**一种融合音频情感和视频表情分析的实时欺诈识别系统**

1. **本专利申请的技术领域**

本发明涉及机器学习和模式识别在反欺诈领域的应用，提出了一种通过对视频用户实时的音频情感分析建模以及微表情特征建模来分析用户潜在的欺诈行为系统。

1. **现有技术状况**
2. **现有技术的客管缺陷和存在的技术问题**

语音情绪分析难度较大，还没有成熟稳定的技术应用于欺诈识别, 同时缺少实际的大数据样本支持。

微表情识别比较完善，但是实际生产样本数据很难收集，提取欺诈识别相关的微表情特征比较困难，实际效果堪忧。

目前市场上还没有同时进行语音和表情分析的欺诈识别系统。

1. **本专利申请的技术方案**

本专利发明了一种融合同步语音情感分析和视频微表情特征分析的实时欺诈识别系统。通过屏蔽底层特征融合抽取等细节，只对外提供统一API接口。模块架构概要设计图如下:

本系统实现同步语音情感和视频微表情特征分析进行欺诈识别的流程如下：

(1) 用户通过web进行语音/视频接入系统，语音视频数据分离;

(2)对于语音流数据，通过特定算法提取情感高阶特征【特定算法是什么算法？如何提取？】; 对于视频流数据，提取面部动作和眼神方向等高阶特征【什么叫高阶特征？低阶特征？】；

(3)将两组特征数据分别代入离线预训练的语音欺诈识别模型和视频表情欺诈识别模型计算欺诈概率【具体模型？如何计算？】；

(4)通过加权模块融合两组模型的输出结果，返回欺诈识别结果【如何融合？】。

**该系统的核心在于离线预训练的语音情感欺诈识别模型和视频表情欺诈识别模型**。

语音情感特征提取和训练流程：【沟通建议：以一个具体的例子贯彻各个步骤：要实现的目的是什么？为实现这个目的，执行了哪些步骤、每一步做了什么操作、操作如何实现的（应用的原理、模型、算法以例子show出来）？】

1. 通过专用的音频分析软件【什么软件？？opensmile】从语音片段中提取大量的音频低阶特征，如MFCC, Pitch, Zero Crossing Rate. 【提取这些特征后续有何作用？】
2. 从低阶特征中提取进一步的动态回归系数，类似于导数，提炼语音的动态特征；
3. 对上述(1) (2)步骤提炼的语音特征，运用大量统计函数（如max，min, skewness， kurtosis）进行高阶特征提取【举例说明如何提取？】，将序列化数据转化为标量特征【什么叫序列化数据？标量特征？】；
4. 运用特征筛选算法，对上述特征进行筛选，选出最优特征子集；
5. 采用分类算法【哪种分类算法？】对筛选特征后的数据进行训练，交叉验证确定最优模型参数【如何交叉验证？】；

视频表情特征提取和训练流程：【沟通建议：以一个具体的例子贯彻各个步骤：要实现的目的是什么？为实现这个目的，执行了哪些步骤、每一步做了什么操作、操作如何实现的（应用的原理、模型、算法以例子show出来）？】

（1）对收集的视频数据进行切割和人工标注【如何切割？如何标注？标注什么？】；

（2）基础特征提取，如眼球朝向，头部朝向，面部动作

（3）高维特征构造，如对多种表情同时出现持续时间，次数等高阶特征进行构造，标准化处理【举例说明构造后、标准化处理后的结果是怎样的？】

（4）特征筛选，选出最优特征组合【跟上面的4一样吗？】

（5）分类算法训练数据，交叉验证确定最优模型参数【同上问：哪种分类算法？如何交叉验证？】

对上述两个模型，根据预测精度分配权重【权重怎么分配？】，融合两个模型的预测结果【如何融合？】。

1. **本专利申请的关键点或欲保护点**

语音情感分析在欺诈识别上的应用技术，语音情感特征的分层抽取融合技术，实时视频流的微表情特征分析在欺诈识别上的应用技术。同步结合实时语音情感分析和视频表情分析验证模型。

1. **本专利申请的优点**

高维特征融合提取技术，从语音和视频多角度进行验证，欺诈识别精度高。

1. **参考文献**