

M 现代计算机

XIANDAI JISUANJI

第28卷第16期（总第760期）

半月刊（1984年创刊）

2022年8月25日出版

主管单位 中山大学
主办单位 广州中山大学出版社有限公司
出版单位 广东现代计算机杂志社有限公司
发 行 广东省报刊发行局（全国公开发行）
印 刷 广州一龙印刷有限公司
社 长 黄少伟
主 编 石玉珍
编 委 邹岚萍 熊锡源 李 文 石玉珍 梁嘉璐
地 址 广州市海珠区新港西路135号
中山大学内（510275）
电 话 020-84112089（编辑部）
网 址 www.moderncomputer.cn
电子邮箱 tougao@moderncomputer.cn

ISSN 1007-1423
CN 44-1415/TP

邮发代码：46-121
定价：30.00元



邮局订刊二维码



现代计算机
官方网站二维码

ISSN 1007-1423



9 771007 142222 16>

M 现代计算机

第28卷 第16期（总第760期）

2022年8月

2022年8月 第28卷

第16期（总第760期）



ISSN 1007-1423
CN 44-1415/TP

MODERN COMPUTER

现代计算机



中山大学出版社 主办

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

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

版权声明

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

目次

研究与开发

- 优化预训练模型的小语料中文文本分类方法····· 陈 蓝, 杨 帆, 曾 楨 (1)
- 异常数据实时检测方法研究综述····· 吴 琼, 李永飞, 李铭洋 (9)
- 基于藏传佛教纹饰符号构造纹理的信息隐藏方法····· 吴秀琪, 张春玉, 丁泽威, 郭晨飞 (16)
- 基于注意力机制 BiLSTM-CharCNN 的药物不良反应监测方法····· 蒯娟霞, 徐 鹏, 叶思维 (22)
- 云端马拉松图像的号牌识别研究····· 黎蕴玉, 丁小波, 蔡茂贞, 钟地秀, 彭 琨 (28)
- 基于电气标准模型的移动端配网图数一体化建模工具设计与应用····· 张旭东, 陈星明 (32)
- 基于集成学习的有害垃圾自动识别方法研究····· 孟德尧, 吴荣海, 杨邓奇 (38)
- 机器学习在心肌梗塞并发症中的预测研究····· 王 蔚, 程 君, 李先杰, 彭 雷 (43)
- 基于边缘提取与多项式拟合的三维位姿解算方法
····· 朱 腾, 郭晓莹, 阎汉生, 梁剑斌, 舒 彬 (48)
- 基于改进的 YOLOv3 算法研究····· 刘 博, 李艳武 (53)
- 5G 网络 NSA 组网共模抗干扰问题的研究····· 陈雪娟, 叶利丽, 楚高峰 (57)

图形图像

- 基于改进 U-Net 模型和 CBAM 的脑肿瘤 MRI 图像分割
····· 张晓倩, 罗 建, 杨 梅, 金芊芊, 朱 熹 (60)
- 一种基于空间预测的运动补偿帧率上转换算法····· 马文丹, 李 凡, 俞文静 (67)

实践与经验

- C4.5 决策树方法在课程教学成绩分析中的应用····· 胡 佳 (73)

开发案例

- 大数据可视化在城市规划中的应用····· 操心慧, 操金金, 许丽娟, 弓佳明, 赵 茜 (77)
- 基于 RFID 和树莓派的物品收纳管理系统····· 卢依思, 盘茂杰, 郑思晴, 梁焕诗, 魏纵横 (84)
- 基于激光雷达的移动配送机器人自主导航研究
····· 施建壮, 张国伟, 卢秋红, 黄 威, 吴松林 (91)
- 基于 Java Swing 的找点软件设计与实现····· 王 勇, 洪 进, 杜兰兰, 王圣波 (96)
- 基于改进协同过滤算法的校园二手交易系统研究····· 许 诺, 高 耀, 周文柱 (101)
- 公共资源交易主体与成交业绩查询系统的设计····· 付宏燕 (106)
- 基于图数据库的阅读行为知识图谱构建研究····· 陈光仪, 陈义明, 吴小慧 (111)
- 一款在线藏文化博物馆平台的设计与实现····· 王 超, 倪胜巧, 普 尺 (114)
- 一个智慧门禁系统的软件设计与实现····· 韩培珊, 郑晓君, 谢 松, 黄翔琪 (118)

Modern Computer

(Vol.28, No.16; Aug.25, 2022)

CONTENTS

Research and Development

A Small-Corpus Chinese Text Classification Method for Optimizing Pre-trained Models	(1)
A Review of Real-Time Detection Methods for Abnormal Data	(9)
Information Hiding of Textures Based on Tibetan Buddhist Decorative Symbols	(16)
Adverse Drug Reaction Detecting Method Based on Attention-BiLSTM-CharCNN	(22)
Racing Bib Number Recognition of Cloud Marathon Images	(28)
Design and Application of Integrated Modeling Tool for Distribution Network Diagram Number Based on Electrical Standard Model	(32)
Research on Automatic Identification Method of Hazardous Waste Based on Ensemble Learning	(38)
The Prediction of Myocardial Infarction Complications with Machine Learning	(43)
3D Object Pose Estimation Method Based on Edge Extraction and Polynomial Fitting	(48)
Algorithm Research Based on Improved YOLOv3	(53)
Research on Common Mode Anti-jamming of NSA in 5G Network	(57)

Graphic and Image

MRI Image Segmentation of Brain Tumor Based on Improved U-Net Model and CBAM	(60)
A Spatial Prediction Based Motion-Compensated Frame Rate Up-Conversion	(67)

Practice and Experience

Application of C4.5 Decision Tree Method in Course Teaching Score Analysis	(73)
--	------

Development Solution

Application of Big Data Visualization in Urban Planning	(77)
Item Storage Management System Based on RFID and Raspberry Pi	(84)
Research on Autonomous Navigation of Mobile Delivery Robot Based on Lidar	(91)
Design and Implement of Marking Software Based on Java Swing	(96)
Research on Campus Second-Hand Trading System Based on Improved Collaborative Filtering Algorithm	(101)
Design of Market Entities and Trading Results Query System in Public Resource Trading	(106)
Research on the Construction of Reading Behavior Knowledge Graph Based on Graph Database	(111)
Design and Implementation of a Platform in Tibetan Culture Museum	(114)
Design and Implementation of Smart Door System for Face Recognition	(118)

研究与开发

文章编号: 1007-1423(2022)16-0001-09

DOI: 10.3969/j.issn.1007-1423.2022.16.001

优化预训练模型的小语料中文文本分类方法

陈 蓝, 杨 帆, 曾 桢

(贵州财经大学信息学院, 贵阳 550000)

摘要: 针对 GloVe、BERT 模型生成的字向量在小语料库中表义不足的问题, 提出融合向量预训练模型, 对小语料中文短文本分类的精确度进行提升。本文以今日头条新闻公开数据集为实验对象, 使用 GloVe、BERT 模型通过领域预训练, 对 GloVe 与 BERT 生成的预训练字向量进行向量融合, 实现语义增强, 从而提升短文本分类效果。结果表明, 当语料库中的数据量为 500 时, 融合字向量的准确度相较于 BERT 字向量的准确度提升了 5 个百分点, 相较于 GloVe 字向量的准确度提升了 3 个百分点。词义选取的维度待进一步加强。本文所提方法能够对小语料库的短文本数据实现精准分类, 对后续文本挖掘工作具有重要意义。

关键词: BERT; GloVe; 向量融合; 小语料; 短文本

基金项目: 教育部产学合作协同育人项目(BZX1902-20): 基于 Jupyter Notebook 的用户信息行为分析整合实验教学设计

0 引言

数字信息资源是指所有以数字形式将文字、数值等多种信息存储在计算机中, 通过网络通信、计算机或终端再现出来的资源。近年来, 数字信息资源的快速增长, 为用户带来便利的同时也导致了“信息爆炸”。数字信息资源的重要组成部分之一就是文本, 针对种类繁多的文本信息资源, 运用现代化的管理手段和管理方法, 将资源按照一定的方式组织和存储起来, 能够使用户在查找海量信息时实现高效检索。

目前, 将文本信息转换为计算机能够识别的数据是自然语言处理的一个重要问题。其中最普及的解决方法是将文本转换为向量的形式, 将一句文本语言转化为一个向量矩阵, 通过相似词具有相近的向量, 对词义进行表示。目前, 由于深度学习的发展及应用, 学者们通过各种神经网络对生成的多个维度的词向量进行特征提取, 降低损失函数值从而对词向量进行优化, 增强向量对词义的表达力。

在现实需求以及自然语言处理技术的基础上, 本文提出了一种中文文本字向量的表示模

型, 使用 GloVe 模型和 BERT 模型生成的字向量进行融合后, 通过文本特征提取得到对应的字粒度向量。

1 研究现状

在中文自然语言处理领域, 计算机无法对非结构化的文本数据进行处理, 因此在对中文文本信息进行处理时, 需要经过分词以及向量化的过程, 也就是将文本信息转化为计算机能够识别的数值数据。其中在文本向量化方面, 最早的文本转换方式为 one-hot(独热)编码形式, one-hot 编码虽然解决了分类器处理离散数据困难的问题, 但是没有考虑词与词之间的相互关系, 并且由 one-hot 生成的特征矩阵较为稀疏, 增加了机器运算的负担。在 2014 年前后, 主要有两种文本向量化方法, 一种是矩阵分类算法, 另一种是基于浅窗口的方法。基于浅窗口方法的代表模型就是 Word2Vec, 为了获得更多的语义信息, Mikolov 等^[1]提出了基于深度表示的模型—Word2Vec, 该模型为输入文本搭建一个具备上下文信息的神经网络, 从而计算得到含有上下文信息的词向量, 该向量也在一定

程度上反映了词与词之间的相关性。虽然 Word2Vec 可以利用上下文信息预测词向量使得生成的词向量包含了语义信息,但由于其构建过程是单向学习,没有充分利用所有语料。而基于矩阵分解算法通过文本共现矩阵表达文本词向量,通过奇异值分解(singular value decomposition, SVD)对共现矩阵进行降维,章秀华等^[2]提出一种奇异值分解域差异性度量的低景深图像目标提取方法,其能够完整提取目标,但 SVD 的计算代价过大,并且难以将新的词汇或者文本合并进去。2014 年 Stanford NLP Group 结合 Word2Vec 以及 SVD 的优点提出了 GloVe (global vectors for word representation)模型^[3],该模型基于全局词频统计将一个词语表达为一个向量,通过单词之间的相似性、类比性等,计算出两个词语之间的语义相似性。方炯焜等^[4]结合 GloVe 词向量与 GRU 模型提高了文本分类性能。石隽锋等^[5]通过并行实现统计共现矩阵和训练学习,从而在中文和英文的词语推断任务上,显著地提高了预测的准确率。FANG 等^[6]以 GloVe 为基础建立情感分析系统,虽然 GloVe 能够最大限度地利用全局和局部信息进行语料库训练,但无法应对一词多义或者新词组合的情况。针对该问题,2019 年 Devlin 等^[7]引入动态词向量 BERT (bidirectional encoder representations from transformers)模型, BERT 模型利用 Transformer 结构的 encoder 部分对文本进行双向学习和处理,主要包含 MLM (masked language model)任务和 NSP (next sentence prediction)任务。其中,核心任务是 MLM 任务,通过对目标单词进行掩码来预测词语的向量,利用自注意力机制学习词与词间关系,使得词向量的表示能够融入句子级的语义信息。段丹丹等^[8]使用 BERT 预训练语言模型对短文本进行句子层面的特征向量表示,并将获得的特征向量输入 Softmax 回归模型进行训练与分类,实验证明 BERT 有效地表示句子层面的语义信息,具有更好的中文短文本分类效果。Chao 等^[9]结合动态掩码与静态掩码,提出新的 MLM 任务与层间共享注意力机制,有效地提高了 BERT 在实体关系提取上的性能。Danilov 等^[10]提出了一种基于双向

编码转换(BERT)和图卷积网络的门上下文感知文本分类模型(GC-GCN),通过使用带有门控机制的 GCN 将图嵌入和 BERT 嵌入集成在一起,以实现上下文编码的获取。虽然 BERT 解决了词向量无法表示一词多义的问题,但通过 BERT 进行向量化的过程中,缺乏了整体的词和词之间的关系。综上所述,现在国内外学者在文本向量化领域做了一些相关工作,但目前的文本向量化在中文文本语料的处理上仍然存在表义不足,因此,中文文本向量化具有研究潜力与价值。

本文在上述研究的基础上提出了基于 GloVe 与 BERT 字向量模型的融合字向量模型。通过 GloVe 领域预训练产生的文本向量无法解决一字多义的情况,但是能够最大限度地利用全局和局部信息进行语料库训练,从而给每一个字都提供一个相对稳定的字向量,在投入的数据量较小的情况下,通过 BERT 领域预训练难以达到训练效果,更多的是依赖初始权重集的选择。因此, GloVe 与 BERT 在训练时各有优劣,本文通过扩充 GloVe 字向量产生的维度,与 BERT 字向量进行向量融合,从而在预训练生成的融合字向量中,既体现了 GloVe 字向量的全局稳定性,也通过 BERT 字向量的展现解决了一字多义的问题。

本文采用今日头条发布的中文新闻数据,该数据属于短文本类型,且包含大量同义或异义字词。新闻文本数据通过 GloVe 及 BERT 模型生成的融合字向量矩阵^[11],该矩阵通过卷积神经网络实现文本特征的提取后进行字向量训练优化^[12],通过全连接层对新闻数据的分类结果的准确率、召回率等一系列指标进行评判,对文本词向量的词义表示能力进行评价。

2 思路与框架

2.1 融合向量模型

融合字向量模型基于 GloVe 字向量模型以及 BERT 字向量模型,主要分为四个部分:输入处理层、融合层、特征提取层以及分类输出层,其具体结构如图 1 所示。

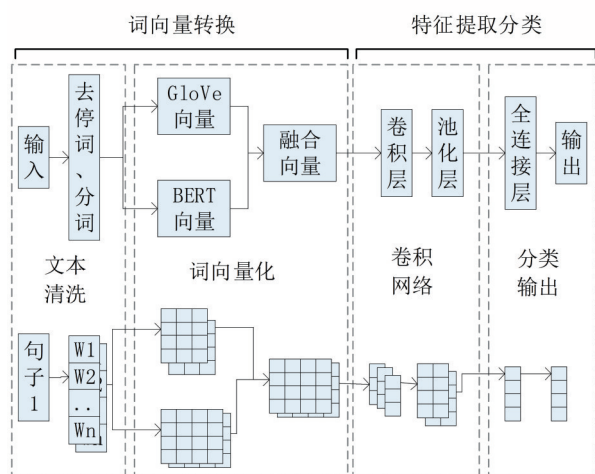


图1 融合字向量模型结构

在该融合字向量模型中，以今日头条发布的 m 条新闻数据作为输入，选取数据中的文本数据以及标签数据，提取文本数据中的中文字词后，为了使 GloVe 与 BERT 的分词结果相同以达到词向量矩阵的数据量相同，从而进行 BERT_tokenize 单字分字处理，得到字粒度的中文文本语料库。

将该语料库输入 GloVe 字向量模型中，本文以 300 维的中文 GloVe 模型作为预训练模型获取字向量，在该字向量的表示中，GloVe 模型通过语料的全局信息进行训练后，相同字有相同的字向量，因此不能表达一词多义。同时，将处理好的语料库输入 BERT 模型，生成 768 维的文本字向量，将 GloVe 向量与 BERT 向量通过点加的方式获得融合字向量。

融合字向量通过文本卷积神经网络^[13-14]获取多层次的语义特征信息，通过训练发现该字向量的关键信息，从而对 768 维的向量进行特征抽取，实现词向量降维。

在进行特征提取后，经过全连接层对文本数据进行分类处理。

2.2 输入处理层

文本卷积神经网络接受的输入数据形式为矩阵形式，因此，需要通过向量化任务将中文文本数据转化为矩阵形式。与英文文本以词组成句不同，中文文本是以字组成句，因此，在中文自然语言处理中，还需要在向量化之前增加分词任务，目前有基于字符的分词方法^[15]、

基于词的分词方法^[16]等。为了使 GloVe 与 BERT 生成的向量矩阵维度相同，从而为后续向量融合奠定基础，本文采用符合 BERT 分词习惯的基于字符的分词方法： $s_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$ ，其中 s_i 为第 i 个句子， w_{i1} 代表第 i 句中的第 1 个字， n 表示文本数据所有句子中所含的最大字数。将 s_i 进行拼接处理后，得到输入文本输入矩阵： $S = [s_1, s_2, \dots, s_m]$ ，该矩阵中共有 $m \times n$ 个字，不考虑字与字之间的重复性，将其进行扁平化处理，作为之后的文本输入。

在 GloVe 词向量模型中，将 $m \times n$ 个字中第一次出现的字挑选出来，若共有 o 个不重复字，这 o 个字组成共现矩阵的坐标标签，那么共现词频矩阵 X 可表示为式(1)：

$$X = \begin{bmatrix} x_{11} & \dots & x_{1o} \\ \vdots & x_{ij} & \vdots \\ x_{o1} & \dots & x_{oo} \end{bmatrix} \quad (1)$$

其中， x_{ij} 代表词 j 在词 i 上下文窗口中出现的次数，由于词 j 在词 i 上下文窗口中出现的次数与词 i 在词 j 上下文窗口中出现的次数相等，所以有 $x_{ij} = x_{ji}$ ，那么 X 就为一个对称矩阵。 x_i 表示为任意词出现在词 i 上下文的次数，词 j 出现在词 i 上下文的概率就为： $P_{ij} = P(j|i) = \frac{x_{ij}}{x_i}$ 。

若存在三个词向量 v_i 、 v_j 和 v_k ，通过构造一个函数来表达 $\frac{P_{ik}}{P_{jk}}$ 是 GloVe 的关键问题：

$$f(v_i, v_j, v_k) = \frac{P_{ik}}{P_{jk}} = \frac{P(k|i)}{P(k|j)} \quad (2)$$

Pennington 等^[3]将该关系函数构造为 $v_i^T \tilde{v}_j + b_i + b_j = \log x_{ij}$ ，其中 b_i 和 b_j 为关于词 i 和词 j 的偏置项，那么该文本共现矩阵的损失函数就为：

$$L = \sum_{ij=1}^v f(x_{ij}) (v_i^T \tilde{v}_j + b_i + b_j - \log x_{ij})^2 \quad (3)$$

其中 $f(x_{ij})$ 为权重函数，且有：

$$\begin{cases} f(x_{ij}) = \left(\frac{x}{c}\right)^\alpha, & x < c \\ f(x_{ij}) = 1, & x \geq c \end{cases} \quad (4)$$

最后，通过 AdaGrad 的梯度下降算法对该函数进行训练，从而获取较优的词向量。GloVe 词向量模型训练过程如图 2 所示。

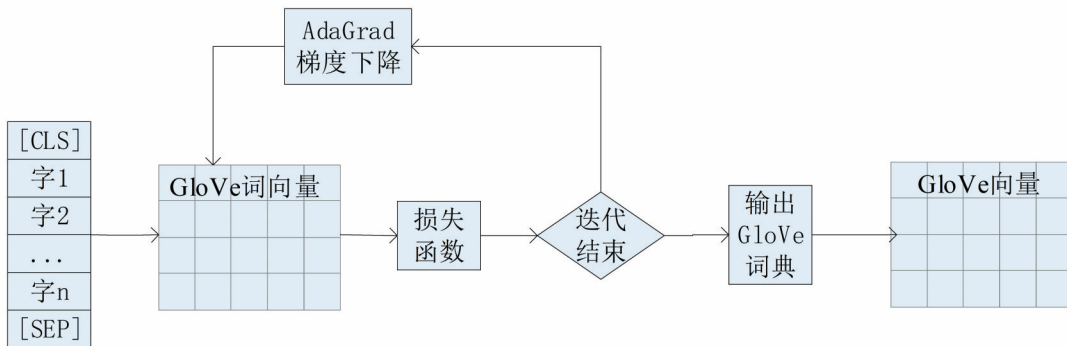


图2 GloVe词向量模型结构

BERT以Transformer的encoder结构为基础，主要包含MLM掩码任务和NSP语句预测任务，BERT在输入嵌入层(input embedding)通过查询词典中每个词语对应的向量表得到句子的向量矩阵，与GloVe不同的是，BERT在输入嵌入层的基础上增加了体现词语在句子中所在位置的位置嵌入层，具体计算方式见式(5)：

$$\begin{cases} PE(pos, i) = \sin\left(\frac{pos}{10000^{2i/d_{model}}}\right), i \text{ 为双数} \\ PE(pos, i) = \cos\left(\frac{pos}{10000^{2i/d_{model}}}\right), i \text{ 为单数} \end{cases} \quad (5)$$

其中， pos 为该词在句子中的位置，根据出现位置的单数或双数，以 \sin 或 \cos 方式生成的位置值交替出现， d_{model} 为模型需要训练的参数。BERT在位置嵌入层的基础上增添了体现句子在文本语料的位置关系的句子嵌入层，根据句子出现的位置，表示为 $[E_{11}, E_{12}, \dots, E_{1o}, E_{21}, E_{22}, \dots, E_{2o}, \dots]$ 的形式， E_{ij} 表示第*i*个句子的第*j*个词，且 $E_{i1} = E_{i2} = \dots = E_{io}$ ，用以区分该句中的词语与其它句子中的词语。这三个嵌入层共同组成encoder结构的输入层 $X_{embedding}$ 。

三个矩阵 Q 、 K 、 W 与 $X_{embedding}$ 相乘得到 $q_i, k_i, v_i, i \in (1, 2, \dots, o)$ ，将 q_i 与 k_j 做点积得到 $\alpha_{i,j}$ ， $\alpha_{i,j}$ 通过全连接层后得到 $0 \sim 1$ 之间的 $\alpha'_{i,j}$ ， $\alpha'_{i,j}$ 与对应位置的 v_i 相乘且求和得到输出 b_i ，这就是多头注意力机制，具体过程如图3所示。

多头注意力机制生成的 b_i 通过前馈神经网络的训练生成BERT词向量，BERT词向量模型结构如图4所示。

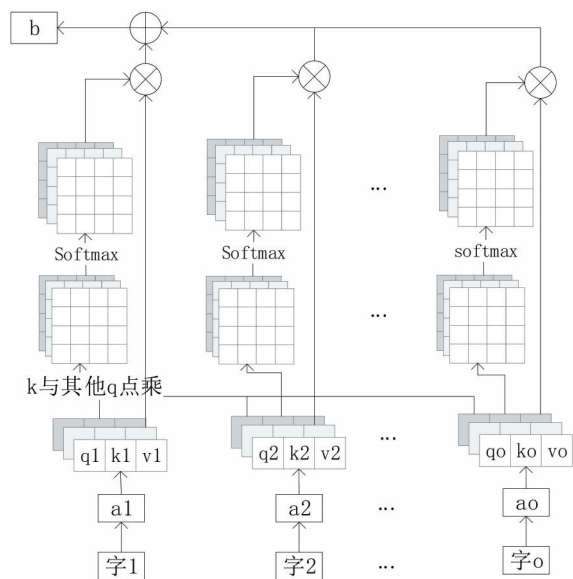


图3 多头注意力机制

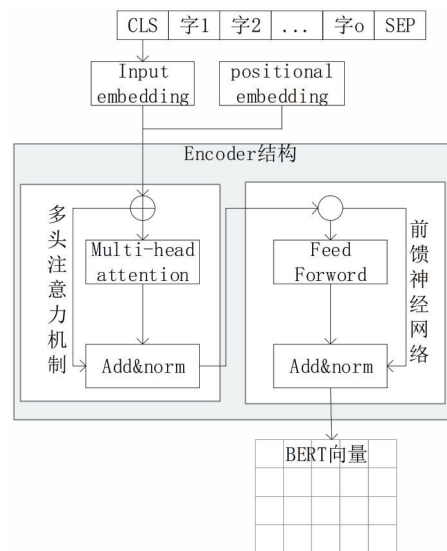


图4 BERT词向量模型结构

2.3 融合层

通过 GloVe 和 BERT 模型生成的向量矩阵 M_{GloVe} 和 M_{BERT} 经过向量融合^[17-18]后生成最终的文本词向量矩阵 M_{GB} 。为了在 M_{GB} 中充分体现出 GloVe 以及 BERT 生成的词义信息, 本文通过 M_{GloVe} 和 M_{BERT} 相加的融合方式生成维度为 $(m \times (n + 2)) * k_{embedding}$ 的 M_{GB} 矩阵, 该维度表示有 m 句话, 每句话有 n 个字, 每个字的字向量维度为 $k_{embedding}$, 融合模型如图 5 所示。

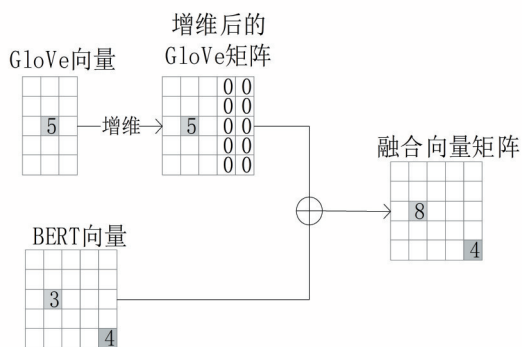


图 5 融合字向量

2.4 特征提取层

在文本向量化表示中, 向量维度过大会造成计算机的运行负担较大, 部分维度的向量可能在最终的分类判别中只占极小的权重, 因此采用特征提取的方式对数据进行降维处理, 目前常见的特征提取方式有主成分分析法^[19]、线性判别分析法^[20]、独立成分分析法^[21]等, 本文采用文本卷积神经网络^[22-24]完成对融合向量矩阵 M_{GB} 的特征提取任务。语料库中的文本数据通过数据清洗以及分词处理后, 输入 GloVe 模型及 BERT 模型, 生成字向量矩阵 M_{GloVe} 和 M_{BERT} , 两个矩阵经过融合层生成字向量矩阵 $M_{GB} = [M_{GB,1}, M_{GB,2}, \dots, M_{GB,i}, \dots]$, 其中 $M_{GB,i}$ 表示第 i 句话的向量矩阵, 维度为 $n * k_{embedding}$ 。将 $M_{GB,i}$ 输入到文本卷积神经网络中, 文本卷积神经网络对字组成的滑动窗口的特征进行组合和筛选, 从而获得不同抽象层次的语义信息。在自然语言处理中, 卷积核滑动过的位置都是完整的一个字, 其宽度是每个字的维度, 使用大小为 d 的卷积核, 从而获取 d 个字之间的关系。输入矩阵 $M_{GB,i}$, 使用大小为 d 的卷积核进行卷积操作后,

得到一个维度为 $(n - d + 1) * 1$ 的特征向量 c_i 。 $c_i = \text{cnn}(M_{GB,i}) + b$, $\text{cnn}(M_{GB,i})$ 代表对 $M_{GB,i}$ 做卷积操作, b 为一个偏置项。此处是判断该句也就是 $M_{GB,i}$ 的标签, 因此采用最大池化的池化方法, 从特征向量 c_i 筛选出一个最大的特征做为这个句子的特征值 a_i , 从而对后续的分类奠定基础。文本卷积神经网络结构如图 6 所示。

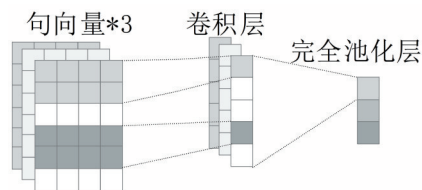


图 6 文本卷积神经网络结构

2.5 分类输出层

每个句子的特征值经过全连接层, 以 sigmoid 函数作为激活函数, 从而输出预测的分类结果。计算公式为: $per_i = 1 / (1 + e^{-a_i})$, 若 $per_i > 0.5$, 则类别为正类; 若 $per_i \leq 0.5$, 那么类别为负类。

3 实证研究

3.1 实验数据

本实验以今日头条发布的开源新闻数据集作为输入数据, 原数据集包含 15 类共 382669 条数据, 每条数据包含文本 id 名、文本类别数值代号、文本类别名、文本内容、关键词五列。从 15 类新闻中抽取“文化”和“娱乐”两个板块的新闻, 其中文化新闻共 28030 条, 娱乐新闻共 39396 条, 满足实验所需。从这两个类别的数据中抽取文本类别数值代号和文本内容两列内容后, 分别抽取 250 条、500 条、2500 条、5000 条, 组成共有 500 条数据、1000 条数据、5000 条数据、10000 条数据的 4 个不同大小的文本语料库。

将文本内容进行文本正则化处理, 仅保存文本中的中文字, 并且按照字粒度对文本进行分词后, 将其按照类别标签存储在各自类别的文件夹里, 每一条文本存储在以索引编号命名的文件里。将文本数据进行打乱顺序处理, 选取其中 20% 的数据作为评估集, 80% 的数据作

为训练集，并且评估集与训练集相互独立。由于单次划分得出的结果并不稳定，因此每个输入数据集进行20次实验，选取20次实验中效果最好的5次实验结果的平均值做为最终的实验结果。

3.2 评价指标

本文实验的评估指标有：评估集的准确率 *accuracy*、评估集的查全率 *recall*、评估集的查准率 *precision*，评估集的 F_1 值四个。

将“文化”类的数据设为负类，“娱乐”类的文本数据设为正类，得到的预测类别与实际类别的情况见表1。

表1 二分类混淆矩阵

		预测类别	
		正类	负类
实际类别	正类	正确肯定 <i>TP</i>	错误否定 <i>FN</i>
	负类	错误肯定 <i>FP</i>	正确否定 <i>TN</i>

将预测为正类且实际类别也为正类的结果记为 *TP*，预测为负类但实际类别为正类的结果记为 *FN*，将预测为正类但实际类别为负类的结果记为 *FP*，预测为负类且实际类别也为负类的结果记为 *TN*，那么有：

$$accuracy = \frac{TP + TN}{all} \quad (6)$$

$$recall = \frac{TP}{TP + FN} \quad (7)$$

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

$$F_1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} \quad (9)$$

其中 *accuracy* 展示了分类正确的数据条数与数据总数的比值，可以在总体上衡量一个预测的性能；*recall* 表示被正确分类的正类数据条数与所有应该被判断为正类的数据条数之间的比值，展示正类样本被误判的程度；*precision* 展示了被正确分类的正类数据条数与所有被判断为正类的数据条数之间的比值，找到所有的正类防止漏掉正类样本。 F_1 将 *recall* 和 *precision* 视为相同重要的两个评估标准，对模型的性能进行了一个综合评价。

3.3 参数设置

GloVe 词向量基于全局词频统计把单词表现为一个300维的向量形式，词典大小为352221；BERT 预训练模型采用 bert-base-chinese 模型，包含12个 encoder 单元，768个隐藏单元数，12个注意力机制的头数，110M 参数，词典大小为21128，生成768维的词向量。在文本卷积神经网络中，选取两层卷积神经，为了查看两个字组成的词语之间的关系，将第一层卷积核的大小设为2；为了查看主谓宾之间的关系，将第二个卷积核的大小设为三；由于是二分类，在全连接层采用 sigmoid 激活函数，将数据特征分为正负两类。具体参数见表2。

表2 卷积层参数设置

参数名称	参数值
卷积核高度	(2, 3)
卷积核数目	128
批尺寸	128
学习率	0.01
优化器	RMSProp

4 结果与分析

为了验证该融合向量算法的有效性，将今日头条的4种不同大小的数据集在 GloVe 预训练算法、BERT 预训练模型以及融合词向量预训练模型中进行训练，得到的准确率见表3。

表3 准确率

准确率	GloVe /%	BERT /%	融合 /%
500条数据	77.0	75.80	80.80
1000条数据	78.60	78.90	82.00
5000条数据	86.82	86.64	87.70
10000条数据	86.82	87.46	87.82

各个模型在不同数据集上的召回率见表4。

表4 召回率

召回率	GloVe /%	BERT /%	融合 /%
500条数据	92.31	85.39	90.00
1000条数据	93.16	92.42	91.22
5000条数据	93.60	92.68	91.85
10000条数据	93.60	92.38	93.79

模型在不同数据集上的查全率如图7所示。

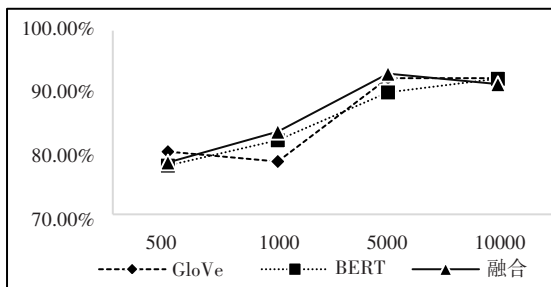


图7 查全率

模型在不同数据集上的 F_1 值如图8所示。

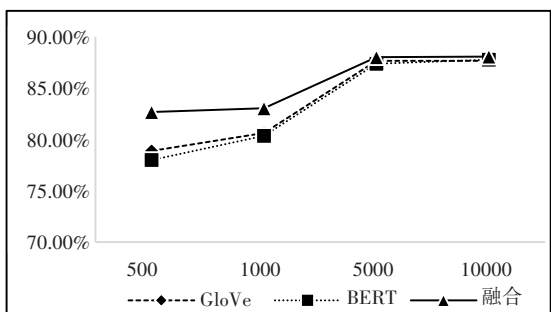


图8 F_1 值

由实验得出如下结论:

(1) 通过表3可发现:模型的训练效果在数据量较小时比数据量较大时要差,当数据量从1000增加到5000时,各个模型训练的准确度都提升了5到8个百分点,其中GloVe模型的准确率从78.60%提升到86.82%,BERT的准确率从78.90%提升到86.64%,融合词向量模型的准确率从82.00%提升到87.70%。

(2) 通过表4中召回率的展示可以发现,GloVe模型的表现不比BERT模型表现差。在表2中,输入数据为500条时,GloVe模型产生的词向量最终训练后的准确率为77.00%,而BERT预训练模型产生的词向量训练后的准确率为75.80%,这是由于小数据量时的同词异义的情况出现的概率较小,采用词频统计的方法构建字典从而生成词向量,相比使用数据进行迭代优化出来的词向量,在数据量较小时能够更加准确地表达词义。

(3) 在图8中,融合向量模型的 F_1 值总是略高于单一模型的 F_1 值,并且数据量越小,融合向量模型的优势越明显。图7显示,相较于单一

模型产生的词向量,融合向量模型的查全率在1000条以及5000条数据组成的语料库中都高于单一模型的训练效果。在表3中,融合向量准确率在数据量较小时有小幅度的提升,当数据量为500时,融合词向量相较于BERT词向量提升了5个百分点,相较于GloVe词向量提升了3个百分点;当数据量为1000时,融合词向量相较于BERT词向量的精确率提升了3.1个百分点,相较于GloVe词向量提升了3.4个百分点。

(4) 上述图表显示,当输入的数据量较大时,各个模型训练出来的词向量的最终准确率都在86%~88%之间,准确率相差一个百分点以内,并没有显著的区别。这是由于输入模型的数据量足够让BERT训练出一个较好的结果,同时也反映了融合词向量模型在大数据量时的训练效果,虽然没有明显优于BERT词向量模型,但也并不逊色于BERT词向量模型。

从5000条输入数据增加至10000条输入数据时,GloVe模型与融合词向量模型的准确率没有明显的变化,而BERT词向量模型的准确率提升了0.82个百分点,但BERT模型依靠增加数据量来提升准确度的收效减小,代价过大。

5 结语

本文的研究价值主要体现在:通过对BERT与GloVe模型生成的词向量进行融合,获取两种词向量的语义信息,从而在小数据量的语料库中,融合词向量的语义信息表示能力有一定的增强,采用浅层文本卷积神经网络对融合词向量进行特征提取,使得该融合词向量能够准确表达词义的同时,也降低了实验中对硬件的要求,降低了模型的部署难度。

本文详细阐述了融合词向量的整体框架、模型结构,以及相关的计算公式,并用不同大小的数据集来验证模型的有效性。实验表明,相较于传统的GloVe模型和单一的BERT预训练模型,融合词向量模型的准确率在较小数据量的情况下都能得到提升。这说明融合不同词向量方法的文本表示能够获取字词的先验知识,同时根据数据特征进行优化,从而获得词向量更好的表示方法。

在未来的研究中,可以尝试采用不同维度的文本信息,比如词粒度文本信息,在词向量融合时引入注意力机制,为BERT词向量以及

GloVe词向量加一个权重之后进行词向量融合,从而进一步提升文本的表达能力。

参考文献:

- [1] MIKOLOV T, CHEN K, CORRADO G, et al. Efficient estimation of word representations in vector space[EB/OL]. arXiv:1301.3781, 2013.
- [2] 章秀华, 程鉴, 洪汉玉, 等. 奇异值分解域差异性度量的低景深图像显著性目标提取方法[J/OL]. 电子与信息学报, 2022: 1-11. <http://kns.cnki.net/kcms/detail/11.4494.TN.20211227.1303.014.html>.
- [3] PENNINGTON J, SOCHER R, MANNING C. GloVe: global vectors for word representation[C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, 2014: 1532-1543.
- [4] 方炯焜, 陈平华, 廖文雄. 结合 GloVe 和 GRU 的文本分类模型[J]. 计算机工程与应用, 2020, 56(20): 98-103.
- [5] 石隽锋, 李济洪, 王瑞波. 一种改进的 GloVe 词向量表示学习方法[J]. 中文信息学报, 2021, 35(4): 16-22.
- [6] FANG B, SUN F C, LIU H P, et al. A glove-based system for object recognition via visual-tactile fusion[J]. Science China (Information Sciences), 2019, 62(5): 11-13.
- [7] DEVLIN J, CHANG M W, LEE K, et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding [C]//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2019(1): 4171-4186.
- [8] 段丹丹, 唐加山, 温勇, 等. 基于 BERT 模型的中文短文本分类算法[J]. 计算机工程, 2021, 47(1): 79-86.
- [9] CHAO L, ZHAO Q. Targeted BERT Pre-training and Fine-Tuning Approach for Entity Relation Extraction[C]//Abstracts of the 7th International Conference of Pioneering Computer Scientists, Engineers and Educators (ICPCSEE 2021) Part II, 2021: 116-125.
- [10] DANILOV G, ISHANKULOV T, KOTIK K, et al. The classification of short scientific texts using pre-trained BERT model[J]. Studies in health technology and informatics, 2021, 281: 83.
- [11] NARESHKUMAR S K, DICKEETA D S, MAMATA D H, et al. A novel approach for dimension reduction using word embedding: an enhanced text classification approach[J]. International Journal of Information Management Data Insights, 2022, 2(1): 100061.
- [12] MA Z Z. Improved convolutional neural network for biomedical text classification[J]. Journal of Physics: Conference Series, 2021, 1883(1).
- [13] 李文亮, 杨秋翔, 秦权. 多特征混合模型文本情感分析方法[J]. 计算机工程与应用, 2021, 57(19): 205-213.
- [14] 邵云飞, 刘东芬. 基于类别特征扩展的短文本分类方法研究[J]. 数据分析与知识发现, 2019, 3(9): 60-67.
- [15] CHEN X, SHI Z, QIU X, et al. Adversarial multi-criteria learning for Chinese word segmentation[EB/OL]. arXiv Preprint arXiv:1704.07556, 2017.
- [16] CAI D, ZHAO H, ZHANG Z, et al. Fast and accurate neural word segmentation for Chinese[EB/OL]. arXiv Preprint arXiv:1704.07047, 2017.
- [17] 苏依拉, 高芬, 仁庆道尔吉. 中文字粒度切分在蒙汉机器翻译的应用[J]. 中文信息学报, 2019, 33(12): 54-60.
- [18] 张丽, 马静. 融合词语统计特征和语义信息的文本分类方法研究[J]. 计算机工程与科学, 2021, 43(7): 1308-1315.
- [19] 李东博, 黄铝文. 重加权稀疏主成分分析算法及其在人脸识别中的应用[J]. 计算机应用, 2020, 40(3): 717-722.
- [20] 刘忠宝. 一种改进的线性判别分析算法在人脸识别中的应用[J]. 计算机工程与科学, 2011, 33(7): 89-93.
- [21] 田甜. 独立成分分析及其在特征提取中的应用研究[D]. 西安: 西安电子科技大学, 2017.
- [22] GAO W Q, HUANG H. A gating context-aware text classification model with BERT and graph convolutional networks[J]. Journal of Intelligent & Fuzzy Systems, 2021, 40(3): 4331-4343.
- [23] 李文亮, 杨秋翔, 秦权. 多特征混合模型文本情感分析方法[J]. 计算机工程与应用, 2021, 57(19): 205-213.
- [24] 陈杰, 马静, 李晓峰. 融合预训练模型文本特征的短文本分类方法[J]. 数据分析与知识发现, 2021, 5(9): 21-30.

作者简介:

通信作者: 陈蓝(1997—), 女, 重庆人, 硕士研究生, 研究方向为自然语言处理, E-mail: chenlan202009@163.com

杨帆(1968—), 男, 贵州贵阳人, 博士, 教授, 研究方向为自然语言处理

曾楨(1982—), 男, 贵州贵阳人, 博士, 教授, 研究方向为信息系统、数据挖掘

收稿日期: 2022-04-05 修稿日期: 2022-04-27

(下转第15页)

文章编号: 1007-1423(2022)16-0009-07

DOI: 10.3969/j.issn.1007-1423.2022.16.002

异常数据实时检测方法研究综述

吴 琼, 李永飞, 李铭洋

(华北科技学院计算机学院, 三河 065200)

摘要: 随着数据价值的不断凸显, 数据可信性日益重要, 异常数据检测逐渐成为数据挖掘领域的研究热点之一。物联网监测所产生的海量数据, 对传统异常检测方法带来了新的挑战。本文分别介绍了传统异常检测方法和针对海量动态数据的实时异常检测方法, 指出结合大数据技术的流处理框架可以加快数据处理速度, 并实时反馈检测结果; 同时一些基于机器学习和神经网络新进展的新型异常数据实时检测算法, 在实现实时异常检测的同时, 还可以提高异常检测准确率。通过结合新的数据处理技术可以更好地实现异常数据的实时检测。

关键词: 异常检测; 分类方法; 实时性; 检测算法

基金项目: 河北省重点研发计划项目(19270318D); 中央高校基本科研业务费资助项目(3142017067); 河北省教育厅科学技术研究项目(Z2019044)

0 引言

异常数据通常也称为离群值、噪声等。Edgeworth把异常值定义为“显然严重偏离样本集合中其他观测值的观测值”^[1]。异常数据检测是指从数据中找出明显与其他数据不同的数据^[2], 最早采用基于统计学的方法, 现在已成为数据挖掘的四大任务之一。异常数据所占比例较小但却可能蕴含丰富的内容, 因此异常数据检测具有重要的研究意义和应用价值^[3], 对保障数据可信性也有重要作用。

从Knox等^[4]提出“基于距离的异常数据”的概念以来, 异常数据检测已经成为数据挖掘中的一个研究热点。异常值可能来源于机械故障、仪器错误、人为错误等, 其产生往往不可避免, 而且会影响数据分析的结果, 甚至可能造成严重后果。目前, 异常数据检测已经广泛应用于诸多领域。在医疗卫生行业, 异常数据检测可以及时发现病人的身体指标异常, 提早治疗, 避免病情恶化。在生态环境监测领域, 可以对各项监测指标的异常情况尽早采取应对措施。在应急管理的事前数据采集和处理阶段,

异常数据实时检测可以对突发事件及时响应, 进而及时调整应对措施。

近年来, 随着大数据时代的到来, 海量数据的产生对数据处理的方法提出了更多的挑战。异常数据检测的实时性和准确性方面有了更高的要求。异常数据检测的核心问题是异常检测算法模型的构建, 它将直接影响异常检测的检测率^[5]。

本文主要介绍传统异常数据检测方法和异常数据实时检测方法。首先介绍常用的异常检测方法以及它们常用的检测算法, 主要分为基于统计学和基于机器学习两大类; 接着介绍异常数据实时检测方法以及采用的算法; 最后对异常数据实时检测的发展趋势进行总结和展望。

1 异常检测方法

1.1 基于统计学的方法

基于统计学的异常检测方法, 一般建立一个数据分布模型, 然后计算对象符合该模型的概率, 将低概率的对象视为异常。最简单的方法有箱线图、 3σ 准则、Grubbs 检验等, 这些方法都需要假设数据服从某种分布, 然后利用数

据去进行参数估计。复杂一些的还有混合高斯建模、基于马尔科夫模型和时间序列建模,等等。基于统计学的方法优点是鲁棒性较好,适合低维度数据。但是处理高维数据受限制,而且受数据分布和模型参数的影响,因此限制了它的应用范围。

1.2 基于机器学习的方法

机器学习算法按照是否需要人工标记可以分为无监督、有监督和半监督模式。无监督模式不需要任何标签,也不依赖完善的先验知识,因此在异常检测领域应用更加广泛。

1.2.1 基于聚类的方法

基于聚类的异常检测方法将数据分为不同的簇,而异常数据是不属于任何一个簇的^[6]。

聚类分析属于无监督模式,不依赖预先对数据的标记和训练,可以根据数据的相似度把数据划分为多个类或类簇^[7]。K-means算法逻辑简单、计算复杂度低、聚类效果也不错,是应用最广泛的聚类算法之一。由于异常数据检测的结果依赖于聚类算法的分析结果^[8],因此要求在使用K-means算法实现异常检测时能够解决聚类结果不稳定的问题。文献[9]提出基于近邻传播算法和最大最小距离算法结合使用的APMMD算法,利用近邻传播算法和最大最小距离思想计算初始聚类中心,并将获得的初始聚类中心应用于K-means聚类算法中,使迭代次数降低,其聚类结果保持稳定且具有较高的异常检测的准确率。

1.2.2 基于近邻度的方法

基于近邻度的方法包括基于距离和基于密度的方法。这种方法不需要假设数据的分布。

(1) 基于距离的方法。基于距离的方法通过计算数据与近邻数据之间的距离来判断出异常数据^[6],一般采用欧式距离和曼哈顿距离,主要应用于全局近邻^[10]。常用的算法是K-最近邻(K-nearest neighbor)算法,它是一种简单易用、有监督模式的机器学习算法。KNN算法首先找到k个最近的邻居,然后根据k个最近的邻居计算异常分数。

(2) 基于密度的方法。基于密度的方法与基于近邻度的异常点检测密切相关^[11],是在基于

距离的基础上,把数据之间的距离和周围邻近的数据个数相结合^[12]。它通过比较每个点和其邻域点的密度,当点与包围其的邻居的密度不同时,被认定为异常点。经典的方法是Breunig等^[7]提出的局部离群点方法LOF(the local outlier factor)。LOF中计算距离采用欧式距离,LOF为每个点分配一个局部离群因子(LOF),LOF的大小表示该点的局部密度与其最近邻居的局部密度之比,LOF越高,则表明是异常点的可能性越大。基于连通性的异常检测算法COF(connectivity-based outlier factor)的局部密度通过链式距离方式计算求出。LOF和COF算法都是借助数据的相对密度计算异常分值,得分越高越可能是异常数据。由于LOF数据量的增多该算法的时间复杂度较高,文献[13]提出改进的LOF算法即K-LOF算法。先利用计算效率高的K-means聚类算法对数据进行检测预处理获得初选异常数据,然后用基于密度的LOF算法再挖掘出最终的异常数据。结果表明,相比LOF算法K-LOF不但提高了检测的精确度,也降低了异常数据检测的计算复杂度。

1.2.3 其他专用的数据异常检测方法

另外,还有一些专用的异常检测方法,如One Class SVM和孤立森林(Isolation Forest)。这类算法的思路是不需要利用统计、距离和密度的量化指标去表达异常数据的疏离程度,而是直接描述正常数据与异常数据的疏离程度。

(1) One Class SVM算法。One Class SVM属于无监督学习算法,不需要标记训练集和输出标签。该算法的思路比较简单,首先找一个超平面将正常数据找出来,再通过正常数据的特征去学习一个决策边界,然后通过这个边界去判断新来的数据是否与训练数据类似,超出边界则为异常。由于该算法只有一个Class,比较适合解决极度不平衡的数据异常检测^[14]。文献[14]得出One Class SVM算法核函数计算时间比较长,因此不适合用于大量数据的处理。

(2) 孤立森林算法。孤立森林算法是2008年由南京大学周志华教授团队首次提出,2012年又进行了改进。孤立森林算法属于无监督学习,它通过样本的疏密程度去描述样本之间的差异。异常点被定义为“容易被孤立的离群

点”。它由多个 iTree(孤立二叉树)构成, 每个 iTree 的构建是从数据特征集合中随机选择一个分割值, 通过这个分割值对数据进行划分然后构造左右子树, 直到所有数据被划分或已经达到树的高度限制, 其中只关心路径较短的点, 它们更可能是异常点。这种划分情况下, 异常数据点在 iTree 中更靠近根节点, 同时低密度的

点远离大多数样本, 将很早被孤立^[14]。孤立森林算法是线性时间复杂度, 与 K-means、LOF 等算法相比较, 它不需要计算有关的距离、密度的指标, 处理异常数据快速且高效。因此适合实时在线异常数据检测。表 1 列出了常用异常检测算法的优缺点比较。

表 1 异常检测算法优缺点比较

方法分类	典型算法	优点	缺点
基于统计的方法	无参数: 直方图	基于数据分布快速且有效、鲁棒性好	需要预先假设数据的分布情况, 而且不适用于高维数据
	有参数: 高斯模型		
基于聚类的方法	K-means	实现简单、聚类效果不错, 不需要标签和先验知识	异常检测结果依赖聚类效果
基于距离的方法	KNN	无数据分布假设、简单方便计算	消耗时间长
基于密度的方法	LOF、COF	可以根据异常程度给一个定量的值, 具有良好的鲁棒性	高维数据计算量很大
其他专用异常检测方法	One-class SVM	无监督不需要标记训练集和输出标签, 适合解决极度不平衡数据	数据需具有高全局相关性, 不适用于大量数据
	Isolation Forest	线性时间复杂度, 处理异常数据快速且高效, 可以满足实时性要求	不适合特别高维的数据

2 基于流处理框架的异常数据实时检测

由于传统方法不能满足海量的动态化数据的实时异常检测的需求。因此出现了基于流处理框架的新方法。动态化的数据, 也称为数据流, 它具有动态产生且事先无法预知的特点。动态数据需要刚到达就及时处理, 同时满足数

据实时性异常检测要求, 数据处理效率要求更高^[15]。目前异常数据实时检测中主要采用基于 Storm、Spark 等实时流处理框架结合机器学习算法和基于滑动窗口的方法等。依靠流数据分析处理方法实现异常检测, 可以用在实时数据分析场景中^[16]。基于流处理框架的异常数据实时检测方法对比如表 2 所示。

表 2 基于流处理框架的异常数据实时检测方法对比

方法	计算方式及效率	特点
MapReduce+LOF	并行化计算, 提高了效率	提高了数据异常检测的速度和准确率。但有计算延迟问题
Spark+孤立森林算法	基于内存的方式, 比 MapReduce 思想计算更快	低延迟, 大量数据下可以加快检测速度和提高检测精度
HDFS+最小闭包球算法	快速准确地存储数据, 进而减少计算时间	降低了时间复杂度, 提高了异常检测率
Spark Streaming+流回归机器学习算法+正态统计技术	实时计算	实时且准确的检测异常数据
Storm+动态 KNN	实时计算	可以节省内存, 动态观察检测结果
Storm+滑动窗口方法	实时计算, 实时分析	延时低, 提高实时处理效率, 对非数值型数据还要进一步研究
Flink+Kafka 预处理+ARIMA 模型	减少了计算时间, 提高了计算效率	先预处理在实时异常检测

2.1 结合机器学习算法的方法

Hadoop 核心是分布式文件系统 HDFS 和并行化计算模型 MapReduce。随着海量数据的产生,为了加快数据的处理速度则采用并行化计算。文献[17]为了解决初始聚类中心敏感问题,利用最大最小距离的思想改进了 K-means 聚类算法,同时采用 MapReduce 并行化实现该算法的分布式聚类。结果表明提高了算法的计算效率,并且降低了算法执行过程的通信开销^[17]。文献[18]采用基于 Hadoop 平台中的 MapReduce 并行计算框架,并使用基于密度的 LOF 算法的分布式财务异常数据分析模型。采用 MapReduce 并行化计算框架,算法在运行时可以在多个计算节点运行。将 MapReduce 框架和加入领域关系的 LOF 算法结合使用,可以并行计算,进而提高了数据处理速度和算法的准确率。

由于孤立森林算法的每棵树随机采样独立生成,具有很好的处理大数据的能力和速度,可以进行并行化处理。同时还存在孤立二叉树间异常能力的差异性,文献[19]采用加权计算测试样本在孤立森林算法中异常值,同时采用基于内存的 Spark 框架不同于 Hadoop 框架中从 HDFS 中读取数据,比较适合迭代次数多的算法。结果表明在大规模数据下可以加快检测速度和提高检测精度。文献[20]提出基于 HDFS 框架的数据异常检测方法,利用分布式 HDFS 框架可以快速准确地存储数据,采用基于支持向量数据描述并结合最小闭包球算法实现实时异常检测。该方法降低了时间复杂度,提高了异常检测率,并减少了运行时间。

由于 K-means 聚类算法在处理大量数据时效率较低,文献[21]提出了基于 Apache Flink 流计算框架结合流处理思想的 SK-means (stream K-means) 方法,提高了算法的执行效率,聚类效果更好并且可以较快地进行异常数据检测。文献[22]提出基于分布式流处理框架 Spark Streaming,采用流回归机器学习算法和正态统计技术相结合的方法进行数据异常检测。该方法可以实时且准确分析瓦斯浓度流数据中的异常数据,解决了流数据中大数据机器学习处理和实时性问题。文献[23]提出基于 Storm 实时处理平台采用动态 KNN 的累积距离的异常检测方

法。该方法适用于实时处理框架,每一组时间序列只用动态地保存 K 个时间点的数值,可以简化操作和节省内存。同时可以动态地观察数据检测结果。

2.2 基于结合滑动窗口的方法

滑动窗口机制可以处理最新到达的数据,文献[24]提出基于 Storm 流数据框架的滑动窗口计算方法。采用 Storm 平台上实现滑动窗口计算方法进行实时分析,并通过增大滑动窗口的吞吐量,提高了数据异常检测的实时处理效率。但是该方法只对数值型数据实现了实时处理,还需要进一步研究。文献[25]提出基于 Storm 流处理的数据实时处理方法,采用基于滑动时间窗口实现异常数据检测。可以实现在 Storm 上实时监测数据预处理、数据异常检测。

为了从海量数据中实时且高效地检测出异常值,文献[26]提出了 Flink 的异常检测方法,针对实时流数据,首先用 Kafka 对数据进行实时预处理,然后在 Flink 平台上利用 ARIMA 模型进行预测。

3 基于新型算法的异常数据实时检测

随着机器学习、神经网络等领域的发展,又出现了一些新颖且有效的异常数据实时检测方法。基于新型算法的异常数据实时检测方法如表3所示。

3.1 基于层级实时记忆的方法

文献[27]提出基于层级实时记忆(hierarchical temporal memory, HTM)的时间序列异常检测算法,HTM 算法是一种仿生物结构的机器学习算法,它不需要采用滑动窗口法批处理数据,就可以实现对数据的实时检测。它是“记忆-预测”的运行模式,将复杂的问题转化为模式识别,可以提前预警数据异常。随着云计算技术的发展,云资源的运行会产生海量的时序数据^[28]。文献[28]将基于分层时间记忆算法用在企业多云时序数据实时监测中,可以实现实时异常检测。由于云资源监测要实现自动化实时异常检测,而 HTM 算法存储大量时序数据符合实时流式分析、无监督以及动态数据在线学习的要求^[27-28]。因为该算法运用到监测系统中可

表3 基于新型算法的异常数据实时检测方法比较

方法分类	算法	优点	缺点
基于层级实时记忆的方法	HTM	不需要采用滑动窗口法批处理数据, 就可以实现对数据的实时检测	计算复杂度较高
	HTM	自适应性强, 该算法应用到企业的云时序数据实时监测中, 实时异常检测并提高了企业的运维效率	计算复杂度较高
	面向多核的并发HTM空间池算法	实现多个核心上并行完成, 提高了执行效率和预测准确率	没有整体提高HTM模型的训练效率
基于长短期记忆神经网络的方法	LSTM	可以捕捉到数据在时间维度上的信息	无法满足实时异常检测的要求
	LSTMs-Autoencoder	可以实时检测并保证检测结果准确, 同时还能应对考虑概念漂移现象	算法的自适应性不高

以高效地检测异常, 并提高企业的运维效率。HTM算法已经应用到许多数据智能处理领域如异常检测、数据预测^[28]。针对数据量的不断增长, 快速处理的需要, 以及无法并行化计算的问题, 文献[29]提出了面向多核的并发HTM空间池算法, 将HTM空间池区域分区, 各区独立完成训练任务并且利用CPU中的计算核心, 实现多个核心并行完成。使用基于多核心和共享内存的大数据平台Phoenix, 避免带来额外的通信开销^[29], 并且提高了算法的执行效率和预测准确率。

3.2 基于长短期记忆神经网络的方法

对于时间序列数据, 直接采用LSTM算法自适应性不高^[27], 而且单一模型检测结果准确率不高^[31]。文献[30]提出通过LSTM网络和自动编码器进行不同组合预测模型, 进而影响检测器的性能。由于流数据数量大、到达快速, 单个平稳模型可能无法满足数据实时异常检测的要求。文献[31]提出了基于LSTMs-Autoencoder的流数据异常检测算法。该算法采用多个LSTM单元, 形成了一个深层递归的神经网络(LSTMs), 然后将递归神经网络与自动编码器相结合, 实现了对流数据的实时检测并保证检测结果准确, 同时还能应对考虑概念漂移现象。

4 总结和展望

基于统计的方法由于数据分布快速且高效, 鲁棒性较好适合低维数据, 但对高维数据处理受限制。基于机器学习的方法克服了传统统计方法不能处理高维数据的问题。随着数据量的

增多、动态数据的产生, 对处理数据的速度、实时性有了更高的要求。大数据中的批处理方式处理速度较慢, 可以采用基于Storm、Spark、Flink等流式处理框架来实现实时计算和分析, 并且高效准确地检测出异常值。随着机器学习、神经网络等领域的发展, 又出现了基于层级实时记忆的异常检测算法, 它不需要采用滑动窗口法批处理数据, 就可以实现对数据的实时检测。

随着海量数据以及动态数据的产生, 除了采用流处理框的并行化和实时的计算方法, 还需要继续改进算法的性能, 进而实时检测更多的异常数据。比如, 虽然HTM算法具有较强的自适应性, 可以实现异常数据实时检测。但它也存在时间复杂度的问题, 要使之应用广泛还需要进一步研究改进。与此同时, 对于数据类型的增多和应用领域的扩大, 可以研究通过丰富数据编码的方式来实现不同类型的数据异常检测, 同时实时数据异常检测效果更好。面对数据不平衡等问题, 可以通过从大量数据中学习获得准确有效的特征, 建立基于深度学习的神经网络模型, 进而提高异常检测的效率^[32]。对于不平衡数据的处理也可以采用基于深度学习的异常检测方法^[33]。在异常数据实时检测的未来发展中, 基于层级实时记忆的和神经网络模型的方法改进, 以及基于深度学习的方法是一种趋势。

5 结语

(1) 新型的异常数据实时检测方法不仅加快了海量动态数据的处理速度, 也可以进行实时

异常数据检测并实时反馈检测结果。与此同时,还提高了异常数据检测的效率和准确率。

(2) 异常检测方法都需要快速且实时地检测出结果,这样才能最大程度地挽回损失或避免发生更大的事故。异常数据实时检测还要继续进一步研究,准确、高效地应用到工业生产、医疗技术、物联网检测、应急管理等领域中才具有实际意义。因此,异常数据实时检测方法具有广阔的应用前景。

参考文献:

- [1] ANSCOMBE F J, GUTTMAN I. Rejection of outliers[J]. *Technometrics*, 1960, 2(2): 123-146.
- [2] FREEMAN J. Outliers in statistical data[J]. *Journal of the Operational Research Society*, 1995, 46(8): 1034-1035.
- [3] 赵曼. 基于数据相关性的异常检测算法研究[D]. 北京: 北京交通大学, 2017.
- [4] KNOX E M, NG R T. Algorithms for mining distance-based outliers in large datasets [C] // *Proceedings of the International Conference on Very Large Data Bases*, 1998: 392-403.
- [5] 金文进, 杨武. 异常检测技术研究综述[J]. *软件导刊*, 2008(1): 10-13.
- [6] 卓琳, 赵厚宇, 詹思延. 异常检测方法及其应用综述[J]. *计算机应用研究*, 2020, 37(S1): 9-15.
- [7] BREUNIG M M, KRIEGEL H P, NG R T, et al. LOF: identifying density-based local outliers [C] // *ACM SIGMOD Record*, Dallas, Texas, USA, ACM, 2000, 29(2): 93-104.
- [8] 江驰, 谭斌, 黄明祥, 等. 污染源在线监控数据异常智能检测方法研究[J]. *环境科学与管理*, 2016, 41(10): 114-117.
- [9] 王美琪, 李建. 一种改进K-means聚类的近邻传播最大最小距离算法[J]. *计算机应用与软件*, 2021, 38(7): 240-245.
- [10] 吕云洁. 基于数据流的异常检测技术及应用的研究[D]. 河北: 华北电力大学, 2008.
- [11] 祁欣妍. 基于深度学习的物联网实时数据异常检测及分类技术研究[D]. 江苏: 东南大学, 2021.
- [12] 刘杰. 海量电能质数据的异常状态实时检测技术研究[D]. 北京: 北方工业大学, 2019.
- [13] 罗翔宇. 面向智能电网的WAMS实时数据处理与异常检测方法研究[D]. 河北: 华北电力大学, 2018.
- [14] 于宁. 云环境中时序数据的异常检测算法研究[D]. 北京: 中国石油大学, 2020. DOI: 10.27643/d.cnki.gsybu.2020.000527.
- [15] 杨力平. 智能电网大数据实时流处理方法研究[D]. 河北: 华北电力大学, 2016.
- [16] 梁汝鹏, 毛泽湘, 郑少秋, 等. 交通流数据实时处理与智能事件检测方法研究[C] // *第八届中国指挥控制大会论文集*, 2020: 501-505. DOI: 10.26914/c.cnkihy.2020.014664.
- [17] 栗国保, 韩青菊. 一种改进K-means聚类算法的MapReduce并行化实现[J]. *数字技术与应用*, 2016(12): 134, 136. DOI: 10.19695/j.cnki.cn12-1369.2016.12.098.
- [18] 王金元, 王宇, 张亚松, 等. 基于Hadoop的分布式财务异常数据分析系统设计[J]. *信息技术*, 2022(1): 21-25, 31. DOI: 10.13274/j.cnki.hdzt.2022.01.005.
- [19] 王诚, 狄莹. 孤立森林算法研究及并行化实现[J]. *计算机技术与发展*, 2021, 31(6): 13-18.
- [20] 陈国瑞, 袁旭华. 基于HDFS开源架构的异常数据实时检测算法[J]. *计算机仿真*, 2021, 38(8): 445-449.
- [21] 姜晓艳, 张伟, 王佳慧, 等. 基于流处理改进的SK-means策略[J]. *北京信息科技大学学报(自然科学版)*, 2021, 36(5): 51-56. DOI: 10.16508/j.cnki.11-5866/n.2021.05.009.
- [22] 吴海波, 施式亮, 念其锋. 瓦斯浓度流数据实时异常检测方法[J]. *计算机与数字工程*, 2019, 47(5): 1086-1090, 1105.
- [23] 宋若宁. 海量数据环境下的网络流量异常检测的研究[D]. 北京: 北京邮电大学, 2015.
- [24] 王德文, 杨力平. 智能电网大数据流式处理方法与状态监测异常检测[J]. *电力系统自动化*, 2016, 40(14): 122-128.
- [25] 孙顺苗. 基于数控机床运行数据的状态监测平台研究与开发[D]. 重庆: 重庆大学, 2019. DOI: 10.27670/d.cnki.gcqdu.2019.000715.
- [26] 刘子豪. 基于Flink的大规模水文传感器数据异常检测系统的研究与应用[D]. 江苏: 江苏科技大学, 2020. DOI: 10.27171/d.cnki.ghdcc.2020.000559.
- [27] 曾惟如, 吴佳, 闫飞. 基于层级实时记忆算法的时间序列异常检测算法[J]. *电子学报*, 2018, 46(2): 325-332.
- [28] 程学林, 郑佳卉, 蒋烁森, 等. 企业多云时序数据实时监测方案研究与实现[J/OL]. *小型微型计算机系统*: 1-9 [2022-03-31]. <http://kns.cnki.net/kcms/detail/21.1106.TP.20211015.0027.006.html>.
- [29] 牛德姣, 周时颖, 蔡涛, 等. 面向多核的并发HTM空间池算法[J/OL]. *小型微型计算机系统*: 1-10 [2022-03-31]. <http://kns.cnki.net/kcms/detail/21.1106.TP.20210906.1437.021.html>.
- [30] 刘庆宇. 信息物理融合环境下电网虚假数据实时检测算法研究[D]. 河北: 燕山大学, 2020. DOI: 10.27440/d.cnki.gysdu.2020.000208.
- [31] 时磊. 基于LSTMs-Autoencoder的流数据异常检测

- 算法[J]. 仪表技术与传感器, 2021(10):120-125.
- [32] 吴镜锋, 金炜东, 唐鹏. 数据异常的监测技术综述[J]. 计算机科学, 2017, 44(S2):24-28.
- [33] 胡姣姣, 王晓峰, 张萌, 等. 基于深度学习的时间序列数据异常检测方法[J]. 信息与控制, 2019, 48(1):1-8. DOI:10.13976/j.cnki.xk.2019.8062.

作者简介:

吴琼(1996—), 女, 陕西宝鸡人, 硕士研究生, 研究方

向为数据分析、应急管理信息化与智能化

通信作者: 李永飞(1978—), 男, 河北邯郸人, 硕士研究生, 副教授, 研究方向为数据挖掘、数据安全等, E-mail:305283277@qq.com

李铭洋(1997—), 男, 内蒙古通辽人, 硕士研究生, 研究方向为数据预测、应急管理信息化与智能化

收稿日期: 2022-04-11 修稿日期: 2022-05-28

A Review of Real-Time Detection Methods for Abnormal Data

Wu Qiong, Li Yongfei, Li Mingyang

(School of Computer Science, North China University of Science and Technology, Sanhe 065200)

Abstract: With the increasing value of data, the reliability of data is becoming more and more important, and abnormal data detection has gradually become one of the research hotspots in the field of data mining. The massive data generated by Internet of Things monitoring brings new challenges to traditional anomaly detection methods. This paper introduces traditional anomaly detection methods and real-time anomaly detection methods for massive dynamic data respectively, and points out that a stream processing framework combined with big data technology can speed up data processing and feedback detection results in real time; at the same time, some new advances based on machine learning and neural networks. The new real-time detection algorithm of abnormal data can not only realize real-time abnormal detection, but also improve the accuracy of abnormal detection. Real-time detection of abnormal data can be better achieved by combining new data processing techniques.

Keywords: anomaly detection; classification method; real-time; detection algorithm

(上接第8页)

A Small-Corpus Chinese Text Classification Method for Optimizing Pre-trained Models

Chen Lan, Yang Fan, Zeng Zhen

(School of Information, Guizhou University of Finance and Economics, Guiyang 550000)

Abstract: Aiming at the problem of insufficient representation of word vectors generated by GloVe and BERT models in small corpora, a fusion vector pre-training model was proposed to improve the accuracy of Chinese short text classification in small corpora. Taking today's headline public data set as the experimental object, using GloVe and BERT models through domain pre-training, vector fusion of pre-trained word vectors generated by GloVe and BERT to achieve semantic enhancement, thereby improving the short text classification effect. When the amount of data in the corpus is 500, the accuracy of the fused word vector is improved by 5 percentage points compared to the accuracy of the BERT word vector, and the accuracy of the GloVe word vector is improved by 3 percentage points. The dimension of word meaning selection needs to be further strengthened. The proposed method can accurately classify short text data in small corpus, which is of great significance for subsequent text mining work.

Keywords: BERT; GloVe; vector fusion; small corpus; short text classification

文章编号: 1007-1423(2022)16-0016-06

DOI: 10.3969/j.issn.1007-1423.2022.16.003

基于藏传佛教纹饰符号构造纹理的信息隐藏方法

吴秀琪, 张春玉, 丁泽威, 郭晨飞

(西藏民族大学信息工程学院, 咸阳 712082)

摘要: 利用藏传佛教纹饰符号生成具有藏文化特征纹理的含密纹理图像, 进而实现信息隐藏。信息隐藏者首先构造一个包含不同特征的纹饰符号库, 对其特征选择或构造特征值进行编码, 以实现二进制数据与纹饰符号之间的映射关系; 根据秘密信息从纹饰符号库中选择对应的纹饰符号, 并确定合适的坐标值进行摆放, 对源含密纹理图像进行可逆形变操作, 生成初步的含密纹理图像, 然后随机选择干扰符号和相近的背景颜色进行填充, 最终生成一副藏文化特色的具备良好视觉美学效果的含密纹理图像。信息提取者对生成的含密纹理图像去除背景颜色和干扰符号, 然后进行逆变换操作, 根据特征解码得到秘密信息。实验证明该方法比其他方法生成类似加密数据的含密纹理图像更具民族特色, 因而不易引起攻击者怀疑, 视觉上更具隐蔽性。

关键词: 信息隐藏; 藏传佛教纹饰符号; 可逆形变; 生成纹理

基金项目: 大学生创新创业训练计划项目(202110695027): 基于藏传佛教纹饰符号构造纹理的信息隐藏方法

0 引言

信息隐藏技术的相关研究最早开始于上个世纪 90 年代中期, 历经多年的发展成为通讯信息安全领域的重要研究方向^[1-2]。传统的信息隐藏方式大多是将秘密信息隐匿于普通的媒体介质中, 信息窃取者发现秘密信息的难度增加, 从而达到信息隐蔽传输的目的^[3]。纵观信息隐藏技术的发展进程, 出现了许多像数字水印技术、隐写术、可视密码技术等十分优秀的隐写方法, 为国家信息安全体系的完善以及信息安全水平的提升打下坚实的基础^[4-6]。

鉴于当前信息隐藏的技术大多采用对常见的数字载体(视音频、数字图像等)进行适当的修改, 而后嵌入秘密信息的方式, 攻击者不断优化隐写分析算法, 通过分析秘密信息与媒体介质在统计特征上的差异, 判别秘密信息在媒体介质中的隐藏位置, 预测秘密信息的数据量, 使得秘密信息暴露的可能性大大增加^[7-11]。此类基于载体修改式的信息隐藏只要修改则一定会留下痕迹, 从而为隐写分析者留下可乘之机。

为了从根本上消除修改载体带来的风险, 零隐藏概念、无载体信息隐藏技术被提出。

无载体信息隐藏发展至今, 大致可以分为编码/映射式和生成式。其中生成式无载体信息隐藏也称为构造式信息隐藏, 是指在不预先指定原始载体的情况下, 由秘密信息按照一定规则直接生成含密载体。该类方法又分为半构造式和完全构造式。半构造式信息隐藏由事先给出的载体构造预设条件, 根据秘密信息并遵循一定的构造规则生成含密载体, 含密载体属于特定类型。Otrori 和 Kuriyama 教授最先提出不提前指定数字载体, 在纹理合成的过程中实现数据嵌入的基本思路^[12-13]。该方法在常规的纹理样本图像中选取多个彩色样本点, 然后将二值数据和彩色样本点之间的关系用 LBP 码进行映射, 最终根据秘密信息的内容从纹理样本图像中选取恰当的纹理图像, 生成一副具有传统美术工艺的含密纹理图像。

上述基于纹理合成的半构造式信息隐藏方法, 生成的纹理图像类似加密数据, 反而容易

引起攻击者怀疑。本文以藏传佛教特有的纹饰符号为基础, 利用模拟水影画(又称湿拓画, marbling)的方法来隐藏秘密信息^[14-15], 该方法需要先在一张白纸上写出需要传递的秘密信息, 利用模拟 marbling 的方法对该图案进行形变, 再用相似的背景图案对其余的空白部分进行填充, 最后生成具有藏文化特色的含密纹理图像。该方法不仅具有一定的隐写能力, 并且生成的藏文化特色纹理不容易引起攻击者怀疑, 视觉隐蔽性更强。

1 总体方案

本文提出的信息隐写方法的总体方案如图 1 所示。在信息隐藏阶段, 首先构建纹饰符号库, 其中包含多种单元, 包括图形的颜色、纹理等

特征, 按照其特征, 选择或构造特征值进行编码。其次, 确定纹饰符号与二进制数据之间的映射关系, 从而形成秘密信息、二进制数据与纹饰符号之间的映射表。再次, 将秘密信息分成两部分, 一部分从纹饰符号库中选择相应的单元, 确定每个单元的图形特征和背景颜色, 同时根据另一部分秘密信息在画布上选择合适的坐标值, 然后将选定的纹饰符号摆放到画布相应的坐标值处形成一幅源含密图像。最后, 在填充过程中利用 marbling(水影画)方法对生成的图像进行多次齿状变换和正弦变换, 随机选择干扰符号摆放到画布上, 剩余空白部分选择与纹饰符号相近的颜色来填充, 得到具有复杂纹理结构的含密纹理图像。

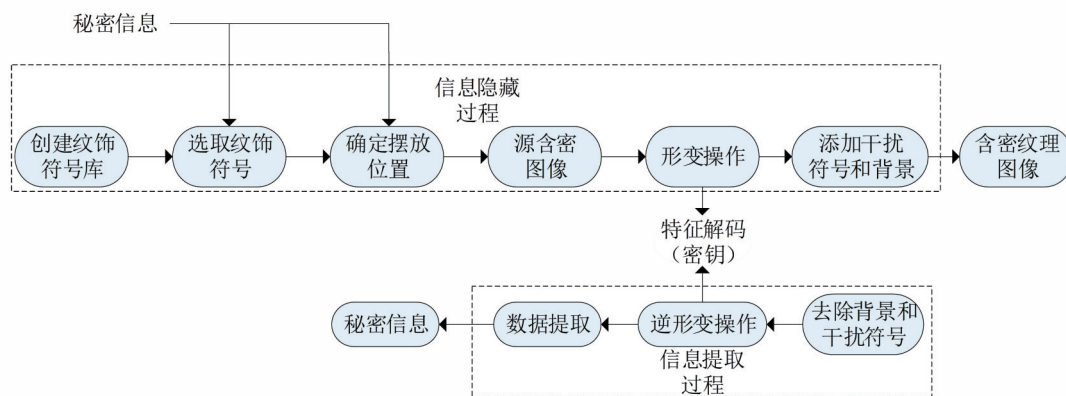


图 1 总体方案流程图

秘密信息的提取是信息隐藏的逆过程, 信息提取者和信息隐写者需要事先约定共享相同的纹饰符号库, 然后信息提取者对含密图像去除背景颜色和干扰符号, 并进行逆变换操作, 进而根据特征解码得到秘密信息。

2 隐写方案

2.1 构建纹饰符号库

根据秘密信息绘制源图像前, 选择具有不同特征的纹饰符号来代表不同的二进制数据, 本文将前期搜集的藏传佛教纹饰符号进行整理分类, 并将其特征种类记为 $E = \{E_1, E_2, E_3\}$, 其中 E_i 可表示为纹饰符号、颜色及坐标等。每种 E_i 可有多种取值, $E_i = \{S_1, S_2, \dots, S_n\}$, 其中 $S_1, S_2,$

\dots, S_n 表示不同纹饰符号, $E_2 = \{C_1, C_2, C_3, C_4\}$, 其中 C_1, C_2, C_3, C_4 分别表示红、绿、蓝、黄四种颜色。将二进制数据与 E_i 的取值空间相互对应, 则 E_i 的每个取值可与多个比特相对应。

表 1 给出一个简单的示例, 此纹饰符号库中包含两个特征种类: 纹饰符号 E_1 和颜色 E_2 。本实例所选的四个纹饰符号均为前期搜集整理的纹饰符号库中随机选取, 每种纹饰符号对应 2 比特的二进制数据, 颜色包含四种, 分别是红色、蓝色、黄色和绿色, 每种颜色也对应 2 比特的二进制数据。此处需要说明的是, 由于信息隐藏者和信息提取者均需要利用该纹饰符号库对秘密数据进行编码, 因此信息提取者和信息隐藏者需要事先约定共享相同的纹饰符号库。

表 1 纹饰符号库及编码示例

二值数据	00	01	10	11
纹饰符号 E_1	7	×	5	8
二值数据	01	10	11	00
颜色 E_2	红	绿	蓝	黄
背景颜色 R	$R_1: [255, 99, 71]$	$R_2: [173, 255, 47]$	$R_3: [135, 206, 235]$	$R_4: [255, 215, 0]$

此外, 该纹饰符号库中应该还包括已定义完成的背景颜色, 如本示例中的 R 值, 此 R 值采用 RGB 颜色值。应注意的是, 选择的背景颜色应包含在纹饰符号库的特征种类颜色 E_2 中, 同时应与选择的纹饰符号的颜色相近, 以便于纹饰符号的隐写。

2.2 嵌入秘密信息

信息隐藏者根据纹饰符号库中定义的各项特征的取值空间决定每个纹饰符号携带秘密信息的数量, 而在信息隐藏过程中, 每个纹饰符号包括多种特征, 因此仅需选择部分特征来表示秘密数据。本文将秘密信息分为三部分, 第一部分秘密信息为 E_1 对应的藏传佛教纹饰符号, 本文选取典型的统计特征(如像素和等)作为纹饰符号的特征值; 第二部分秘密信息为 E_2 对应的不同颜色, 可以采取 2 比特的二进制数据表示纹饰符号颜色; 第三部分 E_3 为对应符号在画布上的坐标。在画布上作一个 $m \times n$ 直角坐标系, 随机选取 K 个点, 并将这些点作为每个纹饰符号的形心坐标。

信息隐藏者需要将每个单元(纹饰符号和颜色确定后的图形)摆放到画布上, 由于目前常见的信息隐藏单位通常是 1 个字节, 因此可以选择 2 比特的秘密数据来表示每个单元在画布上的形心坐标位置。此外, 由于每个单元的形状、颜色及其摆放位置等特征均由秘密数据决定, 为方便后续秘密信息的提取, 一般将每个单元在画布上的位置摆放得较为稀疏, 同时在画布上也会摆放多个干扰符号, 使得信息窃取者难以确定秘密信息隐藏的位置, 继而提高信息隐藏的安全性。

2.3 生成含密图像

在确定 K 个单元纹饰符号的坐标及其特征值后, 将确定的纹饰符号嵌入到画布上, 以画

布的最左上方为坐标原点, x 轴从左向右递增, y 轴从上向下递增。在嵌入纹饰符号前, 需要预设好一定面积内可摆放的纹饰符号的最大数目。信息隐藏者一般通过对源含密图像进行形变操作来达到具有水影画效果的纹理图, 对图像进行形变操作需要用到形变函数, 常见的形变函数有齿状变换、斜变换、正弦变换、正切变换、幂函数变换等。进行信息隐藏的过程中, 为生成较为复杂的纹理图, 可选择多种形变函数对源含密图像进行形变操作, 本文仅选取齿状变换和正弦变换来进行形变操作。

正弦变换是进行形变操作的常用函数之一, 假设源含密图像中某一纹饰符号的摆放位置为 $M(x, y)$, 经过操作后得到 $M_1(x_1, y_1)$, 使用的操作为

$$z = x \sin t - y \cos t \quad (1)$$

$$k_1 = A \sin(wz + \theta) \cos t \quad (2)$$

$$k_2 = A \sin(wz + \theta) \sin t \quad (3)$$

$$x_1 = x + k_1 \quad (4)$$

$$y_1 = y + k_2 \quad (5)$$

以上各式中, w 为正弦变换的波长, A 为振幅, θ 为相位, t 为形变的偏移量。

另一种齿状变换在 x 方向和 y 方向分别进行形变操作, 假设源含密图像中某一纹饰符号的摆放位置为 $M(x, y)$, 经过齿状变换后的映射点为 $M_2(x_2, y_2)$, 则变换过程为

$$x_2 = x \quad (6)$$

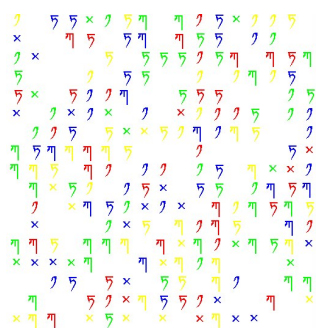
$$y_2 = y - \frac{ay}{b + y} \quad (7)$$

$$d = s/2 - |\text{fmod}(x, s) - s/2| \quad (8)$$

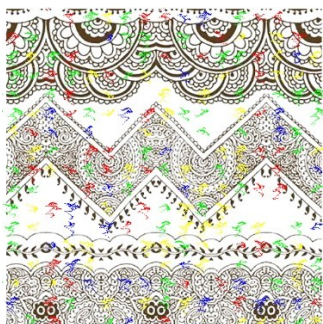
以上各式中, α 表示最大位移量, s 控制齿形间距的大小, $\text{fmod}(\cdot)$ 为求余函数。

对源含密图像进行形变操作后, 信息隐藏者在源含密图像的背景空白处添加一些与已有符号的特征相似的背景图形, 即干扰符号, 而

干扰符号由纹饰符号库中纹饰符号不规则变换产生,同时填充与选定的纹饰符号颜色相近的背景颜色,以增加隐秘性,达到视觉上较为协调的效果,如图2所示。



(a) 源含密图像



(b) 初步绘制后的图像

图2 初次绘制图像

在添加干扰符号后,信息隐藏者对初步绘制完成的源含密图像仍然可以进行多次形变操作,从而获得更为复杂的纹理图像。为了提高秘密信息的安全性,隐写者通过对形变的顺序和参数进行加密,生成读取秘密信息所必须的密钥,通过信任的安全通道将密钥发送给提取者。

2.4 提取秘密信息

提取秘密信息时,提取者与隐写者共享纹饰符号库,根据纹饰符号库中事先定义好的背景色除去图像中填充的背景图案,得到仅包含隐写单元的图像。在接收到发送者发送的密钥后,根据密钥对源含密图像进行形变的反操作。然后,利用卷积滤波确定每个符号所处位置,读取其形心坐标,进而提取其中部分的秘密信息。最后,根据纹饰符号库不同特征对应的秘密数据,提取得到相应的秘密信息。

3 实验结果

本文通过大量实验来验证此隐写算法,在实验中首先建立一个纹饰符号库,库中包含的特征有藏传佛教纹饰符号和颜色, η 代表“00”, \times 代表“01”, ζ 代表“10”, ϖ 代表“11”;颜色特征包括四种:红色、绿色、蓝色和黄色,分别表示二进制数据“01”“10”“11”“00”,即每个单元可表示4比特的秘密信息;同时秘密信息在画布上的摆放位置也代表2比特的秘密信息,也即每个单元可表示6比特的秘密数据。此外,纹饰符号库中还定义了四种背景颜色,分别是[255,99,71]、[173,255,47]、[135,206,235]、[255,215,0]。

随机产生一定比特的二进制数据作为秘密信息,将这些秘密信息划分成每6比特作为一个秘密信息数据单元,根据每个秘密信息数据单元从纹饰符号库中选取对应的纹饰符号,确定在画布上的摆放位置。图3是生成含密纹理图像的实验结果,其中图3(a)是从纹饰符号库中选取对应的纹饰符号摆放到画布上,画布大小为 512×512 ;图3(b)是在图3(a)的基础上进行正弦变换和齿状变换,添加干扰符号,初步生成的含密纹理图像;图3(c)是在图3(b)的基础上添加背景颜色得到含密纹理图像。其中使用的参数为:正弦变换中 $w = 50$ 、 $A = 12$ 、 $\theta = \pi/3$ 、 $t = \pi/2$,齿状变换中 $\alpha = 40$ 、 $\gamma = 84$ 、 $s = 20$ 。由图3(c)可观察到,此方法生成的含密纹理图像具有良好的视觉美学效果。

信息提取者根据接收的具有传统视觉美学效果的含密纹理图像图3(d),首先对其去除背景颜色和干扰符号,可以得到图3(e);再进行正弦逆变换和齿状逆变换,根据特征解码获取含密数据单元和位置坐标,如图3(f),从纹饰符号库中查找对应的纹饰符号的特征信息,最终得到秘密信息。

通过添加不同的干扰符号和背景颜色或采取不同的形变操作,可以得到具有不同视觉效果含密纹理图像。图4是添加不同的干扰符号的一组实验结果,其中图4(a)是在 2048×2048 画布上摆放纹饰符号后的图像;在图4(a)的基础上经过一次正弦变换,然后再进行两次不同

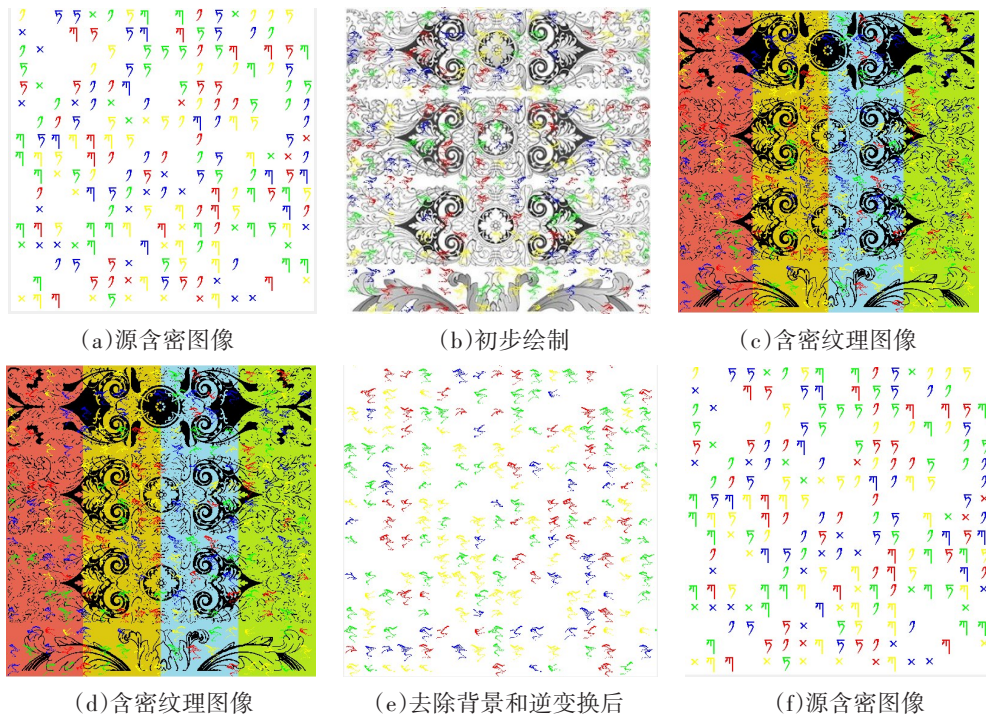


图3 含密纹理图像的生成与秘密信息的提取

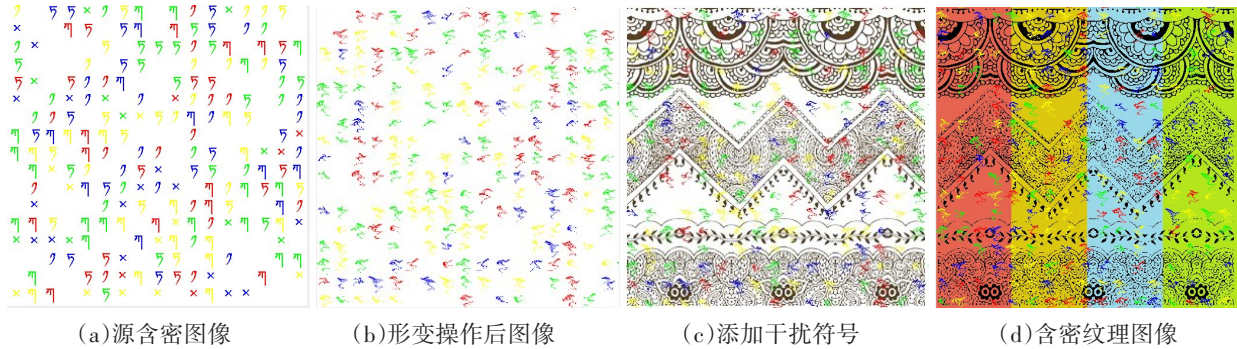


图4 含密纹理图像的生成

方向上的齿状变换后可以得到图4(b)；在图4(b)上添加与上一组实验不同的干扰符号后可获得图4(c)；添加上背景颜色后可获得含密纹理图像，如图4(d)所示。

大量实验表明，采用添加藏文化特征的干扰符号、背景色等方法，能成功地利于藏传佛教纹饰符号进行形变的模拟 marbling 操作，进而生成具有藏文化特色的纹理图。该含密纹理图区别于其他同类方法生成的类似加密数据的纹理图像，因为具有明显的藏文化特色而接近自然图像，不容易引起攻击者的怀疑，因此具有更强的隐蔽性。

4 结语

西藏在信息化建设方面取得了巨大成功，信息安全也成为当前研究重点。为了进行西藏特色的信息安全研究，本文通过模拟水影画的方法实现基于藏传佛教纹饰符号的半构造式信息隐藏。在隐藏过程中充分考虑藏文化特色的图像语义特征，使含密纹理图像的生成与秘密数据的隐藏同步进行。用藏传佛教纹饰符号不同的纹理特征表征二进制数据，建立纹饰符号与二进制数据之间的映射关系，在初次绘制的原图像上添加干扰符号和背景颜色，并进行一系列的可逆形变操作，最终生成一副具有藏文

化特色的含密纹理图像。实验结果表明,该方法较同类方法生成的含密纹理图像不仅具有更好的视觉美学效果,而且隐蔽性更强。

参考文献:

- [1] 潘琳,钱振兴,张新鹏. 基于构造纹理图像的数字隐写[J]. 应用科学学报,2016,9:625-632.
- [2] 张新鹏,钱振兴,李晟. 信息隐藏展望研究[J]. 应用科学学报,2016,9:475-489.
- [3] FRIDRICH J. Steganography in digital media: principles, algorithms and applications [M]. Cambridge: Cambridge University Press,2009.
- [4] 司广文,秦川,姚恒,等. 基于纹理特征分类与合成的鲁棒无载体信息隐藏[J]. 应用科学学报,2020,38(3):441-454.
- [5] 沈淑涛. 基于纹理度划分的藏文古籍图像隐藏信息检测方法[J]. 财富时代,2020(04):55-56.
- [6] 张敏情,时晨曦. 基于互信息和特征融合的图像隐写分析[J]. 武汉大学学报(理学版),2013(6):551-557.
- [7] 张新鹏,殷赵霞. 多媒体信息隐藏技术. 自然杂志,2017,(04).
- [8] WANG H, WANG S. Cyber warfare: steganography vs. steganalysis[J]. Communication of ACM,2004,47(10):76-82.
- [9] FRIDRICH J, SOUKAL D. Matrix embedding for large payloads[J]. IEEE Transactions Information Forensics and Security,2006,1(3):390-395.
- [10] ZHANG X, WANG S. Dynamical running coding in digital steganography[J]. IEEE Signal Processing Letters,2006,13(3):165-168.
- [11] ZHANG W, ZHANG X, WANG S. Near-optimal codes for information embedding in gray-scale signals [J]. IEEE Transactions Information Theory,2010,56(3):1262-1270.
- [12] OTORI H, KURIYAMA S. Data-embeddable texture synthesis [C] //Proceedings of the 8th International Symposium on Smart Graphics, Kyoto, Japan, 2007:146-157.
- [13] OTORI H, KURIYAMA S. Texture synthesis for mobile data communications [J]. IEEE Conference Graphics Applications,2009,29(6):74-81.
- [14] WU K C, WANG C M. Steganography using reversible texture synthesis [J]. IEEE Transactions Image Proceedings,2015,24(1):130-139.
- [15] XU J, MAO X, JIN X. Hidden message in a deformation-based texture [J]. Visual Computer, 2015,31:1653-1669.

作者简介:

吴秀琪(2000—),女,山东临沂人,本科生,研究方向为信息安全方面

张春玉(1979—),女,陕西汉中,人,硕士,副教授,硕士生导师,研究方向为信息安全方面

丁泽威(2001—),男,山西运城人,本科生,研究方向为移动通信与无线技术方面

郭晨飞(1997—),男,陕西西安人,硕士研究生,研究方向为图像处理方面

收稿日期:2022-04-13 修稿日期:2022-05-08

Information Hiding of Textures Based on Tibetan Buddhist Decorative Symbols

Wu Xiuqi, Zhang Chunyu, Ding Zewei, Guo Chenfei

(School of Information Engineering, Xizang Minzu University, Xianyang 712082)

Abstract: The use of Tibetan Buddhist decoration symbols to generate a dense texture image with the texture of Tibetan cultural characteristics to hide information. Information hidden first constructs a decorative symbol library that contains different features, and encodes its feature selection or constructive feature value to achieve the mapping relationship between binary data and decorative symbols; select the corresponding decoration from the decorative symbol library according to the secret information from the secret information. Symbols, and determine the appropriate coordinate value to place it, carry out reversible formats for the source of the dense texture, generate a preliminary dense texture image, and then select the interference symbols and similar background colors to fill it, and finally generate a pair of Tibetan culture. Features with a good visual aesthetic image containing texture. The information extractor removes the background color and interference symbols of the generated texture -containing texture, and then the inverter replacement operation is performed, and the secret information is obtained according to the characteristic decoding. Experiments have proved that this method has more national characteristics than other methods to generate a dense texture image similar to encrypted data, so it is not easy to cause the attacker to doubt, and it is more concealed visually.

Keywords: information hiding; Tibetan Buddhist ornamental symbols; reversible deformation; texture generation

文章编号: 1007-1423(2022)16-0022-06

DOI: 10.3969/j.issn.1007-1423.2022.16.004

基于注意力机制 BiLSTM-CharCNN 的药物不良反应监测方法

蒯娟霞, 徐 鹏, 叶思维

(广东东软学院信息管理与工程学院, 佛山 528225)

摘要: 推特等社交媒体为药物不良反应监测提供了新的信息源。本文结合深度学习技术从海量的社交媒体用户评论中监测出其是否有关药物不良反应推文。为了提高精度, 丰富推文的信息特征, 提出一种端到端的基于注意力机制 BiLSTM-CharCNN 神经网络模型。首先, 使用 CharCNN 模型提取字符级特征: 对文本进行字符数字化, 再做卷积池化和全连接操作, 得到字符级特征向量; 然后, 使用 Attention-BiLSTM 模型提取特征向量: 通过调整权重参数, 去除掉冗余信息, 提取关键部分的信息, 对文本进行优化; 最后, 将两种方法提取到的特征信息进行连接, 再到包含一个全连接和一个 softmax 的输出层进行分类操作。实验结果证明, 该方法可以显著提高分类的准确性, 且效果优于现有的基准方法。

关键词: 药物不良反应; 社交媒体; 深度学习; Attention; BiLSTM-CharCNN

0 引言

药物不良反应(adverse drug reactions, 简称 ADR)是指患者在使用某种药物治疗疾病时所引发的有害反应。及时地知道药物不良反应对医药公司和监管机构非常重要。获得药物不良反应的传统方式具有时效低、更新慢的特点, 因此需要找到一个具有时效性高, 更新快的方法。人们在社交媒体上发表健康情况的推文满足以上两个特点, 因此社交媒体也就成为药物不良反应的研究数据来源。社交媒体的数据是用户用药的第一手资料, 时效性高, 覆盖率广, 但如何在海量的数据里找到与 ADR 相关的推文, 这对 ADR 研究是一项巨大的挑战。

Sarker 等^[1-2]对多种社交媒体资源进行全面挖掘, 实现了多种特征的结合分析, 这些特征包括 n-gram 特征、词典特征、极性特征、情感分数特征和主题模型特征, 通过实验发现 SVM 算法的性能最好。Korkontelios 等^[3]关注到了情绪分析在药物不良反应识别任务中的作用, 他们利用一个新颖的情感分析算法并结合支持向量机(SVM)分类器, 来实现对推特数据集上的

药物不良反应任务的识别及提取, 其在推特数据集上的 $F1$ 值达到了 69.16%。Cocos 等^[4]利用双向长短时记忆网络(BiLSTM)训练推特数据集, 其推文数量为 844, 最后的 $F1$ 值为 0.755。Xia 等^[5]利用迁移学习结合 LSTM 模型, 在迁移学习的实例、特征表示、参数、关系知识 4 个方面进行了实验, 取得了良好的效果, 使用迁移学习解决了数据稀缺的重要难题。Rezaei 等^[6]使用推特的数据集, 通过 CNN、HNN 和 FastText 三类深度学习网络对其进行分类, 这三种模型的输入都采用 word2Vec 向量, 最后证明了其方法的有效性。

本文提出基于注意力机制的 BiLSTM-CharCNN 药物不良反应推文识别方法。通过将字符级向量与词向量相结合的方法来对推文进行识别。以社交媒体上人们讨论健康情况的推文为实验数据, 并将实验结果与 BiLSTM, CharCNN, BiLSTM+Attention, BiLSTM+CharCNN 的结果进行对比, 以召回率(*Recall*)、精确率(*Accuracy*)、 $F1$ 值为评价指标, 验证了本文所提出的基于注意力机制的 BiLSTM-CharCNN 药物不良反应识别的有效性。

1 模型方法

1.1 词向量表示

词向量指的是将高维度离散的文本数据转化为低维度密集向量^[7-10], 本文方法中的词嵌入部分能够准确地抓取词语的语义。接着应用词嵌入模型 Word2vec, 它可以通过机器学习模型将原来不同的词转化为不同的实数向量。Word2vec 可以在大量文本数据集中进行局部训练, 训练得到的结果一词向量, 可以很好地拟出度量词与词之间的相关性。

1.2 长短时记忆网络 LSTM

长短时记忆网络(LSTM)^[11]是一种特殊的RNNs, 能有效地解决传统RNN在处理时间序列长期依赖中的梯度消失和梯度爆炸的问题。LSTM通过在RNN传输状态后加入细胞状态(C_t)来控制传输状态, 由于在传输过程中细胞状态变化较慢, 误差相对稳定, 可以在多个时间步上持续学习且在一定程度上加强LSTM的记忆能力。

LSTM的细胞结构如图1所示。

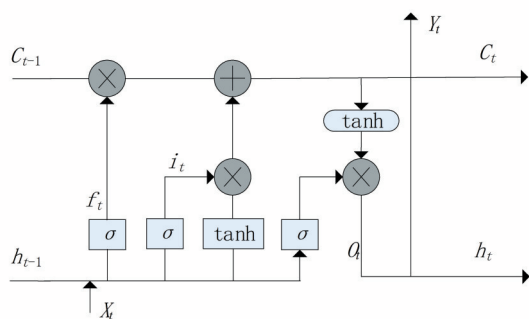


图1 LSTM单元结构

LSTM结构中的核心部分是图1中最上边的线-C, 叫做细胞状态(cell state), 它一直存在于LSTM的整个系统之中, 其中:

$$C_t = f_t * C_{t-1} + i_t * C_t \quad (1)$$

式(1)中 f_t 叫做遗忘门, 表示 C_{t-1} 的需要用作计算 C_t 的特征。 f_t 是一个向量, 向量中的每个元素的范围均介于 $[0, 1]$ 之间。通常使用sigmoid函数作为激活函数。上图中的 \otimes 代表LSTM中最重要的门机制, 遗忘门表示的是 h_{t-1} 和 x_t 之间的单位乘关系, 可由公式(2)表示:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2)$$

其中 \tilde{C}_t 表示细胞状态更新值, 由以下函数得到:

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (3)$$

i_t 叫做输入门, 由 x_t 和 h_{t-1} 经由sigmoid激活函数计算而来, 如下所示:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (4)$$

i_t 通过控制 \tilde{C}_t 的部分特征来更新 C_t , 同 f_t 的关系如下所示:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (5)$$

最后, 为了计算预测值和生成下一个时间片完整的输入, 需要计算隐节点输出, 如下式所示:

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (6)$$

$$h_t = O_t * \tanh(C_t) \quad (7)$$

由式(6)和式(7)可知, 隐节点输出 h_t 取决于输出门 O_t 和细胞状态 C_t , 且 O_t 的计算方式与 f_t 及 i_t 类似。

LSTM是一个信息单向传播的模型, 无法编码从后到前的信息。但是对于一些文本分类任务来说, 后边的信息依然可以影响前边的词句。为了解决这个问题, 提出了BiLSTM模型, 其主要思路是将后向关系和前向关系链接到同一个输出层, 且在其中共享权值。其网络结构如图2所示。

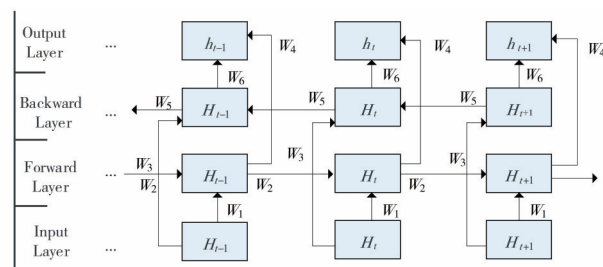


图2 BiLSTM结构图

在前向层从1时刻到 t 时刻正向计算一遍, 然后将结果保存; 同时, 从 t 时刻到1时刻, 逆向计算一遍, 同样将计算结果保存; 最后将正向计算和逆向计算的结果相加, 得到整个BiLSTM的计算结果, 其具体计算过程见下式:

$$h_t = f(w_1 x_t + w_2 h_{t-1}) \quad (8)$$

$$h'_t = f(w_3 x_t + w_5 h_{t+1}') \quad (9)$$

$$o_t = g(w_4 h_t + w_6 h'_t) \quad (10)$$

1.3 注意力机制

注意力机制 (attention mechanism)^[12]是一种模仿人类注意力的网络构架。在注意力资源有限的情况下,能够有效筛选和提炼大量复杂冗余信息中的高质量内容,可以同时聚焦多个细节部分。注意力机制可以弥补信息的长距依赖性这一缺陷。

1.4 注意力机制的 BiLSTM-CharCNN 模型

本文的模型结构如图3所示:主要包括输入层、特征提取层和输出层。

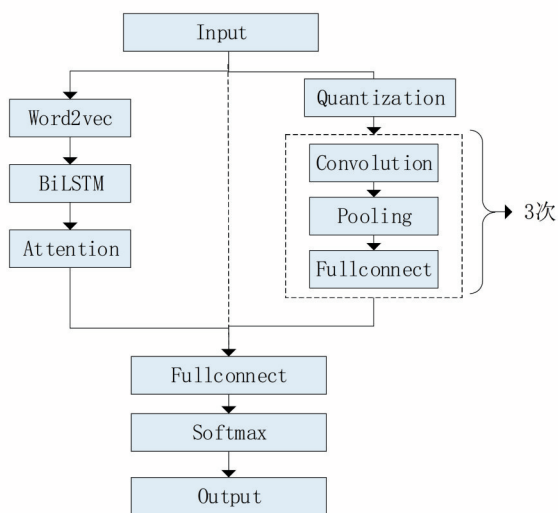


图3 模型框架图

输入层主要对原始的数据进行清洗、分词等操作,转化为较规范的文本形式;接下来,在特征提取阶段,使用了两种特征提取方法。

(1) 字符级特征提取:使用 CharCNN 模型提取字符级特征^[13]。首先对文本进行字符数字化,然后对其做卷积操作,对卷积的结果做 Max-pooling 池化操作,这样的卷积池化操作共执行3次。最后将其输入全连接层,得到字符级特征向量。

(2) 使用 BiLSTM-Attention 模型提取特征向量^[14-15]。首先,使用 word2vec 算法得到文本的词向量^[16],接下来将词向量输入到 BiLSTM 模型中,最后将得到的特征向量做 Attention 操作,这样就能通过调整权重参数去除冗余信息,提取关键部分的信息,对文本进行优化。再使用两种不同的方法进行特征提取,并将提取到的特征信息进行连接,然后输入到输出层。输出

层包含一个全连接层和一个 softmax 层,使用全连接层调整特征向量的维数,然后使用 softmax 分类器对文本进行分类操作。

1.5 算法描述

输入:原始数据集 X ; 类别标记 Y ;
输出:分类模型; 模型的评价指标;

(1) 数据预处理: $p_X = f_1(X)$, 函数 f_1 主要包含数据清洗和分词;

(2) 特征提取:

① BiLSTM + Attention: $F_{BA} = f_2(p_X)$ 函数 f_2 包含多个步骤,具体如下:

第一步:词嵌入: $V_{WE} = f_{\text{skipgram}}(p_X)$

第二步: BiLSTM: $F_{ba} = F_{\text{bi-lstm}}(V_{WE})$

第三步: Attention: $F_{BA} = F_{\text{Attention}}(F_{ba})$

② CharCNN: 函数 f_3 包含多个步骤,具体如下:

第一步: Quantization:

$$V_{\text{quantization}} = f_{\text{Quantization}}(p_X)$$

第二步: 卷积: $F_{\text{Conv}} = f_{\text{Conv}}(V_{\text{quantization}})$

第三步: 池化: $F_{\text{pooling}} = f_{\text{Max-pooling}}(F_{\text{Conv}})$

第四步: 重复计算第二步和第四步3次

③ $F_{\text{final}} = F_{BA} \oplus F_c$

(3) 输出层: softmax: $y = f_{\text{softmax}}(F_{\text{final}})$, 最后得到分类结果 y 。

2 实验与结果分析

2.1 实验数据

本文所采用的数据集是 SMM4H 共享任务评测的数据集。该数据集主要提供了推特用户发表推文的 ID 号和用户的 ID 号。由于存在用户删帖等因素的存在,最后一共收集到 7168 条推文。在实验之前,首先统计本研究的数据集,结果见表 1。

表 1 数据集统计

统计项目	训练集	验证集	测试集
Twitter 评论数	5734	1434	1500

由于本研究数据集的样本数较小,为了充分训练,本文采用 8:2 的数据集划分方式,最后测试集的数据集是在训练结束后在整个数据

集上随机采样得到的。

表2给出了推文的示例。由表2可以看到, 社交媒体的文本是不规范的, 充斥着各种符号和 URL 等信息, 所以需在实验前对数据进行文本预处理。

表2 推文示意表

推文	分类结果
Neutrogena Naturals Multi-Vitamin Nourishing Night Cream Review: Is It Safe & Effective? http://dlvr.it/MvDhgr	noADR
I did this when I had a heartburn saw gaviscon going for nearly a tenner I had to just painkillers and sleep it off https://twitter.com/_ezedamian/status/81082817598	hasADR

2.2 数据处理

因为推文数据都是不规则的、较口语化的文本内容, 所以首先需对推文数据进行文本预处理。表3显示了文本预处理的实例。

表3 推文预处理示意表

推文	预处理结果
Neutrogena Naturals Multi-Vitamin Nourishing Night Cream Review: Is It Safe & Effective? http://dlvr.it/MvDhgr	neutrogena naturals multi-vitamin nourishing night cream review: is it safe & effective? Url
I did this when I had a heartburn saw gaviscon going for nearly a tenner I had to just painkillers and sleep it off https://twitter.com/_ezedamian/status/810828175987052544...	i did this when i had a heartburn saw gaviscon going for nearly a tenner I had to just painkillers and sleep it off url

文本具体的处理方法为: 统一单词的大小写, 将所有单词转化为小写; 将所有出现的 URL、电子邮件地址以及提及的其他用户名分别替换为“url”“email”“atSign”; 推文中有一类标签文本, 表示推文的类型, 在本文的研究模型中未使用到, 所以将标签类型删掉; 将推文中的数字部分修改为一个代表性的数字。

2.3 超参数选择

本实验采用正则化和 Dropout 的方法来降低模型的复杂度, 避免过拟合。在 CharCNN 阶段

和 BiLSTM 阶段均使用 Dropout 的方法来避免过拟合, 在损失函数中引入了 L2 正则化, 既能提升准确度, 又能增加模型的鲁棒性。本文使用 Adam 梯度下降法来加速模型的收敛速度。

在神经网络中隐藏层数、单元个数、批量大小或迭代次数等参数也被称为超参数, 它们决定了特定神经网络结构的总体设置。本实验使用网格搜索的方法进行超参数选择, 即用预定义的超参数搜索空间的不同超参数组合来训练神经网络结构的多个不同版本, 并评估这些参数组合, 选择在验证集上评估性能最佳的配置作为最终的参数组合。

相关超参数设置: 词向量的维度设置为 200 维; CharCNN 阶段字符序列长度为 1014; CharCNN 卷积层层数设置为 3; LSTM 的时间步设置为 128, LSTM 输出神经元个数设置为 12; 训练 batch 大小设置为 64; 迭代 epoch 设置为 5; 学习率设置为 0.001; dropout 值设置为 0.9 (这是 dropKeepProb 的值)。

2.4 评价指标

本文研究内容是社交媒体上药物不良反应检测任务, 是自然语言处理中常见的分类任务。与药物不良反应有关, 标记为 1; 否则标记为 0。预测结果的混淆矩阵如表 4 所示。

表4 混淆矩阵

预测结果	实际结果	
	1	0
1	TP	FP
0	FN	TN

其中, TP 表示预测为与药物不良反应有关, 实际也与药物不良反应有关; FP 表示预测与药物不良反应有关, 而实际无关; FN 表示预测与药物反应无关, 而实际有关; TN 表示预测与药物不良反应无关, 实际也无关。

2.4.1 召回率

召回率表示模型实际为 1 的样本, 预测仍为 1 的样本概率, 其计算公式为:

$$R = \frac{TP}{TP + FN} \quad (11)$$

2.4.2 精准率

精准率是指在所有预测为1的样本中, 实际为1的样本比例, 其计算公式为:

$$P = \frac{TP}{TP + FP} \quad (12)$$

2.4.3 F_1 值

F_1 值是对召回率和精准率的综合评价指标, 是对其进行加权平均的结果, 其计算公式为:

$$F_1 = \frac{2PR}{P + R} \quad (13)$$

2.5 实验结果及分析

为了验证本研究提出模型的有效性, 本文选取 BiLSTM、CharCNN、BiLSTM + Attention 算法做对比实验, 通过与经典模型的对比分析, 能够准确地反映本模型的优越性, 其实验结果如表5所示。

表5 药物不良反应的二分类结果

Method	召回率 / %	精准率 / %	F_1 值 / %
BiLSTM	49.07	100	65.54
CharCNN	48.69	99.4	64.99
BiLSTM+Attention	78.47	84.73	81.31
BiLSTM+CharCNN	82.35	86.68	84.25
our model	85.05	86.27	85.45

表5中所有的数据都是在相同的实验参数和相同数据集中训练得到的结果, 其中单独的 BiLSTM 模型或者 CharCNN 模型在这个任务上的效果都欠佳, 可以看到召回率都在0.5以下。其精准率很高, 而召回率很低, 这是因为精准率和召回率是两个相互矛盾的指标。由公式(11)和(12)可得, 精准率的大小与假正例成反比, 召回率的大小与真正例成反比。这样就导致精准率和召回率相互矛盾, 类似于处于拔河比赛的两端。而 F_1 值能够平衡这两个指标, 衡量模型性能的指标, 但是这两个模型的 F_1 值也都处于一个较低的水准。

比较 BiLSTM + Attention 模型和单纯使用 BiLSTM 模型的 F_1 值和精准率, 会发现 F_1 值提高了约 16%, 召回率提高了约 29%。同样比较 BiLSTM 和 BiLSTM + CharCNN, 发现后者的 F_1 值提高了约 19%, 召回率提高了约 33%。比较 BiLSTM + Attention 和 BiLSTM + CharCNN 模型,

发现相对于前者, 后者召回率提高了4%, F_1 值提高了约3%, 精准率提高了约2%。最后, 本文用到的模型与前几个模型相比, 在精准率几乎不变的情况下, 召回率提高了约3%, F_1 值也提高了约1%。对于药物不良反应的监测问题, 该模型在一定程度上提高有不良反应用户被检测出的概率。

3 结语

本文模型主要是对社交媒体上的药物不良反应进行分类。使用了双向长短时记忆网络 LSTM 和 CharCNN 模型进行特征提取, 同时结合注意力机制对特征向量进行优化。使用双向长短时记忆网络, 能够完美保留 LSTM 处理文本的优势, 处理长距离的依赖特征, 并能保证更全面地考虑上下文信息。注意力机制能够通过调整双向长短时记忆网络中各节点的权重, 使得模型能够识别文本中更加关键的部分, 这样能最大程度地削弱冗余部分对文本的影响, 进而优化特征向量。CharCNN 从字符信息出发, 提取文本在字符粒度上的特征向量。与传统的人为设计的文本特征和基于单一的神经网络提取的特征向量相比, 能够从不同层次更全面地描述推文中的文本特征, 且能识别文本中的关键信息, 进而在药物不良反应的分类任务中取得更好的效果。下一步的研究重点就是如何确定推文中的不良反应所对应的相关药物。

参考文献:

- [1] SARKER A, GINN R, NIKFARJAM A, et al. Utilizing social media data for pharmacovigilance: a review[J]. Journal of Biomedical Informatics, 2015, 54: 202-212.
- [2] SARKER A, GONZALEZ G. Portable automatic text classification for adverse drug reaction detection via multi-corpus training[J]. Journal of Biomedical Informatics, 2015, 53: 196-207.
- [3] KORKONTZELOS I, NIKFARJAM A, SHARDLOW M, et al. Analysis of the effect of sentiment analysis on extracting adverse drug reactions from tweets and forum posts[J]. Journal of Biomedical Informatics, 2016, 62: 148-158.
- [4] COCOS A, FIKS A G, MASINO A J. Deep learning for pharmacovigilance: recurrent neural network ar-

- chitectures for labeling adverse drug reactions in Twitter posts[J]. Journal of the American Medical Informatics Association, 2017, 24(4):813-821.
- [5] XIA L, WANG A, FAN W, et al. A deep learning based transfer learning framework for healthcare text analytics[C]//PACIS. 2019:75.
- [6] REZAEI Z, EBRAHIMPOUR-KOMLEH H, ESLAMI B, et al. Adverse drug reaction detection in social media by deep learning methods[J]. Cell Journal(Yakhteh), 2020, 22(3):319.
- [7] 胡江鹭. 基于社交媒体的药物副作用识别[D]. 黑龙江:哈尔滨工业大学, 2017. DOI: 10.7666/d. D01589771.
- [8] 魏巍. 药物不良反应知识发现与利用模型研究[D]. 武汉:武汉大学, 2017.
- [9] 郭凯. 基于深度学习和语义分析的药物不良反应发现[D]. 辽宁:大连理工大学, 2017.
- [10] 王姝元. 基于社交媒体的药物不良反应发现研究与应用[D]. 辽宁:大连理工大学, 2018.
- [11] SCHMIDHUBER J, HOCHREITER S. Long short-term memory [J]. Neural Comput, 1997, 9(8):1735-1780.
- [12] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [EB/OL]. arXiv Preprint arXiv:1706.03762, 2017.
- [13] 朱晓旭, 林鸿飞, 曾泽渊. 基于社交媒体的药物不良反应检测[J]. 山西大学学报(自然科学), 2020, 43(1):14-21. DOI:10.13451/j.sxu.ns.2019121.
- [14] 赵明珍, 林鸿飞, 徐博, 等. 面向社交网络的潜在药物不良反应发现[J]. 中文信息学报, 2017, 31(5):194-202. DOI: 10.3969/j. issn. 1003-0077.2017.05.027.
- [15] ZHANG X, ZHAO J, LECUN Y. Character-level convolutional networks for text classification [EB/OL]. arXiv Preprint arXiv:1509.01626, 2015.
- [16] LI Z, YANG Z, LUO L, et al. Exploiting adversarial transfer learning for adverse drug reaction detection from texts [J]. Journal of Biomedical Informatics, 2020, 106:103431.

作者简介:

通信作者:蒽娟霞(1992—),女,甘肃白银人,硕士,助教,研究方向为机器学习、智能信息处理, E-mail: jalynnxi@qq.com

徐鹏(2001—),男,广东深圳人,本科生,研究方向大数据管理与应用

叶思维(2001—),男,广东东莞人,本科生,研究方向大数据管理与应用

收稿日期:2022-04-19 修稿日期:2022-05-21

Adverse Drug Reaction Detecting Method Based on Attention-BiLSTM-CharCNN

Xi Juanxia, Xu Peng, Ye Siwei

(School of Information Management and Engineering, Neusoft Institute Gangdong, Foshan 528225)

Abstract: Social media such as Twitter has provided a new source of information for adverse drug reaction monitoring. The article combines deep learning technology to detect whether tweets about adverse drug reactions from massive social media user comments. In order to improve the accuracy and enrich the informative features of tweets, an end-to-end attention mechanism based BiLSTM-CharCNN neural network model is proposed. Firstly, use the CharCNN model to extract character-level features: digitize the text, then perform convolution pooling and full connection operations to obtain character-level feature vectors; secondly, use the Attention-BiLSTM model to extract feature vectors: by adjusting the weight parameters, remove redundant information, extract key parts of information, and optimize the text; finally, the feature information extracted by the two methods is connected, and then the classification operation is performed at the output layer containing a full connection and a softmax. Experimental results demonstrate that the proposed method can significantly improve the classification accuracy and outperform existing benchmark methods.

Keywords: adverse drug reactions; social media; deep learning; Attention; BiLSTM-CharCNN

文章编号: 1007-1423(2022)16-0028-05

DOI: 10.3969/j.issn.1007-1423.2022.16.005

云端马拉松图像的号牌识别研究

黎蕴玉, 丁小波, 蔡茂贞, 钟地秀, 彭 琨

(中移互联网有限公司云产品事业部, 广州 510000)

摘要: 对于马拉松赛事举办方而言, 如何将赛事期间为运动员们拍摄的海量照片进行云端的精准分类、推送和管理是一项具有挑战性的服务工作。为了实现云端海量马拉松图片的智能分类, 本文提出了一种基于目标检测的两阶段号牌识别算法框架, 利用号牌编号的识别实现高效的图像分类功能。本文首先利用 RetinaNet 算法实现号牌检测, 然后通过 DenseNet-CTC 进行端到端的号牌识别, 在保证精度的前提下有效减少冗余信息, 提升推理速度。在 RBNR、FUNNR 和 TEST 数据集上进行了对比实验, 结果证明了本方法的有效性和良好的泛化性。

关键词: 马拉松图片; 号牌检测; 号牌识别; 图像分类

0 引言

由于低门槛和轻装备的特点, 马拉松已成为各城市和企业乐于举办的运动赛事。全球的马拉松赛事数量持续增长, 统计显示^[1]2019 年全国共举办 1828 场次规模赛事, 覆盖了全国 31 个省区市, 参加人次达 712 万。对于马拉松赛事举办方而言, 将赛事期间为运动员拍摄的海量照片进行云端的精准分类、推送和管理是一项具有挑战性的赛事服务工作。相比于效率低、耗时长和成本高的传统人工分类方式, 基于运动员号牌智能识别成了实现云端海量马拉松赛事图片精确分类的重要方法。马拉松运动员号牌字符主要由若干位大小写字母和数字任意组合而成, 通过贴在运动员衣服上进行区分, 不同场次的号牌字符字体、颜色和背景各不相同。由于马拉松号牌图片具有格式多样、光照多变和扭曲变形等特点, 设计一种鲁棒性强、高效的基于号牌识别的云端图片分类方案具有一定的难度。

近年来, 研究者进行了大量运动员号牌识别的工作, 并广泛应用在运动员图片分类的任务上。文献[2]利用人脸检测预估号牌位置, 再使用 SWT 精细定位分割号码, 最后进行 TesseractOCR 文字识别。文献[3]则采用 SVM 人体检

测模型、HOG 文字检测和 TesseractOCR 识别模型的多模型融合方法。文献[4]首先利用可变形部件模型进行人体检测, 再使用图像处理技术进行号牌检测和分割, 最后输入三层 BP 神经网络得到单字符的识别结果。文献[5]利用迁移学习策略实现较高精度的号牌识别。文献[6]将 YOLOv3 人体检测模型、CTPN 文字检测模型和 CRNN 文字识别模型进行串联, 经树过滤后实现号码牌识别。文献[7]则基于 YOLOv4 人体检测结果进行字符级 CRAFT 文字检测和基于注意力机制的号码识别。

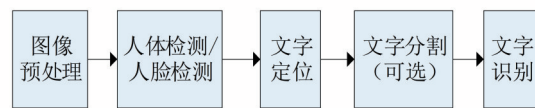


图 1 常见的号牌识别流程

基于马拉松号牌的特殊性, 现有号牌识别方法大体采用图 1 所示的算法流程。通用的 CTPN^[8]、EAST^[9]等文字检测技术主要针对自然场景设计, 会识别图片中的所有文字内容, 因此通过人体/人脸检测和文字检测相结合的方法先检测出人体, 过滤冗余信息, 再基于人体检测结果进行号牌检测, 该流程虽能有效识别出

号牌,但也造成模型数量多、处理复杂度高和运行时间长的问题。

综上所述,本文主要提出一种基于目标检测的两阶段号牌识别云端图片分类系统,去除图1中的人体/人脸检测和文字分割两个模块,将号牌直接当成目标进行检测,有效减少冗余信息,两阶段的号牌检测识别模型在保证识别精度的情况下有效提升推理速度,实现马拉松图像的精准分类和管理。

1 提出方案

本文提出了一种基于两阶段号牌识别马拉松图片分类方法,首先将经过预处理的图片通过 RetinaNet 模型进行号牌检测,将裁剪的号牌检测区域输入号牌识别模型进行文字识别。本方法的具体流程如图2所示,包含举办方终端和云端服务器两部分,具体步骤为:①举办方终端批量上传马拉松赛事图片和赛事的所有号牌值至云端服务器;②在云端服务器对图片进行去重、过滤模糊无效图、获取图片的属性信息和角度矫正等预处理操作;③将预处理的图片输入基于 RetinaNet 的号牌检测模型定位号牌区域,裁剪出号牌区域作为号牌识别模型的输入,获得识别的号牌值;④将识别的号牌结果进行BK树过滤、号牌聚类等后处理,实现马拉松图片智能分类。

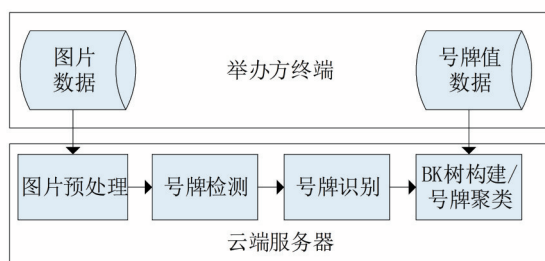


图2 基于号牌识别的马拉松云端图像分类图

1.1 RetinaNet 号牌检测模型

现有的号牌检测模型主要使用人体检测和文字检测的多模型融合或直接使用文字检测模型,前者虽然能提升号牌检测准确率,但处理速度慢、计算复杂度高;而后者则会形成大量

冗余文字信息,影响检测精度。基于以上分析,本文提出一种基于 RetinaNet^[10]的号牌检测网络(如图3所示),主干网络为 ResNet50,利用特征金字塔进行多尺度特征提取,每个特征层对接2个分支:框分支和类别分支,分别预测号牌框4点矩形坐标、是否为号牌;每个特征层对应3个锚点以提取不同尺度目标,类别分支和方向分支均使用 Focal 损失函数缓解类别不平衡问题。该模型仅针对号牌标注数据进行训练,能有效检测出号牌区域,避免冗余文字信息的产生。

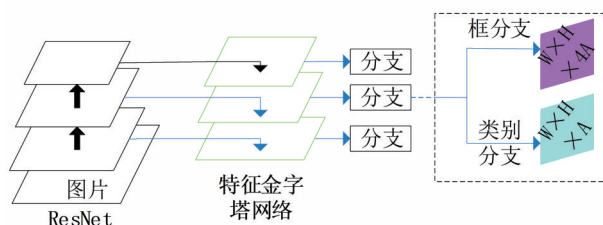


图3 RetinaNet 号牌检测网络

1.2 DenceNet-CTC 号牌识别模型

马拉松的号牌文字序列主要由随机的字母和数字组成,并无明显的上下文结构信息,而常用的循环神经网络和注意力机制的文字识别网络主要针对上下文序列信息进行识别。此外,卷积神经网络结构每一步计算都依赖于前一步的计算和输出结果,存在模型参数多和推理耗时的问题。

为加快推理速度和提高号牌识别准确率,本文的号牌识别模型采用 DenceNet(dense convolutional network)作为主干网络,CTC(connectionist temporal classification)作为损失函数实现非定长序列的端到端号牌识别^[11],具体如图4所示。DenceNet 网络主要由卷积神经网络、Dense 模块和 Transition 层组成,通过相互连接所有层的密集机制减轻梯度消失和加强特征传递及重用,其中 Dense 模块是指第 l 层连接前面所有层的特征作为输入,获得特征 $x_l = H_l([x_0, x_1, \dots, x_{l-1}])$; Transition 层则是批标准化、卷积和池化层的集合。最后通过 CTC 损失函数,利用前向概率求解,直接预测出号牌序列值。

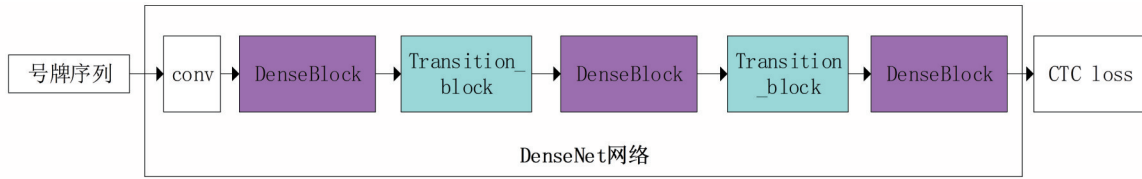


图4 DenseNet-CTC号牌识别模型

2 实验

2.1 数据集

(1) Racing Bib Number Recognition(RBNR)^[2]。该数据集采集了自然场景下三个不同场次的含马拉松号牌图片，217张图片中含290个号牌，号牌均由3~6位纯数字构成。

(2) 趣味运动会数据集(FUNNR)。该部分为自行采集的多个场次趣味马拉松运动会的号牌图片集，号牌由大写字母、部分特殊字符和数字组成，共有4653张图片和9118个号牌，部分为遮挡号牌。

(3) 测试集(TEST)。该测试集为随机采集的马拉松号牌图片126张，共有140个号牌，用于整体识别性能测试，其中清晰号牌图片含106张，有遮挡的图片20张。

2.2 消融实验

实验基于Keras实现并使用英伟达GTX

1080TI显卡进行训练，对RBNR和FUNNR两个数据集按9:1划分训练集和测试集，用于模型的训练和测试，并采用平移、随机裁剪、加噪声和生成样本等多种数据增强方式提升样本的丰富性。

2.2.1 号牌检测实验

该部分进行了号牌检测的RetinaNet和EAST两个算法对比实验，如表1所示，在RBNR和FUNNR两个测试集上，RetinaNet的检测准确率均表现良好，表明对比EAST算法，本方法能有效减少误检测情况。

2.2.2 号牌识别实验

为了证明本文提出的号牌识别算法的有效性，我们将DenseNet-CTC和TesseractOCR进行实验对比。如表2所示，本文提出的号牌识别算法在两个数据集上的编辑距离和准确率指标均高于TesseractOCR。其中准确率是指字符均正确识别的号牌数占总号牌数的比例。

表1 号牌检测算法对比

指标	数据集					
	RBNR			FUNNR		
	准确率	召回率	F1值	准确率	召回率	F1值
RetinaNet	0.76	1.0	0.86	0.867	0.74	0.80
EAST	0.68	0.84	0.75	0.73	0.90	0.80

表2 号牌识别算法对比

指标	数据集			
	RBNR		FUNNR	
	编辑距离	准确率 / %	编辑距离	准确率 / %
DenseNet-CTC	0.991	96.55	0.968	84.4
TesseractOCR	0.144	10.34	0.281	10.75

2.2.3 整体测试

针对 FUNNR 数据集训练的模型进行测试集 (TEST) 的整体识别性能测试, 具体的识别结果如图 5 所示, 对于单号牌图片图 5(a), 本方法能正确检测和识别出号牌内容, 而对于有遮挡的多号牌图片图 5(b), 本文方法也表现出良好的识别效果, 证明本方法的有效性和良好的泛化性。



(a)



(b)

图 5 识别结果图

在本次的测试中共检测出号牌 122 个, 其中正确识别 120 个, 错误识别 2 个, 漏识别 20 个, 准确率为 98.26%, 召回率为 85.71%。其中 CPU 上平均单张图片的处理时间为 0.748 秒, 对比文献[3]中 2.19 秒的单幅图片处理时间, 本方法的处理速度提升了 3 倍。

3 结语

针对日益增长的云端马拉松图片, 本文设计了一种基于目标检测的两阶段号牌识别策略进行马拉松图片分类, 本方法分别利用 RetinaNet 和 DenseNet-CTC 进行号牌检测和识别, 在保证识别准确率的同时有效提升运行速度, 对局部有遮挡的多号牌图片仍能有效识别。

参考文献:

[1] 赵逸凡. 我国马拉松赛事开展的现状与对策研究

[J]. 科教导刊(电子版), 2020(19):1.

- [2] BEN-AMI I, BASHA T, AVIDAN S. Racing Bib Number Recognition[J]. Ami, 2012.
- [3] SHIVAKUMARA P, RAGHAVENDRA R, QIN L, et al. A new multi-modal approach to bib number/text detection and recognition in marathon images [J]. Pattern Recognition, 2017, 67: 479-491.
- [4] 赵丽科, 郑顺义, 马浩, 等. 田径运动员号码牌图像的号码识别[J]. 华东师范大学学报(自然科学版), 2017(3):15.
- [5] IVARSSON E, MUELLER R M. Racing bib number recognition using deep learning[C]//Proceedings of the 25th Americas Conference on Information Systems (AMCIS), 2019.
- [6] WANG X, YANG J. Marathon athletes number recognition model with compound deep neural network [J]. Signal, Image and Video Processing, 2020, 14 (7): 1379-1386.
- [7] NADY A, HEMAYED E E. Player Identification in Different Sports[C]//16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, 2021.
- [8] ZHI T, HUANG W, TONG H, et al. Detecting text in natural image with connectionist text proposal network [C]//European Conference on Computer Vision. Springer, Cham, 2016.
- [9] ZHOU X, YAO C, WEN H, et al. EAST: an efficient and accurate scene text detector[J]. IEEE, 2017.
- [10] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]//IEEE Transactions on Pattern Analysis & Machine Intelligence. IEEE, 2017: 2999-3007.
- [11] ZHAN H, LYU S, LU Y, et al. DenseNet-CTC: an end-to-end RNN-free architecture for context-free string recognition [J]. Computer Vision and Image Understanding, 2021, 204(11): 103168.

作者简介:

通信作者: 黎蕴玉(1990—), 女, 广东广州人, 硕士研究生, 研究方向为人工智能应用实践、计算机视觉, E-mail: 19802021462@139.com

丁小波(1986—), 男, 广东广州人, 本科, 研究方向为系统架构设计、移动端技术

蔡茂贞(1986—), 男, 广东广州人, 本科, 研究方向为分布式系统架构设计、人工智能应用实践

钟地秀(1993—), 女, 广东广州人, 硕士, 研究方向为人工智能应用实践

彭琨(1991—), 男, 广东广州人, 硕士, 研究方向为人工智能应用实践

收稿日期: 2022-03-21

修稿日期: 2022-04-12

(下转第 37 页)

文章编号: 1007-1423(2022)16-0032-06

DOI: 10.3969/j.issn.1007-1423.2022.16.006

基于电气标准模型的移动端配网图数一体化建模工具设计与应用

张旭东, 陈星明

(南京南瑞信息通信科技有限公司, 南京 210000)

摘要: 随着国家电网深入数字化改革, 移动化作业场景越来越普及, 传统的 PC 客户端软件已经不能满足现有的作业需求, 不仅工作效率降低, 而且数据一致性得不到保证。针对这类问题, 本文基于配网 PMS2.5 系统的 PC 客户端设计了一个移动端图形数据一体化建模工具, 该工具提供在移动端图形拓扑建模并实时预览成图效果、设备台账信息现场采集、设备地理坐标获取等一系列功能, 简化 PC 客户端建模流程, 保证了模型及数据的一致性, 提高建模及工作效率, 并在实际应用中取得不错效果。

关键词: 电气标准模型; 图数一体化; 移动建模

基金项目: 南京南瑞信息通信科技有限公司科技项目(5246DR220054)

0 引言

随着科学技术不断发展、智能设备不断普及, 手机、平板等电子产品已经成为人们生活和工作的必需品。国家电网设备部、运检部、互联网部当前重点工作任务是加快治理配网数据, 因此, 配网^[1]中低压^[2]设备数据建设和治理成为了重点工作内容之一。

国家电网数字化班组以业务在线化、作业移动化、信息透明化、支撑智能化为方向聚焦基层班组成员常态业务、核心业务。坚持中台战略, 遵循“促共享, 强应用”的工作路线, 贯通专业信息系统, 深化移动终端应用, 强化智能分析支撑, 挖掘数据资产价值, 开展基于移动应用的数字化班组建设, 支撑公司设备管理、营销服务数字化转型, 全面提升基层班组及员工数字化水平, 实现班组减负和提质增效, 切实打通班组应用的“最后一公里”。数字化班组坚持以班组需求为导向, 从效率、质量、安全等维度, 持续深化移动终端应用功能整合, 推进班组移动业务 APP 适应性改造, 进一步提升实用性、便捷性和可操作性。

数字化班组通过走访基层及移动化应用场景相关调研, 决定开展图模维护工具的移动化作业建设。同时, 在国家电网 PMS2.5 系统建设

过程中, 配网移动端图数一体化是实现移动化作业的关键环节。现有的配网 PC 客户端建模工具已经不能满足当前的移动作业需求, 存在诸多问题。比如: ①对于配网设备台账数据的维护只能依赖原始导入数据, 且仅限于 PC 客户端人工操作, 初始数据来源也依赖人工录入, 数据准确性存在较大的误差; ②设备的图形数据需要专职人员根据纸质图纸在 PC 端进行手动绘制, 无法与现场勘测的实际设备图形做对比, 若专职人员巡检时发现现场图形错误, 并不能及时维护; ③设备地理坐标信息及图片信息无法精准获取, 同样只能专职人员手动维护, 设备坐标的准确度无法保证, 等等。

1 配网图数一体化建模现状

配网中低压设备建模包含以下设备类型: 线路、杆塔、站房(配电室、箱式变电站、环网柜、开关站、电缆分支箱)、柱上设备(柱上变压器、柱上断路器、柱上隔离开关、柱上负荷开关、柱上跌落式熔断器)、电缆终端、电缆段等。每种设备类型拓扑连接关系存在先后顺序约束规则, 例如: 线路下只能跟随杆塔、站房、电缆等设备, 而不能拥有柱上设备; 电缆终端后只能连接故障指示器、杆塔和站房等。除了拓扑关系存在约束规则, 设备与设备之间的连

接也有特定的约束规则,例如:杆塔与杆塔间用档距段相连;杆塔与柱上设备之间用连接线相连;杆塔与站房之间又根据不同的站房类型情况变得复杂,如:杆塔与环网柜之间用电缆相连,而杆塔与开关站之间则用超连接线相连等。这些约束规则增加了系统的复杂度和实现难度,所以,在资源有限的移动端建模难度不言而喻。

对于配网中低压设备维护主要包括图形拓扑关系^[3-4]维护和资源资产台账维护。拓扑关系具体是指配网设备之间的连接、从属、对应等关系信息,而拓扑关系的操作包含设备退役、设备新增、设备更新、设备切改等。设备退役是将选择的设备进行删除,然后重新建立剩余设备间新的图形拓扑关系;设备新增操作同理;而对于设备更新则需要进行设备退役和设备新增两步操作来完成;设备切改则是选择某个特定设备,将此设备之后的所有设备添加到另一条线路上并删除原来线路上设备。设备台账信息记录着设备的所有信息,包括生产日期、生产厂家、投用日期、材质等基本资源信息,也包括设备所属线路、维护班组、电压等级等核心资源信息。

现有的配网PC客户端设备建模工具主要是以图元拖拽、划线连接等操作方式为主。PC客户端具体操作流程如下:①专职人员拿到工程师设计图纸的纸质稿;②按照图纸要求根据设备类型逐一拖拽图元至画布并连接图元;③待所有图元绘制完毕后点击一键成图按钮进行图形拓扑关系的校验,检查是否存在连接规则不匹配或者存在孤立图元等问题;④在图形拓扑关系校验通过后,专职人员还要逐一手工维护每种设备类型的资源资产台账信息;⑤所有设备类型台账维护完成后,再进行台账数据校验,确保数据的基本属性都已得到维护;⑥所有数据校验通过后才可提交图形和台账数据,最终完成一张设计图纸的建模过程。上述的数据校验仅限于数据类型的校验,而无法提供对数据准确性的校验,若校验不通过,则不能进行数据保存操作,此外PC端建模过程与现场施工完全分开,也无法进行现场图模数据核对。图1是PC客户端建模操作界面,右侧区域是电气标准模型图元库,左侧区域是画板,图2是拓扑关

系及台账信息校验界面。

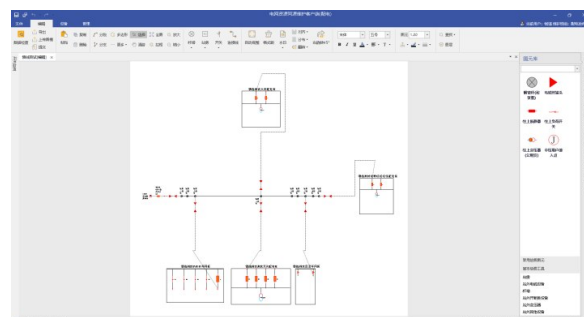


图1 PC客户端操作界面

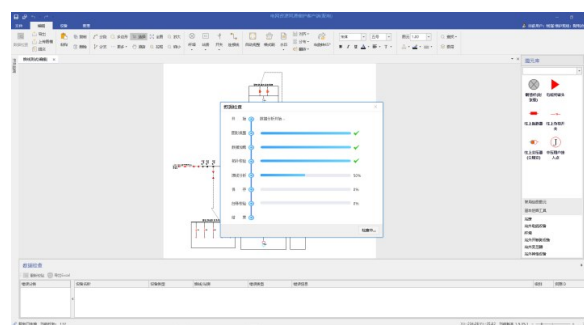


图2 数据校验

这种复杂逻辑业务系统,由于PC客户端有较大的屏幕展示和较强的处理器性能,从设计上和用户使用性上来说具有一定的优势,但数据的准确性都是建立在专职人员手工绘制的基础上,所以仍存在不少问题,例如:①PC客户端建模过程与现场施工完全分开,无法进行现场图模数据核对,数据无法实时校验;②拖拽图元式建模对作业环境有一定要求,只能在有专有网络的办公室,而且需要逐一拖拽,费时费力,效率较低;③人工绘图对数据准确性无法保证,误差较大;④对于现场设备的更换,很难进行实时数据更新,数据闭环管理较难;⑤设备的空间信息无法得到实时的采集,仅限于专职人员手工维护;⑥无法对设备照片进行实时采集上报。

2 配网移动端图数一体化设计

2.1 配网移动端图数一体化设计方案

对于上述诸多问题,本文设计一款移动端配网图数一体化建模工具。移动端优点显而易见,不受地理环境、工作条件等影响,方便用户随身携带。专职人员可以随时开展户外现场

作业，在移动端就可以全部完成整个建模流程及设备台账维护。该工具支持对配网设备的拓扑关系进行现场退役、新增、更新和切改及设备台账信息现场维护，最终通过国网专有网络实时将数据保存至移动作业数据库。

根据前一节分析，设备图形拓扑具有先后顺序关系和连接关系，所以在移动端建模选择采用层次树模式^[5]结合移动端数据库缓存^[6]实现。层次树规定了设备之间的先后顺序关系，数据库缓存存储了设备间的连接关系，二者结合便可以限定一张完整的图形拓扑结构图，层次树结构如图3所示。移动端数据库缓存选择使用Android^[7]系统自带的嵌入式SQLite数据库^[8]，采用SQLite数据库缓存方式还有一个优点，对于一些网络信号不好的偏远地区，设备数据的采集可以采用缓存模式，先将采集到的数据缓存到SQLite数据库中，待网络条件允许后再进行数据提交。数据库设计部分核心属性如图4所示。其中equipmentContainer字段表示当前设备的连接点，connectivityNodes字段表示当前设备的前后连接关系。

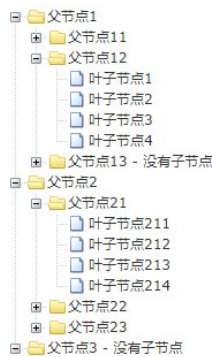


图3 层次树结构

Tables	Data
allmodelsymbolinfo_t	
android_metadata	
collectionnossend_t	
collectionssend_t	
commoncodeidist_t	
graphicsdevice_t	
newconfigs_t	
numproperte_t	
room_master_table	
savegovermodel_t	
sqlite_sequence	
treenode_t	

rpe	equipmentContainer	connectivityNodes	figureType	objType	pov
Root.F4	[Root.F4_com_1]	Insert	32000000	0	
Root.F73	[Root.F3_com_11;Root.F74_com_0]	Polyline	14000000	0	
Root.F74	[Root.F74_com_0]	Insert	20200000	0	
Root.F75		Label	-1	0	
Root.F2	[Root.F2_com_11;Root.F2_conn_2]	Insert	30500000	0	
Root.F71		Label	-1	0	
Root.F83	[Root.F77_conn_0;Root.F79.F84_com_0]	Polyline	14000000	0	
Root.F79		SubDrawing	20400000	0	
Root.F79	[Root.F79.F130_com_0]	Insert	30000000	0	
Root.F79		Label	-1	0	

图4 SQLite数据库部分核心属性

对于层次树的层次关系，需要在服务数据库中编写特定的脚本来规定主干节点上具体有哪些设备，然后根据主干节点逐层递归查询下属子节点。通过用户选择的不同节点类型来限定不同的约束规则，页面底部弹出不同的交互界面。例如选择线路节点时，页面底部弹框新增设备按钮只会存在杆塔、站房、电缆等设备选项，而两种设备间的连接线会根据选择的两种设备类型进行自动补全，并将两设备相连，对于相同设备类型的设备支持批量新增，并支持移动端实时预览新的拓扑关系。随着国家电网实物资产标签化的深入进行，现在对于设备台账的维护可以通过移动端扫描设备唯一标识二维码^[9]来实现数据的实时对比及采集，更方便了设备资源资产信息维护，对于相同设备类型的设备也支持批量台账维护。对于设备坐标位置信息，使用高德地图定位服务^[10]实时获取设备的坐标信息并更新入数据库。对于设备图像的采集，通过移动设备相机拍照功能进行现场拍照，并实时进行数据传输。待采集的馈线或台区采录完成后，最终可以对所采馈线或台区进行采集数据分析与统计。

综上所述，整个配网图数一体化工具涵盖图形拓扑关系维护及同步、数据采集治理、数据质量检查和综合查询统计等功能，具体结构如图5所示。

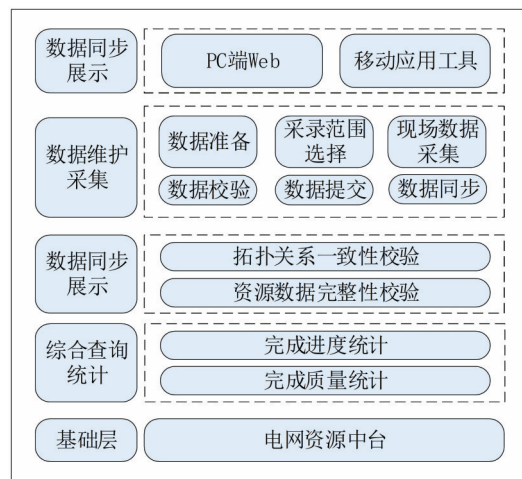


图5 一体化工具功能结构

2.2 配网移动端图数一体化建模流程

配网移动端图数一体化建模工具是通过智

能移动终端完成对现场设备数据采集、图形拓扑维护等操作。工具利用 Android 内嵌的小型数据库 SQLite 结合层次树来实现图形的拓扑连接关系,同时使用高德地图完成设备坐标采集。该工具首先需要在移动端创建维护变更申请单,并关联所需维护的馈线与台区,对于完全新增的馈线或台区,需要从0到1创建整条线路和设备,而对于存量数据,需要先拉取图形拓扑关系,并存入移动端本地 SQLite 数据库缓存中,再通过移动端页面操作来更新维护设备与设备之间的连接关系及台账信息,最终将现场维护好的数据发送到电网资源中台,完成电网设备图数一体化建模任务,其具体实施流程如图6所示。

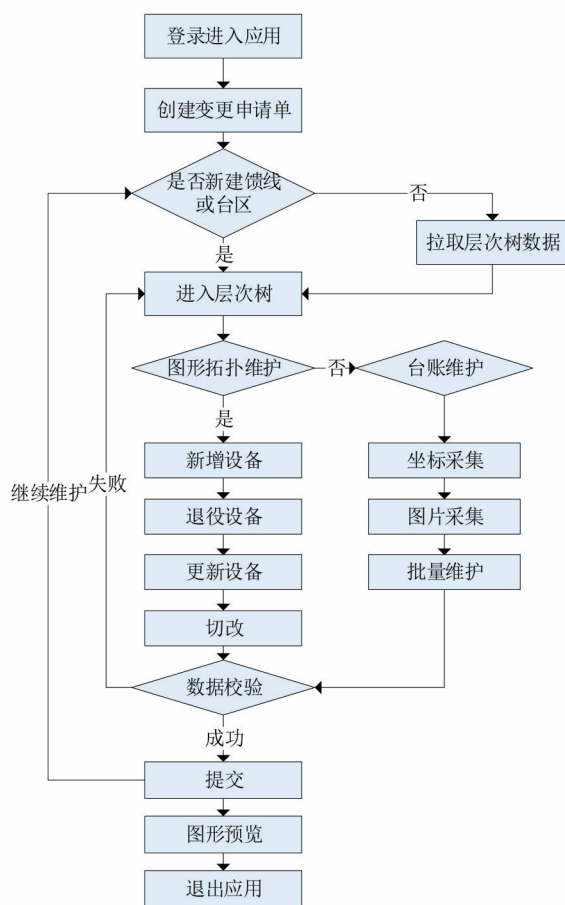


图6 图数一体化流程图

2.3 配网移动端图数一体化建模工具效果

配网移动端图数一体化建模工具设计及开

发基于 Android 操作系统,并针对 Android6 以上系统做了较好的兼容适配,具有较好的用户体验和交互效果。

核心操作步骤及页面如下:首先,根据需要创建任务变更申请单(图7),并关联将要采集的馈线和台区名称(图8),确定采集对象后,便可进入采集范围列表页面(图9),在该页面中选择相应的馈线或台区会拉取层次树拓扑数据,进入对应的馈线和台区层次树操作页面(图10),在该页面可以选择层次树上的设备节点进行新增、退役及台账维护操作,点击不同的设备类型底部弹框界面不完全相同(如图11),在该层次树页面右上角存在四个按钮,从上到下的作用分别是:①实时查看图形拓扑图形(图12);②查看当前节点详细名称,并以 Toast 方式弹出;③对层次树进行折叠和展开操作;④点击提交按钮即可进行数据校验及图形拓扑的重新生成,并可以实时查看图形拓扑,再展示图形拓扑效果页面(图12),可以放大、缩小页面并选择相应的设备图元进入台账维护页面(图13),在该页面允许对设备的资源资产信息进行维护,提供了实物 ID 扫码、地理坐标采集(图14)、图片采集及批量维护等功能。

3 结语

本文首先通过对配网 PC 客户端图模维护工具进行分析,发现其在实际使用场景及数据准确性方面存在诸多缺点,工作效率较低、数据准确度差且无法形成闭环管理,专职人员工作不能保证所有数据全程无错误录入与执行,较大的工作量给其造成一定工作压力。基于此,本文设计一款配网移动端图数一体化建模工具,该工具基于层次树并结合数据库缓存方式,解决了图形复杂的拓扑关系问题。配网移动端图数一体化建模工具的落地改进了配网传统 PC 客户端图形建模方式和数据维护模式,提高了数据的正确性和数据源的可靠性,改变了传统的人工手录模式。在时间有限的情况下,专职人员大大增加了馈线及台区的数据采集量,节省了成本和时间,为国网 PMS2.5 系统的移动化建设奠定了坚实的基础,具有一定的里程碑意义。



图 7 变更申请单



图 8 选择馈线/台区



图 9 采集范围列表



图 10 层次树



图 11 层次树操作

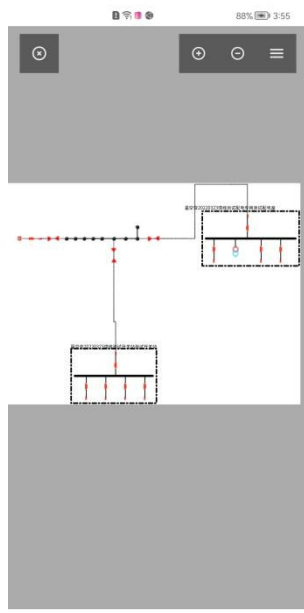


图 12 图形预览



图 13 台账维护



图 14 地理信息定位

参考文献:

[1] 郝睿,孟静静.电力系统中的配网自动化技术研究分析[J].中国设备工程,2021(19):127-129.

[2] 黎海霞.中低压配网详细规划研究[J].机电信息,2016(24):25,27.

[3] 吴新新.面向电网拓扑管理的图数据库关键技术研究[D].南京:南京邮电大学,2020.

[4] 罗群,刘春雨,顾强,等.基于最优匹配回路功率的

配电网拓扑辨识方法[J].电测与仪表,2019,56(19):1-6.

[5] 张晓敏,王茜.基于概念层次树的个性化推荐算法[J].计算机工程,2007(24):57-59,62.

[6] 翁唱玲,杨清.移动数据库缓存模型研究[J].计算机应用,2013,33(11):3267-3270.

[7] 叶永川.基于Android平台软件开发技术分析[J].新型工业化,2021,11(12):112-113,117.

- [8] 黄斯锋. 基于Android的SQLite数据库诊断工具的设计和实现[D]. 北京:北京交通大学,2017.
- [9] 何明,范蕾,赵欢,等. 基于二维码技术的配网设备台帐移动采录工具研究与应用[J]. 计算机产品与流通,2019(01):138.
- [10] 李庭威,王峰,汪进,等. 高精度差分在高德地图上的多点实时定位[J]. 电脑知识与技术,2017,13(23):200-201.

作者简介:

通信作者:张旭东(1991—),男,江苏宿迁人,硕士,工程师,研究方向为计算机软件与理论,E-mail:zhangxudong1@sgepri.sgcc.com.cn

陈星明(1989—),男,江苏无锡人,硕士,工程师,研究方向为移动通信

收稿日期:2022-02-21

修稿日期:2022-04-02

Design and Application of Integrated Modeling Tool for Distribution Network Diagram Number Based on Electrical Standard Model

Zhang Xudong, Chen Xingming

(Nanjing NARI Information & Communication Technology Co. Ltd., Nanjing 210000)

Abstract: With the in-depth digital reform of State Grid, mobile operation scenarios are becoming more and more popular. Traditional PC client software can no longer meet the existing operation requirements, which not only reduces the work efficiency, but also can not guarantee the data consistency. PMS2.5 against this kind of problem, this paper based on the distribution network system of PC client designed a mobile end graphics data integration modeling tool, this tool provides graphic topology modeling and real-time preview on the mobile end mapping effect, equipment parameter field information acquisition, geographic coordinates for a series of functions such as equipment, simplifying the PC client modeling process, It ensures the consistency of model and data, improves the efficiency of modeling and work, and achieves good results in practical application.

Keywords: electrical standard model; graph integration; mobile modeling

(上接第31页)

Racing Bib Number Recognition of Cloud Marathon Images

Li Yunyu, Ding Xiaobo, Cai Maozhen, Zhong Dixiu, Peng Kun

(Cloud Products Division, China Mobile Internet Co., Guangzhou 510000)

Abstract: For marathon organizers, how to accurately classify, recommend and manage the massive photos taken for athletes in the event is a challenging service. In order to realize the automatic classification of marathon natural pictures in the cloud, this paper presents a novel two-stage racing bib number recognition framework, which utilizes the number recognition results to efficiently classify the marathon images. This paper firstly utilizes RetinaNet to localize the number position, and then uses DenseNet-CTC to recognize the bib number, which can effectively reduces redundant information and improve the inference speed on the premise of ensuring accuracy. Comparative experiments on RBNR, FUNNR and TEST datasets demonstrate the effectiveness and good generalization of our proposed method.

Keywords: marathon pictures; number detection; number recognition; image classification

文章编号: 1007-1423(2022)16-0038-05

DOI: 10.3969/j.issn.1007-1423.2022.16.007

基于集成学习的有害垃圾自动识别方法研究

孟德尧¹, 吴荣海², 杨邓奇¹

(1.大理大学数学与计算机学院, 大理 671003; 2.大理大学工程实训中心, 大理 671003)

摘要:深度学习是机器学习提供自动识别图像的工具, 广泛应用于自动识别垃圾图像。受限于数据集, 现有的垃圾分类研究中都是平等对待所有垃圾类别, 没有考虑有害垃圾误分类的危害。本文首先基于网络爬虫和人工拍摄的方式构建了包含有害垃圾的垃圾图像数据集(DTrash), 并基于保守的集成学习策略, 提出了基于集成学习的有害垃圾自动识别方法。集成学习方法在DTrash上的漏判误差、误判误差和总体误差分别为0%, 11.63%和4.57%。

关键词:有害垃圾; 深度神经网络; 垃圾分类; 深度学习; 集成策略

基金项目:国家自然科学基金(31960119)

0 引言

随着全球经济的不断发展, 居民的生活水平不断提高, 生活垃圾的生产量也在逐年增加。由于城市人口多, 环境容量有限, 城市地区的生活垃圾污染问题变得更加严重^[1]。垃圾回收任务是缓解环境和改善整个国家经济的有效途径^[2]。垃圾回收工作的效率和质量在很大程度上取决于垃圾分拣的有效性^[3]。针对人工分拣垃圾存在的工作量大、易出错、分拣效率低等问题, 一些学者提出了智能垃圾识别与分类方法, 通过机器识别垃圾图像实现自动分类, 降低人工成本, 进一步提高资源的再利用率。

近年来, 深度学习在图像识别领域取得了突出的成绩, 被广泛应用于垃圾分类^[4-7]。例如, 文献[8]用DenseNet169模型在自制的垃圾数据集NWNUTRASH(废玻璃、废织物、废纸、废塑料和废金属, 共18911张)上实现垃圾分类, 取得了82%的准确率。文献[9]使用支持向量机(SVM)和ResNet50(SVM)模型在TrashNet标杆数据集(包括玻璃、纸张、纸板、塑料、金属和普通垃圾六个类别)上进行垃圾分类, 准确率分别为63%和87%。文献[10]通过数据增强的方法, 将TrashNet上的图像分别做了水平翻转、垂直翻转和随机25°旋转, 将数据集扩充了四倍,

其中90%作为训练集, 10%作为测试集, 利用遗传算法优化DenseNet121的全连接层, 取得了99.6%的准确率。现有基于深度学习的垃圾分类主要基于TrashNet数据集来实现, 取得了很高的准确率。

然而, 受限于垃圾图像数据集, 现有垃圾分类的研究没有考虑生活垃圾的组成成分问题。生活垃圾往往包括废旧电池、过期药品等有害垃圾。有害垃圾和无害垃圾的处理方式不同, 将有害垃圾误当作无害垃圾处理, 对环境和生命造成严重的威胁。在实现垃圾自动识别和分拣时, 如何尽可能降低有害垃圾的漏判误差(将有害垃圾误识别为无害垃圾的误差)是个亟待解决的问题。

本文采用网络爬虫和手动拍照的方式建了一个包含废旧电池、过期药品等有害垃圾的数据集, 并提出了基于保守集成策略的垃圾自动识别方法, 实现有害垃圾的低漏判自动分拣。本文的主要贡献:

(1) 构建了一个包含有害垃圾和无害垃圾的垃圾图像数据集, 包括3281张垃圾图像, 分为七类: 废旧电池、过期软膏、过期药物、废玻璃、废纸、废塑料和废金属。

(2) 提出了基于保守集成策略的垃圾自动识别方法, 降低了有害垃圾的漏判误差。

1 方法和数据集

1.1 数据集

本文通过网络爬虫和手动拍照构建了生活垃圾(Domestic Trash, DTrash)数据集,包括废旧电池(321张)、过期药物(448张)、过期软膏(387张)、废玻璃(501张)、废纸(522张)、废金属(551张)、废塑料(551张)七个类别,其中

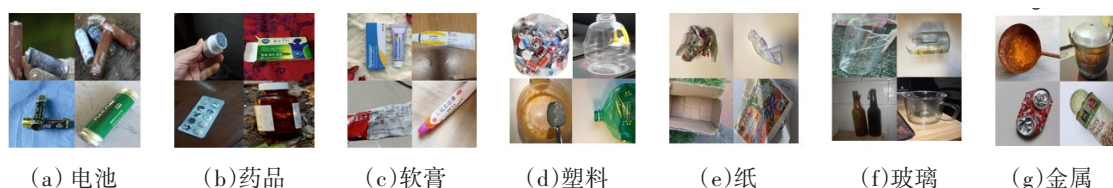


图1 DTrash数据集的图像示例

实际生活中无害垃圾往往比有害垃圾多,因此,本文对数据集的各个类别不做平衡处理。随机地将DTrash数据集按8:1:1的比例划分为训练集、验证集和测试集,见表1。

表1 训练集、验证集和测试集详细信息

数据集	废旧 电池	过期 药品	过期 软膏	废玻 璃	废纸	废金 属	废塑 料
训练集	259	367	305	405	413	437	439
验证集	30	40	41	49	57	56	55
测试集	32	41	41	47	52	58	57

1.2 集成学习方法

1.2.1 DCNN模型的选择

为了提高模型的性能,本文先对VGG-16^[11]、ResNet-50^[12]、ResNext-50^[13]、Vision Transformer (ViT)^[14]和Vision Transformer Hybrid (ViT-Hybrid)在DTrash数据集上进行了测试,结果表明,ViT模型、ViT-Hybrid模型具有更好的性能。因此,本文的集成学习方法采用了这两个模型。ViT和ViT-Hybrid模型的描述见表2。

表2 ViT和ViT-Hybrid模型的信息

模型	块尺寸	层数	简介
ViT	32 × 32	24	无卷积,将图像拆分成小图像块,嵌入到序列中输入到网络
ViT-Hybrid	7 × 7	12	Resnet50的特征提取和Transformer相结合的模型

废旧电池、过期药物、过期软膏是有害垃圾,有害垃圾图像的比例为35.23%。DTrash数据集中所有图像均为RGB格式,大小不一。我们使用Python3.5.3平台的Python Imaging Library中的resize()方法调整了所有图像的大小,以满足DNN模型对输入图像大小的要求。DTrash数据集的图像示例如图1所示。

1.2.2 集成学习方案

本文设计了基于保守策略的垃圾图像自动识别集成学习方法,如图2所示。该方法的目的是在实现垃圾分类时,尽可能减小将有害垃圾预测为无害垃圾的概率。

图2(a)部分是将ViT、ViT-Hybrid在ImageNet上预训练的权重迁移到训练DTrash的模型上。ViT、ViT-Hybrid模型在DTrash训练集上训练后得到集成模型。将测试集输入集成模型得到预测结果。

基于保守策略的集成学习方法使用了迁移学习,将在ImageNet数据集上预训练的ViT和ViT-Hybrid模型的权重参数迁移到DTrash数据集,对全连接层和SoftMax进行微调,再采用保守策略将两个模型集成,得到集成模型。在本文的保守策略中,如果模型ViT和ViT-Hybrid对同一输入图像的预测结果至少有一个是有害垃圾时,集成模型会将该图像预测为有害垃圾(0)。如果模型ViT和ViT-Hybrid对同一输入图像都无法识别时,集成模型会将该图像预测为不确定图像(x)。否则,集成模型会将图像预测为无害垃圾(1)。保守集成策略定义见式(1)。

$$\text{集成模型预测结果} = \begin{cases} 0 & \text{if ViT or ViT - Hybrid} = 0 \\ x & \text{if ViT and ViT - Hybrid} = 1 \\ 1 & \text{else} \end{cases} \quad (1)$$

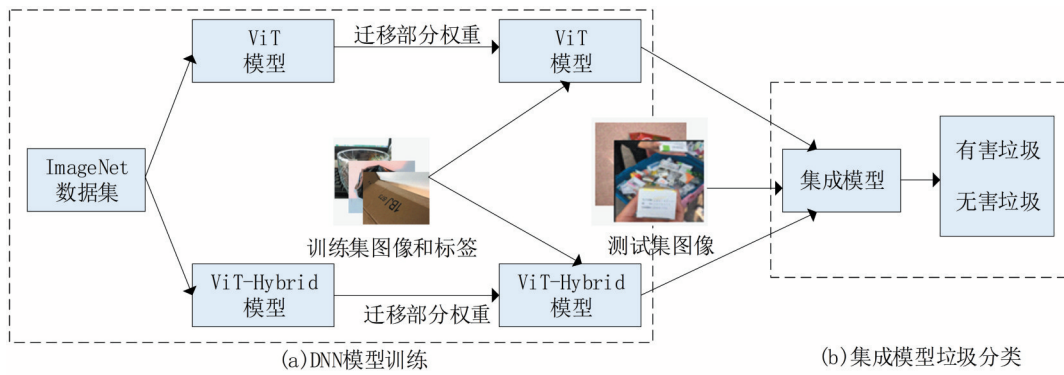


图2 集成学习方法用于自动识别垃圾图像

1.3 模型评价

现有的研究将所有垃圾类别平等对待，通常使用准确率来评价模型性能。本文旨在有效分离有害垃圾和无害垃圾，尽可能地减小有害垃圾被错误地识别为无害垃圾的概率。因此，本文使用三个评价指标：有害垃圾图像的漏判误差(E_{oh})、无害垃圾图像的误判误差(E_{ci})，图像总体误差(E_{ov})，分别定义为公式(2)、(3)和(4)。

$$E_{oh} = \frac{FN}{TP + FN} \quad (2)$$

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

$$E_{ov} = \frac{FN + FP}{TP + TN + FN + FP} \quad (4)$$

其中 TP 表示实际是有害垃圾，模型也将其预测为有害垃圾的照片数量。 TN 表示实际是无害垃圾，模型也将其预测为无害垃圾的照片数量。 FP 表示实际是无害垃圾，模型将其预测为有害垃圾的照片数量。 FN 表示实际是有害垃圾，模型将其预测为无害垃圾的照片数量。记有害垃圾图像集合为 H ，无害垃圾图像集合为 I 。定义有害垃圾标签集合 $L_h = \{\text{过期药品、废旧电池、过期药品}\}$ ；无害垃圾标签集合 $L_i = \{\text{废玻璃、废纸、废金属、废塑料}\}$ 。那么， TP 、 TN 、 FP 和 FN 可分别定义为公式(5)、(6)、(7)和(8)：

$$TP ::= \{\forall im \in H, x \in L_h, f(im) = x\} \quad (5)$$

$$TN ::= \{\forall im \in I, x \in L_i, f(im) = x\} \quad (6)$$

$$FP ::= \{\forall im \in I, x \in L_h, f(im) = x\} \quad (7)$$

$$FN ::= \{\forall im \in H, x \in L_i, f(im) = x\} \quad (8)$$

“ $::=$ ”表示“定义为”，“ $f(im)$ ”表示集成模型为图像“ im ”分配的预测标签。对任意一张有害垃圾图像，模型只要将其预测为集合 L_h 中的任一标签都属于 TP ；对任意一张无害垃圾图像，模型只要将其预测为集合 L_i 中的任一标签都属于 TN ；对任意一张无害垃圾图像，模型只要将其预测为集合 L_h 中的任一标签都属于 FP ；对任意一张有害垃圾图像，模型只要将其预测为集合 L_i 中的任一标签都属于 FN 。

E_{oh} 直观地反映模型错误预测的有害垃圾图像占测试集有害垃圾图像总数的比例。 E_{ci} 直观地反映模型自动预测为有害垃圾图像中无害垃圾图像的比例。

2 实验结果

有害垃圾图像和无害垃圾图像二分类的实验结果如表3所示，集成模型漏判误差(E_{oh})、误判误差(E_{ci})和总体误差(E_{ov})分别0%、11.63%和4.57%。实验结果表明，尽管集成模型的误判误差和总体误差都有所增加，但保守的集成策略有效地降低了有害垃圾的漏判误差，防止了将有害垃圾识别为无害垃圾。图3为集成模型和两个DNN模型对应的二分类混淆矩阵。

表3 单模型与集成模型各指标对比

模型	漏判误差 /%	误判误差 /%	总体误差 /%
集成模型	0	11.63	4.57
ViT	1.75	6.67	3.05
ViT-Hybrid	4.36	7.63	4.27

		真实标签 (集成模型)				真实标签 (ViT)				真实标签 (ViT-Hybrid)	
		有害	无害			有害	无害			有害	无害
模型 预测	有害	114	15	模型 预测	有害	112	8	模型 预测	有害	109	9
	无害	0	199		无害	2	206		无害	5	205

图 3 集成模型(左)、ViT(中)、ViT-Hybrid(右)对有害垃圾、无害垃圾的预测结果

3 讨论

基于保守策略的集成学习方法在识别垃圾时,保证了有害垃圾不会被识别为无害垃圾。但是垃圾分类的总体准确率也是值得关注的指标。图4给出了多分类实验结果的混淆矩阵。混淆矩阵左上方 3 × 3 的矩阵是有害垃圾内部的预测结果,右下方 4 × 4 的矩阵是无害垃圾内部的预测结果。

		真实标签(集成模型)						
		电池	药品	软膏	玻璃	纸	金属	塑料
模型 预测	电池	32	0	1	0	0	1	0
	药品	0	41	3	3	3	0	4
	软膏	0	0	37	0	2	1	1
	玻璃	0	0	0	43	1	5	3
	纸	0	0	0	0	46	1	2
	金属	0	0	0	1	0	50	4
	塑料	0	0	0	0	0	0	43

图 4 多分类实验结果混淆矩阵

多分类的准确率(Accuracy)计算公式如(9)所示:

$$Accuracy = \frac{\sum_{i=1, j=1}^n X_{ij}}{\sum_{i=1, j=1}^n X_{ij}}, (i, j = 1, 2, \dots, n) \quad (9)$$

其中 $X_{i,j}$ 表示实际为第 i 类别,模型将其预测为第 j 类别的标签的图像数, n 为类别数。

从图4可以发现,实际为有害垃圾,集成模型也将其预测为有害垃圾的图像(114张)中,其准确率为 96.49%;实际为无害垃圾,模型也将其预测为无害垃圾的图像(198张)中,其准确率为 91.92%。实验结果表明,集成模型在有效控

制了有害垃圾漏判误差的同时,将有害和无害两类垃圾内部类别间的误识别误差控制在一个较低的水平。

4 结语

针对如何尽可能降低有害垃圾的漏判误差问题,本文设计了基于保守策略的垃圾图像自动识别集成学习方法,在自制 DTrash 数据集上训练集成模型,测试结果显示,集成模型获得了较低的漏判误差(0%),有效地防止了有害垃圾被误识别为无害垃圾,并将有害和无害两类垃圾内部类别间识别误差维持在一个较低的范围。

参考文献:

- [1] KONG L Q, TIAN G J, LIU X J. The spatio-temporal dynamic pattern of urban solid waste emission in China [J]. China Environmental Science, 2017, 37(4): 1408-1417.
- [2] WANG Z, PENG B, HUANG Y, et al. Classification for plastic bottles recycling based on image recognition[J/OL]. Waste Management, 2019, 88: 170-181. DOI:10.1016/j.wasman.2019.03.032.
- [3] TACHWALI Y, AL-ASSAF Y, AL-ALI A R. Automatic multistage classification system for plastic bottles recycling [J]. Resources Conservation & Recycling, 2008, 52(2): 266-285.
- [4] CHU Y, HUANG C, XIE X, et al. Multilayer hybrid deep-learning method for waste classification and recycling[J]. Computational Intelligence and Neuroscience, 2018: 1-9.
- [5] MITTAL G, YAGNIK K B, GARG M, et al. Spot-Garbage: smartphone App to detect garbage using deep learning[C]//ACM International Joint Conference on Pervasive & Ubiquitous Computing, 2016: 12-16.

- [6] 康庄,杨杰. 基于机器视觉的垃圾自动分类系统设计[J]. 杭州:浙江大学学报(工学版),2020,54(7): 1272-1280.
- [7] 袁建野,南新元,蔡鑫,等. 基于轻量级残差网路的垃圾图片分类方法[J]. 环境工程,2021,39(2): 110-115.
- [8] ZHANG Q, YANG Q, ZHANG X, et al. Waste image classification based on transfer learning and convolutional neural network[J/OL]. Waste Management, 2021, 135: 150-157. DOI: 10.1016/j.wasman.2021.08.038.
- [9] YANG M, THUNG G. Classification of trash for recyclability status[R]. CS229 Project Report, 2016.
- [10] MAO W L, CHEN W C, WANG C T, et al. Recycling waste classification using optimized convolutional neural network [J/OL]. Resources, Conservation and Recycling, 2021, 164: 105132. DOI: 10.1016/j.resconrec.2020.105132.
- [11] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. arXiv Preprint arXiv:1409.1556, 2014.
- [12] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C/OL] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 770-778. DOI: 10.1109/CVPR.2016.90.
- [13] XIE S, GIRSHICK R, P DOLLÁR, et al. Aggregated residual transformations for deep neural networks [C/OL] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 5987-5995. DOI: 10.1109/CVPR.2017.634.
- [14] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: transformers for image recognition at scale [EB/OL]. arXiv Preprint arXiv:2010.11929, 2020.

作者简介:

孟德尧(1997—),男,河南安阳人,在读硕士研究生,研究方向为图像识别

吴荣海(1975—),男,云南大理人,硕士,副教授,研究方向为人工智能

通信作者:杨邓奇(1979—),男,云南大理人,博士,教授,研究方向为图像识别, E-mail: dqyang@dali.edu.cn

收稿日期: 2022-03-06 修稿日期: 2022-04-17

Research on Automatic Identification Method of Hazardous Waste Based on Ensemble Learning

Meng Deyao¹, Wu Ronghai², Yang Dengqi¹

(1. College of Mathematics and Computer Science, Dali University, Dali 671003;

2. Engineering Training Center, Dali University, Dali 671003)

Abstract: Deep learning is a tool for machine learning to provide automatic image recognition, which is widely used in automatic recognition of garbage images. Limited by the data set, the existing waste classification studies treat all waste categories equally, without considering the harm of misclassification of hazardous waste. Firstly, a garbage image data set (DTrash) containing harmful garbage is constructed based on web crawler and manual shooting, and an automatic recognition method of harmful garbage based on ensemble learning is proposed based on conservative ensemble learning strategy. The missing judgment error, misjudgment error and overall error of ensemble learning method on DTrash are 0%, 11.63% and 4.57% respectively.

Keywords: harmful waste; deep neural network; waste classification; deep learning; ensemble strategy

文章编号: 1007-1423(2022)16-0043-05

DOI: 10.3969/j.issn.1007-1423.2022.16.008

机器学习在心肌梗塞并发症中的预测研究

王 蔚, 程 君, 李先杰, 彭 雷

(自贡市第一人民医院信息科, 自贡 643000)

摘要: 为了降低心肌梗塞患者伴有并发症的发生率, 利用机器学习方法构建心肌梗塞并发症预测模型, 以心肌梗塞患者的医疗数据作为输入, 以心肌梗塞患者的并发症类型作为输出, 辅助临床医务人员早期判断, 提前采取必要的干预措施。研究表明, Linear_SVM 模型的整体预测性能优于 MLP 模型和 RBF_SVM 模型, 其预测准确率为 76.28%, 特别是在心房纤颤、三度房室传导阻滞、心肌破裂和心肌梗死后综合征上表现出较好的预测效果。

关键词: 机器学习; 多层感知机; 支持向量机; 心肌梗塞并发症

基金项目: 四川省卫生信息学会 2021 年度科研课题项目(2021021)

0 引言

根据 2021 年 7 月发布的《中国心血管健康与疾病报告 2020》^[1], 我国城乡居民总死亡原因当中, 心血管疾病死亡居于首位, 高于肿瘤及其他疾病。心肌梗塞(myocardial infarction, MI)^[2]作为一种常见的心血管疾病, 严重威胁人们的生命健康, 从 2005 年开始, 其死亡率呈快速上升趋势。心肌梗塞患者的不同病程, 可能不存在并发症或者存在并发症但不会严重威胁生命。与此同时, 大约一半的急性和亚急性期心肌梗塞患者有并发症, 如心力衰竭、心脏破裂、心房纤颤等, 其致死率及致残率仍然较高。因此, 提前预测心肌梗塞患者未来可能发生的并发症类型, 能够有效辅助临床医务人员实施必要的预防措施, 也是降低心肌梗塞患者死亡率的重要途径。

1 相关研究

近年来, 大数据、人工智能等技术的快速发展为疾病的诊断和预测提供了强大的技术支持, 并表现出较好的应用效果, 例如糖尿病早期预测^[3]、糖尿病并发症预测^[4]、脊柱手术并发症预测^[5]、脑肿瘤分类^[6]等。

目前, 心肌梗塞患者的并发症研究主要集

中于评价心肌梗塞并发症与影响因子之间的关系, 通过观察影响因子来达到预测心肌梗塞并发症的目的。黄佐贵等^[7]采用免疫比浊法测定 51 例心肌梗塞患者组和 40 例冠心病对照组的高敏 C 反应蛋白水平来预测心肌梗塞并发心力衰竭死亡率, 实验结果表明高敏 C 反应蛋白可作为预测心肌梗塞患者一年内总死亡率和心力衰竭死亡率的一项重要指标。宗敏等^[8]采用统计学方法中的多因素 logistic 回归分析来探讨血小板分布宽度对老年人心肌梗塞患者住院期间心力衰竭的预测价值。实验结果表明血小板分布宽度对 MI 后心衰有独立预测价值, 具有较高的敏感性, 可作为评价 MI 后心衰的独立危险因素。Azwar^[9]使用机器学习算法中的随机森林模型来预测心肌梗塞患者发生心肌破裂并发症的可能性, 该模型以患者的年龄、性别、病死中的心肌梗塞次数、病史中的劳力性心绞痛、高血压等作为模型的输入特征, 实验结果表明该模型具有较好的预测效果。由于心肌梗塞的并发症不同, 其影响因子也不同, 想要利用影响因子来对心肌梗塞并发症的总体分类进行预测, 将会变得十分困难。基于此, 本文通过患者的既往病史, 入院时体征记录以及入院三天的治

疗措施来构建多层感知机(multilayer perceptron, MLP)和支持向量机(support vector machine, SVM)两种心肌梗塞并发症预测模型,为临床医务人员早期判别心肌梗塞患者发生并发症提供风险预警。

2 模型与方法

2.1 多层感知机模型

感知机^[10]作为神经网络和支持向量机的起源算法,最初用于处理简单的二元分类线性可分问题,即能获得一个将训练数据集正实例点和负实例点完全正确分开的分离超平面,却无法构造非线性可分数据集的超平面。多层感知机^[11],也称为多层神经网络,是在单层神经网络的基础上引入了一个或多个隐藏层,并使用非线性激活函数来克服线性模型的限制,提升非线性数据处理能力,其基本结构如图1所示。

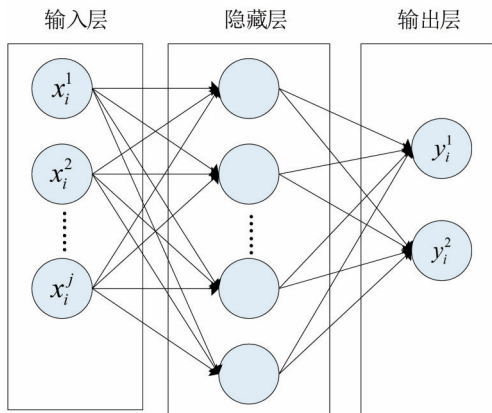


图1 多层感知机模型

由图1可知,MLP是由一个输入层,一个输出层以及若干个隐藏层构成。输入层负责将心肌梗塞患者*i*的*j*个特征作为输入,隐藏层的神经元节点对输入层的输入特征采取加权求和,通过非线性激活函数产生激活响应,使得神经网络可以任意逼近任何非线性函数。在训练过程中,采取向前传播和反向传播算法,不断调整模型权重和偏置参数,从而达到更好的分类效果。

2.2 支持向量机模型

SVM^[12]是由输入空间、特征空间和输出空

间组成,在特征空间中可能存在多个超平面将数据进行划分,依据结构风险最小化原则,从中寻找间隔最大的分类超平面,其基本结构如图2所示。在真实应用数据集中,很难在现有特征空间中找到一个能够正确划分样本的超平面,因此需要使用核函数将原始的特征空间映射到更高维的特征空间中,使得样本在这个高维的特征空间中线性可分,即可求得最优分类超平面。

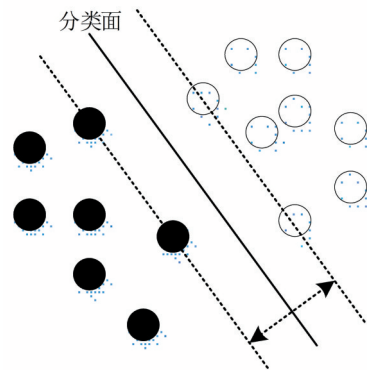


图2 支持向量机模型

2.3 评价指标

为了验证心肌梗塞患者并发症预测模型的有效性,需要对其预测效果进行评价,本文从准确率、召回率、AUC三个维度对模型进行评价。

(1) 准确率 *accuracy*: 心肌梗塞患者发生并发症与不发生并发症的预测准确程度,如表达式(1)所示。

$$accuracy = \frac{TP + TN}{TP + FN + FP + TN} \quad (1)$$

式中, *TP*表示正确预测心肌梗塞患者发生并发症的数量, *TN*表示正确预测心肌梗塞患者不发生并发症的数量, *FN*表示错误预测心肌梗塞患者不发生并发症的数量, *FP*表示错误预测心肌梗塞患者发生并发症的数量。

(2) 召回率 *recall*: 正确预测心肌梗塞患者发生并发症的数量占样本集中有并发症的比例,如表达式(2)所示。

$$recall = \frac{TP}{TP + FN} \quad (2)$$

(3) *AUC*: 表示ROC曲线下的面积,用于

判断模型的优劣,其值一般介于0.5和1之间,AUC值越大,表明模型的分类效果越好。

3 实验设计与结果分析

3.1 实验数据集

本文使用的数据集来源于加州大学欧文分校的UCI机器学习数据库中的心肌梗塞并发症数据集^[13],总计1700例,数据信息包含患者以往病史(如心肌梗死数量,心绞痛的功能等级、肺结核的既往病史、糖尿病的既往病史等),入院时患者体征(如心电图节律、室颤心电图等),入院三天治疗措施(如药品使用频率),共计111个输入特征向量,11种并发症以及无并发症作为样本标签结果,对每一种并发症建立二分类预测模型。

数据集中含有1700例患者的样本数据,每例样本数据包含111个临床数据点,由于部分记录数据存在数据缺失情况,直接影响数据挖掘的结果。针对数据集中缺失的数据,本文采用K最近邻算法^[14]进行数据补齐,即首先采用K-means聚类将所有样本进行聚类划分,随后通过划分的种类的均值对各自类中的缺失值进行填补。

心肌梗塞常见于中老年人,本数据集中心肌梗塞患者的年龄分布如图3所示。从图中可以看出,心肌梗塞患者的年龄多集中于[40,80]之间,平均年龄为61.85岁。

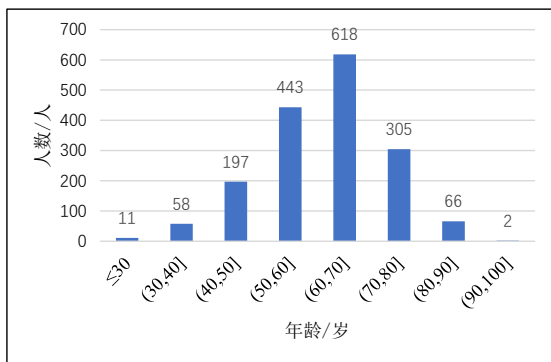


图3 心肌梗塞患者的年龄分布

随着病程的发展,心肌梗塞患者可能没有并发症,也可能有一种或多种并发症,图4展

示了数据集中各并发症的人数分布情况。

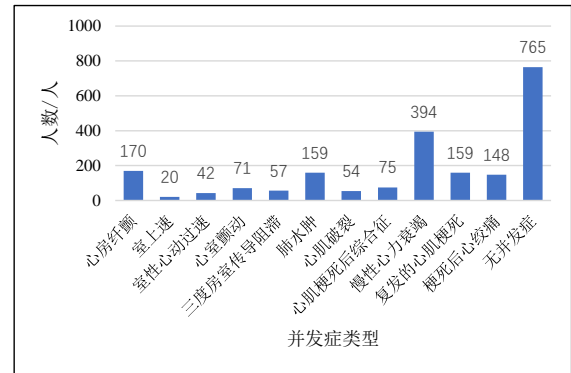


图4 心肌梗塞患者的并发症类型分布

由图4可知,心肌梗塞患者最容易面临慢性心力衰竭并发症,与临床实际经验一致。本数据集中各并发症的样本集,有并发症类别的样本数远低于没有并发症类别的样本数,存在数据类别不平衡的问题。基于此,本文对每一种并发症的样本采用随机过采样^[15]的方法扩充训练集的数据样本量,达到样本平衡的目的。

3.2 模型参数设置

本模型以数据集的80%作为训练集,以此不断地训练模型,优化模型参数,使得整个模型收敛,然后将剩余的20%作为测试集,评估模型的预测效果。由于心肌梗塞患者可能发生多个并发症的情况,违背多分类输出互斥的原则,故对数据集中的每一种并发症分别建立MLP和SVM模型。经过大量的训练拟合测试后,确定了两种模型的使用参数。

针对MLP,本文构建1个输入层,1个隐藏层,1个输出层的网络结构,其中隐藏层的神经元数量为100,损失函数LogLoss,优化函数Adam,激活函数ReLU。

针对SVM,利用python中sklearn.svm.svc()函数,在核函数系数和惩罚松弛变量为默认值的情况下,分别采用Linear和RBF两种核函数,即Linear_SVM和RBF_SVM进行预测分析比较。

3.3 结果分析

利用MLP、Linear_SVM和RBF_SVM三种模型对数据集中的每一种并发症进行预测,其性能如表1~表3所示。

表1 基于MLP的心肌梗塞并发症预测结果

并发症类型	accuracy / %	recall / %	AUC
心房纤颤	85.88	35.29	0.63
室上速	96.24	29.13	0.62
室性心动过速	98.82	50.00	0.74
心室颤动	94.12	16.67	0.57
三度房室传导阻滞	97.06	33.33	0.66
肺水肿	88.82	33.33	0.64
心肌破裂	93.53	33.33	0.65
心肌梗死后综合征	90.58	21.42	0.59
慢性心力衰竭	76.47	41.67	0.64
复发的心肌梗死	81.18	33.33	0.60
梗死后心绞痛	86.47	37.59	0.63

表2 基于Linear_SVM的心肌梗塞并发症预测结果

并发症类型	accuracy / %	recall / %	AUC
心房纤颤	75.30	70.59	0.73
室上速	85.88	60.00	0.68
室性心动过速	85.89	56.12	0.58
心室颤动	85.29	63.14	0.72
三度房室传导阻滞	86.47	66.67	0.77
肺水肿	70.00	68.00	0.69
心肌破裂	78.24	75.00	0.77
心肌梗死后综合征	74.70	71.43	0.73
慢性心力衰竭	66.76	63.28	0.64
复发的心肌梗死	64.12	60.00	0.61
梗死后心绞痛	66.47	63.33	0.65

表3 基于RBF_SVM的心肌梗塞并发症预测结果

并发症类型	accuracy / %	recall / %	AUC
心房纤颤	64.12	41.18	0.54
室上速	64.11	50.00	0.57
室性心动过速	75.88	50.00	0.63
心室颤动	55.29	57.14	0.56
三度房室传导阻滞	74.12	50.00	0.62
肺水肿	57.65	81.25	0.68
心肌破裂	70.58	60.00	0.65
心肌梗死后综合征	54.71	57.14	0.56
慢性心力衰竭	55.29	43.59	0.51
复发的心肌梗死	55.29	43.59	0.61
梗死后心绞痛	63.53	60.00	0.62

为了进一步比较三种模型在心肌梗塞并发症中的预测性能,对表1~表3中的所有并发症的预测结果取均值,将其作为模型的性能评价指标,结果如表4所示。从表4可以看出,相比于Linear_SVM和RBF_SVM,MLP的准确率最

高,召回率和AUC最低,这与原始数据集中的并发症样本数量息息相关,即使扩充样本数量,仍然面临无法准确挖掘出这些并发症与输入特征之间的关联,往往出现过拟合,使得模型的泛化能力较弱。在兼顾准确率、召回率、AUC三种评价指标的情况下,Linear_SVM的整体表现最优,可见在小样本、数据高维的心肌梗塞并发症数据集中,基于Linear核函数的支持向量机表现出更优的模型泛化能力。特别是心房纤颤、三度房室传导阻滞、心肌破裂、心肌梗死后综合征四种并发症均表现出较好的准确率、召回率和AUC,表明该模型对这四种并发症具有较好的拟合能力。

表4 心肌梗塞并发症预测模型性能对比

模型	accuracy / %	recall / %	AUC
MLP	89.92	33.19	0.63
Linear_SVM	76.28	65.23	0.69
RBF_SVM	62.78	53.99	0.59

4 结语

近年来,在医疗领域中利用机器学习算法辅助医务人员完成疾病诊断一直是研究的热点和难点,目前已经广泛应用于心血管疾病、糖尿病、肿瘤、肾病等疾病预测中,并取得了较好的应用效果。在相关应用研究中,研究人员将患者的临床检测数据作为机器学习模型的输入,将疾病的发生风险作为预测输出,有效辅助医务人员提前采取干预措施,减少患病风险,达到早发现、早预防、早治疗的诊疗目的。

目前,研究人员重点关注挖掘诊疗过程的部分特征数据与特定的心肌梗塞并发症之间的关联,忽略了患者诊疗前的病史数据。本文采用公开的真实数据集,根据心肌梗塞患者既往病史,入院时的体征数据以及入院三天的治疗措施,通过机器学习算法中的多层感知机模型和支持向量机模型,建立心肌梗塞患者并发症的预测模型。实验表明基于Linear核函数的支持向量机在本数据集下取得了较好的预测效果,对临床医务人员早发现心肌梗塞并发症有一定的参考价值。

在心肌梗塞并发症的后续研究中,一方面需要持续优化模型参数,尝试使用更多的分类

预测模型,进一步提高模型的预测准确率、召回率、AUC值;另一方面可使用深度学习等方法继续深入分析现有输入特征与心肌梗塞并发症之间的内在联系,对各并发症的影响因子进行重要性排序。

参考文献:

- [1] 《中国心血管健康与疾病报告》编写组.《中国心血管健康与疾病报告2020》要点解读[J]. 中国心血管杂志,2021,3(26):209-218.
- [2] 杨楚诗,张朋柱.基于行为数据的急性心肌梗塞患病风险预测[J]. 计算机仿真,2021,4(38):442-446.
- [3] VIJAYAKUMAR K, LAVANYA B, NIRMALA I, et al. Random forest algorithm for the prediction of diabetes [C]//2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN),2019:1-5.
- [4] 陈思含,张云秋.基于机器学习的2型糖尿病并发症预测模型研究[J]. 中华医学图书馆情报杂志,2020,29(11):31-38.
- [5] HODA M, SADDIK A E, WAI E, et al. Predicting spine surgery complications using machine learning [C]//2019 IEEE International Conference on Multimedia & Expo Workshops(ICMEW),2019:49-53.
- [6] ÇINARER G, EMİROĞLU B G. Classification of brain tumors by machine learning algorithms [C]//2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT),2019:1-4.
- [7] 黄佐贵,杜国伟.高敏C反应蛋白水平对急性心肌梗塞并发心衰死亡率的预测价值[J]. 心血管康复医学杂志,2010,19(2):154-156.
- [8] 宗敏,关晓楠,常晶,等.血小板分布宽度对老年急性心肌梗死患者院内心力衰竭的预测价值[J]. 中华老年心脑血管病杂志,2021,4(23):375-378.
- [9] AZWARI S A. Predicting myocardial rupture after acute myocardial infarction in hospitalized patients using machine learning [C]//2021 National Computing Colleges Conference(NCCC),2021:1-6.
- [10] 徐洪学,汪安祺,杜英魁,等.深度学习的基本模型及其应用研究[J]. 长春师范大学学报,2020,39(12):47-54,93.
- [11] 张振.基于机器学习的心血管疾病风险预测关键技术研究[D]. 成都:电子科技大学,2021.
- [12] 汪海燕,黎建辉,杨风雷.支持向量机理论及算法研究综述[J]. 计算机应用研究,2014,31(5):1281-1286.
- [13] UCI machine learning repository[EB/OL]. <http://archive.ics.uci.edu/ml>.
- [14] 毋雪雁,王水花,张煜东. K最近邻算法理论与应用综述[J]. 计算机工程与应用,2017,53(21):1-7.
- [15] 刘定祥,乔少杰,张永清,等.不平衡分类的数据采样方法综述[J]. 重庆理工大学学报(自然科学),2019,33(7):102-111.

作者简介:

王蔚(1980—),男,四川自贡人,硕士,副主任医师,研究方向为心血管疾病、介入技术

程君(1990—),男,重庆万州人,本科,工程师,研究方向为数据分析与挖掘

李先杰(1982—),男,山东青岛人,硕士,工程师,研究方向为模式识别与智能系统

彭雷(1984—),男,四川自贡人,硕士,工程师,研究方向为数据分析与挖掘、大数据应用

收稿日期:2022-04-06 修稿日期:2022-05-12

The Prediction of Myocardial Infarction Complications with Machine Learning

Wang Wei, Cheng Jun, Li Xianjie, Peng Lei

(Information Section, Zigong First People's Hospital, Zigong 643000)

Abstract: In order to reduce the incidence of myocardial infarction patients accompanied by complications, using machine learning method to build prediction models of myocardial infarction complication in patients with myocardial infarction medical data as input, with myocardial infarction patients with complications of type as the output, auxiliary judgment early clinical medical personnel, take the necessary intervention measures in advance. The results show that the overall prediction performance of Linear_SVM model is better than that of MLP model and RBF_SVM model, with a prediction accuracy of 76.28%, especially for atrial fibrillation, third-degree atrioventricular block, myocardial rupture and post-myocardial infarction syndrome.

Keywords: machine learning; multilayer perceptron; support vector machine; complications of myocardial infarction

文章编号: 1007-1423(2022)16-0048-05

DOI: 10.3969/j.issn.1007-1423.2022.16.009

基于边缘提取与多项式拟合的三维位姿解算方法

朱 腾^{1,2}, 郭晓莹³, 阎汉生^{2,3}, 梁剑斌⁴, 舒 彬⁴

(1. 广东工贸职业技术学院测绘遥感信息学院, 广州 510550;

2. 广东工贸职业技术学院测绘地理信息技术虚拟仿真实训基地, 广州 510510;

3. 广东工贸职业技术学院机电工程学院, 广州 510510;

4. 广东工业大学机电工程学院, 广州 510039)

摘要: 现代智能装配环境对装配指引的要求不断提高, 新型的 AR 装配指引系统以其独有的三维模型叠加显示能力而受到生产单位及研究人员的关注。其中快速鲁棒的目标位姿检测是增强现实(augmented reality, AR)智能装配指引系统的关键步骤之一。为了解决单目图像中复杂纹理物体的注册配准问题, 本文提出一种基于边缘提取与多项式回归拟合的三维物体配准方法。该方法主要采用长板工件的边缘轮廓特征, 通过构建三维模型与二维特征的匹配关系增强算法鲁棒性。同时为提高算法响应速度, 本文通过引入多模板匹配, 减少了位姿求解的拟合时间。最后, 实验结果部分展示了本文算法在实际天线板 AR 布线辅助系统中的工作效果。

关键词: 增强现实; 装配指引; 位姿解算; 边缘特征

基金资助: 广州市基础研究计划基础与应用基础研究项目(20210201298): 多人协同虚拟仿真实验平台三维跟踪注册技术研究; 广东工贸职业技术学院高层次人才专项(2021-gc-06): 无人机测绘虚拟仿真实训系统研究; 广东省科技创新战略专项(pdjh2021a0741): 后疫情时代线上商品展示中的 VR/AR 关键技术研究

0 引言

现代的工业体系愈发复杂化、精细化, 在产品装配、电路板布线等操作要求复杂的生产需求中, 由于操作工艺自由度大, 逻辑要求较高, 难以实现全自动化生产, 必须引入人工操作, 这就对生产装配的指引系统提出了更高的要求。

增强现实(AR)是一种利用计算机和相关显示设备将虚拟信息叠加到用户视觉中来增强真实场景信息的技术^[1]。与早期的文本信息叠加相比, 目前 AR 的研究主要集中在如何准确、稳定地将三维虚拟模型“放置”在真实世界中^[2]。要想获得目标在虚拟场景中的位置信息, 就必须计算出目标在虚拟场景中的位置, 从而将二维投影无偏差地显示出来, 然后将三维虚拟对象无缝地添加到真实环境中^[3]。

目前, 目标姿态检测是计算机视觉领域的一个研究热点^[4-5]。在硬件系统上, 主要有两种

方法: 基于双目视觉结合深度信息的目标重建与匹配^[6], 以及基于单目摄像机与图像边缘^[7]、纹理或特征点的特征匹配^[8]。其中, 双目视觉方法能有效地恢复深度信息, 但需要的计算量大, 难以满足实时跟踪需要^[9]。而实际生产环境对实时性和经济性的要求使得单目视觉方法成为了更受关注的研究方向。

针对经典实验算法与应用要求不匹配的问题, 本文提出了一种基于直线特征和多项式拟合的平板目标三维位置和方向解算方法。实验表明, 该算法能克服实际生产环境下的光照不稳定、自然特性易被遮挡、金属表面材料等问题。

1 理论基础

1.1 研究场景

本文研究对象为大型基站天线板布线过程中的 AR 辅助系统。其主体架构包括 AR 眼镜、

图形处理终端、场景监控摄像头、工作台及目标工件。其中,出于成本控制及实际工作需要,在天线板、工作台等监测场景中均不允许出现人工标记,同时场景监控采用一个普通光学摄像头,如图1所示。

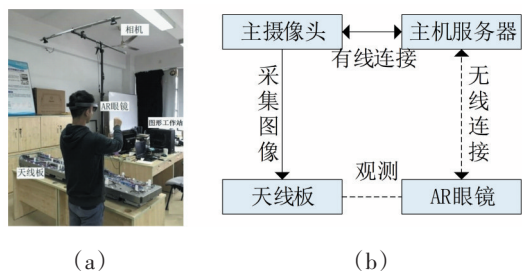


图1 实验工作场景(a)与系统架构(b)

在该研究环境下,头顶摄像机是主要的信息采集设备,它利用全局监控的摄像设备获取目标天线板的位置和姿态变换信息,并控制AR场景中的虚拟模型完成相应的运动。

1.2 单目位姿跟踪原理

单目视觉监测虽然难以获得深度信息,但具有成本低、计算速度快、精度适中等综合优势。单眼透视成像模型如图2所示,它定义了世界坐标系 $O-x_w, y_w, z_w$ 和相机坐标系 $C-x_c, y_c, z_c$ 。

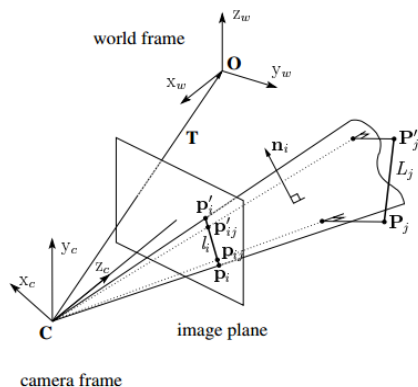


图2 单目视觉位姿检测模型

$$\begin{bmatrix} x_c \\ y_c \\ 1 \end{bmatrix} = K \begin{bmatrix} R_0 & T_0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (1)$$

$$\begin{aligned} R_0 &= R_z R_y R_x \\ &= \begin{bmatrix} \cos \theta_z & -\sin \theta_z & 0 \\ \sin \theta_z & \cos \theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \theta_y & 0 & \sin \theta_y \\ 0 & 1 & 0 \\ -\sin \theta_y & 0 & \cos \theta_y \end{bmatrix} \\ &\quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_x & -\sin \theta_x \\ 0 & \sin \theta_x & \cos \theta_x \end{bmatrix} \end{aligned} \quad (2)$$

其中 K 表示成像矩阵,包括相机内参、畸变系数等参数,可通过棋盘标定法^[10]求解,在固定系统中为定值。姿态矩阵 R_0 和位置向量 T_0 共同构成外参矩阵,及世界坐标系到相机坐标系的变换矩阵,为求解 R_0 和 T_0 ,至少需要4对同名点带入式(1)^[10]。其中,姿态矩阵 R_0 又可以分解为绕 x 、 y 、 z 三个方向的旋转角,由于基站天线板呈长条状,其 z 轴翻转角度 r 可限制在 $(-20^\circ, 20^\circ)$ 之间,而本文算法即研究单目摄像头下天线板二维图像与其姿态角、位移量的解算关系,如图3所示。

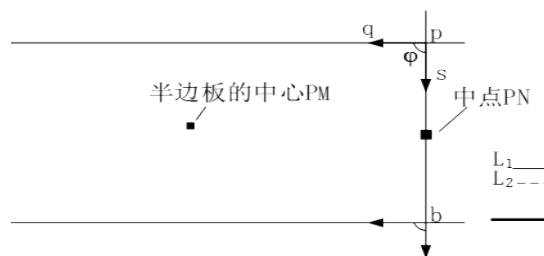


图3 天线板三个方向的旋转及位移示意图

2 基于线特征的三维天线板位姿检测

2.1 基于神经网络的边缘直线提取

天线板表面零件及布线情况将随着工序的进展而不断变化,故本文主要采用天线板边缘直线作为匹配特征,边缘线段的提取方法则借用Mask-RCNN神经网络^[11]出色的目标识别及分割能力来完成。边缘提取的流程包括训练集选取、特征提取、SVM分类、参数训练、神经网络

络分割、边缘拟合，如图4所示。

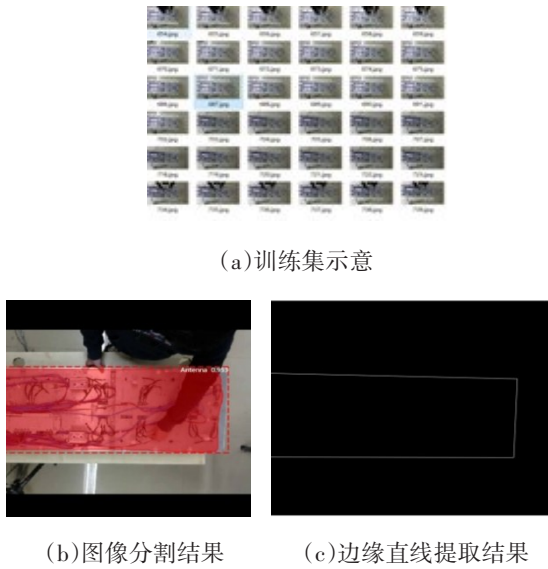


图4 Mask-RCNN边缘提取过程

2.2 基于多项式的翻转角拟合方法

提取得到天线板边缘的三根直线后，即可模拟天线板位姿与边缘直线的对应关系。其中位移向量 T_0 ，平面旋转角均可通过边缘直线快速求得，而对于 z 轴翻转角度 r ，本文提出使用多项式拟合方法求解：

记一组天线板边缘直线的夹角 φ 和与之对应的翻转角度 r 满足多项式(3)：

$$h_\theta(x) = \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \dots + \theta_n x^n \quad (3)$$

则可以得到惩罚函数：

$$J(\theta_0, \theta_1, \dots, \theta_n) = \frac{1}{2m} \sum_{i=1}^m (h_\theta(x^{(i)}) - y^{(i)})^2 \quad (4)$$

由上述两式即可根据给定步长拟合翻转角与边缘夹角在任意角度的对应关系。

3 实验结果

实验部分主要测试算法在实际生产环境下的鲁棒性，同时与 Holcon 的局域模板匹配方法对比了拟合精度与解算速度。在本文研究的智能生产案例中，采样选取了 $[-15^\circ, 15^\circ]$ 间共 7 组采样数据作为多项式拟合参考，如表1所示。

表1 采样数据表

I	1	2	3	4	5	6	7
φ	95.4	93.4	91.4	89.4	87.7	86.6	84.8
r	-15	-10	-5	0	5	10	15

带入惩罚函数式(4)后，解得拟合次数最优值为3。位姿解算方法测试部分主要测试了遮挡、光照、翻转三种干扰因素下的鲁棒性，结果如图5所示。

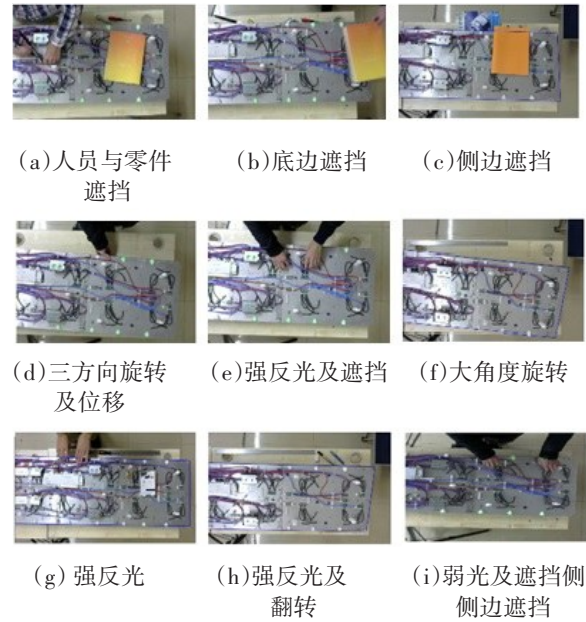


图5 不同工作环境下的形状匹配结果

从图5可以看到，本文采用的神经网络边缘提取与多项式拟合位姿求解算法对于不同的光照和遮挡干扰都有较好的适应能力。同时，为验证本文算法在位姿跟踪过程中的精度，根据 $[-15^\circ, 15^\circ]$ 间的7组采样数据进行了拟合结果与真实值的对比分析，结果如图6所示。

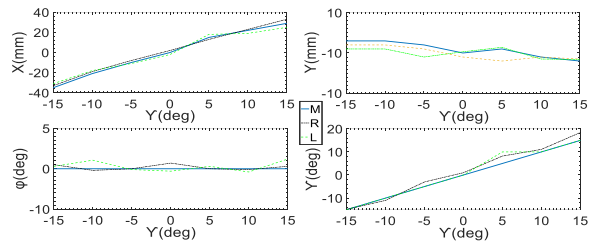


图6 位姿变化量实验数据与真实数据对比

其中 M 为样本真实值， R 为 Holcon 局部轮廓模板匹配方法解算出的翻转角， L 为基于边缘直线与多项式拟合求解出来的结果。对比结果表明，本文的位姿求解方法在 Y 方向的误差不超过5个像素， X 方向不超过2个像素，旋转角

φ 拟合精度可达 1° , 翻转角度可达 2° , 且与真实值误差均小于 Holcon 局部轮廓模板匹配方法。对于位姿解算速度, 局部轮廓模板匹配方法平均耗时 864 ms, 边缘直线多项式拟合方法平均耗时 643 ms, 效率提高 25% 以上。

图 7 展示了 AR 布线辅助系统视角下的虚实融合场景。从图 7 可以看到, 由于目标板材属于大型目标, 多项式拟合算法的误差对人眼产生的视觉上的影响难以察觉, 完全不影响操作人员进行布线操作。

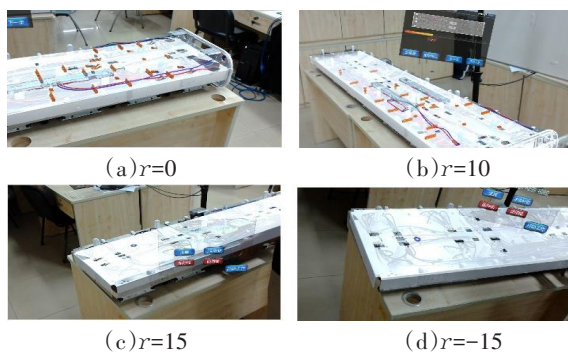


图 7 AR 眼镜端场景展示

4 结语

本文主要利用天线板的边缘直线进行目标工件的三维位姿解算, 结合实际生产环境下的位姿变化量分析, 通过降维、减小搜索空间方法提高位姿解算速度, 利用 Mask-RCNN 神经网络的抗干扰性完成天线板边缘的快速鲁棒提取, 最后通过多项式拟合的方法求出目标的位姿变化量, 并用于天线板 AR 布线辅助系统的虚实融合指引。实验结果表明, 结合直线特征与多项式拟合的位姿解算方法精度可达 2~5 像素, 旋转角拟合精度小于 1° , 均超过商业软件 Holcon 的局部模板匹配算法, 关键帧重定位效率提高 25% 以上。

基于结线特征与多项式拟合的位姿解算方法最终应用于大型天线板智能生产试验线路, 生产实测结果表明, 经过解空间优化后的单目三维注册方法在实时性、准确性、鲁棒性等方面均能够满足实际生产需求。在后续研究中,

本团队将继续结合企业一线生产需求, 从经济性、便携性、鲁棒性等方面进一步优化面向智能生产的 AR 装配指引系统。

参考文献:

- [1] 刘经纬. 增强现实中基于三维标志物的跟踪技术研究[D]. 成都: 电子科技大学, 2011.
- [2] CHEN Y B, HE H W, CHEN H E, et al. Improving registration of augmented reality by Incorporating DCNNs into visual SLAM[J]. International Journal of Pattern Recognition & Artificial Intelligence, 2018:1855022.
- [3] 张春明, 杨天鸿, 姜绍飞. 浅谈虚拟仿真实验教学中心引入无线穿戴式虚拟现实设备的经验体会[J]. 现代计算机(专业版), 2018(27):50-54.
- [4] HARVILLE M, RAHILI A, DARRELL T, et al. 3D pose tracking with linear depth and brightness constraints[C]//Proceedings of the Seventh IEEE International Conference on Computer Vision. IEEE, 1999, 1:206-213.
- [5] LU C, HAGER G, MJOLSNESS E. Fast and globally convergent pose estimation from video images[J]. Transaction on Pattern Analysis and Machine Intelligence(S0162-8828), 2002, 22(6):610-622.
- [6] DAVID P, DEMENTHON D, DURAISWAMI R, et al. Simultaneous pose and correspondence determination using line features[C]//IEEE Computer Society Conference on Computer Vision & Pattern Recognition. Madison, WI, USA, 2003:1-8.
- [7] ZHU M L, DERPANIS K G, YANG Y F, et al. Single image 3D object detection and pose estimation for grasping[C]. IEEE International Conference on Robotics & Automation. Hong Kong, China, 2014: 6907430.
- [8] CHEN S, LIANG L M, LIANG W Z, et al. 3D Pose tracking with multi-template warping and SIFT correspondences[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2015, 26(99):1-10.
- [9] FERRARI V, TUYTELAARS T, GOOL L V. Markerless augmented reality with a real-time affine region tracker[C]//IEEE & ACM International Symposium on Augmented Reality. New York, NY, USA. 2001:970518.

- [10] WANG K, LIU D, LIU Z, et al. A fast object registration method for augmented reality assembly with simultaneous determination of multiple 2D-3D correspondences [J]. *Robotics and Computer-Integrated Manufacturing*, 2020, 63: 101890.
- [11] HE K M, GKIOXARI G, DOLLAR P, et al. Mask R-CNN [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017, PP(99): 1-1.

郭晓莹(2000—),女,广东梅州人,在读专科,从事领域为虚拟现实三维建模

阎汉生(1978—),男,湖北宜昌人,硕士,副教授,研究方向为机械自动化与仿真模拟

梁剑斌(1988—),男,广东佛山人,博士研究生,研究方向为增强现实姿态跟踪

舒彬(1993—),男,湖北黄冈人,硕士,研究方向为增强现实注册配准

收稿日期: 2022-04-14 修稿日期: 2022-04-30

作者简介:

通信作者:朱腾(1989—),男,湖南邵阳人,博士,讲师,研究方向为增强现实场景感知,E-mail:jhgf_1234@sina.com

3D Object Pose Estimation Method Based on Edge Extraction and Polynomial Fitting

Zhu Teng^{1,2}, Guo Xiaoying³, Yan Hansheng^{2,3}, Liang Jianbin⁴, Shu Bin⁴

(1. School of Surveying and Remote Sensing Information, Guangdong Institute of Industry and Commerce, Guangzhou 510550;

2. National Virtual Simulation Training Base of Surveying and Mapping Geographic Information Technology, Guangdong Polytechnic of Industry and Commerce, Guangzhou 510510;

3. School of Mechanical and Electrical Engineering, Guangdong Institute of Industry and Commerce, Guangzhou 510510;

4. School of Mechanical and electrical Engineering, Guangdong University of technology, Guangzhou 510039)

Abstract: With the increasing requirements of modern intelligent assembly environment for assembly guidance, the novel AR assembly guidance system has attracted the attention of production units and researchers due to its unique 3D model overlay display ability. And a fast and robust target pose detection algorithm is one of the key steps in Augmented Reality (AR) assembly guidance system. In order to solve the registration problem of complex texture objects in monocular images, a 3D object registration method based on edge extraction and polynomial regression fitting is proposed in this paper. This method mainly uses the edge contour features of long plate-shaped workpiece, and enhances the robustness of the algorithm by constructing the matching relationship between 3D model and 2D feature. At the same time, in order to improve the efficiency of the algorithm, multi template matching is introduced to reduce the fitting time of pose solution. Finally, the experimental results show the effectiveness of the proposed algorithm in the practical antenna panel AR wiring assembly system.

Keywords: augmented reality; assembly guidance; pose estimation; edge feature

文章编号: 1007-1423(2022)16-0053-04

DOI: 10.3969/j.issn.1007-1423.2022.16.010

基于改进的YOLOv3算法研究

刘 博, 李艳武

(重庆三峡学院电子与信息工程学院, 重庆 404100)

摘要: 在目标检测实际应用领域中, 要求算法能够实现高精度的检测。针对YOLOv3算法检测精度不够高的问题, 本文改变激活函数为Mish、Swish, 同时将Focal loss和Giou、Ciou损失函数融入到损失计算方式中, 并且采用K-means对数据集进行重新聚类得到新的锚框, 在训练时采用数据增强和冻结网络参数的训练方式进行训练。最终在PascalVOC数据集上有2%的提升。

关键词: 目标检测; YOLOv3; 激活函数; 损失函数; K-means

基金项目: 重庆市三峡库区地质环境监测与灾害预警重点实验室开放基金项目(YB2020C0102)

0 引言

近年来, 随着人工智能技术的发展, 这项技术已被应用到无人驾驶、智能家居、智能监控、医疗系统等领域中。类似无人驾驶、智能监控等技术一定需要通过计算机获取图像, 再进行后续判断。目标检测技术作为人工智能领域中的重要分支, 其目的是让机器能够获取环境中重要的信息图像。目前, 目标检测技术分为两类: 传统目标检测算法和基于深度学习的目标检测算法。

传统目标检测算法主要包括Haar特征+Ada-boost算法等。基于深度学习的目标检测算法主要是利用卷积运算构建卷积神经网络对图像进行特征提取, 能够提取到图像更深层的特征和语义信息。基于深度学习的目标检测算法又可分为一阶目标检测算法和二阶目标检测算法。其中一阶目标检测算法包括YOLOv1^[1]、SSD、YOLOv2^[2]、YOLOv3^[3]等。二阶目标检测算法有R-CNN、Fast R-CNN、SPP-Net等。这两大类算法中, 一阶目标检测算法速度快, 但精度低; 二阶目标检测算法则精度高、但速度慢。

为了进一步提升YOLOv3算法的性能, 本文将网络中的LeakyReLU函数改成Swish、Mish^[4]激活函数, 然后采用Giou、Ciou^[5]损失函数优化算法的边框回归方式, 采用Focal loss^[6]优化算法

的分类损失, 利用K-means算法对数据集进行重新聚类得到新的锚框。训练时采用数据增强方法扩大训练的数据量, 通过对图片进行缩放及长和宽的扭曲、色域扭曲、翻转图片等操作, 增强网络的鲁棒性。另外采用了冻结网络参数的训练方法, 此方法是在网络训练的前一半时间对网络中的特征提取网络部分不进行参数更新, 后一半时间才会对整个网络进行参数更新。

1 YOLOv3算法介绍

YOLOv3算法是在YOLOv1和YOLOv2的基础上进行改进的, 同样也是一阶目标检测算法, 直接生成预测结果。YOLOv3的特征提取网络是将YOLOv2中的Darknet-19与残差结构结合提出的Darknet-53, 使得网络能够达到更深的层次。Darknet-53中没有使用池化层而是由 1×1 和 3×3 的卷积层组成, 通过 1×1 的卷积压缩网络。同时还加入了BN层, 主要是为了缓解网络的过拟合问题。网络的颈部分, 借鉴特征金字塔网络结构(FPN)实现三个不同尺度的预测。网络的损失函数有四部分, 矩形框中心点损失、预测框宽高损失、置信度损失和类别损失。预测框中心点损失和宽高损失均采用了平方差损失, 置信度损失和类别损失则是采用了交叉熵损失计算方式。

2 改进的YOLOv3算法

2.1 激活函数

YOLOv3算法的激活函数为LeakyReLU激活函数，其是在ReLU的基础加入了一个很小的负值，无法为正负输入值提供一致的关系预测。因此针对激活函数，本文采用Swish、Mish激活函数分别对算法进行改进。Swish函数是一种自门控激活函数，定义为：

$$\text{Swish}(x) = x\sigma(\beta x) \quad (1)$$

式中， β 为一个正值，当 x 很大时， $\sigma(\beta x)$ 就接近于1。此时激活函数处于“开”状态，函数值近似于 x 本身，克服了函数饱和的问题；当 $\sigma(\beta x)$ 接近0时，激活函数处于“关”状态，且函数值近似为0。Mish函数的表达式为：

$$\text{Mish}(x) = x * \tanh(\ln(1 + e^x)) \quad (2)$$

当 $x \rightarrow \infty$ 时，函数值都不存在饱和的情况； $x < 0$ 时还有轻微负值，不会导致神经元失活，理论上有很好的梯度流。经过学者们的实验，Mish函数在训练稳定性、平均准确率等方面都有了全面的提升。将上述两个激活函数用在YOLOv3模型上，代替原有激活函数来研究算法的性能，以及三种激活函数的性能差异。

2.2 位置回归损失

在YOLOv3算法中，对目标位置计算回归损失是对中心位置偏移量和框的宽高分别计算损失，且都采用均方差损失。这种回归方式没有把anchor当成一个整体，忽略了其他的影响因素。2016年旷视科技提出了计算真实框和预测框的IOU值作为损失函数，不再对坐标和宽高进行独立预测，而是作为一个整体，使预测结果更加准确。但IOU Loss也有不足之处：一是预测框和真实框之间IOU值为0时，不但无法反映两个框的距离，而且无法进行优化；二是IOU无法精确反映两个框的重合度大小。

2019年，斯坦福大学提出了Giou损失函数，相比于IOU新加入了一个包含真实框和预测框的最小的矩形框C。Giou损失函数公式如下：

$$L_{\text{giou}} = 1 - \text{IOU} + \frac{C - (A \cup B)}{C} \quad (3)$$

Giou同样存在缺点，当真实框和预测框属于包含关系时，Giou就成了IOU，依然无法区

分两个框的相对位置。对于垂直方向上的样本，误差很大，基本很难收敛。

针对上述问题，Ciou将Giou中引入的包含预测框和真实框的最小矩形框的面积惩罚项修改成直接计算两个框中心的距离，以此来加速收敛，同时还考虑了Boundingbox的纵横比，进一步提升了回归精度。Ciou损失函数的公式为：

$$L_{\text{ciou}} = 1 - \text{IOU} + \frac{\rho^2(A, B)}{c^2} + \alpha v \quad (4)$$

其中 $\frac{\rho^2(A, B)}{c^2} + \alpha v$ 为损失函数的惩罚项， α 是用作trade-off的参数， v 是用来衡量长宽比一致性的参数。如图1所示， ρ 表示预测框和真实框的中心点距离， c 为包含两者的最小矩形框的对角线距离。

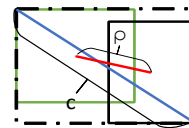


图1 Ciou原理图

针对YOLOv3定位不准确的问题，本文采用Giou、Ciou损失函数作为边框回归损失，研究三种损失计算方式的差异性，同时也研究算法的性能提升。

2.3 分类损失

YOLOv3的分类损失是用交叉熵损失函数来计算的，现改用Focal loss计算分类损失。Focal loss的提出是为了解决一阶目标检测中类别不平衡的问题。主要包含两个方面：一是正负样本不平衡的问题；二是难分类样本和易分类样本不平衡的问题。一张图片中，真实目标的数量远远大于负样本数量，负样本占总损失的大部分，同样易分类样本也占据了样本的大部分，这些样本容易主导模型的优化方向，导致模型分类能力下降。

Focal loss是在交叉熵损失的基础上通过引入 α 和 γ 两个参数来控制不同样本的权重，公式如下：

$$\text{FL}(p_i) = -\alpha_i(1 - p_i)^\gamma \log(p_i) \quad (5)$$

式中 α_i 为权重系数，主要用来平衡正负样本的数量比例不均程度； $(1 - p_i)^\gamma$ 为调制系数，通过

减少易分类样本的权重来使得模型更加关注于难分类的样本。

2.4 K-means 的锚框聚类

在YOLOv3算法中,锚框的大小需要预先设定,而且锚框的设定会直接影响到检测精度。为了能得到一组合理的锚框大小就需要用到K-means聚类算法,其是一种无监督的聚类算法,目的是将相似的框分为一类。使用K-means时,首先获取训练集的标签真实框的大小,然后随机选取 K 个不重复的框,采用IOU指标来对锚框进行聚类。具体流程如下:

- (1) 在所有真实框中随机挑选 k 个作为簇心;
- (2) 计算每个真实框与每个簇之间的 $1-IOU$ 值;
- (3) 计算每个真实框距离最近的簇心,并分配到最近的簇中;
- (4) 重新计算每个簇中的簇心;
- (5) 重复上述操作,直到每个簇中的元素不再改变。

针对YOLOv3算法的多尺度检测特性,需要设置9个锚框大小。首先按照上述步骤得到锚框,然后再引入遗传算法对得到的锚框进行变异,并得到最终的结果为: [17, 19], [22, 52], [50, 39], [47, 97], [94, 81], [95, 164], [225, 124], [163, 230], [324, 254]。

3 实验设置

实验是在Pascal VOC数据集上进行的,采用数据增强以及冻结参数训练,1到50轮冻结特征提取网络参数,不进行权值更新,学习率设置为 $1e-3$,batch_size设置为16;50到100轮不冻结特征提取网络的参数,学习率设置为 $1e-4$,batch_size设置为8。优化器采用Adam,权重衰减设置为 $5e-4$ 。学习率调整方式为StepLR,调节步长设置为1,调整倍数设置为0.94。实验中首先将LeakyRelu激活函数和原损失函数的YOLOv3算法作为基础实验,研究激活函数Mish、Swish和损失函数Giou、Ciou对算法带来的影响,根据对比实验结果找出最优的激活函数和损失函数,再优化算法中的分类损失以及引入经过K-means聚类后的锚框,并得到最终的优化结果。

4 实验结果分析

对于样本分类的情况一共有4种:一是正样本,被分类到正样本的为True Positive,简称TP;二是正样本,被分类到负样本的为False Positive,简称FN;三是负样本,被分类到正样本的为False Positive,简称FP;四是负样本,被分类到负样本的为True Positive,简称TN。评判模型的好坏有两个常见的指标,精确率(Precision)和召回率(Recall)。利用精确率和召回率可以画出Precision-Recall曲线。P-R曲线下的面积即为平均精确度(Average Precision),简称AP。AP值越大,其分类效果越好。mAP是对多个类别的AP求平均值。

表1 改变激活函数和损失函数的实验结果

激活函数	损失函数	Loss / mAP
LeakyRelu	原损失函数	3.16 / 77.44%
LeakyRelu	Giou	1.40 / 77.67%
LeakyRelu	Ciou	1.41 / 78.11%
Swish	原损失函数	3.08 / 77.58%
Swish	Giou	1.30 / 77.75%
Swish	Ciou	1.33 / 78.20%
Mish	原损失函数	3.04 / 78.09%
Mish	Giou	1.29 / 78.32%
Mish	Ciou	1.30 / 78.67%

表1为改变了激活函数和损失函数得到的优化结果。在采用相同损失函数的情况下,激活函数为Swish和Mish时,算法的损失都下降了0.1左右,激活函数为Swish时,算法精度有小幅度的提升,当使用Mish函数时,算法的精度能提升0.6%左右,明显优于Swish函数。将Giou和Ciou作为边框回归函数可以降低平均2.76的损失,对于损失的优化明显,且Giou和Ciou对损失的优化程度都是一样的,对于精度而言,Giou能够提升0.2%左右,Ciou能够提升0.6%左右。最终在使用Ciou和Mish的情况下能够达到78.67% mAP,提升了1.2%。

表2为采用经过K-means聚类后的锚框的结果,本文一共做了两组对比实验,从两组数据可以看出,使用针对数据集合理的锚框会对算法的精度带来一定的提升。

表2 使用K-means聚类锚框结果

激活函数/损失函数	是否使用K-means聚类算法	mAP/%
Swish/Ciou	是	78.25
Mish/Ciou	是	78.71

表3 使用Focal loss实验结果

激活函数/损失函数	K-means	Focalloss	mAP/%
Mish/Ciou	否	否	78.67
Mish/Ciou	是	否	78.71
Mish/Ciou	是	是	79.63

表3为融入Focal loss之后的实验结果,融入Focal loss能给算法带来0.92%的精度提升,提升比较明显,可见Focal loss对于解决样本不平衡问题的有效性。

5 结语

本文针对YOLOv3模型精度不够的问题展开了研究,首先是研究激活函数和损失函数,分别采用Swish和Mish函数对激活函数进行改进,模型的精度获得了不错的提升,同时也体现这两种激活函数的性能差异。损失函数中的边框损失函数采用了Giou和Ciou分别进行优化,使得模型能够对目标进行更好的定位。为了解决训练过程中样本的不平衡的问题,还使用了Focal loss函数来优化分类损失,并采用K-means算法重新对训练集的锚框进行聚类。在训练时,采用数据增强和冻结特征提取网络的方法,最终在Pascal VOC数据集上的测试结果最高能够达到79.63%,相比于原来的精度提升了2.19%。

参考文献:

- [1] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [2] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [C] // Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington DC: IEEE Computer Society, 2017: 7263-7271.
- [3] REDMON J, FARHADI A. YOLOv3: an incremental improvement [C] // The IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2018: 1068-1076.
- [4] MISRA D. Mish: a self regularized non-monotonic neural activation function [EB/OL] (2019-10-2). <https://arxiv.org/abs/1908.08681>.
- [5] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: a metric and a loss for bounding box regression [C] // 2019 IEEE CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2019.
- [6] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, PP (99): 2999-3007.

作者简介:

刘博(1996—),男,重庆人,硕士研究生,研究方向为深度学习

李艳武(1986—),男,湖北天门人,博士,高级工程师,硕士生导师,研究方向为智能优化算法、深度学习

收稿日期: 2022-03-16 修稿日期: 2022-04-05

Algorithm Research Based on Improved YOLOv3

Liu Bo, Li Yanwu

(Department of Electronics and Information Engineering, Chongqing Three Gorges University, Chongqing 404100)

Abstract: In the practical application field of target detection, the algorithm is required to achieve high-precision detection. Aiming at the problem of low detection accuracy of YOLOv3 algorithm, in order to further improve the accuracy of YOLOv3 algorithm, the activation function is changed to Mish and Swish, meanwhile Giou, Ciou and Focal loss are integrated into the loss calculation method, and k-means is used to re-cluster the data set to get new anchors, using data enhancement and freezing network parameters method during the training which resulted in a 2% improvement on the PascalVOC data set.

Keywords: target detection; YOLOv3; activation function; loss function; K-means

文章编号: 1007-1423(2022)16-0057-04

DOI: 10.3969/j.issn.1007-1423.2022.16.011

5G 网络 NSA 组网共模抗干扰问题的研究

陈雪娟¹, 叶利丽¹, 楚高峰²

(1. 乐山职业技术学院电子信息工程系, 乐山 614000; 2. 华为技术有限公司, 成都 610000)

摘要: 近年来, 5G 网络部署进行得如火如荼, NSA 组网作为 5G 部署的一种过渡性方案, 最终会被 SA 组网替代。但是从节省初期投入的角度来看, NSA 组网不可或缺, 甚至将长期与 SA 和 LTE 共存。AAU 共模并发场景作为 NSA 组网方式的一种典型应用, 必然造成 NR 与 LTE 间的小区干扰。经过分析与仿真研究, 本文提出一种通过配置合理的功率配置比来降低干扰, 提升 NSA 组网小区信号性能的方法。

关键词: 5G; NSA; LTE; AAU; LNR; NR

基金项目: 乐山市科技局(21ZRK197)

0 引言

近年来, 我国在移动通信领域的发展速度让世界惊叹, 短短 40 年时间就实现了从 1G 到 5G 的跃迁, 几乎每隔十年产生一代新的移动通信技术。我国的 5G 从 2019 年 11 月正式投入商用后, 广泛应用于广大群众的日常生活中, 预测 5G 带来的直接和间接经济产出将不断增长。目前国内外关于 5G 的部署进行得如火如荼, 各大运营商经过近三年时间的 5G 建设, 很多地区的用户已经可以享受到 5G 带来的全新体验。

5G 部署主要采用 SA (Standalone) 独立组网或者 NSA (Non-Standalone) 非独立组网方式。2019 年 5G 部署之初, 通常采用的是 NSA 非独立组网方式, 主要覆盖一线城市以及部分重点地区, 偏远地区覆盖较少, 实现国内全覆盖需要一段时间^[1]。2020 年 1 月工信部部长表示, 在 2020 年国家将 5G 网络建设的重点放在独立组网基础上, 只有这种独立组网才能够充分显示出 5G 的整体性能。SA 组网相比 NSA 组网投入高、工程量大、周期长, 在相当长的一段时间内必将是二者共存的场景。

1 5G 技术目标

3GPP 对 5G 网络目标进行了定义, 如图 1 所

示, 峰值数据包速率上行 10 Gbps、下行 20 Gbps, 用户终端业务速率上行 50 Mbps、下行 100 Mbps, 移动性满足 500 km/h, 用户终端时延 4 ms。URLLC 场景时延低于 1 ms, 为实现无人驾驶提供重大潜力, 区域流量的容量达到每平米 100 Mbps^[2]。5G 主要实现大带宽、低时延、空口灵活、UCNC、MASSIVE-MIMO 等关键技术。这些关键指标为大规模物联网业务、低时延高可靠业务以及超高清移动宽带业务等应用提供服务与支持。此类业务也必然在生产生活中逐步产生, 并非一蹴而就, 因此需要稳步推进, 逐渐完成完整的 SA 独立 5G 组网。

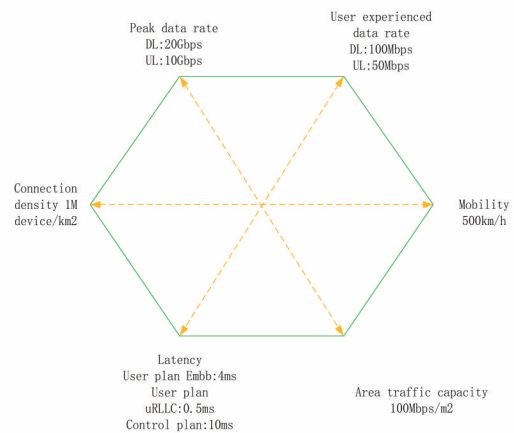


图 1 3GPP 对 5G 目标的定义

2 NSA组网研究

2.1 NSA组网分析

5G部署初期主要采用NSA非独立组网方式，如图2所示。无需建设5G新核心网，借助4G无线空口，利用现有的4G基础设施。5G终端通过NR基站进行数据分流，控制面和用户面数据通过LTE基站向EPC转发，实现快速大面积覆盖，快速提升5G网络普及速度^[3]。初期建设5G核心网5GC暂不需投入，减小网络短期设备更换成本，这种方案在节省5G初期投入成本的同时能让用户得到5G的快速体验。LTE与NR共模并发称为LNR，能够最大程度地提高资源利用率，不仅能满足5G用户的需求，而且能避免前期运营商的巨大投入，解决回报率低的问题^[4]。到目前为止，NSA组网仍然是5G部署的首要之选，非独立组网虽然只是5G商用的一种过渡性方案，但不可或缺，甚至举足轻重。

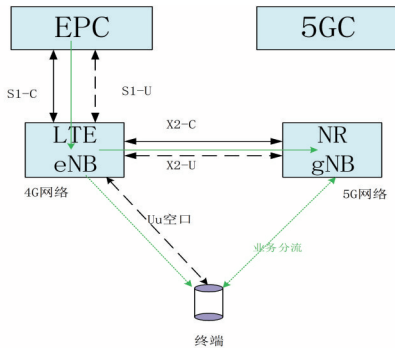


图2 NSA组网模型

2.2 LNR场景的共模建设分析

基于NSA进行组网是各大运营商进行5G部署前期使用最多的组网方式，能在最大程度上解决前期资金投入大的问题，4G与5G的LNR共模并发场景实物组网如图3所示。在原有LTE基站上增加5G主控和基带单板，AAU作为新增的5G有源天线处理单元增加了天线阵子，能够同时处理4G基带单板和5G基带单板传输的数据。RRU作为原本LTE网络的射频单元，仍然处理4G基带单板数据。

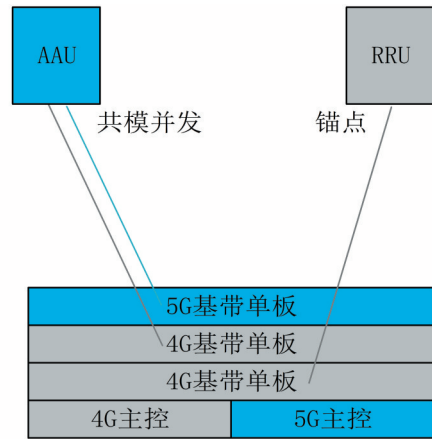


图3 LNR共模并发实物组网示意图

LNR共模并发场景能够在提升资源利用率的同时优化资源，目前4G与5G都利用有限的频谱资源，就频谱资源利用率来说，重叠的频谱资源越多，资源利用率越高。但是基于射频模块的器件性能限制，频谱资源不能无限重叠，需要规划频谱重叠的范围。以中移动2.6G频段来看，2515~2615 MHz的100 MHz用于部署NR，2615~2675 MHz的60 MHz用于部署LTE，当前协议天然支持的频谱资源允许4G与5G小区重叠20 K的频谱资源，如图4所示，重叠越多，干扰越大^[5]。在规划的NR 100 M频谱范围中，运营商可以根据小区实际情况灵活进行LTE部署或者NR部署。4G和5G共模必然带来小区间干扰，需要调整合理的功率控制比来降低干扰影响。

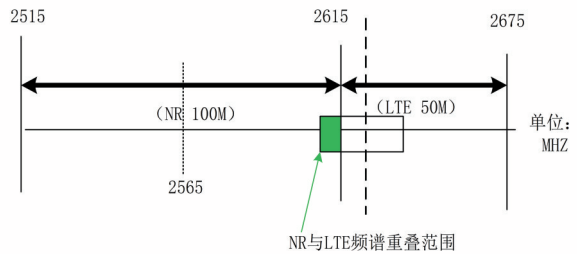


图4 NR与LTE频谱重叠示意图

2.3 功率配置比

将NR新建的天线连接到LTE基站，不仅提升了5G资源的利用率，又提高了4G的覆盖率，实现4G与5G共模。但是，在相同的天线上同

时传输4G与5G信号必然会带来小区间的相互干扰。5G的AAU目前使用较多的是64天线阵,后续也会扩展到128天线阵和256天线阵。在LTE与5G共用频谱的过程当中,必然带来同频干扰。

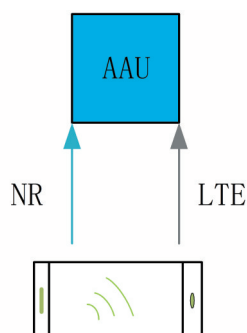


图5 LNR与LTE上行测试场景

LNR共模的AAU场景中,LTE与NR共用同一块射频模块,如图5所示。以2.6G频段TDD上行链路为例,LTE与NR共160 MHz带宽,NR最大带宽为100 M,最小带宽20 M,LTE最大带宽为20 MHz,可以配置的最大带宽为1NR+3LTE。AAU场景固定NR小区带宽为100 MHz,LTE小区带宽为20 MHz,通过调整NR和LTE小区的功率配置比来达到减小干扰的目的,功率配置比参考公式(1):

$$P_{NR/LTE} = \frac{P_{NR}(\max)/W_{NR}}{P_{LTE}(\max)/W_{LTE}} \quad (1)$$

公式(1)将NR小区最大功率与小区带宽的比值除以LTE小区最大功率与小区带宽的比值,所得数据为5G与4G共模的功率配置比。仿真针对最大NR小区带宽100 MHz与最大LTE小区带宽值20 MHz进行,将NR小区功率与LTE小区功率的比值作为仿真中的功率比。通过MATLAB进行仿真,观察误码率情况,如图6所示。

由图6可知,在LNR的AAU共模场景中,5G与4G小区带宽分别为各自最大值100 MHz和20 MHz时,AAU端设定5G与4G功率比值约为6,小区间干扰最低,误码率最小。 P_{NR}/P_{LTE} 比值的减小或增加都将引起误码率的指数式增长,同时发射功率的设定也必须考虑小区带宽值。因此AAU共模并发场景设置合理的小区功率配置比对于获得更优的小区移动信号性能具有不可忽视的重要作用。

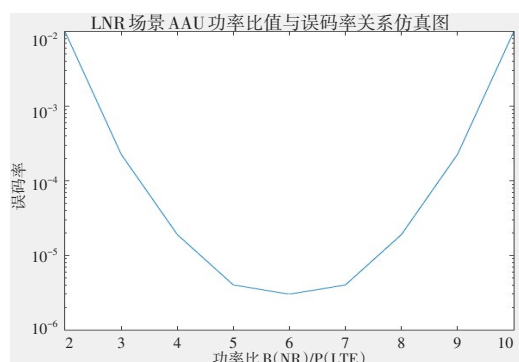


图6 LNR场景5G与4G功率比同误码率之间的仿真关系图

3 结语

5G部署中,NSA组网作为一种过渡性技术不可或缺,甚至举足轻重。LNR共模并发场景作为NSA组网中的一种典型应用,5G与4G信号使用邻近的有限频率资源通过AAU同时发射,必然带来严重的小区间干扰。为了降低相邻小区间干扰带来的影响,通过设置合理的NR与LTE功率配置比能够起到事半功倍的作用,同理功率配置比不合理将严重影响移动信号的性能。

参考文献:

- [1] 周彦. NSA与SA共存网络架构能力对5G平滑演进的必要性研究[J]. 移动通信,2019(1):46-51.
- [2] 3GPP TS 37.340. NR; Multi-connectivity; Overall description; Stage 2[S],2018.
- [3] 张炎炎,李杏,刘娜,等. 4G/5G协同组网关键技术研究[J]. 移动通信,2020,44(5):21-26.
- [4] 李德忠,白波,林琳. 中国移动4G与5G协同组网方案研究[C]//5G网络创新研讨会(2019)论文集. 北京:TD产业联盟,《移动通信》杂志社,2019:176-179.
- [5] 张建国,徐恩,肖清华. 5G NR频率配置方法[J]. 移动通信,2019,43(2):33-37.

作者简介:

陈雪娟(1985—),女,四川乐山人,硕士,讲师,研究方向为移动通信与通信网络领域

叶利丽(1987—),女,四川自贡人,硕士,讲师,研究方向为移动通信领域与学生管理领域

楚高峰(1986—),男,四川成都人,硕士,工程师,研究方向为数据通信网络领域

收稿日期:2022-03-14

修稿日期:2022-05-07

图形图像

文章编号: 1007-1423(2022)16-0060-07

DOI: 10.3969/j.issn.1007-1423.2022.16.012

基于改进 U-Net 模型和 CBAM 的脑肿瘤 MRI 图像分割

张晓倩, 罗建, 杨梅, 金芊芊, 朱熹

(西华师范大学电子信息工程学院, 南充 637009)

摘要: 由于脑肿瘤图像分割常出现缺少上下文信息和分割效率低的情况, 本文采用对 U-Net 模型改进的方法来解决此问题。首先, 在 U-Net 中引入改进的残差块, 其作用是对网络进行深化, 便于提取更多的特征信息; 然后通过加入 CBAM 机制的方法, 使得网络更加关注某些特征层和空间区域, 抑制非病灶区域的特征, 以提高肿瘤分割的准确性。本文利用 MICCAI 提供的公开磁共振成像 MRI 数据集对改进模型进行验证, 并利用 Dice 系数和精准率对模型进行评估, 在整体、核心和增强肿瘤区两者分别达到 0.883、0.80、0.789 和 0.899、0.845、0.804。结果显示, 修改后的模型对脑肿瘤图像的分割精度和准确率均有所提高。

关键词: 深度学习; 脑肿瘤分割; U-Net; CBAM 注意力机制; 深度残差模块

基金项目: 四川省教育厅重点项目(14ZA0123); 西华师范大学英才科研基金项目(17YC157)

0 引言

社会在进步, 计算机技术也在发展, 现已能为大脑研究提供更丰富的脑组织图像和有效信息。磁共振成像 MRI 也成为肿瘤分割的一个常用手段。它具有多参数成像, 肿瘤位置更加敏感, 脑组织成像更清晰等特征。虽然 MRI 能很好地显示脑组织结构和病变区域, 但 MRI 图像存在伪影、偏场效应、灰度偏移场等问题。由于人工筛选 MRI 图像来预判早期脑肿瘤的方法非常耗时, 所以引入深度学习技术是非常有必要的。

目前常用方法分为两大类, 即传统和深度学习的图像方法。

阈值分割和区域生长法在传统方法中占很大比重, 虽然它在脑肿瘤的治疗中取得了一定成果, 但其图像处理算法的整体效果不是很好。这类算法常依赖于人为操作, 因此存在一定的限制, 需要根据实验的差异来设置约束值。实验得到的结果往往不是最精确的, 还需要对其进行优化, 这种情况就会导致实验的鲁棒性差, 脑肿瘤分割效率低^[1]。

近年来, 关于肿瘤分割的深度学习方法有很多。Zhang 等^[2]采用两次分割 MR 图像的方法,

先采用 FCN 模型粗分割训练, 来检测肿瘤的整体区域, 然后再用 FCN 模型进行二次精确分割, 获得脑肿瘤的内部结构。但这会导致脑肿瘤图像局部信息缺失, 分割精度也偏低。Ronneberger 等^[3]提出的对称网络 U-Net, 对医学图像分割适应能力较好, 成为医学图像的常用网络模型, 本文模型基于此进行改进。Vittikop 等^[4]在 U-Net 网络基础上加入了跳跃连接, 将深层次特征信息进行融合, 这使得脑肿瘤图像能很好地弥补缺失的浅层信息, 取得了较好的效果。

由于跳跃连接在 U-Net 模型中能够提高分割准确率, 本文基于 U-Net 模型的同时保留跳跃连接来融合特征信息, 并对模型进行了改进。首先, 将深度残差块替换 U-Net 结构中的卷积块, 弥补丢失的特征信息, 避免梯度消失等问题; 其次, 利用 Dy-ReLU 替换深度残差模块中的 ReLU 激活函数, 提高网络模型的非线性表达能力, 解决梯度退化的问题; 但 U-Net 中通过用跳跃连接来进行简单的图像拼接, 容易丢失上下文特征信息, 影响分割的准确性, 对此, 本文引入 CBAM 注意力机制, 从空间和通道两方面来提取需加以关注的某些特征, 提高分割的准确性。

1 本文方法介绍

1.1 网络模型

由于深度残差块在深度学习图像处理中可以避免梯度消失的问题,同时对浅层网络的分割效果有所提升,所以本文基于深度残差块进行改进并替换掉U-Net基础网络中的卷积块。并将CBAM机制引入U-Net,使关键信息在空间

和通道上都聚焦,提高脑肿瘤分割的效果。

图1所示为本文的网络结构,由多部分组成,包括桥接、跳跃连接、编码、分类器和解码。其中编码和解码区域由四个改进的残差块组成,桥接区域由一个改进的残差块组成。一个改进的深度残差块由批标准化BN、动态激活函数Dy-ReLU、3×3的卷积层和恒等映射部分组成,这部分内容将在后面详细介绍。

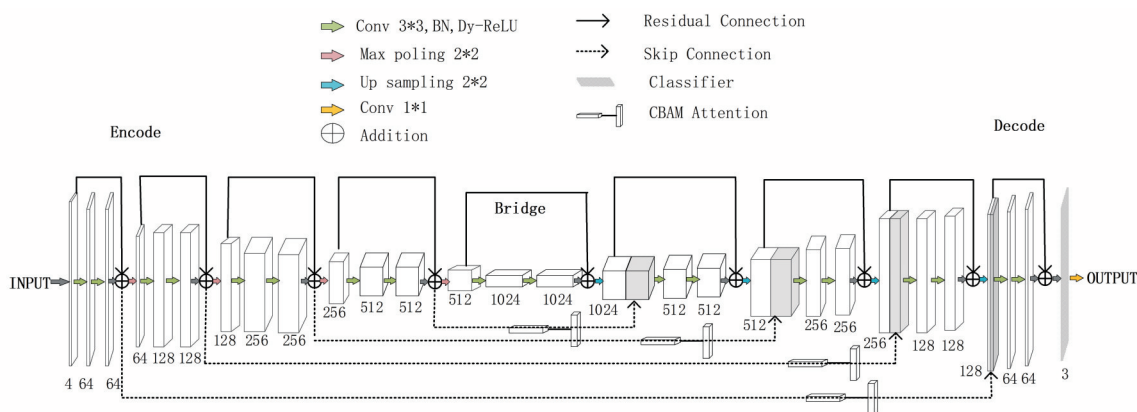


图1 网络结构

编码区域主要由改进的深度残差块和下采样组成,下采样主要采用全局最大池化,其中一共进行了四次操作,每经过一次深度残差块,就将进行一次全局最大池化。经过深度残差块和下采样操作的时候,当图尺寸减小,其相应的通道数将增加。

桥接在网络模型中起着必不可少的作用,主要是用于连接编码和解码部分。解码区域主要由改进的深度残差块和上采样组成,其中一共进行了四次操作。上采样后通道数将会减少一半,图像尺寸扩大一倍。最后获得与输入的特征图像大小近似的图。分类器采用常用的方法,由1×1卷积和Sigmoid搭配而成。1×1卷积主要是用于降低计算量。最后通过Sigmoid对特征图进行映射来显示像素的类别。

跳跃连接实现特征映射的融合的方法是将编码的深、浅层特征级联。但由于编码区域提

取到的特征信息效果很差,带来了大量的冗余特征信息。

针对编码区域提取了大部分的冗余信息,本文在深浅层特征融合之前,引入CBAM机制,从通道和空间两个方面都抑制冗余区域,提高特征提取的效率。

1.2 Dy-ReLU激活函数

激活函数通过引入非线性因素,使得模型具有非线性映射的能力。ReLU激活函数在分割中使用得非常频繁,然而它不会根据实验数据变化而变化,无区别对待所有的输入样本。但动态的激活函数则可以解决这个问题,本文通过引入动态校正单元Dynamic Rectified Linear Unit, Dy-ReLU^[5]来提高网络模型的非线性表达能力。对于输入的全局上下文信息,它通过辅助函数对其进行编码,对后续的分段线性激活函数具有指导性作用。

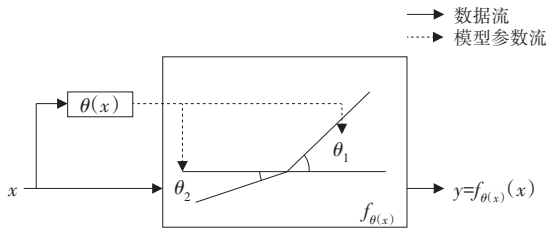


图2 Dy-ReLU示意图

Dy-ReLU示意图如图2所示，将 x 分别流向 $\theta(x)$ 和 $f_{\theta(x)}(x)$ ， $\theta(x)$ 是实验中获得的参数，为了使其对激活函数具有自适应性，采用 x 对上下文信息进行编码， $f_{\theta(x)}(x)$ 输出的是激活函数的值。公式如下：

$$y_c = f_{\theta(x)}(x) = \max_{1 \leq k \leq K} \{a_c^k x_c + b_c^k\}$$

式中 k 为函数数目， c 是通道数， x_c 关于第 C 个通道的输入向量， a_c^k 、 b_c^k 是为辅助函数 $\theta(x)$ 计算得到的参数，表示如下：

$$\theta(x) = [a_1^1, \dots, a_1^c, \dots, a_1^K, \dots, a_c^1, \dots, a_c^c, \dots, a_c^K, \dots, b_1^1, \dots, b_1^c, \dots, b_1^K, \dots, b_c^1, \dots, b_c^c, \dots, b_c^K]^T$$

$\theta(x)$ 可以看作是由类似SE模块的超函数实现，主要是基于注意力机制来获取关于特征信息之间的重要性，随后通过重要度对有用信息和无用信息进行针对性的提取，最后将超参数映射到特征通道中^[6]。

x 先后经过全局池化层和全连接层，其中全连接层经过了两次，随后使用ReLU函数引入更多的非线性因素，让信道间的复杂情况拟合得更好。最后采用Sigmoid标准化输出^[7]。当SE模块中的计算完成后，最后的输出为：

$$a_c^k(x) = \alpha^k + \lambda_a \Delta a_c^k(x), b_c^k(x) = \beta^k + \lambda_b \Delta b_c^k(x)$$

其中， α^k 、 β^k 分别为 a_c^k 和 b_c^k 的开始值， λ_a 、 λ_b 和 λ_c 为控制范围的标量，即每个特征通道所加的权值。

1.3 深度残差模块

在深度学习图像处理中，传统观念认为网络进行更深层次的设计，效果会更好。但这也随之带来梯度消失的问题，而且浅层网络对分割效果的提升也不够明显。基于以上问题，He等^[8]提出一种残差网络。

图3是改进的残差块示意图，其中将Relu

替换为Dy-ReLU激活函数，并对上下文信息进行动态编码，以收集更有利的特征信息。假设输入的是 x ，那么期望输出则为 $H(x)$ ， $H(x)$ 通过将恒等映射添加到浅层网络中，可以表示为 $y = xy = x$ ，然后直接越过中间层，传输到后面的网络层。网络是通过快捷连接来叠加输入输出的，这不仅可以使网络的训练简化，还可以使信息迅速地传递。此时，需要学习的函数变成 $F(x) = H(x) - x$ ，即输入 x 与输出 $H(x)$ 之差，称为残差结构。

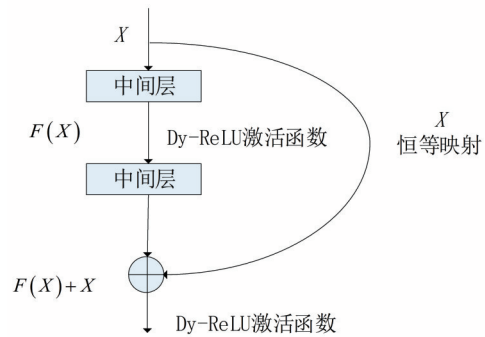


图3 改进的残差块

残差公式如下：

$$y = w_s x + F(x, \{w_i\})$$

x 为输入， y 为输出， w_s 为卷积操作， $F(x, \{w_i\})$ 是残差映射，由网络学习得到。当中间部分由两层网络组成时， $F(x, \{w_i\}) = w_2 \sigma(w_1 x) F(x, \{w_i\}) = w_2 \sigma(w_1 x)$ 。当图的维度出现不同时，零填充方式对于增加维度是非常有效的，并利用 1×1 卷积来保证维数的一致性。

如图4(a)所示是原始的残差块，通常在网络层数较大时，提取的特征信息较好。本文结构如图4(b)所示，是在残差块的基础上，用Dy-ReLU替换ReLU激活函数。为了解决训练中梯度退化的问题，采用恒等映射单元，将输入部分传递到后面的网络层，以促进特征信息的传播，很大程度上可以解决浅层特征在训练中丢失的问题。

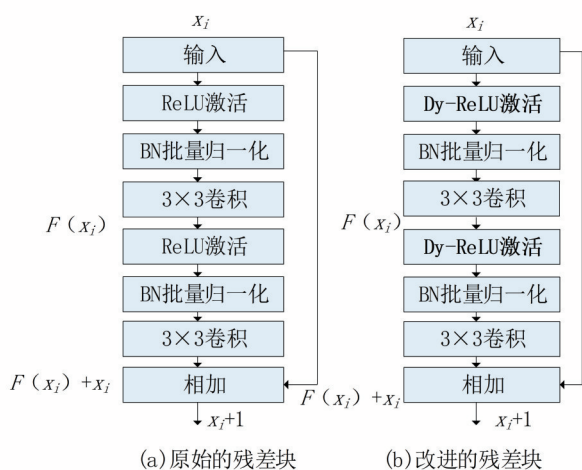


图 4 原始的与改进的残差模块

1.4 注意力 CBAM 模块

Woo 等^[9]首次提出了 CBAM 双重注意力机制, 它是在空间和通道都关注某些特征的注意力模块, 双管齐下, 在提取特征时着重关注一些特征层和空间区域, 增强了特征图中有用的特征, 在图像分割中取得更好的效果。

图 5 即为 CBAM 机制示意图, 特征图先后经过通道注意力机制和空间注意力机制, 为了获得丰富的高层次特征, 双重注意力机制运用全局平均池化和全局最大池化来提取信息。然后, 通过 MLP Multilayer Perceptron 将通道数压缩到 $\frac{C}{r}$, 再扩张回 C , 其中 C 为通道数, r 为衰减比率, 且 $r=8$ 。再将处理过的特征通过 sigmoid 函数, 得到通道注意力模块的最终输出, 随后则对输出中的元素对应相乘^[10]。计算过程如下:

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) = \sigma(W_1(W_0(F_{\text{avg}}^c)) + W_1(W_0(F_{\text{max}}^c)))$$

式中 F 代表输入; σ 常用于表示 sigmoid 激活

函数。

空间注意力机制主要从空间层面探讨特征图的关系来突出空间信息的重要性, 其与通道注意力机制相辅相成。CBAM 机制采用的卷积核大小为 7×7 , 经过空间注意力模块后, 得到最终的特征图。计算过程如下:

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

其中, 7×7 卷积核用 $f^{7 \times 7}$ 表示。

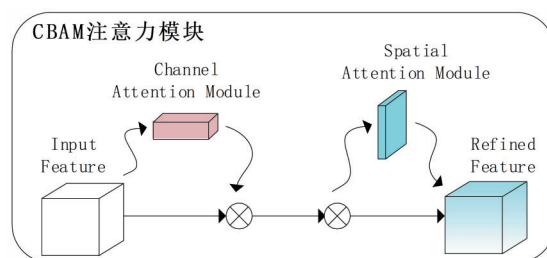


图 5 CBAM 机制示意图

2 实验及结果分析

2.1 数据集

实验使用的是 MICCAI 提供的公开脑肿瘤数据集 BraTS2019 和 BraTS2020^[11], BraTS2019 包括 335 例患者, 这部分数据用于训练。将 BraTS2020 相对于 2019 新增的数据用于测试, 包括高级别胶质瘤 34 例。如图 6 所示, 每个患者的 MR 图像都有 4 种模态 T1, T1ce, Flair 和 T2 和真实分割标签。每一个数据都包含四个方面: 非增强肿瘤和坏死肿瘤、水肿区域、增强肿瘤及背景, 分割中通常将这些不同的区域分为三个部分: ①整体肿瘤, 去除背景的所有部分; ②肿瘤核心, 由坏死、非增强和增强肿瘤部分构成; ③增强肿瘤, 主要由增强肿瘤部分组成。

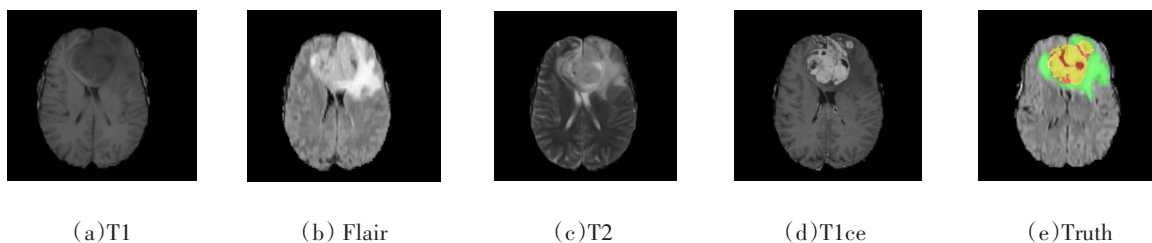


图 6 脑肿瘤四种模态

本文采用Z-score方法对数据进行标准化操作，先计算均值和标准差，再对所有图像采取单位标准化和零均值操作，最后对图像随机切割为160×160的大小。

2.2 损失函数

由于混合损失函数在医学图像中使用效果较好，本文的损失函数由Dice相似系数和交叉熵损失组合而成。交叉熵公式如下：

$$\text{Loss}_{CE} = - \sum_{i=1}^N \sum_{j=1}^L \left[(g_{ij} \log p_{ij}) + (1 - g_{ij}) \log (1 - p_{ij}) \right]$$

式中， L 、 N 分别是真实标签与预测图中像素集合， g_{ij} 代表真实类别， g_{ij} 指第*i*个预测图与第*j*个真实标签的类别， p_{ij} 代表预测值，指第*i*个预测图与第*j*个真实标签中的预测值。在训练的时候，交叉熵通常用于优化，尽管可有效解决梯度消失的问题，但对于类不平衡的图像，它往往有所偏重，尤其对类别数较多的样本会偏向更多，这将使得网络的优化不够好^[12]。

$$\text{Loss}_{Dice} = 1 - \frac{2 \left(\sum_{i=1}^N \sum_{j=1}^L g_{ij} p_{ij} \right) + \varepsilon}{\sum_{i=1}^N \sum_{j=1}^L g_{ij} + \sum_{i=1}^N \sum_{j=1}^L p_{ij} + \varepsilon}$$

对*N*、*L*分别进行计算，其中光滑算子用ε表示，主要用于避免分母为0的情况。Dice常用于学习网络参数，使预测值更接近真实值。混合损失函数的公式如下：

$$\text{Loss} = \text{Loss}_{CE} + \text{Loss}_{Dice}$$

2.3 评价指标

Dice系数、HD距离和查准率常用作图像分割的评价指标，本文也采取相同的策略，公式如下：

$$\text{DSC} = \frac{2TP}{FP + 2TP + FN}$$

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

$$\text{HD} = \max \left[d_{XY}, d_{YX} \right]$$

$$= \max \left\{ \max_{x \in X} \min_{y \in Y} \|x - y\|, \max_{y \in Y} \min_{x \in X} \|y - x\| \right\}$$

公式中，*TP*指预测为正确的正样本个数，*FP*指预测为正确实际是错误的负样本个数，*FN*指预测为错误实际是正确的正样本个数。Dice系数是用于衡量相似性的，用于计算预测值与

真实值的相似度；查准率是指正确预测总数占被预测总数的比例；*Y*表示预测值，*X*代表真实值， d_{XY} 与 d_{YX} 分别表示预测值与真实值的HD距离，HD表示预测与真实值的不匹配的最大程度，是 d_{XY} 和 d_{YX} 中的最大值，这表明值越小，分割越准确。

2.4 参数设置及训练

硬件环境：使用AMD EPYC 7302 CPU，内存大小为64 G，使用NVIDIA GeForce RTX3090 GPU，显卡24 G。软件环境：Pytorch深度学习框架，CUDNN8.0，CUDA11.0，Ubuntu 18.04.5 LTS，Python3.8。在训练时同时输入预处理后的图像和真实标签。25%、75%的数据集分别作为验证集和训练集，BraTS2020中新增部分用于测试集。训练过程中，学习率初始值为0.0003，动量0.9，迭代批量为32，训练迭代期为10000。用衰减系数为0.0001的Adma优化器优化。将早停法的值设置为20来防止过拟合。

图7所示为模型训练过程中损失函数变化的曲线图。连续实线是训练的损失，虚线是验证的损失。由图7可以看出，随着迭代次数的增加，网络的准确率在不断地提高，损失在不断地减少。当值趋近260时，网络趋于稳定。

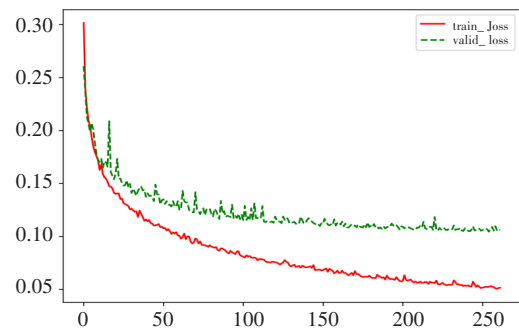


图7 训练损失图

2.5 测试结果

本文选择FCN和U-Net两种常用的基础网络进行实验对比。表1展示了三种模型的分割结果。从表1可以看出FCN的结果很不理想，主要原因是特征信息都没有得到充分利用。U-Net可以看作是基于FCN改进的，融合跳跃连接来提取更多的信息，但由于获取的浅层信息较差，存在冗余，对分割效果产生了很大的影响。本

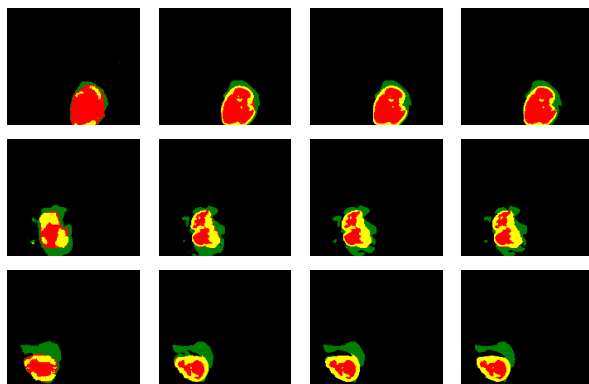
表1 模型对比结果

模型	DSC /%			Precision /%			HD		
	W_T	T_C	E_T	W_T	T_C	E_T	W_T	T_C	E_T
FCN	83.73	76.92	69.81	82.45	82.13	68.31	3.02	2.12	3.26
U-Net	87.35	80.81	77.27	87.34	84.11	77.85	2.67	2.03	2.9
本文方法	88.33	80.01	78.91	89.93	84.48	80.37	2.61	2.02	2.85

注:加粗字体指各列的最优结果

文以U-Net为基础进行改进,改进后的模型在分割中效果较好,DSC较原U-Net模型有小幅提升,查准率在整体、核心和肿瘤增强区域依次提升2.59%、0.37%、2.52%,HD在各区域也有小幅提升,得到了不错的分割结果。

脑肿瘤MR图像分割结果如图8所示。从左到右依次为FCN、U-Net、真实标签和本文方法的分割结果。从图8可以看出,FCN分割效果较差,只能大概定位脑肿瘤的位置,内部分割不够精细。U-Net分割的轮廓清晰,但边缘还是不够细化,分割不是很准确。可以看出本文能够得到与真实标签更加相进的分割结果,效果明显得到改善,性能良好。



(a)FCN (b)U-Net (d)Truth (e)本文方法

图8 对比实验分割结果图

3 结语

利用深度学习技术来分割脑肿瘤图像的方法是很重要的。本文采用改进U-Net模型的方法提高肿瘤的分割结果。利用改进的深度残差块获取特征信息,在跳转连接中引入CBAM注意机制,同时结合混合损失函数对数据进行优化,很大程度上解决了类不平衡的问题。实验

结果显示,本文的方法能较大提升分割的精确度,分割结果也趋近于真实标签,但有一定的局限性。由于本实验使用的是三维数据,采用三维数据切片后的数据进行分割会丢失一些图像信息,影响分割精度。因此,下一步研究的重点是采用三维数据进行图像分割,来弥补图像丢失的不足。

参考文献:

- [1] 庄宇舟. 基于深度学习的脑胶质瘤分割方法研究[D]. 武汉:华中科技大学,2019. DOI:10.27157/d.cnki.ghzku.2019.003553.
- [2] ZHANG C J, FANG M C, NIE H H. Brain tumor segmentation using fully convolutional networks from magnetic resonance imaging [J]. Journal of Medical Imaging and Health Informatics, 2018, 88: 1546-1553.
- [3] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation [J]. Springer International Publishing, 2015, 9351: 234-241.
- [4] VITTIKOP B S, DHOTRE S R. Automatic segmentation of MRI images for brain tumor using unet [C]//2019 1st International Conference on Advances in Information Technology ICAIT. KLS Gogte Institute of Technology Dept of CSE Engg Belagavi Karnataka India, 2019: 507-511.
- [5] CHEN Y P, DAI X Y, LIU M C, et al. Dynamic relu [C]//European Conference on Computer Vision. Springer, Cham, 2020: 351-367.
- [6] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 428: 2011-2023.
- [7] 付顺兵,王朝斌,陈旭,等. 基于Dense U-Net的眼底视网膜血管分割[J]. 内江师范学院学报, 2021, 364: 40-46. DOI: 10.13603/j.cnki.51-1621/z.2021.04.008.

- [8] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition CVPR. IEEE, 2016: 770-778.
- [9] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[M]//FERRARI V, HEBERT M, SMINCHISESCU C, et al. Computer Vision—ECCV 2018. ECCV 2018. Lecture Notes in Computer Science, vol 11211. Springer, Cham. https://doi.org/10.1007/978-3-030-01234-2_1.
- [10] 姚齐水, 别帅帅, 余江鸿, 等. 一种结合改进 Inception V2 模块和 CBAM 的轴承故障诊断方法[J]. 振动工程学报: 1-9[2022-03-19]. <http://kns.cnki.net/kcms/detail/32.1349.TB.20220115.1015.002.html>.
- [11] BAKAS S, AKBARI H, SOTIRAS A, et al. Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features[J]. Scientific Data, 2017, 41: 1-13.
- [12] 付顺兵, 王朝斌, 罗建, 等. 基于改进 U-Net 模型的脑肿瘤 MR 图像分割[J]. 西华师范大学学报(自然科学版), 2021, 422: 202-208. DOI: 10.16246/j.issn.1673-5072.2021.02.016.

作者简介:

张晓倩(1997—), 女, 四川成都人, 硕士研究生, 主要从事医学图像处理研究工作

通信作者: 罗建(1975—), 男, 重庆合川人, 博士, 副教授, 硕士生导师, 主要从事图像处理研究工作, E-mail: luojianz@gmail.com

杨梅(1996—), 女, 四川南充人, 硕士研究生, 主要从事医学图像处理研究工作

金芊芊(1998—), 女, 四川宜宾人, 硕士研究生, 主要从事图像处理研究工作

朱熹(1996—), 男, 四川南充人, 硕士研究生, 主要从事信号处理研究工作

收稿日期: 2022-03-24 修稿日期: 2022-04-26

MRI Image Segmentation of Brain Tumor Based on Improved U-Net Model and CBAM

Zhang Xiaoqian, Luo Jian, Yang Mei, Jin Qianqian, Zhu Xi

(School of Electronic Information Engineering, Xihua Normal University, Nanchong 637009)

Abstract: Due to the lack of context information and low segmentation efficiency in brain tumor image segmentation, this paper is based on a better U-Net model to solve this problem. Firstly, an improved residual block is introduced into U-Net to deepen the network and extract more feature information. Then, by adding CBAM mechanism, the network concentrates more on some feature layers and spatial regions, and suppresses the features of non-lesion areas to improve the accuracy of tumor segmentation. In this paper, the improved model was verified using the public magnetic resonance imaging MRI data set provided by MICCAI, and using the Dice coefficient and precision rate to evaluate the model. The overall, core and enhanced tumor areas reached 0.883, 0.80, 0.789 and 0.899, 0.845, 0.804, respectively. The results show that the segmentation accuracy and accuracy of the modified model for brain tumor images are improved.

Keywords: deep learning; brain tumor segmentation; U-Net; CBAM attention mechanism; depth residual module

文章编号: 1007-1423(2022)16-0067-06

DOI: 10.3969/j.issn.1007-1423.2022.16.013

一种基于空间预测的运动补偿帧率上转换算法

马文丹, 李 凡, 俞文静

(广州软件学院网络技术系, 广州 510990)

摘要: 在物联网高速发展的多媒体时代, 为减少无线信道的带宽消耗, 通常采用运动补偿帧率上转换来支持低比特率的视频通信。本文提出了一种基于空间预测算法, 用于提高运动补偿帧率上转换的性能。该算法的核心是将视频帧分割为基础块和缺失块两种类型的预测模型, 使用改进的双边运动估计计算基础块的运动矢量, 利用运动矢量场的空间相关性, 根据相邻基础块的运动矢量预测缺失块的运动矢量场。实验结果表明, 所提出的空间预测算法能够以较低计算复杂度提高插值帧的主客观质量。

关键词: 运动补偿插帧; 双边运动估计; 空间相关性

基金项目: 广州软件学院科研项目(ky202121)

0 引言

电影通常以每秒 24 帧的速度播放, 而电视节目则以 30 Hz 的标准帧率播放。特别是在多媒体物联网中, 受无线信道带宽的限制, 对视频序列进行编码时要求较低的帧率。低帧率基本可以满足人们的娱乐需求, 但当视频序列中存在大量快速运动时, 会出现运动模糊的现象。高帧率的视频序列相比低帧率视频包含更少的模糊和伪影, 能够带来更好的视觉体验。因此, 在多媒体物联网的接收端, 应提高视频帧率以使视频更加流畅、观看者的舒适度更高。为满足上述需求, 运动补偿帧率上转换(motion-compensated frame rate up-conversion, MC-FRUC)通常可以用来将低帧率视频转换为高帧率视频。

MC-FRUC 是一种在两个相邻的原始帧之间插入多个插值帧的视频处理技术, 近年来得到了学者们的广泛关注^[1-6]。由于插值算法的精度在很大程度上影响插值帧质量, 因此对运动估计(motion estimation, ME)、运动矢量平滑(Motion vector smoothing, MVS)和运动矢量映射(motion vector mapping, MVM)进行了大量研究^[7]。ME 是计算两个相邻原始帧间的运动矢量场

(motion vector field, MVF); MVS 是对得到的运动矢量(motion vector, MV)进行校正, 提高 MV 的准确度; MVM 是得到插值帧的 MVF^[8]。

块匹配算法(block matching algorithm, BMA)是各种 ME 算法中最典型的方法, 与逐像素 ME 相比具有较低的复杂度^[9-10]。一个标准块的大小远远小于视频帧的大小并且大多数目标像素分布在不同的相邻块, 由此提出了基于时空相关性的三维递归搜索算法(3D recursive search, 3DRS)^[11]。为了尽可能真实地获取 MV, MVS 对块匹配进行了平滑约束, 从而有效抑制异常的 MV^[12]。MVS 可以通过中值滤波和惩罚项来实现, 但是这种方法增加了计算复杂度^[13]。MVS 之后, 利用 MVM 从相邻原始帧的 MVF 推导出中间帧的 MVF^[14-15]。前向 MVM 是将 MV 的二分之一沿着其方向映射到目标块, 会出现少量的时域错配问题, 中间帧的某些块可能存在多个 MV 或者没有 MV, 导致重叠和空洞^[16]。根据时间对称假设, 双边 MVM 对中间帧进行双边运动估计(bilateral motion estimation, BME)^[17], 能够有效避免块伪影。然而, 由于视频序列统计偏差, 异常 MV 总是存在, 导致运动补偿插帧过程中出

现边缘模糊和伪影。现有的运动补偿插帧算法，如重叠块运动补偿(overlapped block motion compensation, OBMC)^[18]，可以减少异常MV带来的不良影响。

现有研究校正异常MV耗费大量的计算，但MV精确度的提高不尽人意^[19]。本文提出了一种基于空间预测的ME算法，此算法的核心是一种预测模型，它定义了两种类型的块：基础块和缺失块。为提高BME计算速度，计算基础块的MV时采用连续消除算法(successive elimination algorithm, SEA)^[20]，减少搜索候选并采用下采样模式进行块匹配，进而从相邻基础块的MV准确地恢复出缺失块的MV。实验结果表明，所提算法生成的上转换视频效果良好且计算复杂度较低。

1 所提基于空间预测算法

本文提出基于空间预测算法的流程图如图1所示，首先将视频帧进行块分类：基础块和缺失块，再利用SEA进行全搜索ME获取基础块的MV，然后采用空间预测方法根据基础块的运动信息实现缺失块的ME，结合两类块的MV值，使用两帧参考帧进行插值中间帧。

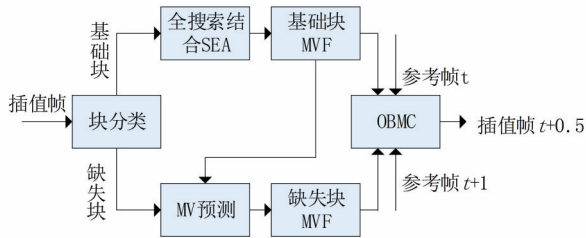


图1 基于空间预测算法流程图

1.1 块分类

空间预测技术的关键在于有效利用空间相关性，而应用该技术的前提在于块分类。空间预测技术模型如图2所示。灰色块表示基本块，其余块表示缺失块，这两种块在水平和垂直方向交替，通过BME算法精确获得基础块的MV，然后利用相邻基础块的运动信息计算出每个缺失块的MV。

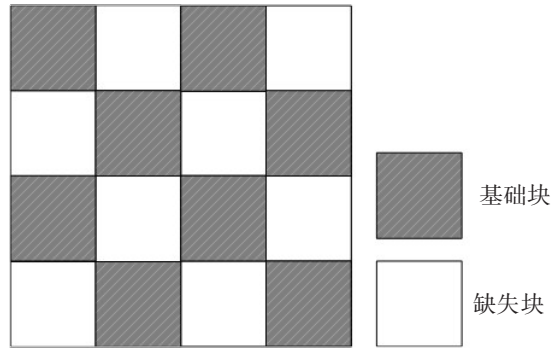


图2 块分类模型

1.2 双边ME

缺失块的MV获得由相邻基本块的MV进行预测，更好的ME算法可以保证基础块MV的精确度。全搜索ME在基础块上进行BME算法。如图3所示，全搜索ME是匹配搜索窗口内最优运动轨迹，但缺点明显。假设搜索半径为8，块尺寸为16×16，则需要搜索289个块才能找到最佳匹配块。以CIF为例，396个16×16尺寸的匹配块，使用全搜索ME意味着巨大的计算量，因此，引入SEA降低计算复杂度。

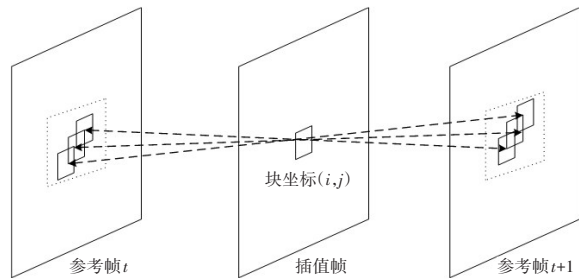


图3 全搜索ME示意图

假设插值帧 $f_{t+0.5}$ 的分辨率是 $M \times N$ ，块的大小为 $s \times s$ ，每帧包含 $M \times N / s^2$ 个标准块， M 和 N 必须能被 s 整除，每个块的左上角像素坐标 (i, j) 作为参考，第 t 帧和第 $t+1$ 帧中块 (i, j) 的亮度累加和 P 计算如下：

$$P_t(i, j) = \sum_{m=0}^{s-1} \sum_{n=0}^{s-1} f_t(i+m, j+n) \quad (1)$$

$$P_{t+1}(i, j) = \sum_{m=0}^{s-1} \sum_{n=0}^{s-1} f_{t+1}(i+m, j+n) \quad (2)$$

其中 $f_t(i+m, j+n)$ 和 $f_{t+1}(i+m, j+n)$ 分别表示 f_t 和 f_{t+1} 中 $(i+m, j+n)$ 坐标处的亮度， $(m,$

n)是块中像素坐标。

当前基础块的左上角像素设置为 $p(p = (i, j))$, 在两个参考帧中对称遍历搜索窗口内的候选块。对于参考帧 f_{t+1} , 第 n 个候选块的偏移量是 $v'_n = (x, y)$, 根据对称原则, 参考帧 f_t 的第 n 个候选块的偏移量是 $-v'_n$, 当搜索半径为 r 时, $x, y \in [-r, r]$, 设置初始偏移量为 (r, r) , 则计算当前基础块初始差值 D_0 的公式如下:

$$D_0 = \left\| \mathbf{B}_t(p - v'_0) - \mathbf{B}_{t+1}(p - v'_0) \right\|_1 \quad (3)$$

其中 $\mathbf{B}_t(i - r, j - r)$ 是 f_t 中块 $(i - r, j - r)$ 所有像素行向量, $\|\cdot\|_1$ 是向量的 l_1 范数。计算 D_0 后, v'_n 在搜索窗口内进行候选块更新, 如果满足以下不等式:

$$\left| P_t(i + x, j + y) - P_{t+1}(i + x, j + y) \right| < D_0 \quad (4)$$

计算第 n 个候选块的差值 D_n :

$$D_n = \left\| \mathbf{B}_t(p - v'_n) - \mathbf{B}_{t+1}(p - v'_n) \right\|_1 \quad (5)$$

更新当前基础块 MV v_s :

$$v_s = v'_n = (x, y) \quad (6)$$

然后更新 $D_0 = \min\{D_n, D_0\}$; 否则, v_s 保持不变, 搜索窗口内的所有候选块根据以上过程对称遍历。

1.3 MV 预测

通过全搜索 ME 获取基础块的 MV, 当前缺失块相邻四个基础块的 MV 作为候选矢量, 候选矢量集组成如下:

$$\mathbf{V}_c = \{v_1, v_2, v_3, v_4\} \quad (7)$$

当前缺失块的左上角像素坐标设为 p , 根据双向绝对差值和计算缺失块的 MV:

$$v_p = \arg \min_{v \in \mathbf{V}_c} \left\| \mathbf{B}_t(p - v) - \mathbf{B}_{t+1}(p - v) \right\|_1 \quad (8)$$

其中 $\mathbf{B}_t(p - v)$ 是 f_t 中块 $(p - v)$ 所有像素行的向量, $\|\cdot\|_1$ 是向量的 l_1 范式, v 是候选向量。

预测得到缺失块 MV, 两类块的所有 MV 形成插值帧 $f_{t+0.5}$ 的 MV $V_{t+0.5}$, 然后利用 OBMC 计算 $f_{t+0.5}$ 中 $p = (i, j)$ 处的像素值:

$$f_{t+0.5}(p) = \frac{1}{2} w(k) (f_t^k(p - v_{i,j}) + f_{t+1}^k(p - v_{i,j})) \quad (9)$$

其中 $v_{i,j}$ 是 $V_{t+0.5}$ 在 p 处的 MV, k 表示三种重叠类型: $k = 1$ 代表没有重叠部分; $k = 2$ 代表两块重叠部分; $k = 3$ 代表四块重叠部分, 根据重叠类型设置 w 值, $k=1, w=1; k=2, w=1/2; k=3, w=1/4$ 。

2 实验仿真结果

所有用于实验的测试序列为标准的 CIF (352×288) 格式和 30 帧/秒, 通过不同视频序列测试空间预测算法性能, 并与文献[21]和文献[22]ME 算法进行比较。从主客观两个角度对插值帧进行评价, 客观评价采用峰值信噪比 (peak signal-to-noise ratio, PSNR) 和结构相似度 (structural similarity, SSIM)。

2.1 主观评估

选择 Foreman 和 Stefan 视频序列进行主观视觉质量对比实验, 其中 Foreman 包含适度的场景运动, 而 Stefan 有大量的剧烈运动。图 4 和图 5 为对比算法和所提算法的插值帧, 分别以 foreman 的第 78 帧和 Stefan 的第 88 帧为例。观察插值帧效果, 可以看出对比算法人脸出现明显模糊和伪影, 所提算法对人脸部分插值效果更好, 对比算法在 Stefan 中存在大量的背景模糊和广告牌字体错位问题, 而所提算法可以减少背景模糊并提供更多的细节。总体而言, 所提算法可以提供更好的主观视觉质量。



图 4 三种算法的 Foreman 序列视觉对比

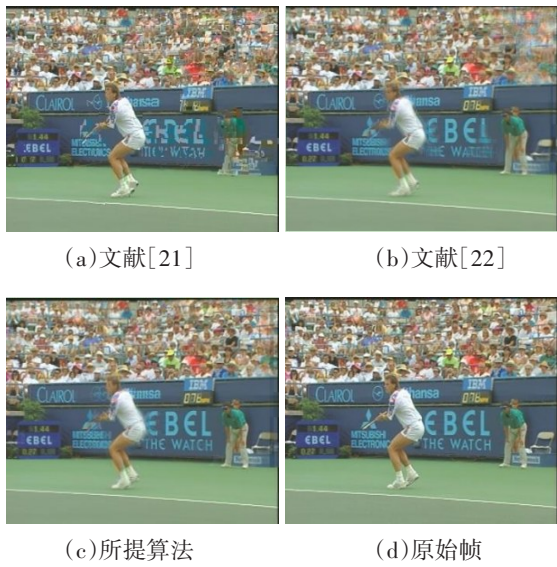


图5 三种算法的Stefan序列视觉对比

2.2 客观评估

表1显示了不同FRUC技术的PSNR值。对于9个测试视频序列，与文献[21]方法相比，本文算法显著提高了PSNR值，最大提高了2.72dB，达到了提高插值帧质量的目的。文献[22]方法可以有效地应用于相对静态或包含较少运动的视频序列中，如Foreman和Mother。但对于包含大量运动的视频，如Bus, City, Football, Mobile, Stefan，所提算法可以更精确地跟踪目标的轨迹，PSNR值比文献[22]提高了3.21dB。不同FRUC技术的SSIM值如表2所示。通过对比看出，本文提出的算法具有明显的优势，在处理静态视频方面只比文献[22]效果差。如表3所示，所提算法运行时间比文献[21]方法和文献[22]方法短，说明空间预测算法的计算复杂度较低。

表1 不同ME方法在测试序列上PSNR值比较

视频序列	文献[21]	文献[22]	本文算法
Bus	25.02	23.4	26.61
City	31.94	33.43	33.58
Football	22.78	22.43	22.77
Foreman	32.37	35.17	34.93
Flower	28.75	32.14	31.47
Mobile	25.33	27.54	27.69
Mother	40.81	43.04	42.49
Soccer	25.96	28.68	28.35
Stefan	27.03	27.55	28.56

表2 不同ME方法在测试序列上的SSIM值比较

视频序列	文献[21]	文献[22]	本文算法
Bus	0.8332	0.7771	0.9117
City	0.8645	0.9330	0.9403
Football	0.5532	0.6271	0.6426
Foreman	0.9107	0.9519	0.9587
Flower	0.9446	0.9795	0.9843
Mobile	0.8633	0.9522	0.9412
Mother	0.9617	0.9802	0.9809
Soccer	0.8661	0.8988	0.9145
Stefan	0.8333	0.8449	0.9005

表3 不同ME方法在测试序列上的平均处理时间比较

视频序列	文献[21]	文献[22]	本文算法
Bus	5.8	6.76	4.63
City	5.7	6.78	4.62
Football	6.1	6.79	4.35
Foreman	5.2	6.80	3.88
Flower	5.6	6.76	4.20
Mobile	5.5	6.80	4.39
Mother	5.0	6.80	3.98
Soccer	5.6	6.75	4.60
Stefan	5.5	6.78	4.30

综上所述，所提算法的计算资源配置更有效，结合全搜索ME和空间相关性，保证精度的同时节省计算时间。

3 结语

本文提出了一种基于空间预测的ME算法，以提高插值帧的MV精度。设计了一个预测模型，将视频帧划分为基础块和缺失块。首先用BME估计基础块的MV，然后由相邻基础块的MV推导出缺失块的MV。为了减少冗余计算，设计了双边SEA来改进BME。基于MVF的空间相关性，设计MV预测来计算缺失块的MV。实验结果表明，本文所提算法能够在较低计算复杂度下提高插值帧的客观和主观质量。

参考文献:

- [1] LU Q, XU N, FANG X. Motion-compensated frame interpolation with multiframe-based occlusion handling[J]. Journal of Display Technology, 2016, 12(1):45-54.
- [2] KAVIANI H R, SHIRANI S. Frame rate up-conversion using optical flow and patch-based recon-

- struction [J]. IEEE Transactions on Circuits & Systems for Video Technology, 2016, 26 (9) : 1581–1594.
- [3] DAR Y, BRUCKSTEIN A M. Motion-compensated coding and frame rate up-conversion: models and analysis [J]. IEEE Transactions on Image Processing, 2015, 24(7):2051–2066.
- [4] XIA M, YANG G, LI R, et al. Detecting video frame rate up-conversion based on frame-level analysis of average texture variation [J]. Multimedia Tools & Applications, 2017, 76(6):1–23.
- [5] GUO D, LU Z. Motion-compensated frame interpolation with weighted motion estimation and hierarchical vector refinement [J]. Neurocomputing, 2016, 181:76–85.
- [6] TSAI T, SHI A, HUANG K. Accurate frame rate up-conversion for advanced visual quality [J]. IEEE Transactions on Broadcasting, 2016, 62(2):626–635.
- [7] HUANG Y, CHEN F, CHIEN S. Algorithm and architecture design of multi-rate frame rate up-conversion for ultra-HD LCD systems [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2016, 27(12):2739–2752.
- [8] PAN Z, LEI J, ZHANG Y, et al. Fast motion estimation based on content property for low-complexity H. 265/HEVC encoder [J]. IEEE Transactions on Broadcasting, 2016, 62(3):675–684.
- [9] LIN Y C, TAI S C. Fast full-search block-matching algorithm for motion-compensated video compression [J]. IEEE Transactions on Communications, 2002, 45(5):527–531.
- [10] DIKBAS S, ALTUNBASAK Y. Novel true-motion estimation algorithm and its application to motion-compensated temporal frame interpolation [J]. IEEE Transactions on Image Processing, 2013, 22 (8) : 2931–2945.
- [11] HAAN G D, BIEZEN P W A C, HUIJGEN H, et al. True motion estimation with 3-D recursive search block matching [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1993, 3(5):368–379.
- [12] KIM D Y, LIM H, PARK H W. Iterative true motion estimation for motion-compensated frame interpolation [J]. IEEE Transactions on Circuits & Systems for Video Technology, 2013, 23(3):445–454.
- [13] ALPARONE L, BARNI M, BARTOLINI F, et al. Adaptively weighted vector-median filters for motion fields smoothing [C] // Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, IEEE, Atlanta, GA, USA, 1996, 2267–2270.
- [14] CHOI B T, LEE S H, KO S J. New frame rate up-conversion using bi-directional motion estimation [J]. IEEE Transactions on Consumer Electronics, 2000, 46(3):603–609.
- [15] ZHANG Y, ZHAO D, LIU H, et al. Side information generation with auto regressive model for low-delay distributed video coding [J]. Journal of Visual Communication & Image Representation, 2012, 23 (1):229–236.
- [16] JEON B-W, LEE G-I, LEE S-H, et al. Coarse-to-fine frame interpolation for frame rate up-conversion using pyramid structure [J]. IEEE Transactions on Consumer Electronics, 2003, 49(3):499–508.
- [17] CHOI B D, HAN J W, KIM C S. Motion-compensated frame interpolation using bilateral motion estimation and adaptive overlapped block motion compensation [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2007, 17 (4) : 407–416.
- [18] ORCHARD M T, SULLIVAN G J. Overlapped block motion compensation: an estimation-theoretic approach [J]. IEEE Transactions on Image Processing, 1994, 3(5):693–699.
- [19] KANG S J, CHO K R, KIM Y H. Motion compensated frame rate up-conversion using extended bilateral motion estimation [J]. IEEE Transactions on Consumer Electronics, 2007, 53(4):1759–1767.
- [20] GAO X, DUANMU C J, ZOU C. A multilevel successive elimination algorithm for block matching motion estimation [J]. IEEE Transactions on Image Processing, 2000, 9(3):501–504.
- [21] KANG S J, YOO S, KIM Y H. Dual motion estimation for frame rate up-conversion [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2010, 20(12):1909–1914.
- [22] YOO D G, KANG S J, KIM Y H. Direction-select motion estimation for motion-compensated frame rate up-conversion [J]. Journal of Display Technology, 2013, 9(10):840–850.

作者简介:

通信作者:马文丹(1992—),女,河南安阳人,硕士,研究方向为图像处理、计算机视觉等, E-mail: dandy0112@163.com

李凡(1981—),男,河南鹤壁人,硕士,研究方向为计算机视觉、网络安全

俞文静(1982—),女,湖南岳阳人,硕士,研究方向为算法优化、计算机视觉

收稿日期:2022-03-21 修稿日期:2022-05-06

A Spatial Prediction Based Motion-Compensated Frame Rate Up-Conversion

Ma Wendan, Li Fan, Yu Wenjing

(Department of Network Technology, Guangzhou Software University, Guangzhou 510990)

Abstract: In Multimedia Internet of Things (IoT), in order to reduce the bandwidth consumption of wireless channel, Motion-Compensated Frame Rate Up-Conversion (MC-FRUC) is often used to support the low-bitrate video communication. In this paper, we propose a spatial predictive algorithm which is used to improve the performance of MC-FRUC. The core of the proposed algorithm is a predictive model to split a frame into two kinds of blocks: basic blocks and absent blocks. Then an improved bilateral motion estimation is proposed to compute the Motion Vectors (MVs) of basic blocks. Finally, with the spatial correlation of Motion Vector Field (MVF), the MV of an absent block is predicted based on the MVs of its neighboring basic blocks. Experimental results show that the proposed spatial prediction algorithm can improve both the objective and the subjective quality of the interpolated frame, with a low computational complexity.

Keywords: motion-compensated frame interpolation; bilateral motion estimation; spatial correlation

(上接第 59 页)

Research on Common Mode Anti-jamming of NSA in 5G Network

Chen Xuejuan¹, Ye Lili¹, Chu Gaofeng²

(1. Department of Electronic Information Engineering, LeShan Vocational and Technical College, Leshan 614000;

2. Huawei Technology Co., Ltd, Chengdu 610000)

Absrtact: NSA as a 5G deployment of a transitional scheme, will eventually be replaced by SA network with full swing of 5G network deployment in recent years. But the NSA network is indispensable and even long-term coexistence with SA and LTE from the point of view of saving the initial investment. As a typical application of NSA, AAU common-mode concurrency will inevitably cause cell interference between NR and LTE. After analysis and simulation research, a method is proposed to reduce interference and improve the performance of NSA cell network by configuring a reasonable power allocation ratio.

Keywords: 5G; NSA; LTE; AAU; LNR; NR

实践与经验

文章编号: 1007-1423(2022)16-0073-04

DOI: 10.3969/j.issn.1007-1423.2022.16.014

C4.5 决策树方法在课程教学成绩分析中的应用

胡 佳

(南昌师范学院数学与信息科学学院, 南昌 330029)

摘要: 本文主要研究了数据挖掘中决策树技术在课程成绩模型中的应用, 运用了决策树技术 C4.5 算法在《数据结构》课程教学过程中进行数据的数据挖掘, 并得出了具体的结论, 以辅助教师在今后教学中的决策。

关键词: 数据结构; C4.5; 预处理; 成绩挖掘

0 引言

数据挖掘是一种从大量数据中提取潜在的和有价值的知识的过程。所谓的分类是把数据库的记录映射到一个已经定义好的类中的过程。实现分类的方法有很多, 包括机器学习方法、神经网络和统计方法等, 其中机器学习的决策树方法是一个比较热门的方向^[1-2], 其是通过信息论中的信息增益去寻找数据库中具有最大信息量的属性, 建立决策树的一个可以分裂的结点, 再根据这个属性字段的不同取值去建立树的分支^[1]。决策树便于企事业单位管理部门做出决策。决策树学习算法的一个最大的优点就是只要在建模过程中训练事例能够用“属性-值”的方式表达出来, 就能使用该算法来进行建模^[3-7]。

1 决策树分类的基本方法

决策树模型的内部结点是测试的属性, 叶子结点是最后测试的类别, 分支是测试属性的值。从决策树的根到叶子结点的路径就是一条分类规则。分类的过程分为两步: 建立模型和使用模型。建立模型是通过给定的训练集找到适合的映射函数 $f: D \rightarrow C$ 的表示模型, (该步也称为有指导的学习), 这里建立模型阶段也分为两步: 树的构造和树的剪枝, 其中树的剪枝包括预先剪枝和后剪枝两种方法; 使用模型是使用上一步训练完成的函数模型预测新数据的类别, 或者对数据集中的每一类别进行描述, 形成分类规则。这里

我们一般采用交叉验证法来评估模型, 也就是在训练开始前将一部分数据保留, 在训练后, 利用这部分数据对模型进行验证, 一般认为准确率在 60% 的就比较合理^[9-11]。

决策树数据挖掘的一般步骤为: 数据预处理, 采用决策树算法进行挖掘, 评价与表示。

2 C4.5 算法决策树的构造过程(用一个算法表示)

整个 C4.5 的工作流程如图 1 所示。

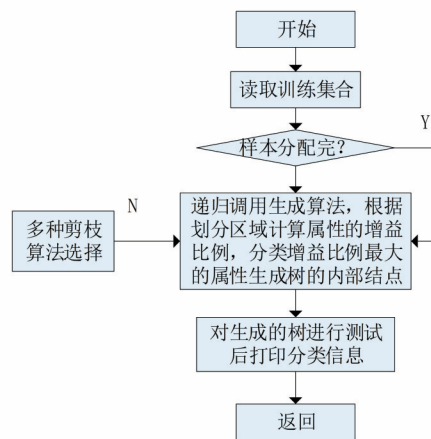


图 1 C4.5 算法流程图

3 决策树技术在数据结构成绩分析中的应用

为了配合学校的发展我们构建了数字化校

园,目前已经有了自己的信息门户平台,在这个平台中可以进行网上辅助教学,也可以完成一般的教务管理,比如在这个系统中可以输入学生成绩并生成成绩分析表,通过成绩分析表得出学生成绩的分布情况,方便老师总结学生本次考试成绩的基本情况,但是这种成绩分析仅仅停留在表面,更多的还是出于教师的教学经验判断。

因此,本文采用决策树方法中的C4.5算法,挖掘出考试成绩与影响教学的各种因素之间隐藏的内在联系。比如,教学改革的对象是计算机专业的学生,教学改革的课程是数据结构,在这门计算机专业的重要基础课程教学过程中,影响学生最终教学效果的因素有学生自身的基础、课堂教学方法和模式等,本次挖掘就是针对学生是否有畏惧情绪,先续课程是否扎实,教学的手段是否采用了多媒体,是否开辟了第二课堂等多方面因素对数据结构成绩的影响进行挖掘,得出内在的规则,这里的第二课堂主

要是采用多媒体的形式,为教学改革提供有利的理论支撑,并且在今后的教学中“对症下药”。本文首先对收集到的数据进行预处理。

3.1 数据的预处理

本文挖掘的对象为接受过数据结构教学的计算机系的大一学生,其中包括进行教改的普通班学生和未进行教改的实验班学生;挖掘的数据集包括学生的成绩数据库,以及采用调查问卷形式收集的学生相关信息情况表,将他们进行对比最后得出影响学生成绩的潜在因素以及教学改革的效果,对调查问卷中的空白项进行忽略,对学生成绩进行规范化,将数据集得出表1。表1中的畏惧情绪有两个属性:是和否;采用网上学堂的两个属性:是和否;教学方式分为两种属性:传统教学全板书和采用1/3多媒体;C语言成绩有三个属性:良好、一般和不合格;数据结构分类的两个属性:良好和一般(其中分数 ≤ 75 分为一般,分数 > 75 为良好)。

表1 课程教学以及调查信息预处理后

学号	畏惧情绪(是/否)	是否采用了网上学堂	教学方式	C语言成绩	数据结构成绩
01	否	是	1/3多媒体	良好	良好
02	否	是	1/3多媒体	一般	一般
03	是	是	1/3多媒体	良好	良好
04	否	否	板书	一般	一般
05	是	是	1/3多媒体	良好	良好
06	否	否	板书	一般	一般
07	否	是	1/3多媒体	一般	一般
08	是	否	板书	良好	一般

3.2 数据结构成绩的挖掘

本文采用C4.5算法进行数据挖掘,计算其信息增益率。训练集的样本数有350个,其中数据结构成绩良好的有152个样本,数据结构成绩一般的有198个样本。其信息增益率计算过程如下:

首先根据公式计算分类属性的信息熵, $Info(D) = -152/350 \times \log_2(152/350) - 198/350 \times \log_2(198/350) = 0.9875$, 根据公式,依次计算每个测试属性的信息量。

(1) 畏惧情绪属性中属性值为“是”的有200个,数据结构成绩良好的有94个,数据结

构成绩一般的有106;畏惧情绪属性中属性值为“否”的有150个,数据结构成绩良好的有58个,数据结构成绩一般的有92个,

$Info_{畏惧情绪}(D) = 200/350 \times (-\log_2(94/200) - \log_2(106/200)) + 150/350 \times (-\log_2(58/150) - \log_2(92/150)) = 2.0356$ 。

(2) 网上学堂教学属性中属性值为“是”的有173个,其中数据结构成绩良好的有101个,数据结构成绩一般的有62;网上学堂属性中属性值为“否”有177个,数据结构成绩良好的有51个,数据结构成绩一般的有126个,

$Info_{是否采用网上学堂}(D) = 173/350 \times (-\log_2$

$(101/173) - \log_2(62/173) + 177/350 \times (-\log_2(51/177) - \log_2(126/177)) = 2.2713$ 。

(3) 教学方法属性中属性值为“1/3多媒体”有213个, 其中数据结构成绩良好的有121个, 数据结构成绩一般的有92; 网上学堂属性中属性值为“板书”有137个, 数据结构成绩良好的有31个, 数据结构成绩一般的有106个,

$\text{Info}_{\text{教学方法}}(D) = 213/350 \times (-\log_2(121/213) - \log_2(92/213)) + 137/350 \times (-\log_2(31/137) - \log_2(106/137)) = 2.2176$ 。

(4) C语言成绩属性中属性值为“良好”的有144个, 其中数据结构成绩良好的有95个, 数据结构成绩一般的有49; C语言成绩中属性值为“一般”有206个, 数据结构成绩良好的有57个, 数据结构成绩一般的有149个,

$\text{Info}_{\text{C语言成绩}}(D) = 144/350 \times (-\log_2(95/144) - \log_2(49/144)) + 206/350 \times (-\log_2(57/206) - \log_2(149/206)) = 2.2528$ 。

因此:

$\text{Gain}(\text{畏惧情绪}) = \text{Info}(D) - \text{Info}_{\text{畏惧情绪}}(D) = 0.9875 - 2.0356 = -1.0481$;

$\text{Gain}(\text{是否采用网上学堂}) = \text{Info}(D) - \text{Info}_{\text{是否采用网上学堂}}(D) = 0.9875 - 2.2713 = -1.2838$;

$\text{Gain}(\text{教学方法}) = \text{Info}(D) - \text{Info}_{\text{教学方法}}(D) = 0.9875 - 2.2176 = -1.2301$;

$\text{Gain}(\text{C语言成绩}) = \text{Info}(D) - \text{Info}_{\text{C语言成绩}}(D) = 0.9875 - 2.2528 = -1.2653$;

$\text{SplitInfo}(\text{畏惧情绪}) = -200/350 \times \log_2(200/350) - 150/350 \times \log_2(150/350) = 0.9852$;

$\text{SplitInfo}(\text{是否采用网上学堂}) = -173/350 \times \log_2(173/350) - 177/350 \times \log_2(177/350) = 0.9999$;

$\text{SplitInfo}(\text{教学方法}) = -213/350 \times \log_2(213/350) - 137/350 \times \log_2(137/350) = 0.9657$;

$\text{SplitInfo}(\text{C语言成绩}) = -144/350 \times \log_2(144/350) - 206/350 \times \log_2(206/350) = 0.9772$;

$\text{GainRatio}(\text{畏惧情绪}) = \text{Gain}(\text{畏惧情绪}) / \text{SplitInfo}(\text{畏惧情绪}) = -1.0481 / 0.9852 = -1.0638$;

$\text{GainRatio}(\text{是否采用网上学堂}) = \text{Gain}(\text{是否采用网上学堂}) / \text{SplitInfo}(\text{是否采用网上学堂}) = -1.2838 / 0.9999 = -1.2839$;

$\text{GainRatio}(\text{教学方法}) = \text{Gain}(\text{教学方法}) / \text{SplitInfo}(\text{教学方法}) = -1.2301 / 0.9657 = -1.2738$;

$\text{GainRatio}(\text{C语言成绩}) = \text{Gain}(\text{C语言成绩}) / \text{SplitInfo}(\text{C语言成绩}) = -1.2653 / 0.9772 = -1.2948$ 。

此时, 得出的最大信息增益率是: C语言成绩, 采用网上学堂和教学方法的两个属性可以合并, 进行分裂后, 再按相同的方法进行分裂。

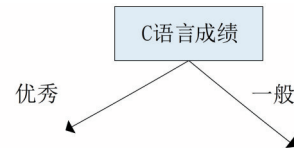


图2 第一次分裂属性

采用递归的思想计算优良分支中的教学方法和畏难情绪的信息增益率, 得到教学方法的信息率更大, 因此将优秀分支进行教学方法的划分, 以期只剩下畏难情绪属性, 可以根据样本的比率进行分配。

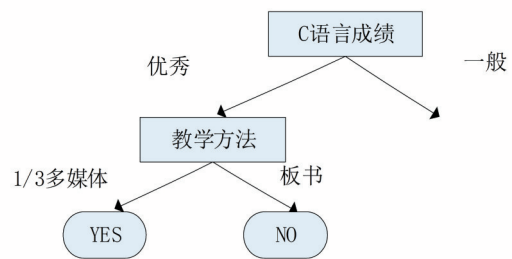


图3 再一次分裂属性

同理, 通过迭代完成分支的划分, 最终建立C4.5决策树模型, 如图4所示。

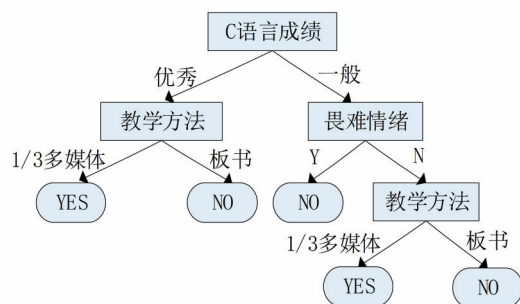


图4 最终的决策树模型

因为数据清理过程中已经对数据进行过处理, 所以这里不进行剪枝。

采用IF-THEN表示分类规则为:

IF C语言成绩优良 AND 教学方法为多媒体, 那么数据结构成绩为良好。

IF C语言成绩好AND采用板书,那么数据结构成绩为一般。

IF C语言成绩一般,AND有畏难情绪,那么数据结构成绩为一般。

IF C语言成绩一般AND没有畏难情绪AND教学方法为多媒体,那么数据结构成绩为良好。

IF C语言成绩一般AND没有畏难情绪AND教学方法采用板书,那么数据结构成绩为一般。

4 结语

本文通过预处理,然后对数据结构成绩进行挖掘,得到了不同教学方法的重要性,数据结构是一门比较抽象的计算机课程,采用一定的多媒体教学对学生学习数据结构知识是有帮助的。在今后的教学过程中,我们可以选择数据挖掘的各种方法和算法进行教学方面的挖掘,以达到辅助教师教学的决策。

参考文献:

- [1] QUINLAN J R. Induction of decision tree [J]. Machine Learning, 1986, 1(1): 81-106.
- [2] QUINLAN J R. C4.5 programs for machine learning [M]. SanMateo: Morgan Kaufmann Publishers, Inc., 1993.

- [3] 邵峰晶,于忠清.数据挖掘原理与算法[M].北京:中国水利水电出版社,2003,2:126-170.
- [4] 刘小虎,李生.决策树的优化算法[J].软件学报,1998,9(10):797-800.
- [5] 胡佳.几种典型关联规则算法的分析与比较[J].现代计算机,2011(15):15-17.
- [6] HAN J, KAMBER M. 数据挖掘概念与技术[M].范明,孟小峰,等译.北京:机械工业出版社,2007.
- [7] 邵峰晶,于忠清.数据挖掘原理与算法[M].北京:中国水利水电出版社,2003.
- [8] RUD O P. 数据挖掘实践[M].朱扬勇,左子叶,张忠平,等译.北京:机械工业出版社,2003.
- [9] 刘小虎,李生.决策树的优化算法[J].软件学报,1998,9(10):797-800.
- [10] 王晓国,黄韶坤,朱炜,等.应用C4.5算法构造客户分类决策树的方法[J].计算机工程,2003,29(14):89-91.
- [11] 栾丽华,吉根林.决策树分类技术研究[J].计算机工程,2004(9):94-96.

个人简介:

胡佳(1982—),女,江西南昌人,硕士,副教授,主要研究方向为数据挖掘、计算机教育

收稿日期:2022-04-16 修稿日期:2022-05-09

Application of C4.5 Decision Tree Method in Course Teaching Score Analysis

Hu Jia

(School of Mathematics and Information Science, Nanchang Normal University, Nanchang 330029)

Abstract: This article mainly for the data mining in the decision tree technology in the curriculum result model application, and it has used the decision tree technology C4.5 algorithm in data structure in the curriculum teaching process carries on the data the data mining, and has drawn the concrete conclusion, and it will assist the teacher in the next teaching decision-making.

Keywords: data structure; C4.5; pretreatment; achievement mining

开发案例

文章编号: 1007-1423(2022)16-0077-08

DOI: 10.3969/j.issn.1007-1423.2022.16.015

大数据可视化在城市规划中的应用

操心慧¹, 操金金², 许丽娟¹, 弓佳明¹, 赵 茜¹

(1. 广州华商学院, 广州 511300; 2. 第二航务工程局有限公司设计研究院, 武汉 430000)

摘要: 城市规划是一座城市发展的重要保障, 是一座城市一个时间段内的发展目标。而数据的大规模、多维度等特点使传统的城市规划方法难以负此重任。同时, 智慧城市和信息化建设持续迅速发展, 大数据可视化在城市规划行业的实践和研究中得到了广泛应用。本文首先指出大数据可视化对于城市规划的必要性和适用性; 其次通过实例看出应用大数据可视化的城市规划方案, 其质量和效率都得到了显著提升; 最后讨论了大数据可视化未来发展方向以及面临的挑战。

关键词: 大数据; 城市规划; 大数据可视化; 交通大数据可视化; 环境音大数据可视化

基金项目: 2019 年广东省普通高校特色创新类项目(2019KTSCX236): 基于深度学习的楼宇巡查机器人视觉图像识别分类研究; 广州华商学院“创新强校工程”: 2019 一流专业建设——计算机科学与技术; 校级青年学术科研项目(2021HSQX48): 基于结构增强的极大团社区发现算法

0 引言

近年来, 随着城市化速度不断加快, 各个城市在进行城市规划的过程中所遇到的问题不断增多, 冲突也不断激化。大数据的 4V (Volume, Velocity, Variety, Value) 特性使规划人员无法通过传统的城市规划方法来很好地解决现代城市涌现的一系列问题。而对于不同城市, 由于其社会结构、经济增长方式等差异, 规划人员又难以找到能够适用所有城市的解决方案。同时, 以大数据为典型的新一代信息与通讯技术不断融入社会的各个方面, 大数据在城市规划行业的实践和研究得到越来越多的应用^[1]。大数据可视化是大数据生命周期的最后一步, 其将数据直观地展现出来, 帮助人们观察数据, 快速理解数据并发现其中的特点与隐藏的意义。2016 年“可视化推动大数据平民化”被中国计算机学会大数据专家委员会评选为最有潜力的大数据发展方向^[2]。在大数据可视化广泛应用的同时, 可视化的缺点也暴露出来。首先, 可视化对于信息展现效果、数据融合效率、工具和对话机制影响、知识阐述能力的研究依然不够^[3]。其次, 国内外对于大数据

可视化停留在分析与理解的层面^[4], 没有完整的可视化治理体系, 缺乏对城市规划工作的理论指导和配套支撑^[5]。

传统城市规划存在着数据不足、过于依赖规划师个人经验、规划效果实现的蓝图对未来缺乏预测等诸多问题, 所以传统的城市规划方法并没有真实的依据集来支撑随后的分析和判断^[6]。而城市规划与大数据可视化结合是城市规划发展中里程碑式的转变。其主要表现在三个方面: 首先, 在建设设计方案和评价比较方案时运用大数据可视化技术将大幅度提高工作效率以及项目的总体水平, 同时也使收益最大化, 为项目的可行性分析提供科学依据; 其次, 对于城市整体而言, 一个微小的误差也许会使城市发展止步不前, 甚至后退; 最后, 在规划设计领域运用可视化技术, 提高资源利用率, 实现更高的经济效益。由此可见, 将大数据可视化应用到城市规划这一资源密集型的行业, 将会很大程度提高城市的信息化和数字化的程度, 使城市发展实现一个全新的快速模式^[7]。

本文共分为以下五节: 第一节通过对大数

据可视化的相关概述来引申出应用于城市规划中大数据可视化的技术及框架；第二节和第三节是交通大数据可视化和环境音大数据可视化应用于城市规划的实例；第四节，提出了更好的应用城市大数据可视化要注意的问题；第五节，综合前文以及城市发展的角度说明大数据可视化对于城市规划的必要性和重要性。

1 大数据可视化在城市规划中的驱动机理

1.1 大数据可视化的基本概念

数据是一种物理符号或这些物理符号的组合，描述了客观事物的性质，状态和相互关系^[8]。大数据被定义为可大量快速生成和(或)多种类型数据的信息资产，需要新的解决方案，使决策增强、深入的见解和流程优化成为可能^[9]。随着大数据产业链的愈加完善，未来发展的重点是大数据挖掘的分析和应用^[10]。大数据可视化由 Jim Thomas 教授在 21 世纪初提出，是大数据挖掘分析和应用中必不可少的技术^[11]，被定义为以图形或图像的形式在屏幕上显示数据，更直观地表达大数据的特征。对于数据，可视化在呈现数据以及其相关的处理结果的基础上，更是帮助用户进行数据分析和数据理解的有效途径^[12]。对于数据可视化而言，我们透过现象看本质的话，它关联着多个方面，从数据收集与处理、信息到可视的转化、知识到决策的变换等，并且在各个反面均有相对应的目标和框架^[13]。

1.2 城市规划涉及的大数据类型

城市规划是一个需要考虑的内容和需要理清的关系错综复杂的一个过程^[14]，越来越多的相关技术与方法正在被运用到城市规划工作中^[15-16]。这里介绍了三种应用大数据可视化的城市规划所涉及的大数据。

1.2.1 城市交通大数据

现在城市规划运用大数据可视化技术的领域众多，交通规划就是最普遍的领域之一^[17]；借助可视化分析技术，城市规划者可以清晰明了地观察城市交通的运行轨迹，知道市民的外出方式与状态以及城市经济活动模式^[18-19]。借

助可视化分析技术，对交通流量以及流速进行实时追踪，能够快速了解道路情况以及应对一些突发事件，营造一座交通顺畅安全的城市环境^[20]。从 2004 年以来，浮动车数据、公交车 GPS 数据、公交 IC 卡数据、ETC 数据、手机信令数据、停车数据等纷纷在交通规划领域得到了应用^[21-22]。

1.2.2 环境音大数据

近年来，城市环境噪音的污染对人们生活的干扰越来越严重，这让噪声污染问题成为政府和百姓最关心的话题之一^[23-24]。在我国，香港是针对城市部分地区绘制噪音地图起步较早的城市，2002 年绘制了 2D 交通噪音地图覆盖了 1100 km² 城区；2010 年通过对车辆源强和最新道路交通状况等对地图进行了更新，还绘制了 3D 交通噪声地图。而在内地地区，具有代表性的首张噪音地图是 2002 年上海市环科院为了提高徐汇区的环境治理而提出的^[25]。2008 年，深圳洛赛声学技术有限公司绘制了中国第一张城市区域噪声地图^[26]，覆盖了滨河路以南、深圳河以北的深圳福田南片约 12 km² 区域。在那之后，北京市劳动保护科学研究所于 2009 年基于已有的交通噪声数据，针对 12.7 km² 的城市区域完成了北京市的首张具有里程碑意义的城区交通噪声地图^[27]。

1.2.3 城市形态与结构大数据

应用于城市形态和结构方面的主要数据有百度热力图、互联网开放数据、手机信令等，其依据监测人群的散布以及实时的改变来达到体现功能以及空间散布的效果。吴志强和叶钟楠^[28]在百度热力图数据的基础上，对上海中心地区的人口分布程度、分布位置等数据进行了持续一周的观察，并对相应的数据进行分析，得到想要的结果。李娟等^[29]基于人的集聚指标，借助百度热力图的数据，在全国范围内实现多中心城市分类。钮心毅等^[30]结合手机信令数据，借助局部空间自相关的知识实现聚类分析，对上海中心地区的公共中心体系进行了识别与分类。郑晓伟^[31]以西安主城区为例，提出基于开放数据的城市中心体系识别与优化方法。

1.3 城市规划中的大数据可视化技术

可视化技术是将复杂数据以图形及图像的方式呈现,让人们可以直接看到数据隐含的规律,掌握其发展趋势,更为科学地做出决策^[32]。信息可视化、科学可视化和可视分析是可视化技术研究的三个主要方向。信息可视化解决的是数据的潜在相关性,提取并以合理的展现方式呈现给用户。科学可视化主要分析具有空间特性和高维度性的数据,提取有用的信息,研究数据深层次的规律。而实现科学可视化和信息可视化的基础是可视分析技术,其通过分析推理技术、转换技术、数据表征等数据处理技术从海量数据中提取有用信息,得到对复杂问题的进一步了解^[33]。

当可视化技术应用到城市规划中时会与更多技术相结合。GIS(地理信息系统)是与可视化技术结合相当广泛的系统,这个系统依靠计算机技术综合管理空间地理信息数据^[34],让用户更为清晰地了解数据的位置信息和相关属性,其应用的城市大数据包括交通大数据和环境音大数据等。随着数据的发展,城市大数据的维度也不断增加,例如日常开销、社交网络、数字地图等维度,其可以借助叠加式可视化技术结合相关的维度进行显示与分析,在年轻用户群体中十分受欢迎。

1.4 基于大数据可视化的城市规划的基本模型以及框架

1.4.1 可视化分析模型

可视化分析是交叉研究领域的典范之一,包含的内容有脑认知、人机交互、数据分析、信息可视化等^[35-36]。其模型如图1所示,首先通过数据变换将可用数据保存到数据库;然后将数据通过相应算法和可视化变换方法映射到定义好的可视化结构上,根据需要对结果进行变化设计,最终呈现在用户界面上。若有设计或数据的不合理之处用户都可以对系统进行反馈,如此反复修改展示就可实现可视分析的目的^[33]。数据可视化的准则是“信,达,雅”,即准确真实,通顺练达和表示优美,所以选定合理的可视化展示方法是可视分析模型的核心。

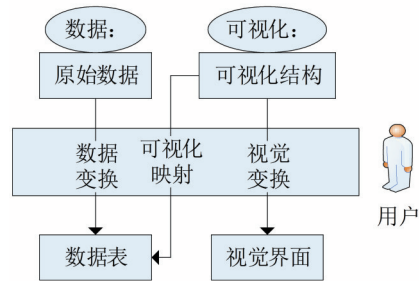


图1 可视化分析模型

1.4.2 可视化交互模型

用户在可视化的过程中也有着关键的作用,因为可视化的最终结果就是为广大用户所呈现。而能够协调系统与用户的重要手段就是使用交互模型^[37],其模型如图2所示。

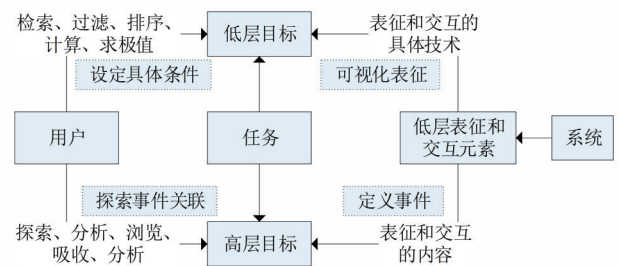


图2 可视化交互模型

在大数据可视化过程中,同一个城市地理空间或人文风俗在不一样的观察角度下所呈现的结构特征是不一样的,所以依据不同的需求,即使是对同一地点或事物的观察和分析也不能脱离交互式的可视化分析操作^[38]。人机交互中,用户根据可视化给出的结果结合自身的需求进行操作,将本身目的反馈给系统,从而得出最终结果。交互模型为可视化分析提供了明确目标,有效提高了分析过程的效率和准确率。

1.4.3 基于城市规划的移动数据可视化框架

城市移动数据与人们出行息息相关,具有海量、高维度的特征。通过对移动数据进行研究,除了能发现市民的行为特征,还能及时看到城市发展中的弊端,提早修正,使城市的发展更为快速且平稳。

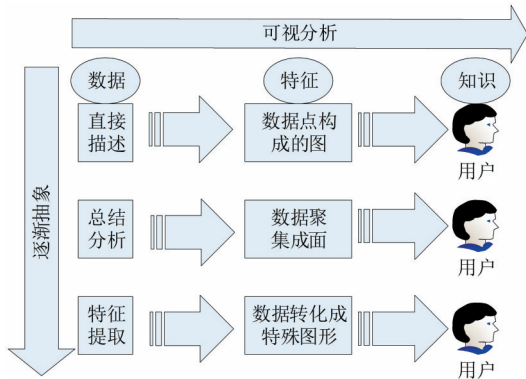


图3 移动数据可视化主要方法框架

图3是Andrienko等^[39]总结的移动数据可视化的三种方法的逻辑关系和流程。直接描述法不需要复杂的总结，只需将处理后的数据直接展示，其能在数据简单时给用户最简明直观的展示。而总结分析可在数据复杂程度不能用直接分析法时，通过计算对数据进一步解释，再呈现给用户，降低了理解成本。特征提取又是在总结分析的基础上，对数据进一步抽象，发现并分析其隐含特征。因此，在可视化过程中需要我们结合多种方法才能更充分地理解数据。

2 基于交通大数据可视化的城市规划运用

近年来在交通大数据的领域中，研究人员展示了一类将时间和空间数据完美融合到地图上的技术，称为路径缩放，这种技术让时空信息可视化的遮挡问题得到了解决^[40]。

2.1 基于路径缩放的交通可视化系统概述

如图4所示，下面部分是系统创建的一个就地的时空数据无遮挡可视化的流程，其核心问题是道路网络优化和层叠展示。在道路网优化部分，利用可以优化约束的能量函数把路径缩放演变成非线性最小二乘优化问题，有效地解决了优化问题。在层叠展示部分，用户可以与系统交互，在拓宽的路线上叠加各种信息。这个步骤利于实现焦点+上下文可视化，且完美地将时间数据和空间数据结合，以促进时间和空间数据的分析。

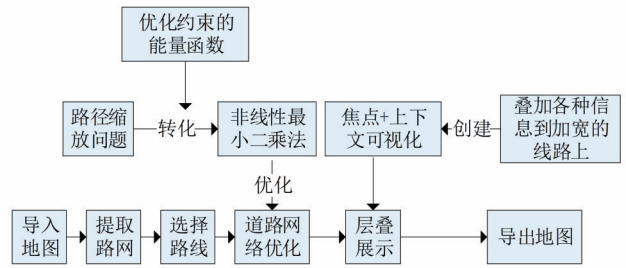


图4 系统创建流程以及核心问题

2.2 基于路径缩放的交通可视化实例论证

通过城市旅游的交互式焦点+上下文地铁地图来展示时空可视化的可用性。图5(a)是日本东京旅游的一个复杂地铁线路，图中有一条人们关注的最优线路的扩宽线路。拓宽线路上还保持其本身线路的信息并包含线路上的相关景点。这样人们就可以在拓宽的路线上看到关键信息，并且必要的上下文信息不受任何遮挡。另一种方法是创建一个静态的说明性地图，它可以直观、清晰地向广大观众传达重要信息。但通常不允许用户交互，重要信息直接显示，且不阻塞其他上下文信息。图5(b)可以清楚地说明路口周围交通模拟的结果。图5(c)显示了具有更宽连接道路的同一路交叉口，以显示具有双向交通流信息的密度图。

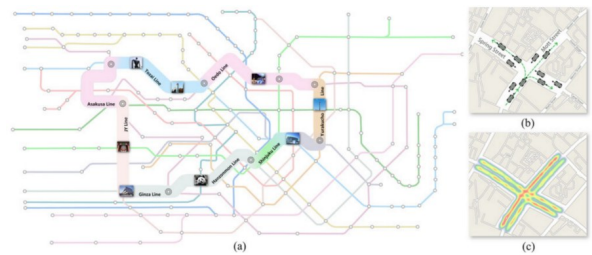


图5 路线缩放方法的实例聚焦+上下文无遮挡可视化

图5(a)具有突出显示的关键信息(例如，突出显示旅游景点和中转站的图像)的加宽旅游路线；(b)用于清晰显示交通模拟信息的加宽道路交叉口；(c)加宽道路交叉口覆盖密度图，以可视化双向交通流信息。

2.3 基于交通大数据的城市规划运用总结

基于路径缩放技术，研究人员可以将时空

数据作为一个不可分割的整体进行分析和表示。通过在拓宽的道路上直接叠加时间序列可视化,实现了空间和时间的无缝结合。拓宽路线对于交通问题有很好的改善,可以营造更好的城市出行环境。

3 基于环境音大数据的城市规划运用

以广州越秀区选定区域的噪声地图的绘制为例,其通过GIS软件和噪声模拟软件模拟预测数据,同时结合实地测量数据完善地图,最终以GIS电子地图的形式展现。

3.1 噪声地图的数据收集与统计

研究选定的区域在建筑数量超过千栋的城市中心区,由于噪声扩散传播的特性,研究范围长约1400 m,宽约1000 m,总面积约为1.6 km²。如图6所示,数据收集时首先将道路分为四级,前两级通过设置测量点以及根据车流量和重车比设置不同的录像采样时间来共同采集数据,而后两级则直接接收实测值来导入模拟软件反求噪声值,其车流录像记录时间固定为二十分钟。

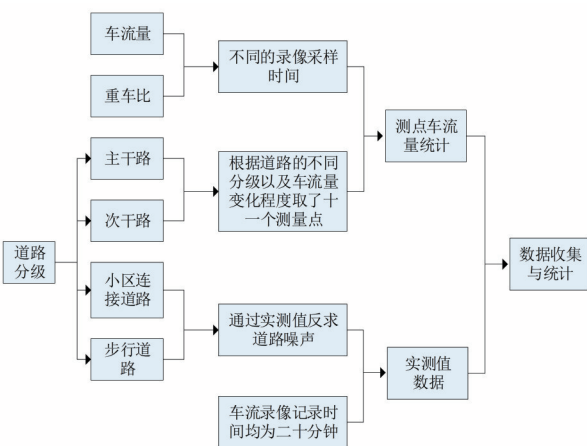


图6 数据收集与统计流程

3.2 噪声地图的数据模拟及导入GIS平台

在Cadna/A3.7XL噪声软件中首先输入了建筑和道路矢量数据以及GIS数据精度,坐标系等数据,然后在车流量数据和实测值的导入后经过噪声模拟得到噪声模拟图。模拟时,通过测点噪音和声场分布的计算得到该区受声点声级

和噪音地图。然后将通过Cadna/A软件模拟出来的噪声模拟地图数据以及接收点位置和实测的环境噪声数据都导入GIS平台软件,得到的噪音地图,如图7所示。图中不仅噪声测点和等值线都标出了对应的LAeq值,还可以直接看到建筑、道路等的位置信息和相关属性,使观察和查询变得更加便利。

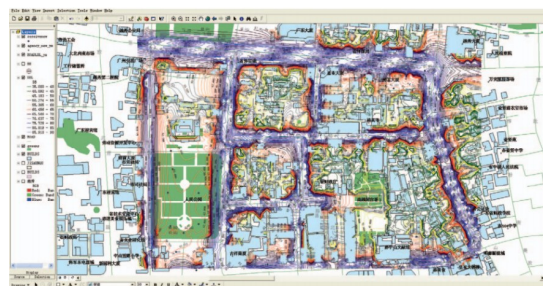


图7 GIS环境中的噪声地图

3.3 基于环境音大数据的城市规划运用总结

利用GIS已有的数据资源与专业软件的集成,从技术层面摆脱了计算大面积噪声图时数据难以获取和编辑复杂的问题,同时降低了噪声地图的制作成本,提高了效率和准确性^[41-42]。噪声地图能够快速判断主要的道路噪声污染位置,从而促进决策过程。同时根据噪声污染分布,有重点地布设监测点位,相应地增加或减少点位数量,提高点位代表性。

4 未来发展展望

现阶段大数据可视化在城市规划中的应用越来越多,但面对城市的快速发展和转型,大数据可视化这一技术仍需进一步提升和完善。首先由于大数据可视化有限的表现维度,用户面对愈发海量和多维的数据时,高效且准确地从中提取相关兴趣点是大数据可视化得以应用的基础,然后通过合适的渲染方式表达,是大数据可视化应用于城市规划的首要问题。其次,城市规划中的管理和决策繁重而复杂,而现有的简单图形及图像并不能使其得到充分且有效的支撑,所以其相关实践还是迫切需要科学的理论指导。再次,数据是在特定语境下才具有特定具体的含义,其本身不会产生知识。因为大数据具有价值密度低的特点,所以它展现的

过程也十分艰难,需要借助日益成熟的业务语义模型,才可以广泛化、多形式地辅助大数据可视化的使用。最后,城市规划的有序执行并不仅仅依赖于对数据的处理和展示,人机交互才是城市规划过程中的纽带。借助大数据可视化技术,加强规划者和计算机之间的交互是规划者做出行之有效的决策的重要环节之一。而为可视化应用实践提供技术和平台保障的是交互式系统的建设,这一部分也有待更深层次的研究。

5 结语

伴随着城市生活的日益丰富与便捷,涉及到城市规划的相关因素也越来越错综复杂,从交通出行到物流、通讯,从噪音管理到生态平衡,从医疗社保到商业活动等。各个方面的城市大数据的多细节、多维度特性不仅给城市带来了机遇,也增加了挑战。显而易见的是,之前城市规划惯例的方法,也就是依赖于理论知识和以往经验的方法,已经不能解决现有的城市规划的大部分问题了。大数据的分析以及可视化的出现为现代城市规划提供了新的方法。可视化作为有效且灵活的沟通手段和信息表达,不仅仅为大数据探索分析提供了有效途径,还提供了更加简洁明了又精准的数据、认知过程理解及问题发现和解决,是营造更好的城市运行系统不可或缺的部分。

参考文献:

- [1] 王腾. 大数据在城市总体规划编制中的应用方法研究[D]. 武汉:武汉大学,2017.
- [2] 赵颖,张卓,袁晓如. 数据可视分析挑战赛三年回顾[J]. 网络与信息安全学报,2018,4(2):55-61.
- [3] BILLGER M, THUVANDER L, WASTBERG B. In search of visualization challenges: the development and implementation of visualization tools for supporting dialogue in urban planning processes[J]. *Environment and Planning B: Urban Analytics and City Science*, 2017, 44(6):1012-1035.
- [4] ZHENG Y, WU W, CHEN Y, et al. Visual analytics in urban computing: an overview[J]. *IEEE Transactions on Big Data*, 2016, 2(3):276-296.
- [5] 孙轩,孙涛. 基于大数据的城市可视化治理:辅助决策模型与应用[J]. *公共管理学报*, 2018, 15(2):120-129, 158-159.
- [6] 尹稚. 价值观·逻辑链·依据集[R]. 北京:第四届清华同衡学术周,2016.
- [7] 谢波. 关于城市设计中三维可视化关键技术的研究[D]. 武汉:华中科技大学,2004.
- [8] 黄正东,于卓,黄经南. 城市地理信息系统[M]. 武汉:武汉大学出版社,2010.
- [9] JOHNSON B W. DATA WAREHOUSES: they aren't just for BIG GUYS anymore[J]. *Behavioral Health Management*, 2002(1):26.
- [10] 付长军,乔宏章. 大数据产业发展现状研究[J]. *无线电通信技术*, 2016, 42(4):1-4, 28.
- [11] MAY R, HANRAHAN P, KEIM D A, et al. The state of visual analytics is and where it is going[J]. 2010 IEEE Symposium on Visual Analytics Science and Technology(VAST), 2010:257-259.
- [12] PARSONS P, SEDIG K. Handbook of human centric visualization[M]. New York: Springer, 2014: 671-691.
- [13] CHEN M, JAENICKE H. An information-theoretic framework for visualization[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2010, 16(6):1206-1215.
- [14] 宋刚,张楠,朱慧. 城市管理复杂性与基于大数据的应对策略研究[J]. *城市发展研究*, 2014, 21(8):95-102.
- [15] TYRVAJAINEN L, GUSTAVSSON R, KONIJN-ENDIJK C, et al. Visualization and landscape laboratories in planning, design and management of urban woodlands[J]. *Forest Policy and Economics*, 2006, 8(8):811-823.
- [16] MANSMANN F, KEIM D, NORTH S, et al. Visual analysis of network traffic for resource planning, interactive monitoring, and interpretation of security threats[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2007, 13(6):1105-1112.
- [17] CHEN W, GUO F, WANG F. A survey of traffic data visualization[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2015, 16(6):2970-2984.
- [18] FERREIRA N, POCO J, VO H, et al. Visual exploration of big spatio-temporal urban data: a study of New York City taxi trips[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2013, 19(12):

- 2149-2158.
- [19] ZENG W, FU C, ARISONA S, et al. Visualizing mobility of public transportation system [J]. IEEE Transactions on Visualization and Computer Graphics, 2014, 20(12):1833-1842.
- [20] SHI Q, ABDEL-ATY M. Big Data applications in real-time traffic operation and safety monitoring and improvement on urban expressways [J]. Transportation Research Part C: Emerging Technologies, 2015, 58:380-394.
- [21] 龙瀛,张宇,崔承印. 利用公交刷卡数据分析北京居住关系和通勤出行[J]. 地理学报, 2012, 67(10):14.
- [22] 姚凯,钮心毅. 手机信令数据分析在城镇体系规划中的应用实践:南昌大都市区的案例[J]. 上海城市规划, 2016(4):91-97.
- [23] 耿瑞珍. 建筑施工噪声何时休[J]. 环境保护, 2011(15):54-55.
- [24] 贾玉鹤. 沈阳市当前环境信访与投诉案件的对策研究[J]. 环境科学导刊, 2007, 26(6):13-17.
- [25] 乐晓妍. 城市道路噪声特性研究及噪声地图信息系统构建[D]. 杭州:浙江大学, 2015.
- [26] 胡笑汗,周鑫,赵晶,等. 我国噪声地图绘制现状及建议[J]. 环境保护, 2018, 46(9):66-68.
- [27] 刘磊,户文成. 噪声地图技术的发展及应用实践[J]. 噪声与振动控制, 2010(81):175-179.
- [28] 吴志强,叶鍾楠. 基于百度地图热为图的城市空间结构研究:以上海中心城区为例[J]. 城市规划, 2016(4):33-40.
- [29] 李娟,李苗裔,龙瀛,等. 基于百度热力图的中国多中心城市分析[J]. 上海城市规划, 2016(3):30-36.
- [30] 钮心毅,丁亮,宋小冬. 基于手机数据识别上海中心城市的城市空间结构[J]. 城市规划学刊, 2014(6):61-67.
- [31] 郑晓伟. 基于开放数据的西安城市中心体系识别与优化[J]. 规划师, 2017, 33(1):57-64.
- [32] 陈文兵,宋玛君,王廷春. 一种利用极限学习机的数据可视化方法[J]. 计算机工程与科学, 2017, 39(5):912-918.
- [33] 宫睿. 海量城市交通数据可视化与人类出行行为的可视分析研究[D]. 成都:电子科技大学, 2018.
- [34] LI Y. Study on the applications of 3D visualization system of geographic information [C] // 4th International Conference on Electrical & Electronics Engineering and Computer Science (ICEECS 2016). Advances in Computer Science Research, 2016.
- [35] 任磊,杜一,马帅,等. 大数据可视分析综述[J]. 软件学报, 2014, 25(9):1909-1936.
- [36] 曾悠. 大数据时代背景下的数据可视化概念研究[D]. 杭州:浙江大学, 2014.
- [37] 董亮. 基于出租车数据的交通数据可视化研究[D]. 兰州:兰州交通大学, 2016.
- [38] MCKENZIE G, JANOWICZ K, GAO S, et al. POI pulse: a multi-granular, semantic signature-based information observatory for the interactive visualization of big geosocial data [J]. Cartographica: The International Journal for Geographic Information and Geovisualization, 2015, 50(2):71-85.
- [39] ANDRIENKO G L, ANDRIENKO K N, DYKES J, et al. Geovisualization of dynamics, movement and change: key issues and developing approaches in visualization research [J]. Information Visualization, 2008, 7(3):173-180.
- [40] SUN G, LIANG R, QU H, et al. Embedding spatio-temporal information into maps by route-zooming [J]. IEEE Transactions on Visualization and Computer Graphics, 2017, 1:1506-1519.
- [41] 周旻. 基于GIS的噪声地图研究[D]. 广州:华南理工大学, 2009.
- [42] ZHOU M. GIS-based city central area noise mapping [C] // 2011 International Conference on Remote Sensing, Environment and Transportation Engineering, 2011: 2192-2195. DOI: 10.1109/RSETE.2011.5964743.

作者简介:

通信作者:操心慧(1995—),女,安徽安庆人,硕士,助教,研究方向为数据可视化, E-mail: 945152118@qq.com

操金金(1994—),女,安徽安庆人,硕士,研究方向为信息工程

许丽娟(1980—),女,青海人,硕士,副教授,研究方向为数据挖掘算法、计算机视觉

弓佳明(1993—),男,陕西咸阳人,助教,研究方向为社区发现算法

赵茜(1995—),女,辽宁大连人,助教,研究方向为金融科技

收稿日期: 2022-03-20 修稿日期: 2022-04-16

(下转第110页)

文章编号: 1007-1423(2022)16-0084-07

DOI: 10.3969/j.issn.1007-1423.2022.16.016

基于 RFID 和树莓派的物品收纳管理系统

卢依思, 盘茂杰, 郑思晴, 梁焕诗, 魏纵横

(广东技术师范大学计算机科学学院, 广州 510665)

摘要: 随着科技的发展和社会的进步, 对物品收纳的需求也日益增加, 如何进行高效的物品管理成为大家关注的问题。针对该问题, 本文以树莓派为控制模块, 结合 RFID 和 4G 等技术开发了物品收纳管理系统, 以实现简单高效的物品收纳。系统由以下几部分组成: 由树莓派与 RFID 阅读器和标签结合组成的感知层, 具有 4G 通信模块的传输层, 网页系统与云数据库结合的应用层。目前该系统已投入使用, 在一定程度上减少了人力、物力的支出, 提高了物品利用率。

关键词: RFID 技术; 物品管理; 树莓派; 4G 技术

基金项目: 广东普通高校重点项目(2019KZDXM063); 国家自然科学基金(2019KZDXM063); 广州科技计划项目(202102080348); 2022 “攀登计划” 专项资金项目(pdjh2022b0298); 2021 广东省大学生创新创业训练计划项目(202110588064)

0 引言

随着社会经济的发展和人们生活水平的提高, 日常生活中需要使用的物品种类日益繁多, 人们时常会遗忘或无法寻得自己所需物品。除此之外, 一个人的生活态度也可从他所处环境的整洁程度反映出来。如何做好物品的存储和管理工作, 不仅是一个简单的收纳问题, 更是一项人生的重要课题, 收纳对于提高人们生活便利是一件利大于弊的事情。因为物品都是有价值且被需要的, 有必要将它们分类整理, 并放置在合适的位置上。但是, 传统的物品收纳只是依靠纸笔和便签纸进行手工记录, 工作效率较低, 而且容易出错。在此背景下, 如何利用智能设备进行物品收纳管理成为人们关注的一个重要问题。

在物联网技术的支持下, 智能设备可以为物品添加相关信息属性, 在提供查找物品便利的同时, 让用户能及时了解某些物品的重要信息, 例如: 家居智能设备的运行情况, 药品或者食品是否过期, 等等。为了实现以上的物品收纳管理功能, 本文通过调研获得物品收纳管理系统的相关需求, 并开发出了基于 RFID 和树莓派的物品收纳管理系统。

1 国内外研究现状

1.1 物品收纳管理系统

目前市面上的物品收纳管理系统, 主要是针对贵重的、易丢失的物品, 而这种系统大部分是面向需要管理物品较多的企业、行政机关单位或者是藏书较多的图书馆, 只有极少部分是针对家庭收纳管理的, 而且成本较高, 灵活性不足。

其中, 鲁金屏等^[1]设计的资产管理系统是用于事业单位的系统, 支持手持阅读器手动盘点、自动盘点设备远程盘点这两种资产盘点方式。唐赞淞等^[2]提出一种家庭物品智能管理系统方案, 主要通过手机 APP 和 RFID 技术结合的模式, 实现物品位置检测及物品有效期提醒等功能。

1.2 4G

4G 技术, 即第四代无线通信技术, 它拥有比 3G 更快的数据传输速度, 以满足传输多样信息的需求, 为人们的日常生活带来极大的便利^[3]。

从 2009 年 4G 面世以来, 信息技术行业在 4G 技术的推动下发生了翻天覆地的改变。高效的 GPS, 手机游戏的社交, 汽车的智能化, 电子媒体, 短视频, 云计算和云存储等技术在我

国都得到了发展,这些技术被广泛应用在人们生活的方方面面^[4],是一种成熟的远距离无线通信技术。

1.3 RFID

RFID技术使用电子标签来标志某个物品,再通过读写器对标签上数据进行收集和处理。作为非接触性自动识别技术,射频识别(RFID)已在国内外迅速发展。通过射频信号可以捕获静态或者快速移动的物体相关信息,用来加以识别。与条形码识别和直接标记技术相比,RFID技术有更长的读取距离,更大的信息储存容量,更高的防污染性能和可读可写的读写性能,并且能够重复使用,在一个读取过程中,可以同时阅读多个电子标签^[5]。目前,RFID技术在国内外的物品管理中得到了广泛的应用,特别是在图书管理和物流仓储管理上。

2 需求分析

基于上述问题,本文从硬件和软件两方面对物品收纳管理系统进行了详细分析和设计,利用RFID技术,采用无线电信号识别的方式,无接触式地读取特定目标对应的数据信息。在不影响管理效率的同时,最大程度地降低物品管理难度^[6]。

为应对不同的应用场景,我们将系统分为两大功能模块:单物品操作功能、多物品操作功能。

在物品管理过程中,系统会对单个物品进行操作,面对日常的添加、借出、归还、盘点,本系统有:添加物品(入库)、拿出物品、放入物品、查询物品等功能,实现对单个物品的日常操作;面对多物品操作,本系统有:多物品识别、拿出物品、放入物品、盘点物品等功能。多物品操作可实现批量操作,在一定程度上提高了物品的识别效率和准确率。同时,当我们需要对所有物品进行日常盘点时,多物品操作功能的多物品识别可以识别出周边粘贴RFID标签的物品,在显示屏显示所识别出的若干物品,仓库盘点人员只需点击该功能几次,进行对比即可,大大提高盘点效率^[7]。

同时,为应对无区域网覆盖的情况,我们在该系统加入了4G通信模块,保证网络的供

给。用户可根据需求,选择加入或不加4G通信模块。

3 系统设计

3.1 系统架构设计

基于RFID和树莓派的物品收纳管理系统整体架构如图1所示,包括感知层、传输层和应用层。

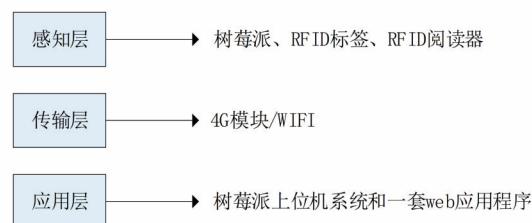


图1 按层次划分的系统架构图

感知层由RFID电子标签、RFID阅读器芯片以及主控板树莓派组成,用于识别物品并采集相关数据;网络层使用4G模块将感知层采集数据上传至云端;应用层则提供了一套树莓派上位机系统和一套Web应用程序,用户使用相应系统即可实现对物品的各种操作。

3.2 系统功能模块划分

本系统共提供了两套应用软件,系统整体的功能划分如图2所示,分为登录注册模块、物品管理模块、借还记录模块和盘点统计模块,其中Web端实现了登录注册模块、物品管理模块的借出物品、归还物品、借还记录和盘点统计等功能;上位机软件则实现了物品管理的所有功能以及借还记录和盘点统计等功能。

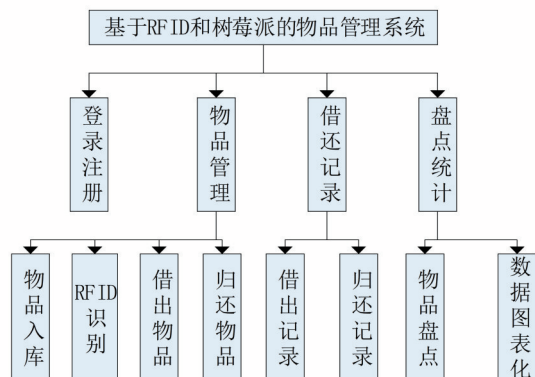


图2 系统功能划分图

4 硬件设计

4.1 硬件选型

此系统所需的硬件设备主要有：树莓派、RFID模块、4G模块。下面将详细介绍各硬件特点及选择原因。

4.1.1 树莓派

树莓派是一款基于ARM的微型电脑主板，以MicroSD卡为内存硬盘，拥有多个USB接口和I/O口，具有所有PC的基本功能，可以安装系统，做服务器等，价格却比计算机便宜许多，体积小。因此，本系统选用了型号为PI4B-2GB的树莓派。

4.1.2 RFID模块

RFID模块包括了阅读器芯片、天线和电子标签。RFID具有电子标签存储信息量大、信息可修改、安全性高、无线传播，以及信息传输范围广等优点。阅读器芯片可以通过天线实现与电子标签的无线通信，实现对标签识别码和内存数据的读出和写入操作。

4.1.3 4G模块

4G模块是基于4G网络来进行通信的，是指支持TD-LTE和FDD-LTE等LTE网络制式的统称，具有通信速度快、网络频谱宽、通信灵活等特点。本系统采用的4G模块为SIM7600CE 4G HAT。

4.2 硬件模块搭建

树莓派主板拥有多个USB接口和I/O口，因此本系统硬件层搭建只需将4G模块通过树莓派USB口相连，RFID阅读器通过GPIO引脚口与树莓派相连，硬件层具体架构如图3所示。

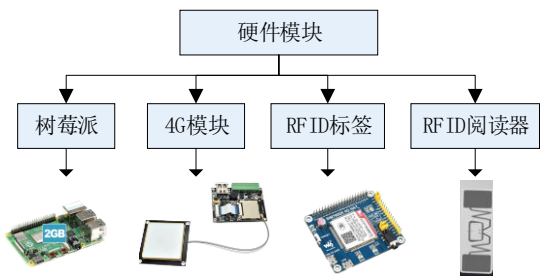


图3 硬件架构图

5 软件设计

5.1 软件系统整体架构

基于RFID和树莓派的物品管理系统在软件方面主要由三部分组成：第一部分为树莓派终端的以Python为脚本的物品管理程序，是本系统的核心软件；第二部分为B/S架构的Web应用程序，提供给用户更加便利的物品管理页面，辅助树莓派终端软件，提升用户体验感；第三部分为组织、存储和管理数据的云数据库。系统整体架构图如图4所示。

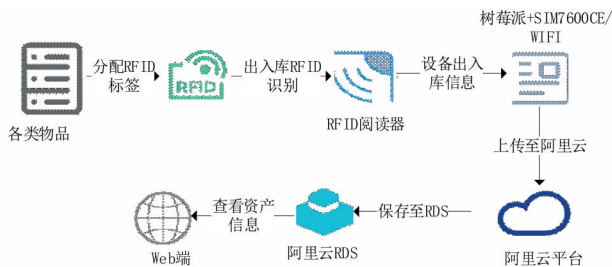


图4 系统整体架构图

云数据库具有低成本、高可靠和轻松部署等优点，在本套系统中起到了软件中间件的作用。云数据库是一个非常可靠又重要的中间件，它不仅能对树莓派上传的数据进行加工处理和保存，还可以与树莓派终端软件、Web端软件实时连接，进行数据交互。通过使用云数据库，两套软件系统之间形成了一条高效传输的数据链路。

树莓派终端和Web网页通过共同连接到云数据库，实现物品数据共享，从而相辅相成地实现有序、合理的物品管理，实现物品管理的数据化、可视化。

5.2 树莓派终端软件设计

树莓派终端软件采用Qt设计UI界面，基于RFID模块、4G模块和树莓派等硬件基础，使用Python语言进行开发，通过感知层采集数据，并实时与云端进行数据交互，实现物品管理各项功能。

5.3 B/S架构的Web应用程序设计

Web采用了Spring+SpringMVC+Mybaits的框架加上Maven管理项目结构，实现前后端分离，

并实时获取云端数据,对云端数据进行查看修改,实现对物品管理数据的可视化。

5.4 数据库设计

本系统采用MySQL语言进行数据库设计,根据系统的需求分析与总体设计,为实现物品的各项管理操作,共设计有用于存储物品信息的data表、用于存储借出记录的table_out_storage表、用于存储归还记录的table_in_storage表共三张数据表。各表详细结构如图5、图6和图7所示。

序号	列名	类型	长度	可空	主键	备注
1	deviceId	varchar	100	<input type="checkbox"/>	<input checked="" type="checkbox"/>	
2	deviceName	varchar	50	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
3	deviceUser	varchar	50	<input type="checkbox"/>	<input type="checkbox"/>	
4	note	varchar	255	<input checked="" type="checkbox"/>	<input type="checkbox"/>	

图5 data表

序号	列名	类型	长度	可空	主键	备注
1	outStorageId	varchar	128	<input type="checkbox"/>	<input checked="" type="checkbox"/>	出库单号 uuid
2	deviceId	varchar	100	<input type="checkbox"/>	<input type="checkbox"/>	出库的设备id
3	deviceName	varchar	128	<input type="checkbox"/>	<input type="checkbox"/>	出库的设备名称
4	newDeviceUser	varchar	128	<input type="checkbox"/>	<input type="checkbox"/>	新的设备使用者

图6 table_out_storage表

序号	列名	类型	长度	可空	主键	备注
1	inStorageId	varchar	128	<input type="checkbox"/>	<input checked="" type="checkbox"/>	入库单号 uuid
2	deviceId	varchar	100	<input type="checkbox"/>	<input type="checkbox"/>	入库的设备id
3	deviceName	varchar	128	<input type="checkbox"/>	<input type="checkbox"/>	入库的设备名称
4	deviceUser	varchar	128	<input type="checkbox"/>	<input type="checkbox"/>	使用人

图7 table_in_storage表

6 系统实现

6.1 硬件功能实现

6.1.1 RFID模块实现

通过RFID模块实现的功能主要是识别RFID标签的EPC以及个数。利用射频信号和空间耦合,实现对被识别物体的自动识别。RFID

模块通过天线与RFID电子标签进行无线通信,可以实现对标签识别码和内存数据的读出或写入操作。当一台粘贴RFID标签的物品进入某一特定区域,可立即被RFID模块自动识别,同时串口将对应读卡器号码以及相应标签信息内码实时传递发送至树莓派,从而展开批量入库、出库、清点等工作^[8]。RFID模块识别到的标签效果如图8所示。

```
main_window_controller.py[line:100] - DEBUG: 调试窗口
[{"01203040506070809101112131415161718192021222324252627282930313233343536373839404142434445464748495051525354555657585960616263646566676869707172737475767778798081828384858687888990919293949596979899100101102103104105106107108109110111112113114115116117118119120121122123124125126127128129130131132133134135136137138139140141142143144145146147148149150151152153154155156157158159160161162163164165166167168169170171172173174175176177178179180181182183184185186187188189190191192193194195196197198199200201202203204205206207208209210211212213214215216217218219220221222223224225226227228229230231232233234235236237238239240241242243244245246247248249250251252253254255256257258259260261262263264265266267268269270271272273274275276277278279280281282283284285286287288289290291292293294295296297298299300301302303304305306307308309310311312313314315316317318319320321322323324325326327328329330331332333334335336337338339340341342343344345346347348349350351352353354355356357358359360361362363364365366367368369370371372373374375376377378379380381382383384385386387388389390391392393394395396397398399400401402403404405406407408409410411412413414415416417418419420421422423424425426427428429430431432433434435436437438439440441442443444445446447448449450451452453454455456457458459460461462463464465466467468469470471472473474475476477478479480481482483484485486487488489490491492493494495496497498499500501502503504505506507508509510511512513514515516517518519520521522523524525526527528529530531532533534535536537538539540541542543544545546547548549550551552553554555556557558559560561562563564565566567568569570571572573574575576577578579580581582583584585586587588589590591592593594595596597598599600601602603604605606607608609610611612613614615616617618619620621622623624625626627628629630631632633634635636637638639640641642643644645646647648649650651652653654655656657658659660661662663664665666667668669670671672673674675676677678679680681682683684685686687688689690691692693694695696697698699700701702703704705706707708709710711712713714715716717718719720721722723724725726727728729730731732733734735736737738739740741742743744745746747748749750751752753754755756757758759760761762763764765766767768769770771772773774775776777778779780781782783784785786787788789790791792793794795796797798799800801802803804805806807808809810811812813814815816817818819820821822823824825826827828829830831832833834835836837838839840841842843844845846847848849850851852853854855856857858859860861862863864865866867868869870871872873874875876877878879880881882883884885886887888889890891892893894895896897898899900901902903904905906907908909910911912913914915916917918919920921922923924925926927928929930931932933934935936937938939940941942943944945946947948949950951952953954955956957958959960961962963964965966967968969970971972973974975976977978979980981982983984985986987988989990991992993994995996997998999100010011002100310041005100610071008100910101011101210131014101510161017101810191020102110221023102410251026102710281029103010311032103310341035103610371038103910401041104210431044104510461047104810491050105110521053105410551056105710581059106010611062106310641065106610671068106910701071107210731074107510761077107810791080108110821083108410851086108710881089109010911092109310941095109610971098109911001100110021003100410051006100710081009101010111012101310141015101610171018101910201021102210231024102510261027102810291030103110321033103410351036103710381039104010411042104310441045104610471048104910501051105210531054105510561057105810591060106110621063106410651066106710681069107010711072107310741075107610771078107910801081108210831084108510861087108810891090109110921093109410951096109710981099110001100111002110031100411005110061100711008110091101011011110121101311014110151101611017110181101911020110211102211023110241102511026110271102811029110301103111032110331103411035110361103711038110391104011041110421104311044110451104611047110481104911050110511105211053110541105511056110571105811059110601106111062110631106411065110661106711068110691107011071110721107311074110751107611077110781107911080110811108211083110841108511086110871108811089110901109111092110931109411095110961109711098110991110001110011110021110031110041110051110061110071110081110091110101110111110121110131110141110151110161110171110181110191110201110211110221110231110241110251110261110271110281110291110301110311110321110331110341110351110361110371110381110391110401110411110421110431110441110451110461110471110481110491110501110511110521110531110541110551110561110571110581110591110601110611110621110631110641110651110661110671110681110691110701110711110721110731110741110751110761110771110781110791110801110811110821110831110841110851110861110871110881110891110901110911110921110931110941110951110961110971110981110991111000111100111110021111003111100411110051111006111100711110081111009111101011110111111012111101311110141111015111101611110171111018111101911110201111021111102211110231111024111102511110261111027111102811110291111030111103111110321111033111103411110351111036111103711110381111039111104011110411111042111104311110441111045111104611110471111048111104911110501111051111105211110531111054111105511110561111057111105811110591111060111106111110621111063111106411110651111066111106711110681111069111107011110711111072111107311110741111075111107611110771111078111107911110801111081111108211110831111084111108511110861111087111108811110891111090111109111110921111093111109411110951111096111109711110981111099111110001111100111111002111110031111100411111005111110061111100711111008111110091111101011111011111101211111013111110141111101511111016111110171111101811111019111110201111102111111022111110231111102411111025111110261111102711111028111110291111103011111031111110321111103311111034111110351111103611111037111110381111103911111040111110411111104211111043111110441111104511111046111110471111104811111049111110501111105111111052111110531111105411111055111110561111105711111058111110591111106011111061111110621111106311111064111110651111106611111067111110681111106911111070111110711111107211111073111110741111107511111076111110771111107811111079111110801111108111111082111110831111108411111085111110861111108711111088111110891111109011111091111110921111109311111094111110951111109611111097111110981111109911111100011111100111111100211111100311111100411111100511111100611111100711111100811111100911111101011111101111111012111111013111111014111111015111111016111111017111111018111111019111111020111111021111111022111111023111111024111111025111111026111111027111111028111111029111111030111111031111111032111111033111111034111111035111111036111111037111111038111111039111111040111111041111111042111111043111111044111111045111111046111111047111111048111111049111111050111111051111111052111111053111111054111111055111111056111111057111111058111111059111111060111111061111111062111111063111111064111111065111111066111111067111111068111111069111111070111111071111111072111111073111111074111111075111111076111111077111111078111111079111111080111111081111111082111111083111111084111111085111111086111111087111111088111111089111111090111111091111111092111111093111111094111111095111111096111111097111111098111111099111111100011111110011111111002111111100311111110041111111005111111100611111110071111111008111111100911111110101111111011111111012111111101311111110141111111015111111101611111110171111111018111111101911111110201111111021111111102211111110231111111024111111102511111110261111111027111111102811111110291111111030111111103111111110321111111033111111103411111110351111111036111111103711111110381111111039111111104011111110411111111042111111104311111110441111111045111111104611111110471111111048111111104911111110501111111051111111105211111110531111111054111111105511111110561111111057111111105811111110591111111060111111106111111110621111111063111111106411111110651111111066111111106711111110681111111069111111107011111110711111111072111111107311111110741111111075111111107611111110771111111078111111107911111110801111111081111111108211111110831111111084111111108511111110861111111087111111108811111110891111111090111111109111111110921111111093111111109411111110951111111096111111109711111110981111111099111111110001111111100111111110021111111100311111111004111111110051111111100611111111007111111110081111111100911111111010111111110111111110121111111101311111111014111111110151111111101611111111017111111110181111111101911111111020111111110211111111102211111111023111111110241111111102511111111026111111110271111111102811111111029111111110301111111103111111111032111111110331111111103411111111035111111110361111111103711111111038111111110391111111104011111111041111111110421111111104311111111044111111110451111111104611111111047111111110481111111104911111111050111111110511111111105211111111053111111110541111111105511111111056111111110571111111105811111111059111111110601111111106111111111062111111110631111111106411111111065111111110661111111106711111111068111111110691111111107011111111071111111110721111111107311111111074111111110751111111107611111111077111111110781111111107911111111080111111110811111111108211111111083111111110841111111108511111111086111111110871111111108811111111089111111110901111111109111111111092111111110931111111109411111111095111111110961111111109711111111098111111110991111111110001111111110011111111100211111111100311111111100411111111100511111111100611111111100711111111100811111111100911111111101011111111101111111110121111111110131111111110141111111110151111111110161111111110171111111110181111111110191111111110201111111110211111111110221111111110231111111110241111111110251111111110261111111110271111111110281111111110291111111110301111111110311111111110321111111110331111111110341111111110351111111110361111111110371111111110381111111110391111111110401111111111041111111111042111111111104311111111110441111111111045111111111104611111111110471111111111048111111111104911111111110501111111111051111111111105211111111110531111111111054111111111105511111111110561111111111057111111111105811111111110591111111111060111111111110611111111111062111111111110631111111111106411111111111065111111111110661111111111106711111111111068111111111110691111111111107011111111111071111111111110721111111111107311111111111074111111111110751111111111107611111111111077111111111110781111111111107911111111111108011111111111108111111111111108211111111111
```

端软件拥有 Web 软件所没有的 RFID 模块，因此终端软件拥有特有的入库和 RFID 识别功能。当使用入库、RFID 识别或借还功能时，树莓派将通过串口向 RFID 模块发送读取等命令，而

RFID 阅读器则将接收的标签信息返回给程序，程序对采集信息进行处理和再加工后，与云数据库进行交互，从而实现各个功能。树莓派终端的软件流程见图 10。

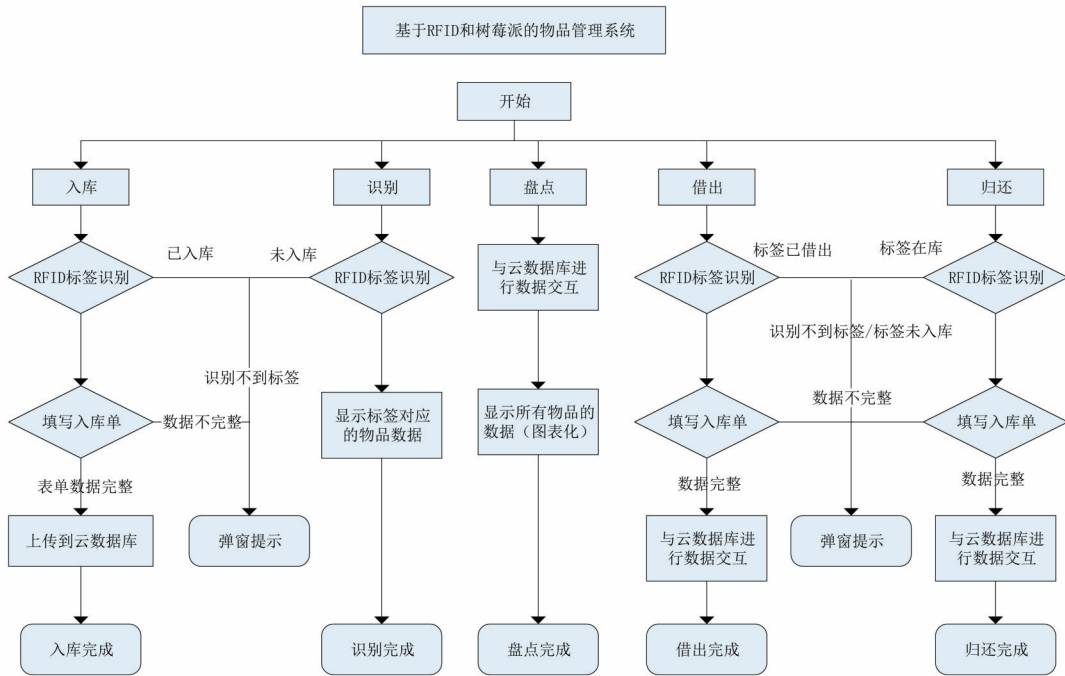


图 10 树莓派终端的软件流程图

树莓派终端软件主界面如图 11 所示，主界面包含了各个功能按钮，供用户进行相应的物品管理操作。



图 11 主界面(菜单)

若用户点击“添加物品”按钮，则系统识别一定范围内的标签，并生成如图 12 所示的入库表单，用户补充物品信息即可完成物品入

库。物品借出、物品归还与入库功能的实现基本一致。



图 12 入库表单

若用户点击“多物品识别”按钮，则系统识别一定范围内的标签，并将各标签及其对应物品信息生成如图 13 所示的资产情况表格。若用户点击“盘点物品”按钮，则系统将显示如

图 14 所示的现有物品情况表格,该盘点界面还提供数据图表化的按钮,供用户直观地查看各类物品数据。



当前rfid识别资产情况

物品id	物品名称	当前使用人	备注	创建时间	当前所在位置	状态
1 e28011606000...	test	b		2022-03-13 09...	404	在库
2 010203056000...	test	b		2022-03-13 09...	404	在库
3 e18011606000...	test	b		2022-03-13 09...	404	在库

图 13 RFID识别功能示意图



现有物品情况 所有物品数据化显示 油壶

物品id	物品名称	当前使用人	备注	创建时间	当前所在位置	状态
28 1440696889...	家秤	陈诚	厨房	2021-10-28 1...	206	在库
29 1440696893...	微波炉	陈诚	厨房	2021-10-28 1...	206	在库
30 1440696897...	挂钩	卢杰	厨房	2021-11-09 ...	206	在库
31 1440696902...	消毒柜	卢杰	厨房	2021-10-28 1...	206	在库
32 1440696906...	冰箱	卢杰	厨房	2021-10-28 1...	206	在库
33 1440696910...	抹布	陈诚	厨房	2021-10-28 1...	206	在库
34 1440696914...	STM32精英板	陈诚	厨房	2021-10-28 1...	206	在库
35 1440696918...	洗菜盆	陈诚	厨房	2021-10-28 1...	206	在库
36 1440696923...	花生2斤	卢杰	厨房	2021-10-28 1...	206	在库
37 1440696927...	铁锹	卢杰	阳台	2021-10-28 1...	206	在库
38 1440696965...	油壶	卢杰	厨房	2021-10-28 1...	206	在库
39 1455922311...	过滤器	卢依思	卧室	2021-11-17 2...	101	在库

图 14 盘点功能示意图

当用户只需查看物品、修改物品或删除物品等非RFID识别相关功能时,可直接使用本应用程序,从而达到简单快捷的管理效果。

Web端的主要页面如图16至图18所示。



图 16 登录页面



收的物物品管理系统

物品id	物品名称	物品介绍	入库时间	入库地点	物品状态	操作
143373217941819320	水盆	卢依思	陈诚录入	2021年10月26日	206	已出库 修改删除
1433732366270925624	厨房纸	卢依思	陈诚录入	2021年10月26日	206	已出库 修改删除
143373237667922432	椅子	卢	陈诚录入	2021年10月26日	204	未出库 修改删除
143373238066420736	芝麻糖	卢依思	陈诚录入	2021年10月26日	206	已出库 修改删除
143373238509113344	苹果	卢依思	陈诚录入	2021年10月26日	206	已出库 修改删除
143373241444018560	7寸显示屏	卢依思	陈诚录入	2021年10月26日	206	已出库 修改删除
1433732416644516664	RFID阅读器	卢依思	陈诚录入	2021年10月26日	206	已出库 修改删除
1433732422843015168	笔记本	卢依思	陈诚录入	2021年10月26日	206	已出库 修改删除
1433732427037319168	鼠标	卢依思	陈诚录入	2021年10月26日	206	已出库 修改删除
14337324541424384	鸡蛋	卢依思	陈诚录入	2021年10月26日	206	已出库 修改删除

图 17 首页(物品管理页面)

6.2.2 B/S架构的Web应用程序

由于B/S架构具有无需安装,只要有能上电脑、手机就可使用的特点,本系统也设计了B/S架构的Web应用程序,该应用程序提供了物品管理功能、借还记录的查询功能和物品数据图表化功能等。Web端根据需求划分的功能框架图如图15所示。

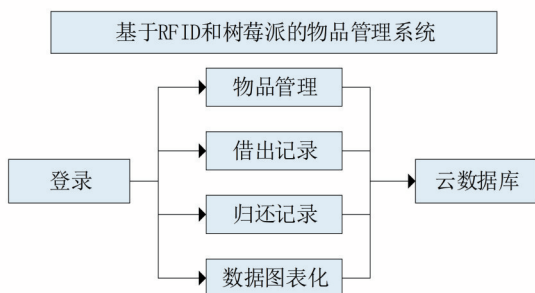


图 15 Web端软件整体架构图

Web端软件意在为用户提供一条更方便的管理物品的渠道,辅助树莓派终端软件的使用。

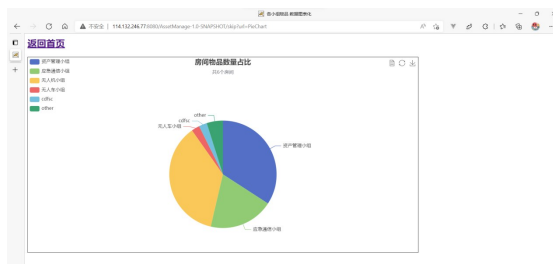


图 18 物品数据图表化

7 结语

本系统充分利用了RFID标签成本低、寿命长和数据存储量大等特点,解决了当前市面使用较多的条形码存在的问题。本研究为用户提供了两套应用软件,用户可以通过树莓派终端、PC端和手机端等设备对物品进行有序、高效的管理。

经过实践,本系统对用户的物品收纳管理起到了一定的作用,能够做到对物品数据的有效监控,大大提高了物品管理的效率,并解决

了用户对物品的闲置、遗忘问题,减少物品浪费,减轻用户压力。

参考文献:

- [1] 鲁金屏,王小锋,杨旭光. RFID技术在行政事业单位资产管理中的应用[J]. 电子制作,2021(2):80-81,66.
- [2] 唐赞淞,王红梅. 家庭物品智能管理系统方案[J]. 电信技术,2017(2):72-73,76.
- [3] 陈敏. 无线通信工程发展趋势与4G技术的简要分析[J]. 中国新通信,2019,21(6):5-6.
- [4] 郭晓玲. 移动通信4G技术的发展与展望[J]. 科技展望,2015,25(25):1,3.
- [5] 王文强. 基于射频识别的刀具信息采集与管理系统设计[D]. 廊坊:北华航天工业学院,2021.
- [6] 李飞. 基于RFID技术的设备资产管理系统设计[J]. 现代电子技术,2020,43(23):130-133. DOI: 10.16652/j.issn.1004-373x.2020.23.029.
- [7] 成志平. RFID射频技术在皮革仓储管理中的应用

研究[J/OL]. 中国皮革:1-5[2022-04-22]. <http://kns.cnki.net/kcms/detail/11.2649.TS.20220414.1047.002.html>.

- [8] 沈业,丁国刚,宋伟. 基于RFID的设备管理系统研制[J]. 电脑编程技巧与维护,2021(12):90-92. DOI:10.16184/j.cnki.comprg.2021.12.034.

作者简介:

卢依思(2001—),女,广东汕尾人,本科生,研究方向为物联网系统设计

盘茂杰(1984—),男,广东阳江人,实验师,硕士,研究方向为物联网系统设计、实验室管理技术

郑思晴(2000—),女,广东汕头人,本科生,研究方向为物联网应用软件设计

梁焕诗(1999—),女,广东广州人,本科生,研究方向为物联网应用软件设计

通信作者:魏纵横(1984—),男,广东梅州人,讲师,博士,研究方向为5G/6G通信技术、物联网技术, E-mail: ppwei17@163.com

收稿日期: 2022-05-10 修稿日期: 2022-06-22

Item Storage Management System Based on RFID and Raspberry Pi

Lu Yisi, Pan Maojie, Zheng Siqing, Liang Huanshi, Wei Zongheng

(School of Computer Science, Guangdong Polytechnical Normal University, Guangzhou 510665)

Abstract: With the development of science and technology and the progress of society, the demand for goods storage is also increasing. How to manage goods efficiently has become a problem of concern. To solve this problem, we take raspberry pie as the control module, combined with RFID and 4G technology, and developed an item storage management system to realize simple and efficient item storage. The system consists of the following parts: the perception layer composed of raspberry pie, RFID reader and tag, the transmission layer with 4G communication module, and the application layer composed of web system and cloud database. At present, the system has been put into use, which reduces the expenditure of human and material resources to a certain extent and improves the utilization rate of goods.

Keywords: RFID technology; goods management; raspberry pie; 4G technology

文章编号: 1007-1423(2022)16-0091-05

DOI: 10.3969/j.issn.1007-1423.2022.16.017

基于激光雷达的移动配送机器人自主导航研究

施建壮¹, 张国伟^{1, 2}, 卢秋红², 黄威¹, 吴松林¹

(1.上海电力大学自动化工程学院, 上海 200082; 2.上海合时智能科技有限公司, 上海 201100)

摘要: 室内移动配送机器人是一种新型机器人, 路径规划是移动机器人研究的关键技术。针对传统 A* 算法在室内移动配送机器人路径规划中存在转弯次数和遍历节点数过多, 容易发生抖动和配送速度过慢的问题, 提出一种对 A* 算法的改进策略。首先改进启发函数来减少遍历节点数、缩短搜索时间, 再结合跳点搜索理论选出关键点代替传统 A* 算法中 Openlist 和 Closelist 的点, 删除多余的拐点来提高运算效率。为了验证改进 A* 的性能, 在栅格地图上模拟出公司地图环境, 模拟室内移动配送机器人从前台配送咖啡到会议室和展厅的办公桌旁。结果表明改进的 A* 算法拐点数量更少, 搜索所耗时间更短, 能够满足公司室内场景下移动机器人配送咖啡的要求。

关键词: 室内配送机器人; A* 算法; 启发函数; 跳点搜索

0 引言

随着人工智能和物联网技术的不断发展, 服务机器人产品作为智能设备也在进行智能化转型升级, 不断实现更强大的功能和更丰富的应用^[1]。作为服务机器人的一个分支, 室内移动配送机器人, 实现了一键配送餐饮功能, 代替很多原有的重复性劳动工作。在国内外疫情爆发情况下, 室内移动配送机器人可以避免人员接触引起的感染, 实现既安全又高效的配送餐饮。并且该室内移动机器人所应用的智能导航技术能够解决在商场、酒店、餐厅、会议室等不同场所时, 遇到地毯或坡面等不平坦道路的运行稳定性和行走能力问题, 环境或者人为因素发生改变后的适应性问题等^[2]。

室内移动配送机器人利用激光雷达进行自主定位和构建地图, 再通过导航算法实现机器人自主路径规划功能。全局路径规划是移动配送机器人实现自主导航的关键技术之一^[3], 路径规划目的是在机器人获得所处环境的全局信息之后, 按照一定评价标准(如转弯次数、时间及代价等), 寻找出一条从起始节点到目标节点的最合适路径。目前, 路径规划算法已经非常成熟, 使用较多的有 A* 算法、人工势场法、

Dijkstra 算法、遗传算法和蚁群算法等^[4]。

由于 A* 算法的可塑性和移植性较强, 被应用于制造业、仓储物流等领域^[5]。但传统 A* 算法存在拐点数量和遍历节点较多、寻路时间较长等问题。目前, 国内外相关学者都对移动机器人全局路径规划方法展开了研究, 且获得一定的成果。本文以上海合时智能科技公司的室内移动配送机器人作为研究硬件平台, 对移动机器人进行室内导航的研究。

1 路径规划算法

1.1 环境地图的表示

机器人在路径规划前, 要对所处环境进行地图构建。按照维数, 可将地图划分为二维和三维地图。三维地图能够真实地展示外部环境, 但也更容易被环境影响, 且计算量大。二维地图可以分成栅格、几何地图等^[6]。栅格地图由 Elfes 和 Moraves 最先提出, 将外部环境分成多个单元格, 称为栅格, 用概率值表示栅格被占据的可能性。栅格地图是对静态环境的近似表征, 易于创建和维护, 因此本文采用栅格法来构建地图。栅格地图中路径规划如图 1 所示, 白色表示可以通行的区域; 黑色栅格表示障碍物^[7]。

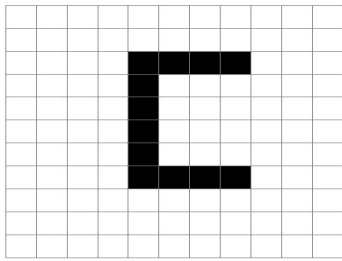


图1 栅格地图

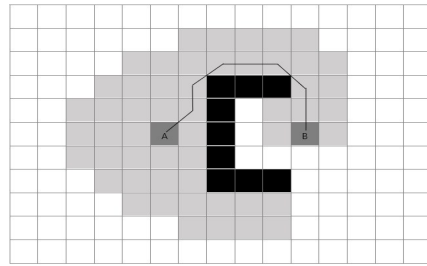


图2 传统A*寻路

1.2 传统A*算法

A*算法是在指定地图上根据移动代价来找到最优路径的启发式算法。其评价函数为：

$$f(n) = g(n) + h(n) \quad (1)$$

式中， $f(n)$ 代表从起始节点经过当前节点 n 到目标节点的估算成本； $g(n)$ 代表起始节点到 n 的实际成本； $h(n)$ 是A*算法的启发函数，代表当前节点 n 到目标节点的估算成本；在传统A*算法中都会使用欧氏距离计算 $g(n)$ 和 $h(n)$ 。

$$g(n) = \sqrt{(x_n - x_0)^2 + (y_n - y_0)^2} \quad (2)$$

$$h(n) = \sqrt{(x_k - x_n)^2 + (y_k - y_n)^2} \quad (3)$$

式中， (x_n, y_n) 是当前节点坐标， (x_0, y_0) 和 (x_k, y_k) 分别是起始点和目标点坐标。 $h(n)$ 大小可以决定A*算法的搜索效率和搜索路径的长短， $h(n)$ 的估值越小，A*需要计算的节点越多，算法搜索的效率也就越低，当 $h(n)$ 趋于0时，A*就变成了Dijkstra算法；但是当 $h(n)$ 估值远超 $g(n)$ 或者 $g(n)$ 等于0时，A*就变成了最佳优先搜索(BFS)，此时算法只追求速度，不能得到最合适的路径^[8-9]。

首先，在规划路径前需要定义两个列表：一个是Openlist列表，用来存放即将要遍历的节点；另一个是Closelist列表，用来存放已遍历过的节点。在寻路过程中，先要挑出Openlist列表中代价最低的节点添加进Closelist列表中，再把代价最低节点周围的邻节点添加进Openlist列表中。与此同时更新起始节点到各节点的代价 $g(n)$ 和各节点到目标点的代价 $h(n)$ ，依次进行迭代，直到目标点加入了Openlist列表，寻路完成。传统A*算法寻路如图2所示，黑色和白色分别表示不可通行和可通行区域，黑色线段表示路径；A表示起点、B表示目标点，浅灰色区域为规划中遍历的节点^[10]。

2 改进A*算法

传统A*算法存在遍历节点和转折点多，内存消耗严重等问题。本文首先提出改进启发函数的计算方式，再结合选取关键点替换传统A*中Openlist和Closelist的两个列表，剔除多余的冗余节点，寻找出一条更优路径。

2.1 改进启发函数

传统的A*通常是采用欧氏距离计算 $h(n)$ ，定义移动机器人是八个方向，因此欧氏距离计算出 $h(n)$ 的代价值都小于当前节点 n 到目标节点的实际代价，从而导致寻路过程中遍历的节点较多^[11-13]。

本文采用改进曼哈顿距离计算函数，传统曼哈顿距离是指两个点在标准坐标系上绝对轴距之和，其基本距离公式为：

$$h(n) = |x_n - x_k| + |y_n - y_k| \quad (4)$$

但是在本次实验地图中允许对角线移动，假设移动一个栅格的代价是 D ，对角线移动一个栅格的代价是 $\sqrt{2} * D$ 。改进后的启发函数为：

$$h1(n) = \min(|x_n - x_k|, |y_n - y_k|) \quad (5)$$

$$h2(n) = |x_n - x_k| + |y_n - y_k| \quad (6)$$

$$h(n) = \sqrt{2} * D * h1(n) + D * (h2(n) - 2h1(n)) \quad (7)$$

式中， (x_n, y_n) 是当前节点的坐标， (x_k, y_k) 是目标节点的坐标； $h1(n)$ 代表沿着对角线移动的步数； $h2(n)$ 代表曼哈顿直线距离。将 $h1(n)$ 和 $h2(n)$ 结合在一起可以得到等价于实际的路径成本，改进后的距离函数 $h(n)$ 启发性强，可以解决欧氏距离遍历的节点数多的问题，同时也可以寻找出最优的路径。

两种启发函数仿真结果对比如图3所示，欧

氏距离函数遍历节点数为75个,曼哈顿距离函数遍历节点数是48个,大大提高了搜索效率。

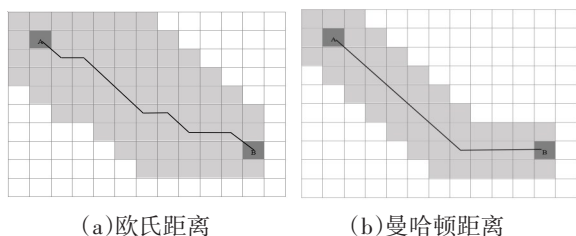


图3 两种启发函数仿真结果

2.2 结合JPS改进A*

在A*算法规划出路径前采用JPS算法跳点搜索原理进行预处理,JPS算法寻路可以分成两个部分:第一是从Openlist列表中取一个最佳节点,从几个特定方向进行搜索,将各个方向得到的跳跃点加入Openlist中;第二部分就是找到一个跳跃点。对于起始点是向所有方向进行搜索;对于其它节点,需要看父节点指向该节点的方向,再面向所有自然邻节点和被迫邻节点方向进行搜索^[14-17]。

2.2.1 无障碍物邻节点

如图4所示,x节点周围是无障碍物的,由于从父节点p(x)出发到所有灰色扩展节点的移动代价都小于或等于p(x)经过x再到灰色节点,因此不用计算灰色节点,只用裁剪规则剪去灰色节点即可。

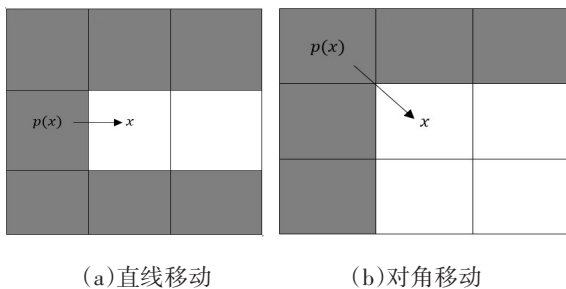


图4 无障碍邻节点

当节点附近无障碍物时,直线移动和对角移动满足不同的邻居修剪原则L()所表示的每步移动代价相同时两个节点之间的距离; $L(\langle p(x), x, n \rangle)$ 表示p(x)到达n经过x的距离; $L(\langle p(x), \dots, n \rangle | x)$ 表示p(x)到达n不经过x的

距离。

直线移动:

$$L(\langle p(x), \dots, n \rangle | x) \leq L(\langle p(x), x, n \rangle) \quad (8)$$

斜线移动:

$$L(\langle p(x), \dots, n \rangle | x) < L(\langle p(x), x, n \rangle) \quad (9)$$

2.2.2 有障碍物邻节点

如图5所示,当节点x的相邻节点有障碍物时,就不能修剪灰色节点。如果每一条从p(x)出发经过x节点后到达节点n的路径都比不经过x节点的路径短,那么节点n就是强制邻节点,强制邻节点的公式为:

$$L(\langle p(x), x, n \rangle) < L(\langle p(x), \dots, n \rangle | x) \quad (10)$$

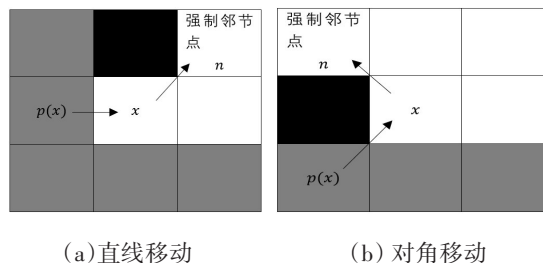


图5 障碍邻节点

2.2.3 选取跳点

一个节点是否为跳点可以从三个方面来判断:①节点y是起点或终点;②节点y在当前搜索方向的前提下有强迫邻居;③当前搜索方向是斜向,且在y节点处经过该斜向的水平分量或垂直分量方向移动可以找到跳点,那么当前节点y是跳点。

3 仿真分析

在栅格地图上模拟公司环境搭建仿真地图,将前台设置为室内移动机器人的起点A,收到送咖啡指令后将咖啡送到会议桌和展厅等不同的目标点B。在栅格地图上通过采用改进前后的A*算法进行对比,在CPU为i7-8565U@1.80Ghz的电脑上对其搜索路径的效果进行验证。从图6所示的仿真结果可以看出,改进后的A*算法拐点数量和寻路过程中遍历的节点数都大大减少,

并且在寻路时间上也有所改善。

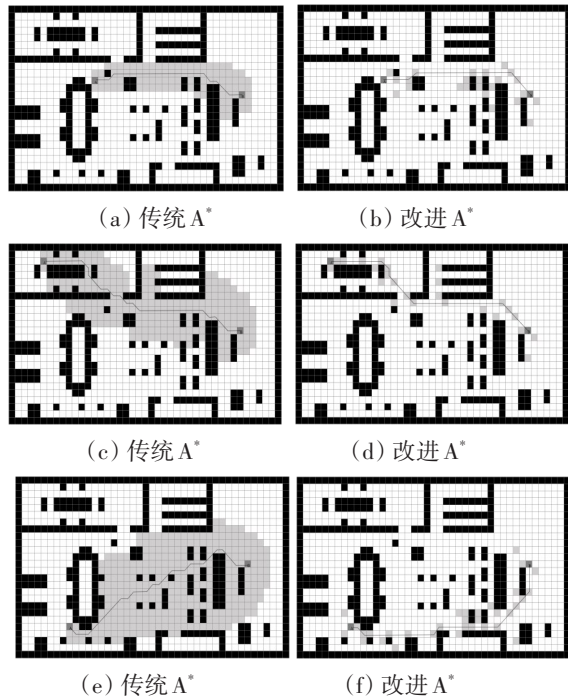


图6 改进A*前后仿真结果

表1列出了三组仿真结果中各项指标的对比,改进前后两种算法路径长度基本一样,三条路径的拐点数量相较于传统A*依次减少了50%、64%和60%,可以解决咖啡机器人运送过程中拐弯次数多的问题;遍历的节点数依次减少了75%、89%和89%,可以很大程度提高算法的搜索效率和内存的占用率。

表1 仿真结果指标对比

		拐点个	遍历节点数	搜索时间/ms	路径长度
第一组	传统A*	6	97	1.3	24.66
	改进A*	3	26	0.9	24.66
第二组	传统A*	14	235	4.2	35.73
	改进A*	5	27	2.7	35.73
第三组	传统A*	15	325	8.2	34.21
	改进A*	6	35	4.6	34.9

4 结语

针对传统A*算法在室内移动配送机器人中存在转折点和遍历节点过多等问题,通过使用改进曼哈顿作为启发函数,即在传统曼哈顿距离函数中允许对角移动,这样既可以解决欧氏距离的路径最短但运行时间长和传统曼哈顿距离搜索时间短但路径不是最短的问题,再结合

跳点搜索改进A*算法转折点过多的问题,避免运送过程中咖啡洒落和机器多次转弯发生电机堵转的问题,提高了室内移动咖啡机器人导航中的路径搜索效率,符合室内移动机器人的实际运用。

参考文献:

- [1] 王春颖,刘平,秦洪政. 移动机器人的智能路径规划算法综述[J]. 传感器与微系统,2018,37(8):5-8.
- [2] 徐美清,孙晨亮. 基于栅格地图的遗传算法路径规划[J]. 科技信息,2011(31):76-77.
- [3] 张贺,胡越黎,王权,等. 基于改进D*算法的移动机器人路径规划[J]. 工业控制计算机,2016,29(11):73-74.
- [4] 张飞,白伟,乔耀华,等. 基于改进D*算法的无人机室内路径规划[J]. 智能系统学报,2019,14(4):662-669.
- [5] 刘生伟,马钺,孟树峰,等. 改进A*算法的AGV路径规划[J]. 计算机应用,2019,39(z2):41-44.
- [6] KAVRAKI L E, SVESTKA P, LATOMBE J C, et al. Probabilistic roadmaps for path planning in high-dimensional configuration spaces [J]. IEEE Transactions on Robotics and Automation, 1996, 12(4): 566-580.
- [7] MA O, ANGELES J. Architecture Singularities of Platform Manipulators[J]. IEEE International Conference on Robotics & Automation, 2002, 10(15): 1542-1547.
- [8] 王殿君. 基于改进A*算法的室内移动机器人路径规划[J]. 清华大学学报(自然科学版),2012,52(8):1085-1089.
- [9] 林黄智,田柳. 基于改进A*算法的无人驾驶车辆路径规划方法研究[J]. 湖南文理学院学报,2021,33(2):17-20.
- [10] 张志文,张鹏,毛虎平,等. 融合改进A*算法和动态窗口法的全局动态路径规划[J]. 电光与控制,2021,28(9):39-43.
- [11] 刘子豪,赵津,刘畅,等. 基于改进A*算法室内移动机器人路径规划计算机工程与应用[J],2021,57(2):186-190.
- [12] ZAFAR M A, ZHANG Z, YU W. Mobile robots path planning based on A* algorithm improved with jump point search [C] // Proc. of the 18th International Bhurban Conference on Applied Sciences and Technologies (IBCAST). Islamabad, Pakistan: IEEE Press,2021:536-544.

- [13] 陈若男,文聪聪,彭瑜,等.改进A*算法在机器人室内路径规划中的应用[J].计算机应用,2019,39(4):1006-1011.
- [14] 杨凤满,张奇志,周亚丽,等.家庭服务机器人路径规划的跳点搜索算法[J].北京信息科技大学学报(自然科学版),2018,33(3):85-89.
- [15] 赵晓,王铮,黄程侃,等.基于改进A*算法的移动机器人路径规划[J].机器人,2018,40(6):903-910.
- [16] LIU L, YAO J, HE D, et al. Global dynamic path planning fusion algorithm combining jump-A* algorithm and dynamic window approach [J]. IEEE Access, 2021, 9:19632-19638.
- [17] 马小陆,梅宏.基于JPS策略的ACS移动机器人全局路径规划[J],2020,42(4):494-502.

作者简介:

施建壮(1995—),男,江苏连云港人,硕士,研究方向为slam路径规划

张国伟(1970—),男,山西太原人,博士,副教授,研究方向为信息检测、移动机器人

卢秋红(1973—),女,山西襄汾人,博士,正高级工程师,研究方向为计算机视觉、移动机器人

黄威(1998—),男,湖北宜昌人,硕士,研究方向为计算机视觉

吴松林(1996—),男,河南商丘人,硕士,研究方向为计算机视觉与slam构图

收稿日期:2022-03-01 修稿日期:2022-04-14

Research on Autonomous Navigation of Mobile Delivery Robot Based on Lidar

Shi Jianzhuang¹, Zhang Guowei^{1,2}, Lu Qihong², Huang Wei¹, Wu Songlin¹

(1. School of Automation Engineering, Shanghai University of Electric Power, Shanghai 200082;

2. Shanghai Heshi Intelligent Technology Co.Ltd, Shanghai 211000)

Abstract: Indoor mobile delivery robot is a new kind of robot, and path planning is the key technology of mobile robot research. Aiming at the traditional A* algorithm in the path planning of indoor mobile delivery robot, there are too many turns and traversal nodes, which easily lead to the problems of shaking and slow delivery speed. An improved strategy for A* algorithm is proposed. Firstly, the heuristic function is improved to reduce the number of traversal nodes and the search time. Then, the key points are selected to replace the Openlist and Closelist points in the traditional A* algorithm by combining the Jump Point Search, and the redundant inflection points are deleted to improve the operation efficiency. In order to verify the performance of improved A*, the company map environment was simulated on the grid map, and the robot was simulated to deliver coffee from the front desk to the desk in the conference room and exhibition hall. The results show that the improved A* algorithm has fewer inflection points and shorter search time, which can meet the requirements of coffee delivery by mobile robots in the indoor scene of the company.

Keywords: indoor delivery robot; A* algorithm; heuristic function; jump point search

文章编号: 1007-1423(2022)16-0096-05

DOI: 10.3969/j.issn.1007-1423.2022.16.018

基于 Java Swing 的找点软件设计与实现

王 勇, 洪 进, 杜兰兰, 王圣波

(空军预警学院, 武汉 430019)

摘要: 标示目标航线的过程可以简单概括为找点、连线, 其基础和核心就在于准确找到目标点迹。目前, 无论是在院校教学还是部队实训都缺少找点软件。为了弥补这一空缺, 按照由易到难、循序渐进的教学要求和认知规律, 本软件采用 C/S 架构, 利用 Java 的 GUI 图形用户接口工具 Swing, 通过绘制雷达扫描的极坐标界面, 刻画方位距离辅助线, 监听鼠标点击事件来实现在计算机上完成标示目标位置的找点训练。

关键词: 找点训练; C/S 架构; Java Swing

0 引言

标图课程包括在标图桌和标图软件中标示目标航迹, 标图软件标示目标航迹也可以叫做计算机标图, 两者的核心教学内容为找点、连线, 最终掌握标示多批目标航迹的技能。标图桌上是手写标注位置和连接多个目标点, 最终在标图板上呈现完整的航迹。计算机标图是通过鼠标点击完成找点, 由软件自动连接目标点, 最终在电脑屏幕上呈现完整的航迹, 但无论是在标图桌还是计算机上标示航迹, 首先都需要找到目标点的位置, 然后才能标注目标属性、高度、发现时间等其它相关信息, 进而连线形成航迹, 这就意味着找点训练是基础。

在标图桌上组织找点训练比较容易, 找到航迹点并进行标注, 不连线即可。但是当前计算机标图采用的软件不具备单独标示目标位置的找点训练功能, 所以教学进程安排通常是先进行标图桌上的找点训练, 达到要求后推进到标图桌上连续航迹的标示训练, 熟悉标图桌连续航迹标示之后直接进入计算机标示连续航迹的训练。从标图桌到计算机, 手工标示到软件标示是一个比较大的跨越, 需要熟悉软件各功能所在区域, 掌握软件操作技巧, 灵活使用鼠标在屏幕中快速找到航迹点, 补充其它相关信息。另外, 跨到计算机标图后只能从难度较大的标示航迹入手, 缺少了找航迹点这一基础训练环节, 计算机标图的两大跨越对于很多基础

薄弱的学员来说门槛高、起步难、训练受阻、倍受打击, 畏难情绪与日俱增, 后续的提升训练越发力不从心、恶性循环, 学员成绩不理想, 不能完成培养目标, 教员压力增加。所以, 无论是教学考核要求还是个人技能提升, 不管是从教员还是学员的角度, 降低计算机标图的门槛, 缩小标图桌标图到计算机标图的跨度, 对学员掌握标图技能至关重要。

本文旨在设计开发计算机标图找点训练软件, 为计算机标示目标航迹提供基础训练, 在软件找点训练中熟悉计算机操作。此外, 找点训练简单、易上手, 学员学习的积极性得到提升, 掌握之后再过渡到计算机标示目标航迹也不至于手足无措。

1 Java Swing 技术及软件设计

1.1 Java Swing 技术

Java Swing 技术主要是用于构建跨平台桌面窗口程序, 它是在 AWT (abstract window toolkit) 基础上采用 Java 语言编写的一种图形组件库技术^[1]。Java Swing 组件除了是轻量级的, 即在运行时不需要产生相应的“对等(Peer)”组件^[2], 实际由自身外观类支持, 没有本地代码, 不依赖操作系统, 而且它的外观是可插拔的^[3] (plug-gable look-and-feel)。Swing 组件包括各种容器和基本组件, 容器主要用来容纳和布置基本组件, 也可以包含相应的容器^[4]。常见的容器有

JFrame、JDialog、JPanel 等; 常见的基本组件包括按钮 JButton、标签 JLabel、单选按钮 JRadioButton。组件之间的关系如图 1 所示。

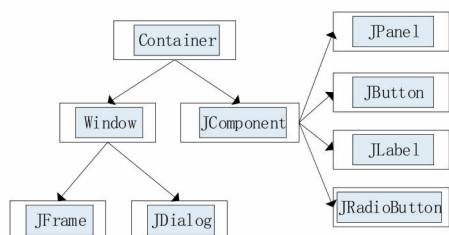


图 1 组件关系图

1.2 找点软件设计

图 2 所示为本软件的系统架构图, 分为功能界面层、基础组件层、数据存储层, 每一层的实现都依赖下一层提供的支持, 用户交互界面使用 Swing 提供的丰富基础组件来构建, 通过监听控件的鼠标事件来完成具体功能的业务处理。鉴于软件没有网络请求的需求, 可以做成单机版, 因此持久化数据采用本地文件进行存储, 访问本地文件系统避免了网络传输的耗时, 运行速度更快。

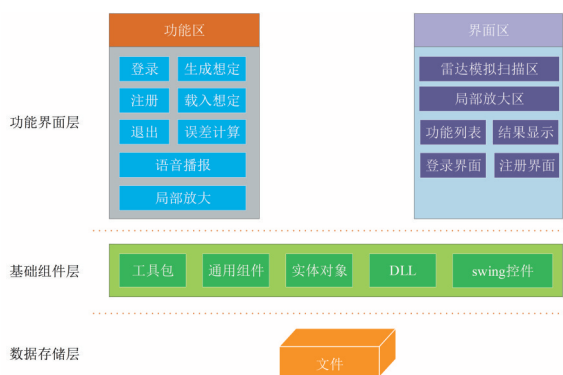


图 2 找点软件系统架构图

2 关键技术及实现

2.1 软件找点

以雷达位置为圆心, 通过编程绘图模拟雷达扫描区域, 半径是雷达最大探测距离, 软件语音播报表示目标坐标位置的方位距离数据, 根据扫描区域大小将方位距离的数值按比例转换成坐标点, 学员在模拟的扫描区域内找到该

坐标点的大致位置, 鼠标点击即可放大点击位置附近的区域, 放大区中显示方位距离辅助线, 在放大区中找到并点击方位距离的准确位置。

2.2 坐标回显

在放大区中点选目标位置之后, 放大区消失, 将放大区中点选的坐标点经过等比缩小变换后回显到极坐标中。

2.3 重绘标记

软件找点训练需要实时显示找到的点迹和对应的正确点迹, 找点过程是动态的, 所以绘制点迹也是动态的。Swing 组件绘图会先把组件中之前的内容清空, 然后开始绘制传入的新内容, 所以每次绘制标记目标的坐标值后都需要保存当前内容用于下一次重绘。

2.4 语音播报

该功能在 Java 项目中引入第三方依赖 jacob.jar, 其内部使用 JNI 封装了对动态链接库 jacob-1.18-x86.dll 中 C++ 操作系统语音功能接口的方法调用, 所以首先需要将 dll 文件拷贝到 C:\Windows\System32 目录, 文件拷贝代码如下图 3 所示, 将该段代码的执行放在软件入口处, 软件一经启动就完成 dll 文件拷贝。另外, 还需要把 dll 文件放到 Java 环境的 jre\bin 目录下, 在制作 .exe 文件时会将 jre 内置进去。

```

public void copyDLLFile2System32() {
    String sourcePath = System.getProperty("user.dir") + File.separator + "jacob-1.18-x86.dll";
    File sourceFile = new File(sourcePath);
    String outputPath = "C:\\Windows\\System32\\jacob-1.18-x86.dll";
    File outputFile = new File(outputPath);
    FileChannel inputChannel = null;
    FileChannel outputChannel = null;
    inputChannel = new FileInputStream(sourceFile).getChannel();
    outputChannel = new FileOutputStream(outputFile).getChannel();
    outputChannel.transferFrom(inputChannel, 0, inputChannel.size());
}
    
```

图 3 拷贝 dll 文件代码

想定报底中的方位距离是阿拉伯数码, 军事通信都使用军用数码, 所以需要将其转换成军用数码, 阿拉伯数码与军用数码的对应发音见表 1。转换之后, 表示目标位置的“084, 290”就变成了“洞八四 两沟洞”。另外, 为适应雷达兵部队, 语音播报要使用男声, 因此在运行软件之前还需要安装专门的语音包, 避免使用系统自带的默认女声, 安装文件 TTS\NextUp.com-NeoSpeech Chinese Wang16 Voice.msi, 选择语音

包和进行其它配置都可以通过在播报内容中添加xml标签进行设置,比如指定语音包之后播报内容就是“<voice required='Name=VW Wang'>洞八四 两沟洞</ voice >”,Name的取值对应语音包在系统注册表中的Name属性值。

表 1 军用数码对应表

数码	军用数码
0	洞
1	腰
2	两
3	三
4	四
5	五
6	六
7	拐
8	八
9	沟

根据教学进程、学习情况、熟练程度需要对想定播报的速度进行调整,播放速度通过两个参数进行控制,一个是报读文本本身内容的速度,实现方式为“<rate speed=' '></rate>”,speed的取值为-10到10,一个是报读每个目标位置之间的间隔时间,实现方式为“<silence msec=' '>”,msec单位是毫秒,比如“<rate speed='5'>洞八四 两沟洞<silence msec='3000'></rate>”表示以速率5读完“洞八四 两沟洞”之后,间隔3秒才读下一个文本内容。

2.5 坐标变换

2.5.1 坐标值→方位距离

有时找点训练结束后会对软件判定结果有所怀疑,所以每次训练都会把找点结果保存到文件中,这样可以通过校对结果与想定报底来进行验证,但是界面上找到的点迹是鼠标位置在坐标系中相对坐标原点在x轴、y轴方向上的偏移量,而报底中的数据都是方位距离格式,为了便于校对,需要把坐标值转换成方位距离数据。此外,当鼠标在界面中移动时,实时显示鼠标位置对应的方位距离信息,有助于更好地找点,这种情况也需要把坐标值转换成方位距离。

如图4所示,其中 $O\left(\frac{ScreenHeight}{2}, \frac{ScreenHeight}{2}\right)$ 表示极坐标的中心原点,即雷达所在位置; $A(X,Y)$ 表示找到的点迹; $X、Y$ 表示A在坐标系中的横纵坐标值,转换成方位距离之后为 (θ,L) ; θ 表示OA与正北方向的夹角,即方位角; L 表示线段OA的长度,即目标与雷达的距离,单位km,最大距离400km。

点O与点A横纵坐标值的差值的平方加和开根号就是L的大小,图中 ΔY 表示点O与点A在纵坐标上的差值, ΔY 与L比值的反余弦的结果就是方向角 θ ,需要注意的是,方向角从正北方向沿顺时针逐渐增加,OA与正北方向的夹角先是从 0° 增加到 180° ;当超过 180° 时,其夹角会重新从 0° 增加到 180° ,所以超过 180° 时,也就是 $X < \frac{ScreenHeight}{2}$ 时, ΔY 与L比值的反余弦结果加上 180° 才是方位角。

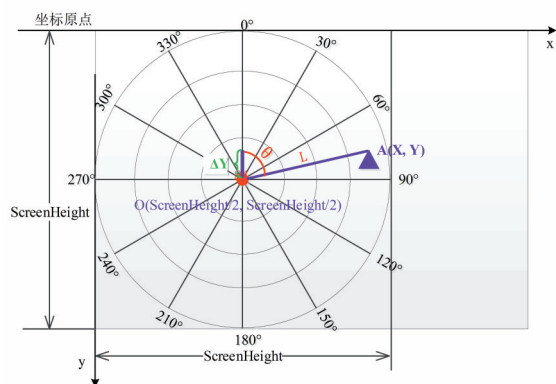


图 4 坐标值转换成方位距离

$\theta、L$ 的计算公式如下:

$$L = \sqrt{\left(X - \frac{ScreenHeight}{2}\right)^2 + \left(Y - \frac{ScreenHeight}{2}\right)^2} \quad (1)$$

$$\theta = \cos^{-1} \frac{\frac{ScreenHeight}{2} - Y}{L}, X \geq \frac{ScreenHeight}{2} \quad (2)$$

$$\theta = \cos^{-1} \frac{Y - \frac{ScreenHeight}{2}}{L} + \pi, X < \frac{ScreenHeight}{2} \quad (3)$$

2.5.2 方位距离→坐标值

在找点训练中,为了在软件界面上直观显示找到的点迹与正确报底之间的偏差,除了在界面上显示找到的点迹,还要把想定报底标注在界面上。但由于想定报底用方位距离表示目标位置,软件界面用坐标系中的坐标值表示位置,因此需要把方位距离转换成坐标值。

如图5所示, ΔX 表示点O与点A在横坐标上的差值,其它符号的含义与图4所表示的含义相同。显而易见, ΔX 与L的比值是方位角 θ 的sin值, ΔY 与L的比值是方位角 θ 的cos值, ΔX 、 ΔY 很容易计算,中心原点O的坐标值加上 ΔX 、 ΔY 的偏移量就是A点的坐标值,这样就把用方位距离(θ, L)表示的目标位置转换成了坐标值表示的A(X, Y)。需要注意的是, L表示的距离单位是km,最大取值400 km,而坐标系中L应该表示长度,单位是像素,最大取值 $\frac{ScreenHeight}{2}$,所以还要将表示km的L转换成像素,假设L在坐标系中表示的像素大小为 μ ,那么

$$\frac{L}{400} = \frac{\mu}{\frac{ScreenHeight}{2}}, \text{ 所以 } \mu = \frac{L \times \frac{ScreenHeight}{2}}{400}.$$

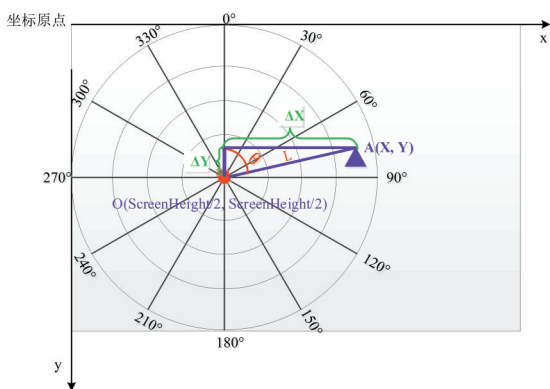


图5 方位距离转换成坐标值

ΔX 、 ΔY 、X、Y的计算公式如下:

$$\Delta X = \frac{\left(L \times \frac{ScreenHeight}{2} \times \sin \theta \right)}{400} \quad (4)$$

$$\Delta Y = \frac{L \times \frac{ScreenHeight}{2} \times \cos \theta}{400} \quad (5)$$

$$X = \frac{ScreenHeight}{2} + \Delta X \quad (6)$$

$$Y = \frac{ScreenHeight}{2} - \Delta Y \quad (7)$$

2.6 误差计算

通过三角和反三角函数将方位距离数据转换成屏幕坐标值以及坐标转换计算的结果都是浮点数类型,而Swing获取的鼠标位置都是整数数据类型,浮点数强制转换成整数会丢失小数部分的精度,所以会导致找到的点迹与想定中的对应的目标位置存在一定误差。

对于找到的点迹集合中的每一个数据,循环遍历想定中的目标位置集合,逐个计算与点迹坐标之间的距离,距离小于2 km的都算正确,以此消除找点误差。

3 运行效果

如图6所示,软件界面布局分为左右两部分,左侧是模拟的雷达扫描界面,同心圆的中心是雷达所在位置,四个同心圆分别为与雷达相距100 km、200 km、300 km、400 km的距离圈,每偏转30°绘制一条方位辅助线。右侧最上面是鼠标移动时在屏幕中的坐标实时转换成方位距离之后的信息和系统当前时间,想定属性选择区用来指定生成想定时的参数,包括生成的方位数据的要求、距离数据的要求、想定个数,想定播报区可以调整想定播放的速度和选择训练模式,贯彻由易到难、循序渐进的教育训练方针,功能区提供使用者操作的功能按钮。

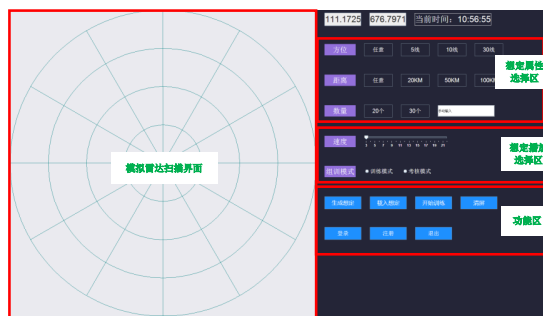


图6 软件主界面

听到语音播报的方位距离信息之后,在模拟扫描界面可以大致定位其位置,点击该位置就会显示放大区。如图7所示,放大区中

1~360°每一度都有辅助线,间隔10 km就有距离圈辅助线,按住鼠标在放大区中拖动时,放大区中黑色数据是鼠标在屏幕上的坐标值实时转换成的方位距离,随鼠标同步移动,便于快速准确找点。

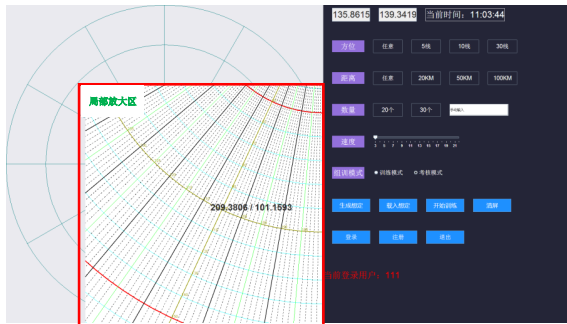


图7 局部放大

在放大区中准确找到目标位置后,鼠标点击,放大区消失,放大区中的坐标等比缩小转换后显示在模拟扫描界面,如图8所示。图中红色等边三角形表示找到的点迹,蓝色表示目标真实的位置。找点训练结束后,左下角面板中显示正确、错误、遗漏点数。

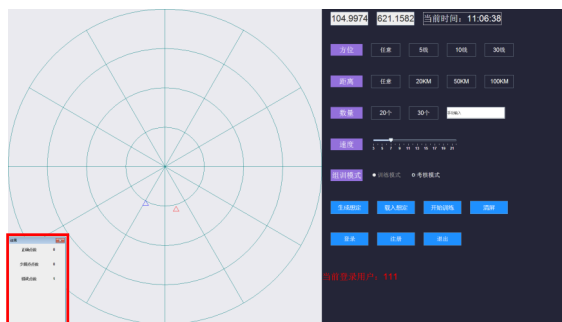


图8 结果显示

4 结语

采用Java Swing技术开发的找点训练软件,涵盖了生成想定、载入和语音播报想定、找点、误差计算等主要功能,贯穿计算机找点训练的整个过程,为计算机标示目标完整航迹提供基础训练。但是语音播报线程与结果显示线程不能异步进行,不能同步通信,存在界面显示的当前目标点与播报点不一致问题,另外,当前使用的语音包只支持Win7 32位系统,这限制了软件的使用范围,下一步考虑更换引入DLL播报语音的方式。

参考文献:

- [1] 林智扬,范明翔,陈景辉.深入浅出Java Swing 程序设计[M].北京:中国铁道出版社,2005:5-6.
- [2] 童争雄,刘特,童璧.利用Java Swing轻量级组件开发的网络流量账务系统[J].计算机工程,2002,28(12):77-79.
- [3] 杨阳.基于JAVA的护理技能大赛抽签平台的设计与实现[J].电脑知识与技术,2017,13(21):80-83.
- [4] 杨阳,汤光恒,胡明亮,等.基于Java Swing技术的学籍管理系统的设计与实现[J].福建电脑,2018,34(5):127-128.

作者简介:

王勇(1990—),男,湖北荆州人,硕士研究生,助教,研究方向为数据挖掘与大数据处理

洪进(1982—),男,湖北鄂州人,硕士研究生,讲师,从事领域为通信专业教学

杜兰兰(1994—),女,河南驻马店人,硕士研究生,助教,从事领域为英语教学

王圣波(1990—),男,湖北天门人,硕士研究生,助教,研究方向为计算机软件与理论

收稿日期:2022-03-08 修稿日期:2022-04-16

Design and Implement of Marking Software Based on Java Swing

Wang Yong, Hong Jin, Du Lanlan, Wang Shengbo

(Air Force Early Warning Academy, Wuhan 430019)

Abstract: The process of plotting target course can be summarized as marking and plotting, among which the key point is to find the target plot accurately. At present, there isn't a proper marking software neither in academy teaching nor unit training. To solve this problem, the author develops a software which can realize a step-by-step marking training. Using C/S architecture and Swing, a GUI tool of Java, the software can map the interface of polar coordinates scanned by radar, draw azimuth-range auxiliary line, and monitor mouse clicking events. In this way, students can have the marking training in computer.

Keywords: mark-training; C/S architecture; Java Swing

文章编号: 1007-1423(2022)16-0101-05

DOI: 10.3969/j.issn.1007-1423.2022.16.019

基于改进协同过滤算法的校园二手交易系统研究

许 诺¹, 高 耀^{1,2}, 周文柱¹

(1. 湖南工业大学电气与信息工程学院, 株洲 412007; 2. 湖南工业大学商学院, 株洲 412007)

摘要: 针对当下高校范围内二手物品交易领域的市场空白与现实需求, 设计并开发了一款校园二手物品交易系统“校易集市”, 通过该系统可以发布闲置物品信息、浏览他人发布的商品并购买。系统基于微信小程序云开发实现, 使用腾讯云服务取代第三方服务器, 具有无需安装和注册、开发效率高、运营成本低等优点。对传统协同过滤算法进行改进, 采用标签评价方式结合K近邻模型, 计算推荐商品, 提升了用户体验。系统在湖南工业大学上线运行数月, 用户数量持续增长, 受到广泛好评。

关键词: 协同过滤; K近邻; 微信小程序; 云开发; 校园二手市场; 交易系统

基金项目: 2021年国家级大学生创新创业训练计划项目(202111535003X): 基于Web的新型校园二手交易平台实践研究——以‘校易集市’为例

0 引言

大学生群体具有旺盛的消费需求, 但对于已购置物品的新鲜感也呈现出明显的边际递减现象, 随着网络购物成为大学生日常生活的重要部分, 校园内也随之产生了大量的闲置物品, 二手交易资源丰富^[1]。对于二手交易领域的高校细分市场, 其用户特征相似度高、市场前景可观, 开发推广针对于特定高校范围的二手交易系统可以提高交易效率, 有利于闲置资源的二次利用^[2]。

然而当下很多高校都没有专门的校园二手交易系统, 大多只能依靠QQ群、微信群等网络媒介进行交易, 而这些交易渠道无法及时进行信息过滤和整理, 导致交易过程耗时长、效率低^[3]。而传统的校园“跳蚤市场”没有固定的时间和地点, 不能理想地满足在校大学生日常的闲置物品交易需求。因此, 开发一款专门的校园二手交易系统很有必要, 它不仅能随时随地为在校大学生提供便捷的二手物品交易支持服务, 还能够培养出绿色环保的校园文化氛围^[4]。

1 系统架构

本系统采用微信小程序开发实现, 微信小

程序相较于手机APP, 具有无需安装和注册、开发成本低、效率高等优点。“校易集市”微信小程序采用云开发技术, 使用腾讯云服务取代第三方服务器, 借助云开发提供的丰富的云端API, 编写并上传部署云函数到云端, 实现数据的存储传递与调用。小程序云开发是集成于小程序控制台的原生Serverless云服务, 为开发者提供了完整的原生云端支持和微信服务支持, 底层由腾讯云提供技术支持, 以确保数据安全、服务稳定^[5]。

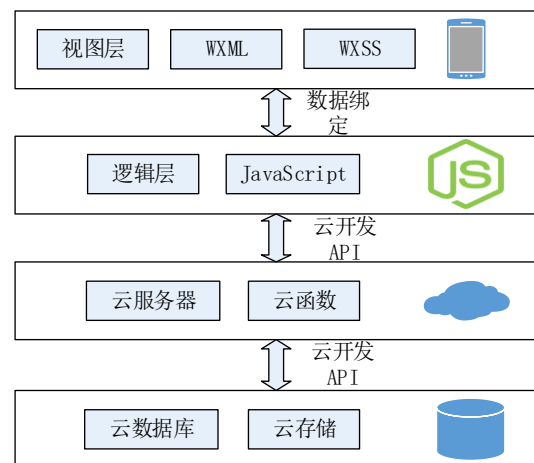


图 1 系统架构图

系统架构如图1所示。小程序的视图层由WXML与WXSS编写实现，该层的作用是将逻辑层的数据渲染成前端视图，同时将视图层的事件通过数据绑定发送给逻辑层。逻辑层接收视图层的事件反馈，并使用JavaScript进行逻辑运算，返回给视图层；逻辑层还可以通过云开发API调用部署在云服务器上的云函数，云函数运行在云端的Node.js环境中，不占用终端内存，各个云函数之间相互独立且可以相互调用，云函数SDK中的数据库和存储API提供了对数据库和存储的相关操作。云开发提供了一个JSON数据库，该数据库支持多种数据类型，可以通过数据库API进行增删改查操作；云存储支持视频、图片等文件的上传、下载、删除、移动、搜索等操作，由相应的存储API进行调用^[6]。

2 改进算法

2.1 面向标签评价

传统的协同过滤算法存在许多不足，例如因用户评价打分的主动性较低、基于微信小程序的电商平台规模较小等原因，在实际情况中很难获得大量用户评价，导致推荐系统的准确度较低^[7]；以及忽略商品属性，不能精确地推荐具有某一属性的商品^[8]。针对上述问题，设计一种面向商品标签的评价方式，提取出商品类别、价格范围、卖家信誉等若干商品标签，若用户对商品的某一方面满意则点赞相应标签，否则点踩相应标签。该评价方式取消了打分制，简化了评价流程，同时又实现了针对性评价。

用户 U_i 对 m 个商品 I_1, I_2, \dots, I_m 的 l 个标签 T_1, T_2, \dots, T_l 评价，规定点赞为1分，点踩为-1分，不评价为0分，第 j 个商品的第 k 个标签评价分值为 s_{jk} ，得到商品-标签评价表，如表1所示。

表1 商品-标签评价表

	T_1	T_2	...	T_k	...	T_l
I_1	s_{11}	s_{12}	...	s_{1k}	...	s_{1l}
I_2	s_{21}	s_{22}	...	s_{2k}	...	s_{2l}
...
I_j	s_{j1}	s_{j2}	...	s_{jk}	...	s_{jl}
...
I_m	s_{m1}	s_{m2}	...	s_{mk}	...	s_{ml}

用户 U_i 对商品 I_j 的标签向量为 $(s_{j1}, s_{j2}, \dots, s_{jl})$ ， e_{ij} 为标签向量的单位向量，如式(1)所示； N_{ij} 为标签向量的模，如式(2)所示。

$$e_{ij} = \left(\frac{s_{j1}}{N_{ij}}, \frac{s_{j2}}{N_{ij}}, \dots, \frac{s_{jl}}{N_{ij}} \right) \quad (1)$$

$$N_{ij} = \sqrt{\sum_{k=1}^l s_{jk}^2} \quad (2)$$

由式(1)及商品标签表计算 e_{ij} ，得到全体用户 U_1, U_2, \dots, U_m 对全部商品的评价，评价以不同方向的单位向量表征，组成商品-用户评价表，如表2所示。

表2 商品-用户评价表

	U_1	U_2	...	U_i	...	U_n
I_1	e_{11}	e_{21}	...	e_{i1}	...	e_{n1}
I_2	e_{12}	e_{22}	...	e_{i2}	...	e_{n2}
...
I_j	e_{1j}	e_{2j}	...	e_{ij}	...	e_{nj}
...
I_m	e_{1m}	e_{2m}	...	e_{im}	...	e_{nm}

E_j 是关于商品 I_j 的 n 个单位向量的和，如式(3)所示。其单位向量 $\frac{E_j}{|E_j|}$ 表征在 l 个标签维度上，全体用户对商品 I_j 的综合评价。

$$E_j = \sum_{i=1}^n e_{ij} \quad (3)$$

2.2 商品推荐策略

对于两件商品 I_1, I_2 ，设其商品评价分别为 E_1, E_2 ，采用余弦相似性方法计算两件商品的相似度 $\text{sim}(E_1, E_2)$ ，如式(4)所示^[9]。由于采用了基于商品标签的评价方式，不存在常规打分方式中不同用户可能使用不同评分标准的情况，提高了评价数据的可信度。

$$\text{sim}(E_1, E_2) = \cos(E_1, E_2) = \frac{E_1 \cdot E_2}{\|E_1\| \cdot \|E_2\|} \quad (4)$$

对于商品 I_j ，选取 k 个与其最相似的商品构成 K 近邻集合 T ，如式(5)所示^[10]。对于未评价过商品 I_j 的用户 U_i ， P_{ij} 表示 U_i 对 I_j 评价的预测向量，如式(6)所示，其中 $\frac{E_j}{|E_j|}$ 为全体用户对商

品 I_j 综合评价的单位向量, $\frac{\mathbf{E}_s}{|\mathbf{E}_s|}$ 为全体用户对 K 近邻集合 T 中商品 I_s 综合评价的单位向量, $\text{sim}(j, s)$ 为 I_j 与 K 近邻集合中 I_s 的相似度, \mathbf{e}_{is} 为 U_i 对 I_s 评价的单位向量。

$$T = \{I_1, I_2, \dots, I_s, \dots, I_k\} \quad (5)$$

$$P_{ij} = \frac{\mathbf{E}_j}{|\mathbf{E}_j|} + \frac{\sum_{s \in T} \text{sim}(j, s) \cdot \left(\mathbf{e}_{is} - \frac{\mathbf{E}_s}{|\mathbf{E}_s|} \right)}{\sum_{s \in T} |\text{sim}(j, s)|} \quad (6)$$

由式(6)可知 $\frac{P_{ij}}{|P_{ij}|}$ 为 l 个维度的单位向量, 将其 l 个维度取平均值, 得到表征用户 U_i 对商品 I_j 评价的预测值 r_{ij} , 且 $r_{ij} \in (-1, 1)$, 设置推荐阈值 R , 当 $r_{ij} > R$ 时, 将商品 I_j 推荐给用户 U_i 。

3 系统设计

系统以高校大学生为服务对象, 具体高校为服务范围, 主要提供二手物品交易支持。根据需求分析可以确定系统主要功能如图2所示。

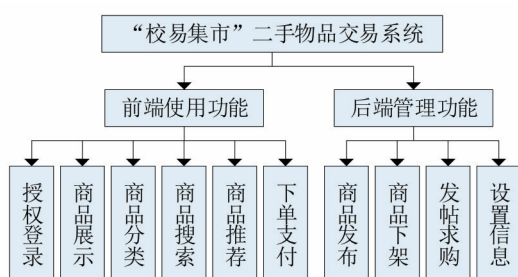


图2 系统功能设计

3.1 发布商品

如图3所示, 用户在商品发布页面填写商品信息, 点击“发布商品”后, 视图层将数据传给逻辑层进行处理, 逻辑层调用云函数将商品数据插入商品信息表, 创建一条商品记录。商品图片以文件形式储存在云存储中, 上传成功后返回一个文件唯一标识符, 也即文件ID, 文件ID被存储在数据库中, 可根据文件ID访问或删除商品图片。

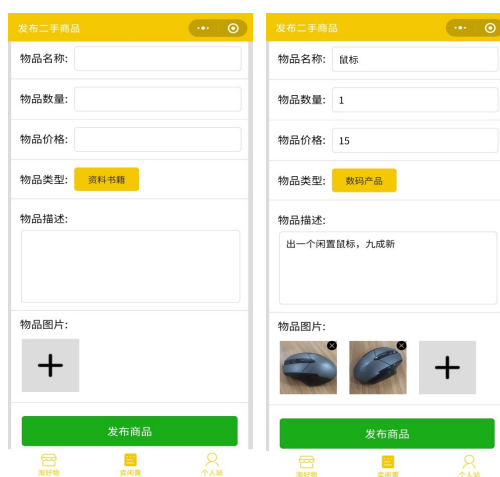


图3 商品发布界面

3.2 搜索与分类商品

如图4所示, 用户在搜索框输入关键字进行查询, 通过调用云函数将关键词与商品信息表的“goods_name”字段进行匹配, 通过模糊查询的方式, 返回所有与关键字有关的结果。

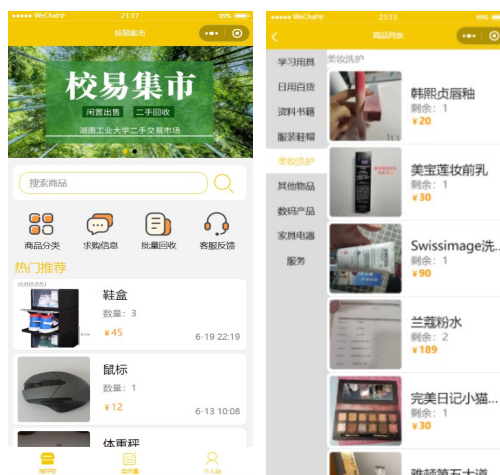


图4 商品搜索与分类界面

在商品分类界面, 通过云函数将商品信息表中的记录按不同的“goods_type”字段进行查询, 并分别存放在对应类别的列表中并返回到逻辑层, 最终按类别显示在视图层中。

3.3 推荐商品

采用简化后的标签评价方式获取用户对商品的评价信息, 根据改进协同过滤推荐算法计算全体用户对每个商品的综合评价 $\frac{\mathbf{E}_j}{|\mathbf{E}_j|}$, 继而采

用余弦相似性方法计算商品间的相似度 $\text{sim}(E_1, E_2)$ ，选取相似度最高的 k 个商品构成 K 近邻集合 T ，计算 U_i 对 I_j 评价的预测向量 P_{ij} 与预测值 r_{ij} ，若 r_{ij} 大于推荐阈值 R ，则将商品 I_j 推荐给用户 U_i 。

3.4 购买商品

买家发现心仪的商品，在与卖家沟通咨询后便可下单，买家支付购物款暂存到系统账户中。待买卖双方线下交接后，买家确认收货，系统将购物款扣除一定的服务费后转给卖家，若订单未被确认，系统会提醒买家确认交易进度；若双方确认交易终止，系统会将购物款退回家账户。

4 系统应用

4.1 后台管理

微信开发者工具的云开发控制台中提供了协助运营分析的后台管理模块，可以在该模块下查看任意时间段的流量资源使用情况、用户访问量、数据库读写次数等数据指标。云开发控制台也提供了对数据库、存储、云函数等模块进行相应操作的功能。为了进一步提高后台管理效率，云开发服务提供了CMS内容管理系统，帮助开发者管理数据库以及云存储。如图5所示，在内容管理系统上，后台管理员可以快捷地设计数据库表，新增、检索或批量导入导出数据记录。

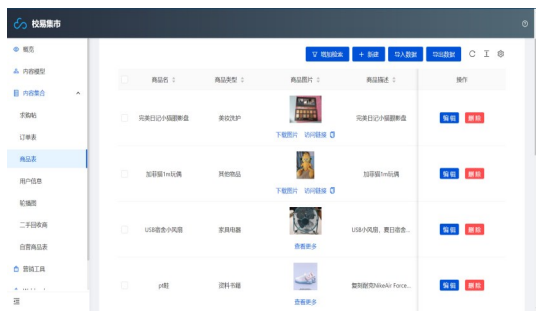


图5 CMS内容管理系统

4.2 应用成效

“校易集市”二手物品交易系统自上线以来，已在湖南工业大学范围内运营数月，累计用户数百人，图6为2021-05-15至2021-06-15

期间用户增长统计图。为了了解用户对推荐系统的满意度情况，将用户满意度量化为1至5星标准，随机对200名用户进行问卷调查，回收问卷167份，用户评价统计如图7所示，由图7可以看出系统基本满足了在校大学生进行二手物品交易的实际需求，受到普遍好评。

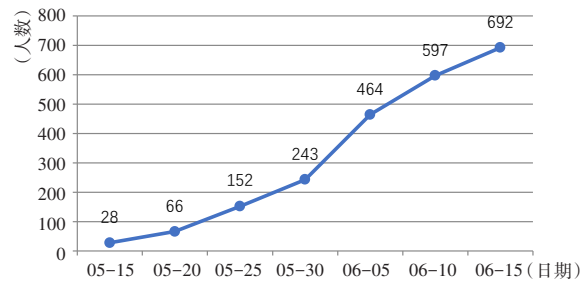


图6 用户增长曲线

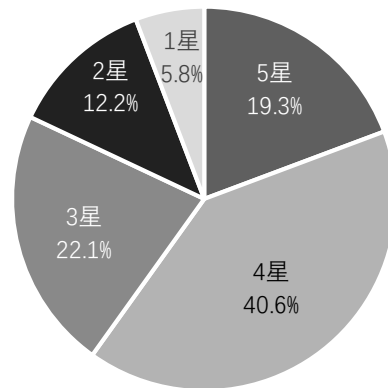


图7 用户评价统计图

5 结语

本文从高校二手市场需求出发，提出了一种便捷安全的高校二手物品交易系统。从系统架构、系统功能等方面论述了该系统的可行性，同时针对商品标签评价方式，提出了一种改进的协同过滤算法，为用户推荐其可能喜爱的商品，提升了用户体验，最终由系统的实际应用证实了其可靠性与现实意义。基于改进协同过滤算法的二手物品交易系统有效地降低了交易门槛，提供了一种更加高效的交易途径，可以满足在校大学生的日常需求，减少闲置资源的浪费，促进校园二手交易生态链的发展。

参考文献:

- [1] 任晓莉,雷洪铸.二手电商平台“闲鱼”信用问题及发展对策研究[J].商场现代化,2021(11):17-19.
- [2] 王祖维,孟彦霖,李晖.构建校园二手交易平台的可行性分析[J].现代商贸工业,2018,39(35):81-82.
- [3] 张子欣,高敏,王斯宇,等.校园内部二手交易平台研究[J].合作经济与科技,2020(9):84-87.
- [4] 刘子义,王旭泽,钟泽耀,等.绿色发展背景下高校二手交易市场现状发展研究:以中国民航大学二手平台为例[J].中国商论,2021(2):57-60.
- [5] 廖勇,于金法.基于微信小程序的教学设备维护系统设计与实现[J].现代计算机,2022,28(2):112-116.
- [6] 马静.基于微信小程序的购物商城系统的设计与实现[J].微型电脑应用,2021,37(3):31-34.
- [7] 魏浩,刘小豫,张伟.一种改进用户相似度的协同过滤推荐算法[J].电子设计工程,2021,29(17):30-34.
- [8] 谢锐兵.基于协同过滤推荐的高校智慧党建预警机制构建[J].现代计算机,2021,27(22):60-65.
- [9] 刘勇,李永杰.基于协同过滤算法的微信小程序智能助手[J].计算机系统应用,2019,28(5):71-76.
- [10] 王颖,王欣,唐万梅.融合用户自然最近邻的协同过滤推荐算法[J].计算机工程与应用,2018,54(7):77-83.

作者简介:

许诺(2001—),男,河南信阳人,在读本科,研究方向为软件开发

高耀(2001—),男,江西抚州人,在读本科,研究方向为市场营销

周文柱(2001—),男,河北张家口人,在读本科,研究方向为软件开发

收稿日期:2022-03-09 修稿日期:2022-04-14

Research on Campus Second-Hand Trading System Based on Improved Collaborative Filtering Algorithm

Xu Nuo¹, Gao Yao^{1,2}, Zhou Wenzhu¹

(1. School of Electrical and Information Engineering, Hunan University of Technology, Zhuzhou 412007;

2. Business school, Hunan University of Technology, Zhuzhou 412007)

Abstract: Aiming at the current market gap and real needs in the field of second-hand goods trading within the scope of colleges and universities, a campus second-hand goods trading system ‘School Market’ is designed and developed. Through this system, you can publish information on idle items, browse the products released by others and buy them. The system is developed and implemented based on the WeChat applet cloud, using Tencent cloud services to replace third-party servers, which has the advantages of no installation and registration, high development efficiency, and low operating costs. The traditional collaborative filtering algorithm is improved, the label evaluation method is adopted, and the K-nearest neighbor model is used to calculate the recommended products, which improves the user experience. The platform has been running for several months in Hunan University of Technology, and the number of users has continued to grow, which has been widely praised.

Keywords: collaborative filtering; K-nearest-neighborhood; WeChat applet; CloudBase; campus second-hand market; trading system

文章编号: 1007-1423(2022)16-0106-05

DOI: 10.3969/j.issn.1007-1423.2022.16.020

公共资源交易主体与成交业绩查询系统的设计

付宏燕

(国家信息中心, 北京 100045)

摘要: 主体关联信息、成交业绩信息在公共资源交易业务中的流动和合理使用, 能够简化招投标流程, 提升交易效率。从交易业务应用出发, 突出关联信息资源的整合共享并服务于交易业务, 设计公共资源交易主体与成交业绩查询的整体架构与主要功能, 总结创新点与应用成效。

关键词: 公共资源交易; 主体关联信息; 成交业绩信息; 查询系统

0 引言

《中共中央国务院关于加快建设全国统一大市场的意见》提出, 营造稳定公平透明可预期的营商环境是全国统一大市场建设进程中的关键一环, 也是主要目标之一^[1]。充分发挥市场在资源配置中的决定性作用, 是社会主义市场经济的基本要求, 也是营商环境建设的首要目标。市场主体在市场三要素中居于主导地位, 是市场经济的主要参与者。聚焦市场主体关切, 激发市场主体发展活力, 是优化营商环境、提升服务效能的有效措施。

公共资源交易市场主体指在公共资源交易市场从事交易活动的组织和个人。交易市场主体类型以企业为主, 据统计超过 138 万企业参与过交易活动, 另有超过 55 万的事业单位、机关、社团等组织一并活跃在交易市场。随着区域市场一体化的推进和全国统一大市场的提出, 跨地域交易行为越来越多, 而省外交易信息的获取渠道较少, 各省交易机构对省外主体的投标信息、资质和成交业绩证明难以评判; 同时, 市场主体也因获取省外交易信息不及时, 影响到参与省外投标的积极性。全国公共资源交易市场主体基本信息和成交业绩面向社会公开, 可很大程度解决上述问题。本文从交易业务应用出发设计, 充分考虑了关联信息资源的整合共享并服务于业务, 为政务信息资源从业务角

度促进整合共享提供一些参考。

1 公共资源交易市场主体的关联范围

1.1 交易市场主体业务特征

公共资源交易市场主体是交易活动的发起者、组织者, 也是交易行为的直接参与者、实现者, 而交易行为通过交易流程不同阶段留痕的数据体现。按照招标、投标、评标、中标的交易过程, 将不同交易角色的市场主体分类, 可以分为招标人、交易机构、投标人、中标人^[2]。在以往的交易流程中, 招标人、投标人等基本信息、交易信息等均由交易机构人工录入, 错误较多质量较差; 同时, 投标业绩由投标人自行提供, 评标专家难以判断真伪。

组织机构类型中的企业, 在交易中占比高、影响大, 其企业基本信息在市场监管总局均有备案登记, 并且在政务信息共享平台上提供共享接口; 部分有资质的企业在主管部门也有备案登记的企业资质信息。这些权威部门的登记信息若用于公共资源交易的招投标中, 可提供权威依据, 简化交易流程, 提升招标、投标、投标效率。

1.2 交易市场主体关联查询范围

公共资源交易较成熟的领域有工程建设招投标、政府采购、土地使用权和矿业权出让、国有产权(实物)交易、国有产权(股权)交易,

其中工程建设招投标和政府采购是政府支出的主要途径，这两个领域的成交主体数量约占总主体数量的81%，成交项目数量约占总项目数量的83%。根据公共资源交易分类类目，工程建设招投标领域的交易项目包括房屋建筑、市政、公路、铁路、民航、水运、水利、能源、邮电通信、桥梁、城市轨道交通、信息网络、工业制造等；政府采购领域的交易项目包括货物类、工程类、服务类。由分类知，建筑类企业在两个领域的交易活动中有绝对高占比，所以主体

关联信息应包括住建部备案登记建筑类企业资质信息、市场监管总局的企业基本信息、信用中国的企业信用信息，以及最主要的分交易领域主体成交记录信息。

2 系统整体架构与主要功能设计

2.1 系统整体架构设计

系统整体架构分为数据来源层、交换层、数据治理层、应用层，如图1所示。

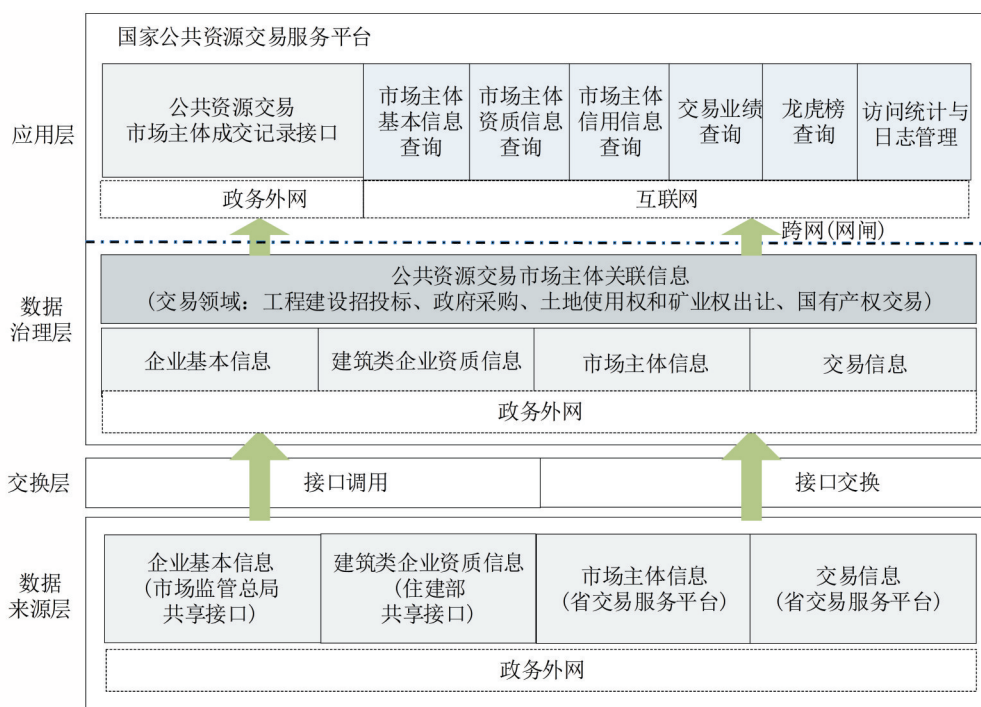


图1 系统整体架构图

数据来源层包括公共资源交易市场主体信息和交易信息(31省和兵团交易服务平台)，以及政务信息共享平台上的企业基本信息(市场监管总局接口)、建筑类企业资质信息(住建部接口)。其中，公共资源交易市场主体信息和交易信息由省交易服务平台经统一的API和Webservice对接接口交换至国家公共资源交易服务平台；政务信息共享平台上的接口调用，将公共资源交易主体的统一社会信用代码作为输入条件，将相关信息调取到国家公共资源交易服务平台对应数据表中。在数据治理层，对交

易信息中的交易额进行大额数值的核验和异常处理后，以市场主体统一社会信用代码为唯一主键关联所有表的字段，实时获取各表数据，形成公共资源交易市场主体关联信息表，用于综合查询服务，避免大量查询时不同表相互关联的时间影响。应用层使用治理后的关联信息，查询功能部署在互联网，嵌入全国公共资源交易平台的“数据服务栏目”<http://data.ggzy.gov.cn>面向社会公众公开；同时，在政务外网部署了公共资源交易市场主体成交记录接口，面向政务系统用户提供接口服务。

2.2 功能结构设计

2.2.1 主要功能设计

系统主要功能设计如表1所示,由一级功能、二级功能、三级功能组成,上一级功能可往下钻,业务查询逐级细化,这些主要功能部署在互联网。

表1 主要功能设计

一级功能	二级功能	三级功能
主体信息一站式查询	主体基本信息查询	
	主体资质信息查询	
	主体信用信息查询	
	主体交易信息查询	
主体交易业绩查询	主体成交总量查询	交易领域查询
		交易量查询
		交易额查询
	主体分领域成交总量查询	分领域交易量查询
		分领域交易额查询
	主体成交项目信息查询	成交项目概要信息查询
交易龙虎榜	分领域分年TOP100主体排名	成交项目中标公示查询
访问设计与日志管理	访问量统计	任意时间段用户量统计
		任意时间段页面访问量统计
	日志管理	

2.2.2 主体信息一站式查询

系统提供市场主体基本信息、资质信息、信用信息、交易信息的一站式查询,输入主体名称或统一社会信用代码即可支持模糊查询、精准查询。交易信息主体范围为公共资源交易市场主体;主体基本信息、资质信息、信用信息的主体范围并不局限于公共资源交易市场主体,覆盖了企业、事业单位、政府机关、社会组织。其中市场主体登记信息,包含统一社会信用代码、法定代表人、注册资金、经费来源、住所、业务范围等基本信息;资质信息包含了住建部核准的建筑类资格资质企业;信用信息与信用中国网站一致,采用了对接信用记录数量接口、详细内容跳转信用中国网站的方式。

2.2.3 交易业绩查询

交易业绩查询针对公共资源交易市场主体。系统对已归集的全中国公共资源交易信息,治理后统一对外提供服务,包括该主体成交总量(交易量、交易额)情况,及该主体作为中标人在工程建设招投标、政府采购、土地使用权出让、矿业权出让、国有产权交易五大领域参与的成交项目情况,如招标方、项目名称、中标日期、中标金额,支持下钻至成交项目中标公示,支持原文链接到该中标公告首次发布的地址。

2.2.4 交易龙虎榜

交易龙虎榜功能分年度分领域(工程建设招投标、政府采购、土地使用权出让、矿业权出让、国有产权交易领域),按照中标或成交次数进行了交易主体前100的排名,并且提供对每个TOP主体的成交项目概要,支持每笔成交信息的下钻,直到成交详细信息。当年的TOP主体排序根据当年主体成交项目数量的增加而动态变化。

2.2.5 访问统计与日志管理

系统可统计任意时间段内的UV、PV情况,支持总值、均值及导出,方便对比,也支持每个页面的访问情况统计、按主体访问情况的统计等。通过UV、PV的统计值,可以分析用户关注领域、关注热点等行为。同时,日志记录了用户访问时间、操作细节,支持按时间搜索和关键字的搜索等。

2.2.6 数据接口服务补充

除了互联网的查询功能,针对地方交易部门在政务外网有批量使用主体交易信息的需求,设计了部署在政务外网的公共资源交易市场主体成交记录接口服务,无条件向地方交易部门开放。接口以主体名称或统一社会信用代码为输入条件,提供市场主体在五大交易领域的成交记录信息,同时设计了分用户动态调用策略、日调用量监测、分时间段调用量统计等功能,并及时汇总地方使用接口数据的业务场景。

2.3 创新点设计

(1) 设计了可回溯的多源异构数据整合与共享并实时同步,实现了全国市场主体基本信

息、建筑类资质信息、信用信息、交易信息的一站汇聚,两部委两平台多来源数据整合(市场监管总局、住建部、信用中国、国家公共资源交易服务平台),31省交易部门多流向共享应用,可回溯源头、可统计流向,提升数据跨省流动的价值。

(2)设计了四大交易领域成交信息的一键核查,涵盖了全国市场主体在工程建设招投标、政府采购、土地使用权和矿业权出让、国有产权交易领域所有成交信息,促进交易公开透明,提供评标业绩依据,提升招投标业务效率。交易业绩龙虎榜激励市场主体积极参与,一并提供了大额交易监管窗口。

(3)设计了结合业务需求的安全策略,如互联网的反爬策略,根据正常交易业务查询时间,设计为控制5分钟内访问量不得超过100次;当天访问量不得超过1000次(每日0点起算)。否则将进入实时IP黑名单。IP黑名单每分钟更新一次,每天超过1000次进入黑名单的IP当天不允许继续访问。

3 系统应用成效

公共资源交易主体与成交业绩查询系统整合了2019年以来全国公共资源交易五大领域(工程建设项目招投标、政府采购、土地使用权和矿业权出让、国有产权交易)超过400万个成交项目和80万家市场主体信息,为社会公众和政务部门提供综合查询服务。据统计,互联网功能日最大访问量513.4万次,2021年月度日均访问量如图2所示。

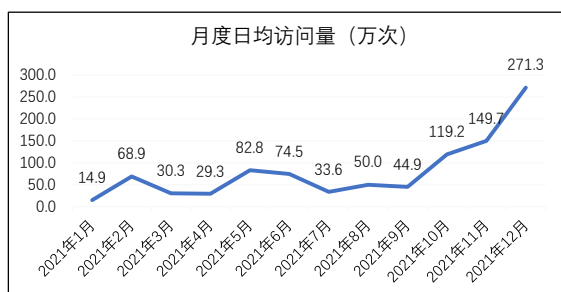


图2 2021年月度日均访问量趋势图

政务外网的数据接口服务2021年支撑政务部门调用超100万次,涉及部委用户3个、国家级平台用户1个、地方用户41个(含省级和地市级)。其中全国信用信息公共服务平台将成交业绩信息用于“信易贷”业务,2021年实际支持邮储银行的中小企业融资信用评价超5000次;南京、云南等地已将接口调用结果用于招投标,简化投标企业的业绩证明材料提交。

4 结语

公共资源交易市场是社会主义市场经济体系中的重要组成部分,市场主体是最积极最活跃的因素。主体关联信息、成交业绩信息在公共资源交易业务中的流动和合理使用,促进了交易业务的公开透明,能够简化招投标流程,提升交易效率,推进投标材料无纸化的发展。本文的设计充分利用信息资源整合服务于交易业务应用,以主体组织的统一社会信用代码或名称作为输入条件,系统简单易用并产生了一定的社会价值;而鉴于主体中自然人数量少且身份信息敏感,从安全考虑不宜作为查询对象,期待后续有合适的方法将自然人信息与业绩纳入使用范围。

参考文献:

- [1] 马晓白,陈功,张力.建设全国统一大市场关键在优化营商环境[EB/OL].(2022-04-28) http://www.jjckb.cn/2022-04/28/c_1310577041.htm.
- [2] 付宏燕.公共资源交易数据仓库的构建[J].现代计算机,2022,28(5):46-51.

作者简介:

付宏燕(1974—),女,河北石家庄人,硕士,高级工程师,从事领域为国家电子政务工程项目的技术管理与应用推进

收稿日期:2022-05-10

修稿日期:2022-06-12

Design of Market Entities and Trading Results Query System in Public Resource Trading

Fu Hongyan

(State Information Center, Beijing 100045)

Abstract: The entities related information and trading results information can simplify the bidding process and improve trading efficiency, if they are used reasonably in public resource trading. Based on the application, the integration and sharing of related information resources is proposed for trading business. The structure and functions of market entities and trading results query system in public resource trading are designed, and the innovation and application effects are summarized.

Keywords: public resources trading; entities related information; trading results information; query system

(上接第 83 页)

Application of Big Data Visualization in Urban Planning

Cao Xinhui¹, Cao Jinjin², Xu Lijuan¹, Gong Jiaming¹, Zhao Qian¹

(1. Guangzhou Huashang College, Guangzhou 511300;

2. Design and Research Institute of Second Navigation Engineering Bureau Co., Ltd., Wuhan 430000)

Abstract: Urban planning is an important guarantee for the development of a city and the development goal of a city in a certain period of time. The large-scale and multi-dimensional characteristics of the data make the traditional urban planning method difficult to bear this task. At the same time, smart city and information construction continue to develop rapidly, and big data visualization has been widely used in the practice and research of urban planning industry. First, it highlights the necessity and applicability of big data visualization for urban planning. Secondly, through examples, we can see that the quality and efficiency of urban planning scheme based on big data visualization have been significantly improved. Finally, the future development direction and challenges of big data visualization are discussed.

Keywords: big data; urban planning; big data visualization; traffic big data visualization; environment sound big data visualization

文章编号: 1007-1423(2022)16-0111-04

DOI: 10.3969/j.issn.1007-1423.2022.16.021

基于图数据库的阅读行为知识图谱构建研究

陈光仪, 陈义明, 吴小慧

(湖南农业大学信息与智能科学技术学院, 长沙 410128)

摘要: 智慧图书馆是全社会形成良好智慧学习环境的重要基础, 其应能为读者提供个性化和智能化的阅读服务。论文研究了如何获取读者在线阅读时的行为习惯并基于图数据库构建对应知识图谱的相关内容。首先, 对比分析了目前主流的知识图谱构建方法及层次构架; 然后结合自底向上的构建方法, 按照知识获取、知识表示、知识存储和知识可视化的步骤, 研究并逐步实现了基于图数据库的读者阅读行为知识图谱的构建。所得成果为后续进一步研究图书馆如何根据读者实际需求提供智慧化的阅读服务奠定了重要基础。

关键词: 知识图谱; 阅读行为; 图数据库

基金项目: 湖南省教育厅科学研究项目(18C0143)

0 引言

语义网络作为人工智能的重要应用领域之一, 可以给用户提供一个更加准确、更加智能的知识获取环境。而知识图谱是实现语义网络的技术基础, 是通向语义网络环境的鲜明道路^[1]。在智慧学习的大环境下, 叠加近年来新冠疫情的防控需求, 在线阅读已越来越多地成为广大读者的首选阅读方式。如果能够有效获取读者的阅读行为并构建对应的知识图谱, 对于图书馆而言, 可以及时了解其在阅读过程中的实际需求, 继而进行针对性的阅读指导并为读者推荐个性化的阅读内容; 对于出版商而言, 可以及时调整、改进电子出版物的内容编排及后续再版工作, 以更好地适应目标读者群体的实际需求。因而, 此项研究工作对于进一步提升读者的阅读学习效果, 完善图书馆的智慧化阅读服务, 推动促进全社会形成良好的智慧学习环境大有裨益。

1 知识图谱构建技术

构建知识图谱有自顶向下和自底向上两种方式。前者通常是指基于百科类网站等高质量的结构化数据源, 从中提取本体和模式信息后再加入到知识库中, 因而适用于那些内容明确、

关系清晰的领域知识图谱构建; 而后者是指通过借助特定的技术手段从公开采集的数据中提取模式信息, 选择其中置信度较高的新模式, 经人工审核后再加入到知识库中^[2]。目前大部分知识图谱的构建都采用自底向上的方式, 其层次架构按照知识获取的过程可分为信息抽取、知识融合和知识加工。

信息抽取是指从多源异构的数据源中提取出实体、属性以及实体之间的关系, 在此基础上形成本体化的知识表达, 它是知识图谱构建技术的关键^[1]。早期信息抽取主要是基于预定规则的抽取技术, 工作量庞大且仅适用于特定的专业领域, 后来人们开始尝试使用统计机器学习的方法, 通过标注部分数据得到训练集, 在此基础上再使用均方根误差算法(root mean squared error, RMSE)或多项式回归算法(polynomial regression, PR)等有监督学习算法识别命名实体。

从开放领域中抽取信息所得到的结果, 可能具有较高的数据冗余度且包含大量错误内容, 数据内在的层次性和逻辑性也缺失严重, 这就需要通过有效的知识融合技术来清洗并整合数据, 主要工作包括实体对齐和知识合并等^[3]。

经过融合处理后, 所得到的数据、信息或

事实表达还必须经过进一步的知识加工才能形成最终结构化、网络化的知识体系。此过程中涉及的主要技术包括本体构建、知识推理和质量评估^[4]。

上述层次架构可用图1所示的模型来表示^[1]。

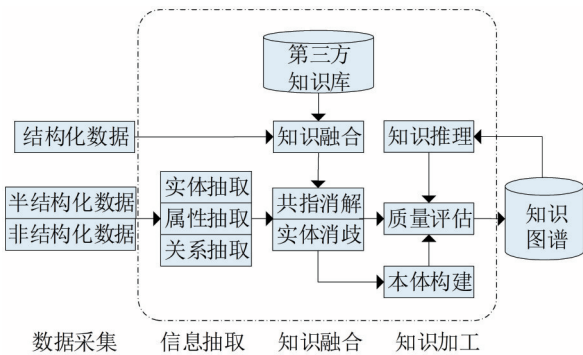


图1 自底向上构建知识图谱的层次架构

2 阅读行为知识图谱构建

就本文所研究的读者阅读行为知识图谱构建而言，采用自底向上的方法更为合适。这种方法将知识图谱的构建过程分为四步：知识获取、知识表示、知识存储和知识可视化错误!未找到引用源。。

结合项目的实际需求，本文设计出阅读行为知识图谱构建系统的实现流程如下：首先从存档的电子出版物中抽取读者的阅读行为数据，然后对数据进行清洗并进行格式调整，再将数据加载到图数据库中，创建知识节点及节点之间的关系以得到完整的知识图谱，最后使用图数据库操纵语言对知识进行查询推理，并将结果可视化地呈现在页面上。完整流程如图2所示。

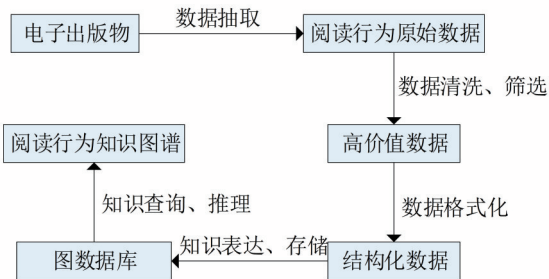


图2 阅读行为知识图谱构建流程

2.1 知识获取

构建知识图谱的首要任务是获取知识。从各种类型的数据源中提取出实体(概念)、属性以及实体之间的相互关系，在此基础上形成本体化的知识表达^[1]。

本文构建阅读行为知识图谱所需的数据主要来源于读者在阅读过程中所生成的各种标注和注释信息，这是一种简明且方便获取的、用以了解读者阅读行为的数据。读者在阅读电子出版物的过程中，会随手在文档中添加一些附注、标记和注释信息，这些信息真实准确地反映了读者个人的阅读习惯，及其对所阅读内容的认识、理解和掌握程度。为保证数据来源的随机性和真实性，本文收集整理了数十位学生读者在阅读不同类型电子出版物后所形成的文档材料，编写程序自动提取出其中所包含的读者阅读行为数据。部分原始数据如图3所示。

```
<?xml version="1.0" encoding="UTF-8"?>
<xfdf xmlns="http://ns.adobe.com/xfdf/" xml:space="preserve">
<annots>
<text rect="366,195007,575,077026,386,195007,595,077026"
creationdate="D:20180723094116+08'00'" name="3f8aa3ae-e0f8-4cd8-a595-
4a8a91022b70" opacity="1" color="#FFFF00" flags="print,nozoom,norotate"
date="D:20180723094215+08'00'" title="anby1" subject="备注" rotation="0"
page="54" icon="Comment" width="1" state="" statemodel="">
<contents-richtext><body xmlns="http://www.w3.org/1999/xhtml"
xmlns:xfa="http://www.xfa.org/schema/xfadata/1.0/"
xfa:APIVersion="Acrobat.11.0.0" xfa:spec="2.0.2"><p dir="ltr"><span style="text-align:left;font-size:13pt;font-style:normal;font-weight:normal;color:#000000;font-family:Arial">Omicron[</span><span style="text-align:left;font-size:13pt;font-style:normal;font-weight:normal;color:#000000;font-family:Arial">ø</span><span style="text-align:left;font-size:13pt;font-style:normal;font-weight:normal;color:#000000;font-family:Arial">ma</span><span style="text-align:left;font-size:13pt;font-style:normal;font-weight:normal;color:#000000;font-family:Arial">1</span><span style="text-align:left;font-size:13pt;font-style:normal;font-weight:normal;color:#000000;font-family:Arial">D</span><span style="text-align:left;font-size:13pt;font-style:normal;font-weight:normal;color:#000000;font-family:Arial">n</span></p>
</body>
</contents-richtext>
<popup rect="384,962006,470,743988,564,961975,590,744019">
```

图3 阅读行为原始数据局部

2.2 知识表示

自动提取的原始数据中往往会包含一些信息噪音^[5]。因而本文设计了专门的数据清洗程序以去除其中的噪音，然后再将数据格式化为知识表示的形式。具体过程为：

(1) 首次清洗。构造如下正则表达式，对抽取的数据进行完整清洗，去除价值密度低的数据，保留重要的标记注释文本、生成时间和创建位置等内容。

(1) 首次清洗。构造如下正则表达式，对抽取的数据进行完整清洗，去除价值密度低的数据，保留重要的标记注释文本、生成时间和创建位置等内容。

```
(r'<(.*?)\w+="\S+"?\w+=("S+")?(\w+="S+"?)\{8}\w+=("S+")?")'
```

(2) 再次清洗。将元组转换为列表数据,构造正则表达式,采用循环扫描的方式对数据进行再次清洗。其中的少部分数据可能发生错置现象,需要抽取相应数据并重载至正确位置。

(3) 格式化数据。构造正则表达式,整理清洗后的数据,对它们进行适当的切分和合并,最后将所有数据格式化为知识表示的形式。

2.3 知识存储

关系型数据库在结构化数据的存储和处理方面拥有绝对优势,但对知识存储及语义检索的支持却不够友好。而以图论为基础的图数据库在这些方面却恰好拥有与生俱来的长处,尤其在保持数据语义及处理复杂关系等方面,图数据库明显优于关系型数据库^[6]。目前在学术研究和商业领域,主要的图数据库产品包括 ArangoDB、FlockDB 和 Neo4j 等。根据 DB-Engine 排名,其中最为活跃的当属开源产品 Neo4j,它不仅支持严格的事务处理,还提供强大的图搜索能力和极好的横向扩展能力。

本文通过编写程序,实现了将格式化后的读者阅读行为数据自动加载到 Neo4j 图数据库中的功能。程序能根据格式化数据的结构创建相应节点,添加属性并标注关系,自动完成知识图谱的存储和构建。

2.4 知识可视化

Neo4j 数据库支持强大的图操纵语言 Cypher,可以快捷高效地实现知识图谱的查询和推理。在配置好所需模块后,调用浏览器打开图数据库,输入相应 Cypher 命令便能查询所需内容并将结果可视化呈现在页面上。本文所创建的读者阅读行为知识图谱的部分查询结果如图 4 所示。



图 4 读者阅读行为知识图谱局部

3 结语

本文介绍了读者阅读行为知识图谱的完整构建过程,设计并实现了一个自动化的开放知识图谱构建系统。测试结果表明:本文所得成果能正确高效地实现对读者阅读行为数据的自动提取、清洗、筛选和格式化,并能在此基础上将融合后的数据表达为知识再存储到图数据库中。后续研究工作重点在于:一方面对获取的阅读行为数据和读者阅读习惯、阅读情感之间的关联性进行深入研究;另一方面对如何将所构建的知识图谱用于帮助图书馆提供更好的智慧阅读服务进行分析^[7-8]。作者将从这两个方面着手,扎实开展后期的理论研究和实践创新工作,以期取得更有价值的成果,为推动全社会形成良好的智慧学习环境添砖加瓦。

参考文献:

- [1] 张吉祥,张祥森,武长旭,等.知识图谱构建技术综述[J].计算机工程,2022,48(3):23-37.
- [2] 赵君珂,张振宇,蔡开裕.基于自然语言处理的医学实体识别与标签提取[J].计算机技术与发展,2019,29(9):18-23.
- [3] CIFARIELLO P, FERRAGINA P, PONZA M. Wisser: a semantic approach for expert finding in academia based on entity linking [J]. Information Systems, 2019, 82.
- [4] 杨玉基,许斌,胡家威,等.一种准确而高效的领域知识图谱构建方法[J].软件学报,2018,29(10):2931-2947.
- [5] 文利情.知识图谱赋能的阅读推广研究[J].农业图书情报学报,2021,33(8):63-70.
- [6] 王红,张青青,蔡伟伟,等.基于Neo4j的领域本体存储方法研究[J].计算机应用研究,2017,34(8):2404-2407.
- [7] 高凡,欧阳娟,吴秀明.高校图书馆阅读推广对读者个体影响力的评价研究[J].大学图书馆学报,2021,39(6):62-70.
- [8] 李春.新媒体时代我国国民阅读行为变化及对策研究[J].河南图书馆学刊,2021,41(7):126-128.

作者简介:

陈光仪(1980—),男,湖南湘潭人,硕士,讲师,研究方向为数据库技术、信息组织与检索

陈义明(1980—),男,湖南南县人,博士,副教授,研究方向为人工智能、信息组织与检索

吴小慧(1996—),女,湖南长沙人,本科,研究方向为信息组织与检索

收稿日期:2022-03-04

修稿日期:2022-04-14

(下转第117页)

文章编号: 1007-1423(2022)16-0114-04

DOI: 10.3969/j.issn.1007-1423.2022.16.022

一款在线藏文化博物馆平台的设计与实现

王 超¹, 倪胜巧^{1,2}, 普 尺¹

(1. 西藏大学信息科学技术学院, 拉萨 850000; 2. 四川大学计算机学院, 成都 610065)

摘要: 藏文化作为我国极具特色的少数民族文化之一, 其保留、传承和发扬对于我国的中华民族文化多样性具有重要意义。本文详实地描述了一款藏文化博物馆平台的设计与实现, 采用前后端分离的开发模式, 基于 Spring Boot 框架, Vue 等技术, 实现了一款包含藏历史、藏语言、藏服饰、藏戏藏舞等多个内容模块的藏文化博物馆平台; 通过后台管理界面实现了平台内容的动态添加和修改。

关键词: 藏文化; 线上平台; 前后端分离; Spring Boot; Vue

0 引言

随着藏民族地区经济和社会的发展, 藏民族文化在图书、报刊、电视等纸质媒介和普通电子媒体上开始向外部传播, 但是民族文化流传度仅局限于固定的区域。如此模式下的藏文化传播, 无论是传播内容或传播形式都被主流媒体把控, 受众一直处于被动接受状态, 易导致外界对藏文化产生刻板的印象^[1]。迄今为止国内外还不曾有系统、深入的对藏网络文化的研究成果和著述, 国内藏文化网络内容重复, 特色不够鲜明。尽管各个藏文化网络所开设的频道栏目不一样, 内容却大致相同, 相互之间缺乏差异性^[2]。而且在国内的一些藏文化网站, 内容冗杂, 多是线下博物馆的线上产品, 以展示藏品为主, 没有对藏文化做一个比较全面、统一的整理, 对其进行保留、展示和宣传, 没有文化内涵, 不能满足大部分浏览用户的需求。本文针对该现象设计实现了一款具有藏历史、藏语言、藏服饰、藏戏藏舞等多个内容模块的藏文化博物馆平台, 促进藏文化的保留、传承和发扬。

1 平台选用框架技术介绍

在平台的前后端框架技术方面, 后端开发选用了 Spring Boot, 主要为前端提供数据接口及服务器支撑; 后台管理和前端模块选用 Vue, 实现了数据及页面视图的双向绑定更新。

1.1 Spring Boot 介绍

Spring Boot 是由 Pivotal 团队提供的开源后端框架, 设计目的在于简化新 Spring 应用的初始搭建以及开发过程, 提供了一种新的编程范式, 可以更加快速、便捷地开发 Spring 项目, 具有控制反转和面向切面两大核心, 框架内部能有效管理所有的对象创建, 并维护着对象之间的依赖关系^[3], 是现在后端服务器应用开发的主要框架。

1.2 Vue 介绍

Vue 是于 2014 年在 Github 上公开的一款用于创建用户界面的渐进式 JavaScript 框架, 也是用于创建单页应用 SPA^[4]的 Web 应用框架。基于 MVVM 原理, 使得 Vue 可以对页面上需要使用的数据进行统一管理, 并能够很方便地获取数据的更新, 实现页面视图与数据模型的交互^[5]。

2 平台的设计与实现

2.1 平台总体设计

平台通过 SSM 的设计模式和前后端分离的方式进行开发。对于前端, 使用 VUE 框架将平台各个功能模块设计为一个组件, 这样具有相同功能的模块直接调用组件就可以起到相同的效果, 减少了重复代码, 不但节约开发的时间成本, 也有利于平台内容动态的实现。对于

后端,使用Spring Boot做整合,并且在持久层结合使用Mybatis框架,实现对数据的存储和访问管理。

平台的整体结构如图1所示,其中面向用户的功能主要有通过文化百科了解藏族起源、藏医、藏历等相关历史文化,认识学习藏语语言,了解藏族的特色服饰文化和传统节日,欣赏藏

族的藏戏、藏舞等;同时,用户还可以对平台内容进行收藏,并提供日程安排和提醒功能。面向后台管理员的主要功能有用户权限管理以及通过可视化接口实现对平台内容的动态增删、修改;还可以在用户信息模块通过饼图、地图、柱状图了解用户的部分基本信息,以此来了解藏文化的受众。

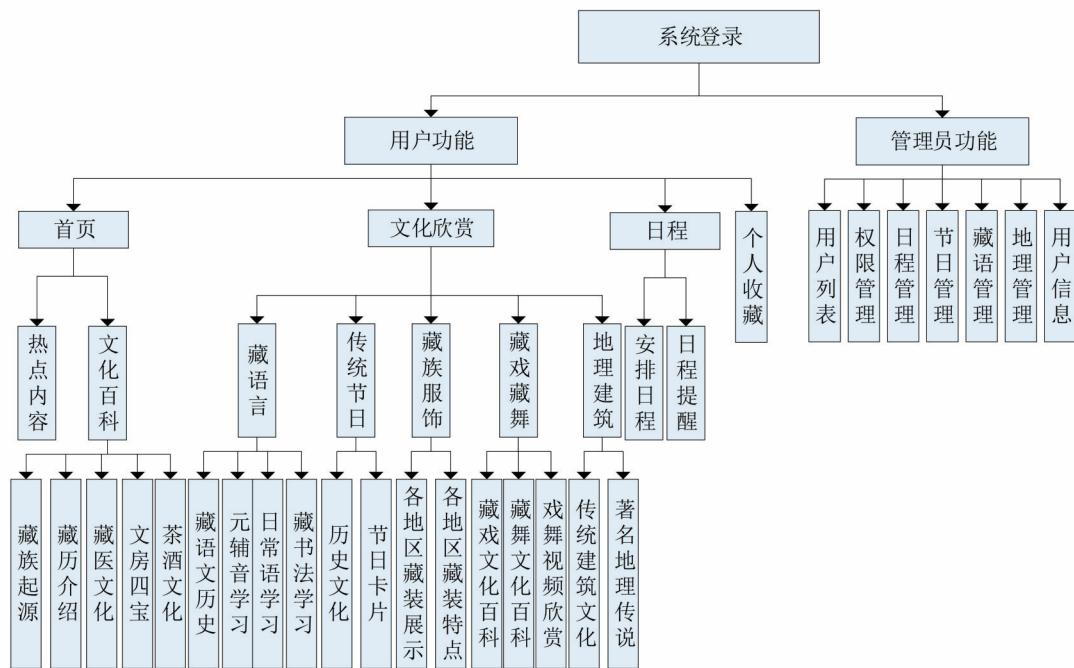


图1 平台整体结构

2.1.1 用户模块

用户登录后在当天第一次进入平台后,系统会为用户进行还未完成日程的提醒。首页上,提供当前点击率较高的平台内容并且提供藏族起源、藏历、藏医等的介绍。在文化欣赏模块,用户可以选择进入自己当前感兴趣的栏目,学习藏语言、了解藏族服饰等,用户对于自己喜欢的内容可以进行收藏,以便于下次直接浏览。

2.1.2 管理员模块

管理员登录后可以在管理模块对已注册用户和平台基本内容进行增删、修改操作;并且可以在用户信息模块借助一款基于JavaScript的百度开源可视化图表库Echarts生成的

图表,了解用户基本信息,分析藏文化受众,如图2所示。



图2 用户信息模块

2.2 数据库设计

本平台采用MySQL数据库,为了保证数据库设计的合理性和减少数据不必要的冗余,数据库表结构严格遵循数据库表的三大范式,保证属性的原子性、记录的唯一性^[6]。平台根据图1的整体结构设计数据库各表,详细信息见表1。

表1 数据库设计

序号	表名称	说明
1	user_info	用户信息表: 用户id、用户昵称、密码、地区、头像、性别、出生日期、创建时间
2	user_collect	用户内容收藏表: 存储用户id、内容名称、内容地址
3	language_words	藏语表: 存储汉语名称, 藏文名称, 音频编码
4	geography	地理表: 地理名称、内容描述、内容地址、图片地址
5	festival	藏族节日表: 节日名称, 藏历日期, 背景知识, 图片地址
6	dance_video	藏戏藏舞表: 名称, 图片地址、音频编码
7	content_record	热点内容记录表: 内容名称, 内容路径, 内容等级, 内容类型
8	arrange_date	日程安排表: 用户id, 日期, 安排详情

2.3 重要技术模块实现

2.3.1 图片存储

SM.MS图床是一个在线图片存储平台,支持的图片格式包括JPG、JPEG、GIF、PNG及BMP,单档不超过5MB,单次最多可上传10张图片。本平台将本地所有的图片存储到SM.MS中,通过将生成的在线地址存储到数据库,实现平台加载图片时为加载在线链接,减轻了平台服务器的存储压力。

2.3.2 音频存储

阿里云视频点播服务(ApsaraVideo VoD,简称VOD)是集视频采集、编辑、上传、媒体资源管理、自动化转码处理、视频审核分析、分发加速于一体的一站式音视频点播解决方案^[7]。本平台所有音频通过该服务设置转码方案和工作流,上传视频生成音频编码保存到数据库,通过在Spring Boot的pom文件中引入依赖com.

aliyun.oss.aliyun-sdk-oss调用GetPlayInfoRequest和GetPlayInfoResponse接口获取解密后的音频地址。

2.3.3 网站加速

CDN(Content Delivery Network,即内容分发网络)是构建在现有网络基础之上的虚拟网络,依靠部署在各地的服务器,通过中心平台的负载均衡、内容分发等功能模块,使访问的用户就近获取到自己所需要的内容,降低网络的拥堵,提高用户访问的响应速度^[8]。本平台通过借助阿里云上的CDN技术,将用户访问时服务器上存储的平台静态内容缓存到CDN节点上,当用户访问静态内容时,就可以不直接通过访问服务器,而是访问就近的CDN节点。以这样的方式达到用户访问平台时加速内容加载的效果,也减轻了服务器源站的压力。

2.3.4 数据访问

Axios支持Promise API,可以在客户端支持防御XSRF攻击、转换前端请求的数据和后端响应的数据等的特性。本平台通过VUE挂载Axios,使用axios.defaults.baseURL绑定访问根路径,简化了代码和向客户端发送请求和转换数据的过程。

3 结语

本文主要论述了基于Spring Boot和VUE框架的一款藏文化博物馆平台的设计与实现,重点介绍了平台架构设计、数据库设计等内容。通过使用阿里云的视频点播技术、SM.MS、CDN等方式来减小平台存储压力、提升性能,达到了预期目标。

参考文献:

- [1] 希娜吉. 新媒体环境下藏文化传播研究[J]. 西藏科技, 2018(8): 34-36.
- [2] 王万宏. 中国涉藏网络文化发展的现状研究[J]. 阿坝师范高等专科学校学报, 2013, 30(2): 58-60, 79.
- [3] 孙岩, 李晶. 基于SpringBoot的旅游资源管理网站的设计与实现[J]. 信息技术与信息化, 2021(1): 37-39.
- [4] 沈剑翘, 陈泽椿. Vue.js在构建系统前端SPA的应用[J]. 科技创新与应用, 2020(3): 181-182.

- [5] 沈佳棋,倪珊,王杰,等. 基于Vue+SpringBoot的分类学科竞赛管理系统设计[J]. 无线互联科技, 2020, 17(17): 74-77.
- [6] 宋龙生,倪胜巧. 校园失物招领微信小程序的设计与实现[J]. 计算机时代, 2021(10): 86-89. DOI: 10.16644/j.cnki.cn33-1094/tp.2021.10.022.
- [7] 杨力乙. 基于微信公众平台的交通综合服务系统设计与开发[D]. 杭州:浙江工业大学, 2019.
- [8] 赵顶. 基于QoS分析的云服务用户效用模型研究[D]. 湖南:湖南大学, 2015.

作者简介:

王超(2000—)男,甘肃张掖人,在读本科

通信作者:倪胜巧(1982—),男,浙江金华人,博士,副教授,副院长,教师,主要研究方向为计算机专业教育、机器学习, E-mail: qiaoger@126.com

普尺(2000—),女,西藏日喀则人,在读本科

收稿日期: 2022-03-08 修稿日期: 2022-04-13

Design and Implementation of a Platform in Tibetan Culture Museum

Wang Chao¹, Ni Shengqiao^{1,2}, Pu Chi¹

(1. School of Information Science and Technology, Tibet University, Lhasa 850000;

2. School of Computer Science, Sichuan University, Chengdu 610065)

Abstract: Tibetan culture as one of my country's extremely distinctive ethnic culture, its reservation, inheritance and carry forward is of great significance for my country's national cultural diversity. The article describes the design and implementation of a Tibetan Cultural Museum platform. It adopts the development model of the front and rear ends, based on the Spring Boot Framework, Vue and other technologies, and realizes a history, Tibetan language, Tibetan clothing, Tibetan opera and Tibetan dance. The Tibetan Cultural Museum platform of multiple content modules; and achieve dynamically added and modifications of platform content through the background management interface.

Keywords: Tibetan culture; online platform; front and rear end separation; Spring Boot; Vue

(上接第113页)

Research on the Construction of Reading Behavior Knowledge Graph Based on Graph Database

Chen Guangyi, Chen Yiming, Wu Xiaohui

(College of Information and Intelligence, Hunan Agricultural University, Changsha 410128)

Abstract: Smart library is an important foundation for society to form an intelligent learning environment. It should be able to provide readers with personalized and intelligent reading services. This paper studies how to obtain the behavior and habits of readers when reading online, and the content of constructing knowledge graph based on graph database. Firstly, it compares and analyzes the current knowledge graph construction methods and framework; Then, combined with the bottom-up construction method, according to the steps of knowledge acquisition, knowledge representation, knowledge storage and knowledge visualization, the construction process of reader reading behavior knowledge graph based on graph database is studied and gradually realized. The results have laid an important foundation for further research on how to provide intelligent reading services according to the needs of readers.

Keywords: knowledge graph; reading behavior; graph database

文章编号: 1007-1423(2022)16-0118-03

DOI: 10.3969/j.issn.1007-1423.2022.16.023

一个智慧门禁系统的软件设计与实现

韩培珊, 郑晓君, 谢松, 黄翔琪

(广东外语外贸大学南国商学院计算机学院, 广州 510545)

摘要: 为了在保障室内安全的同时应对当前疫情防控的需要, 本文设计了一种智慧门禁系统, 该系统采用百度 AI 开放平台实现人脸信息采集和识别等功能, 所检测到的人脸数据信息将实时上传到云服务器, 传输到所开发的微信小程序界面, 并告知使用者来访者信息, 从而确保来访者身份, 保障来访安全。所有数据通过云物联网平台实时传输到小程序, 方便用户实时控制和监测。

关键词: 物联网; 人脸识别; 小程序; 活体检测; 云开发后端技术

基金项目: 2021 年度广东省级大学生创新训练项目(S202112620030)

0 引言

随着社会的快速发展和科技不断进步, 人们对生活质量水平的追求越来越高。近年来, 来访安全问题被大多数人忽略, 仅仅只有独居年轻人群体重视。然而疫情爆发后这几年, 由于封闭居所隔离, 逐渐出现不法分子冒充防疫人员进行入室抢劫以及对屋主进行不法行为的案例, 来访安全问题逐渐被大众重视。为了解决疫情下的来访安全问题, 本文提出了一个智慧门禁系统。

智慧门禁系统项目在百度 AI 开放平台创建人脸识别应用, 获得 AppID、API Key 和 Secret Key 等, 并搭建人脸数据库, 用于存放采集到的人脸数据信息。百度人脸识别平台提供人脸注册、更新和删除等功能, 将采集和监测到的人脸信息数据实时上传至后台巴法云服务器, 方便管理员查询。录入的人脸信息数据会即时在微信开发者工具上转变成 Base64 格式编码并在小程序显示人脸详细信息, 同时与人脸库中的人脸数据信息进行比对, 比对结果同步到小程序界面告知用户来访者信息。针对可能使用任务图片识别蒙混软件的情况, 本系统实现了活体检测、人脸库与检测到的人脸比对等功能。

1 研究方法

1.1 百度 AI 人脸识别

1.1.1 创建人脸识别应用

进入百度 AI 开放平台右侧导航的控制台-产品服务-人工智能-创建人脸识别应用, 应用是调用 API 服务的基本操作单元, 创建成功后点击“管理应用”, 获取 API Key 和 Secret Key 等设备信息, 进行接口调用操作及相关配置。用户需要保存好 AppID、API Key 和 Secret Key, 因为它们程序运行的唯一标识。获取 token 用于校验, 因为在调用 API 时必须要在 URL 中带上 access_token 参数。

1.1.2 下载 SDK

基于本项目我们选择 Java Http SDK, 下载解压后在桌面创建对应这 4 个 jar 文件的项目文件夹, 将这 4 个文件拖入项目文件夹。需要注意, 安装的是 AIP 不是 API。将包解压执行解压命令就可以了。

1.1.3 创建人脸库

在创建好的人脸识别应用界面中点击人脸管理, 点击人脸库名称, 新建一个组, 登记好组 ID, 在编写代码的时候要用到, 然后点进去创建用户, 添加人脸 ID 和照片。

1.2 编写程序并运行

本项目人脸识别原理是获取图片,将其转成 Base64 编码格式,上传数据到调用的百度 API 接口上,实现活体检测,防止使用人物图片识别蒙混进入软件。在确认是活体后,从中截取图片,调用百度人脸识别接口进行人脸识别,最后将得到的返回结果输出为 JSON 格式^[1]。下载官方 API,导入官方提供的四个工具类。先拍摄照片,然后通过 SDK 来检测照片里面的人是谁,如果这个人在人脸库里面,而且有权限的话,那么就开门,否则不开门。

1.2.1 主函数

String auth = getAuth();语句是获取 token 地址,该 token 有一定的有效期,有效期 30 天左右,需要自行管理,失效后需重新获取或者更新。

faceSearch(auth);语句是调用人脸库比对函数,将待检测的图片与人脸库中已有人脸数据进行数据比对。

1.2.2 人脸检测代码

String accessToken=token;这里仅为了简化编码每一次请求都去获取 access_token,线上环境 access_token 有过期时间,客户端可自行缓存,过期后重新获取。请求的图片需经过 Base64 编码,图片的 Base64 编码是指将图片数据编码成一串字符串,使用该字符串代替图像地址。首先得到的是图片的二进制,然后用 Base64 格式编码即可。图片的 base64 编码不包含图片头^[2]。

1.2.3 与人脸库比对

```
public static String faceSearch(String token) {
    Gson gson = new GsonBuilder().create();
    String url = "https://aip.baidubce.com/rest/2.0/face/v3/search";
    String s = ImageToBase64("地址\\1.jpg");
    try { Map<String, Object> map = new HashMap<>();
        map.put("image", s);
        map.put("liveness_control", "NORMAL");
        map.put("group_id_list", "xh 0713");
        map.put("image_type", "BASE64");
        map.put("quality_control", "LOW");
        String param = gson.toJson(map);
        String accessToken = token;
        String result = HttpUtil.post(url, accessToken, "application/json", param);
        System.out.println("与人脸库比对" + result);
        return result;
    } catch (Exception e) {e.printStackTrace();}
    return null;
}
```

图 1 人脸信息识别部分代码示例

1.3 小程序开发

微信小程序(Mini Program)相对于 APP,具备无需安装、无需注册、开发费用低等优点。基于微信平台超 9 亿的用户流量入口以及小程序“用完即走”的理念,同时可以节省用户的手机空间,成为取代手机 APP 的最佳选择^[3]。小程序依靠微信平台,近三年来发展迅速,涉及生活娱乐、医疗交通、健身阅读等各个领域,但对于实现实时查询公共场所进出人流量需求的小程序较稀缺。对于高校学生,此类小程序新颖且为日常生活提供便利。对于开发者,小程序开发门槛相对较低,能够满足简单的基础需求,如查询、通知、统计、打卡等^[4]。

本设计主要使用了小程序云开发的云函数、云存储和云数据库三大功能^[5]。监测到的人脸数据信息会实时同步在小程序转变成 Base64 格式编码并与上传至百度智能云人脸库进行比对,随后将比对结果发送至小程序界面告知使用者来访者人脸信息,所有人脸信息数据都会上传至巴法云服务器,从而保障来访安全和来访信息的具体情况。后台开发者也可以通过巴法云服务器获取实时数据,从而保障使用者的权益。

所监测到的数据同步到小程序上的误差时间不会太大。也就是说,不会因为延时太长而导致室内人因为来不及弄清来访者身份而造成人员伤亡的情况。

2 研究结果

2.1 人脸录入结果

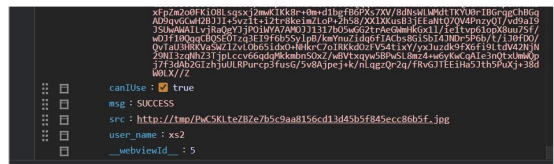


图 2 人脸信息录入成功结果

2.2 人脸识别结果

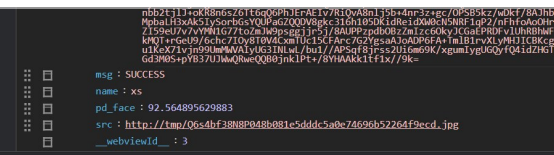


图 3 人脸信息识别成功结果

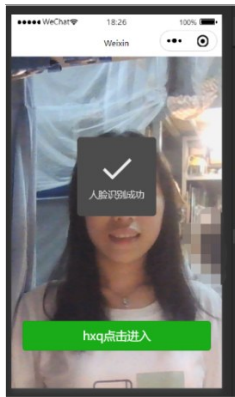


图4 人脸识别结果

2.3 小程序界面

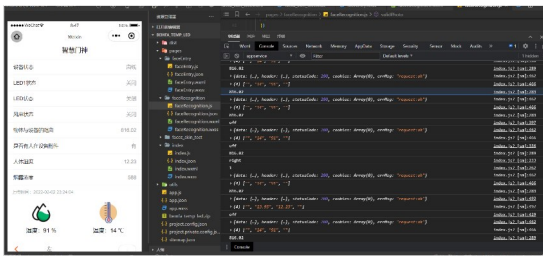


图5 小程序实时数据界面截图

3 总结与讨论

针对当前新冠疫情爆发下出现的来访安全问题，本研究设计了基于百度AI开放平台和小程序的前端技术(WXML+WXS+JavaScript)与云开发后端技术^[4]用于保障来访安全。通过百度AI开放平台创建人脸识别应用，获得AppID、API Key和Secret Key，并搭建人脸库，用来存放人脸数据。百度人脸识别提供人脸注册、更

新和删除以及用户信息查询，对来访者自动进行活体检测，获取用户人脸列表，复制、删除用户等，实现了用户通过手机微信端，可以实时查看来访者信息从而保障疫情时代下来访安全的设想。

参考文献:

- [1] 顾德喜,李贵霖,时祥凯,等.基于百度AI开放平台的人脸识别门禁系统[J].单片机与嵌入式系统应用,2019,19(6):33-36.
- [2] 朱琳,张凤,申泽轩,等.树莓派3B系列的人脸识别实验室门禁系统[J].单片机与嵌入式系统应用,2019,19(4):69-71,76.
- [3] 李昊.基于微信小程序的智能推荐点餐系统的设计与实现[D].南京:南京邮电大学,2020.
- [4] 李茹平,合子琦,张林燕,等.基于微信小程序云开发的人流量监测系统[J].电子技术与软件工程.2021(8):68-70.
- [5] 华伟,杨菲,韩斌.多平台疫苗查询管理和疫苗知识推广系统的开发与应用[J].电子技术与软件工程,2018(19):73-74.

作者简介:

韩培珊(2001—),女,广东中山人,在读本科,专业为物联网工程

郑晓君(2001—),女,广东深圳人,在读本科,专业为物联网工程

谢松(2003—),男,广东湛江人,在读本科,专业为计算机科学与技术

黄翔琪(2000—),女,广东汕头人,在读本科,专业为物联网工程

收稿日期:2022-03-15 修稿日期:2022-04-09

Design and Implementation of Smart Door System for Face Recognition

Han Peishan, Zheng Xiaojun, Xie Song, Huang Xiangqi

(School of Computer Science, South China Business College, Guangdong University of Foreign Studies, Guangzhou 510545)

Abstract: In order to ensure indoor safety and meet the needs of the current epidemic prevention and control, this paper designs a smart door god system. The system uses Baidu AI open platform to realize the functions of face information collection and recognition. The monitored face data information Real-time transmission to the developed WeChat applet interface and inform users to ensure the identity of the visitor and ensure access security. All data is transmitted to the applet in real time through the cloud IoT platform, which is convenient for users to control and monitor in real time.

Keywords: internet of things; face recognition; applets; liveness detection; cloud development back-end technology

