项目编号

**2020年本科生创新创业训练项目**

**申 请 书**

项目名称 情感语音库的建立

项目类型 √创新训练 □创业训练 □创业实践

项目周期 √ 1年 □ 2年

项目负责人 冯璇

学 号 41824061

项目立项学院/单位 ­­­ 计算机与通信工程学院

负责人手机 18811562608

北京科技大学教务处

2019年11月

填 写 须 知

一、项目分类说明

1．创新训练项目是本科生个人或团体，在师指导下，自主完成创新性实验方法的设计、设备和材料的准备、实验的实施、数据处理与分析、总结报告和科研论文的撰写等工作。

2．创业训练项目是本科生团队，在导师指导下，团队中每个学生在项目实施过程中扮演一个或多个具体的角色，通过编制商业计划书、开展可行性研究、模拟企业运行、进行一定程度的验证试验，撰写创业报告等工作。

3. 创业实践项目是在学校导师和企业导师共同指导下，采用前期创新训练项目成果或团队成员已有的研究成果，提出一项具有市场前景的创新性产品或者服务，以此为基础开展实际的创业实践活动。

二、申请书请按顺序逐项填写，填写内容必须实习求是，表达明确严谨。空缺项要填“无”。

三、创业类项目填写内容说明：项目简介主要说明项目来源、产品和服务介绍、项目优势、市场需求等；申请理由包括项目背景和意义、团队优势等；项目方案主要阐述执行思路；项目特色与创新主要说明项目特色和可行性；项目预期成果包括商业计划书、发表论文，调研报告，竞赛获奖、创建和运营公司等相关成果。

四、填写时可以改变字体大小等，但要确保表格的样式没有改变；填写完后用A4纸张双面打印，不得随意涂改。

五、申报过程有不明事宜，请向教务处教学科办公室联系咨询，电话62332202。

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 项目名称 | 情感语音库的建立 | | | | | | | | |
| 项目负责人 | 姓 名 | 性别 | | 学 号 | 班 级 | | 所在学院 | | 联系电话 |
|  |  | |  |  | |  | |  |
| 项目组成员 |  |  | |  |  | |  | |  |
|  |  | |  |  | |  | |  |
|  |  | |  |  | |  | |  |
|  |  | |  |  | |  | |  |
| 指导教师 | 姓 名 | | 职 称 | | | 所在学院 | | 联系电话 | |
| 张德政 | | 教授 | | | 计算机科学与技术系 | | 13301179533 | |
| 一、项目简介  [人工智能](http://ai.qianjia.com/)核心诉求应该是让机器实现对人类的意识、思维的信息过程的模拟，能以人类智能相似的方式做出反应。这就需要从人类IQ、EQ两个维度对人工智能进行审视，而目前大部分人工智能所从事的研究仅限于倾向IQ方向的模式识别，通过自然语言处理、机器学习、模式识别、物联感知、逻辑推理等技术的综合应用，使机器具备一定的逻辑思维判断能力。但是离开人类的干预，人工智能目前仍处于婴儿期，尤其缺乏对复杂情感的理解和表达能力，可以说离开情感的赋能，人工智能依旧是简单的机器。  **下一步人工智能的发展，需要加强对情感，情绪的了解，要走进认知学，心理学。**不仅是脑科学，而是认知学。因为我们目前对人的情感理解非常少，而这对于人工智能来说是很重要的。情感人工智能的应用价值可以深入到人们日常工作和生活的方方面面。例如，视频情感人工智能具有连续、实时、远程、非接触的采集生理和心理信息特点;分析授课人和听课人在教学过程中的心理状态变化，实时反映课程的关注度、兴趣点和授课老师的职业精神状态。  情感语音数据库是语音情感识别研究的数据基础,为语音情感识别模型的建立提供训练和测试数据。情感语音数据库包括情感语音和情感语言两个组成部分。  我们将利用大量的影视作品，做情感语音分析，通过说话和语音判断情绪，建立情感语音数据库。语音库目前分为三种形式：  第一种数据库来自专业或业余演员的表演，朗读预先准备的句子或段落。由于这种方法操作简单，目前大部分情感语音数据库都是用这种方法获得的。  第二种数据库是来访者在知情并签署同意书的情况下置身于一个虚拟场景，从虚拟环境中诱引出语音。  第三种数据库来自现实生活，是人们在现实生活中表现出最真实情感的语音。用三种方法获取的数据库其自然度各不相同，通过试验发现，在使用同样特征参数的情况下，用不同方法获得的数据库其情感识别率不同。本项目希望利用深度学习获取语音特征，进行自动化建库。 | | | | | | | | | |
| 二、选题背景  当今世界的两大难题“情感是什么？它是如何工作的呢？”，它困扰着一代又一代的科学工作者。现在我们的 AI 都是用逻辑的方法来判断情感。因为逻辑代表 IQ，而情感代表 EQ。人类的情感是非常丰富的。**未来，从情绪到情感，最是人工智能未来前进的方向。**  1872年达尔文阐述了《人与动物的情感表达》，达尔文很明确地认为情感只存在于人与动物之中。这是一个确定的事实，正常人能明确区分出动物和人有情感，而树木和石头是没有情感。动物与人有情感，既有生理上的痛苦与快乐感受，也有心理上的喜悦与忧虑感受，而树木与石头是没有这种感受的。无论是人们谈论意义、意图、自我、意识、智能、记忆、思维，还是谈论喜悦、开心、愤怒、嫉妒、傲慢、贪欲、仇恨、愉快、厌恶等等，这些概念都是以情感为基础的，没有情感就没有这些概念存在。因此，毋庸置疑，人类与动物的所有行为，都是源于情感归于情感。  声音是由物体振动产生的声波，是通过介质（空气或固体、液体）传播并能被人或动物听觉器官所感知的波动现象。声音（语音消息）的基本模拟形式是一种称为语音信号的声学波。语音信号可以通过麦克风转化成电信号，转换成语音波形图，。声音作为波的一种，频率（声源在一秒内振动的次数）和振幅是描述波的重要属性，频率的大小与我们通常所说的音高对应，而振幅影响声音的大小。这些特性与人的情感息息相关，它毫不保留的显示了人的当前情感。  目前对于情感识别有两种方式，一种是检测生理信号如呼吸，心律和体温等，另一种是检测情感行为如面部表情识别，语音情感识别和姿态识别。  故我们决定通过对中文情感语音进行特征参数提取的方法，来对语音中angry、fear、happy、sad四种情感语句,提取发音持续时间、平均发音速率、基音周期的最大值、平均值等8个特征参数,完成对语音情感的分析，来对人的心理状态进行分析，完成对情感的判断。其中，语音情感数据集是研究语音情感识别的重要基础，按照情感描述类型可将数据集分为离散情感数据库和维度情感数据库，前者以离散的语言标签（如高兴，悲伤等）作为情感标注，后者以连续的实数坐标值表示情感。我们打算利用影视，做情感语音分析，通过说话和语音判断情绪，完成主要情感语音数据库的建设。 | | | | | | | | | |
| 三、项目方案  为了实现语音情感识别的目的，我们将该项目分为五个部分，包括语音特征提取方法，语音情感识别模型的选择，分类器的训练，以及最终的计算机自动化识别语言情感进而建立语音情感数据库。以下为详细说明。  1.语音特征提取  （1）声学特征是语音情感识别中最常用的一类特征，主要包含三大类：音高音强等韵律特征、频谱特征以及声音质量特征。  对于声学特征计划采用开源软件 openFILE进行帧的水平的低层次基础声学特征LLDs的提取，应用全局统计函数得到全局特征。计划采用广泛使用的特征提取配置文件 emobase2010提取出的声学特征作为特征集，emobase2010特征集包含音高、MFCC、F0等特征系数，并使用均值、最大值、最小值、斜率、百分位数等函数进行统计，最终得到包含1582维声学特征的特征集。  （2）利用文本或音素信息的方法文本信息反映了人说话的内容，而音素是人发声的基本单元，这两种信息对于帮助理解人的说话语音有着重要帮助，因此可以使用说话人或音素信息来辅助训练语音情感识别模型，以提高情感识别准确率。由于比较复杂，我们采取文本信息来辅助语音识别模型  2.文本情感分析：  对带有情感色彩的主观性文本进行分析、处理、归纳和推理的过程。互联网(如博客和论坛以及社会服务网络如大众点评)上产生了大量的用户参与的、对于诸如人物、事件、产品等有价值的评论信息。这些评论信息表达了人们的各种情感色彩和情感倾向性,如喜、怒、哀、乐和批评、赞扬等。基于此,潜在的用户就可以通过浏览这些主观色彩的评论来了解大众舆论对于某一事件或产品的看法。  情感分析的最底层的任务,它旨在抽取情感评论文本中有意义的信息单元，情感信息抽取可提炼出对情感分析有贡献的词或短语元素，其结果对特征降维、提高系统性能有重要作用，常用的统计分析方法有基于信息增益、互信息、期望交差熵、词频、文档频次等。  评价词语的抽取和判别，即为评价词语的识别和极性及度量判断，评价词语的抽取和判别往往是一个一体化的工作，主要分为基于语料库和基于词典两种方法。  基于语料库的评价词语抽取和判别：主要是利用大语料库的统计特性,观察一些现象来挖掘语料库中的评价词语并判断极性。其最重要的优点在于简单易行，缺点在于评论语料库有限，同时评论词语在大语料库中的分布等现象不易归纳。  基于词典的评价词语抽取及判别方法：主要是使用词典中的词语之间的词义联系来挖掘评价词，其中难度最大的是词典的更新程度决定词义分析。  3.语音情感识别模型  计划选择CNN作为语音情感识别模型。  卷积神经网络（CNN）是深度学习中一种非常流行的前向神经网络，有很高的计算速度和很好的特征降维和选择能力，被广泛应用于图像和语音领域.CNN有三个特性：稀疏连接、权值共享和相等表示。局部感受域指一个神经元只与部分邻层神经元连接，即只与一部分输入数据相连，然后在更高层将局部信息组合起来形成高级抽象表征，得到相应的全局信息。这是一种自下而上的提取特征的过程，即层层依次接收局部的输入信息，最后进入全连接层聚合。与每层具有相当大小的全连接网络相比，CNN能够有效降低网络模型的学习复杂度，更容易训练。参数共享即CNN存在多个特征平面，同一特征平面的神经元共享权值，即共享卷积核参数。有了局部感受域和参数共享这两个特点，CNN大大减少了需要训练的参数个数，同时降低了模型的复杂度，减少计算时间，使模型不易过拟合。同时，多特征图能抽取更多特征，也能获得更好的识别率。  4.训练分类器  我们计划人工对大约数百条语音进行识别，建立用于训练的样本集（或者直接采用中科院自动化所的语音库）。获得样本集之后，我们将样本集中的部分样本作为训练样本用于训练该神经网络。所谓训练就是使得神经网络的参数结构逐步符合实际的数学映射关系。通过训练能够获得实现自动识别语音情绪的分类器。  5.建库  在分类器训练完成后，对于大量的数据，这一过程可以由计算机自行完成。如果分类器性能比较高，对于海量数据的分类就会成功率比较高。计算机自动分类的语音就有更高的可信度。 | | | | | | | | | |
| 四、项目特色与创新  近年来,国内外研究者们以各自的科研任务为背景,建立了若干面向语音情感识别研究的数据库。然而,由于情感的复杂性以及缺乏统一的数据库建立标准,只有少数的高质量的数据库得到了大多数研究者们的认可。  我们希望建立一个高质量的情感语音数据库，可以识别使用者的语音语调，分析判断使用者的情绪。其创新点在于我们所建立的数据库是一个面向自然人机交互和人工智能研究的数据库，实现了人机交互，而不是简单的对于语音的识别分析。我们的特色在于这个数据库可以识别不同情绪的人的声音并判断出他们所蕴含的情感，判断该情感是正面情感还是负面情感，并向他们推荐一些可以缓解负面情绪的音乐来平复使用者的心情。 | | | | | | | | | |
| 五、项目进度与安排  第一阶段：2019年12月——2020年1月  文献调研、设计方案。通过网络下载、图书文献检索等方式搜集大量关于情感语音数据库的文献，对搜集到的资料进行归纳整理，制定详细的实施方案。  第二阶段：2020年2月——2020年5月  进行前期准备，搜集大量歌曲，结合歌曲创作背景，分析其歌词、韵律等所蕴含的情感，进行语音特征提取。搜集不同人在不同情景环境下说话的声音，并根据语音、语调、语速等分析其蕴含的情感，建立语音、语言、语义的对用关系。根据不同情感对语料进行整理分类。我们计划人工对大约数百条语音进行识别，建立用于训练的样本集。该样本集中的部分样本将被作为训练样本用于训练该神经网络，获得实现自动识别语音情绪的分类器。  学习python、卷积神经网络（CNN）等知识，为后期进一步语音情感识别做准备。  第三阶段：2020年6月——2020年9月  在分类器训练完成后，将大量的数据用于建立一个高质量的情感语音数据库。  取得初步研究成果，撰写中期项目进展报告。研究论文大体方向、结论已初步形成雏形。  第四阶段：2020年10月——2020年12月  对系统进行测试，进行细节的修改及完善，撰写研究论文和设计文档初稿。  完善论文及研究报告，进行研究总结，发表相关论文。 | | | | | | | | | |
| 六、项目经费使用计划（500字以内，作为评审参考，具体划拨经费方案见立项通知。）  项目预算总经费为：4000元  1.图书资料费：购买图书、订阅有关报刊、杂志、档案、文献等材料的费用，搜集资料，印刷、复印费用500  2.外出调研，参观讨论的费200  3.计算机使用费：包括录入费，上网费、维护费，以及研究过程中发生的信息检索费，数据调查费500  4.专家咨询、报告费200  5论文版面费1000  6.购置课题研究所需设备以及文具纸张等费用500  7.申请专利费1000  7.其他费用：除上述费用之外与项目研究开发有关的其他费用100 | | | | | | | | | |
| 七、项目成果形式  体现在主要研发经费和工作用于创新性理论研究或应用研究的实验，并最终达到能够申请专利。  成果特征： 硬件口 软件口 研究报告口 发表论文/申请专利√ 经济效益口 | | | | | | | | | |
| 指导教师意见：    签 名： 年 月 日 | | | | | | | | | |
| 学院意见：  是否同意立项： 是 口 否 口  建议项目级别： 国家级 口 市级 口 院级 口      教学院长签字：（公章）  年 月 日 | | | | | | | | | |
| 学校意见：  是否同意立项： 是 口 否 口  项目级别： 国家级口 市级口 院级口      负责人签字：（公章）  年 月 日 | | | | | | | | | |