

基于 CNN 智能 AI 助手的早期教育系统设计

沈梦婷^a, 岑 岗^a, 周 闻^b, 朱润锴^a, 梁娅勤^a

(浙江科技学院 a. 信息与电子工程学院; b. 机械与能源工程学院, 杭州 310023)

摘 要: 随着国家对教育的重视,中国国民平均受教育水平不断提高,越来越多的家庭意识到早期教育的重要性。为了解决现存的城乡之间早期教育资源差距大和人们对早期教育 App(application)信赖度不高的问题,设计了基于卷积神经网络(convolutional neural networks,CNN)深度学习技术的早期教育系统。系统将 CNN 技术运用于图像识别和儿童语音识别;内置人工智能助手小铃同学来引导儿童对事物的探索;设置了家长端,通过采集儿童端的儿童使用数据,系统将其制成图表的形式反馈给家长。本系统将语音识别和 CNN 技术相结合时,着重研究儿童声音的特点并以此为基础进行声音模型训练,提高了识别的精准度;通过数据挖掘分析技术,在家长端生成儿童的使用报告并获得基于此报告的儿童兴趣方向推荐和相关课程信息,为家长提供了了解儿童兴趣方向的途径。

关键词: 卷积神经网络;深度学习;人工智能;早期教育;系统设计

中图分类号: TP302.1

文献标志码: A

文章编号: 1671-8798(2020)06-0590-05

Design of early education system based on CNN intelligent AI assistant

SHEN Mengting^a, CEN Gang^a, ZHOU Wen^b, ZHU Runkai^a, LIANG Yaqin^a

(a. School of Information and Electronic Engineering; b. School of Mechanical and Energy Engineering,
Zhejiang University of Science and Technology, Hangzhou 310023, Zhejiang, China)

Abstract: As the country attaches importance to education, the average education level of Chinese citizens continues to increase, and more and more families are aware of the importance of early education. In order to solve the existing problems of the large gap in early education resources between urban and rural areas and people's low trust in early education apps (applications), an early education system was designed on the basis of convolutional neural networks (CNN) deep learning technology. The system used CNN technology for image recognition and children's speech recognition; it had built-in artificial intelligence assistant Xiaoling to guide children's exploration of things; it was equipped with a parent terminal, to which the system fed back in the form of a chart the children's usage data collected from the children terminal. When the system combined speech recognition and CNN technology, it focused on the characteristics of children's voice and conducted voice model training based on this, which improved the accuracy of recognition. Through data mining and analysis, the

收稿日期: 2020-01-14

基金项目: 国家级大学生创新创业训练计划项目(201911057011)

通信作者: 岑 岗(1959—),男,浙江省象山人,教授,主要从事教育信息科学与技术研究。E-mail: gcen@163.com。

system generated a children's usage report on the parent terminal and obtained children's interest direction recommendations and relevant course information based on this report, providing parents with a way to understand the direction of children's interests.

Keywords: convolution neural network(CNN); deep learning; artificial intelligence; early education; system design

近年来,随着国内人均受教育水平的提升,越来越多的家庭开始重视对孩子的早期教育,中国的早期教育市场也随之迎来蓬勃发展。2017年中国科学早期教育大会上的报告表明,新一代家长消费观念已升级,尤其是90后新生代父母,他们更愿意在孩子的早期教育上做更大的经济投入。2017年中国早期教育市场总体规模已经达到4891亿元左右,同比增长17.9%^[1]。2018年中国教育行业融资在儿童早期教育领域的分布已超20.00%^[3]。随着“互联网+教育”模式的发展,早期教育逐步向线上教育和移动教育转型,教育部、中央网信办等六部门联合发布《关于规范校外线上培训的实施意见》,首次从国家层面为教育行业的线上培训提供了政策依据,为其良好的发展奠定了基础^[2]。姜皓文^[4]提出人工智能在儿童教育领域中的应用有助于推动因材施教方法的实现。王奕宸^[5]对人工智能教育应用的观点和看法进行了问卷调查,发现有72.65%的受访者赞成人工智能教育的应用。这说明社会大众对人工智能应用在教育领域持积极态度。目前中国早期教育多为线下教育,早期教育在不同地区的发展程度差距大并且可能存在收费不规范的情况。因此,设计早期教育人工智能助手,提供线上教育服务,以缓解目前线下早期教育存在的问题。良好的早期教育需要结合成熟的早期教育理论,意大利幼儿教育家蒙特梭利提出的蒙特梭利教育理论强调儿童通过自身的实践进行自我教育,重点培养孩子的专心、独立与手眼协调能力^[6];中国著名教育家陈鹤琴^[7]在教学过程中提出“做中学,做中教,做中求进步”的主张,强调实践的重要性,让儿童在学习的过程中激发主体意识和创造精神。人工智能渗透于社会生活的各个方面,而教育人工智能是国内外均具发展前景的一大领域^[8-9]。但目前市面上的早教产品同质化严重,虽提供语音识别、视觉分析等技术来提升儿童的使用感受,但教育内容泛化,难以提供个性化教育服务。因此,课题组设计了一款早期教育系统,通过CNN(convolutional neural networks,卷积神经网络)深度学习技术对儿童提供个性化教育服务,通过数据的积累和分析,在家长端生成儿童的个性教育分析报告,做到对早期教育的因材施教。

1 系统总体设计

课题组设计的早期教育系统主要回答儿童在日常生活中遇到的疑问及解决儿童想了解新事物时却无法及时得到实时性解答的问题。本系统将早期教育与人工智能深度学习相结合,利用卷积神经网络算法所具有的表征学习能力,对儿童向AI助手进行提问时所输入的语言或图片进行识别,将所得信息进行精简处理后以语音播报的形式反馈给儿童,使其更贴合儿童的认知程度,使他们在日常生活中获得早期教育。

为使早期教育有较好的质量保障,本系统设计家长端与儿童端双向早期教育应用。

儿童端设有智能AI助手小铃同学,智能问答与智能识图为其核心模块,帮助儿童解答日常生活中提出的问题,识别儿童拍摄上传的图片,通过语音播报内容来帮助儿童了解更多日常生活中的事物。为使儿童得到全面发展,系统设计了包含诗词、算术等不同方面的拓展模块来培养儿童的学习能力。儿童端设计如图1所示。

家长端具有可视化数据分析。陆永婷^[10]指出,家长是早期教育过程中的重要资源,家长的参与度对儿童未来的发展有巨大影响力。在该系统中,儿童日常的提问记录、拍摄照片的识别结果及儿童在其余模块所分配的时间,在系统经过处理之后以图表的形式展现于家长端,帮助家长了解孩子日常的学习情况,辅助家长给出孩子个性化培养的方案。家长端设计如图2所示。

在儿童端与家长端同时设有安全模块,其中位置共享与通信功能保障了孩子在日常生活中的安全。

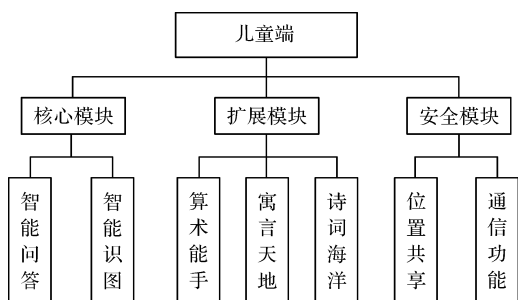


图 1 儿童端设计

Fig. 1 Design of children terminal

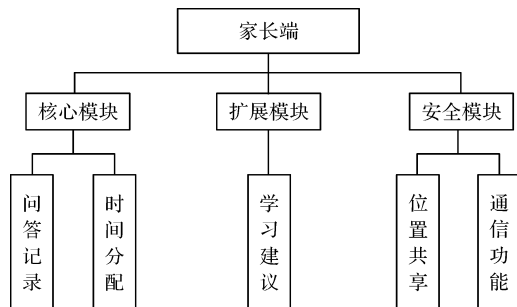


图 2 家长端设计

Fig. 2 Design of parent terminal

1.1 核心模块

在儿童端,核心模块有智能问答和智能识图两个子模块。根据儿童的语言及行为特点,交互通用型设计^[11]尤为重要,儿童端设计注重提升儿童的使用感受。儿童可通过语音输入向 AI 助手提问,助手小铃在对输入的语音进行识别与搜索后,以语音播报的形式反馈给儿童;儿童也可以在智能识图模块中拍摄并上传图片,小铃在对图片进行信号提取与搜索后,将图片中的具体事物语音反馈给儿童。儿童端核心模块流程图如图 3 所示。

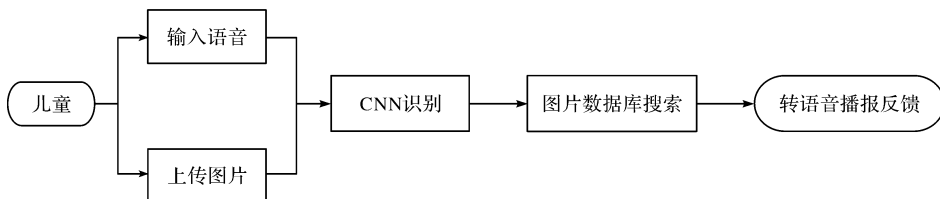


图 3 儿童端核心模块流程图

Fig. 3 Core flowchart of children terminal

儿童端的智能语音识别基于时频分析后的语音谱完成,语音识别过程中需要克服说话者发音的差异性并对环境杂音进行降噪处理。与成年人相比,儿童未发育完全的声带会产生较高的共振峰频率,在成长过程中的儿童的音色也有诸多变化,加之儿童的语言表达能力相对较弱,有时会出现语言表述逻辑混乱的情况,因此,建立完善的儿童声学模型和语言模型尤为重要。在对儿童的语音进行特征提取时,将深度自编码器应用于传统的语音信号特征提取过程中,对提取到的浅层特征梅尔频率倒谱系数(Mel frequency cepstrum coefficient,MFCC)进行线性判别分析^[12]。深度自编码器包含多个隐含层,编码器中的 CNN 分层提取功能将传入下一层的参数进行优化,依次训练整个网络,从而得到识别率较高的语音深度特征。在构建儿童声学模型上,参考 CNN-TDNN-OPGRU 模型^[13],利用 CNN 的降采样特性去除无关语音信息,降低声学特征数量;运用时延神经网络(time delay neural network,TDNN)在语音输入时根据时延单元大小考虑多帧输入,进而获得较长的上下文信息;运用更适合语音任务的门控神经网络(output-gate projected gated recurrent unit,OPGRU)来对语音数据进行训练,相比较于门控限制网络(gated recurrent unit,GRU),OPGRU 的网络计算复杂度更低,解码速度更快,能有针对性地解决儿童语音所表现出的更高基频和共振频率所带来的识别难度。本系统使用 Python 语言和 TensorFlow 框架,采用中国社会科学院语言所普通话儿童语音语料库^[14]进行儿童声学模型训练。

儿童端的图像识别功能主要应用于物品图像分析,利用 CNN 来提升图像识别和检索的准确率。CNN 卷积层主要用于提取图像特征、增强原始图像信息,其局部感受野和权值共享的特点既解决了参数冗余问题又省去了人工手动选取特征的步骤;池化层将图像中局部区域的特征进行聚合,在保持特征不变性的同时还可降低特征图的维度。因此,本系统利用 CNN 的局部感受野、权值共享和池化三大特征来提升对图像的识别效率。在对图像内容进行检索的过程中,首先对输入的图像进行预处理,再运用特征提取算法提取图像特征,最后将提取到的特征向量与已建立好的图像特征数据库进行特征向量匹配,输出检索结果。为保障儿童在使用该功能时能够得到及时的图片信息反馈,优化图像检索结果和提升图像

检索效率至关重要。参考吴斌^[15]所提出的对图像检索的优化方法,将半监督阶梯网络与四值哈希检索方法相结合。半监督机器学习通过训练少量有标签数据和大量无标签数据来达到机器学习优化的目的;同时,将阶梯网络运用其中,在编码层增加横向连接计算阶梯损失函数,解决在进行深度自编码过程中由于网络深度的增加,顶层难以学习数据原始特征的问题;在提高检索效率方面,利用四值哈希方法检索,即在CNN的全连接层的第二层后添加改写的Sigmoid激活函数,将哈希码细化,提高检索时的匹配度和匹配效率。图像识别基本流程如图4所示。图像识别训练时,系统使用Python语言和TensorFlow框架,采用CIFAR-10图像数据集进行图像识别的训练。

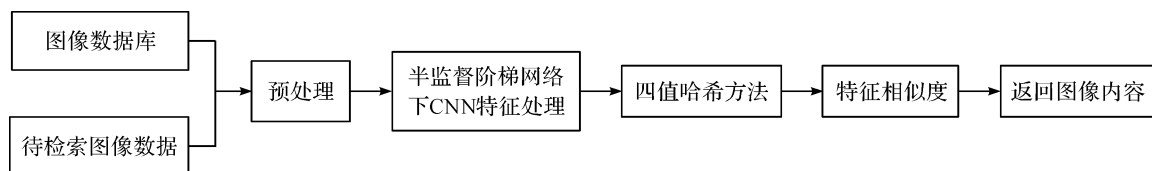


图4 图像识别基本流程

Fig. 4 Basic process of image recognition

在家长端,核心模块给出孩子使用本系统的数据分析。系统将上传孩子对各个模块的使用时间分配数据,在云端将数据进行处理后,以图表的形式展现于家长端。系统将孩子与小铃同学的问答记录实时传至云端数据库,家长可通过查看问答记录来了解孩子感兴趣的事物,这有助于增进家长与孩子的沟通交流。

1.2 拓展模块

在儿童端,拓展模块中的“算术能手”“寓言天地”“诗词海洋”子模块分别通过做题目、讲故事、朗诵录音等方式来培养孩子的学习兴趣和学习能力,同时设有奖励制度以增强孩子的学习动力。

在家长端,家长可参考系统智能分析得出的孩子学习报告来了解孩子的强弱项。系统给出科学的学习时间分配来帮助家长实施针对自己孩子的个性化学习方案,避免盲目给孩子报补习班、兴趣班的弊端。

1.3 安全模块

在儿童端与家长端都设有安全模块,家长与孩子可进行位置共享与语音通话,由此实现家长与儿童的双向交流。系统通过蓝牙定位自动获取位置信息后,建立传输控制协议(transmission control protocol, TCP)实时传输交换数据,系统解析数据后可以给出双方的实时位置^[16]。家长不仅可查看孩子和自己的实时位置,还可开启栅栏模式,自定义孩子的安全距离,此模式下,如果孩子超出家长设定的安全距离,系统将同时在儿童端和家长端发出提醒,以保障孩子的安全。

2 技术路线与数据模型

2.1 技术路线

为做到及时解决儿童所提出的问题,课题组将本系统设计为一款App,针对本系统的特点,选择Hybrid App开发技术,可同时使用原生语言和网页语言,兼具Native App良好的用户体验感和Web App跨平台的优势^[17]。

客户端使用HTML5、JavaScript、CSS、Ajax及主流Vue框架实现系统的整体架构和布局。

服务端使用PHP+MySQL技术组合为动态数据存储标准,根据文献[18],在PHP的动态网页中使用MySQL数据库,可以大大提升网页的应用效果和运行效率。利用PHP+MySQL成熟的框架、稳定的性能、嵌入式开发方式、简洁的语法,将有利于系统的快速开发。

儿童端主界面、家长端主界面如图5和图6所示。

2.2 数据模型

根据儿童与家长双向信息交流的需求,构建一个与系统功能相对应的数据模型,分为儿童端与家长端,如图7所示。儿童端采集儿童的使用数据,将数据反馈至数据库,在经过对数据的分析处理后,以图表形式作为可视化状态输出至家长端。家长端可在安全模块中操作对孩子安全范围的数据设定,数据传至数据库后在儿童端的相应模块中给予反馈。



图 5 儿童端主界面

Fig. 5 Main interface of children terminal



图 6 家长端主界面

Fig. 6 Main interface of parent terminal

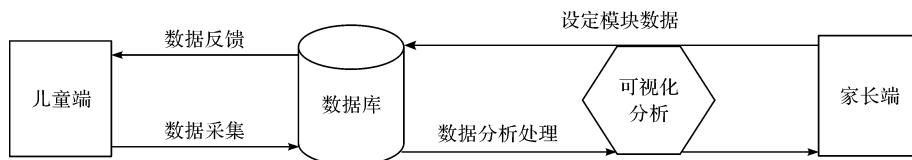


图 7 基于 CNN 深度学习技术的早教系统数据模型

Fig. 7 Data model of early education system based on CNN deep learning technology

3 结 语

本研究将人工智能和 CNN 深度学习技术运用于早期教育系统中。系统通过基于 CNN 深度学习技术的 AI 助手小铃同学来提高儿童在早期的提问能力和思维能力,同时重视家长在早教过程中的参与度,通过数据分析让家长了解孩子的视角和孩子的兴趣方向,帮助家长给出适合孩子的个性化教育方案。本系统设计上仍存在一些问题和不足,如缺少算法之间的对比与数据库的备份问题等,在后续的研发中我们将加以完善和提高。

参考文献:

- [1] 蒋平. 2018 年早教行业现状分析与前景预测 二三四线城市发展空间巨大[EB/OL]. (2018-04-19)[2019-12-02]. <https://www.qianzhan.com/analyst/detail/220/180419-4babaa2f.html>.
- [2] 赵婀娜. 让“互联网+教育”健康发展[N]. 人民日报, 2019-07-24(5).
- [3] 2017—2022 年中国早教中心行业市场发展现状及投资前景预测报告[EB/OL]. (2017-10-24)[2019-12-02]. <http://www.chinabgao.com/report/3515699.html>.
- [4] 姜皓文. 人工智能技术在教育领域中的应用[J]. 电子技术与软件工程, 2018(5):255.
- [5] 王奕宸. 基于人工智能的教育应用及其对中国教育公平的意义[J]. 教育教学论坛, 2019(31):54.
- [6] 殷玉霞. 蒙特梭利教育理论观念在家庭中的应用[J]. 教育教学论坛, 2014(53):113.
- [7] 杨庆, 田景正. 陈鹤琴与蒙特梭利教学法比较[J]. 湖南第一师范学报, 2006, 6(1):46.
- [8] 高丹阳, 张双梅. 人工智能在教育领域的研究现状与特征分析[J]. 中国教育信息化, 2019(13):28.
- [9] 张晨晨, 殷雪峰, 李鹤. 早教机器人:教育与人工智能的绝妙融合[J]. 湖北农机化, 2019(17):15.
- [10] 陆永婷. 中美家长参与早期学校教育的差异研究[J]. 当代学前教育, 2009(2):45.
- [11] 陆超. 交互通用性设计原理在儿童早教生态产品中的应用研究[J]. 包装工程, 2017, 38(10):146.
- [12] 王嘉伟. 基于卷积神经网络的语音识别研究[J]. 科学技术创新, 2019(31):72.
- [13] 徐高鹏. 基于深度学习的中文儿童语音识别声学模型研究[D]. 兰州:兰州理工大学, 2019.
- [14] 高军. 中国社会科学院语言所普通话儿童语音库: CASS Mandarin Child Speech Corpus[C]//第十届中国语音学学术会议论文集. 上海:中国声学学会;中国中文信息学会, 2012.
- [15] 吴彬. 基于深度学习的特征提取及其在图像检索中的应用[D]. 南京:南京邮电大学, 2019.
- [16] 李红卫, 蒋祖星. 一种战术通信系统 TCP 协议双端代理加速方法设计[J]. 舰船电子工程, 2019, 39(7):67.
- [17] 汪正刚. 基于认知发展理论的儿童早教类 App 设计开发研究[D]. 沈阳:沈阳航空航天大学, 2017.
- [18] 梁琰. MySQL 数据库在 PHP 网页中的动态应用研究[J]. 电脑知识与技术, 2019, 15(9):7.