



(12) 发明专利

(10) 授权公告号 CN 117809354 B

(45) 授权公告日 2024. 06. 21

(21) 申请号 202410223747.7

(22) 申请日 2024.02.29

(65) 同一申请的已公布的文献号

申请公布号 CN 117809354 A

(43) 申请公布日 2024.04.02

(73) 专利权人 华南理工大学

地址 510640 广东省广州市天河区五山路
381号专利权人 人工智能与数字经济广东省实验
室(广州)

(72) 发明人 张通 吴梦琪 王锦炫 陈俊龙

(74) 专利代理机构 广州市华学知识产权代理有
限公司 44245

专利代理师 霍健兰

(51) Int.Cl.

G06V 40/16 (2022.01)

G06N 3/0442 (2023.01)

G06N 3/045 (2023.01)

G06N 3/0464 (2023.01)

G06V 10/80 (2022.01)

G06V 10/82 (2022.01)

A61B 5/16 (2006.01)

A61B 5/00 (2006.01)

(56) 对比文件

CN 109409222 A, 2019.03.01

US 2016350801 A1, 2016.12.01

审查员 黄子怡

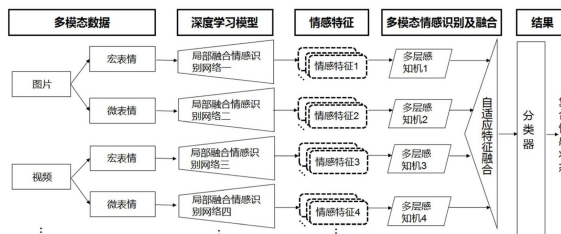
权利要求书3页 说明书8页 附图5页

(54) 发明名称

基于头部可穿戴设备感知的情感识别方法、
介质及设备

(57) 摘要

本发明涉及情感识别技术领域,具体提供了一种基于头部可穿戴设备感知的情感识别方法、介质及设备;其中,方法包括佩戴者情感识别方法,步骤为:采集佩戴者的多模态情感数据;采用局部融合情感识别网络进行处理;左眼输入和右眼输入进行深度卷积;左下脸部输入和右下脸部输入分别通过嵌入层将提取到的动作单元嵌入,然后与面部行为编码一起输入到空域图卷积中;采用多层感知机中进行空间映射,计算空间注意力和通道注意力后进行特征图融合得到情感特征;将情感特征进行融合,通过分类得到复合情感识别结果。该方法在局部多视角脸部数据中采用面部动作单元信息辅助情感感知,提升情感信息从佩戴者身体外观中投射的鲁棒性,提升情感判别精度。



1.一种基于头部可穿戴设备感知的情感识别方法,其特征在于:头部可穿戴设备包括:设备本体和设置在设备本体上的多模态数据采集装置;多模态数据采集装置包括用于分别采集佩戴者四个视角的图片 and 视频的四个摄像模块一;所述四个视角是指:左眼、右眼、左下脸部、右下脸部;

基于头部可穿戴设备感知的情感识别方法,包括佩戴者情感识别方法;所述佩戴者情感识别方法包括如下步骤:

步骤X1、采集佩戴者的多模态情感数据;佩戴者的多模态情感数据包括通过四个摄像模块一获取的图片数据一和视频数据一;图片数据一和视频数据一均包括四个视角数据;

步骤X2、对佩戴者的多模态情感数据进行情感特征提取:

对图片数据一和视频数据一,分别采用局部融合情感识别网络进行处理;其中,图片数据一分别以四个视角数据作为局部融合情感识别网络的四个视角输入;视频数据一分别对四个视角数据提取起始帧和峰值帧,作为局部融合情感识别网络的四个视角输入;

所述四个视角输入在局部融合情感识别网络中的处理方式是:左眼输入和右眼输入进行深度卷积以提取局部视角特征;左下脸部输入和右下脸部输入分别通过嵌入层将提取到的动作单元嵌入,然后与面部行为编码FACS一起输入到空域图卷积中以提取局部视角特征;将四个视角输入提取得到的局部视角特征同时输入到多层感知机中进行空间映射,计算空间注意力和通道注意力后进行特征图融合,得到最终的情感特征;

步骤X3、将步骤X2得到的情感特征进行融合,通过分类得到复合情感识别结果;

所述局部融合情感识别网络为四个,分别为局部融合情感识别网络一、局部融合情感识别网络二、局部融合情感识别网络三和局部融合情感识别网络四;

所述局部融合情感识别网络一对图片数据一进行处理,得到图片数据一的宏表情情感特征;局部融合情感识别网络二对图片数据一进行处理,得到图片数据一的微表情情感特征;局部融合情感识别网络三对视频数据一进行处理,得到视频数据一的宏表情情感特征;局部融合情感识别网络四对视频数据一进行处理,得到视频数据一的微表情情感特征;

四个局部融合情感识别网络均包括分别四个视角输入四个局部特征提取单元;针对左眼输入和右眼输入的两个局部特征提取单元均包括深度卷积网络一;针对左下脸部输入和右下脸部输入的两个局部特征提取单元均由嵌入层、空域图卷积网络依次连接组成;嵌入层还与动作单元提取器连接;空域图卷积网络还与面部动作编码系统连接;左下脸部输入和右下脸部输入;四个局部特征提取单元的输出同时与多层感知机连接,并通过通道注意力和空间注意力进行融合;

其中,所述局部融合情感识别网络三和局部融合情感识别网络四,针对左眼输入和右眼输入的两个局部特征提取单元分别还包括动作放大网络;左眼输入和右眼输入分别通过动作放大网络将微笑表情放大,再输入深度卷积网络一,提取局部视角特征。

2.根据权利要求1所述的基于头部可穿戴设备感知的情感识别方法,其特征在于:所述步骤X2中,对图片数据一和视频数据一,在采用局部融合情感识别网络进行处理之前,先分别进行预处理;

对图片数据一,预处理包括利用串联的预处理卷积神经网络一进行人脸检测;利用串联的预处理卷积神经网络一进行人脸检测,是指:生成候选框,并对候选框进行初步筛选,对人脸关键点进行检测;经过卷积、激活函数、池化、全连接处理后,输出每个候选框的置信

度、坐标偏移量和五个关键点的坐标,以实现人脸检测;

对视频数据一,预处理包括利用串联的预处理多层深度卷积神经网络二进行人脸检测;利用串联的预处理多层深度卷积神经网络二进行人脸检测,是指:使用视频流方式,逐帧读取视频数据一;对视频数据一的每一帧图像,利用串联的预处理多层深度卷积神经网络二对图像改变尺寸后的金字塔形数据进行操作,得到人脸框、关键点坐标和人脸分类,以实现人脸检测;预处理多层深度卷积神经网络二包括依次连接的图像改变尺寸层、卷积神经单元一、卷积神经单元二、最大池化层一、全连接层一、卷积神经单元三、最大池化层二和全连接层二,以及连接卷积神经单元三与最大池化层二之间的空间注意力层。

3. 根据权利要求1所述的基于头部可穿戴设备感知的情感识别方法,其特征在于:所述步骤X3,是指:采用多模态自适应融合模块将步骤X2得到的情感特征进行融合:多模态自适应融合模块的输入为情感特征 $X = \{X_1, \dots, X_n\}$,其中 X_i 为第 i 个情感特征, n 为情感特征的数量;利用注意力机制并且迭代地进行特征融合,最终得到融合特征;将融合特征输入到分类器中进行学习,得到复合情感识别结果;复合情感识别结果采用情感状态的复合表示;情感状态的复合表示形式为:情感类别及对应比例。

4. 根据权利要求1所述的基于头部可穿戴设备感知的情感识别方法,其特征在于:所述头部可穿戴设备,多模态数据采集装置还包括音频采集模块;所述步骤X1中,多模态情感数据还包括音频数据;所述步骤X2中,还对音频数据进行情感特征提取:对音频数据进行滤波、平滑、分帧操作;提取梅尔倒谱系数特征;利用梅尔倒谱系数特征并载入为特征向量形式,输入到基于注意力机制的BiLSTM神经网络中进行情感特征提取;

所述步骤X1中,多模态情感数据还包括文本数据;所述步骤X2中,还对文本数据进行情感特征提取:采用word2vec模型对文本数据进行处理得到序列上下文词向量表示,再经过基于LSTM情感分析网络进行情感特征提取。

5. 根据权利要求1所述的基于头部可穿戴设备感知的情感识别方法,其特征在于:所述头部可穿戴设备,多模态数据采集装置还包括用于采集被观测者图片和视频的摄像模块二;

基于头部可穿戴设备感知的情感识别方法,还包括被观测者情感识别方法;被观测者情感识别方法包括如下步骤:

步骤Y1、采集被观测者的多模态情感数据;被观测者的多模态情感数据包括通过摄像模块二获取的图片数据二和视频数据二;

步骤Y2、对被观测者的多模态情感数据进行情感特征提取:

对图片数据二,分别采用表情卷积神经网络一、表情深度卷积网络二和图神经网络进行情感特征提取,得到图片数据二的宏表情情感特征、微表情情感特征、姿态情感特征;

对视频数据二,分别采用三维卷积神经网络三、基于峰值帧光流的卷积神经网络四进行情感特征提取,得到视频数据二的宏表情情感特征、微表情情感特征;

所述基于峰值帧光流的卷积神经网络四进行情感特征提取,是指:首先对视频数据二进行旋转、裁剪、人脸对齐的预处理,利用峰值帧检测算法提取出峰值帧;接着,计算视频数据二的起始帧与峰值帧之间的光流矢量 u, v 以及光应变 ε ;灰度化处理后,分别作为RGB图像的三个通道,合成为一张RGB图像;对RGB图像进行情感特征提取;

光流矢量 u, v 以及光应变 ε 采用如下公式组进行计算:

$$I(x, y, t) = I(x, y, t) + \frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial t} dt + \varepsilon;$$

$$u = \frac{dx}{dt};$$

$$v = \frac{dy}{dt};$$

其中, $I(x, y, t)$ 表示像素点在起始帧的光强度; t 表示所在的时间维度; dx, dy 分别表示起始帧到峰值帧的横坐标和纵坐标的移动值; dt 表示起始帧移动到峰值帧所用时间; 令

$$I_x = \frac{\partial I}{\partial x}, I_y = \frac{\partial I}{\partial y}, I_t = \frac{\partial I}{\partial t}, \text{分别表示像素点的灰度沿 } x, y, t \text{ 方向的偏导数; } I_x u + I_y v + I_t = 0,$$

其中, I_x, I_y, I_t 均由图像数据求得;

步骤Y3、将步骤Y2得到的情感特征进行融合,通过分类得到复合情感识别结果。

6. 根据权利要求5所述的基于头部可穿戴设备感知的情感识别方法,其特征在于:还包括个体数据存储数据库的建立方法:在个体数据存储数据库构建个体ID;在佩戴者情感识别方法或被观测者情感识别方法开始前,先获取佩戴者或被观测者的个体ID;在佩戴者情感识别方法或被观测者情感识别方法完成后,将佩戴者情感识别方法或被观测者情感识别方法得到的情感识别结果以及时间、地点存储到对应的个体ID中。

7. 一种可读存储介质,其特征在于,其中所述存储介质存储有计算机程序,所述计算机程序当被处理器执行时使所述处理器执行权利要求1-6中任一项所述的基于头部可穿戴设备感知的情感识别方法。

8. 一种计算机设备,包括处理器以及用于存储处理器可执行程序存储器,其特征在于,所述处理器执行存储器存储的程序时,实现权利要求1-6中任一项所述的基于头部可穿戴设备感知的情感识别方法。

基于头部可穿戴设备感知的情感识别方法、介质及设备

技术领域

[0001] 本发明涉及情感识别技术领域,更具体地说,涉及一种基于头部可穿戴设备感知的情感识别方法、介质及设备。

背景技术

[0002] 精神疾病往往会影响患者日常生活、学习、工作还有社交等活动,长期焦虑、抑郁会影响患者个人发展,甚至出现自伤、伤人等行为。因此,对精神疾病进行筛查和监测,以实现精神疾病的及时确诊和治疗,对精神疾病患者是非常重要的。

[0003] 现有精神疾病诊断方式一般采用填表答题方式,或采用外部设备采集被诊断者的生物信号数据或表情数据、进行数据处理和情感识别,从而得到精神疾病的诊断结果。但是采用外部设备采集被诊断者的生物信号数据或表情数据,一般需要被诊断者留在特定的检测环境,采集一段较短时间内的数据;被诊断者处于特定环境时的精神状态,并不能完全代表日常生活中的精神状态;并且采样时间不长,采样数据量有限,也对诊断结果的准确度产生影响。若采用头部可穿戴设备采集情感数据,之后对情感数据进行处理和分析,则可解决检测场所局限和采样时间限制,是一种理想的方式。但是采用头部可穿戴设备获取脸部图像,由于距离面部太近,难以获取全脸图像,需要采用多个摄像头配合才能把脸部各个位置的图像获取完毕。而现有方式在进行表情倾向识别时只能提取全脸(全局)的情感特征,不具备多个视角表情的情感特征提取融合的能力。

[0004] 表情分为宏表情和微表情。宏表情是一些比较明显、持续时间较长的面部表情,容易被检测和识别,在情感识别中扮演着非常重要的角色,也是现有情感识别技术中主要采用的情感数据;但是由于人们有时会隐藏或不自觉地压抑自己的情感,刻意控制和改变宏表情,因此仅通过宏表情来识别情感并不准确。微表情在无意识状态下自发产生,难以掩饰或伪装,通常与真实情感直接相关,因此在情感分析中加入微表情数据更为可靠;但微表情持续时间短,通常只有 $1/25s \sim 1/3s$;动作强度低,难以察觉和捕捉,在现有情感识别技术中应用较少。

发明内容

[0005] 为克服现有技术中的缺点与不足,本发明的目的在于提供一种基于头部可穿戴设备感知的情感识别方法、介质及设备;该方法可采集佩戴者局部多视角脸部数据,在局部多视角脸部数据中采用面部动作单元信息辅助情感感知,协调多区域情感状态的一致性和整体性,提升情感信息从佩戴者身体外观中投射的鲁棒性,从而提升情感判别精度。

[0006] 为了达到上述目的,本发明通过下述技术方案予以实现:一种基于头部可穿戴设备感知的情感识别方法,头部可穿戴设备包括:设备本体和设置在设备本体上的多模态数据采集装置;多模态数据采集装置包括用于分别采集佩戴者四个视角的图片和视频的四个摄像模块一;所述四个视角是指:左眼、右眼、左下脸部、右下脸部;

[0007] 基于头部可穿戴设备感知的情感识别方法,包括佩戴者情感识别方法;所述佩戴

者情感识别方法包括如下步骤:

[0008] 步骤X1、采集佩戴者的多模态情感数据;佩戴者的多模态情感数据包括通过四个摄像模块一获取的图片数据一和视频数据一;图片数据一和视频数据一均包括四个视角数据;

[0009] 步骤X2、对佩戴者的多模态情感数据进行情感特征提取:

[0010] 对图片数据一和视频数据一,分别采用局部融合情感识别网络进行处理;其中,图片数据一分别以四个视角数据作为局部融合情感识别网络的四个视角输入;视频数据一分别对四个视角数据提取起始帧和峰值帧,作为局部融合情感识别网络的四个视角输入;

[0011] 所述四个视角输入在局部融合情感识别网络中的处理方式是:左眼输入和右眼输入进行深度卷积以提取局部视角特征;左下脸部输入和右下脸部输入分别通过嵌入层将提取到的动作单元嵌入,然后与面部行为编码FACS一起输入到空域图卷积中以提取局部视角特征;将四个视角输入提取得到的局部视角特征同时输入到多层感知机中进行空间映射,计算空间注意力和通道注意力后进行特征图融合,得到最终的情感特征;

[0012] 步骤X3、将步骤X2得到的情感特征进行融合,通过分类得到复合情感识别结果。

[0013] 优选地,所述局部融合情感识别网络为四个,分别为局部融合情感识别网络一、局部融合情感识别网络二、局部融合情感识别网络三和局部融合情感识别网络四;

[0014] 所述局部融合情感识别网络一对图片数据一进行处理,得到图片数据一的宏表情情感特征;局部融合情感识别网络二对图片数据一进行处理,得到图片数据一的微表情情感特征;局部融合情感识别网络三对视频数据一进行处理,得到视频数据一的宏表情情感特征;局部融合情感识别网络四对视频数据一进行处理,得到视频数据一的微表情情感特征。

[0015] 优选地,四个局部融合情感识别网络均包括分别四个视角输入的四个局部特征提取单元;针对左眼输入和右眼输入的两个局部特征提取单元均包括深度卷积网络一;针对左下脸部输入和右下脸部输入的两个局部特征提取单元均由嵌入层、空域图卷积网络依次连接组成;嵌入层还与动作单元提取器连接;空域图卷积网络还与面部动作编码系统连接;左下脸部输入和右下脸部输入;四个局部特征提取单元的输出同时与多层感知机连接,并通过通道注意力和空间注意力进行融合;

[0016] 其中,所述局部融合情感识别网络三和局部融合情感识别网络四,针对左眼输入和右眼输入的两个局部特征提取单元分别还包括动作放大网络;左眼输入和右眼输入分别通过动作放大网络将微笑表情放大,再输入深度卷积网络一,提取局部视角特征。

[0017] 优选地,所述步骤X2中,对图片数据一和视频数据一,在采用局部融合情感识别网络进行处理之前,先分别进行预处理;

[0018] 对图片数据一,预处理包括利用串联的预处理卷积神经网络一进行人脸检测;利用串联的预处理卷积神经网络一进行人脸检测,是指:生成候选框,并对候选框进行初步筛选,对人脸关键点进行检测;经过卷积、激活函数、池化、全连接处理后,输出每个候选框的置信度、坐标偏移量和五个关键点的坐标,以实现人脸检测;

[0019] 对视频数据一,预处理包括利用串联的预处理多层深度卷积神经网络二进行人脸检测;利用串联的预处理多层深度卷积神经网络二进行人脸检测,是指:使用视频流方式,逐帧读取视频数据一;对视频数据一的每一帧图像,利用串联的预处理多层深度卷积神经

网络二对图像改变尺寸后的金字塔形数据进行操作,得到人脸框、关键点坐标和人脸分类,以实现人脸检测;预处理多层深度卷积神经网络二包括依次连接的图像改变尺寸层、卷积神经单元一、卷积神经单元二、最大池化层一、全连接层一、卷积神经单元三、最大池化层二和全连接层二,以及连接卷积神经单元三与最大池化层二之间的空间注意力层。

[0020] 优选地,所述步骤X3,是指:采用多模态自适应融合模块将步骤X2得到的情感特征进行融合:多模态自适应融合模块的输入为情感特征 $X=\{X_1, \dots, X_n\}$,其中 X_i 为第 i 个情感特征, n 为情感特征的数量;利用注意力机制并且迭代地进行特征融合,最终得到融合特征;将融合特征输入到分类器中进行学习,得到复合情感识别结果;复合情感识别结果采用情感状态的复合表示;情感状态的复合表示形式为:情感类别及对应比例。

[0021] 优选地,所述头部可穿戴设备,多模态数据采集装置还包括音频采集模块;所述步骤X1中,多模态情感数据还包括音频数据;所述步骤X2中,还对音频数据进行情感特征提取:对音频数据进行滤波、平滑、分帧操作;提取梅尔倒谱系数特征;利用梅尔倒谱系数特征并载入为特征向量形式,输入到基于注意力机制的BiLSTM神经网络中进行情感特征提取;

[0022] 所述步骤X1中,多模态情感数据还包括文本数据;所述步骤X2中,还对文本数据进行情感特征提取:采用word2vec模型对文本数据进行处理得到序列上下文词向量表示,再经过基于LSTM情感分析网络进行情感特征提取。

[0023] 优选地,所述头部可穿戴设备,多模态数据采集装置还包括用于采集被观测者图片和视频的摄像模块二;

[0024] 基于头部可穿戴设备感知的情感识别方法,还包括被观测者情感识别方法;被观测者情感识别方法包括如下步骤:

[0025] 步骤Y1、采集被观测者的多模态情感数据;被观测者的多模态情感数据包括通过摄像模块二获取的图片数据二和视频数据二;

[0026] 步骤Y2、对被观测者的多模态情感数据进行情感特征提取:

[0027] 对图片数据二,分别采用表情卷积神经网络一、表情深度卷积网络二和图神经网络进行情感特征提取,得到图片数据二的宏表情情感特征、微表情情感特征、姿态情感特征;

[0028] 对视频数据二,分别采用三维卷积神经网络三、基于峰值帧光流的卷积神经网络四进行情感特征提取,得到视频数据二的宏表情情感特征、微表情情感特征;

[0029] 所述基于峰值帧光流的卷积神经网络四进行情感特征提取,是指:首先对视频数据二进行旋转、裁剪、人脸对齐的预处理,利用峰值帧检测算法提取出峰值帧;接着,计算视频数据二的起始帧与峰值帧之间的光流矢量 u, v 以及光应变 ε ;灰度化处理后,分别作为RGB图像的三个通道,合成为一张RGB图像;对RGB图像进行情感特征提取;

[0030] 光流矢量 u, v 以及光应变 ε 采用如下公式组进行计算:

$$[0031] \quad I(x, y, t) = I(x, y, t) + \frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial t} dt + \varepsilon;$$

$$[0032] \quad u = \frac{dx}{dt};$$

$$[0033] \quad \mathbf{v} = \frac{dy}{dt};$$

[0034] 其中, $I(x, y, t)$ 表示像素点在起始帧的光强度; t 表示所在的时间维度; dx, dy 分别表示起始帧到峰值帧的横坐标和纵坐标的移动值; dt 表示起始帧移动到峰值帧所用时间;

令 $I_x = \frac{\partial I}{\partial x}$, $I_y = \frac{\partial I}{\partial y}$, $I_t = \frac{\partial I}{\partial t}$, 分别表示像素点的灰度沿 x 、 y 、 t 方向的偏导数;

$I_x u + I_y v + I_t = 0$, 其中, I_x, I_y, I_t 均由图像数据求得;

[0035] 步骤Y3、将步骤Y2得到的情感特征进行融合, 通过分类得到复合情感识别结果。

[0036] 优选地, 还包括个体数据存储数据库的建立方法: 在个体数据存储数据库构建个体ID; 在佩戴者情感识别方法或被观测者情感识别方法开始前, 先获取佩戴者或被观测者的个体ID; 在佩戴者情感识别方法或被观测者情感识别方法完成后, 将佩戴者情感识别方法或被观测者情感识别方法得到的情感识别结果以及时间、地点存储到对应的个体ID中。

[0037] 一种可读存储介质, 其中所述存储介质存储有计算机程序, 所述计算机程序当被处理器执行时使所述处理器执行所述的基于头部可穿戴设备感知的情感识别方法。

[0038] 一种计算机设备, 包括处理器以及用于存储处理器可执行程序存储器, 所述处理器执行存储器存储的程序时, 实现所述的基于头部可穿戴设备感知的情感识别方法。

[0039] 与现有技术相比, 本发明具有如下优点与有益效果:

[0040] 1、本发明基于头部可穿戴设备, 可实时采集佩戴者的数据, 进而实现佩戴者情感感知和识别; 可采集佩戴者的局部多视角表情数据, 采用面部动作单元信息辅助宏表情和微表情的局部融合情感识别网络, 提高宏表情和微表情特征的表达能力;

[0041] 2、本发明能够识别复合性的情感状态, 这符合神经学中对情感的认知, 情感是多维、多层次的体验;

[0042] 3、本发明支持对多模态情感数据进行情感特征提取, 融合, 具备多模态情感数据相互支持、相互补充的能力;

[0043] 4、本发明实现了跨个体的情感信号采集, 能够同时采集和处理使用者与多个被测者的情感信号, 扩展了应用场景; 例如, 在辅助精神疾病问诊、日常交流、精神疾病筛查以及精神疾病患者日常情绪变化监测等场景下均可使用; 从而使本发明具备提高辅助筛查、诊断、监测效率的能力;

[0044] 5、本发明将采集到的数据和情感分析结果建立个体数据存储数据库; 有助于挖掘家居生活中的情感需求、辅助心理疾病诊断、记录日常情感状态; 通过人脸、设备的历史活动和数据记录结合实时数据进行情感状态的回顾和分析, 建立用户的情感个人画像, 了解个体情感在时间、环境和经历下的变化; 同时通过设备采集到的地点、背景图片、穿着和面部表情等信息进行处理, 分析影响情感状态的因素。

附图说明

[0045] 图1是本发明头部可穿戴设备的结构示意图;

[0046] 图2是本发明佩戴者情感识别方法的流程示意图;

[0047] 图3是本发明预处理多层深度卷积神经网络二的结构图;

- [0048] 图4是本发明局部融合情感识别网络一和局部融合情感识别网络二的结构图；
[0049] 图5是本发明局部融合情感识别网络三和局部融合情感识别网络四的结构图；
[0050] 图6是本发明多模态自适应融合模块的结构图；
[0051] 图7是本发明佩戴者情感识别方法的优选方式流程示意图；
[0052] 图8是实施例二头部可穿戴设备的结构示意图；
[0053] 图9是实施例二被观测者情感识别方法的流程示意图。

具体实施方式

[0054] 下面结合附图与具体实施方式对本发明作进一步详细的描述。

[0055] 实施例一

[0056] 本实施例一种基于头部可穿戴设备感知的情感识别方法,头部可穿戴设备包括:设备本体1和设置在设备本体上的多模态数据采集装置;本实施例中,设备本体1采用眼镜,如图1所示;多模态数据采集装置包括用于分别采集佩戴者四个视角的图片和视频的四个摄像模块一3、4、5、6;四个视角是指:左眼、右眼、左下脸部、右下脸部;其中,摄像模块一3用于采集佩戴者右眼图片和视频;摄像模块一4用于采集佩戴者左眼图片和视频;摄像模块一5用于采集佩戴者右下脸部图片和视频;摄像模块一6用于采集佩戴者左下脸部图片和视频。

[0057] 基于头部可穿戴设备感知的情感识别方法,包括佩戴者情感识别方法。

[0058] 佩戴者情感识别方法,如图2所示,包括如下步骤:

[0059] 步骤X1、采集佩戴者的多模态情感数据;佩戴者的多模态情感数据包括通过四个摄像模块一获取的图片数据一和视频数据一;图片数据一和视频数据一均包括四个视角数据;

[0060] 步骤X2、对佩戴者的多模态情感数据进行情感特征提取:

[0061] 首先,优选对图片数据一和视频数据一,分别进行预处理:

[0062] 对图片数据一,预处理包括利用串联的预处理卷积神经网络一进行人脸检测;利用串联的预处理卷积神经网络一进行人脸检测,是指:生成候选框,并对候选框进行初步筛选,对人脸关键点进行检测;经过卷积、激活函数、池化、全连接处理后,输出每个候选框的置信度、坐标偏移量和五个关键点的坐标,以实现人脸检测;

[0063] 对视频数据一,预处理包括利用串联的预处理多层深度卷积神经网络二进行人脸检测;利用串联的预处理多层深度卷积神经网络二进行人脸检测,是指:使用视频流方式,逐帧读取视频数据一;对视频数据一的每一帧图像,利用串联的预处理多层深度卷积神经网络二对图像改变尺寸后的金字塔形数据进行操作,得到人脸框、关键点坐标和人脸分类,以实现人脸检测;预处理多层深度卷积神经网络二,如图3所示,包括依次连接的图像改变尺寸层、卷积神经单元一、卷积神经单元二、最大池化层一、全连接层一、卷积神经单元三、最大池化层二和全连接层二,以及连接卷积神经单元三与最大池化层二之间的空间注意力层。

[0064] 然后,对图片数据一和视频数据一,分别采用局部融合情感识别网络进行处理;其中,图片数据一分别以四个视角数据作为局部融合情感识别网络的四个视角输入;视频数据一分别对四个视角数据提取起始帧和峰值帧,作为局部融合情感识别网络的四个视角输

入。

[0065] 具体地说,局部融合情感识别网络为四个,分别为局部融合情感识别网络一、局部融合情感识别网络二、局部融合情感识别网络三和局部融合情感识别网络四。

[0066] 局部融合情感识别网络一对图片数据一进行处理,得到图片数据一的宏表情情感特征;局部融合情感识别网络二对图片数据一进行处理,得到图片数据一的微表情情感特征。

[0067] 局部融合情感识别网络一和局部融合情感识别网络二,结构如图4所示,均包括分别四个视角输入的四个局部特征提取单元。针对左眼输入和右眼输入的两个局部特征提取单元均包括深度卷积网络一;针对左下脸部输入和右下脸部输入的两个局部特征提取单元均由嵌入层、空域图卷积网络依次连接组成;嵌入层还与动作单元提取器连接;空域图卷积网络还与面部动作编码系统连接;左下脸部输入和右下脸部输入;四个局部特征提取单元的输出同时与多层感知机连接,并通过通道注意力和空间注意力进行融合。

[0068] 局部融合情感识别网络三对视频数据一进行处理,得到视频数据一的宏表情情感特征;局部融合情感识别网络四对视频数据一进行处理,得到视频数据一的微表情情感特征。

[0069] 局部融合情感识别网络三和局部融合情感识别网络四,结构如图5所示,均包括分别四个视角输入的四个局部特征提取单元。针对左眼输入和右眼输入的两个局部特征提取单元均包括动作放大网络和深度卷积网络一;左眼输入和右眼输入分别通过动作放大网络将微笑表情放大,再输入深度卷积网络一,提取局部视角特征。针对左下脸部输入和右下脸部输入的两个局部特征提取单元均由嵌入层、空域图卷积网络依次连接组成;嵌入层还与动作单元提取器连接;空域图卷积网络还与面部动作编码系统连接;左下脸部输入和右下脸部输入;四个局部特征提取单元的输出同时与多层感知机连接,并通过通道注意力和空间注意力进行融合。

[0070] 四个视角输入在局部融合情感识别网络中的处理方式是:左眼输入和右眼输入进行深度卷积以提取局部视角特征;左下脸部输入和右下脸部输入分别通过嵌入层将提取到的动作单元嵌入,然后与面部行为编码FACS一起输入到空域图卷积中以提取局部视角特征;FACS (Facial Action Coding System) 即面部行为编码,它特指一组面部肌肉运动状。将四个视角输入提取得到的局部视角特征同时输入到多层感知机中进行空间映射,计算空间注意力和通道注意力后进行特征图融合,得到最终的情感特征。

[0071] 四个局部融合情感识别网络采用卷积神经网络进行情感特征的提取,同时采用基于特权动作单元信息的面部表情识别网络,将AUs (AUs是面部运动单元,用于描述人脸的最小运动单位和表情特征) 作为特权信息,指导情感识别;同时,将AUs作为网络浅层的辅助输出标签,以此辅助模型中的浅层特征表达。

[0072] 步骤X3、采用多模态自适应融合模块将步骤X2得到的情感特征进行融合:多模态自适应融合模块的结构如图6所示,输入为情感特征 $X=\{X_1, \dots, X_n\}$,其中 X_i 为第 i 个情感特征, n 为情感特征的数量;利用注意力机制并且迭代地进行特征融合,最终得到融合特征;将融合特征输入到分类器中进行学习,得到复合情感识别结果;复合情感识别结果采用情感状态的复合表示;情感状态的复合表示形式为:情感类别及对应比例。例如,开心-75%、悲伤-1%、惊讶-5%、害怕1.5%、厌恶-1.5%、生气-1%、中立-15%。

[0073] 本实施例中,头部可穿戴设备,多模态数据采集装置优选还包括音频采集模块7,例如麦克风;

[0074] 佩戴者情感识别方法,优选方案如图7所示:

[0075] 步骤X1中,多模态情感数据还包括音频数据和文本数据;文本数据可通过小程序/移动端APP获取;

[0076] 步骤X2中,还对音频数据进行情感特征提取:对音频数据进行滤波、平滑、分帧操作;提取梅尔倒谱系数特征;利用梅尔倒谱系数特征并载入为特征向量形式,输入到基于注意力机制的BiLSTM神经网络中进行情感特征提取;

[0077] 还对文本数据进行情感特征提取:采用word2vec模型对文本数据进行处理得到序列上下文词向量表示,再经过基于LSTM情感分析网络进行情感特征提取。

[0078] 本实施例,优选还包括个体数据存储数据库的建立方法:在个体数据存储数据库构建个体ID;在佩戴者情感识别方法或被观测者情感识别方法开始前,先获取佩戴者或被观测者的个体ID;在佩戴者情感识别方法或被观测者情感识别方法完成后,将佩戴者情感识别方法或被观测者情感识别方法得到的情感识别结果以及时间、地点存储到对应的个体ID中。地点可通过头部可穿戴设备的GPS模块2获取。

[0079] 通过使用数据库技术将设备采集到的个体ID、时间、地点、情感状态、语音高频词、图片背景信息、情感状态结果、穿着(颜色)等数据记录保存下来,通过数据统计以及结合实时数据进行情感状态的回顾和分析,建立用户的情感个人画像,了解个体情感在时间、环境和经历下的变化。分析影响情感状态的因素。

[0080] 通过建立个体时间维度的情感状态画像,统计个体语音高频词汇,通过柱状图进行可视化,针对不同时间下的情感状态,给出情绪变化曲线图可视化以及高频情绪出现的频次统计图,对于出现消极情绪较多的情况,给出负面情绪预警,同时给出相应的参考建议。

[0081] 通过建立空间维度的个体情感状态画像,使用位置数据、图片背景、情感状态、穿着、高频词汇等数据的统计和分析,分析影响情感的因素,包括环境、地点对情绪的影响,例如统计在同一个环境中的情感状态,并给出简单的统计报告。

[0082] 实施例二

[0083] 本实施例一种基于头部可穿戴设备感知的情感识别方法,与实施例一的区别在于:本实施例中,头部可穿戴设备,多模态数据采集装置还包括用于采集被观测者图片和视频的摄像模块二8,如图8所示。

[0084] 基于头部可穿戴设备感知的情感识别方法,还包括被观测者情感识别方法;被观测者情感识别方法,如图9所示,包括如下步骤:

[0085] 步骤Y1、采集被观测者的多模态情感数据;被观测者的多模态情感数据包括通过摄像模块二8获取的图片数据二和视频数据二;

[0086] 步骤Y2、对被观测者的多模态情感数据进行情感特征提取:

[0087] 对图片数据二,分别采用表情卷积神经网络一、表情深度卷积网络二和图神经网络进行情感特征提取,得到图片数据二的宏表情情感特征、微表情情感特征、姿态情感特征;

[0088] 对视频数据二,分别采用三维卷积神经网络三、基于峰值帧光流的卷积神经网络

四进行情感特征提取,得到视频数据二的宏表情情感特征、微表情情感特征;

[0089] 所述基于峰值帧光流的卷积神经网络四进行情感特征提取,是指:首先对视频数据二进行旋转、裁剪、人脸对齐的预处理,利用峰值帧检测算法提取出峰值帧;接着,计算视频数据二的起始帧与峰值帧之间的光流矢量 u, v 以及光应变 ε ;灰度化处理后,分别作为RGB图像的三个通道,合成为一张RGB图像;对RGB图像进行情感特征提取;

[0090] 光流矢量 u, v 以及光应变 ε 采用如下公式组进行计算:

$$[0091] \quad I(x, y, t) = I(x, y, t) + \frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial t} dt + \varepsilon;$$

$$[0092] \quad u = \frac{dx}{dt};$$

$$[0093] \quad v = \frac{dy}{dt};$$

[0094] 其中, $I(x, y, t)$ 表示像素点在起始帧的光强度; t 表示所在的时间维度; dx, dy 分别表示起始帧到峰值帧的横坐标和纵坐标的移动值; dt 表示起始帧移动到峰值帧所用时间;

令 $I_x = \frac{\partial I}{\partial x}$, $I_y = \frac{\partial I}{\partial y}$, $I_t = \frac{\partial I}{\partial t}$, 分别表示像素点的灰度沿 x, y, t 方向的偏导数;

$I_x u + I_y v + I_t = 0$, 其中, I_x, I_y, I_t 均由图像数据求得。

[0095] 在步骤Y1和步骤Y2中,被观测者的多模态情感数据还可以包括音频数据二;还对音频数据二进行情感特征提取。

[0096] 步骤Y3、将步骤Y2得到的情感特征进行融合,通过分类得到复合情感识别结果。

[0097] 实施例三

[0098] 本实施例一种可读存储介质,其中所述可读存储介质存储有计算机程序,所述计算机程序当被处理器执行时使所述处理器执行实施例一或实施例二所述的基于头部可穿戴设备感知的情感识别方法。

[0099] 实施例四

[0100] 本实施例一种计算机设备,包括处理器以及用于存储处理器可执行程序存储器,所述处理器执行存储器存储的程序时,实现实施例一或实施例二所述的基于头部可穿戴设备感知的情感识别方法。

[0101] 上述实施例为本发明较佳的实施方式,但本发明的实施方式并不受上述实施例的限制,其他的任何未背离本发明的精神实质与原理下所作的改变、修饰、替代、组合、简化,均应为等效的置换方式,都包含在本发明的保护范围之内。

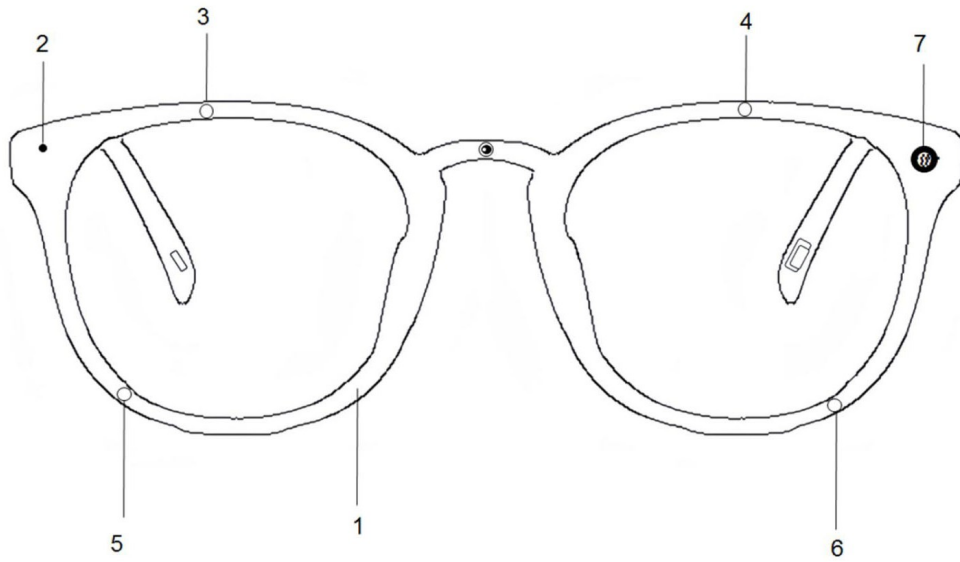


图1

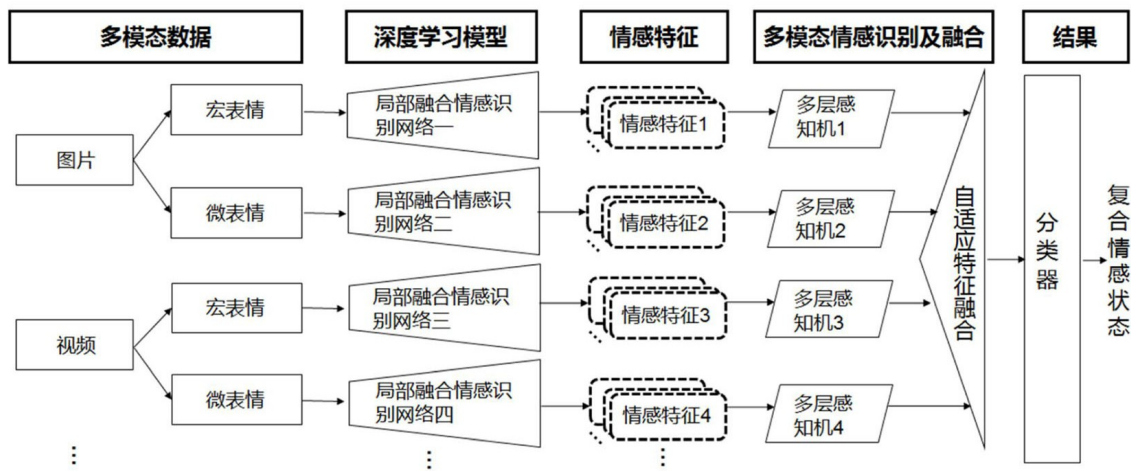


图2

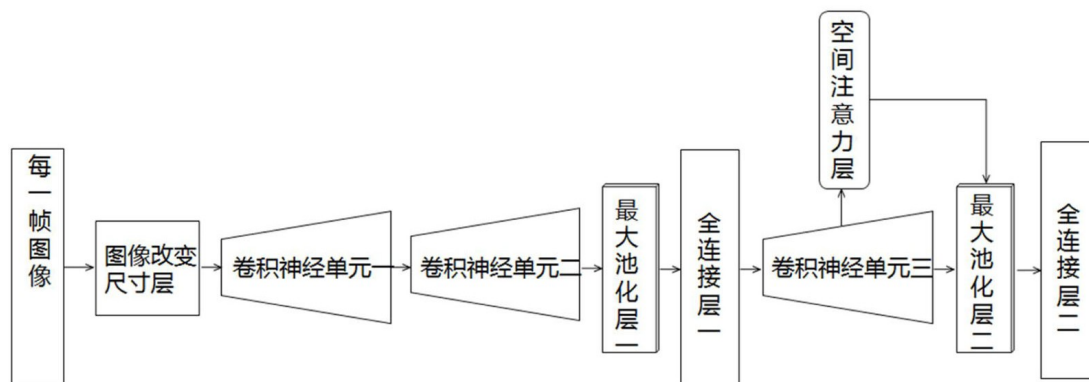


图3

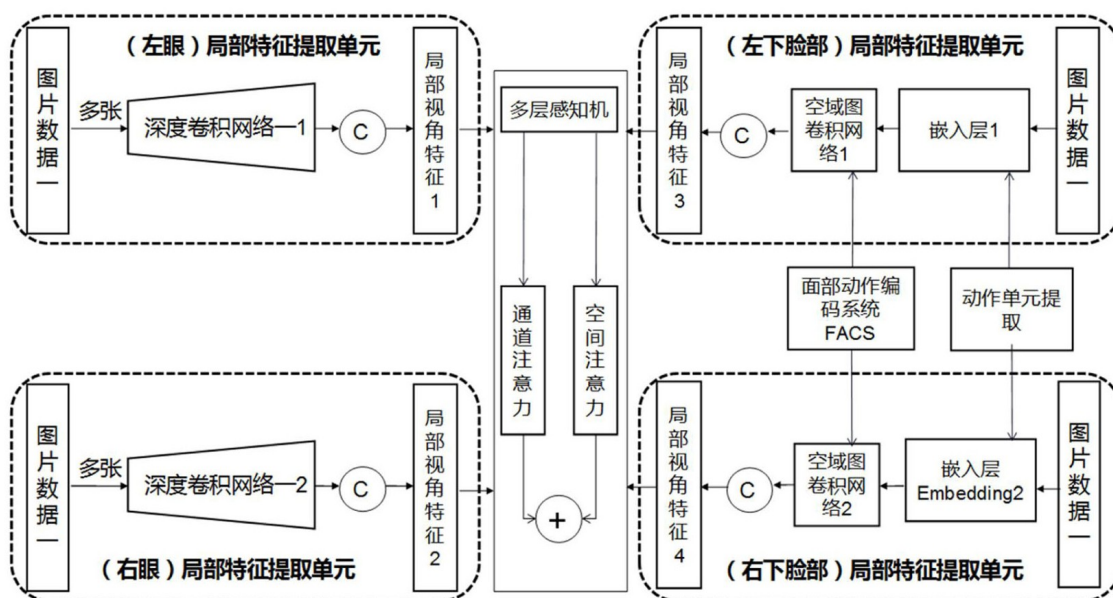


图4

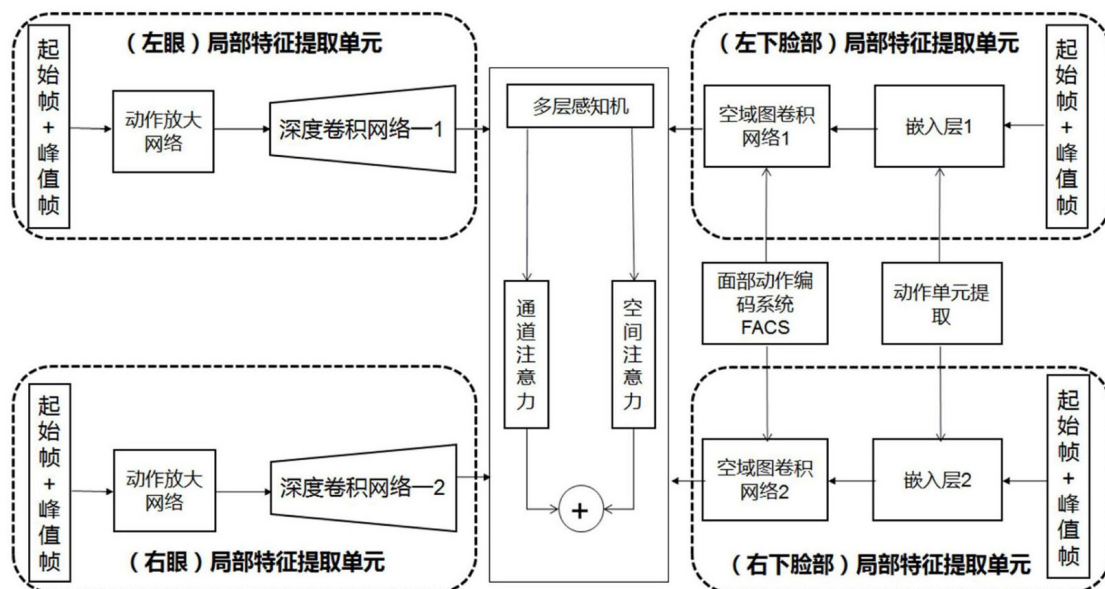


图5

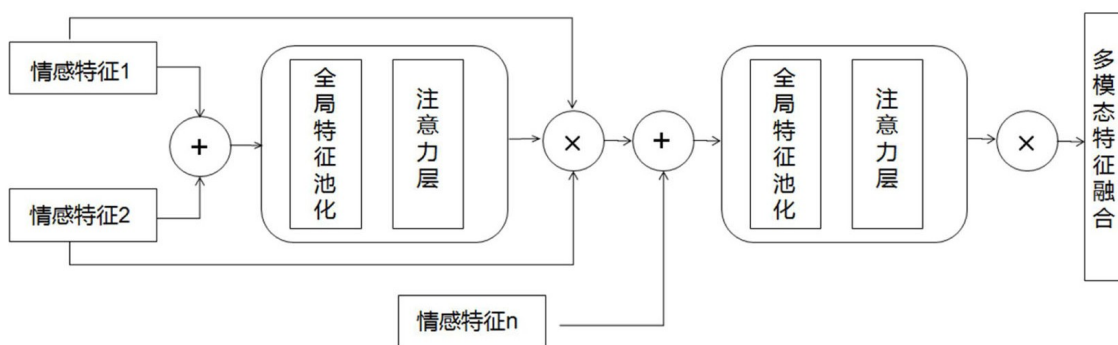


图6

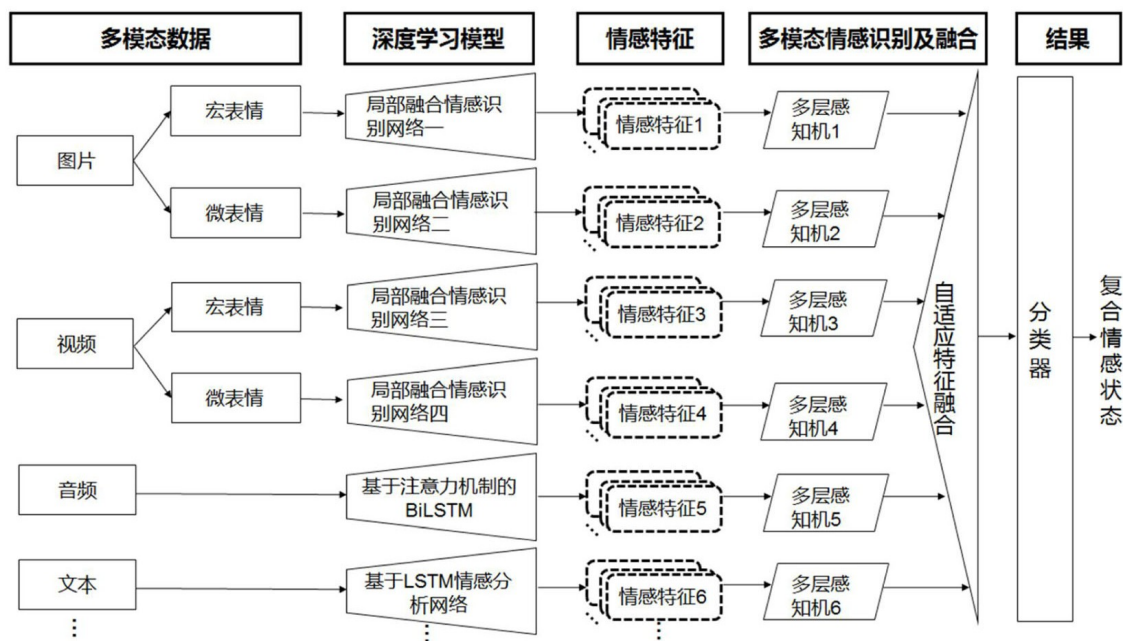


图7

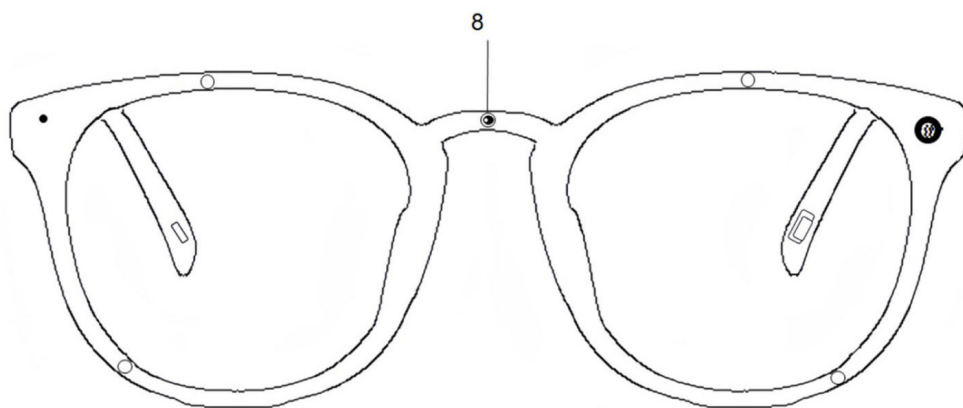


图8

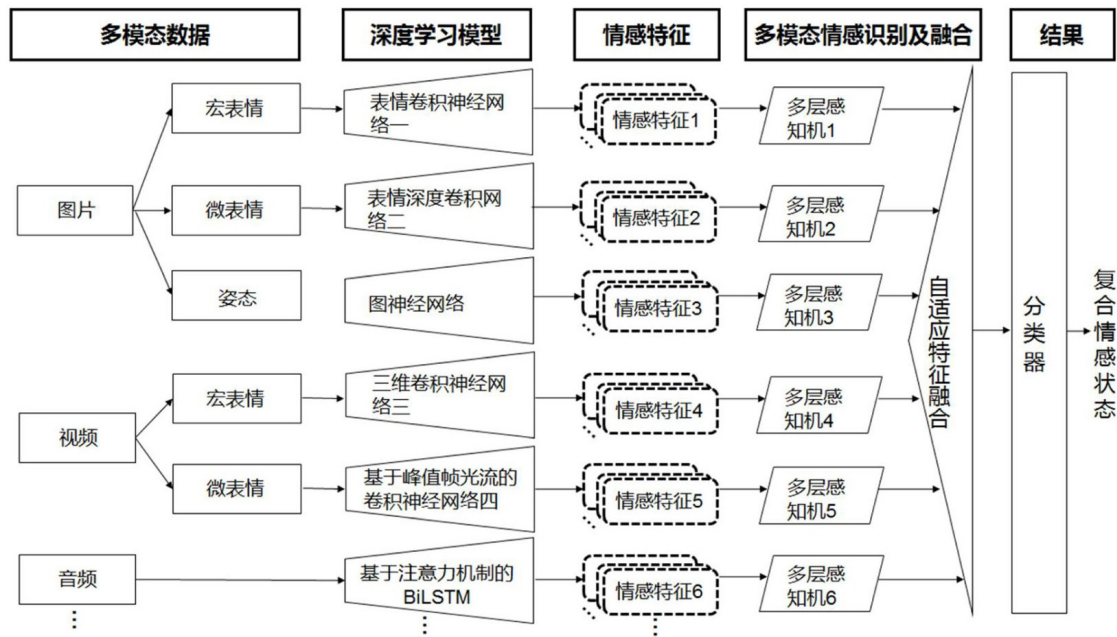


图9