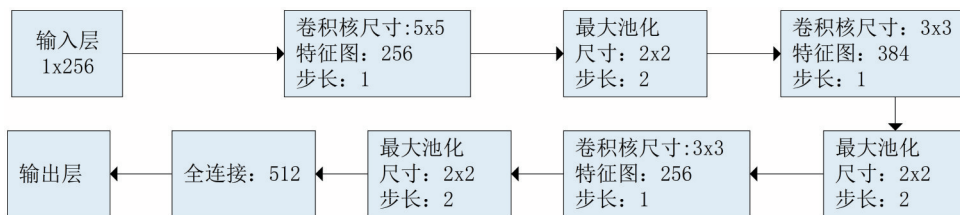


图3 CNN黄瓜病害识别模型网络结构图

另外，在前两层卷积中，加入归一化层，为防止图像在网络训练时数据参数的不稳定及提高模型特征提取性能，将输入的图像数据归一化到 $[-1, 1]$ 之间，并加快模型收敛速度进行处理，以提高模型的鲁棒性和泛化能力。

2.2.2 模型训练

模型训练过程具体如下：首先，建立数据集。采集相关黄瓜病害图像，进行图像处理，录入对应数据库并作为模型训练数据；其次，参照模型网络结构，设定网络模型参数，对训练数据进行训练。具体参数设定如表1所示。

表1 模型结构参数设置

结构参数	具体设置
池化方法	最大池化方法
激励函数	ReLU函数, Tanh函数、Sigmoid函数
算法优化	使用Adam算法、SGD算法；学习率=0.001、训练轮数=200、正则化系数=0.001。

对照相同条件下的训练效果，选择最优输出效果的模型并保存，执行器接收传入的程序。在反复迭代的训练过程中，不断优化网络模型中算法、函数等参数，得到最小收敛值作为最终输出^[8]。

2.2.3 模型测试

对训练后的模型进行测试，同样需要对图像进行预处理操作，再选择对应的模型进行识别^[1]。识别过程中对训练图片进行增强处理，加权求均值即为预测值。

2.3 系统设计

病害识别系统采用B/S架构模式及SSM框架技术，结合MySQL数据库，运用Java语言开发实现。①系统架构包括应用支持模型层、数据层、使用层。应用支持层主要提供AI深度学习模型及AI算法库、中间件、服务件等；数据层主要用于存储病害图像数据、病害知识库、病害特征数据、病害防治数据等；使用层主要包括病害诊断、病害查询及防治措施等服务。②功能结构主要包括病害诊断、查询、系统管理等，其中病害诊断包括图像上传、图像预处理、AI模型图像识别，功能结构如图4所示。③业务流程设计具体为：用户登录系统后进入

主界面，选择采集的病害图像，调用图像预处理的各项操作后，接着选择算法模型对图像进行识别，最后返回识别结果；还可直接查询病害信息、发病规律、防治措施等。

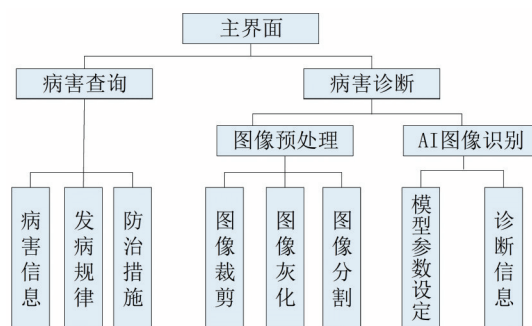


图4 黄瓜病害识别系统功能图

3 结果与分析

3.1 算法对模型性能影响分析

本研究的网络结构中算法设置使用Adam算法和SGD算法。SGD算法在训练过程中保持学习率不变。而Adam算法通过计算梯度的一阶矩估计和二阶矩估计来设计不同参数的自适应学习率^[8]。在保持网络结构层数不变及正则项系数和初始学习率均为0.001等相同条件下，对两种算法进行实验，实验数据显示：当迭代次数在1500次时，设置Adam算法后病害识别率为40.34%；设置SGD算法后病害识别率为40.26%；当迭代次数在4500次时，设置Adam算法后病害识别率为90.84%；设置SGD算法后病害识别率为81.67%。结果表明，随着迭代次数的增加，使用Adam算法模型的准确率比SGD算法更高。

3.2 模型性能对比分析

为验证基于AI的CNN网络模型的可行性，将其与传统的CNN模型进行对比实验，分别对黄瓜四种病害(白粉病、斑点病、炭疽病、霜霉病)图像进行识别。选择识别率、识别时间、损失率三个指标对模型性能进行评估^[8]。实验数据如下：传统CNN模型的平均识别率为45.68%，平均识别时间为7.8s，损失率为30.32%；基于AI图像处理的CNN网络模型平均

识别率为95.66%，平均识别时间为5.7 s，损失率为16.99%。结果表明，基于AI的CNN网络模型黄瓜病害识别率最高，识别时间最短，识别效果最好。

4 结语

本文在基于AI图像处理技术的CNN网络模型基础上，构建病害识别模型并应用到开发的病害识别系统上，对黄瓜的四种病害图像进行识别。实验数据表明，模型结构中采用Adam算法和SGD算法，在相同条件下前者对病害图像的识别率更高些。采用基于AI的CNN网络模型与传统的CNN模型相比，前者对病害图像识别具有更好的识别性能。本文只研究和分析了四种常见的黄瓜病害，具有局限性，后续为了进一步推广模型，将收集更多的病害图像，调整和优化模型。

参考文献：

- [1] 吴建伟,黄杰,熊晓菲,等.基于AI的桃树病害智能识别方法研究与应用[J].中国农业科技导报,2022,24(5):111-118.
- [2] 邵明月,张建华,冯全,等.深度学习在植物叶部病害检测与识别的研究进展[J].智慧农业(中英文),

2022,4(1):29-46.

- [3] 孟亮,郭小燕,杜佳举,等.一种轻量级CNN农作物病害图像识别模型[J].江苏农业学报,2021,37(5):1143-1150.
- [4] 贾鹤鸣,郎春博,姜子超.基于轻量级卷积神经网络的植物叶片病害识别方法[J].计算机应用,2021,41(6):1812-1819.
- [5] 廖经纬,蔡英,王语晨,等.基于卷积神经网络的植物病害识别技术[J].现代计算机,2018,(19):43-48,53.
- [6] 刘小红.基于预处理图像压缩的病害诊断应用研究[J].计算机技术与发展,2020,30(7):215-220.
- [7] 胡进坤,郭晓洁,李建平,等.基于深度学习的多模光纤通信系统的模式与模群识别[J].光学学报,2022,42(4):46-53.
- [8] 王国伟,刘嘉欣.基于卷积神经网络的玉米病害识别方法研究[J].中国农机化学报,2021,42(2):139-145.

作者简介：

通信作者:刘小红(1979—),女,湖南长沙人,硕士研究生,高级工程师,研究方向为计算机教育、移动互联网开发

收稿日期:2022-04-12 修稿日期:2022-06-28

Research on Disease Identification Method Based on AI Deep Learning

Liu Xiaohong

(Software Engineering Department of Hunan University of Information Technology, Changsha 410151)

Abstract: Rapid and accurate identification of cucumber disease types is the premise of disease control. In view of the low accuracy of cucumber disease image recognition, a convolution neural network model method based on AI image processing technology is proposed to improve the disease image recognition rate. Firstly, AI image processing technology is used to preprocess the disease image; Secondly, the convolutional neural network model and parameters are selected, trained by the number of samples and loaded into the web program; Finally, the accuracy of recognition is more than 90% in the database of four common cucumber disease images. The results show that the convolution neural network model based on AI image processing technology can accurately identify the types of diseases, which provides a basis for the rapid identification of diseases in the open field environment.

Keywords: disease identification; AI; deep learning; convolutional neural network

实践与经验

文章编号: 1007-1423(2022)19-0075-06

DOI: 10.3969/j.issn.1007-1423.2022.19.013

利用遗传算法进行智能家居设备识别特征选择

王靖翔, 刘翔宇, 尹钟舒, 刘丹丹, 谢镛镛, 杜彦辉

(中国人民公安大学信息安全学院, 北京 100038)

摘要: 现有的智能家居设备识别方式主要是基于流量特征进行的, 然而流量特征种类繁多, 如何确定强特征和有效的特征集成是智能家居设备识别的一个重要问题。文章提出了一种基于遗传算法的智能家居设备识别特征选择方法, 该方法从网络传输数据包中收集不同的协议特征和载荷特征, 并利用遗传算法进行特征选择确定最佳特征集。通过实验论证, 所选择的特征集在公开数据集 Aalto 与 CIC IOT 2022 上分别取得了 0.79 和 0.89 的准确率, 表明该特征集能够对大多数智能家居设备如 IP 摄像头、智能电灯、家用智能传感器等实现较为准确的分类识别。

关键词: 物联网; 设备识别; 遗传算法; 数据包信息; 特征选择

基金项目: 中国人民公安大学 2021 年基本科研业务费科技类项目(2021JKF105); 中国人民公安大学 2021 年高水平非在编机构建设项目(2021FZB13)

0 引言

近年来, 随着通信技术的快速革新和物联网技术的蓬勃发展, 越来越多的物联网设备连接在网络空间上, 研究表明到 2026 年, 网络空间的物联网设备数目将超过 800 亿^[1]。智能家居作为物联网技术的一大广泛应用场景, 已经给人们的生活带来巨大变化, 如 IP 摄像头、智能家电、家用智能传感器等设备接入互联网, 让人们可以切实享受到物联网技术带来的便利。

然而, 智能家居设备在给人们生活带来便利的同时, 其组成的复杂网络环境也带来了新的问题与挑战^[2]。许多智能家居设备往往存在一些公开或非公开的安全漏洞, 将用户置于长久的安全威胁之下, 例如不法攻击者可以利用某个特定类型摄像头的漏洞, 获得该类型设备的管理员权限, 窃取、挟持用户流量, 侵害用户的隐私等^[3-4]。由于特定的漏洞往往只威胁某一特定品牌型号的设备, 因此, 对智能家居设备的分类和识别, 成为学术界与工业界关注的热点问题。

现有的智能家居设备识别方式主要是基于

流量特征进行的, 然而流量特征种类繁多, 如何确定强特征和有效的特征集成为智能家居设备识别的一个重要问题。

基于以上原因, 本文依托网络流量提出一种基于遗传算法的物联网设备识别最优化特征选择方法, 该方法从网络数据包中的头部数据和载荷数据中提取各类流量特征, 构建特征池, 然后通过遗传算法进行特征选择, 确定最优特征集, 通过实验论证, 所选择的特征集能够对大多数物联网设备如 IP 摄像头、智能电灯、家用智能传感器等实现较为准确的分类识别。

1 面向包特征选择的遗传算法

遗传算法^[5]是较为经典的一种自然启发式算法, 其借鉴达尔文生物演变进化的理念, 让数据从初始随机混沌的状态, 根据一定的规则, 逐渐演变进化到合理解。在遗传算法中, 每个个体表示问题的一个解决方案, 个体之间通过交叉、变异等方式产生新个体, 组成新的种群, 然后由适应度函数对新种群中的每个个体进行评价, 优胜劣汰, 最后得到最佳个体, 即问题的最佳解决方案。其基本过程如图 1 所示。

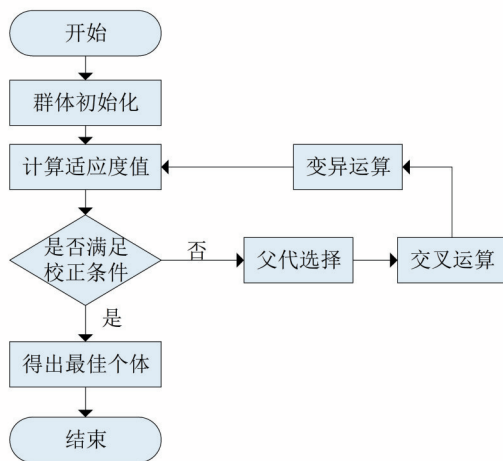


图1 遗传算法流程图

面向智能家居设备识别的特征选择其实是一个约束目标优化问题，相比于粒子群算法^[6]、模拟退火算法^[7]、蚁群算法^[8-9]等其他优化算法，遗传算法的优点是全局搜索能力较强，可以在较短的时间内达到全局最优解的90%以上，并且具有较高的鲁棒性。因此本文尝试用遗传算法解决面向智能家居设备识别的包特征选择问题。

一般说来，遗传算法GA可以表示为 $GA = (M, C, P, S, I, V, A, T)$ ，其中 M 表示种群大小； C 表示染色体编码方式； P 表示初始种群； S 、 I 、 V 分别代表选择算子、交叉算子、变异算子； A 和 T 分别表示适应度函数和终止条件。利用遗传算法对面向设备识别的包特征选择问题进行求解之前，要设定以上8个变量。

其中， M 表示种群中个体的数量， M 的取值越小，遗传算法所需的计算时间就越短，但算法的全局搜索能力越弱，陷入局部最优解的可能性越高。相反， M 的取值越大，算法得到全局最优解的可能性越大，但是其计算时间也会显著增加。目前，还没有可靠的理论计算方法来确定 M 的取值，其取值一般需要凭借个人经验和反复调试的结果来确定。在本文中，种群大小 M 的取值设为80。

染色体编码方式 C 是将问题的解决方案抽象到染色体上的具体方法，在本文中，利用One-hot编码方式进行染色体编码，即假设共有 N 个包特征，那么种群中个体的基因值个数为 N ，每个基因代表是否选择该特征，0代表不选

择，1代表选择。

初始种群 P 是算法的起始点，在本文中，随机产生与特征总量相同的0、1值来表示个体的基因值，得到初始种群。

选择算子 S 的作用是模仿自然选择的优胜劣汰，即保留种群中质量较高的个体，淘汰质量较低的个体，确保优良的基因能够保留到下一代种群中。在本文中，使用轮盘赌选择法^[10]，即每个个体的保留概率与其适应度值成正比，适应度越高，保留的概率就越大。除此之外，还采用精英保留策略^[11]，把种群在进化过程中出现的适应度最高的个体完整保留到下一代中，这样可以确保最优个体不会被其他算子破坏，从而加速算法的收敛。

交叉算子 I 的作用是随机交换两个个体的某些基因位，从而产生新个体。进行交叉算子的目的是将可能有益的基因组合在一起，产生更加优秀个体，增强遗传算法的全局搜索能力。在本文中，交叉算子的实现方式是在染色体中设置固定的交叉点，然后将两个染色体交叉点前后的部分基因进行交换，从而形成全新个体。

变异算子 V 的作用是对个体的一小部分基因值进行改动，目的同样是为了增强遗传算法的全局搜索能力。在本文中，设置基因的变异率为0.03，变异的方向有两个，即0变异为1、1变异为0。

适应度函数 A 是群落中个体的评价方式，适应度越高代表个体的质量越高。 A 的选取将直接影响算法的性能，通常是需结合实际问题来确定，本文中的适应度函数 A 为染色体所确定的候选特征向量的分类效果(决策树分类器的 $F1$ 值)。

遗传算法的终止条件 T 的设定方式主要有两种：设定种群迭代次数和达到某一阈值，设定迭代次数是指种群迭代到一定的次数后算法即停止；达到某一阈值指的是设定一个目标值，当进化的结果达到这个值时，即终止算法。在本研究中，由于没有明确的目标值，因此只能采用设定种群迭代次数的方法。

2 构建面向智能家居设备识别的特征池

现有的面向物联网设备识别的流量特征主

要分为三大类:协议特征、统计特征和载荷特征。

2.1 协议特征

即依靠协议的强特征对不同的物联网设备进行分类识别,例如Miettinen等^[12]在一个新的物联网设备引入到一个网络后收集前 n 个数据包,每个设备的指纹是基于从每个设备的前12个数据包中提取的23个特征创建的,其中19个特征是表示以下协议是否存在的二进制值,包括链路层、网络层、传输层、应用层的各种协议、数据负载、IP选项(Padding、RouterAlert)等。作者使用随机森林进行设备分类,27种设备中有17种识别准确率在95%以上,10种识别准确率在50%左右。

2.2 统计特征

即流量数据包的长度、数量,请求的时间间隔等特征,基于统计的方法对不同的物联网设备进行分类识别。例如Sivanathan等^[13]在六个月的时间内收集了28种物联网设备(如相机、灯、插头、运动传感器、电器和健康监测器)的流量数据来对物联网设备进行分类,使用的特征包括设备睡眠时间、端口号、活动周期、休眠时间、流量持续时间等。虽然通过阶段分类准确率达到99%,但所使用的特征鲁棒性不强,部分特征提取困难。

2.3 载荷特征

即深度挖掘载荷数据中的特征对不同的物联网设备进行分类识别。例如陈庆港等^[14]提出了一种基于深度可分离卷积的识别方法,基于网络数据流利用卷积层深度挖掘物联网设备流量的载荷特征。实验结果证明,即使在资源有限的情况下,该方法也能够有效识别设备类型。Bezawada等^[15]参考了IoT Sentine^[12]中的一部分特征集,再加上其他来自于载荷的特征,共选择了17个基于协议的特征和3个基于载荷的特征。设备的指纹来自于连续的5个数据包,共形成了100维的特征向量。研究人员使用了多种算法,包括决策树、kNN和梯度增强,取得了较高的准确率。

本文参考前人的研究,对所捕获的pcap格式网络数据包文件进行特征提取,所构造的特征池中包含94个特征,囊括了主要的协议特

征和载荷特征,如表1所示。

表1 特征池

类型	特征
LLC	LLC_dsap、LLC_ssap、LLC_ctrl
EAPOL	EAPOL_version、EAPOL_type、EAPOL_len
IP	IP_version、IP_len、IP_flags、IP_id、等15种
ICMP	ICMP_id、ICMP_type、ICMP_code等13种
TCP	TCP_seq、TCP_ack、TCP_ACK等20种
UDP	UDP_len、UDP_chksum等5种
DHCP	DHCP_options
BOOTP	BOOTP_op、BOOTP_flags等10种
DNS	DNS_length、DNS_id等6种
EtherType	Ether_Type
载荷熵	entropy
包长度	pkc_size
负载长度	payload_bytes
端口号类	port_class1、port_class2

其中87个是直接从数据包头部字段中提取的,这些特征来自LLC、EAPOL、IP、ICMP、TCP、UDP、DHCP、BOOTP和DNS等九种常见的网络协议。除此之外,本文所构造的特征池中还包含载荷熵、包长度、负载长度、端口号类等可能对物联网设备分类有用的特征,载荷熵提供了关于载荷特征的线索,端口号类特征提供了关于源端口号和目标端口号的摘要信息。

关于端口号的特征,不同的端口号所提供的服务是不同的,端口号的类别可按照知名端口(0-1023)、注册端口(1023-49151)、动态端口(49151-65535)划分,在知名端口中还有一些常用服务的端口,例如53-DNS、80-HTTP和123-NTP等。因此本文除了直接将源、目的端口号值作为特征之外,还对端口号类进行了两种不同的形式化描述,方法如表2所示。由于不确定哪一种端口号类的描述方法更为有效,本文将这两种方法都收集到特征池中。

表2 两种端口号类描述方法

port_class1(13类)	0-Reserved、53-DNS、67-BOOTP server、68-BOOTP client、80-HTTP、123-NTP、443-HTTPS、190-ssdp、5353-mDNS、49153-ANTLR、0:1023-知名端口、1023:49151-注册端口和49151:65535-动态端口
port_class2(1026类)	将每个知名端口(0-1023)单独划分一类,注册端口(1023-49151)、动态端口(49151-65535)

本文在特征提取阶段共提取总结了94个特征，因为不确定它们是否全部有助于设备的分类识别，所以将它们划到特征池中，为下一步的特征选择提供素材。

3 实验与分析

3.1 数据集

实验采用了Aalto与CIC IOT 2022两个公开数据集，首先在Aalto数据集上建立模型，进行实验最终的特征向量，再在CIC IOT 2022数据集上验证模型的有效性。由于Aalto与CIC IOT 2022两个数据集的采集使用的是完全不同的物联网设备，所以可以证明所建立的模型在物联网设备识别中具有较好的泛化性。

Aalto数据集包括智能照明灯、传感器、物联网摄像头和智能咖啡机等27种常见的物联网设备，研究人员收集每个设备初次引入到网络时所产生的流量数据，通过mac地址过滤后以pcap格式的文件进行存储，每个设备重复此类操作20次。

对于CIC IOT数据，本文使用了类似Aalto采集方式的部分数据集，其中包含40种不同于Aalto的物联网设备所产生的流量数据。

3.2 实验过程

(1) 划分训练集和测试集。在进行特征提取之前，为了排除干扰，首先将pcap文件分为测试集和训练集。Aalto数据集每个设备有20个pcap文件，本文将其分为两部分：16个作为训练数据和4个作为测试数据(80%:20%)。对于CIC IOT 2022数据集，也同样遵循这个比例进行分割。

(2) 特征提取。本文通过scapy包完成对包头部字段特征的提取，编写相应的代码完成对其余特征的提取。共形成包含94个特征的初始特征池。

(3) 遗传算法进行特征选择。设置遗传算法初始群落80，染色体随机，变异率0.03，适应度函数为候选特征向量的分类效果(决策树分类器的F1值)，种群迭代次数为100次，对训练集进行训练，训练过程如图2所示。

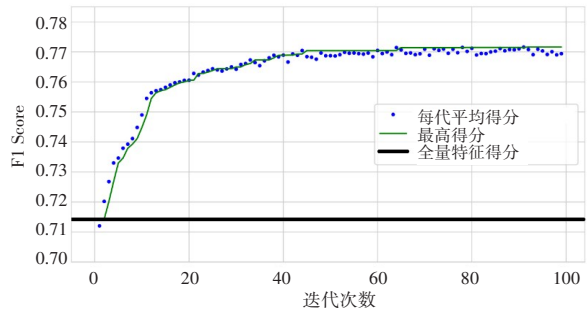


图2 遗传算法训练过程

可以看出，相较于使用全量特征进行分类后的F1得分(0.714)，使用遗传算法进行特征选择后分类效果有着较大的提升，其F1得分可以达到0.772。最终选择的最优子代，即最佳特征集共包含35个特征，如表3所示。

表3 遗传算法选择特征集

类型	特征
LLC	LLC_ssap
EAPOL	EAPOL_type、EAPOL_version
IP	IP_ttl、IP_ihl、IP_flags、IP_chksum、IP_options
ICMP	ICMP_chksum
TCP	TCP_dataofs、TCP_window、TCP_options、TCP_RST
UDP	UDP_len、UDP_chksum、UDP_dport
DHCP	DHCP_options
BOOTP	BOOTP_htype、BOOTP_hlen、BOOTP_secs、BOOTP_flags、BOOTP_sname、BOOTP_file、BOOTP_options
DNS	DNS_arcount、DNS_nscount、DNS_qdcount、DNS_rd、DNS_id
EtherType	Ether_Type
载荷熵	entropy
包长度	pkc_size
负载长度	payload_bytes
端口号	sport_bare、dport_bare

3.3 验证分析

本文采用所选择的35个特征进行分类试验，使用决策树分类器，在Aalto与CIC IOT 2022两个数据集的分类结果如表4所示。

对比同一数据集不同特征集的分类结果，Aalto数据集在使用优化特征集后的分类准确率由0.73提升到了0.79，CIC IOT 2022数据集的分

类准确率由0.77提升到了0.89,可以看出,使用优化后特征集的分类效果有着较为显著的提升,证明本文提出的基于遗传算法的智能家居设备识别最优化特征选择方法是切实有效的。

表4 实验结果

数据集	特征集	准确率	召回率	F1得分
Aalto	全量特征集	0.73	0.71	0.72
Aalto	优化特征集	0.79	0.77	0.78
CIC IOT 2022	全量特征集	0.77	0.79	0.76
CIC IOT 2022	优化特征集	0.89	0.85	0.86

对比同样使用优化特征集的不同数据集的分类效果,可以看出,CIC IOT 2022数据集的分类效果更好,由于在特征选择阶段只使用了Aalto数据集,而在CIC IOT 2022数据集上仍达到不错的分类效果,表明所选的特征集将很好地适用于其他物联网环境。

4 结语

本文提出了一种基于遗传算法的智能家居设备识别最优化特征选择方法。该方法从网络传输数据包中收集不同的协议特征和载荷特征,并通过遗传算法进行特征选择确定最佳特征集。通过实验论证,本文确定的特征集能够较好地适用于不同的物联网环境,能够对大多数物联网设备如IP摄像头、智能电灯、家用智能传感器等实现较为准确的分类识别。

参考文献:

- [1] TAHAEI H, AFIFI F, ASEMI A, et al. The rise of traffic classification in IoT networks: a survey [J]. *Journal of Network and Computer Applications*, 2020, 154: 102538.
- [2] REN J, DUBOIS D J, CHOFFNES D, et al. Information exposure from consumer IoT devices: a multi-dimensional, network-informed measurement approach [C] // *Proceedings of the Internet Measurement Conference*, 2019: 267-279.
- [3] SHAHID M R, BLANC G, ZHANG Z, et al. IoT devices recognition through network traffic analysis [C] // *2018 IEEE International Conference on Big Data (Big Data)*. IEEE, 2018: 5187-5192.
- [4] HUSSAIN F, HUSSAIN R, HASSAN S A, et al. Machine learning in IoT security: current solutions and future challenges [J]. *IEEE Communications Surveys & Tutorials*, 2020, 22(3): 1686-1721.
- [5] HOLLAND J H. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence* [M]. MIT Press, 1992.
- [6] KENNEDY J, EBERHART R. Particle swarm optimization [C] // *Proceedings of ICNN'95-International Conference on Neural Networks*. IEEE, 1995, 4: 1942-1948.
- [7] KIRKPATRICK S, GELATT JR C D, VECCHI M P. Optimization by simulated annealing [J]. *Science*, 1983, 220(4598): 671-680.
- [8] DORIGO M, GAMBARDELLA L M. Ant colony system: a cooperative learning approach to the traveling salesman problem [J]. *IEEE Transactions on Evolutionary Computation*, 1997, 1(1): 53-66.
- [9] COLORNI A, DORIGO M, MANIEZZO V. Distributed optimization by ant colonies [C] // *Proceedings of the First European Conference on Artificial Life*, 1991, 142: 134-142.
- [10] HAMMOOD D A, OMRAN S S, AL-KHALID A S. Using genetic algorithm to cryptanalyse a simple substitution cipher [C] // *11th Middle Eastern Simulation Multiconference, MESM 2010-1st GAMEON-ARABIA Conference, GAMEON-ARABIA 2010*, June. 2010: 93-97.
- [11] BHATEJA A, KUMAR S. Genetic algorithm with elitism for cryptanalysis of vigenere cipher [C] // *2014 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT)*. IEEE, 2014: 373-377.
- [12] MIETTINEN M, MARCHAL S, HAFEEZ I, et al. Iot sentinel: automated device-type identification for security enforcement in IoT [C] // *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*. IEEE, 2017: 2177-2184.
- [13] SIVANATHAN A, GHARAKHEILI H H, LOI F, et al. Classifying IoT devices in smart environments using network traffic characteristics [J]. *IEEE Transactions on Mobile Computing*, 2018, 18(8): 1745-1759.
- [14] 陈庆港, 杜彦辉, 韩奕, 等. 基于深度可分离卷积的物联网设备识别模型 [J]. *信息安全*, 2021, 21(9): 67-73.
- [15] BEZAWADA B, BACHANI M, PETERSON J, et al. Behavioral fingerprinting of IoT devices [C] // *Proceedings of the 2018 Workshop on Attacks and Solutions in Hardware Security*, 2018: 41-50.

作者简介:

王靖翔(1998—),男,山东枣庄人,硕士,研究方向为物联网安全

刘翔宇(1995—),男,安徽马鞍山人,博士,研究方向为信息安全

尹钟舒(1998—),男,山东济南人,硕士,研究方向为物联网安全

刘丹丹(1998—),女,天津人,硕士,研究方向为物联网安全

谢榕榕(1998—),女,浙江温州人,硕士,研究方向为物联网安全

通信作者:杜彦辉(1969—),男,北京人,博士,教授,研究方向为网络安全,E-mail:dyh6889@126.com

收稿日期:2022-06-06 修稿日期:2022-07-22

Genetic Algorithm is Used to Select the Features of Smart Home Devices

Wang Jingxiang, Liu Xiangyu, Yin Zhongshu, Liu Dandan, Xie Rongrong, Du Yanhui

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

Abstract: Existing smart home device identification methods are mainly based on traffic characteristics, but there are many kinds of traffic characteristics, how to determine strong features and effective feature integration for smart home device identification is an important issue. This paper proposes a feature selection method for smart home device recognition based on genetic algorithm. The method collects different protocol features and load features from network transmission data packets, and uses genetic algorithm to select the best feature set. Through experimental demonstration, the accuracy of the selected feature set is 0.79 and 0.89 on the public data set Aalto and CIC IOT 2022 respectively, indicating that the feature set can achieve relatively accurate classification and recognition of most smart home devices such as IP cameras, smart lights, smart sensors and so on.

Keywords: internet of things; device identification; genetic algorithm; packet information; feature selection

(上接第66页)

作者简介:

李首政(1999—),男,广东广州人,本科,研究方向为机器学习、模式识别

通信作者:王琪(1987—),女,河南南阳人,硕士,讲师,研究方向为图像处理、深度学习,E-mail:wangqi_

8766@163.com

王力(1983—),男,河南南阳人,博士,讲师,研究方向为大地测量

收稿日期:2022-05-23 修稿日期:2022-06-19

Research on Weibo Sentiment Analysis Method Based on Support Vector Machine

Li Shouzheng¹, Wang Qi¹, Wang Li²

(1.School of Information Engineering, Nanyang Institute of Technology, Nanyang 473000;

2. School of Civil Engineering, Nanyang Institute of Technology, Nanyang 473000)

Abstract: Weibo is currently one of the most popular social platforms in China. The sentiment analysis of Weibo text is helpful for its media value. Therefore, a weibo data sentiment analysis method based on support vector machine is proposed. First, the weibo data is crawled through Weibo Spider, and the weibo text data set is constructed by manual annotation. Then, combined with TF-IDF algorithm and traditional bag of words model, a new bag of words model based on keywords is proposed to obtain the text feature matrix to solve the problem of high sparsity and high redundancy of weibo text, and finally the Gaussian kernel support vector machine method is used to perform sentiment analysis on the crawled weibo data. Compared with the methods such as naive Bayes and decision trees, the experimental results show that the method in this paper obtain a higher accuracy rate, and has obvious advantages on small sample data.

Keywords: Weibo text; sentiment analysis; support vector machine; machine learning

文章编号: 1007-1423(2022)19-0081-06

DOI: 10.3969/j.issn.1007-1423.2022.19.014

轻量化卷积神经网络在药片检测分类中的应用

黄开坤, 徐 兴

(南华大学机械工程学院, 衡阳 421001)

摘要: 在药片残缺检测分类任务中, 为衡量分类模型的精度与模型大小, 保证模型精度的同时, 减少计算量和参数量。提出一种改进的 GhostNet 轻量化神经网络, 保证在加工过程中能够有效地检测出残缺药片。构建数据集并增强处理, 压缩网络模型, 将注意力机制升级为 ECA 模块。实验结果显示, 改进后的网络模型能够达到 98.85% 的分类精度, 模型参数量仅为 0.635×10^6 。与其它轻量化网络进行了对比分析, 取得了更高的分类精度, 更少的参数量和计算量, 更优的分类性能。

关键词: 图像处理; GhostNet; 注意力机制; 轻量化网络

0 引言

药品质量检测是医药领域十分重要且不可缺少的环节, 在保障人民健康、社会安全稳定方面起着重要作用。在医药生产领域, 人们对药片的品质追求不断提高, 药片生产的各个环节都受到严格的管制。药片生产过程中难免出现缺陷次品, 严重影响产品质量, 在药片包装之前进行有效的剔除, 是保证药片品质的关键步骤。

传统人工检测方法效率低, 不确定因素影响大。随着技术的不断提升, 机器视觉检测技术应用于很多行业领域。药物质量检测技术一般包括人工检测、激光检测、机器视觉检测等^[1], 其中基于机器视觉的医疗检测包括经典传统检测^[2-3]和深度学习检测^[4-5]。卷积神经网络(CNN)是在图像处理领域优异的神经网络^[6], 长期以来是图像识别领域的核心算法之一, 比传统图像处理算法具有更强大的特征学习能力。在机器视觉缺陷检测应用领域, 越来越多的检测方法从经典的图像处理和机器学习转到深度学习^[7]。

本文提出改进后的 GhostNet 网络模型, 对网络模型进行了进一步的压缩, 保证网络在取得较高的分类精度的同时达到更少的参数量和计算量, 更少的内存占用与计算消耗。在网络模

型中引入 ECA 注意力机制模块, 并分别与原始网络、主流轻量化神经网络进行性能对比分析。

1 GhostNet 网络

GhostNet 网络是华为 2020 年提出的分类网络^[8]。该研究提出了一个新的 Ghost 模块, 通过廉价操作来生成更多的特征图, 在压缩网络的同时确保分类准确性, 降低模型参数, 利用 Ghost 模块构建 Ghost bottleneck 结构, 利用 Ghost 瓶颈结构叠加主干神经网络结构。

1.1 Ghost 卷积模块

GhostNet 是以 MobileNet 系列为基础的改进的网络结构, 从图像特征冗余的角度出发, 避免了大量计算量和参数量去生成这些相似特征图, 并用一系列简单的线性操作来生成特征图。操作过程如图 1 所示。

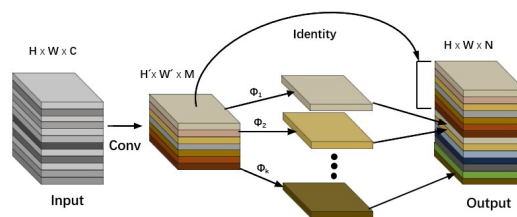


图 1 Ghost Module 模块

Ghost模块首先通过普通卷积对输入特征图进行特征浓缩，将输入特征图 $X \in \mathbf{R}^{C \times H \times W}$ 压缩为 M 通道的特征图 $Y' \in \mathbf{R}^{H' \times W' \times M}$ ，其中 C 是输入通道数，卷积运算为

$$Y' = X * f' + b \tag{1}$$

公式 $*$ 为卷积运算， b 为偏置项， $f' \in \mathbf{R}^{c \times k \times k \times m}$ 为卷积核大小为 k 的卷积核， $M \leq N$ ，保证输出通道数小于输入通道数，对原始特征进行必要的特征浓缩。此过程的参数量 P_1 和计算量 F_1 分别为

$$P_1 = C \times k \times k \times M \tag{2}$$

$$F_1 = C \times k \times k \times M \times H' \times W' \tag{3}$$

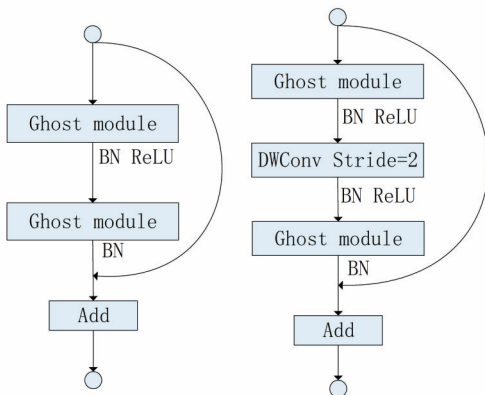
保证最后获得 N 个特征图，对 Y' 特征图进行简单的线性运算，来生成 S 个 Ghost 特征图。线性操作为图 1 中的 $\Phi_1, \Phi_2, \dots, \Phi_k$ ，线性运算为逐层的深度可分离卷积，卷积核大小 k' ，Identity 为恒等映射。 M 次线性运算能够得到 $M(S - 1)$ 的输出通道数，参数量 P_2 和计算量 F_2 为

$$P_2 = k' \times k' \times (S - 1) \times M \tag{4}$$

$$F_2 = 1 \times k' \times k' \times (S - 1) \times M \times H' \times W' \tag{5}$$

总共得到的输出通道数为 $N = M \times S$ ，最终输出即 N 个特征图。

Ghost Bottleneck 是由 Ghost Module 组成的瓶颈结构，参考 ResNet 中的基本残差块^[9]，构成了多个卷积层和残差结构，并用 Ghost 模块代替普通卷积模块。Ghost 瓶颈分为两部分，分别是主干部分和残差边部分，结构如图 2 所示。



(a) Stride=1 bottleneck (b) Stride=2 bottleneck

图 2 Ghost 瓶颈结构

图 2(a) 步长 stride=1，包含两个 Ghost 模块，首先增加通道数，接下来减少网络通道数并连

接到残差路径。图 2(b) 步长 stride=2，在两个 Ghost 模块之间加入深度可分离卷积进行特征层的宽高压缩。

1.2 整体网络模块分析

GhostNet 网络主干由 Ghost 瓶颈模块叠加而成，借鉴了 MobileNetV3^[10] 的基本结构的优势。网络第一层为 16 通道的卷积层，然后由多个瓶颈模块构成，依据输入特征图的大小分为不同的阶段，不同阶段的末位瓶颈模块步长为 2，其它步长都为 1。最后经过全局平均池化层、卷积层和全连接层对输出进行分类。

在 G-bneck 瓶颈模块加入了通道注意力机制 (Squeeze-and-Excitation, SE)^[11] 模块。SE 机制模块网络架构图如图 3 所示。首先对输入进来的特征层进行全局平均池化，接着通过两次全连接，再进行一次 Sigmoid 将值固定到 0~1 之间，获得输入特征每一个通道的权值，最后将权值乘上原始输入特征层，得到新的加权后的特征图。

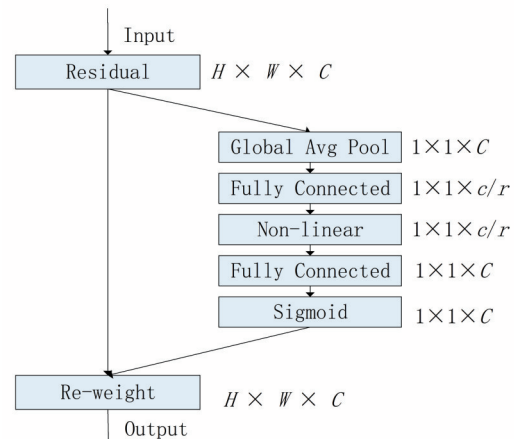


图 3 SE 注意力机制网络架构

2 改进 GhostNet 网络模型设计

2.1 网络模型的优化

GhostNet 网络模型在公开数据集上进行的测试，类别数量达 1000 类，而此药片完整性分类任务数据集为类别数量较小的 6 类，药片特征并非复杂多样的图案特征，减少输出层的分类数量 classes，减半整体网络的通道数。网络模

型复杂会造成过拟合风险, 去掉网络结构中的重复操作层, 进一步减少网络的参数量。添加 Dropout 层^[12], 在神经网络训练过程中, 当训练数据量较小, 模型复杂度较高时, 模型在训练集上的准确率较高, 在测试集上的准确率相对较低, 出现了过拟合, 为了提高模型的泛化能力, 在全连接前添加 Dropout 层, 丢弃一些神经元, 减少神经元之间的相互作用, 提高网络模型分类精度。

2.2 引入 ECANet 模块

ECANet 是一个非常轻量级的即插即用模块^[13], 可以提高神经网络架构的性能, 结构如图 4 所示。ECA 模块可以被看作是 SE 模块的改进版, SENet 采用的降维操作会对通道注意力机制产生负面作用。卷积具有良好的跨通道信息获取能力, 轻量 ECA 模块是将 SE 模块的全连接层替换成卷积核大小为 k 的 1D 卷积进行学习, 避免了降维, 有效地实现了跨通道交互。卷积核大小可由通道维数的非线性映射自适应确定, 如公式(6)所示。

$$C = \mathcal{O}(k) = 2^{(\gamma * k - b)} \quad (6)$$

在通道维数 C 给出时, 卷积核大小 k 便可由公式(7)得到:

$$k = \psi(C) = \left\lfloor \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right\rfloor_{\text{odd}} \quad (7)$$

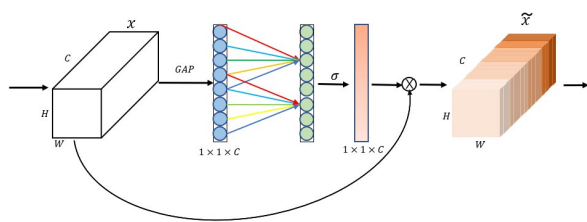


图 4 ECA 模块结构图

2.3 优化后的网络模型

优化后的网络结构如表 1 所示, 首先通过第一层 Conv2d3x3, 将输入特征转换为包含 8 个卷积核的标准卷积层, 压缩了网络, 降低了参数量。中间为一系列的瓶颈层。接着通过全局平均池化层和卷积层将特征图转化为 100 维特征向量, 并进行最终分类。将原网络的 SENet 替换为

轻量化 ECANet, 保留原网络主要特征提取 G-bneck 模块, 在全连接层 FC 网络之前添加 Dropout 层, 提升模型分类准确性。表中 Exp 表示增大比率, Out 表示输出通道数, ECA 为注意力机制模块, Str 为步长, cls 为分类数量。

表 1 改进后的 GhostNet 网络结构

Input	Operator	Exp	Out	ECA	Str
224 ² x 3	Conv2d3x3	—	8	—	2
112 ² x 8	G-bneck	8	8	—	1
112 ² x 8	G-bneck	24	12	—	2
56 ² x 12	G-bneck	36	12	—	1
56 ² x 12	G-bneck	72	20	1	2
28 ² x 20	G-bneck	60	20	1	1
28 ² x 20	G-bneck	120	40	—	2
14 ² x 40	G-bneck	100	40	—	1
14 ² x 40	G-bneck	92	40	—	1
14 ² x 40	G-bneck	240	56	1	1
14 ² x 56	G-bneck	336	56	1	1
14 ² x 56	G-bneck	336	80	1	2
7 ² x 80	G-bneck	480	80	—	1
7 ² x 80	G-bneck	480	80	1	1
7 ² x 80	Conv2d1x1	—	480	—	1
7 ² x 240	AvgPool7x7	—	—	—	—
1 ² x 240	Conv2d1x1	—	100	—	1
1 ² x 100	FC	—	cls	—	1

3 实验方案与结果分析

3.1 实验环境配置

本次实验在 Windows10 环境下, 采用 python 语言, 在 pycharm2020.2 版本编程环境下进行, 在 Anaconda 中创建 python 虚拟环境, python 版本为 3.8; 利用 Keras、Pytorch 深度学习架构, 使用 cuda11.4 提升模型训练和推理的速度。实验设备为: GPU 为 RTX 3090, 显存 24 G; CPU 为 Intel (R) Xeon(R) Silver 4210R CPU @ 2.40 GHz。

3.2 数据集及模型训练

实验数据集是由工业相机 MV-CE120-10UC 拍摄所得, 根据生产厂商提供的散装药片和相关资料制作残缺药片, 药片总共分为 6 种类别, 包括污染、破损、完好、粗糙、划痕、斑点。

药片图像总共为 3360 张，每个类别为 560 张，部分药片样本如图 5 所示。

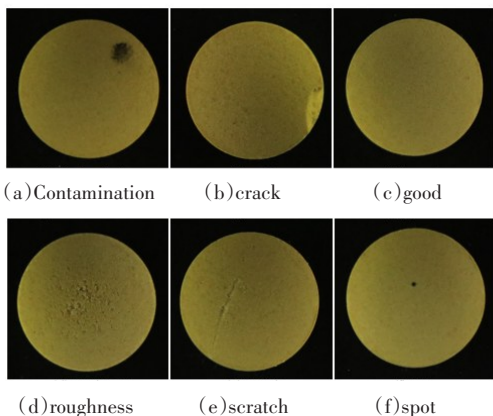


图 5 部分药片图样示例

为提高网络模型的泛化能力，避免出现拟合风险，对数据集进行增强操作，增加数据集数量^[14]。增加后的图像共 12000 张，将打乱后的数据集分为 70% 的训练集、15% 的验证集和 15% 的测试集。原始特征图片的大小为 800 × 800 × 3。送入网络图片大小为 224 × 224 × 3，初始学习率为 0.001，Batch Size 大小为 32，使用 Adam 优化器训练，迭代次数(epoch)为 50 代。

3.3 评价指标

从图像分类评价指标方法出发，对模型的分类精度、运算量、参数量等进行评估。

模型分类精度(Precision)，表示预测出来准确结果占有所有预测结果的准确率，分类精度越高说明模型预测能力越强。

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

其中，TP 表示正样例中分类为正样例的数量，FP 表示标签是负样例中分类为正样例的数目。模型运算量常用 FLOPs 来表示，浮点运算量即为计算量，是用来表示模型的计算复杂度。

模型参数量 parameters，是用来表示模型内部总的参数量，衡量模型所消耗内存的大小，参数越大所占用的内存就越大。

3.4 模型评估

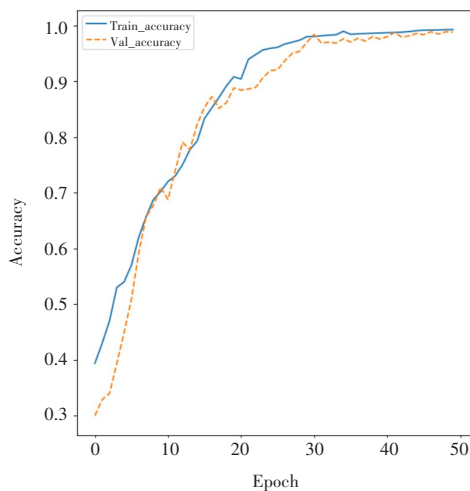
为了验证改进后的网络模型性能的优异，并与原始网络模型进行性能比较，从药片分类精度、模型参数量、模型运算量方面进行对比，

如表 2 所示。

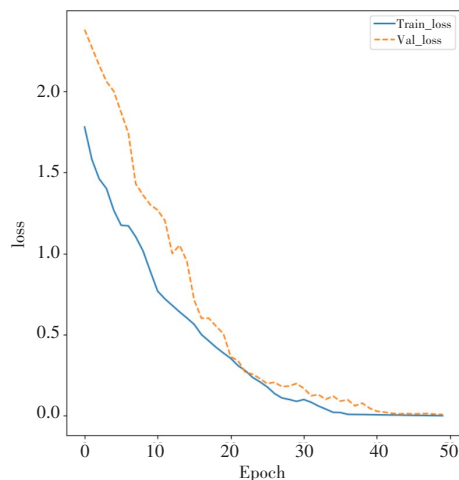
表 2 网络评价指标对比

Model	Flops/M	Params/MB	Acc/%
原始 GhostNet	147.66	10.94	98.89
改进后的 GhostNet	40.15	2.42	98.85

由实验结果可知，原始网络模型参数约为改进后 GhostNet 模型的 4.5 倍，计算量约为改进后的 3 倍，改进后的网络模型精度与原始网络相差仅为 0.04%，在大量减少网络模型参数的前提下，保持较高的模型精度。优化后的网络模型在训练过程中的精度和损失曲线如图 6 所示。



(a) Training and validation accuracy



(b) Training and validation loss

图 6 改进后的网络训练曲线

从图 6 能够看到网络模型共训练 50 代，随

着训练次数的增加收敛速度更快,在迭代40代后曲线上升趋势渐渐平缓,整个过程中未出现过拟合。结果表明,改进后的GhostNet网络模型具有较高的分类精度。

3.5 不同网络结构对比实验

在相同实验条件下,为了验证改进的网络模型的性能,与主流轻量化神经网络 MobileNetV、ShuffleNetV2^[15]和 EfficientNet^[16]进行对比实验,结果如表3所示。

表3 不同网络评价指标对比

Model	Weights/M	Accuracy/%
EfficientNetB0	12.6	97.46
ShuffleNetV2	2.5	97.95
MobileNetV3	6.3	98.36
改进的 GhostNet	1.8	98.85

由实验结果得知,改进后的GhostNet网络模型的精度高于其它网络,权重模型大小仅为1.8 MB,明显低于其它轻量化网络模型,运算速率也相应提高,计算量和参数量明显降低,说明改进网络模型具有更高的分类精度,更少的参数量和计算量。

4 结语

在药片分类实际应用中,在提高分类精度的同时,保证更低的网络存储空间占用和功能消耗,使得轻量化神经网络模型更容易在移动端或嵌入式设备上存储和计算。实验过程中,采用药片分类数据集对模型进行训练和测试,对数据集增强处理,压缩网络模型通道数,优化网络结构,降低参数量和计算量,引入ECA注意力机制。实验结果表明,改进后的网络模型具有更低的参数量与运算量,模型权重文件仅为1.8MB,能够取得的分类精度为98.85%,参数量和计算量低于原始网络;网络模型分类精度高于其它轻量化神经网络,取得较优的分类性能。

参考文献:

[1] 张辉,易俊飞,王耀南,等.医药质量检测关键技术及其应用综述[J].仪器仪表学报,2020,41(3):1-17.

[2] 高体红,周凯汀,郑力新.药片视觉系统中图像处理单元设计与验证[J].计算机仿真,2015,32(1):420-424.

[3] 疏义桂.基于机器视觉的铝塑泡罩包装药品缺陷检测[D].武汉:华中科技大学,2013.

[4] LIU R, GU Q, WANG X, et al. Region-convolutional neural network for detecting capsule surface defects[J]. Boletin Tecnico/Technical Bulletin,2017,55(3):92-100.

[5] 王振亚,赵继红,王艳鹏,等.基于改进EfficientNet网络的药片检测分类的研究[J].现代计算机,2021,27(28):27-32.

[6] RAZAVIAN A S, AZIZPOUR H, SULLIVAN J, et al. CNN features off-the-shelf: an astounding baseline for recognition [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. Columbus, OH, USA: IEEE Computer Society, 2014:806-813.

[7] 肖书浩,吴蕾,何为,等.深度学习在表面质量检测方面的应用[J].机械设计与制造,2020(1):288-292.

[8] HAN K, WANG Y, TIAN Q, et al. GhostNet: more features from cheap operations[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA, IEEE Computer Society, 2020:1050-1069.

[9] HE K, ZHAN 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, NV, USA, IEEE Computer Society, 2016:770-778.

[10] 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): IEEE, 2019:1314-1324.

[11] IANDOLSA F N, HAN S, MOSKEWICZ M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size[J]. Computer vision and pattern recognition, 2016, 20(3) 1-22.

[12] SRIVASTAVA N, HINTON G, KRIZHEVSKY A, et al. Dropout: a simple way to prevent neural networks from overfitting [J]. The Journal of Machine Learning Research, 2014, 15(1): 1929-1958.

[13] WANG Q, WU B, ZHU P, et al. ECA-Net: efficient channel attention for deep convolutional neural networks [C] //Proceedings of the IEEE Conference

- on Computer Vision and Pattern Recognition. Seattle, WA, USA, IEEE, 2020: 13–19.
- [14] 高友文, 周本君, 胡晓飞. 基于数据增强的卷积神经网络图像识别研究[J]. 计算机技术与发展, 2018, 28(8): 62–65.
- [15] MA N N, ZHANG X Y, ZHENG H T, et al. ShuffleNet V2: practical guidelines for efficient CNN architecture design [C]//Proceedings of the European Conference on Computer Vision (ECCV). Munich, Germany, 2018: 116–131.
- [16] TAN M X, LE Q. Efficientnet: rethinking model scal-

ing for convolutional neural networks [C]//Proceedings of the 36th International Conference on Machine Learning, PMLR, 2019: 6105–6114.

作者简介:

通信作者: 黄开坤(1996—), 男, 河南信阳人, 硕士研究生, 研究方向为计算机视觉与图像处理等, E-mail: 1769818459@qq.com

徐兴(1996—), 男, 安徽宣城人, 硕士研究生, 研究方向 ROS 移动机器人

收稿日期: 2022-06-09 修稿日期: 2022-08-20

Application of Lightweight Convolution Neural Network in Tablet Detection Classification

Huang Kaikun, Xu Xing

(School of Mechanical Engineering, University of South China, Hengyang 421001)

Abstract: In the task of tablet defect detection and classification, in order to measure the accuracy and size of the classification model, ensure the accuracy of the model, and reduce the amount of calculation and parameters. In this paper, an improved GhostNet lightweight neural network is proposed to ensure that incomplete tablets can be effectively detected in the processing process. Build data sets and enhance processing, compress network models, add dropout layer to prevent over fitting, and upgrade attention mechanism to ECA module. The experimental results show that the improved network model can achieve 98.85% classification accuracy, and the model parameter is only 0.635×10^6 . Compared with other lightweight networks, it achieves higher classification accuracy, less parameters and computation, and better classification performance.

Keywords: image processing; GhostNet; attention mechanism; lightweight network

(上接第 70 页)

Detection System for Sheet Metal Hole Defect of Washing Machine Shell Based on U-net Algorithm and Fcos Target Detection Algorithm

Deng Qiuju, Zhang Zhenlong

(Chongqing College of Mobile Communication, Hechuan 401520)

Abstract: The hole detection in the punching process of shell sheet metal is an indispensable link in the automatic production line of washing machine. A detection system based on u-net image segmentation algorithm and FCOS target detection algorithm is developed and designed to solve the problem of hole detection of washing machine shell sheet metal. According to the requirements of image segmentation of sheet metal holes, the u-net algorithm is introduced to realize it. According to the detection characteristics of sheet metal holes in the shell of washing machine, the FCOS algorithm is used to solve the problem of hole target detection. The final detection data show that the detection accuracy of the system for the sheet metal hole defects of the washing machine shell on the automatic washing machine production line meets the industry standard, has good detection accuracy, and can fully meet the detection requirements of normal production.

Keywords: shell sheet metal hole; object detection; image segmentation; defect detection

文章编号: 1007-1423(2022)19-0087-05

DOI: 10.3969/j.issn.1007-1423.2022.19.015

轮播图导航制作在 Web 前端中的实践与应用

赵素萍¹, 杨璐², 康晶晶¹

(1. 晋中信息学院大数据学院, 晋中 030800; 2. 晋中信息学院信息工程学院, 晋中 030800)

摘要: 无论是在电脑端还是手机端的网页, 无论是手机 APP 还是各种 APP 内嵌的小程序, 主题 banner 部分都有轮播图, 轮播图是一种观赏性强的信息展示方式。轮播图主要是实现周期性切换图片的效果, 其中轮播图中的焦点导航可以提示和控制当前播放的图片。同时还可让用户通过箭头实现上一张和下一张图片的控制效果。本文使用 HTML5、CSS3 和 JavaScript 技术来实现。是综合性较强的案例, 希望为网站设计人员提供一种思路。

关键词: 轮播图导航; 焦点; 箭头

0 引言

网站的首页, 除了有常规的导航栏, 通常还有一个轮播图导航。轮播图也叫焦点图, 无论是购物网站、政府网站、教育网站, 还是新闻网站, 无论是在电脑端还是手机端的网页, 无论是手机 APP 还是各种 APP 内嵌的小程序, 主题 banner 部分都可见到轮播图。轮播图是一种观赏性强的信息展示方式。

在默认情况下, 轮播图是循环向右轮播, 轮播图的两侧各有一个箭头, 用户可以通过单击箭头切换上一张或下一张图片。轮播图图片也是一种导航, 单击图片会跳转到相应的网页。轮播图的下方区域有焦点, 用户悬浮焦点可实现图片切换。

本着一切以用户为中心的原则, 当用户悬浮焦点或单击箭头后, 暂停周期定时器, 同时当用户的鼠标离开焦点和箭头区域后, 触发周期定时器让其自动切换图片。本案例主要使用 HTML5 结构, CSS3 样式和 JavaScript 脚本来实现轮播图导航。这些技术是前端工程师必须要掌握的基本技能。

1 HTML 内容

网页中主要显示的内容为一张超链接图片,

4 个无序列表项实现用户想看随机一张轮播图的效果, 两个箭头实现上一张和下一张的效果, 这里需要注意的是, 大于号和小于号在 HTML 中属于标签的一部分, 在正文中不允许使用, 需要使用特殊字符对应的代码才能实现, 其中“<”表示小于号, “>”表示大于号, 主要代码如下:

```
<div class="banner">
  <a href="" id="imga"><img src="" id="bnimg"></a>
  <ul>
    <li class="dot" onmouseover="overchange(0)"
      onmouseout="settime(0)">1</li></ul>
  </ul>
  <div class="lt" onclick="backimg()" onmouseover="stopchange()"
    onmouseout="startchange()">&lt;</div>
  <div class="rt" onclick="nextimg()" 其余同上>
  &gt;</div>
</div>
```

2 样式

样式主要通过相对定位、绝对定位和偏移将 dot 导航和箭头导航放到合适的位置。通过过渡和 hover 伪类实现导航悬浮特效。

2.1 dot 导航

通过绝对定位的定位设置，将 dot 导航放置在合适的位置，需要注意的是，设置绝对定位的元素，需要将父元素设置为已经定位的一种，主要包括绝对定位 absolute、相对定位 relative 和固定定位 fixed。由于绝对定位会使其失去在源文档流的位置，所以如果要保留父元素在源文档流的位置，父元素不能设置为绝对定位。由于固定定位始终是相对于浏览器进行定位的，而这种方式使用的情况少之又少，所以常用的方式是将父元素定位为 relative 相对定位，这样父元素既是已经定位的元素，同时也不会失去源文档流中的位置。

通过偏移进行精确定位，偏移包括四种，分别是 left、right、top 和 bottom，一般情况下只需要使用两个属性就可以达到我们想要的效果，如需将元素设置在父元素的左下角，那么只需要设置 left 和 bottom 两个属性即可。如果用户同时设置了 left 和 right，浏览器会以 left 为准。同理，如果用户同时设置了 top 和 bottom，浏览器会以 top 为准。主要样式代码如下：

```
.banner{position: relative; /*父元素*/
.banner ul {position: absolute; bottom: 10px; left: 800px; /*子元素*/
```

设置每个 dot 导航的背景为圆形、灰色、透明度为 0.5 的样式，由于列表项为块状元素，特点是独占一行，为了让其共享一行，先需要将其设置为行内块式，行内块式的特定是可以共享一行，同时还可以设置其宽度、高度。一般情况下，横向的导航都是这样设置的。

背景颜色是通过 rgba 函数来实现的，rgb 是三原色，a 是透明度，通过一个函数即可定义元素的背景颜色和透明度。为了增加网站元素的立体效果，一般元素都会加圆角和阴影。初学者可以多次尝试选择合适的圆角半径、阴影大小。主要样式代码如下：

```
.dot{ display: inline-block;
background-color: rgba(11,11,11,0.5);
border-radius: 50%;
}
```

2.2 箭头导航

盒子模型的整体宽度从外向内分别包括外边距、边框、内填充、宽度。具体计算如图 1 所示。

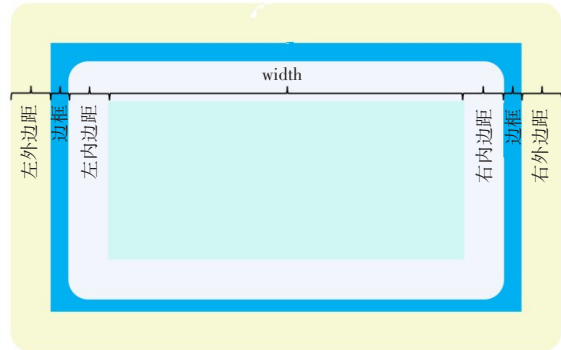


图 1 盒子模型总宽度

为了将箭头放到合适的位置，内填充设置为离 top 有 100 px 的位置，需要注意的是，banner 的高度为 300，内填充为 100 时，内容的高度只能 200。元素的整体高度如公式 1 所示：

$$\text{元素高度} = \text{上/下 margin} + \text{上/下 border} + \text{上/下 padding} + \text{height} \quad (1)$$

其中左右导航没有设置外边距 margin，没有设置外边框 border，只设置了 padding-top 和 height。所以要想达到 300 的高度，只需要使 padding-top 加 height 等于 300 即可。这是初学者较难掌握的知识点，需要多练习达到熟练运用的效果。

先设置看不到的左右两边，当鼠标悬浮 banner 时，左右导航定位在 banner 的左右两边可以看到的位置，主要通过绝对定位和偏移实现，为了增加动画效果，使用过渡实现，背景颜色透明度为 0.2，这里只展示右边的箭头，学习者可根据右箭头自主添加左箭头的样式。主要代码如下：

```
.rt, .lt{
background-color: rgba(11,11,11,0.2);
padding-top: 100px; height: 200px;
transition: all 0.5s; /*过渡*/
}
.rt {right: -1000px; /*初始放在离右边很远的位置*/
.banner: hover .rt {right: 0; /*鼠标悬浮 banner 时,箭头放在 banner 右边*/
```

3 脚本

脚本的引入方式主要包括三种,行内式、内嵌式和外链式。这三种引入方式各有特点,分别应用于不同的场合。

行内式只用于控制当前元素,如滚动标签marquee,其鼠标悬浮暂停滚动,鼠标离开继续滚动的效果只需要使用行内式即可。由于滚动元素一般在网页中的使用情况并不多,所以常使用引入式。

嵌入式的脚本一般放在head中和body中两种。由于加载的顺序是先head,再body。所以Head内的脚本一般为暂时不执行的方法和全局变量。Body内的脚本一般为暂时执行的JavaScript脚本。该方法一般用于初学者初次接触脚本,以及教授基础知识的讲师。在实际的网站中推荐使用外链式。

外链式通过script标签的src属性来引入“*.js”脚本文件。该方法利于后期维护,同时同一个脚本文件可用于多个网页文件,所以减轻了文件体积,可加快页面的加载速度。

本文主要使用的是行内式和外链式。轮播图主要通过脚本实现以下几种效果。

3.1 图片自动切换

自动切换需要用到周期性定时器的函数setInterval(周期性触发时间,触发周期)。通过该函数,让浏览器自动每隔一定时间触发修改图片的函数即可。为了修改图片,需要先获取到网页中的元素,获取元素需要使用DOM文档对象模型,获取元素的方法主要有四种,主要包括getElementById、getElementsByName、getElementsByTagName、getElementsByClassName。这里初学者需要注意,由于ID具有唯一性,所以通过ID获取元素的方法是getElementById,其中Element表示一个元素,可以直接用,而其他三种情况没有唯一性,即使只有一个,返回的元素类型仍然是数组,也需要通过下标[0]来取出元素。这是初学者经常会犯的一个错误。本案例中的4个dot导航就需要通过类名获得,否则通过ID获取的方法会有一定的代码冗余,也会降低浏览器加载的速度。本案例引入的元素和定义的变量主要如下:

```
var sj=setInterval(changeimg,3000);//3s 切换照片的
周期定时器
var imgshow=document.getElementById('bning');
var imga=document.getElementById("imga");//通过
ID 获取超链接和图片
var dots=document.getElementsByClassName('dot');
//获取所有 DOT
var lt=document.getElementsByClassName("lt")[0];
//获取箭头元素
```

为了在同一个图片元素中显示不同的图片,只需要修改图片元素路径——src属性即可。此时多张图片的路径主要是通过数组来预先设置,然后再将其循环设置为图片的src属性。超链接的链接目标同上。主要代码如下:

```
var imgspath=[ "images/b1.jpg", ..... ]; //4 张图片
路径
var hrefs=[ "http://www.baidu.com", ..... ]; //4 个
超链接网址
```

为了让用户了解到当前播放的是第几张图片,通过dot导航的文字颜色变化来实现。由于计算机具有指令性操作的特性,即程序员要让计算机执行的每一步操作都必须明确说明,否则会出错,如当前播放的是第一张图片,我们如果要想让其播放第二张图片,就需要将第一个dot导航先设置为白色,再将当前的第二个导航设置为蓝色。由于切换图片在多个地方都会用到,为了降低代码的冗余,这里主要通过定义函数来实现,函数体功能为:将所有列表项的文字颜色都改成白色,修改图片和超链接,修改当前dot项的颜色。主要代码如下:

```
var i=-1;//设置全局变量,记录当前播放的图片
function change(k){
clearcolor();//清除所有颜色
imgshow.src=imgspath[k];//通过下标修改显示的
图片
imga.href=hrefs[k];//通过下标修改超链接的链接
目标
dots[k].style.color="blue";//重新设置当前列表项
的文字颜色
}
```

经过上述的基本操作后,只需要设置周期性触发操作的函数,即可实现自动切换。为防止数组溢出显示错误,需要确定全局变量i的值

始终在0~3之间。主要代码如下：

```
function changeimg () {
    if(++i>3) i=0;//如果下一项超出范围,回到第一项
    change(i);修改图片
}
```

3.2 dot 导航

由于计算机的每一步操作都要求由程序员给出,所以为了实现鼠标悬浮更改图片,鼠标离开继续切换的效果,需要通过 onmouseover 鼠标悬浮事件和 onmouseout 鼠标离开事件,分别设置以下效果:

鼠标悬浮 overchange 事件实现轮播图切换到当前图片和超链接,并暂停周期切换,获取当前 dot 导航的下标可以通过参数传递。主要代码如下:

```
function overchange (k) {
    change(k);
    clearInterval(sj);
}
```

鼠标离开 settime 事件继续从当前图片的下一张图片开始周期切换。为了获取当前图片的下标,这里使用了最简单的方法,也是初学者最容易接受的方法,通过参数传递当前导航的下标,因为数组的下标是从0开始的,所以第一个 dot 导航的参数为0。

```
function settime (index) {
    i=index; //通过参数获取当前的 dot 导航下标
    sj=setInterval(changeimg, 3000);
}
```

初学者一定要注意的是,已经定义的变量不需要重复定义,所以这里的周期定时器不能重新定义,否则会发生切换错乱,这里的错乱主要表现为两个定时器同时定时,切换的速度不再是3s一张,而是更快。

3.3 箭头导航

当用户单击左右箭头时,分别切换上一张和下一张,需要注意的是,最后一张的下一张不能切换时,需要将颜色改为灰色,让用户有个心理准备,第一张的前一张同理;否则,实现切换并停止周期定时器。直到鼠标离开时重

新设置周期定时器即可。主要代码如下:

```
function nextimg() {
    if(++i>3) {i=3; rt.style.color= "gray";} //最后一个不能切换,文字设为灰色
    else {change(i); clearInterval(sj);} //切换并停止自动翻页
}
```

最终的轮播图效果如下,除了能正常切换外,图2表示鼠标悬浮时的实现效果;图3表示已经是最后一张还单击下一张的效果。



图2 鼠标悬浮DOT2效果



图3 鼠标单击NEXT效果

4 结语

本案例主要使用了HTML5的盒子模型、列表、超链接、图片等标签;CSS3的过渡、伪类、透明背景、定位和圆角等特效;JavaScript的鼠标单击、悬浮、离开事件,Window对象的周期定时器等功能;是综合性较强的案例,也是初学者比较感兴趣的案例,希望为初学者提供学习素材。

参考文献:

- [1] 张颖. 网页中简易轮播图案例解析[J]. 信息记录材料, 2021, 22(9): 231-233.
- [2] 罗才华. 一种基于原生JavaScript的焦点轮播图简易实现[J]. 南方职业教育学刊, 2019, 9(3): 103-109.
- [3] 黑马程序员. HTML5+CSS3 网站设计基础教程[M]. 北京:人民邮电出版社, 2019.
- [4] 赵建保. JavaScript 前端开发模块化教程[M]. 北京:人民邮电出版社, 2019.

(下转第100页)

文章编号: 1007-1423(2022)19-0091-04

DOI: 10.3969/j.issn.1007-1423.2022.19.016

基于“碳中和”的高性能计算集群组网建设方法探析

陈 阳, 陈坚泽

(广东液冷时代科技有限公司, 佛山 528000)

摘要: 围绕国家人工智能战略布局和发展需求, 基于“碳中和”政策目标, 介绍了高性能计算服务器集群建设方案, 包括高性能计算集群、节能建设、分布式存储集群、高性能计算网络等方案建设。分析高性能计算服务器集群管理需求, 并探析了高性能计算服务器集群管理方法。

关键词: 高性能计算; 服务器集群; 建设与管理

基金项目: 佛山市高新区高技术产业化创业团队专项(2020197000110); 佛山市“蓝海人才计划”创新创业团队项目(2030032000083)

0 引言

2021 年 3 月 13 日, 新华社公布了《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要》(以下简称《纲要》), 《纲要》指出以人工智能为代表的新一代信息技术, 将成为我国“十四五”期间推动经济高质量发展、建设创新型国家, 实现新型工业化、信息化、城镇化和农业现代化的重要技术保障和核心驱动力之一^[1]。围绕国家人工智能战略布局和产业发展需求, 各地政府积极建设人工智能技术支撑平台, 如人工智能基础研究^[2]、智慧交通^[3]、区块链金融、物联网技术^[4]等, 同时启动智慧城市大数据平台^[5]等建设, 为人工智能研究和应用提供健全完善的基础平台服务。这些平台建设对基础计算算力提出了更高的要求, 亟需建设一个高性能计算服务器集群公共服务平台, 为人工智能技术支撑平台以及人工智能科学研究提供高性能、高通量的算力科研保障。同时, 高计算平台也意味着需要更高的能耗才能支撑高算力, 在国家“碳中和”的政策下, 如何利用新型技术, 解决高性能数据中心所带来的高能耗问题也是建立高性能数据中心的潜在研究问题。

本文以华南某科研机构建立高性能计算服

务器集群建设项目为例, 重点介绍高性能计算服务器集群建设中的高性能计算集群、节能建设、分布式存储集群、高性能计算网络等方案建设内容, 并讨论高性能计算服务器集群管理需求以及相应管理方法。

1 高性能计算服务器集群建设

华南某科研机构在前期投资建设项目中已基本完成实验室数据中心基础条件建设, 包括完成基础机房环境、电气系统、基础综合布线铺设等方面的建设, 现针对人工智能关键技术攻关的算力需求, 搭建十个小计算集群, 每个计算小型集群配置算力 1PFLOPS 计算资源, 构建一个柔性的高性能计算集群, 同时采用液冷散热技术提升集群服务器散热问题, 减少空调机组装配, 实现“碳中和”节能目标。

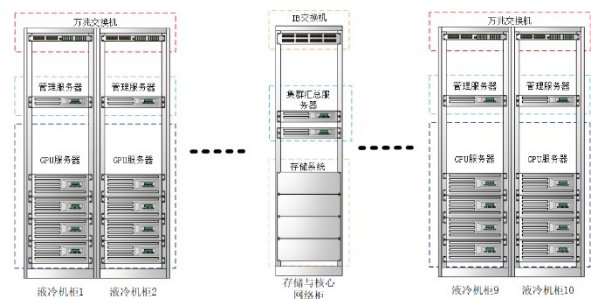


图 1 高性能计算服务器集群建设机柜置放示意图

下面从高性能计算集群、节能建设、分布式存储集群、高性能计算网络等方面阐述高性能计算服务器集群建设内容。

1.1 高性能计算集群方案

高性能计算集群建设基于高性能计算(High Performance Computing, HPC)技术基础构建计算集群平台,其中十台管理服务器,分别用于管理十个集群,同时配备两台管理服务器,用于做统一集群管理。高性能计算集群总体业务架构分为以下四层递进建设。

(1) 基础设施层:利用已建成的模块,方便快捷组装,本项目在现有机房基础设施上部署服务器集群、搭建高性能存储集群、搭建内部集群网络。

(2) 平台核心集群模块:搭建核心 GPU 计算集群、分布式存储集群、高性能运算网络。计算方面,搭建高性能 GPU 计算服务器集群,支撑高密度算力运算需求;存储方面,搭建分布式海量存储集群,支撑海量数据高吞吐访问及大容量存储;网络方面,搭建高速运算网络、高速存储网络、管理网路。

(3) 硬件上构建核心软件中台:主要实现集群管理与作业调度功能,其中,集群管理通过安装部署集群管理软件,构建服务器集群体系,实现算力资源虚拟化管理;作业调度则在集群基础上部署作业调度管理软件,构建算力资源调度体系,实现算力资源弹性调度管理。

(4) 算力业务应用层:通过构建高性能计算开放平台,对各用户提供适用于各种运算业务的运算资源和数据,如大数据计算、基因测序、多模态数据模型、图像识别等。

1.2 节能建设方案

从最新的国家政策导向可知,数据中心低碳节能已经是数据中心建设和运营很重要的一个指标,数据中心节能降耗成为国家“碳中和”“碳达峰”战略的重要一环。因此,基础计算平台建设应该满足 IDC 能耗政策要求。

本次高性能计算服务器集群节能方案采用液冷热管理节能技术,该技术利用超高导热系数的液冷导热模组将服务器高热流密度的核心芯片热量通过液冷模块带到服务器外,进一步通过水循

环冷却系统排到室外,液冷循环采用自然冷却无需压缩机,实现节能效果。图2为本次建设采用的间接液冷系统热管理架构图,间接液冷系统由导热液冷模组、快速接头、液冷分配单元、冷量温控单元、一次冷却环路、二次冷却环路、封闭气冷通道、自然冷却单元等构成,具有低能耗、高功率密度、高可靠性等优点。

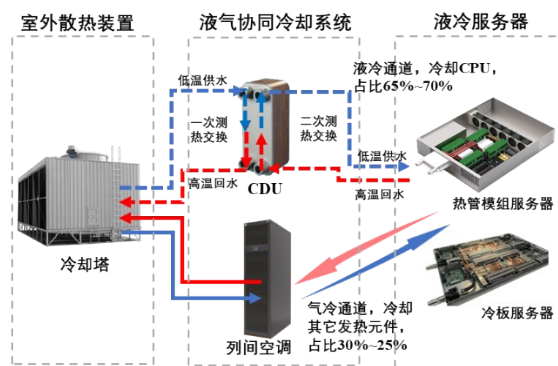


图2 间接液冷技术原理图

1.3 分布式存储集群方案

为构建分布式高性能存储集群,集群建设采用基于BeeGFS并行文件系统分布式存储方案,其主要优势表现在:①分布式文件内容和元数据,有效避免架构瓶颈,一方面可跨多个服务器的条带化文件内容,另一方面则可使文件系统的元数据存放于多个元数据服务器内。②兼容性好,BeeGFS存储服务基于横向扩展(Scale-Out)设计。每个BeeGFS文件系统实例可以具有一个或多个存储服务组件,方便提高性能与空间。一个存储服务实例具有一个或多个存储服务组件。③缓存优化能力强,由于BeeGFS自动使用存储服务器上的所有可用RAM自动进行缓存,因此它还可以在将数据写入磁盘之前将较小的IO请求聚合到较大的块中。④优化高并发访问,BeeGFS用于在高I/O负载的情况下提供最佳的稳健性和性能,优化解决简单的文件系统(比如NFS)在高并发访问的情况下存在严重的性能问题,以及在多个客户端写入同一个共享文件时会损坏数据等典型问题。

1.4 高性能计算网络方案

在高性能计算场景下,由于集群之间需要相互通信,所以对网络的带宽和时延要求比较

高(应用之间带宽>40 Gbps, 时延<10 us 微秒), 现有的TCP/IP软硬件结构无法满足该需求, 因此需要使用RDMA(Remote Direct Memory Access)技术远程直接内存访问, 构建IB(Infiniband)网络实现高性能场景下高速度、高吞吐网络传输需求。RDMA模式对数据包的加工都在网卡内完成。因此就跳过了操作系统, 直接把数据发送到网卡内, 少了应用内存与内核数据之间的交互, 所以速度上更快, 时延更短。IB网络: 基于无限带宽技术, 这种网络有很高的带宽(100 Gb/s以上)和非常低的时延(毫秒级)。

2 高性能计算服务器集群管理方法

2.1 高性能计算服务器集群管理需求分析

本次高性能计算服务器集群管理通过集群管理软件, 构建服务器集群体系, 实现算力资源虚拟化管理。其中集群管理需求主要表现在如下方面:

(1) 满足对多种深度学习、机器学习及大数据任务的资源调度和管理需求, 要求提供大规模GPU集群调度、集群监控、任务监控、分布式存储等功能。

(2) 实现集群资源调度与服务管理统筹, 提供针对GPU优化的调度算法, 实现集群资源调度高效管理。

(3) 提供面向用户的可视化接口或应用接口, 网页端可视化界面、客户端SDK、集成开发环境(IDE)拓展接口等。

(4) 提供丰富的用户管理, 集群、任务监控, 任务调度, 任务错误分析, 任务监控等服务功能, 提高运维人员的工作效率。

(5) 实现容器化和微服务化, 使得运行环境可以在开发和运维达到统一。软件需支持任何形式的计算任务以及大部分计算框架, 包括各种深度学习框架和机器学习框架(如PyTorch、Tensorflow)等。

2.2 高性能计算服务器集群管理方法探析

针对以上分析的高性能计算服务器集群管理需求, 设计图3的高性能计算服务器集群管理架构图, 共分为用户管理、集群管理、业务管理三大方面。

2.2.1 用户管理

及对用户组群集资源参数进行管理如用户组的GPU数量、存储配额、组名、最大运行作业数、等待作业数等; 支持同时在多个已分配资源的租户空间执行任务, 各用户资源互不影响。设置不同的资源分配和服务访问权限; 不同用户组间实现数据隔离。

2.2.2 集群管理

包括分布式管理、并行训练作业管理、集群总计显示等内容。分布式管理是集群管理的重点, 包括如下方面:

(1) 分布式计算集群监控: 包括集群资源总体监控人GPU资源监控。集群资源总体监控, 监控和显示群集GPU、内存和存储总体使用情况。

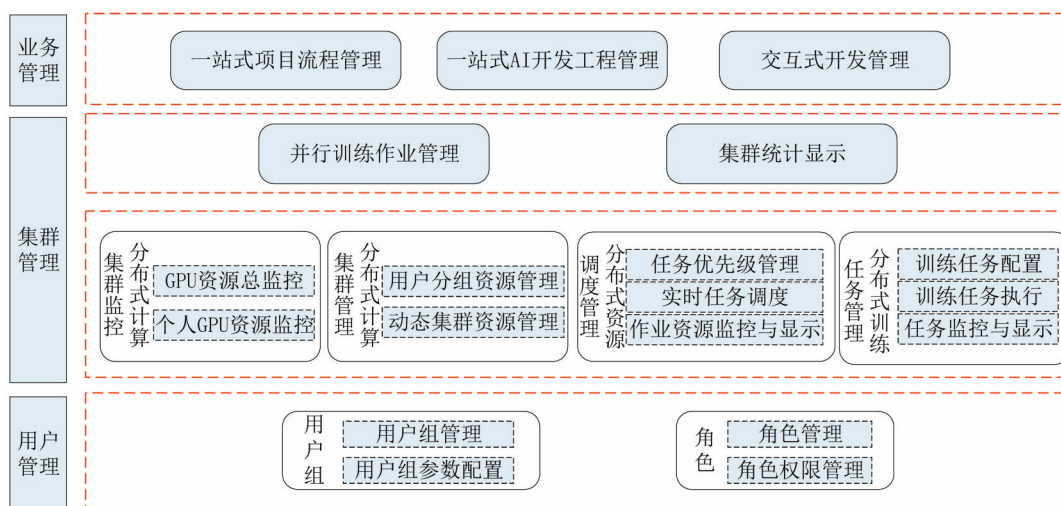


图3 高性能计算服务器集群管理架构图

(2) 分布式计算集群管理: 对用户组进行集群资源配额管理, 对组内用户进行集群资源配额管理; 集群资源管理, 动态添加基础资源, 动态分配和管理集群资源。

(3) 分布式资源调度管理: 执行任务优先级管理; 实时任务资源分配和调度管理; 以任务方式根据优先级分配计算资源, 任务完成进行计算资源回收; 计算作业资源监控和执行情况显示。

(4) 分布式训练任务管理: 训练任务配置管理, 基础任务参数管理; 训练任务执行管理; 训练任务监控, 查看模型训练和资源使用情况。

2.2.3 业务管理

基于高性能计算服务器集群应用的一站式业务流程管理, 包括一站式项目流程管理、一站式AI开发工程管理、交互式开发管理等。

(1) 一站式项目流程管理: 实现项目流程构建、项目流程表单设计、任务分配、团队管理、项目流程可视化、项目检索等项目流程管理支持。

(2) 一站式AI开发工程管理: 包括数据集管理、数据集推荐、模型训练、模型部署、API调用示例、关联用户训练任务与部署任务等内容。

(3) 交互式开发管理: 支持用户通过平台内置AI镜像进行创建交互式开发环境, 环境实例可以使用CPU资源也可以使用GPU资源; 平台支持开发实例的持久化。

3 结语

在国家人工智能战略布局和产业发展大背景下, 高性能计算服务器集群公共服务平台建设需求日渐突出。本文提供一种切实可行、低碳高效的高性能计算服务器集群建设方案, 并探析高性能计算服务器集群管理方法, 借助自主研发集群管理软件, 构建服务器集群体系, 实现算力资源虚拟化管理。相关管理方法有待在实践中进一步优化和深化。

参考文献:

- [1] 《国家及各地区国民经济和社会发展第十四个五年规划和2035年远景目标纲要》[J]. 中国信息界, 2022(5):110.
- [2] 朱磊明. 云计算平台基础服务[J]. 信息与电脑(理论版), 2018(4):30-31.
- [3] 杨燕艳, 朱春燕. 基于大数据技术的智慧医疗平台设计[J]. 信息记录材料, 2022, 23(6):173-176.
- [4] 上海建立类脑芯片与片上智能系统研发与转化功能型平台[J]. 华东科技, 2019(3):11.
- [5] 崔昌云. 智慧城市大数据平台的研究与应用探究[J]. 中国信息化, 2021(4):112-113.

作者简介:

陈阳(1984—), 男, 广东潮州人, 硕士研究生, 高级工程师, 研究方向为计算机应用技术研究

陈坚泽(1983—), 男, 广东潮州人, 硕士研究生, 研究方向为电气自动化、计算机应用技术研究

收稿日期: 2022-05-04 修稿日期: 2022-06-27

Analysis of High Performance Computing Server Cluster Construction and Management Methods

Chen Yang, Chen Jianze

(Guangdong Liquid Cooling Technology Co., Ltd., Foshan 528000)

Abstract: Based on the national artificial intelligence strategy layout and development needs, and the policy goal of “carbon neutrality”, the HPC server cluster construction scheme was introduced in this paper, including HPC cluster, energy saving construction, distributed storage cluster, HPC network construction scheme. The requirements of high performance computing server cluster management were analyzed and the methods of high performance computing server cluster management were discussed.

Keyword: high performance computing; server cluster; construction and management

开发案例

文章编号: 1007-1423(2022)19-0095-06

DOI: 10.3969/j.issn.1007-1423.2022.19.017

基于设备规则的变电站自动成票推理系统

顾明煜, 丁雪峰, 江红成, 花 冉, 赵 军

(国网江苏省电力公司南通供电公司, 南通 226001)

摘要: 变电站的检修工作对整个电力系统供电的质量与效率有着重要的影响, 检修工作其中一个重要的环节是开具工作票, 工作票的重点内容是安全措施。开具工作票与变电站接线图、设备特征、工作内容等因素相关, 传统的工作票开具方式主要依靠人工或部分机器参与, 已不能很好满足日益繁忙的检修要求。本文利用计算机和人工智能技术, 结合实际工作需求研发智能开具工作票的算法, 从电力系统相关的规则知识入手, 将工作计划停电间隔中的设备所涉及的安全措施一一映射为成票规则, 建立一种新型工作票的自动成票推理算法系统, 经测试具有较大推广价值。

关键词: 设备规则; 关系映射; 人工智能; 算法推理; 自动成票

0 引言

变电站是电力能源传输系统中的重要一环。变电站所涉及的电力生产是一项复杂、高综合性的业务流程, 所需处理的数据内容数量多、离散性强^[1]。变电站进厂检修的必要条件是工作票, 工作票是指在进场工作前由项目负责人按照工作任务要求对现场进行风险辨识, 按工作内容要求确定检修人员、范围、工作时间等, 并将采取的安全措施填写所成的票据。工作票可防止盲目性误操作, 确保检修过程顺利进行, 提高检修作业的安全可靠性。变电检修工作票是使检修全过程各工序之间科学协调和有序推进, 真正达到预防为主, 实现事故为零的目标^[2]。工作票所涉及的知识体量十分庞大, 利用传统技术仅仅依靠键值对匹配的方式去数据库中筛选开具工作票的方式, 不仅对服务器性能的要求十分苛刻, 而且匹配一个变电站每条线路的数据关系就需要花费很多的时间。

当前已有的工作票系统虽然已经能够开具一些简单通用的工作票, 但对于高压电力电缆工作所需要的第一种工作票, 目前还不能有效包含站内的具体安全措施, 需要人工介入增补

变电检修第一种工作票, 使得开票和许可流程变得复杂, 快速入站检修需求与严谨工作票管理要求之间的矛盾日益突出。

本文系统利用规则映射、人工智能、全文检索等相关技术, 结合全文检索技术具备的高并发处理能力、快速检索性能^[3], 将繁琐庞大的工作票开具流程交由全流程自动化方式去实现, 用一种基于数据表的新方法实现作业票管理系统的数据流转^[4]。在实际生产工作中, 变电站检修负责人只需要提供工作计划中对应停电间隔的一次接线图纸, 由机器智能映射出各个设备的安全措施, 生成对应的工作票, 以实现工作票的快速生成、生产管理系统的自动填报, 协助配网运维人员高效、便捷地完成工作票系统录入工作, 保障工作票规范、准确^[5]。

1 模型框架

1.1 工作流程及推理策略

成票推理算法的核心是将设备规则和工作计划表中设备一一映射, 从而推理成工作票安全措施。工作票自动成票的过程最重要的一个环节是根据设备规则表推断出每条工作计划中

停电间隔重要设备关系及其属性、电气设备表的安全措施。即变电检修人员上传一条工作计划后，从工作计划中解析对应的变电站名称、停电间隔范围、工作设备和工作任务，结合设备规则表推理得出标示牌，断开开关等安全措施，最后将所有推理出来的安全措施整理到工作票页面上。

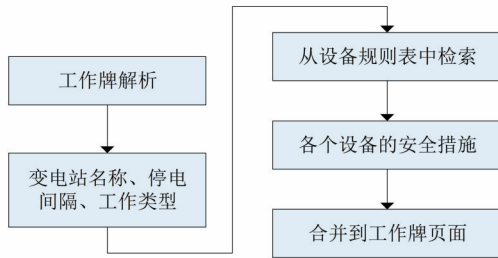


图 1 工作票成票算法流程图

1.2 规则表

电气常量规则表，是指将电站不同种类的设备和不同种类的工作，按其所属的类别和名称进行排列，方便后续进行检索和分词时能够快速匹配到。结构见表1。

表 1 电气常量规则表

编号	关键字	名称	别名	属性
----	-----	----	----	----

停电间隔规则表，是指变电站某个停电间隔内所有设备组成的关系集合映射，方便后续对线路进行解析时，加载该停电间隔内的所有设备。结构见表2。

工作票规则表，是指不同设备和安全措施文本组成的关系集合映射，方便后续整理工作票页面的安全措施。结构见表3。

表 2 停电间隔规则表

别名	一重名称	类型	属性	子间隔	主标识	设备	左侧节点	左侧地刀	左侧地刀	右侧节点	右侧地刀
----	------	----	----	-----	-----	----	------	------	------	------	------

表 3 工作票规则表

编号	对象	安措文本	是否合并	停电间隔类型	停电间隔属性	类型	个数	设备类型	关键字	变量取值	检修内容个数	检修内容类型
----	----	------	------	--------	--------	----	----	------	-----	------	--------	--------

2 算法实现

2.1 全文检索技术

全文检索大致分为索引创建和搜索索引两个过程。索引创建是利用分词组件将文档切分(split)一个个单词、去除停词(通过与停词集合的比对，存在于分词组件中)、去除标点、最后获取词元。搜索索引指当用户输入一个检索短语后，对检索短语进行的词法分析：①关键字比对(内置关键字库)，如不是关键字则作为普通词处理；②语法分析，将语句根据语法规则形成语法树。

本文利用文件系统中已经存在的变电站规则表数据，通过 POI 工具首先加载表格中表单的内容，分词组件对不同文档进行切分，根据常见的切分规则对表格中相应的数据进行切分，

加载到全文检索引擎内存中，开票时利用工作计划中的“变电站+停电间隔+工作内容”形成查询语句，进行后续搜索过程。

2.2 基于统计的机器学习分词算法

2.2.1 算法流程

该分词方法基于人工标注的词性和统计特征，对文字进行建模，即根据观测到的数据(标注好的语料)对模型参数进行训练，在分词阶段再通过模型计算各种分词出现的概率，将概率最大的分词结果作为最终结果。机器学习是通过让计算机模拟人对句子的理解，达到识别词的效果^[6]，机器学习应用到分词算法中即衍生了基于统计的机器学习分词算法。算法描述如下：

假定句子 $S(W_1, W_2, \dots, W_n)$ ， W_i 代表词，其概率为 $P(S) = P(W_1, W_2, \dots, W_n)$ 根据条件概

率公式, 每个词出现的概率等于之前每个词出现的条件概率相乘。

$$P(W_1, W_2, \dots, W_n) = P(W_1) \cdot P(W_2 | W_1) \cdot P(W_3 | W_1, W_2) \cdots P(W_n | W_1, W_2, \dots, W_{n-1}) \quad (1)$$

但是这样计算量太大, 句子越长越麻烦, 因此Andrey Markov提出了一种简化方法, 即马尔可夫假设: 假设任意一个词出现的概率只与它前面的一个词有关。

$$P(S) = P(W_1) \cdot P(W_2 | W_1) \cdot P(W_3 | W_2) \cdots P(W_n | W_{n-1}) \quad (2)$$

这就是二元模型(Bigram Model), 相应的高阶语言模型即任意一个词 W_i 出现的概率只与它前面的 $i-1$ 个词有关。元数越高越准确, 但相应越复杂、越难实现。

2.2.2 具体实现

本文采用三元模型, 在统计量足够大的情况下, 任意词的出现概率描述如下:

$$P(W_i | W_{i-1}) = P(W_{i-1}, W_i) / P(W_{i-1}) = \text{联合概率} / \text{边缘概率} = \text{两个词一起出现的次数} / \text{单个词出现的次数}$$

系统中需要对设备名称、设备类型等名词进行分词, 以获取最简短和最精准的词语, 方便检索出最匹配的规则表, 就需要根据上述算法分词, 去除干扰检索效率的词语, 通常形容词、地名、符号等信息可以通过该分词算法去除, 得到最大概率的短词。例如: “110 kV 华庆 745 开关” 通过分词算法, 拆分成 “110 kV” “华庆 745 开关” 短语, 把这两个任务分开处理, 也就是不同任务使用不同的处理策略或两个任务按时间顺序先后被处理^[7], 后续通过检索技术可以大大缩小检索的时间。

2.3 基于设备规则的成票推理算法

工作票内容包括停电间隔内工作地点及设备双重名称、工作内容, 应拉断路器(开关)、隔离开关(刀闸), 应设遮拦、应挂标示牌及防止二次回路误碰等措施^[6]。

技术路线描述如下: ①建立停电间隔内设备的映射规则: 首先将每种设备类型对应的关键字、名称、别名、属性的关系一一映射, 将工作类型的关键字、名称一一映射, 对于一个

变电站解析出来的停电间隔应该有其详细的设备映射; ②建立设备的工作票规则。

具体实施步骤如下: ①任务开始时, 初始化设备规则表, 利用文本检索的方式, 从规则表中用POI工具读取表格中每个设备类型, 依次加载到JVM内存中; ②初始化工作规则表, 利用文本检索的方式, 从规则表中以POI工具读取表格中每个工作类型, 依次加载到JVM内存中; ③初始化电气常量规则表, 将上述两步初始化出来的设备规则表和工作规则表封装到电气常量规则表中; ④解析用户提交的工作计划表项之工作内容, 主要是检查: 之前的顿号是不是停电间隔, 如果是, 再次拆分; ⑤初始化停电间隔规则表, 对停电间隔内容的设备名称、相邻名称以及属性进行解析; ⑥构造停电间隔内设备的设备属性和相邻关系表, 依次遍历该停电间隔内的设备属性和相邻关系; ⑦汇总电气表内所有设备, 根据设备全称获取设备对象信息体, 封装成设备的集合, 同时需要对重复的设备进行去重; ⑧构造相邻关系表项的相邻设备信息, 根据前面得到的停电间隔内的设备相邻关系表, 再次迭代出相邻设备的信息; ⑨将电气表内的所有设备信息以及相邻设备信息封装到电气表字典中; ⑩得到电气表字典中信息后, 通过每个设备所属的设备类型从工作票规则表中关联对应的安全措施文本, 将所有安全措施文本一一填充到工作票的内容页面, 再通过POI工具替换标记符, 生成工作票文本。

2.4 软件实现

算法系统软件流程图如图2所示。

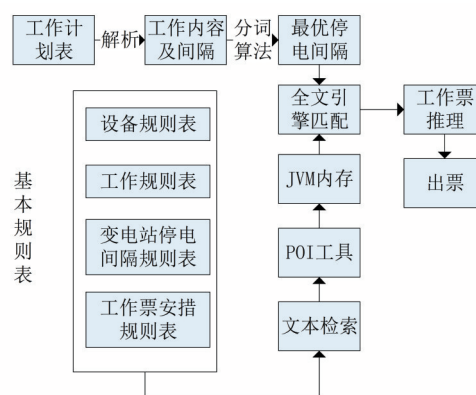


图2 软件流程图

3 展示及测试

为验证工作票生成算法的有效性，需要输入实际工作计划表进行测试。下面以某 110kV 变电站为例，进行说明。变电站局部接线图如图 3 所示。

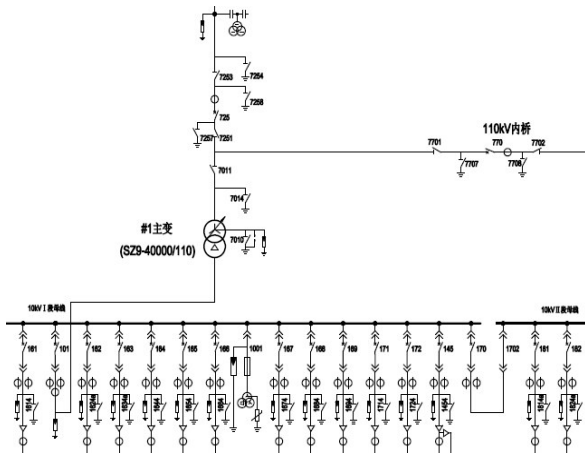


图 3 变电站局部接线图

3.1 电气常量规则表

根据电气图，整理生成电气常量规则表，部分如表 4 所示。

表 4 电气常量规则表

编号	关键字	名称	关键字	名称
1	开关保护	开关保护	缺陷处理	缺陷处理
2	主变保护	主变保护	渗油处理	渗油处理
3	开关	开关	发热处理	发热处理
4	流变	流变	小修	小修
5	避雷器	避雷器	预试	预试
6	压变	压变	校验	校验
7	保护	保护		
8	刀闸	刀闸		

3.2 停电间隔规则表

根据电气图生成 110 kV 变电站的设备类型表，生成后的部分规则如表 5 所示。

表 5 停电间隔规则表

别名	一重名称	类型	属性	子间隔	主标识	设备	左侧点	左侧地刀	右侧点	右侧地刀
1	110 kV 锋灶 725 线	线	110 kV 锋灶	线	AIS、线路	725	7251、7253、7254、7257	流变、开关保护	7258	7253
2	110 kV 锋灶 725 开关	开关	110 kV 锋灶	开关	AIS、线路	745	7451、7453、7454、7457	流变、开关保护	7458	7453
3	110 kV 内桥 770 开关	开关	110 kV 内桥	开关	AIS、内桥	770	7701、7702、7707、7708	流变、开关保护	7707	7701

3.3 工作票规则表

生成工作票规则表，生成后的部分规则如表 6 所示。

表 6 工作票规则表

编号	对象	安措文本	是否合并	停电间隔类型	间隔属性
1	刀闸	在 X 刀闸线路侧装设接地线一组	否	线	AIS
2	刀闸	合上 X 接地刀闸	是	开关	AIS* 线路
3	刀闸	在 X 刀闸开关侧装设接地线一组	否	开关	内桥

3.4 工作计划

上传变电站对应实际检修工作计划，列出部分测试数据如表 7 所示。

表 7 工作计划表

工作类型	停电范围	工作内容
检修	10 kV1 号电容器及 162 开关	1 号电容器小修、预试；162 开关、流变小修、预试
检修	1 号主变；110 kV 锋灶 725 开关；110 kV 内桥 770 开关	7701 刀闸小修
检修	10 kV 华庆 167 线	开关、流变、避雷器小修、预试，保护校验

3.5 测试指标

考虑测试结果的准确性和有效性, 提供10个变电站, 每个变电站提供20条工作计划进行测试。通过统计测试项目的分词准确率, 工作任务推理准确率, 安全措施推理准确率, 平均生产时间, 得出如下结果。

表8 工作票规则表

指标	结果
分词准确率	96%
工作任务推理准确率	97%
安全措施推理准确率	97%
平均生产时间	5 s

3.6 生成工作票

在前端界面上点击开票, 生成该条工作计划表所对应的工作票, 依次可以从生成的工作票界面上找到以上工作计划表中的变电所, 以及所解析出来的工作设备和安全措施, 在上述的规则表中都能够一一找到对应映射关系。

停电范围	工作内容	计划开始时间	计划结束时间
1号主变; 110kV锋灶725开关; 110kV内桥770开关	7701刀闸小修	2021-04-19 09:30:00	2021-04-25 18:30:00

图4 工作计划图

4.工作任务:

工作地点及设备双重名称	工作内容*
110kV内桥7701刀闸	7701刀闸小修

5.计划工作时间: 自 2021-04-19 09:30:00 至 2021-04-25 18:30:00

6.安全措施(必要时可附页绘图说明, 红色表示有电):

应拉断路器(开关)、隔离开关(刀闸)

拉开101、725、770开关

拉开7011、7251、7253、7701、7702刀闸

将101开关手车拉至试验位置

将101、725、770开关方式开关切至就地位置

拉开7010接地刀闸, 断开7010接地刀闸操作电源

断开7701刀闸操作电源

应装接地线、应合接地刀闸 (注明确实地点、名称及接地线编号)

在7701刀闸主侧装设接地线一组

在7701刀闸开关侧装设接地线一组

图5 工作票图

4 结语

本文针对电力行业的变电站检修工作任务, 从工作计划表着手, 基于“工作设备+工作任务”逻辑推导, 设计实现了基于设备规则的开票算法。系统算法通过预定义的设备规则表, 通过解析设备类型(开关、流变、压变、避雷器、主变、刀闸等设备), 及设备类型的安全措施, 通过检索的方式自动生成工作计划所需安全措施。系统测试表明, 本算法具备开具变电站检修任务工作票的功能, 已达到了算法和软件所设定目标。

参考文献:

- [1] 冷超,徐良,唐大勇,等. 新生产管理系统(PMS)在静安变电站的应用[J]. 华东电力, 2010, 38(4): 561-563.
- [2] 赖湘华. 工作票在检修工作中的应用[C]//《中国建材科技》安全生产技术交流会专刊(二), 2015: 25-27.
- [3] 宋玉红,吴琼贵. 基于Elasticsearch大数据搜索引擎全网流监测系统设计与实现[C]//第十四届全国信号和智能信息处理与应用学术会议论文集, 2021: 483-487.
- [4] 赵叶. 供电系统作业票计算机自动生成与管理系统设计[D]. 石家庄:河北科技大学, 2011.
- [5] 陈稷露,丁鹏. 配电网智能工作票系统的设计与应用[J]. 机电信息, 2020(36): 99-100.
- [6] 刘武. 基于统计机器学习算法的汉语分词系统的研究[D]. 北京:北京邮电大学, 2006.
- [7] 袁彰. 基于自然语言处理技术及智能成票规则的配网工作票系统[J]. 电子元器件与信息技术, 2021, 5(2): 190-192.

作者简介:

顾明煜(1989—),男,江苏南通人,高级工程师,工程硕士,从事领域为变电一次检修技术管理

丁雪峰(1980—),男,江苏如皋人,高级工程师,从事领域为变电运检技术管理

江红成(1976—),男,江苏如东人,高级工程师,工程硕士,从事变电运检技术管理

花卉(1985—),男,江苏如东人,高级工程师,从事领域为变电运检技术管理

赵军(1983—),男,江苏南通人,高级工程师,工学硕士,从事领域为变电运检技术管理

收稿日期: 2022-05-10 修稿日期: 2022-07-01

Reasoning Algorithm of Substation Automatic Ticket Generation Based on Equipment Rules

Gu Mingyu, Ding Xuefeng, Jiang Hongcheng, Hua Ran, Zhao Jun

(State Grid Jiangsu Electric Power Company Nantong Power Supply Company, Nantong 226001)

Abstract: The maintenance work of substation has a very important impact on the quality and efficiency of power supply of the whole power system. However, one of the important links of maintenance work is the issuance of work ticket. However, the traditional issuance method of work ticket can not meet the requirements of actual production due to various reasons. Using computer and artificial intelligence technology, combined with the actual work needs, it is of great significance to develop the algorithm of intelligent work ticket issuance. In particular, starting with the relevant rule knowledge of power system, the safety measures involved in all equipment in the outage interval of work plan are mapped into ticket making rules one by one, and a new automatic ticket making reasoning algorithm system of work ticket is established.

Keywords: relational mapping; artificial intelligence; algorithmic reasoning; automatic ticket

(上接第90页)

作者简介:

赵素萍(1986—),女,山西太原人,硕士研究生,讲师,研究方向为网站设计、移动互联

杨璐(1984—),女,河北唐山人,硕士研究生,讲师,

研究方向为移动互联

康晶晶(1984—),女,山西太原人,硕士,副教授,研究方向为软件工程、网站设计

收稿日期:2022-05-19 修稿日期:2022-06-24

The Practice and Application of the Navigation Production of the Carousel Map in the Web Front-End

Zhao Suping¹, Yang Lu², Kang Jingjing¹

(1. School of Big Data, Jinzhong College of Information, Jinzhong 030800;

2. School of Information Engineering, Jinzhong College of Information, Jinzhong 030800)

Abstract: No matter on the web page of the computer or mobile phone, no matter on the mobile APP or the mini programs embedded in various APPs, the theme banner has a rotation chart. The rotation chart is a highly ornamental way to display information. The rotation chart is mainly used to realize the effect of periodically switching pictures. The focus navigation in the rotation chart can prompt and control the currently playing pictures. At the same time, the user can control the previous and next pictures through the arrows. This article uses HTML5, CSS3 and javascript technology to implement. It is a comprehensive case, hoping to provide a way of thinking for website designers.

Keywords: carousel map navigation; focus; arrow

文章编号: 1007-1423(2022)19-0101-05

DOI: 10.3969/j.issn.1007-1423.2022.19.018

机载气吸式抓取一体装置设计与试验

吴子彬¹, 吴志坤², 许绍聪¹, 余乃浩², 林嘉乐¹, 殷惠莉¹

(1. 华南农业大学电子工程学院(人工智能学院), 广州 510642; 2. 华南农业大学工程学院, 广州 510642)

摘要: 为了解决传统机载设备抓取不同形状物体难度大, 作业效率低的问题, 设计一种以负压发生器、泄气阀、吸附吸盘为核心的气吸式机械手, 实现具有无人直升机特色的抓取装置。该抓取装置与远程视觉系统通信, 配合无人直升机上的双目摄像头进行图像识别, 将目标对象的图像坐标转换为世界坐标, 使装置完成准确搜索、拾取以及投放等相应任务。

关键词: 作业型无人机; 气吸式机械手; 抓取一体装置; 机器视觉

基金项目: 2018 年第二批教育部产学研合作协同育人项目(201802070138、201802070088); 校级大学生创新创业训练计划项目(X202110564138)

0 引言

无人直升机, 因其结构紧凑、转场灵活、机动性和安全性高, 且可以实现空中悬停、垂直起降等特点, 在搜救、自然灾害监测等领域都有着广阔的应用前景^[1-2]。然而当下国内对于无人机搜救抓取方面的研究尚不成熟, 薛彩梁等^[3]提出一种铁丝网抓取方案, 虽然抓取面积大, 但容易掉落。章东莞等^[4]提出一种搜救抓取一体化无人机的方案设计, 但传统机械手抓取不同形状物体难度大, 作业效率低, 且抓取精度不高。因此本文设计了一种机载半自动气吸抓取装置, 该装置采用吸附吸盘接触抓取物体, 可以克服物体形状问题, 有效提高了精度和作业效率。同时附有保护装置, 确保物品不会因飞行姿态改变而掉落。

1 设计思路

无人直升机机载半自动气吸抓取装置主要由两部分构成, 无人直升机与气吸机械抓手。无人直升机采用直升机为亚拓 700 级航模遥控直升机, 该直升机在正常飞行中可以提供 100 N~600 N 的力, 因此可搭载 10 kg~20 kg 的负载工作。气吸式机械抓手部分又分为主控系统、机械系统、视觉系统三个方面。在抓手部分搭建

电控图传系统, 上位机搭建视觉系统, 使用图传设备实现超视距识别并实时回传视频, 可直接在视觉系统界面采取手动点击画面, 系统会把点击处的坐标信息以算法解读反馈给电控系统, 电控系统读取坐标信息, 得到物体在现实世界中位置距离信息, 并控制机械部分完成搜索、拾取以及投放等相应任务。使其在灾害领域可以发挥安全、高效等作用。

2 主控系统

在控制系统中, 双目摄像头通过图传模块将图像信息传输至计算机上位机端, 上位机经过一系列信息处理, 获得目标的坐标信息, 并通过数传模块将信息发送至装置主控系统, 主控系统分别对舵机、继电器、泄气阀进行操作以此控制机械手完成相应任务, 控制示意图如图 1 所示。

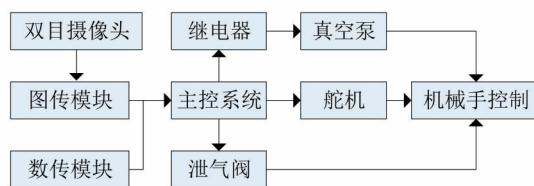
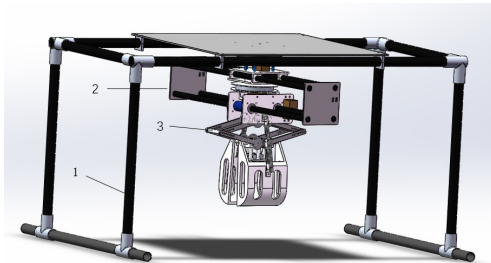


图 1 控制示意图

3 机械部分

机械手的设计包括主体架、传动部分、机械爪三部分，完整建模图如图2所示。



1—主体架 2—传动部分 3—机械爪

图2 机械手的机械建模图

3.1 机械手主体架

主体架采用碳管以及三通铝管搭建，具有质量轻、强度高、耐腐蚀等特点^[5]，由于700级直升机施加给传动轴的扭矩不超过1000 N·m，所以在进行仿真分析时外部载荷施加的扭矩为1000 N·m。运用SolidWorks Simulation进行静应力分析，得出传动轴受到的最小应力为4.154 N/m²，最大应力为6.677e+07 N/m²，仍小于碳纤维板的拉伸强度3500 Mpa和弯曲强度1700 MPa。故可以使脚架在无人机下保持稳定。

3.2 机械手传动部分

传动部分分为旋转部分和水平位移部分。抓取装置旋转结构设计机械手的旋转360°舵机直接驱动。在本设计中，舵机机身作为动力输出部分，机身与机械手整体固定一起转动。与舵机齿轮配合的舵盘直接固定在脚架支撑板上，这样更加充分利用有效空间，使系统更加紧凑，旋转结构的转动范围为-90°~+90°，即可覆盖全部抓取区域。

装置水平位移设计选用软传动设计方案。水平位移结构由四根8×12×850 mm碳纤维圆管做导轨与结构框架，整个水平位移结构上层两个碳管和旋转结构固定，下层的两根碳管用来实行位移的导轨，导轨上装有一对直线轴承，可以在导轨上顺滑运动。机械手水平位移原理，皮带轮与皮带啮合，而皮带从结构框架一端起，经过压带轮，再环绕皮带轮3/4圈，行成包角，再通过另一个压带轮回到框架另一端，相当于

皮带与框架都相对静止，从正上方看，形成字母Ω。720°舵机带动皮带轮，皮带轮在皮带上的位移就是水平位移的距离。

皮带轮的直径=

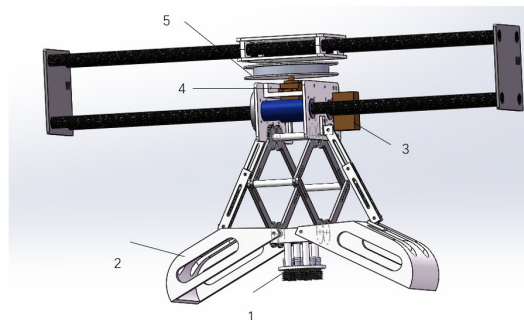
$$\frac{\text{水平位移总距离} - 20 \text{ mm (左右两端安全距离)}}{2\pi} \quad (1)$$

这样得出的皮带轮直径，从复位原点到最大移动距离舵机行程刚好为一圈。

3.3 机械爪抓取

该机械爪的抓取系统主要是由负压发生器、泄气阀、吸附吸盘、真空吸盘缓冲支架、气管、气动三通管组成。真空吸盘是气吸式抓取的执行结构。工作时首先将真空吸盘通过接管与真空泵接通，然后与待抓取物体等接触，此时泄气阀开始工作，真空泵开始抽吸，使吸盘产生负压，被吸材料将被牢固地抬起，可以开始移动被抓取物。当被抓取物体抓取成功时，升降系统到达上止点，泄气阀停止工作，使真空吸盘内由负气压变成零气压或稍为正的的气压。由于重力被抓取物体就会脱离吸盘，从而完成了抓取和释放动作。

本设计中，五个真空泵分别对应五只吸盘，相互独立，因此在吸附抓取时即使个别吸盘未能与待抓取物体接触，也能完成抓取任务。并附加一个联动抓手结构，通过一个120°舵机、在下降过程中抓手张开，上升过程中抓手合拢，避免因飞行姿态的改变而导致抓取失败。抓取建模图如图3所示。



1—吸附吸盘 2—联动抓手 3—120舵机
4—720舵机 5—皮带轮

图3 机械手的抓取建模图

4 视觉系统

4.1 视觉处理流程

本文中,视觉系统需要实时显示图传的图片,并通过神经网络进行数字图像几何畸变矫正^[6],计算机端根据图像建立直角坐标系,操作者点击位置后通过程序将其转化为极坐标,再发送指令控制机械手精准地抓取目标,流程图如图4所示。

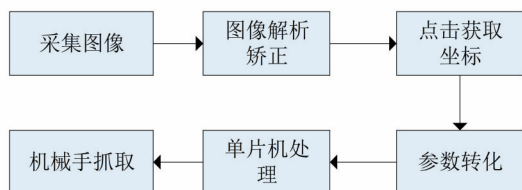


图4 抓取流程图

4.2 下位机端视觉系统搭建

下位机端在无人直升机脚架上构建电控图传系统,本文采用Ratel2摄像头搭配PandaRC VT5802图传,该搭配具有传输图像清晰、稳定、抗干扰能力强等一系列优点。具体硬件连接图如图5所示。



图5 硬件连接图

4.3 上位机端视觉系统搭建

上位机端的视觉系统搭建,要求能够与无人直升机上的电控图传系统通信,接收其发送的数据和图片,传输过程具有快速、实时、准确等特点。并且视觉系统需要友好的GUI界面,方便操作者观察数据和发送指令,本文用Python搭建上位机,利用Python的serial模块、

QT和OpenCV库完成搭建。上位机界面如图6所示。

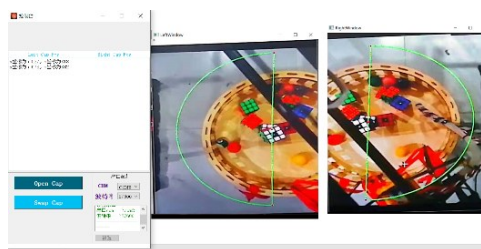


图6 上位机软件图

4.4 窗口点击参数转化

双目视觉可以从两个视点观察同一物体,然后通过三角测量原理计算图像像素间的位置偏差(视差)来获取景物的三维信息^[7]。如图7所示,此时 $X_{left} = X_{right} = Y$, f 是摄像机的焦距。由几何关系可得:

$$\begin{cases} X_{left} = f \frac{x_c}{Z_c} \\ X_{right} = f \frac{x_c - B}{Z_c} \\ Y = f \frac{y_c}{Z_c} \end{cases} \quad (2)$$

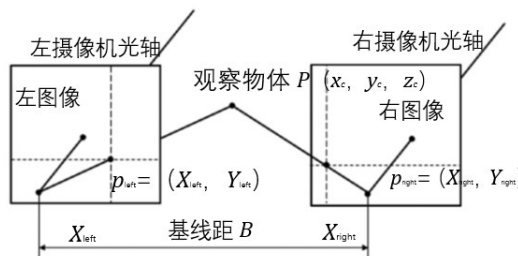


图7 双目视觉成像原理

视差 $D = X_{left} - X_{right}$,由此可计算出特征点P在摄像机坐标下的三维坐标为

$$\begin{cases} x_c = \frac{B \cdot X_{left}}{D} \\ y_c = \frac{B \cdot Y}{D} \\ z_c = \frac{B \cdot f}{D} \end{cases} \quad (3)$$

在视觉系统中,世界坐标与像素坐标是不同的,为了准确抓取该目标需要确定两者的关系^[8]。需要对双目摄像机标定,本文采用张正友标定法,得到相机内部参数矩阵M1,外部参

数矩阵 M_2 ，由如图 8 所示的小孔成像模型，在摄像机的光轴中心建立坐标系。

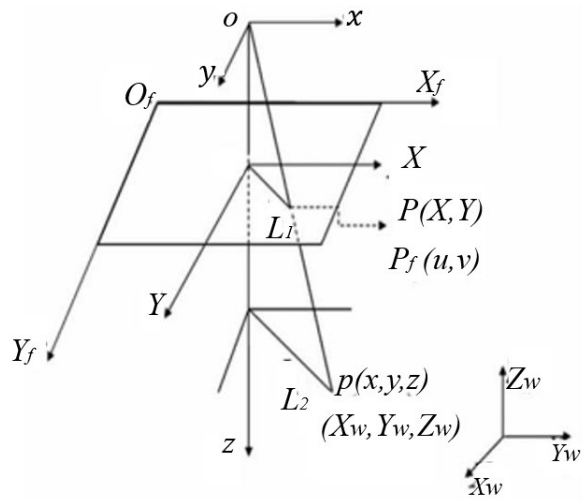


图 8 小孔模型

其中 xoy 为相机坐标， $X_fO_fY_f$ 为图像坐标， XOY 为像素坐标， $X_wY_wZ_w$ 为世界坐标，各坐标系之间的转换关系如下。

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{f}{dx} & 0 & u_0 & 0 \\ 0 & \frac{f}{dy} & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = M_1 M_2 \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (4)$$

视觉系统坐标为 (x, y) 坐标系，机械手运动轨迹为极坐标，由以下公式进行转化。

$$r = \sqrt{x^2 + y^2} \quad (5)$$

$$\theta = \arctan \frac{y}{x} (x \neq 0) \quad (6)$$

5 机载机抓取试验

5.1 试验目的

实物制作完毕之后，对机械手进行试验测试，进行模拟搜救任务，验证直升机机载抓取的效果是否符合设计要求，并对试验过程中装置出现的问题进行进一步优化，为设计提供参考。

5.2 试验内容

无人直升机从起降区起飞，飞行至救援物资存放区，领航员通过机载图像设备寻找、选定并自动拾取某一特定的救援物资(魔方与斯诺

克彩球)。飞越障碍后将该救援物资投放给待救区内特定的待救对象，完成一趟搜救任务，试验现场图如图 9 所示。在规定时间内，循环试验抓取性能，试验设备稳定性与准确性。

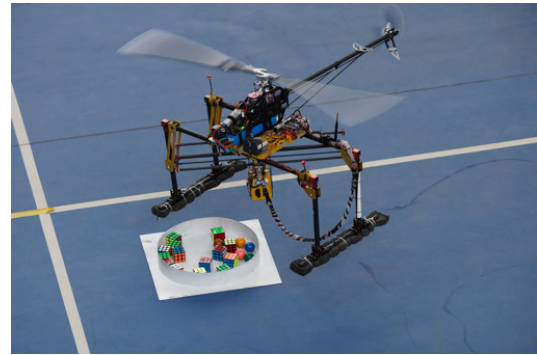


图 9 试验现场图

5.3 试验数据

在 5 min 内进行抓取试验，试验所得到的数据见表格 1。

表 1 模拟搜救试验数据

数据类别	第一组	第二组	第三组	第四组
魔方抓取个数/个	6	8	7	9
彩球抓取个数/个	6	6	6	6
直升机往复次数/次数	12	14	13	15
直升机飞行时间/秒	14	10.9	12	10.5
二次抓取次数/次	1	1	1	0
空中掉落次数/次	0	0	0	0
平均抓取时间/秒	6	5	5.5	5
平均释放时间/秒	5	5.5	5	4.5
抓取成功率/%	92	93	93	100

6 结语

本文设计了一种无人直升机机载气吸式抓取机械手，机械手满足抓取(吸取)类方体和球体以及不规则表面的物体，能够在各种恶劣环境下进行搜救任务。本设计对搜救机器人的研究和推广有一定的参考价值。试验结果显示，对机械手的结构和控制系统进行优化后，可以满足三维空间作业需求，机械手动作响应较快，误差较小，满足高精度快速运作的的需求。

参考文献:

[1] 于进勇,王超.垂直起降无人机技术发展现状与展望[J].飞航导弹,2017(5):37-42.

- [2] 刘泽宇,蒋克全,程兴国,等. 无人直升机的价值和未来发展趋势[C]//2019年(第四届)中国航空科学技术大会论文集. 中航出版传媒有限责任公司(China Aviation Publishing & Media Co., 2019: 410420.
- [3] 薛彩梁,乔栋. 某空中飞行器抓取装置结构优化设计[J]. 现代国企研究,2016(10):123.
- [4] 章东莞,张扬,田文强,等. 搜救抓取一体化无人机的方案设计[J]. 科技创新与应用,2019(22):105-107.
- [5] 黎小平,张小平,王红伟. 碳纤维的发展及其应用现状[J]. 高科技纤维与应用,2005(5):28-34,44.
- [6] 陆懿,陈光梦,程松. 基于神经网络的数字图像几何畸变矫正方法[J]. 计算机工程与设计,2007(17):4290-4292.
- [7] 章毓晋. 计算机视觉教程[M]. 北京:人民邮电出版社,2011:220-242.
- [8] 张纯纯,冯创意,高统林. 基于机器视觉的目标识别与抓取研究[J]. 农业装备与车辆工程,2019,57(12):93-96.

作者简介:

吴子彬(2000—),男,广东汕头人,本科,主要研究方向为电子信息工程

吴志坤(2000—),男,广东揭阳人,本科,主要研究方向为机械自动化

许绍聪(2001—),男,广东阳江人,本科,主要研究方向为电子科学与技术

余乃浩(2002—),男,福建龙岩人,本科,主要研究方向为机械自动化

林嘉乐(2002—),女,广东广州人,本科,主要研究方向为电子信息工程

通信作者:殷惠莉(1975—),女,山东东营人,硕士,讲师,主要研究方向为电子技术、农业信息化,E-mail: huiliyin@scau.edu.cn

收稿日期: 2022-05-23 修稿日期: 2022-06-16

The Design of an Airborne Air Suction Grab Integrated Device

Wu Zibin¹, Wu Zhikun², Xu Shaocong¹, Yu Naihao², Lin Jiale¹,
Yin Huili¹

(1. College of Electronic Engineering (College of Artificial Intelligence), South China Agricultural University, Guangzhou 510642;
2. College of Engineering, South China Agricultural University, Guangzhou 510642)

Abstract: In order to solve the difficulty and low efficiency of traditional airborne equipment in grasping objects of different shapes, an air suction manipulator with negative pressure generator, vent valve and suction cup as the core was designed to realize the grabbing device with characteristics of unmanned helicopter. The grasping device communicates with the remote vision system, coordinates with the binocular camera on the unmanned helicopter for image recognition, and converts the image coordinates of the target object into the world coordinates, so that the device can accurately search, pick up and drop corresponding tasks.

Keywords: operational unmanned aerial vehicle; pneumatic manipulator; grab integrated device; machine vision

文章编号: 1007-1423(2022)19-0106-05

DOI: 10.3969/j.issn.1007-1423.2022.19.019

临床分子诊断与基因检测信息平台的构建

莫远明, 卢瑞鸿, 陈 康

(中山大学附属第一医院信息数据中心, 广州 510080)

摘要: 为提高分子诊断与基因检测中心工作效率和精细化管理水平, 完善项目报告管理流程, 杜绝多收漏收费用等情况, 设计一套基于互联互通数据平台的信息系统, 实现业务全流程的数据闭环, 同时针对项目报告的复杂性, 研发基于不同项目报告模板的知识库与辅助决策模块, 使得中心信息化达到互联互通和无纸化要求, 为实现智慧医疗提供有力的信息平台。

关键词: 信息平台; 互联互通; 无纸化; 智慧医疗

0 引言

随着医疗技术的进步, 分子诊断技术和分子遗传学、基因测序技术等发展, 新型分子基因检测技术已经获得不错的成绩, 在临床中也逐渐被广泛应用^[1]。基因检测技术运用各种分子生物学方法, 并结合生物化学、遗传学、临床医学、信息学等多学科综合研究, 逐渐成为临床疾病分子诊断精确与否的判定依据^[2]。因此, 成立专门的分子诊断与基因检测中心势在必行。中山大学附属第一医院分子诊断与基因检测中心成立于 2019 年, 前身为病理科分子室, 是集分子检测、教学和科研为一体的现代化检测中心, 实验室面积为 800 平米。该中心拥有高素质专业技术队伍和先进的专用仪器设备, 是 PQCC 示范实验室和广东省病理质控中心下属的分子病理诊断示范基地, 也是广东省内率先获得卫生部临检中心认证的临床基因扩增检验实验室之一。

在原来的分子诊断与基因检测系统中, 存在较多的缺陷, 例如: 信息孤岛较多, 申请单系统与收费闭环程度不够, 导致患者已经退费但仍然可以发报告等情况时有发生, 给医院带来不必要的经济损失; 系统无法直接从检查设备获取数据, 基本依赖技师手工录入, 效率较低且容易出错, 随着医院业务的逐年增长, 科室工作压力陡增; 系统缺少必要的知识库以及

辅助决策模块, 使得技师在发报告时对于部分复杂的基因检测报告, 需要人工对照资料以及计算、判断, 容易导致医疗差错, 增加患者医疗风险。因此, 设计一套全业务流程互联互通、数据闭环程度高、具有专业知识库以及辅助决策功能的分子诊断与基因检测信息平台, 具有重要的意义。

1 信息平台整体设计

1.1 信息平台整体架构(图 1)

分子诊断与基因检测信息平台的架构是以医院互联互通平台为基础进行设计的, 整体上可以分为三个层次。最底层是挂号系统、医嘱系统、电子病历系统、物价系统等临床和管理业务系统, 通过 ESB 总线服务跟信息平台进行交互^[3], 为平台提供患者信息、检测项目信息、费用信息等基础数据; 中间层是互联互通交互平台, 其拥有强大的数据交互能力, 通过完善的基础设施、多样的接口支撑以及高效的并发服务和完善的安全体系, 实现数据实时交互、系统实时通讯, 为解决信息孤岛问题打下强有力的基础; 最上层为分子诊断与基因检测平台, 通过下面两层的支持, 实现申请单管理、标本运送、报告发布、知识库管理等具体系统功能, 为临床和管理部门实现质控分析、绩效管理、科研管理、辅助决策等打下良好基础, 同时支

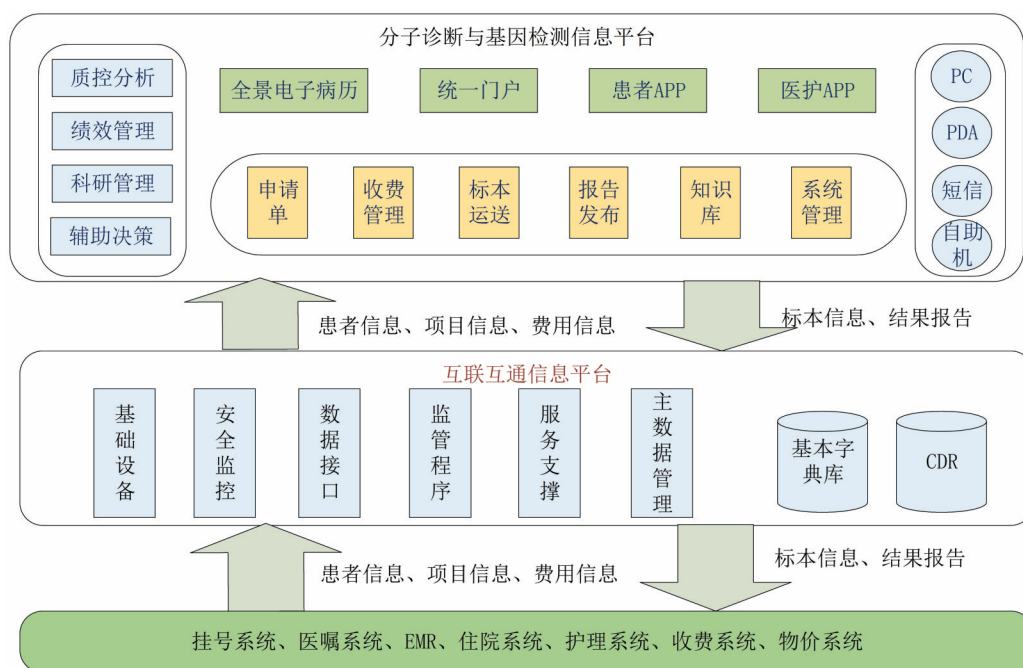


图1 信息平台整体架构图

持电脑PC端、PDA、自助机等多种客户端,方便系统的部署与推广。

1.2 主要业务流程

根据医院管理的要求,结合临床的习惯,对分子诊断与基因检测流程做了进一步的梳理和优化。首先,当医院需要启动新项目时,经医务部门批准后,财务部门在物价系统新增项目,并传到分子诊断与基因检测基础项目维护列表中。项目启用后医生可进行开单,患者执收费单到收费处缴费,护士执行医嘱并采样、打包,护工运送标本至分子诊断与基因检测中心交接,检测中心对标本进行上机、发报告。在此流程中,如果上机前患者进行了退费操作或者未缴费,则系统自动进行拦截并提示未缴费不可进行下一步操作;或者标本已经上机,则不可进行退费申请,除非有特殊原因,需要检测中心管理员权限的用户退检后方可退费,杜绝少收费漏收费的情况。

2 主要功能模块设计

基于互联互通以及业务流程闭环的要求,系统主要功能模块设计为申请单模块、收费模块、标本管理模块、报告管理模块、知识库与

决策支持模块、系统管理模块等,各个模块采用高内聚、低耦合的设计原则,使得模块之间依赖程度降低,并通过接口服务实现模块之间的数据通信。系统主要功能模块设计如图2所示。

2.1 申请单及收费功能

申请单与收费功能分门诊和住院两种情况,在门诊,医生开立检查申请单,选择项目、标本类型、执行科室等内容,并通过接口获取项目费用和耗材费用,形成申请单并传到收费系统。患者执申请单可以到人工窗口缴费,或者在医院APP或者自助机上自行缴费。如需退费,则需打印退费申请单,取消检查申请方可退费成功。住院患者有两种情形,一种是医生下医嘱,护士审核分解自动扣费;另一种情形是护士直接通过护理综合收费补录费用。系统遵循严格的数据闭环流程,例如未缴费或已退费的项目,不能执行打印条码、标本运送等流程,如果标本已经上机,则不能再进行退费申请,除非进行退检方可退费。

2.2 标本管理功能

标本管理包括标本采集、打包、运送、接收、上机等功能。系统支持通过扫描枪等设备

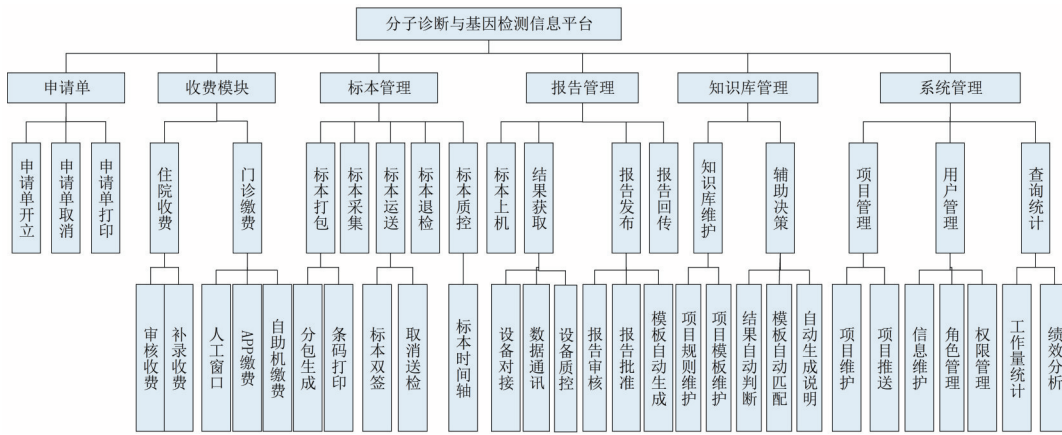


图2 分子诊断与基因检测信息平台功能模块

识别条形码并传至系统^[4]。护士采集标本后扫描条码，记录采集时间点、采集人，采集完成后统一进行标本登记、打包，系统针对设定标本包容量自动进行分包，同时记录打包时间、打包人、运送地点等信息。护工必须跟打包护士进行双签后方可取走标本包，送到接收地点后再与接收人员进行双签，确保对标本运送全程有迹可循，避免标本丢失以及进行标本追踪。系统记录标本全流程的各个重要时间节点，形成标本数据库，为后续质控分析与效率统计提供条件。

2.3 报告管理功能

报告管理包括原始结果获取、报告审核、报告批准、报告上传等功能。系统支持直接从设备采集原始数据，通过采集盒子支持网口、串口等不同的数据接口连入内网，也支持从第三方设备中央站系统获取数据；系统采用统一采集服务实现设备数据的采集和监控，能够实时监控数据采集状态，一旦发现数据缺失，会马上进行报警，避免数据大量丢失。对于报告

发布，必须经过审核以及批准双签，而且审核和批准不能由同一个人完成，否则不予通过。报告发布后，将通过互联互通平台上传到临床业务系统，医护人员可在全景电子病历查看历次检测报告；如果产生危急值，也会将危急值推送至病区并提示医护人员处理。

2.4 知识库管理功能

系统支持知识库维护以及基于该知识库的辅助决策功能。分子诊断与基因检测项目报告具有特殊性，例如常规药物基因组中，每组项目对应的基因有不同的点位，每个点位不同的结果有相应的结果说明；另外对于他汀、华法林、氯吡格雷等药物基因组，需要根据不同的规则对各个点位的结果进行综合判定。因此知识库必须支持维护不同类别、不同规则的基因项目，从而在得到结果后自动根据规则计算判定最终结果，并且匹配相应的结果说明及报告格式。知识库后台的数据库界面如图3所示。

项目编号	项目名称	结果判断规则
787300	罗库溴铵剂型，疗效，药代动力学预测	IF[06408]>TC:该受检者ABC81(T1236C)基因型为TT型，相比TC和CC基因型，TT型患者在相同初始和维持剂量下，麻醉起效时间长，恢复时间延长。IF[06408]...
787400	司琼类药物疗效和毒副作用	IF[06410]>CT:该受检者ABC81(C3435T)基因型为TT，与TC或CC基因型患者相比，TT基因型患者更能减少术后恶心与呕吐。IF[06410]>CT:该受检者ABC81(C343...
787600	氯巴占疗效药代动力学预测	IF[06412]<1/2;06413;IF[06413]<1/3;06412;2/3)
787600	氯巴占疗效药代动力学预测	IF[06414] = "1/1":该受检者CYP2C9基因型综合判型为正常代谢，与其他基因型相比，正常代谢型的癫痫患者的氯巴占代谢增强，不良反应风险降低，但治疗...
788100	丙咪嗪不良反应预测	IF[06416]>GA:该受检者CYP2C19(2G681A)为GG与具有G或AA基因型的患者相比，具有GG基因型的重度抑郁症患者可能具有较低的丙咪嗪血浆浓度。IF[064...
788200	聚乙二醇干扰素α-2a类药物疗效预测	IF[06418]>CT:该受检者IFNL3(rs12979860)位点基因型为TT，TT携带者聚乙二醇干扰素α-2a、聚乙二醇干扰素α-2b和利巴韦林治疗HCV感染的疗效差。IF[06418]...
789100	钙离子通道拮抗剂类药物疗效预测	IF[06422]>GG:该受检者ABC81(G2677T/A)基因型为GG型，应用钙离子通道拮抗剂类药物(硝苯地平、尼群地平、氨氯地平、非洛地平)治疗时，GG基因型...
789100	ARB类药物疗效预测	IF[06422]>CT:该受检者ABC81(C3435T)基因型为TT，应用钙离子通道拮抗剂类药物(硝苯地平、尼群地平、氨氯地平、非洛地平)治疗时，TT基因型患者相...
789300	阿替洛尔类药物疗效及不良反应预测	IF[06426]>CT:该受检者CYP2C9(3)基因型为CC:该患者使用氯沙坦，有效药物浓度明显降低，降压效果明显减弱，不建议使用；或使用缬沙坦、厄贝沙坦的...
789300	阿替洛尔类药物疗效及不良反应预测	IF[06428]>AC:该受检者CYP3A5(3)基因型为3/3(CCI):代谢茶碱的活性是CA和AA型的1/3，使用茶碱马尿酸时，肝毒性风险升高。IF[06428]>AC:该受检者CYP...
783900	放疗副作用(不良反应)预测	IF[06256]>AC:该受检者TANC1(rs10497203)基因型为CC型。接受放射治疗的CC基因型患者具有增加的晚期不良反应风险。IF[06256]>AC:该受检者TANC1(rs104...
784200	甲氧蝶呤药物剂量，疗效及不良反应预测	IF[06269]>CT:该受检者ABC81(C3435T)基因型为TT型，使用甲氧蝶呤治疗的TT基因型和淋巴瘤或白血病患者具有较大的药物浓度毒性风险增加。IF[06269]>C...
784400	顺铂不良反应预测	IF[06276]>GA:该受检者TPMT(3C)INM_000367.3.c.719A>G基因型为AA型，为正常代谢者。具有AA基因型的患者接受顺铂治疗时，发生听力丧失的风险正常。IF...
782400	他汀类药物不良反应预测	IF[06163]<1类:该受检者SLCO1B1基因型为1类，肌病和横纹肌溶解症风险显著增加，他汀类药物耐受剂量最低。IF[06163]>1类:该受检者SLCO1B1基因型为1...
5601	P450(CYP 2C19)基因检测	IF[04308] = "1/1" (636GG,681GG):酶活性高，对药物代谢速度快。IF[04308] in ("1/2" (636GG,681GAY), "1/3" (636GA,681GG)):酶活性中等，对药物代谢速度...
5608	乙酰胆碱酯酶(AChE)基因检测	IF[04314] = "Glu504Glu":Glu504Glu阳性；对酒精代谢能力差，可用硝酸甘油作为急救药物。IF[04314] = "Glu504Lys":Glu504Lys阳性；对酒精代谢能力差，慎用硝...

图3 知识库后台数据库界面

2.5 系统管理功能

系统管理包括项目维护、用户管理、查询统计等功能。在字典维护中,可同步物价系统的收费项目,并且设置对应的项目明细,同时推送给医嘱系统;用户管理包括用户角色管理、用户账户管理以及人员信息维护;统计查询包含工作量统计、质控数据查询、科研数据统计、绩效统计等数据查询和导出功能。报表维护除了能自定义字段外,还支持模块配置,在新建报表时进行模板套用和保存^[5]。

3 系统开发实现

3.1 程序开发

系统采用面向对象开发语言C#进行开发,开发环境为Visual Studio,采用C/S架构。服务器采用Windows Server操作系统,接口开发采用

Web Service技术并部署在IIS上。报告模板采用水晶报表(Crystal Report)进行设计,利用水晶报表强大的展示功能,能够根据不同的分子诊断与基因检测项目设计相应的模板,目前已经有16种报告模板并实现自动匹配。系统主要运行界面如图4所示。

3.2 数据库设计

系统采用Microsoft SQL-sever数据库,利用其强大的T-SQL语言,能够进行系统的快速开发。系统核心数据表包括患者信息表(Patient)、申请单信息主表(MastSurvey)、申请单信息明细表(Mastest)、标本信息主表(MastHead)、标本信息明细表(SpecimenTestItem)、结果信息表(Mastresult)、报告信息表(Mastrbdt)等,其实体-关系(E-R)图如图5所示。

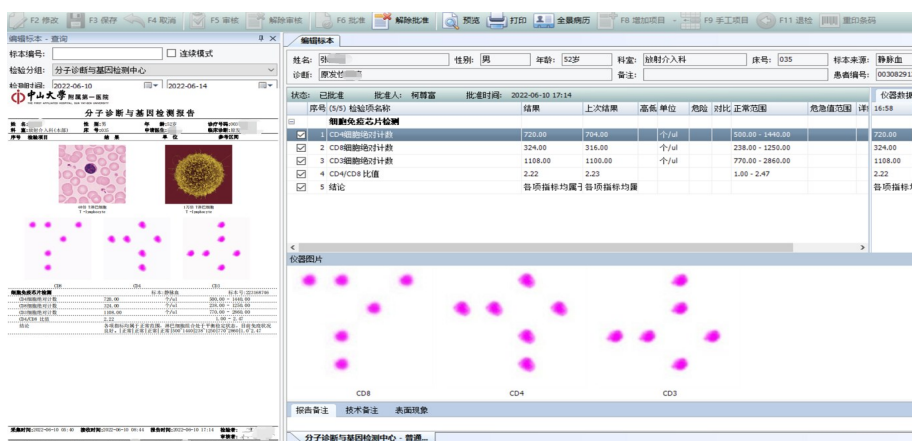


图4 结果审核以及报告预览界面

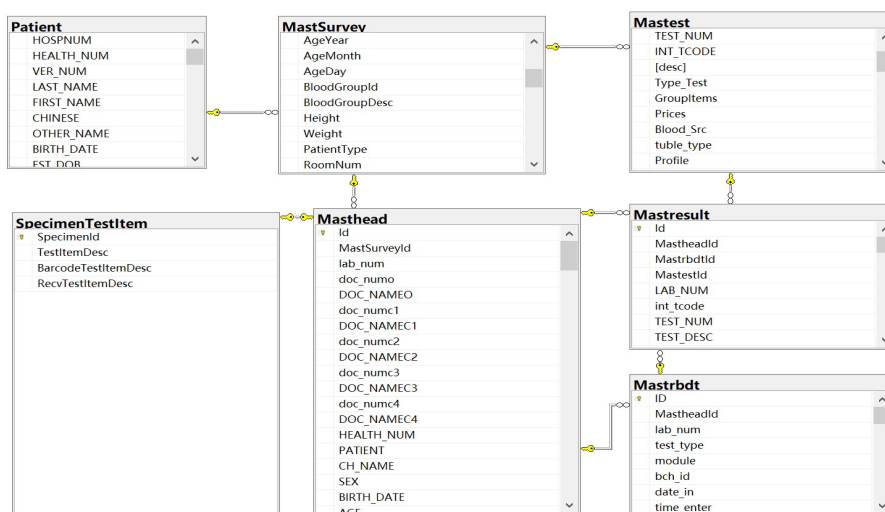


图5 核心数据表实体-关系图

4 建设效果

分子诊断与基因检测信息平台经过近一年的开发实施以及后期完善,先后完成科室培训、仪器接入、系统功能调试等工作,并在2021年11月正式上线运行,目前科室人员操作熟练,系统稳定运行,系统功能满足业务需求。自上线以来,共开立申请单8471例,其中住院6563例,门诊1634例,体检274例,开单数量前五的科室分别为:心内三科(833例)、呼吸与危重症医学科一区(712例)、放射治疗科(547例)、神经二科(438例)、肿瘤科(384例)。通过互联互通平台的上传,将检查报告推送到医护患各个客户端,使得医护人员在病房可以通过全景电子病历查看历次检查报告,患者在医院APP上也可以第一时间知道检测结果,提高患者就诊满意度。

5 结语

分子诊断与基因检测信息平台的建设,一方面实现数据的闭环,减少费用错收漏收问题,同时实现设备数据自动采集以及知识库和辅助决策应用,提高了临床医务人员的工作效率,减少医疗差错,有效地支撑了医院业务的开展。同时与信息平台的对接,有助于形成临床数据库,为科研、质控、管理提供数据基础,为建

设无纸化医院以及智慧医疗提供助力。

参考文献:

- [1] 常新剑,任俊宏.分子基因组检测方法在罕见病分子诊断中的应用及其意义[J].全科口腔医学电子杂志,2019,6(24):171.
- [2] 侯英勇,钱梦佳,程韵枫,等.基因检测在精准医疗中的应用与管理[J].中国临床医学,2022,29(1):7-10.
- [3] 莫远明,王毅,刘小洲,等.基于互操作信息平台的院区检验信息平台的构建[J].现代医院,2020,20(3):407-410.
- [4] 林思阳.基于健康信息交换平台的院区标本运送系统设计与应用[J].中国医院统计,2021,28(4):346-350.
- [5] 莫远明,王毅,林琳,等.基于数据集成平台的医院智能报表系统的构建[J].现代医院,2020,20(4):568-571.

作者简介:

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

卢瑞鸿(1988—),男,广东梅州人,本科,工程师,研究方向为医疗信息化

陈康(1973—),男,海南琼海人,本科,工程师,研究方向为医疗信息化

收稿日期:2022-05-20 修稿日期:2022-06-16

Construction of Clinical Molecular Diagnosis and Gene Detection Information Platform

Mo Yuanming, Lu Ruihong, Chen Kang

(Information data center of the First Affiliated Hospital of Sun Yat-sen University, Guangzhou 510080)

Abstract: In order to improve the work efficiency and fine management level of the molecular diagnosis and genetic testing center, improve the project report management process, and eliminate the situation of overcharging and missing fees, an information system based on the interconnected data platform is designed to realize the data closed loop of the whole business process. At the same time, in view of the complexity of the project report, the knowledge base and auxiliary decision-making module based on different project report templates are developed, Make the center informatization meet the requirements of interconnection and paperless, and provide a powerful information platform for the realization of intelligent medical care.

Keywords: information platform; interconnection and interworking; paperless; smart medicine

文章编号: 1007-1423(2022)19-0111-05

DOI: 10.3969/j.issn.1007-1423.2022.19.020

基于 Unity 3D 酷跑游戏的设计与实现

李爱军

(苏州职业大学计算机工程学院, 苏州 215104)

摘要: 酷跑游戏类似于其他的酷跑类游戏, 有简单的操作、较低的技术性和真实的游戏画面等特点。目前, 酷跑游戏实现的操作是基本的运动功能, 包括自动奔跑、跳跃、翻滚。道具系统有双倍积分、二次跳跃、加速奔跑、吸铁石等动能。游戏设计操作简单, 界面清晰, 即使是新玩家也能很容易熟悉游戏方法。

关键词: 手机软件; 游戏; 酷跑游戏; Unity 3D

0 引言

主要游戏方式便是酷跑。以一个成功的酷跑游戏《神庙逃亡》为例, 我们可知酷跑类游戏的主要功能。玩家不能控制角色的移动速度, 随着游戏时间的推移, 角色移动的速度越来越快, 玩家操作的时间越来越短, 考验玩家的反应能力来达到增加游戏难度的目的。

1 游戏主要设计

1.1 游戏类型定位

游戏有很多种类, 主要有策略游戏、动作游戏、竞速游戏、卡牌游戏、冒险游戏、体育游戏、休闲游戏、音乐游戏、模拟游戏、射击游戏等。

本游戏为酷跑游戏, 应归为竞速游戏或休闲游戏。

1.2 运行环境

本游戏的运行环境为手机 Android 6.0 及以上系统版本, 它的优点是游戏界面非常简洁, 且占用的内存资源少, 操作起来很方便。

1.3 游戏玩法

本游戏主要靠滑动屏幕控制角色移动来操作游戏, 控制角色躲避障碍拾取金币。游戏地图分为三个跑道, 向左滑为向左跑道移动, 向右滑为向右跑道移动, 向上滑为跳跃, 向下滑为滚动, 当角色在空中时向下滑动可以使角色

快速落地滚动。

1.4 游戏流程

玩家运行主程序进入游戏画面, 出现开始菜单。开始菜单包含两个按钮, 开始和退出。点击开始按钮, 游戏开始运行, 点击退出按钮, 退出主程序。开始游戏后弹出暂停按钮, 点击暂停按钮游戏暂停, 出现暂停菜单。暂停菜单包含两个按钮, 返回游戏和退出游戏。点击返回游戏, 游戏继续; 点击退出按钮, 游戏退出。当游戏角色死亡时, 出现结束菜单。结束菜单包含两个按钮, 再次开始游戏和退出游戏。点击再次开始游戏, 游戏重新开始; 点击退出游戏, 退出主程序。

1.5 开发技术

本游戏是采用 3D 虚拟技术和游戏物理引擎进行开发的, 游戏的现场布景具有很强的真实感、立体感以及逼真的光影效果, 用户能够在游戏中感受到绚丽多彩且富有冲击感的感官体验。

2 游戏界面搭建

2.1 酷跑角色与场景搭建

2.1.1 酷跑角色与场景搭建

在 Scene 界面通过导入模型, 拖拽搭建等步骤完成场景的初步搭建。

2.1.2 金币组

将8个金币设为一组，布置在跑道上。

2.1.3 障碍物与碰撞监测器

障碍物有木板、小汽车、大巴，分别添加碰撞检测器。

2.1.4 道具

将道具模型导入，制作碰撞动画粒子特效。

2.1.5 游戏文本菜单

游戏菜单用来开始或退出游戏以控制游戏进程，右侧为道具buff剩余时间，右下角为音乐控制按钮，左上角为收集的金币数。

2.2 游戏逻辑实现

2.2.1 道路设置

游戏只有一个游戏场景，为一条无限长的马路。通过交替移动两段足够长的道路实现“无限”道路的功能。

```
void Update()
{
    if (transform.position.z > floorOnRunning.transform.position.z + 32)
    {
        RemoveItem(floorOnRunning);
        AddItem(floorOnRunning);
        floorOnRunning.transform.position = new Vector3(0, 0, floorOnRunning.transform.position.z + 32);
        GameObject temp = floorOnRunning;
        floorOnRunning = floorForward;
        floorForward = temp;
    }
}
```

2.2.2 角色奔跑

作为一个酷跑类游戏，首先要让游戏主角跑起来，我们通过在主角的模型上挂载一个主角控制脚本来实现。主角的控制脚本名为 PlayerController。

通过在z轴正方向给主角添加一个正值，使主角可以沿z轴正方向匀速跑动。然后在y轴负方向，也就是默认的地面方向添加正值，使主角可以稳定在地面上，不会违背物理常识。这里没有选用unity3d自带的物理系统，因为其系统在应用时发生明显的错误效果，所以进行了变量的赋值使其实现较真实的物理效果。

游戏主角以默认速度奔跑：

```
moveDirection.z = speed;
moveDirection.y -= gravity * Time.deltaTime;
characterController.Move((xDirection * 10 + moveDirection) * Time.deltaTime);
```

酷跑类游戏在游戏过程中会通过缓慢增加主角的奔跑速度来减少躲避障碍物的反应时间，从而起到加大游戏难度的效果。本游戏中主角每奔跑一段固定长度的路程时增加一点奔跑速度，同时为了防止某些实力过于强的玩家在游戏中长时间不失败导致的主角奔跑速度过高，游戏中还为主角的奔跑速度设置了最大值，使其速度不能无限制增长。

限制游戏主角的最大奔跑速度：

```
private void SetSpeed(float newSpeed)
{
    if (newSpeed <= maxSpeed)
        speed = newSpeed;
    else
        speed = maxSpeed;
}
```

2.2.3 滑动检测

本游戏的操作非常简单，用指针在屏幕上进行上下左右的滑动即可控住主角的移动，一共设置了4个操作检测，分别为“向上”“向下”“向左”“向右”。

首先创建一个枚举类来存放上下左右空5种滑动状态。

滑动方向枚举类：

```
public enum InputDirection
{
    NULL,
    Left,
    Right,
    Up,
    Down
}
```

当玩家没有触摸屏幕时，屏幕没有检测到指针，滑动状态默认为空值，即游戏主角以默认速度匀速奔跑。当玩家向上滑动屏幕时，屏幕检测到指针，将状态更新为“开始滑动”并记录当前的指针位置。

当指针以某一个反向持续滑动20个单位时，记录此位置与开始滑动位置时的向量，并将其转化为角度，根据角度分析此次滑动的方向，同时将滑动状态更新为“未开始滑动”。

2.2.4 玩家移动

当游戏中检测到玩家的控制输入后，角色要根据控制指令进行相应的移动。当玩家向左滑动时，如果角色不在最左边的道路上，那么角色向左移动一格。当玩家向右滑动时，如果角色不在最右边的道路上，那么角色向右移动一格。

当玩家向上滑动时，角色进行一次跳跃。通过给角色y轴正方向添加正值来改变角色的竖直方向的位置。

当玩家向下滑动时，如果角色此时在地面上，那么角色进行一次翻滚，如果角色测试在跳跃中，那么角色迅速落到地面。与跳跃类似，通过给角色y轴负方向添加正值来改变角色在竖直方向的位置。

翻滚动作代码：

```
if(inputDirection == InputDirection.Down)
{ AnimationManager.instance.animationHandler = AnimationManager.instance.PlayRoll; }
```

当主角已经到达最左或最右的位置时不能再往左或往右移动，需要限制主角的移动以免跑出地图。

2.2.5 道具效果

本游戏一共设置了4种道具，分别为磁铁道具、加速道具、双倍积分道具和二连跳道具。

磁铁道具的效果为：角色自动吸取道路上的金币，持续一段时间。实现原理为，为角色添加一个大小等于磁铁作用范围的碰撞检测器，当碰撞检测器碰撞到金币时，触发金币的事件。拾取到磁铁道具时激活该碰撞检测器，持续时间结束时自动关闭。

二连跳道具效果为，玩家可以在跳跃过程中再次跳跃，当角色拾取到二连跳道具时，可以在道具持续时间内进行二次跳跃。

二连跳道具代码：

```
void JumpDouble()
{ AnimationManager.instance.animationHandler = AnimationManager.instance.PlayDoubleJump;
  moveDirection.y += jumpValue * 1.3f; }
```

双倍积分道具的效果为：将拾取金币时获

得的积分翻倍，持续一段时间。实现原理为：增加积分的数值。

加速道具的效果为：大幅度提升角色的速度，持续一段时间。实现原理：增加奔跑速度的数值。

2.2.6 相机跟随

本游戏为第三人称视角游戏，为了使角色始终处于画面中央，需要将显示游戏画面的相机对象设置为始终跟随角色移动，位置处于角色身后的斜上方。

相机跟随效果代码：

```
void LateUpdate()
{ if (! isShaking
  && GameController.instance.isPlaying
  && ! GameController.instance.isPause)
{ pos.x = target.transform.position.x;
  pos.y = Mathf.Lerp (pos.y, target.transform.position.y + height, Time.deltaTime * 5);
  pos.z = target.transform.position.z - distance;
  transform.position = pos; }
```

2.2.7 碰撞检测

碰撞检测是非常重要的功能，游戏内几乎所有的物体交互都要依靠碰撞检测来触发。在游戏制作中，给需要碰撞检测功能的物体添加碰撞检测器组件boxcolliger便可实现碰撞检测。

当角色触碰到4个具有特殊效果的道具时，激活特殊效果，销毁道具，开启效果持续时间计时器。

当角色触碰到金币时，销毁金币，增加玩家的积分。如果角色此时拥有磁铁道具，那么当代表磁铁效果返回的碰撞检测器触碰到金币时，先使金币平滑地飞向角色，制作磁铁吸引的效果，当金币触碰到角色时，金币销毁，增加玩家的积分。

当角色触碰到障碍物时，角色死亡，执行死亡动画。

障碍物碰撞。障碍物用来阻挡游戏角色的前进，并且在游戏角色碰撞到障碍物时结束游戏。对于大部分障碍物，都可以直接添加碰撞

器使其在碰撞时触发死亡效果。有两种障碍物是特例，一种是车辆，游戏中的两种车辆在玩家碰撞到时会触发死亡效果，但是当游戏角色跳跃到车的顶部时，是可以在上面行走的，所以在添加碰撞器时要注意这一点；而另一种是下方有空间的挡板，游戏角色在没有二连跳时无法越过，只能通过从下方空间翻滚经过，在游戏制作中，翻滚动作虽然会播放翻滚动画，但是碰撞器并不会播放翻滚动画，所以要在播放翻滚动画时短暂地关闭碰撞器，防止碰撞到障碍物，在经过障碍物后再次将碰撞器开启。

2.2.8 路面随机生成障碍物

酷跑类游戏通常没有通关功能，所以只要玩家可以保持游戏不失败，就可以无限进行游戏。而前面讲到本游戏实现无限道路的方法为交替移动两块足够长的道路，玩家需要在道路上使用道具或躲避障碍物。为了增加游戏的趣味性，游戏制作中需要在道路上随机生成各种障碍物或道具。但是在制作过程中发现，如果道具的生成完全随机化，那么会出现大量的物品重叠或没有通路等问题。所以为了保证在相对随机的环境下也能正常游戏，解决办法为手动制作多组道路上的障碍物排列，每当角色移动到新的道路时，随机选取一组预先制作好的障碍物排列。

2.2.9 界面显示

除了游戏主要的环境和角色以外，还需要多个UI界面来显示一些游戏内数值。本游戏设置了积分显示，道具持续时间显示。积分显示为记录玩家拾取金币的数值，普通情况下玩家每拾取1个金币，积分+1，在双倍积分道具效果持续时间内，玩家每拾取一个金币，积分+2。道具持续时间显示为记录玩家当前已激活道具的效果剩余时间，当玩家触碰到道具时开始计时，当计时时间结束时道具效果结束。

金币积分计数代码：

```
public override void HitItem()
```

```
{ base.HitItem();  
  GameAttribute.instance.AddCoin(1); }
```

2.2.10 游戏进程控制

所有单机都必不可少地设置开始暂停等游戏进程控制。本游戏一共设置了开始游戏、暂停游戏、恢复游戏、重新开始游戏和退出游戏五个选项。第一次打开游戏时弹出开始游戏和退出游戏。点击开始游戏按钮，游戏开始并弹出暂停按钮；点击退出游戏按钮，游戏关闭；点击暂停按钮，游戏暂停并弹出恢复游戏按钮和退出按钮；点击恢复游戏按钮，游戏继续；点击退出游戏按钮，游戏关闭。当角色死亡时，弹出重新开始游戏按钮和退出游戏按钮。点击重新开始按钮，游戏重新开始，开始界面与第一次进入时一致；点击退出游戏按钮，游戏退出。通过封装方法到按键，由点击事件激活。

当游戏重新开始时，需要将一切数值清空，包括玩家已获得的积分，当前已激活道具的剩余时间等，还要重置角色的位置和道路的位置。

2.2.11 动画与声音播放

游戏过程中角色需要切换不同的动画，正常跑步时播放跑步动画，向左或向右移动时播放移动动画，跳跃时播放跳跃动画，翻滚时播放翻滚动画，碰撞障碍物时播放倒地动画。

本游戏一共有五种不同的音效，分别为菜单键点击音效，金币拾取音效，触发道具音效，碰撞障碍物音效和滑动操作音效。

3 游戏测试

在游戏制作完成后，需要进行游戏测试，用来测试游戏中是否有与预期不符的异常。因为Unity3D引擎可以实时测试的优点，我们可以在游戏发布前先进行初步的流程测试。

首先测试游戏主角是否能根据玩家的操作进行正确的动作，然后是游戏中的4个道具是否可以正确地激发效果。同时，道具在激活时，游戏界面右上角的道具效果计时界面也要正确

地为道具效果计时，在倒计时结束时道具效果也要同时消失。然后是游戏主角在拾取金币时是否能正确地记录金币数，在双倍积分效果的持续时间内拾取金币是否可以正确地增加双倍金币数。还有游戏场景内的各个障碍物测试，不同的障碍物拥有不同的碰撞区域，游戏角色在碰撞到障碍物时能否顺利触发死亡效果。最后是游戏菜单，对所有的游戏菜单按键测试，例如开始、暂停、退出等按键能否正确地实现。

4 结语

游戏无论是在汽车的选择、关卡的选择还是在音效这一感官体验上，都能给玩家带来不一样的视觉感受与听觉感受。主菜单界面的帮助界面能够帮助新手玩家更好地了解本游戏的

玩法，选择不同的车辆进行游戏会给玩家带来焕然一新的感觉，不同的关卡具有不同的难易程度，能够适合不同的玩家群体，震撼的游戏背景音乐能够带来刺激的听觉享受。

参考文献：

- [1] 宣雨松. Unity3D 游戏开发[M]. 2版. 北京: 人民邮电出版社, 2018.
- [2] 马遥, 陈虹松. Unity 3D 完全自学教程[M]. 北京: 电子工业出版社, 2018.

作者简介：

李爱军(1973—), 男, 甘肃高台人, 硕士, 讲师, 教师, 研究方向为面向程序对象设计、数据库技、IOS 开发技术

收稿日期: 2022-06-05 修稿日期: 2022-07-27

Design and Implementation of Virtual Parking Lot Based on Unity 3D

Li Aijun

(School of Computer Engineering, Suzhou Vocational University, Suzhou 215104)

Abstract: The parkour game introduced in this article is similar to other parkour games, with simple operation, low technology and real game pictures. At present, parkour games to achieve the operation is the basic movement function, including automatic running, jumping, rolling. The prop system has double integral, double jump, accelerated running, magnet and other kinetic energy. The game design is simple to operate, the interface is clear, even new players can be easily familiar with the game method.

Keywords: mobile software; game; parkour game; Unity 3D

文章编号: 1007-1423(2022)19-0116-05

DOI: 10.3969/j.issn.1007-1423.2022.19.021

地方高校班级管理系统的设计与实现

刘思云, 陈贵松

(岭南师范学院电子与电气工程学院, 湛江 524048)

摘要: 人才培养方案修订是为了培养优秀学生, 成为工作上的佼佼者, 而实际上在校优秀和在职优秀会出现一些“倒挂”现象。班级管理系统通过记录学生的重要在校成长轨迹, 来帮助梳理学生过去在校表现与未来工作成长之间的关系, 为统计在校优秀和在职优秀的学生分布情况积累大量基础数据, 从而为人才培养方案修订提供重要参考价值, 更有利于培养适合于社会需求的人才。

关键词: 班级管理; 地方高校; 人才培养方案修订

基金项目: 广东省高等教育教学改革项目(电气信息类“一流课程”新模式的改革与实践)(LSJGZH1903)

0 引言

人才培养方案修订是高校的一项重要基础工作^[1]。对于全国重点高校来说, 因为学校资源丰富、生源佳、学生自主学习能力强, 人才培养方案只要有一定框架, 不太与社会时代脱节, 学生毕业都会有一定的竞争力。而对于地方高校, 由于招生范围有一定局限, 加上学校资源匮乏, 学生自律意识较弱, 想让学生有一定可媲美的竞争力, 人才培养方案就显得更为重要了, 必须具有一定的针对性和可执行性。

地方高校人才培养方案调整一般是 2~3 年修订一次, 修订时需要先进行调研论证。调研主要是找几家用人单位谈谈自己的需求和对自己毕业学生的评价, 或者联系近几年已毕业的学生谈谈他们自己对学校人才培养的看法, 然后以此为基础信息和依据, 进行人才培养方案修订。由于此法采集的数据样本少, 有一定的依赖性, 且个人看法不一定具有代表性和典型性, 也谈不上前瞻性, 以此为支撑调整人才培养方案必然具有很大的局限性和非适配性。

人才培养根本目标是鼓励学生成为社会所需人才, 成为可能为社会创造最大价值的人。但事实上, 经常发现一些优秀学生工作后并没

能很好地获得职业成长, 一些表现平平的学生反而有很好的职业表现, 从而出现在校优秀和工作优秀的“倒挂”现象。企业人才需求已经在挑战大学人才培养^[2], 但由于缺乏具体数据支持, 无法进行进一步的统计和分析。

如果学习优秀无法转换为工作优秀, 那么为什么还要让学生追求学习优秀呢? 凭什么让大家去主动争取学习优秀呢? 这里面是不是培养方案(包括学校对学生的评价机制)出了问题呢? 学习成绩一般的人, 如果工作表现十分突出, 是不是也反证培养方案有一定瑕疵呢? 一个人人都不愿意去争取优秀的学校, 这会是一个好学校吗? 这样的学校能不断发展吗?

针对上述问题, 如果能从学生的在校表现与工作表现之间找到某些联系, 必然会对培养方案的科学修订有很大的帮助和借鉴价值。

本文设计的班级管理系统能记录班级每个学生的基本信息、学业表现、社会实践情况, 并追踪其毕业去向和表现, 可勾画学生在大学时期的成长轨迹, 通过积累大量基础数据样本, 帮助梳理学生过去在校表现与未来工作成长之间的关系, 从而为人才培养方案修订提供重要的参考价值。

1 系统设计

1.1 架构设计

本系统使用 Java Web 开发架构,采用 Spring + SpringMVC + Mybatis (SSM) 框架^[3],如图1所示。

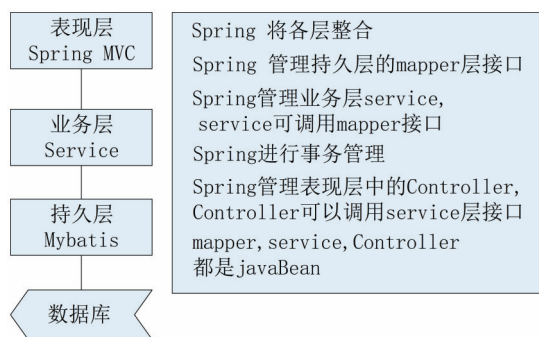


图1 系统架构图

1.2 功能设计

本系统功能设计分为前台与后台,用户分别是学生和管理员。前台分五个功能模块,包括学生基本信息、学业信息、社会实践、毕业去向和学生修改个人密码。后台分五个功能模块,包括学生基本信息管理、学业信息管理、社会实践管理、毕业去向管理,及修改学生和和管理员密码,整个系统功能框架如图2所示。

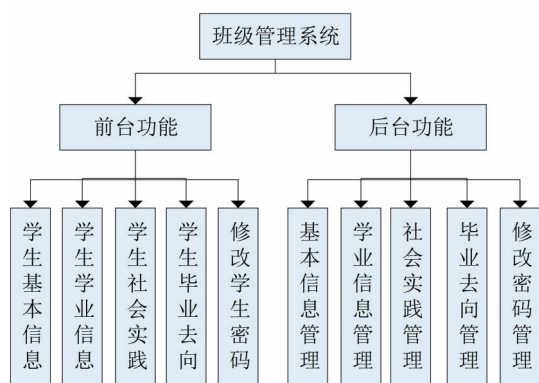


图2 系统功能框架图

1.2.1 前台设计

前台模块主要面向学生,包括学生基本信息、学生学业信息、学生社会实践、学生毕业

去向和修改学生个人密码,如图3所示。

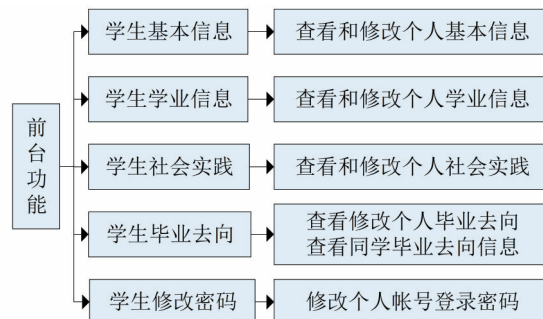


图3 前台功能模块图

1.2.2 后台设计

后台模块面向对象是管理员,包括学生基本信息管理、学业信息管理、社会实践管理、毕业去向管理、修改学生和和管理员密码,如图4所示。



图4 后台功能模块图

1.3 数据库表设计

本系统主要用以下四个数据库表存储相关数据。

(1) 基本信息表(s_basic_information), 主要用于保存学生基本信息,如学号、姓名、性别、生日、生源、联系方式、账号密码和系统权限等。

(2) 学业信息表(s_academic_information), 主要用于保存学生学业信息,如学业绩点、考级考证、学业奖励和课程补考等。

(3) 社会实践表(social_pracite), 主要用于保存学生社会实践情况,如社团职务、竞赛获

奖、实习情况和其他荣誉等。

(4) 毕业去向表(graduation_to)，主要用于保存学生毕业有关信息，如毕业设计题目、就业单位、工作地址和单位联系方式等。

2 系统实现

用户登录系统时需要输入帐号、密码和选择身份，登录界面如图5所示。系统账号分两种，一种是学生账号，为班级名+学号，系统默认密码为LNNU1111，登录后可查看个人信息和同学部分信息，也可修改个人信息和密码；另一种是管理员账号，具有整个系统管理权限，可查看班级所有同学信息，可修改管理员和学生密码。以下对部分功能实现进行阐述。



图5 用户登录界面图

2.1 前台主要功能实现

2.1.1 基本信息模块实现

学生登录帐号后，首页是基本信息模块，可以查看个人基本信息，包括：学号、姓名、性别、生日、照片、生源和联系方式，如图6所示。点击修改，可修改个人基本信息，如图7所示。

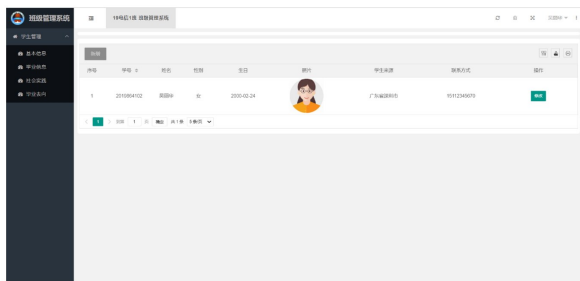


图6 学生基本信息查看界面图

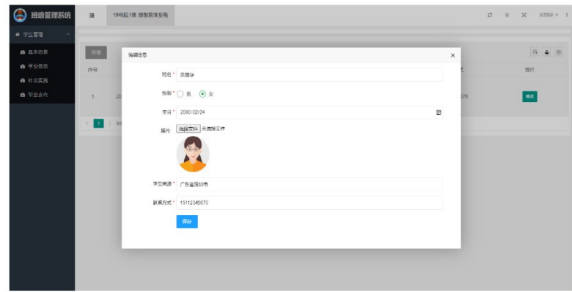


图7 学生基本信息修改界面图

2.1.2 学业信息模块实现

该模块记录了学生在校期间的学业表现，包括学业绩点、考级考证、学业奖励和课程补考情况等，如图8所示。学生查看个人学业信息后，根据个人在校学习进展，可随时对学业信息进行修改，如图9所示。



图8 学业信息查看界面图

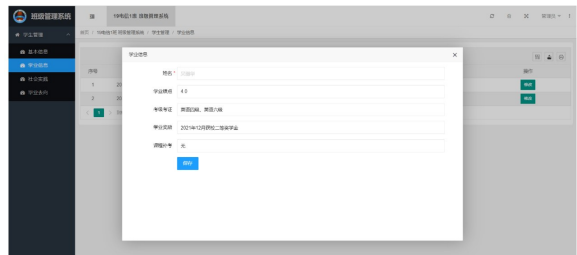


图9 学生学业信息修改界面图

2.1.3 社会实践模块实现

该模块记录了学生在校期间参加社会实践情况，包括担任社团职务、竞赛获奖、实习情况和其他荣誉等，如图10所示。点击修改，可更新社会实践信息。



图10 社会实践查看界面图

2.1.4 毕业去向模块实现

该模块记录了学生毕业设计和工作去向信息,包括毕业设计题目、就业单位、工作地址、单位联系电话等,如图11所示。学生可查看其他同学的有关此模块信息,但只能修改个人的有关此模块内容。



图 11 学生毕业去向查看界面图

2.2.2 学业信息管理实现

此模块管理员可查看修改所有学生在校学业信息,如学业绩点、考级考证情况、学业奖励、课程补考情况等,如图14所示。

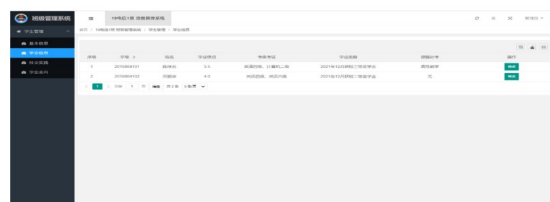


图 14 学业信息管理查看界面图

2.2 后台主要功能实现

管理员直接在数据库表中插入管理员账号和密码,便可使用管理员账号和密码登录系统后台,登录后默认跳转到学生信息管理功能页面,即可对学生信息进行查看修改等操作。

2.2.1 基本信息管理实现

此模块管理员可查看修改所有学生基本信息,如学号、姓名、性别、生日、照片、学生来源和联系方式等,如图12所示。并可修改、增加或删除学生,如图13所示。

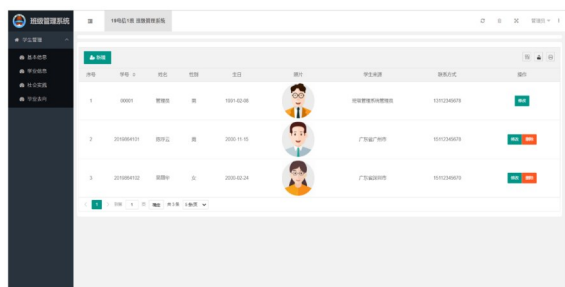


图 12 基本信息管理查看界面图

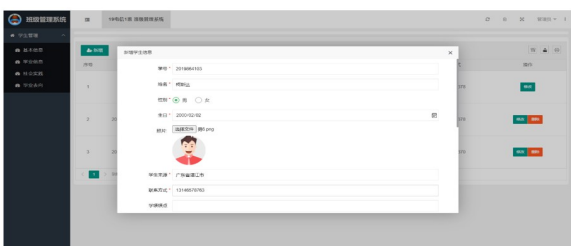


图 13 基本信息管理增加学生界面图

2.2.3 社会实践管理实现

此模块管理员可查看修改所有学生的社会实践情况,如社团职务、竞赛获奖、实习情况、其他荣誉等,如图15所示。



图 15 学生社会实践管理界面图

2.2.4 毕业去向管理实现

此模块管理员可查看修改所有学生毕业设计题目、就业单位、工作地址、单位联系方式等,如图16所示。



图 16 毕业去向管理界面图

2.2.5 修改密码管理实现

为保证系统安全性,管理员可点击页面右上角管理员标签右边三角形,选择修改密码,在弹出密码修改界面后进行有关操作,如图17所示。

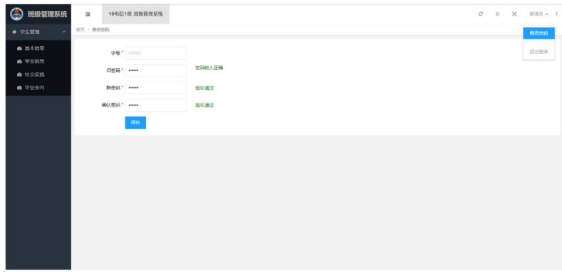


图 17 管理员修改密码界面图

3 结语

此系统能动态记录并更新学生的基本信息、学业信息、社会实践和毕业去向等情况，还可以根据实际需要增添部分功能。此系统能帮助寻求学生的在校表现与工作成长的关联度，为人才培养方案修订积累大量详细的第一手数据，而不是单一地、费时费力地、被动地仅仅依靠访谈调研来获取有限的人才培养数据，借助系统可以主动梳理学生过去在校表现与在职表现

之间的关系，从而更加科学地对人才培养方案进行修订。

参考文献：

- [1] 许彪,李三福.基于“四位一体”目标体系的人才培养方案信息模型构建探索[J].高等职业教育探索,2022(3):67-72.
- [2] 袁靖宇.高校人才培养方案修订的若干问题[J].中国高教研究,2019(2):6-9.
- [3] 李唯.基于SpringBoot+Mybatis的驾校预约系统设计与实现[J].电脑编程技巧与维护,2022(3):10-12.

作者简介：

刘思云(1976—)男,湖南长沙人,硕士研究生,讲师,研究方向为智能系统应用

陈贵松(2000—)男,广东揭阳人,本科生,研究方向为信息管理系统应用

收稿日期:2022-05-14 修稿日期:2022-06-19

Design and Implementation of Class Management System in Local Universities

Liu Siyun, Chen Guisong

(School of Electronic and Electrical Engineering, Lingnan Normal University, Zhanjiang 524048)

Abstract: The purpose of revising the talent training plan is to cultivate excellent students and become outstanding in work. In fact, there will be some "upside down" phenomenon between excellent students in school and in-service. By recording students' important growth tracks in school, the class management system can help sort out the relationship between students' past performance in school and future work growth, and accumulate a large amount of basic data for statistics on the distribution of excellent students in school and in-service, so as to provide important reference value for the revision of talent training programs, which is more conducive to cultivating talents suitable for the society

Keywords: class management; local universities; revision of talent training program