



(12)发明专利申请

(10)申请公布号 CN 110516571 A

(43)申请公布日 2019. 11. 29

(21)申请号 201910756936.X

(22)申请日 2019.08.16

(71)申请人 东南大学

地址 211102 江苏省南京市江宁区东南大
学路2号

(72)发明人 郑文明 夏万闯 宗源 江星洵
路成 刘佳腾

(74)专利代理机构 南京苏高专利商标事务所
(普通合伙) 32204

代理人 冯艳芬

(51)Int.Cl.

G06K 9/00(2006.01)

G06K 9/62(2006.01)

G06N 3/04(2006.01)

G06N 3/08(2006.01)

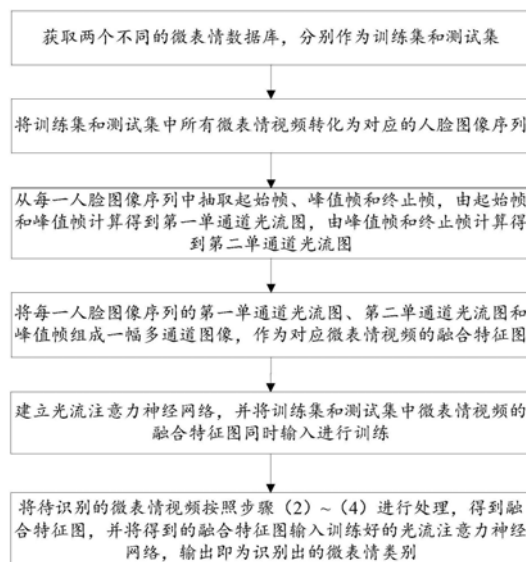
权利要求书2页 说明书6页 附图3页

(54)发明名称

基于光流注意力神经网络的跨库微表情识别方法及装置

(57)摘要

本发明公开了一种基于光流注意力神经网络的跨库微表情识别方法及装置,方法包括:(1)获取两个不同的微表情数据库作为训练集和测试集;(2)将微表情视频转化为人脸图像序列;(3)从每一人脸图像序列中抽取起始帧、峰值帧和终止帧,由起始帧和峰值帧计算得到第一单通道光流图,由峰值帧和终止帧计算得到第二单通道光流图;(4)将每一人脸图像序列的第一单通道光流图、第二单通道光流图和峰值帧组成融合特征图;(5)建立光流注意力神经网络,并将训练集和测试集对应融合特征图作为输入进行训练;(6)将待识别的微表情视频处理得到融合特征图,输入光流注意力神经网络,得到微表情类别。本发明泛化能力强,识别准确率高。



1. 一种基于光流注意力神经网络的跨库微表情识别方法,其特征在于该方法包括:

(1) 获取两个不同的微表情数据库,分别作为训练集和测试集,其中,每个数据库中包含有若干个微表情视频及其对应的微表情类别标签;

(2) 将训练集和测试集中所有微表情视频转化为对应的人脸图像序列;

(3) 从每一人脸图像序列中抽取起始帧、峰值帧和终止帧,由起始帧和峰值帧计算得到第一单通道光流图,由峰值帧和终止帧计算得到第二单通道光流图;

(4) 将每一人脸图像序列的第一单通道光流图、第二单通道光流图和峰值帧组成一幅多通道图像,作为对应微表情视频的融合特征图;

(5) 建立光流注意力神经网络,并将训练集和测试集中微表情视频的融合特征图同时输入进行训练,训练时采用的整体损失函数为:训练集对应的神经网络输出和测试集对应的神经网络输出的最大均值差异,加上训练集对应的神经网络输出经过Softmax处理得到的微表情类别和训练集中存储的微表情类别标签的交叉熵;

(6) 将待识别的微表情视频按照步骤(2)~(4)进行处理,得到融合特征图,并将得到的融合特征图输入训练好的光流注意力神经网络,输出即为识别出的微表情类别。

2. 根据权利要求1所述的基于光流注意力神经网络的跨库微表情识别方法,其特征在于:训练集和测试集中,同一类别的微表情视频的分类标签相同。

3. 根据权利要求1所述的基于光流注意力神经网络的跨库微表情识别方法,其特征在于:步骤(2)具体包括:

(2-1) 获取训练集和测试集中所有微表情视频的起始帧;

(2-2) 对起始帧进行人脸检测,得到人脸位置信息;

(2-3) 根据起始帧的人脸位置信息,从微表情视频的每帧图像中裁剪出矩形人脸图像,得到人脸图像序列。

4. 根据权利要求1所述的基于光流注意力神经网络的跨库微表情识别方法,其特征在于:步骤(3)具体包括:

(3-1) 从所获取的人脸图像序列中抽取起始帧、峰值帧和终止帧;

(3-2) 计算起始帧和峰值帧之间、峰值帧和终止帧之间的水平方向光流场和垂直方向光流场;

(3-3) 对于起始帧和峰值帧之间的水平方向光流场和垂直方向光流场,将其由笛卡尔坐标转换为极坐标,将相应坐标的值转换为HSV颜色空间的值,之后转换为灰度图像,作为第一单通道光流图;

(3-4) 对于峰值帧和终止帧之间的水平方向光流场和垂直方向光流场,按照步骤(3-3)做同样处理,得到第二单通道光流图。

5. 根据权利要求1所述的基于光流注意力神经网络的跨库微表情识别方法,其特征在于:步骤(5)具体包括:

(5-1) 建立光流注意力神经网络,所述光流注意力神经网络包括五个卷积层、六个线性整流层、四个最大池化层和两个全连接层,前四个卷积层后都依次连接一个线性整流层和一个最大池化层,第五个卷积层后依次连接一个线性整流层、一个全连接层、一个线性整流层和一个全连接层,最后一个全连接层的输出为神经网络输出;

(5-2) 将训练集和测试集中微表情视频的融合特征图同时输入光流注意力神经网络进

行训练,训练时采用的整体损失函数 L_{all} 为:

$$L_{all} = L_s + \lambda MMD(D_s, D_t)$$

$$L_s = \frac{1}{n} \sum_x -p(x) \log q(x)$$

$$MMD(D_s, D_t) = \frac{1}{M^2} \sum_{i=1}^M \sum_{j=1}^M k(d_i^s, d_j^s) + \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N k(d_i^t, d_j^t) - \frac{2}{MN} \sum_{i=1}^M \sum_{j=1}^N k(d_i^s, d_j^t)$$

式中, L_s 为训练集输出经过Softmax处理得到的微表情类别 $q(x)$ 和训练集中存储的微表情类别标签 $p(x)$ 的交叉熵, x 表示对应样本, n 表示Softmax处理后的输出个数, λ 表示权重系数, $MMD(D_s, D_t)$ 为 D_s, D_t 的最大均值差异, $D_s = [d_1^s, d_2^s, \dots, d_M^s]$ 为训练集输出,元素为 M 个, $D_t = [d_1^t, d_2^t, \dots, d_N^t]$ 为测试集输出,元素为 N 个, $k(\cdot)$ 表示高斯核函数。

6.一种基于光流注意力神经网络的跨库微表情识别装置,其特征在于:包括存储器和处理器,所述存储器中存储有计算机程序,所述处理器用于执行所述计算机程序时实现权利要求1-5中任意一项所述的方法。

基于光流注意力神经网络的跨库微表情识别方法及装置

技术领域

[0001] 本发明涉及图像处理技术,尤其涉及一种基于光流注意力神经网络的跨库微表情识别方法及装置。

背景技术

[0002] 微表情(Micro-Expression)是一种人类在试图隐藏某种情感时无意识做出的短暂的面部表情。微表情是一种重要的真实情感信息,通常能够有效地反应一个人的真实心理状态,被认为是识别谎言的关键信息,对理解人类真实情感状态,有着重要的作用。因此,有效准确识别微表情,对日常生产生活有着重大的意义。例如,在刑侦时,经过微表情识别训练的审讯员可以更有效地判断嫌疑人话语的真实性,更有针对性地获取可靠信息;在临床医疗中,医生可以通过微表情识别有效推测患者的真实状态,从而更有效地与患者交流,改进治疗方案。与普通表情相比,微表情持续非常短,研究表明微表情仅持续 $1/25 \sim 1/3s$ 。与此同时微表情面部肌肉动作幅度也非常小,通常只发生在人脸中很小的一部分区域,不会同时出现在上半脸和下半脸。因此对于没有经过专业训练的普通人,捕捉并正确识别微表情有着相当的难度。但人工识别微表情的经济成本和时间成本较高,难以大规模推广,并且人工识别易受环境的影响。受益于计算机视觉技术的快速发展,当前通过计算机手段自动识别微表情具有一定的可行性和商业价值。

[0003] 微表情识别指将给定的微表情样本识别为某一具体的情绪类别,相关研究主要基于传统机器学习方法,依靠LBP相关的时空描述子特征或光流相关的特征。LBP-TOP特征是局部二值模式(Local Binary Pattern,LBP)在三维空间上的拓展,可有效描述图片序列的动态纹理特征,最早由Pfister等人应用于微表情识别,随后其大量改进模式被广泛应用于微表情识别,如Huang等人提出的完备局部量化模式(Spatial Temporal Completed Local Quantized Pattern,STCLQP);Wang等人提出的六交点局部二值模式(LBP with Six Intersection Points,LBP-SIP)等等。另一主流特征是光流(Optical Flow,OF),时序动态信息被证明是微表情的有效描述手段,而有关光流的研究证实了图像序列的时序动态信息可以有效改进微表情识别的效果,如Xu等人提出的面部动力谱特征(Facial Dynamics Map,FDM);Liu等人提出的主方向平均光流特征(Main Directional Mean Optical flow feature,MDMO)等等。

[0004] 随着深度学习的飞速发展,研究人员也在着手将其应用于微表情识别工作。Kim等人提出了一个基于卷积神经网络(Convolutional Neural Network,CNN)和长短时记忆循环神经网络(Long Short Term Memory Recurrent Neural Network,LSTM RNN)的时空特征学习网络,该方法使用所有微表情视频样本的峰值帧训练CNN学习空间特征,而后使用CNN的输出作为特征训练LSTM网络学习时间特征,这是深度学习在微表情识别工作上的首次尝试。当前绝大部分有关微表情识别的研究仅可以取得60%以上的识别率,距离实际应用仍有一段距离;此外随着微表情数据集的扩充,基于深度学习的方法将是微表情识别在未来的主要研究方向之一,但目前相关研究较少,整体进展也较为缓慢。

[0005] 传统微表情识别往往在单个微表情库上进行训练和测试,而同一微表情数据库数据通常建立在相同的实验范式下,训练集和测试集选自同一数据库,往往不能有效的识别其他微表情数据库中的样本,泛化能力差。

发明内容

[0006] 发明目的:本发明针对现有技术存在的问题,提供一种基于光流注意力神经网络的跨库微表情识别方法,泛化能力更强,准确率更高。

[0007] 技术方案:本发明所述的基于光流注意力神经网络的跨库微表情识别方法包括:

[0008] (1) 获取两个不同的微表情数据库,分别作为训练集和测试集,其中,每个数据库中包含有若干个微表情视频及其对应的微表情类别标签;

[0009] (2) 将训练集和测试集中所有微表情视频转化为对应的人脸图像序列;

[0010] (3) 从每一人脸图像序列中抽取起始帧、峰值帧和终止帧,由起始帧和峰值帧计算得到第一单通道光流图,由峰值帧和终止帧计算得到第二单通道光流图;

[0011] (4) 将每一人脸图像序列的第一单通道光流图、第二单通道光流图和峰值帧组成一幅多通道图像,作为对应微表情视频的融合特征图;

[0012] (5) 建立光流注意力神经网络,并将训练集和测试集中微表情视频的融合特征图同时输入进行训练,训练时采用的整体损失函数为:训练集对应的神经网络输出和测试集对应的神经网络输出的最大均值差异,加上训练集对应的神经网络输出经过Softmax 处理得到的微表情类别和训练集中存储的微表情类别标签的交叉熵;

[0013] (6) 将待识别的微表情视频按照步骤(2)~(4)进行处理,得到融合特征图,并将得到的融合特征图输入训练好的光流注意力神经网络,输出即为识别出的微表情类别。

[0014] 进一步的,训练集和测试集中,同一类别的微表情视频的类别标签相同。

[0015] 进一步的,步骤(2)具体包括:

[0016] (2-1) 获取训练集和测试集中所有微表情视频的起始帧;

[0017] (2-2) 对起始帧进行人脸检测,得到人脸位置信息;

[0018] (2-3) 根据起始帧的人脸位置信息,从微表情视频的每帧图像中裁剪出矩形人脸图像,得到人脸图像序列。

[0019] 进一步的,步骤(3)具体包括:

[0020] (3-1) 从所获取的人脸图像序列中抽取起始帧、峰值帧和终止帧;

[0021] (3-2) 计算起始帧和峰值帧之间、峰值帧和终止帧之间的水平方向光流场和垂直方向光流场;

[0022] (3-3) 对于起始帧和峰值帧之间的水平方向光流场和垂直方向光流场,将其由笛卡尔坐标转换为极坐标,将相应坐标的值转换为HSV颜色空间的值,之后转换为灰度图像,作为第一单通道光流图;

[0023] (3-4) 对于峰值帧和终止帧之间的水平方向光流场和垂直方向光流场,按照步骤(3-3)做同样处理,得到第二单通道光流图。

[0024] 进一步的,步骤(5)具体包括:

[0025] (5-1) 建立光流注意力神经网络,所述光流注意力神经网络包括五个卷积层、六个线性整流层、四个最大池化层和两个全连接层,前四个卷积层后都依次连接一个线性整流

层和一个最大池化层,第五个卷积层后依次连接一个线性整流层、一个全连接层、一个线性整流层和一个全连接层,最后一个全连接层的输出为神经网络的输出;

[0026] (5-2) 将训练集和测试集中微表情视频的融合特征图同时输入光流注意力神经网络进行训练训练时采用的整体损失函数 L_{all} 为:

[0027] $L_{all} = L_s + \lambda MMD(D_s, D_t)$

[0028] $L_s = -\frac{1}{n} \sum_x p(x) \log q(x)$

[0029] $MMD(D_s, D_t) = \frac{1}{M^2} \sum_{i=1}^M \sum_{j=1}^M k(d_i^s, d_j^s) + \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N k(d_i^t, d_j^t) - \frac{2}{MN} \sum_{i=1}^M \sum_{j=1}^N k(d_i^s, d_j^t)$

[0030] 式中, L_s 为训练集输出经过Softmax处理得到的微表情类别 $q(x)$ 和训练集中存储的微表情类别标签 $p(x)$ 的交叉熵, x 表示对应样本, n 表示Softmax处理后的输出个数, λ 表示权重系数, $MMD(D_s, D_t)$ 为 D_s, D_t 的最大均值差异, $D_s = [d_1^s, d_2^s, \dots, d_M^s]$ 为训练集输出,元素为 M 个, $D_t = [d_1^t, d_2^t, \dots, d_N^t]$ 为测试集输出,元素为 N 个, $k()$ 表示高斯核函数。

[0031] 本发明所述的基于光流注意力神经网络的跨库微表情识别装置包括存储器和处理器,所述存储器中存储有计算机程序,所述处理器用于执行所述计算机程序时实现上述方法。

[0032] 有益效果:本发明与现有技术相比,其显著优点是:本发明提供一种基于光流注意力神经网络的跨库微表情识别方法及装置,通过光流法有效地关注与某种微表情强相关的部分,使卷积神经网络能有效地提取与微表情相关的特征,有效地提高了识别率;并且利用最大均值差异来降低不同数据库对于算法结果的影响,有效地提高了模型的泛化能力。

附图说明

[0033] 图1是本发明提供的基于光流注意力神经网络的跨库微表情识别方法的一个实施例的流程示意图;

[0034] 图2是CASME2数据库的一个样本的起始帧、峰值帧和终止帧;

[0035] 图3是起始帧和峰值帧的光流图以及峰值帧和终止帧的光流图;

[0036] 图4是光流图和峰值帧的融合特征图;

[0037] 图5是光流注意力神经网络训练过程的示意图;

[0038] 图6是光流注意力神经网络识别过程的示意图。

具体实施方式

[0039] 本实施例提供了一种基于光流注意力神经网络的跨库微表情识别方法,如图1所示,包括:

[0040] (1) 获取两个不同的微表情数据库,分别作为训练集和测试集,其中,每个数据库中包含有若干个微表情视频及其对应的微表情类别标签。

[0041] 其中,训练集和测试集来自不同数据库,可能存在标签不统一的情况,因此会对训练集和测试集中微表情类别标签,根据标签的定义,统一两个数据库的微表情类别,使得同一类别的微表情视频的类别标签相同,对于无法统一的微表情视频,进行删除。本实施例在

CASM2微表情数据库、SMM微表情数据库和SMIC-HS数据库之间做跨库微表情识别,由于类别不统一,对不同微表情库标签进行处理: CASME2中, happy类的标签归为positive, sadness、disgust和fear类归为negative, surprise类归于surprise; SMM数据库处理和CASME2相同, SMIC库标签不做处理。

[0042] (2) 将训练集和测试集中所有微表情视频转化为对应的人脸图像序列。

[0043] 该步骤具体包括:

[0044] (2-1) 获取训练集和测试集中所有微表情视频的起始帧;

[0045] (2-2) 对起始帧进行人脸检测, 得到人脸位置信息;

[0046] (2-3) 根据起始帧的人脸位置信息, 从微表情视频的每帧图像中裁剪出矩形人脸图像, 得到人脸图像序列。本实施例中, 采用face_recognition的face_landmark函数去检测起始帧的人脸关键点, 即人脸位置信息, 并按照起始帧的位置信息裁剪出整个视频的人脸, 人脸框大小依据为: 水平方向为左脸颊的最左点的x坐标点到右脸颊最右点的x坐标点, 垂直方向为下巴最下面点的y坐标点到眉毛最上点加上三十个像素点的y坐标点。

[0047] (3) 从每一人脸图像序列中抽取起始帧、峰值帧和终止帧, 由起始帧和峰值帧计算得到第一单通道光流图, 由峰值帧和终止帧计算得到第二单通道光流图。

[0048] 该步骤具体包括:

[0049] (3-1) 从所获取的人脸图像序列中抽取起始帧、峰值帧和终止帧, 如图2所示; 如果在数据库中直接标有峰值帧, 则直接提取即可, 若无, 就取最中间的一帧作为峰值帧;

[0050] (3-2) 计算起始帧和峰值帧之间、峰值帧和终止帧之间的水平方向光流场和垂直方向光流场; 光流场通过Gunnar-Farneback算法计算;

[0051] (3-3) 对于起始帧和峰值帧之间的水平方向光流场和垂直方向光流场, 水平光流场为笛卡尔坐标系中x坐标值, 垂直光流场为笛卡尔坐标系中y坐标值, 将其由笛卡尔坐标转换为极坐标, 得到幅值与角度, 根据HSV颜色空间格式定义, 将所得角度归一化到区间[0, 255], 并赋值给HSV图像的色调(H), HSV图像的饱和度(S)设为255, 上述所得幅值归一化到[0, 255]后, 赋给HSV图像的明度(V), 之后将HSV图像转换为灰度图像, 作为第一单通道光流图, 如图3所示; 其中, 图像转换可以在opencv完成, 也可编程实现;

[0052] (3-4) 对于峰值帧和终止帧之间的水平方向光流场和垂直方向光流场, 按照步骤(3-3)做同样处理, 得到第二单通道光流图。

[0053] (4) 将每一人脸图像序列的第一单通道光流图、第二单通道光流图和峰值帧组成一幅多通道图像, 作为对应微表情视频的融合特征图。

[0054] 本实例中将第一单通道光流图、峰值帧的灰度图和第二单通道光流图依次对应图像的R、G、B通道, 组成一幅多通道图像, 结果如图4所示; 之后还可以对融合特征图进行归一化处理, 增强模型的鲁棒性。

[0055] (5) 建立光流注意力神经网络, 并将训练集和测试集中微表情视频的融合特征图同时输入进行训练, 训练时采用的整体损失函数为: 训练集对应的神经网络输出和测试集对应的神经网络输出的最大均值差异, 加上训练集对应的神经网络输出经过Softmax 处理得到的微表情类别和训练集中存储的微表情类别标签的交叉熵。

[0056] 如图5所示, 该步骤具体包括:

[0057] (5-1) 建立光流注意力神经网络, 所述光流注意力神经网络包括五个卷积层

(conv1~conv5)、六个线性整流层、四个最大池化层和两个全连接层(Fc1、Fc2)，前四个卷积层后都依次连接一个线性整流层和一个最大池化层，即卷积层、线性整流层、最大池化层、卷积层、线性整流层…这样类推，图5中省略了线性整流层和最大池化层，到第五个卷积层时，其后面依次连接一个线性整流层、一个全连接层、一个线性整流层和一个全连接层，最后一个全连接层的输出为神经网络输出；线性整流层具体为线性整流激活函数ReLU；

[0058] (5-2) 将训练集和测试集中微表情视频的融合特征图同时输入光流注意力神经网络进行训练，训练时采用的整体损失函数 L_{all} 为：

$$[0059] \quad L_{all} = L_s + \lambda MMD(D_s, D_t)$$

$$[0060] \quad L_s = \frac{1}{n} \sum_x -p(x) \log q(x)$$

$$[0061] \quad MMD(D_s, D_t) = \frac{1}{M^2} \sum_{i=1}^M \sum_{j=1}^M k(d_i^s, d_j^s) + \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N k(d_i^t, d_j^t) - \frac{2}{MN} \sum_{i=1}^M \sum_{j=1}^N k(d_i^s, d_j^t)$$

[0062] 式中， L_s 为训练集输出经过Softmax处理得到的微表情类别 $q(x)$ 和训练集中存储的微表情类别标签 $p(x)$ 的交叉熵， x 表示对应样本， n 表示Softmax处理后的输出个数， λ 表示权重系数， $MMD(D_s, D_t)$ 为 D_s, D_t 的最大均值差异， $D_s = [d_1^s, d_2^s, \dots, d_M^s]$ 为训练集输出，元素为 M 个， $D_t = [d_1^t, d_2^t, \dots, d_N^t]$ 为测试集输出，元素为 N 个， $k()$ 表示高斯核函数。

[0063] 本实例中卷积层的卷积核大小都为 3×3 ，步长为1，五个卷积层的通道数分别为64, 64, 128, 128, 128，最大池化层大小为 2×2 ，步长为2，第一个全连接层的通道数为1024，最后

一个通道数为3，即微表情类别；高斯核函数为 $k(x, y) = e^{\frac{-\|x-y\|^2}{\sigma}}$ ， $\sigma = [2.0, 5.0, 10.0, 20.0, 40.0, 80.0]$ ；权重系数 $\lambda = 2$ 。

[0064] (6) 将待识别的微表情视频按照步骤(2)~(4)进行处理，得到融合特征图，并将得到的融合特征图输入训练好的光流注意力神经网络，如图6所示，输出即为识别出的微表情类别。

[0065] 本实施例还提供了一种基于光流注意力神经网络的跨库微表情识别装置，包括存储器和处理器，所述存储器中存储有计算机程序，所述处理器用于执行所述计算机程序时实现上述方法。

[0066] 为了验证本发明的优越性能，本发明在CASME2微表情库、SAMM微表情库和SMIC_HS微表情库之间进行跨库微表情识别，实验结果如表1所示。本实验选用准确率(Accuracy)和F1_score作为实验的评价标准，从结果可以看出，本发明取得了最好的结果，以此证明了本发明的有效性。

[0067] 表1

[0068]

训练数据库	测试数据库	本算法Accuracy/F1_score
CASME2	SMIC_HS	62%/60%
CASME2	SAMM	48%/47%
SMIC_HS	CASME2	74%/71%
SMIC_HS	SAMM	45%/45%

SAMM	CASME2	53%/50%
SAMM	SMIC_HS	55%/54%

[0069] 以上所揭露的仅为本发明一种较佳实施例而已,不能以此来限定本发明之权利范围,因此依本发明权利要求所作的等同变化,仍属本发明所涵盖的范围。

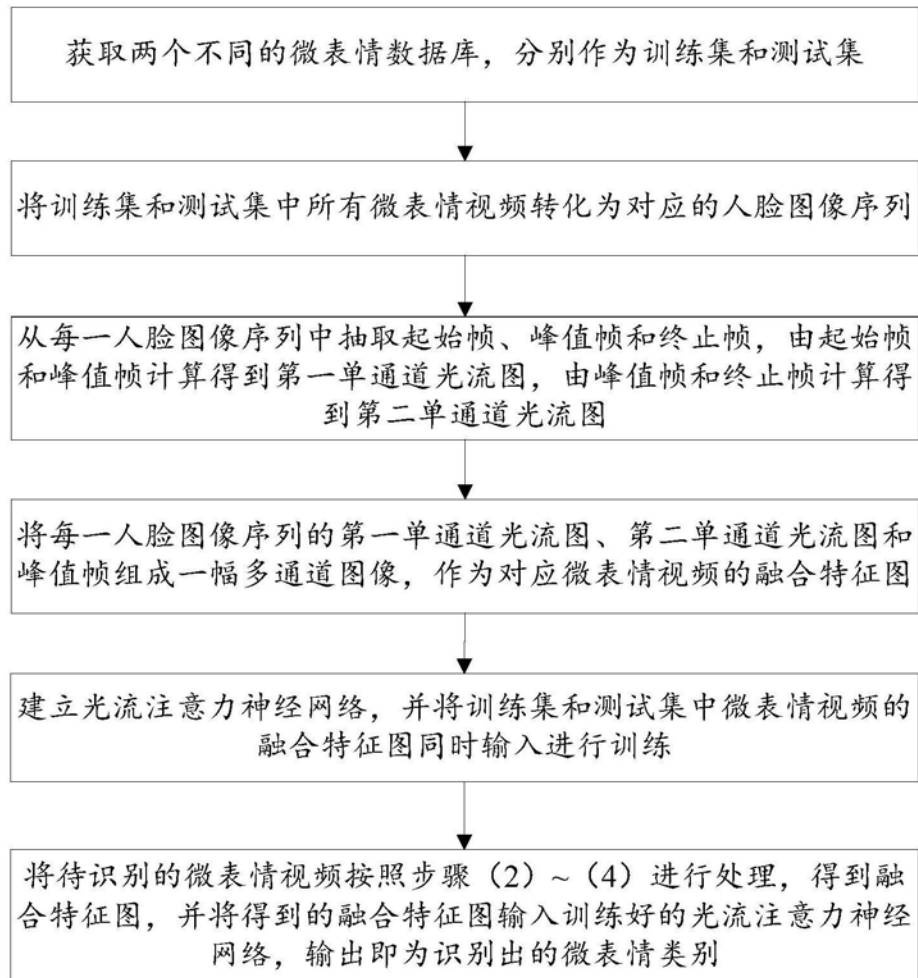
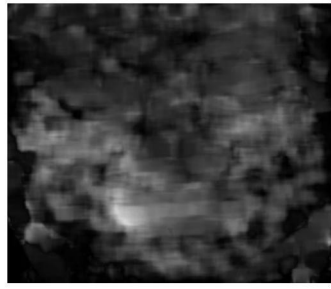


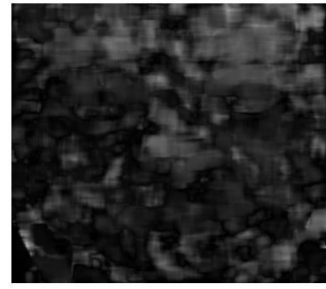
图1



图2



起始帧与峰值帧的光流图



峰值帧与终止帧的光流图

图3



图4

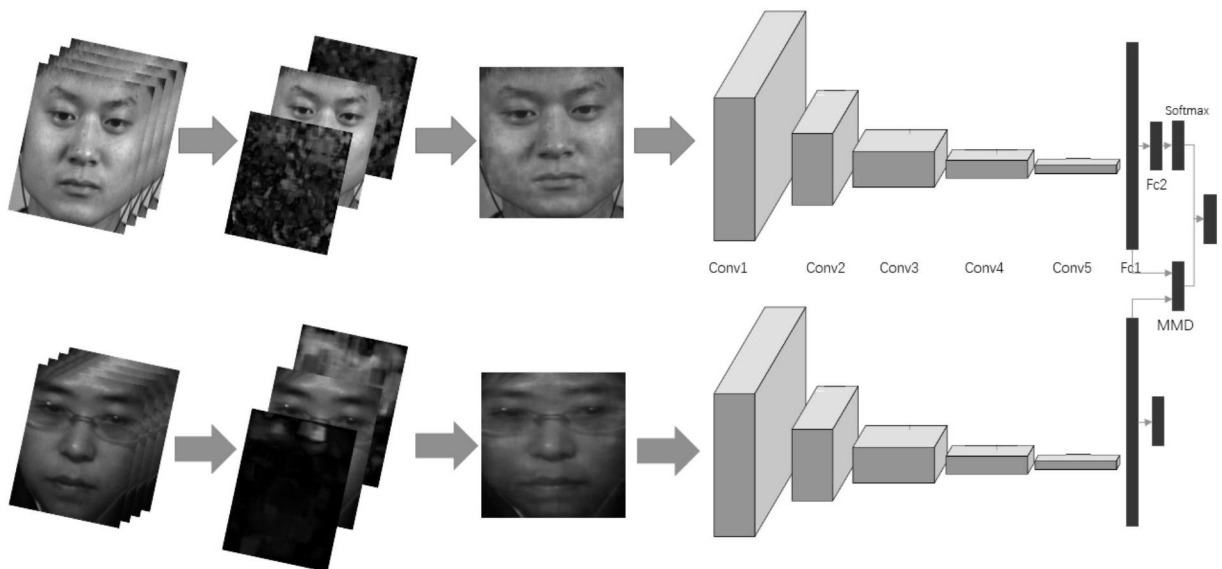


图5

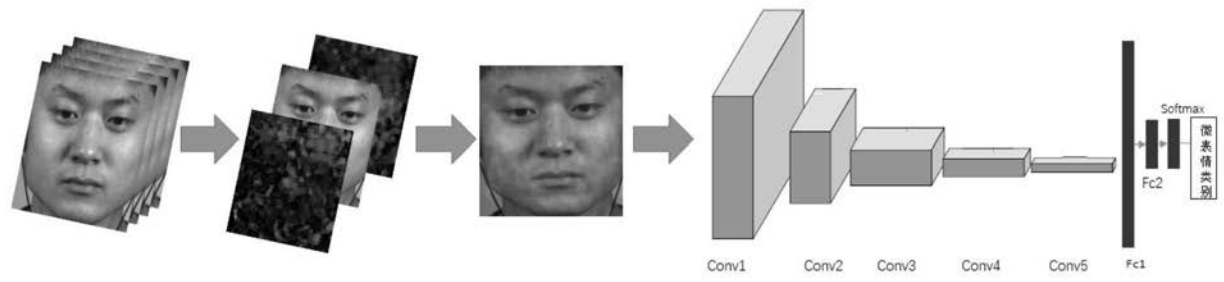


图6