

基于 Fasttext 模型和树莓派的图书馆机器人

李白华

合肥经济学院 信息工程学院, 安徽 合肥 230011

DOI: 10.61369/TACS.2025070025

摘 要 : 图书文本数据的增加对图书管理、图书分类提出了新的要求。本文设计基于 Fasttext 模型和树莓派的图书馆机器人。设计图书馆机器人硬件平台, DSI 显示屏减少硬件设备设置的复杂度。设计符合用户需求的用户管理界面、AI 智能平台和数据库管理软件。AI 智能体中的人脸识别模型识别用户身份, 百度智能体完成 OCR 文字识别功能、AI 语音合成功能。文本分类模型对比常用的分类模型, 择优部署训练后的 Fasttext 模型在树莓派上, 根据分类结果驱动机器人电机运动。使用中文图书数据集测试模型, 图书分 21 个类别, 模型分类宏准确率 93.8%, 为嵌入式设备构建端云协同结构提供新的思路。

关 键 词 : 图书分类; 树莓派; Fasttext 模型; AI 智能平台

Library robot based on Fasttext model and Raspberry Pi

Li Baihua

Hefei University of Economics, Hefei, Anhui 230011

Abstract : The increase of book text data puts new requirements on library management and Library classification. This paper designed a library robot based on Fasttext model and Raspberry Pi. Hardware platform of library robot was designed, using display screen to reduce the complexity of configuration of hardware equipment. The interface of user management, AI intelligent platform and database management software were designed to meet the user's needs. Face recognition model in the AI intelligent was used to recognize the user's identity, Baidu intelligent was used to complete the OCR text recognition and AI speech synthesis. After comparing the commonly used classification models the trained Fasttext model was deployed on the Raspberry Pi, the motors were drove to move according to the calculation results of AI intelligent. Using Chinese book dataset to test the model, the book was divided into 21 categories, and macro accuracy rate of the classification model is 93.8%, which provides new ideas for embedded devices to build end-cloud cooperative structure.

Keywords : library classification; Raspberry Pi; Fasttext model; AI intelligent platform

引言

随着图书文本的增加, 信息量的加大对图书馆自动识别管理提出新的要求, 自动化、智能化成为技术发展趋势。应用在图书馆的小型管理机器人可以减轻工作人员在读者身份核查、图书上架归类等环节的重复性工作, 推进图书馆服务智能化。图书文本识别分类的特殊性主要表现在: 图书文本内容丰富, 涵盖的专业术语、学科知识多, 分类体系复杂, 要求智能体模型能够理解并捕捉丰富的语义特征, 同时对图书进行分类并根据结果将图书载至相应书架位置。

朱德意等^[1]在 Tesseract OCR 引擎的基础上, 结合 OpenCV 进行图像预处理, 完成灰度处理、降噪等工作, 实现芯片 IMEI 码的自动识别, 这为文本分类前图像处理提供思路。姚腾军等人^[2]使用 Fasttext 算法, 结合 n-gram 特征和层次 Softmax 理解处理评论数据, 实现快速文本分类, 在文本分类的实时场景中应用广泛。在中文评论数据集上, FastText 的 F_1 分数达 0.9286, 这证明 n-gram 对捕捉多标签语义是有效的。曹琪^[3]在智慧图书馆的文本分类研究中, 以知识图谱为基础做语义增强, 并以文献主体全文本为对象实现期刊文献二分类, 分类准确率为 0.63。

常用的文本分类方法包括基于向量空间模型 (VSM)^[4]、朴素贝叶斯 (NB)^{[5][6]}、k 最近邻 (kNN)^[7]、支持向量机 (SVM)^[8]、神经网络方法^[9]等。由于这些方法大多使用共现矩阵作为文本表示, 导致文本表示存在维度爆炸问题。高维表示往往无法捕捉文本的主要语义, 使得基于这些算法的文本分类模型性能通常较差。Fasttext 模型^[10]能够在低维连续空间中表示文本, 并捕捉输入文本的主要语

基金项目: 安徽省 2022 年高等学校省级自然科学研究重点项目 - (项目编号: 2022AH052622)。

作者: 李白华 (1984.12-), 女, 河南项城人, 讲师, 硕士, 嵌入式开发, 276368350@qq.com

义。文中以树莓派以及 Fasttext 模型作为基础来设计机器人。运用训练过的 OpenCV 模型识别读者身份，通过网络通信，借助百度云 OCR 模型把图书图片转变成文字识别分类。在对四种文本分类模型进行综合对比之后，选择准确率较高的模型来设计分类智能体。在硬件资源有限的处理器上实现端到端的图书文本实时分类分拣，利用端云协同架构让图书馆服务变得更加智能化。

一、系统硬件设计

Raspberry Pi 4 Model 主处理器是 BCM2711，集成 VI GPU 支持视频解码和图形渲染。文中设计配合 DPI 接口，驱动显示屏。处理器内置的千兆以太网控制器和网络通信模块实现与白云度的通信，减轻机器人运算负担，使设备能集中算力计算图书分类。

Pi 摄像头模块主要时拍摄读者和图书的照片。系统主板通过 USB 口连接摄像头。本设备中使用的是树莓派摄像头 V2 型号，用于捕获图像，带有语音播报功能。

主板使用微型 SD 卡来引导基于 Linux 内核的操作系统 (OS)。该卡还用于存储捕获的图书数据和用户脸部图像。4.3 寸 MIPI DSI 触摸屏通过 USB 口接入主板，通过屏幕配置网络参数；扩展驱动板用来连接和控制外部的电机、传感器等配件。系统整体框架图如下图 1 所示：

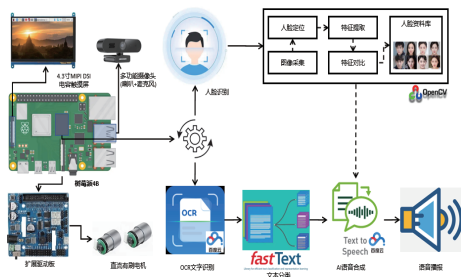


图 1 系统框架图

二、软件设计

软件系统设计分为四个部分，如图 2 所示。

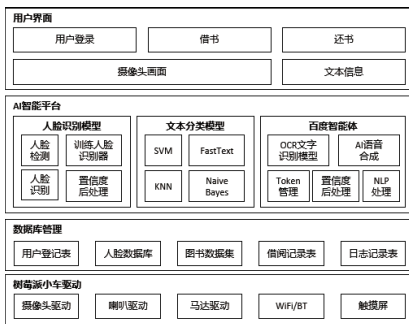


图 2 软件系统框架图

（一）用户界面

提供用户交互操作界面，三个主要功能有对应的三个按钮“用户登录”，“借书”，“还书”。界面底部分别是摄像头画面预览窗口和状态显示栏位，用户在树莓派小车上可以透过触摸屏直接操作，系统运行时的程序主界面如图 3 所示：



图 3 用户界面图

（二）AI 智能平台

三个子系统分别为：人脸识别；文本分类；百度智能体。人脸识别子系统提供人脸检测，识别器训练，置信度后处理。其中识别器训练为模型训练环节，需要使用相对大的 CPU/GPU，树莓派计算资源有限，因此训练环节是在电脑上运行并把训练好的人脸识别模型保存下来 (face_model.yml)，上传到树莓派后运行。训练后的模型返回脸部识别结果会包含脸部识别结果的置信度，置信度的取值范围为 0-1，1 表示识别结果 100% 可信，0 表示完全不可信。置信度越高代表识别的准确度高。通过置信度后处理过滤出高置信度的识别结果，置信度低的识别结果被当作噪声过滤掉。置信度阈值设置为 0.8，置信度过高或者过低会影响脸部识别的准确度。

百度智能体子系统主要提供 OCR 文字识别和 AI 语音合成，OCR 文字识别识别图书图片中的文字；AI 语音合成则是把文字合成对应的语音文件。置信度后处理与人脸识别的置信度后处理类似，区域识别结果置信度低于 0.8，识别结果就会被过滤掉。

文本分类子系统有支持多个不同的文本分类模型 (SVM, NaiveBayes, KNN, Fasttext)，根据测试集对比结果。Fasttext 模型的输入层不仅将句子中每个词对应的词表示作为输入，还将句子的 n-gram 特征作为附加特征输入。

在隐藏层，Fasttext 对输入层的词表示和 n-gram 特征进行平均。在输出层，Fasttext 使用层次化 softmax 预测输入文本的标签。在 Softmax 回归方法中，给定训练集可表示为：

$$\{(x^{(1)}, y^{(1)}), \dots, (x^{(m)}, y^{(m)})\}, y^{(i)} \in \{1, 2, \dots, K\} \quad (1)$$

对于一个输入 x ，模型输出一个 K 维向量，向量的每个元素值表示为 x 属于当前类别的概率。模型可表示为 $h_{\theta}(x)$ ，公式如 (2) 所示：

$$h_{\theta}(x) = \begin{bmatrix} P(y=1|x;\theta) \\ P(y=2|x;\theta) \\ \vdots \\ P(y=K|x;\theta) \end{bmatrix} = \frac{1}{\sum_{j=1}^K e^{\theta^{(j)T}x}} \begin{bmatrix} e^{\theta^{(1)T}x} \\ e^{\theta^{(2)T}x} \\ \vdots \\ e^{\theta^{(K)T}x} \end{bmatrix} \quad (2)$$

模型对应的损失函数如公式 (3) 所示：

$$J(\theta) = - \left[\sum_{i=1}^m \sum_{k=1}^K 1\{y^{(i)} = k\} \log \frac{e^{\theta^{(k)T} x^{(i)}}}{\sum_{j=1}^K e^{\theta^{(j)T} x^{(i)}}} \right] \quad (3)$$

数据库集合用来保存用户、图书以及借阅记录，以便能够供上层 AI 智能平台以及用户界面来使用。在人脸数据库里存有全部获得授权用户的照片，其用途在于对识别模型加以训练。在用户登录的时候，要依靠这些照片来开展特征方面的对比工作。

三、实验分析

（一）数据集准备

数据集是从百度飞桨模型的中文图书分类数据集，原始的数据量 13.3 万条。数据预处理中把样本数在 100 以下的数据过滤，将这个数据集按照训练集和测试集 8 比 2 的比例来进行随机切分。

（二）文本分类模型训练

训练集里面的实际语料标签包含了书名、作者、关键词、摘要、分类标签等诸多方面。文中所采用的文本语料具体实例如表 1 所示。

表 1 数据集语料案例

书名	关键词	摘要	分类
智能家居机器人设计与控制	智能机器人设计	本书详细介绍了嵌入式系统应用程序设计的方法与技巧，结合视觉、触觉、听觉等传感器将机器人推向智能化。本书以服务机器人的各种复杂功能下手，从简单的基本运动、机器人的视觉系统、机器人的触觉系统及机器人的听觉系统逐步进行详细的讲解。	机器人技术

实验中由于 n-gram 会生成诸多特征，所以采用哈希方法予以处理。在此过程中，多个 n-gram 有可能会被划分至同一个“桶”当中，实验中将桶的数量设定为 200 万。文中开展实验对不同的学习率、训练轮数、n-gram 以及向量维度参数组合加以测试，实验结果如图 4 所示。

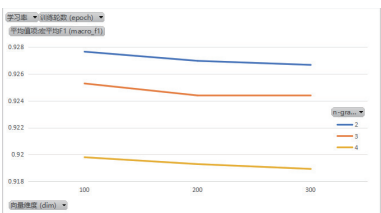


图 4 模型训练参数 n_gram 和 dim 设置实验图

图的纵轴所呈现出来的是平均 F1 值，通过对实验结果分析可知，当把 n_gram 设定为 2，将 dim 确定为 100，训练轮数 epoch 达到 150，学习率 lr 设定为 0.1 的时候，F1 值是最高值。

为了更客观合理地评估各类文本分类模型的性能，本文将支持向量机（SVM）、KNN 和朴素贝叶斯算法（NB）应用于文本分类进行对比实验。实验期间，每个模型的实验独立重复 10 次。每次实验将数据集随机划分，然后通过训练集训练模型，通过测试集评估模型性能。每个模型的精确率、召回率和 F 值为 10 次独立实验的平均值。具体的数据如表 2 所示。

表 2 文本分类模型结果对比

模型名称	宏准确率	宏召回率	Macro-F1	Accuracy
SVM	0.93	0.899	0.913	0.911
NB	0.876	0.739	0.785	0.842
KNN	0.89	0.857	0.871	0.865
Fasttext	0.938	0.921	0.929	0.929

本文设计图书分类为 21 个类别，实验中的宏准确率、宏召回率为 21 个类别准确率和召回率的平均值，并根据这两个参数计算 Macro-F1。Accuracy 时按照测试集的总量取的准确率的平均值。从表中可以看出 Fasttext 模型的宏准确率为 0.938，宏召回率为 0.921。SVM 虽然在宏准确率达到 0.93，但在宏召回率方面下降到 0.899，说明 SVM 在部分类别的数据集推理中漏掉正样本。而 Fasttext 模型在类别样本平均值和全体样本测试平均值中均表现优异，说明模型处理这类的样本数据中具有较好的语义捕捉效果。

将训后的 AI 智能体部署在树莓派后，综合测试的结果如图 5 所示。



图 5 机器人实物图

四、结语

为实现图书馆智能管理功能，提出并设计实现基于 Fasttext 和树莓派的图书馆机器人，在树莓派基础硬件的基础上，通过 MIPI DSI 接口接入触摸屏，SD 卡存储用户脸部数据集，网络通信模块接入百度云，获取百度智能体计算的结果。结合机器人端云协同的框架，将提前训练的人脸识别模型、百度智能体、文本分类模型部署在树莓派机器人上。文本分类模型的训练数据集使用中文图书数据集训练测试，对比 KVM、KNN、NB 和 Fasttext 常用的分类模型，选择部署宏准确率为 0.938 的 Fasttext 模型。测试表明，机器人能够快速识别用户身份、识别图书并分类，根据分类结果驱动机器人运动到相应的图书存放位置。

参考文献

- [1] 朱德意, 孙晴艺, 董思凡, 等. 基于 OpenCV 的芯片 IMEI 码的检测与识别 [J]. 计算机技术与应用, 2022(008): 032.
- [2] 姚腾军, 翟振刚, 高炳涛. 基于 fastText 的文本分类模型 [C]. 2020 年 IEEE 人工智能与信息系统国际会议 (ICAIS). 中国大连, 2020: 154-157.
- [3] 曹琪. 基于知识图谱语义增强的文本分类研究 - 以“智慧图书馆”主题为例 [D]. 武汉: 华中师范大学信息管理学院, 2024.
- [4] 刘少辉, 董明楷, 张海军, 等. 基于向量空间模型的多层次文本分类方法 [J]. 中文信息学报, 2002, 16(1): 9-15, 27.
- [5] 何明, 孙建军, 程颖. 基于朴素贝叶斯的文本分类研究综述 [J]. 信息科学, 2016, 34: 147-154.
- [6] 狄鹏, 段立国. 新型朴素贝叶斯文本分类算法 [J]. 数据采集与处理, 2014, 29(1): 71-75.
- [7] 张宁, 贾子彦, 史忠植. 基于 KNN 算法的文本分类 [J]. 计算机工程, 2005, 31(11): 171-172.
- [8] 范康信. 基于 SVM 的网络文本情感分类系统的研究与设计 [J]. 计算机时代, 2015(12): 34-37.
- [9] 刘腾飞, 于双元, 张洪涛, 等. 循环神经网络和卷积神经网络在文本分类中的应用 [J]. 计算机工程与软件, 2018, 39(5): 64-69.
- [10] 朱美瑶, 张寅昊, 王宇喆, 钟美君. 基于 FastText 模型的匿名数据文本分类研究 [J]. 统计学与应用, 2023, 12 (2): 563-568.