

2015年全国大学生信息安全竞赛作品简介

——（基于语音内容检测和情感分析的老年人通话安全监护系统）

（团队名）



**2015年全国大学生信息安全竞赛**

**作品简介**

**作品名称：基于语音内容检测和情感分析的老年人通话安全监护系统**

**电子邮箱：** [**Catherine.dai@foxmail.com**](mailto:Catherine.dai@foxmail.com)

**提交日期：**

填写说明

1. 所有参赛项目必须为一个基本完整的设计。参赛作品简介旨在能够清晰准确地阐述（或图示）该参赛队的参赛项目（或方案）。

2. 参赛作品简介采用A4纸撰写。除标题外，所有内容必需为宋体、小四号字、1.5倍行距。

3. 参赛作品简介不超过6页A4纸。

4. 参赛作品简介模板里已经列的内容仅供参考，作者也可以多加内容。

5. 为保证网评的公平、公正，作品简介中应避免出现作者所在学校、院系和指导教师等泄露身份的信息。

1. **摘要（总览）**
2. **相关工作（选题背景）**

截止2015年，我国65岁以上人口数量达到1.38亿，占比超过10%。我国已经成为世界上唯一一个老年人口过亿的国家。老年人的身心健康以及人身财产安全已经成为不容忽视的重要问题。老年群体的特殊性决定了老年人对电子产品的接受程度相对较弱，社交方式往往只有通过电话和面对面聊天。同时老年人社交范围较窄，日常生活大多局限在室内，除了与家人、邻居、朋友接触之外，很少与他人交流。如今中国“四二一”的人口结构决定了年轻人赡养老人的负担越来越重，年轻人忙于工作和照顾小家庭，与老年人的交流缺乏，难以全面及时地了解老年人的身心健康和人身财产安全状况。虽然我国养老服务设施建设正在逐步推进，但主要侧重于保证老年人的物质生活，对于老年人的心理健康和人身财产安全，社区的一些义工服务虽然能提供一定帮助，但显然难以做到照顾周全。

基于以上情况，实施诈骗的不法分子往往选择白天拨打电话，白天年轻人都上班了，家里老年人较多，不法分子抓住老年人资信度比较闭塞，容易受骗的情况实施作案。根据调查显示，年龄上看中老年人超过70%。世界各国针对老年人屡屡被骗的情况，采取了许多政策制度上的措施，但却没有便捷的技术上的实现手段。当今市场上仅仅有部分根据举报的电话号码的拦截诈骗电话的产品，无法实现对于所有陌生电话的拦截和监控。因此开发一款专门针对老人的防诈骗系统，实现全方面的电话来电监控和防护。

同时，许多年长成人因活动受限、慢性疼痛、体弱或者其它精神或者躯体问题而失去独立生活的能力，需要获得某种形式的长期护理。此外，老年人更有可能遇到诸如丧亲之痛、因退休或者出现残疾而使社会经济地位下降等事件。所有这些因素都可能使老年人出现孤立、失去独立性、孤独和心理困扰。

精神卫生对身体健康带来影响，反之亦然。例如，存在诸如心脏病等躯体健康疾病的年长成人罹患抑郁症的比率比医学状况良好的人更高。相反，对伴有心脏病老年人的抑郁症不做出治疗，就可能对躯体性疾病的结局带来负面影响。

为什么基于通话？老年人通话频率占社交频率

1. **设计与实现 （系统设计、算法）**
2. **系统方案**

根据上述有关老年人通话安全的调研结果和对有关语音情感分析和关键字检测方法的研究。为了方便用户使用和软件开发，本小组在结合了相关的原理和程序设计的基础上，将整个系统分为通话监听模块、语音检索模块、情感检测模块、高频词统计模块和亲属通知模块5个模块进行开发，从而能够在实际情况下对本系统做到更加合理化的应用。

1. **系统需求分析**

### 3.2.1 需求概述

基于语音内容检测和情感分析的老年人通话安全监护系统为用户提供了情感检测以及关键字检测两大功能模块。情感检测模块下可细分成语音韵律特征提取、语音情感匹配两个部分；关键字检测模块下可细分成建立语音模型、语音模型匹配和关键字定位三个部分。接下来两大模块的几个关键部分进行需求分析。

### 3.2.2 各模块需求分析

声音特征提取及初步声学模型建立的IPO图如图2-1所示：

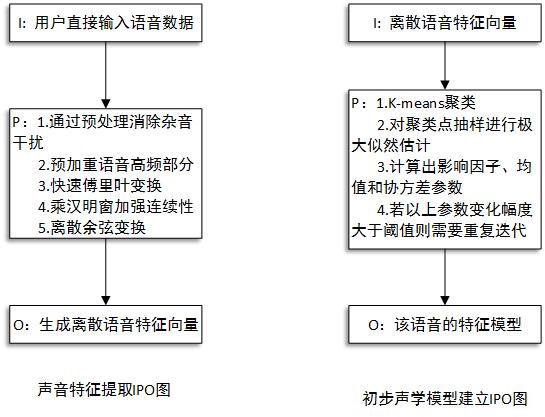


图2-1 声音特征提取及初步声学模型建立的IPO图

提高模型识别概率及基于I-SVM分类器提升识别率的IPO图如图2-2所示：

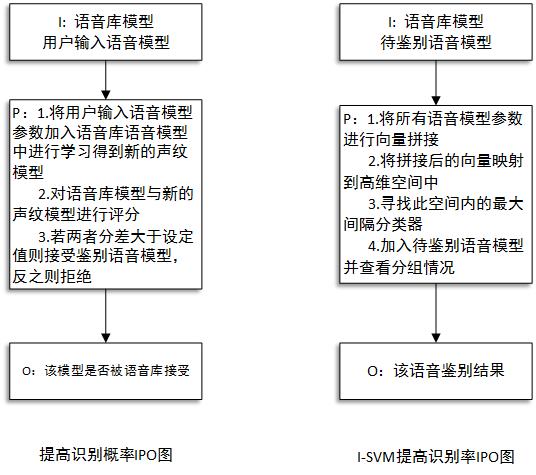


图2-2 提高模型识别概率及基于I-SVM分类器提升识别率的IPO图

关键字监测与定位及人脸识别的IPO图如图2-3所示：

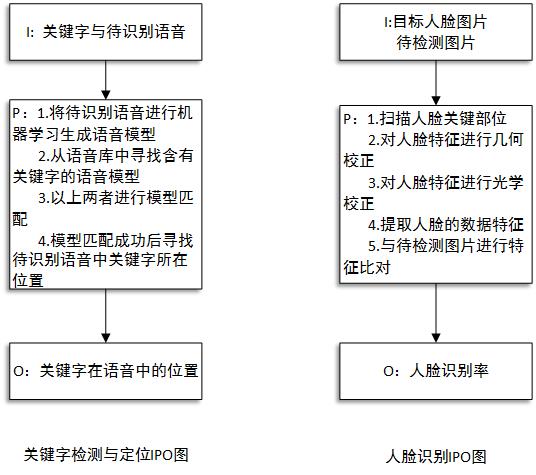


图2-3 关键字检测与定位及人脸识别的IPO图

### 3.2.3 性能需求

1、开启服务后系统实时监控通话信息

2、通话过程中系统能实时分析通话内容

3、系统年非计划宕机时间不高于8小时

4、语音识别准确率≥80%，情感识别准确率≥80%

<公式>

5、所支持语音格式：采样频率16000Hz、单声道16位的wav格式

### 3.2.4 硬件、框架

整个系统由移动端、服务器、关键词数据库等组成。为了保证用户的隐私，采集到的音频数据仅用作特征匹配，匹配之后即删除。用户可以上传疑似诈骗电话到服务器，服务器会进行高频词统计分析，但只保留高频词，不保留音频本身。整个系统硬件的拓扑图如图2-4所示：

**<tupian>**

1. **系统模块设计**

系统根据功能层次分为3个大模块，如图2-5所示。用户在开启相关服务后，后台服务会记录通话内容并发送至服务器处理，处理过程包括：根据用户选择的功能，选择是进行语音内容识别或是进行语音情感分析；如果是进行语音内容识别，则通过调用深度学习生成的语音模型对语音进行解码；调用高频词统计库来判断语音内容是否安全；如果是进行语音情感分析，则先提取语音特征，再对特征进行归类。最后，不管是语音内容识别还是语音情感分析，都会将匹配的结果返回给用户，并且提供直观的参考数值为使用者提供更直观的参考依据。

<图片>

音频特征提取模块：主要用于提取用户输入的实时语音进行语音特征提取，在提取结束后结果不会返回至用户界面中，将提取出的特征参数直接传递给SVM分类建模模块进行下一步生成模型的操作。

SVM分类模块：当语音特征提取之后，将其送到训练好的SVM分类模型中进行分类。SVM分类模块利用核函数将输入特征的样本点映射到高维特征空间中，在高维空间中被超平面分割，判断其情感类别。SVM分类模块会记录一次通话中每小段音频的情感分类，如果通话中某些负面情感占比到达一定阈值，则将结果反馈给亲情账号。

Kaldi语音内容识别模块：用户在提供了语音之后，系统将会把语音送入Kaldi工具箱中进行解码，解码后获得语音内容。将语音内容和关键词数据库中的关键词对比，如果有关键词出现，则将结果反馈给亲情账号，并提醒老人。

关键词统计模块：老人可以在客户端举报诈骗、骚扰等不良通话内容，服务器收到举报后将通话提交给高频词统计模块。高频词统计模块先对通话录音进行解码，统计通话中各词语的出现频率并记录。当某一日常使用频率不高的词语在不同举报记录中都出现且频率较高时，将其记录至关键词数据库。

1. **测试与分析（初步测试结果）**
2. **软件测试方案**
3. **软件测试环境与测试设备**
4. **软件测试数据**
   1. **gg**

**五．创新点总结**

**1.情感分析对老年人心理状态进行监控。**老年人的健康问题，诸如疾病、跌倒等显性的健康问题很容易受到关注，市面上也有许多成熟的产品。然而对老年人心理状态、情绪状态的关注则较少，不健康的心理是诱发各种老年病的直接或间接因素。

**2.具有针对性的电信诈骗事前预警与事后取证。**当前电信诈骗作为热点问题，引起了广泛的社会关注，市面上的产品也较多。然而基于语音的产品很少，而且没有针对性，效果很差。我们的作品针对老年人的许多诈骗场景专门提取关键词，动态更新词库。并提供亲情账户，可以及时发现，及时取证，从最大程度上减少老年人人身财产安全损失的可能。

**3.深度学习在海量数据集上的应用。**因为我们的系统基于深度学习和语音识别，在海量数据集上的表现明显更好。系统有着自学习的功能，随着语音数据集的扩大，关键词数据库也会更符合当前老年人电信诈骗的形势实时更新，效果会越来越好。

**六．未来工作**