

视频放大和深度学习在微表情识别任务上的应用

刘汝涵, 徐 丹*

(云南大学信息学院 昆明 650504)
(danxu@ynu.edu.cn)

摘 要: 针对微表情动作过于微弱不利于识别和目前主流方法合并情绪类别不利于微表情在现实任务中的应用 2 个问题, 提出一种基于眼部干扰消除的视频放大方法, 并利用卷积神经网络实现微表情识别任务. 首先, 利用基于相位的视频动作处理技术对微表情数据集 CASME 和 CASME II 中的视频数据进行放大; 然后利用特征点定位获取眼部坐标, 并将原始眼部视频替换到放大视频中进行图像融合, 以实现眼部干扰的消除操作; 最后利用 VGG16 的思想设计卷积神经网络模型网络, 实现对放大后的微表情数据情绪类别的识别. 实验在不同方法下分别对 2 个数据集的准确率进行对比, 并用几种调优策略下的模型分别就原始数据集和放大数据集的准确率进行对比. 结果表明, 文中方法能够更好地提升真实情绪分类状态下的微表情识别准确率.

关键词: 微表情; 视频放大; 卷积神经网络; 情绪识别

中图分类号: TP391.41 **DOI:** 10.3724/SP.J.1089.2019.17568

Video Amplification and Deep Learning in Micro-Expression Recognition

Liu Ruhan and Xu Dan*

(School of Information, Yunnan University, Kunming 650504)

Abstract: To address the difficulty in identification of micro-expressions and low practical significance of consolidated emotional categories, a video amplification method based on eye interference elimination is proposed, and convolutional neural network (CNN) is used to realize micro-expression recognition. First, we amplify the video data of CASME, CASME II using phase-based video motion processing technology. Then, to eliminate eye interference, feature point location is used to obtain eye coordinates, and replace the original eye video into an enlarged video with fusion processing. Finally, the idea of VGG16 is used to construct CNN model, and identify emotion categories in the enlarged micro-expression data. Experiments compared the accuracy of two datasets under different methods as well as the accuracy of original and enlarged datasets under several models with different tuning strategies. The results show that the method can effectively improve the recognition accuracy under real emotion categories.

Key words: micro-expression; video amplification; convolutional neural network; emotion recognition

人的脸部作为人际交往中最为关键也是最为直接的沟通渠道, 可以传输各种各样的信息. 当我

们观察一张人脸时, 除了看到最为基本的结构特征外, 更多的是观察到对方的面部表情, 并基于此

收稿日期: 2018-09-02; 修回日期: 2019-06-06. 基金项目: 国家自然科学基金(61163019, 61271361, 61540062); 云南省应用基础研究计划重点项目(2014FA021); 云南省教育厅科学研究基金产业化培育项目(2016CYH03). 刘汝涵(1990—), 女, 硕士研究生, 主要研究方向为计算机视觉、表情识别; 徐 丹(1968—), 女, 博士, 教授, 博士生导师, CCF 会员, 论文通讯作者, 主要研究方向为图形图像处理、计算机视觉、认知计算.

展开交流。在刑侦破案时,根据目击者的证言,人们能够发现,歹徒往往在行凶前就已经引起周围人的恐惧。通常人们把这些归因于本能或者直觉,但事实上,研究人员发现,人们之所以能够提前感知到对方的隐藏属性,正是因为观察到了对方的微表情信息。

人脸微表情最早在 1966 年由 Haggard 等^[1]发现并提出,它是一种极为短暂、极为微小的面部表情动作,是人们在试图隐藏自己的真实情绪时产生的,无法伪造。正因如此,相比于人们有意识做出的普通面部表情,微表情能够更加直接地体现人们真实的情绪感受。

相关研究^[2-3]表明,微表情有 3 个区别于普通面部表情的显著特点:

(1) 持续时间短。一个完整的微表情从出现到结束最短只需 1/25 s。

(2) 动作幅度小。若非经过专业训练,普通人员很难正确识别微表情的存在。

(3) 不会同时出现在上半脸和下半脸。由于是在人们抑制情绪时出现的自发式表情,微表情呈现出局部化的特点,不会同时出现在多处面部区域。

由于以上 3 个特点,很长时间以来,国内外针对微表情的研究都仅仅局限于心理学层面^[4],而利用计算机实现自动识别微表情的工作很难顺利地展开。直到近年来,随着计算机技术的成熟和发展,该项工作^[5-10]重新受到众多科研工作者的重视。

通过对近年来微表情检测和识别方法的研究能够发现,目前基于计算机自动识别微表情的方法主要分为以下几类^[7]:基于局部二值模式(local binary pattern, LBP)^[8]及其改进方法的识别、基于光流场特征^[9]的检测和识别,以及基于张量分析^[10]的识别等。各类识别与检测方法依旧局限于利用传统技术和方法对微表情进行任务实现,工作计算量大、耗时长且普适性差。更为重要的是,由于微表情持续时间短且动作幅度小的特点,传统技术和方法很难在识别精度上有进一步的提升空间,这已经成为了目前微表情识别技术发展的最大瓶颈。而神经网络,尤其是卷积神经网络(convolutional neural network, CNN)作为近年来在图像识别领域取得巨大成功的新兴技术,其在微表情识别方面的应用尚未见报道,因此本文注意到利用 CNN 实现微表情识别任务的可能性。

本文针对 CNN 在图像识别及分类任务上的巨大优势,提出了基于视频局部放大技术和深度学

习的人脸微表情识别方法,成功地实现了神经网络在微表情识别任务上的应用。

1 相关工作

本文以中国科学院微表情数据集 CASME^[11-12]和 CASME II^[13]为数据基础,通过基于相位的视频放大技术对微表情视频进行放大处理,以放大表情动作幅度;随后,通过裁剪等一系列预处理操作对视频图像序列进行处理,并将数据送入 CNN 中进行特征提取,最终实现微表情的识别。下面介绍现有 3 类微表情识别方法中代表性的方法。

1.1 现有微表情识别方法

基于 LBP-TOP 的识别方法是一种基于 LBP^[8]的改进方法,也是最早实现人脸微表情自动识别的方法,它为后续一系列的微表情识别技术提供了良好的对比依据。

基于面部动力谱(facial dynamics map, FDM)^[9]的方法是一种基于光流场特征识别的代表方法。

基于判别式张量子空间分析(deterministic tensor subspace analysis, DTSA)^[14]的方法通过判别式张量的子空间学习最优的特征表达,实现微表情的识别,是基于张量分析的识别方法。

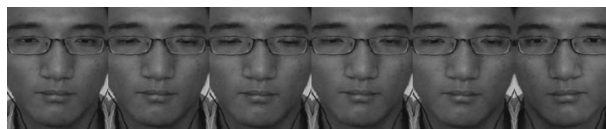
1.2 实验数据集

1.2.1 CASME

CASME 是由中国科学院提出的自发式微表情数据集。为了解决微表情持续时间短的问题,该数据集采用 60 帧/s 的高速摄像头对微表情进行视频捕捉^[11]。图 1 中展示了 CASME 中一些微表情的示例图。



a. 取自 CASME: EP05_2 中紧张的表情



b. 取自 CASME: EP07_4 中开心的表情

图 1 CASME 微表情数据集示例图

CASME 团队从超过 1500 个面部动作表情中选出了 195 个微表情视频,分别利用动作单元^[15]和情绪标签对每一个微表情视频序列进行相应的标注,标出了每一个视频中微表情的起始(onset)和结

束(offset)帧, 以及微表情达到最大幅度的帧(apex frame). 表 1 所示为数据集各表情的具体分类名称及个数.

表 1 CASME 数据集表情分类及个数统计

情绪标签	动作单元标签	表情个数
蔑视(contempt)	R12; L12	2
厌恶(disgust)	4; 9; 1+4+9; 4+7+9+10+25+26	46
恐惧(fear)	20; L20	2
开心(happiness)	12; 6+12+13; L10; L12	9
压抑(repression)	14; 15; 17; 24; 14+18; 15+17	40
悲伤(sadness)	1; 1+4; 1+4+15	6
惊讶(surprise)	1+2; 16; 34; 1+2+25; 1+2+4+34	21
紧张(tense)	4; 7; 9; 15; 16; 18; 14+16; 24	69

其中, 动作单元标签是基于面部动作编码系统(facial action coding system, FACS)^[15]给出的实现每个表情所对应的动作单元号, 每个数字代表一个特定的面部动作, 通过这些动作单元的单独或组合出现, 会在人脸上显示出各种各样的面部表情. 例如, 序号 4 代表皱眉(降低眉毛); 序号 9 代表皱眉鼻, 而序号 4 和 9 表情的组合往往代表愤怒、厌恶等负面情绪. 在序号前面加上“R”或“L”则代表特定表情只出现在右半边或左半边脸.

1.2.2 CASMEII

CASMEII^[13]是 CASME 数据集的改进版本, 为了改善现有数据集数据量少、不足以支持相关研究中训练和测试工作的现状, CASME 团队提供了一个更大规模的 CASMEII 数据集. 该数据集采用与 CASME 相同的标注方式, 从 26 位参与者的接近 3000 个面部动作表情中选出了 255 个微表情样本. 而与之之前的 CASME 数据集不同的是, 该数据集采用 200 帧/s 的超高速摄像头代替原有的 60 帧/s 摄像头, 以获取更为精准的微表情视频数据. 表 2 所示为数据集各表情的具体分类名称及个数.

表 2 CASMEII 数据集表情分类及个数统计

情绪标签	动作单元标签	表情个数
厌恶(disgust)	4; 4+10; 4+5+7+9; 4+L9+14	63
恐惧(fear)	20; L16	2
开心(happiness)	12; 14; L14+L15; 6+12; 6+12+7	32
压抑(repression)	15; 12+15; 14; 17; 12+14+17	27
悲伤(sadness)	1; L1; 1+4	7
惊讶(surprise)	1+L2; 1; L2+12; R1+R2; L2+26+38	25
其他(others)	18; 4; 14; 17; 15; 4+7; 4+38	99

1.3 基于相位的视频放大

在万千世界中, 很多运动会由于过于微弱而无法被肉眼所察觉; 例如, 人类脉搏的跳动, 手指按压木板造成的微弱形变, 收音机放出的音乐在

周边绿植上产生的细小振动. 为了能直观地看到这些运动, 人们往往需要借助视频放大技术对微弱运动进行精密的计算, 并根据速度矢量来扭曲视频帧, 以达到放大局部微小运动的目的.

基于相位的视频放大运动^[16]是一种基于复值可操作金字塔的运动分析方法. 随着时间的推移, 其复值可操作金字塔系数的相位变化与运动相对应, 因此可以通过对其进行时域滤波和放大处理, 以达到放大微弱运动的目的. 对比早前的欧拉线性视频放大(Eulerian video magnification, EVM)技术^[17], 基于相位的视频放大技术不涉及精密的光流计算, 通过计算局部的相位变化对运动进行测定并放大, 很好地降低了视频放大过程中同时放大噪声的可能, 拥有更好的鲁棒性, 同时也支持更大的动作放大倍数^[18]. 图 2 所示为基于相位的视频放大方法的处理流程.

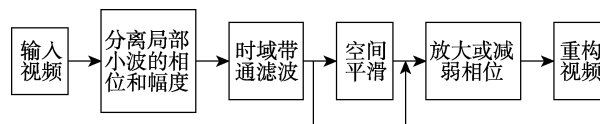


图 2 基于相位的视频放大方法流程图

如图 2 所示, 基于相位的视频放大方法^[16]首先在复值可操作金字塔的每一个不同的空间尺度和方向上计算其局部相位, 以分离出局部小波的相位和幅度. 其次, 对分离出的所有相位进行时域带通滤波, 以求分离出所需处理的特定时间频率. 然后, 为了实现视频放大操作, 其对进行带通后的相位加上特定的放大因子, 并将该操作应用于每一帧视频. 最后, 对放大后的视频序列进行重构以完成放大操作.

本文基于该方法, 提出一种眼部干扰消除的视频动作放大方法, 在放大微表情动作幅度的同时, 有效地消除了由于眨眼带来的眼部模糊现象, 以降低微表情识别的难度.

2 本文方法

为解决人脸微表情动作幅度小, 难以被察觉的问题, 本文引入了基于相位的视频运动处理技术^[16]对微表情视频进行放大操作, 以便提高识别准确率.

然而, 由于微表情的动作幅度过于微小^[2]且平均持续时间仅保持在不到 1s 的时间内, 其动作幅度和持续时间甚至均小于人在眨眼时的幅度, 而标准的基于相位的视频运动处理技术^[16]的局限性,

导致很难同时保证眨眼动作的清晰程度和人脸微表情动作的放大幅度。也就是说,当目标视频中脸部微表情被放大到可见程度时,眨眼动作会导致人眼部位清晰度下降,甚至完全模糊;而当人眼在眨眼过程中仍能保持清晰时,脸部的微表情则不会被放大或放大效果不明显。图3给出了视频放大前后眼部的差别对比图,其视频图像为视频数据中第5帧图像,是微表情发生前的基准脸(中性脸)图像,但由于产生了眨眼动作,被放大后的视频明显发生了眼部模糊现象。

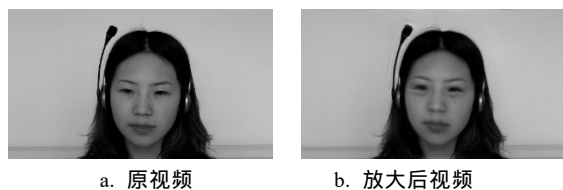


图3 视频放大前后眼部对比图

综合前文分析,为了解决视频放大技术在微表情数据集上的局限性,本文提出在放大处理过程中消除眼部干扰,并利用 CNN 模型对放大后的微表情数据集进行训练,最终实现对微表情的识别。

2.1 基于干扰消除的视频放大

为了消除眨眼对微表情视频放大结果的干扰,本文首先在视频中对人脸进行关键点定位,以找出眼部的坐标位置;然后在利用基于相位的视频放大技术对视频进行处理后,将未经过放大处理的原始视频中的眼睛部位进行裁剪并替换放大后的眼部视频,以消除眨眼的干扰。图4所示为本文中视频放大处理的流程图。

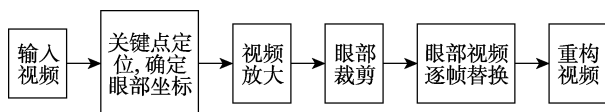


图4 本文视频放大处理流程

在消除眼部干扰后,本文对处理后的视频进行了脸部裁剪等预处理操作,以便后期使用。此外,经过分析可以发现,在2个数据集共447个视频数据中,仅有14个数据涉及眼部相关动作,且所有与眼部有关的微表情动作均非独立出现,即在眼部出现微表情动作的情况下均会伴随脸颊或眉毛等周边面部肌肉的运动。此种情况下,由于眨眼动作的幅度远远大于眼部微表情的动作幅度,且对所有微表情数据均产生影响,其对识别结果的负面

影响将远远超过放大该部位所产生的正面影响。因此,为了不影响放大及后期模型训练的效果,本文对进行干扰消除处理后的眼部视频不做特殊处理,使用包括消除干扰的眼部图像在内的整幅面部图像进行后期识别任务。

原视频直接放大及消除眼部干扰后的视频放大效果对比图如图5所示,数据出自CASME数据集中视频EP01_5的开心微表情。其中,第105帧为微表情起始阶段的其中一帧图像,由于眨眼动作的出现,视频经原始放大技术处理而未消除眼部干扰时有明显模糊,不利于后期的识别与检测;第120帧图像则显示了该微表情顶点(动作幅度最大点)的表情图像,放大了嘴部上扬幅度,同时消除了眼部模糊干扰。

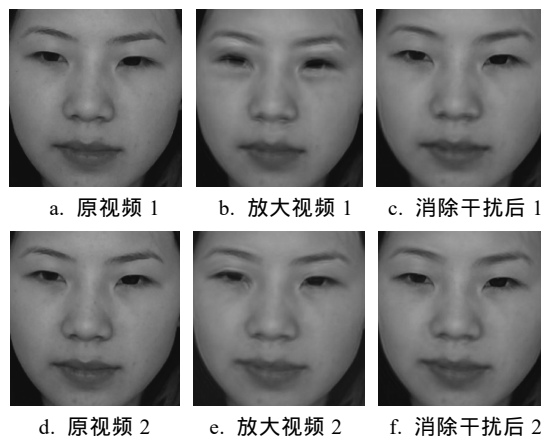


图5 视频放大效果对比图

2.2 CNN 模型设置

本文将CASME数据集中的视频图像序列按照情绪标签分为如表1所示8类,分别为蔑视(contempt)、厌恶(disgust)、恐惧(fear)、开心(happiness)、压抑(repression)、悲伤(sadness)、惊讶(surprise)和紧张(tense)。将CASMEII数据集中的视频图像序列按情绪标签分为如表2所示7类,分别为厌恶(disgust)、恐惧(fear)、开心(happiness)、压抑(repression)、悲伤(sadness)、惊讶(surprise)及其他(others);同时将所有视频图像序列统一裁剪为 110×110 大小的灰度图像,以备后期使用。

以CASME数据集为例,实验根据数据集所给标签,从视频图像序列中取出所有被标记为“微表情”的视频帧,并分别按所给标签标定情绪类别,以扩充样本量,共4001幅图像,取出其中的10%作为测试集,剩余90%的图像作为训练集。

为了完成针对微表情数据的识别和分类工作,

本文设计使用了如表 3 所示的 CNN 模型, 主要包括 5 个卷积层、最大池化层、丢包层, 以及 3 个全连接层; 其具体模型架构如表 3 所示。

表 3 CNN 模型架构

层名	输出大小	核大小
conv2d_1	110×110	1×1, 128
batch_normalization_1	110×110	
activation_1	110×110	
conv2d_2	110×110	1×1, 64
batch_normalization_2	110×110	
activation_2	110×110	
max_pooling2d_1	55×55	2×2, stride2
dropout_1	55×55	
conv2d_3	53×53	3×3, 64
batch_normalization_3	53×53	
activation_3	53×53	
max_pooling2d_2	26×26	2×2, stride2
dropout_2	26×26	
conv2d_4	26×26	1×1, 32
batch_normalization_4	26×26	
activation_4	26×26	
conv2d_5	24×24	3×3, 32
batch_normalization_5	24×24	
activation_5	24×24	
max_pooling2d_3	12×12	2×2, stride2
dropout_3	12×12	
dense_1		512
batch_normalization_6		512
activation_6		512
dropout_4		
dense_2		128
batch_normalization_7		128
activation_7		128
dropout_5		
dense_3		8
activation_8		8

其中, 在模型的参数设置方面, 本文将丢包层 1~5 的丢包率分别设置为 0.6, 0.6, 0.5, 0.4, 0.5; 损失函数方面, 实验中选用了多分类对数损失函数, 即 categorical_crossentropy. 该损失函数公式可表示为 $L(p^{\text{true}}, p^{\text{predict}}) = -\sum_i p_i^{\text{true}} \log p_i^{\text{predict}}$.

模型采用 5 层卷积层的架构, 为了达到降维的目的, 本文在每层卷积层后加入了 2×2 , 步长为 2 的最大池化层, 同时加入 batch normalization 以加快收敛速度、提升模型性能; 加入 dropout 丢包层,

以防止模型出现过拟合, 提升模型泛化能力。

另外, 为了进一步提升模型泛化能力, 降低过拟合出现的可能, 本文在后续 CASMEII 数据集的处理和训练过程中还加入了 L_2 正则化, 同时调整了 dropout 层的丢包率参数. 实验表明, 此类参数调整对实验结果的优化有较大影响。

3 实验结果和分析

本文采用对数据集进行整体准确率判定的方式来评价模型的最终性能, 在准确率判定的过程中, 所有对微表情进行正确识别的图像序列都将被计算其中. 具体计算公式为

$$\text{总体准确率} = \frac{\text{所有被正确识别的微表情图像序列总数}}{\text{微表情图像序列总数}} \quad (1)$$

本文方法在 CASME 和 CSAME II 数据集进行了整体准确率计算, 并分别给出了与第 1.1 节中提到的 3 类微表情识别技术中代表方法的对比结果, 如表 4 所示。

表 4 几种微表情识别方法的准确率对比

识别方法	CASME		CASME II	
	准确率/%	情绪标签	准确率/%	情绪标签
FDM	42.02	4	41.96	5
LBP-TOP	36.05	4	39.22	5
DTSA	32.45	5	N/A	N/A
本文	41.80	8	46.35	7

由于 CASME 和 CASMEII 数据集中各表情样本数量不均衡的缘故, 目前大多数微表情识别方法采用将样本数量较少的表情进行合并, 或不计入测试范围, 或者直接将表情数据划分为积极、消极、中性及其他等类别的方式进行测试. 以文献[9]为例, 在 CASME 的实验过程中, 该文献将微表情划分为厌恶、压抑、惊讶以及紧张 4 类以进行测试; 而在 CASMEII 中则分为厌恶、开心、压抑、惊讶以及其他等 5 类. 但这样的做法往往容易忽略很多潜在的面部微表情, 不利于提高微表情识别技术的真实应用效果。

为了提高微表情识别技术的实际应用效果, 本文在对完整数据集进行消除干扰的视频放大操作, 加大表情动作幅度后, 利用 CNN 对放大后的数据进行训练, 以达到更好的识别效果. 本文在测试过程中分别将数据集 CASME 和 CASMEII 按照

情绪标签分为 8 类和 7 类, 实现了对于微表情各个情绪类别的全分类, 且获得了较好的识别率, 如表 4 所示.

此外, 为了验证 CNN 模型参数设置对训练结果的影响, 本文以 CASMEII 数据集为例, 针对是否加入 L_1 , L_2 正则化, 调整 dropout 层丢包率等参数调优策略下的训练结果的优劣进行了对比, 在该次实验中, 本文将 dropout 层初始参数统一设置为 0.3, 以方便对比, 对比结果如表 5 所示. 其中, 策略 1 为不添加 L_2 正则化并使用初始 dropout 层参数效果; 策略 2 为添加 L_2 正则化并使用初始 dropout 层参数效果; 策略 3 为添加 L_2 正则化并调优各 dropout 层参数后效果; 策略 4 为添加 L_2 正则化并调优各 dropout 层参数, 同时修改初始化函数后效果.

表 5 几种参数调优策略下的准确率对比 %

策略	准确率
1	22.51
2	29.57
3	46.35
4	37.00

通过实验结果可以看出, 在初始 dropout 参数下, 训练后结果产生严重过拟合, 而通过增加 L_2 正则化项可以较好地解决该问题; 在各 dropout 层参数调优过程中, 各层参数间形成一个较大参数差的情况下, 更容易找到最优参数.

最后, 为了验证经过视频放大处理的数据集对于人脸微表情识别效果的积极作用, 本文以 CASME 数据集为例, 分别利用原始数据集和放大数据集(mag_CASME)在同一网络模型下进行训练, 并对比二者的微表情识别准确率, 2 个数据集在各模型下的准确率结果如表 6 所示. 其中, 模型 1 为未进行参数调优策略时的效果; 模型 2 为完成参数调优步骤, 同时使用默认优化器后的效果; 模型 3 为进行参数调优, 同时修改优化器至最佳准确率后的效果.

表 6 2 个数据集在各模型下的准确率 %

数据集	模型		
	1	2	3
CASME	22.22	29.29	36.81
mag_CASME	31.82	35.15	41.80

在该对比实验中, 本文以 mag_CASME 数据集下的训练结果高低为依据, 依次选取了由低到

高 3 组不同结果的参数模型, 再以原始数据集代替经过视频放大处理的数据集进行重新训练, 并对二者最终的准确率进行对比.

由表 6 对比结果可知, 在相同的训练模型下, 经过视频放大处理的数据集(mag_CASME)准确率普遍高于原始数据集(CASME), 能够进一步提升人脸微表情识别的准确率.

4 结 语

本文从解决微表情识别三大难题的角度出发, 针对持续时间短的问题, 选择使用利用高速摄像头采集完成的 CASME^[11-12]和 CASMEII^[13]微表情数据集作为基本处理数据; 针对动作幅度小的问题, 利用基于相位的视频放大算法^[16]对原始数据进行处理, 同时进行眼部干扰消除操作, 实现适用于人脸微表情的视频放大技术, 放大微表情动作幅度, 以达到易于识别的目的. 更为重要的是, 本文利用 CNN 实现对于微表情的识别工作, 成功地实现了神经网络在自发式表情数据识别工作上的尝试.

未来的研究工作将主要从以下 2 个方面进行:

(1) 与其他图像分类数据集相比, 专业的微表情数据集少之又少, 且每个数据集的数据数量有限, 很难真正发挥神经网络模型的巨大作用. 因此, 要想解决微表情识别率不高的问题, 首先要实现对数据集的扩充.

(2) 实验表明, 一个网络模型及其参数设定的好坏很大程度上将直接决定实验的结果. 然而, 目前各种神经网络模型层出不穷, 且更新速度极快, 因此, 在众多模型中摸索并改进出一个适合于人脸微表情识别的模型将会大大提升实验的准确率.

参考文献(References):

- [1] Haggard E A, Isaacs K S. Micromomentary facial expressions as indicators of ego mechanisms in psychotherapy[M] //Methods of Research in Psychotherapy. Heidelberg: Springer, 1966: 154-165
- [2] Porter S, ten Brinke L. Reading between the lies: identifying concealed and falsified emotions in universal facial expressions[J]. Psychological Science, 2008, 19(5): 508-514
- [3] Yan W J, Wu Q, Liang J, et al. How fast are the leaked facial expressions: the duration of micro-expressions[J]. Journal of Nonverbal Behavior, 2013, 37(4): 217-230
- [4] Wu Qi, Shen Xunbing, Fu Xiaolan. Micro-expression and its

- applications[J]. *Advances in Psychological Science*, 2010, 18(9): 1359-1368(in Chinese)
(吴奇, 申寻兵, 傅小兰. 微表情研究及其应用[J]. *心理学进展*, 2010, 18(9): 1359-1368)
- [5] Li J T, Soladie C, Segurier R. LTP-ML: micro-expression detection by recognition of local temporal pattern of facial movements[C] //Proceedings of the 13th IEEE International Conference on Automatic Face & Gesture Recognition. Los Alamitos: IEEE Computer Society Press, 2018: 634-641
- [6] Barros J M D, Mirbach B, Garcia F, *et al.* Fusion of keypoint tracking and facial landmark detection for real-time head pose estimation[C] //Proceedings of the IEEE Winter Conference on Applications of Computer Vision. Los Alamitos: IEEE Computer Society Press, 2018, 1: 2028-2037
- [7] Xu Feng, Zhang Junping. Facial microexpression recognition: a survey[J]. *Acta Automatica Sinica*, 2017, 43(3): 333-348(in Chinese)
(徐峰, 张军平. 人脸微表情识别综述[J]. *自动化学报*. 2017, 43(3): 333-348)
- [8] Pfister T, Li X B, Zhao G Y, *et al.* Recognising spontaneous facial micro-expressions[C] //Proceedings of the IEEE International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2011: 1449-1456
- [9] Xu F, Zhang J P, Wang J Z. Microexpression identification and categorization using a facial dynamics map[J]. *IEEE Transactions on Affective Computing*, 2017, 8(2): 254-267
- [10] Wang S J, Yan W J, Sun T K, *et al.* Sparse tensor canonical correlation analysis for micro-expression recognition[J]. *Neurocomputing*, 2016, 214: 218-232
- [11] Yan W J, Wu Q, Liu Y J, *et al.* CASME database: a dataset of spontaneous micro-expressions collected from neutralized faces[C] //Proceedings of the 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition. Los Alamitos: IEEE Computer Society Press, 2013, 1: 1-7
- [12] Yan W J, Wang S J, Liu Y J, *et al.* For micro-expression recognition: database and suggestions[J]. *Neurocomputing*, 2014, 136: 82-87
- [13] Yan W J, Li X B, Wang S J, *et al.* CASME II: an improved spontaneous micro-expression database and the baseline evaluation[J]. *PLoS One*, 2014, 9(1): e86041
- [14] Wang S J, Chen H L, Yan W J, *et al.* Face recognition and micro-expression recognition based on discriminant tensor subspace analysis plus extreme learning machine[J]. *Neural Processing Letters*, 2014, 39(1): 25-43
- [15] Ekman P, Friesen W V. Facial action coding system (FACS): a technique for the measurement of facial action[M]. Palo Alto: Consulting Psychologists Press, 1978
- [16] Wadhwa N, Rubinstein M, Durand F, *et al.* Phase-based video motion processing[J]. *ACM Transactions on Graphics*, 2013, 32(4): Article No.80
- [17] Wu H Y, Rubinstein M, Shih E, *et al.* Eulerian video magnification for revealing subtle changes in the world[J]. *ACM Transactions on Graphics*, 2012, 31(4): Article No.65
- [18] Wang Xia, Fu Xiaojing, Wang Nan, *et al.* Pre-decision pyramid layers of Eulerian video magnification algorithm for subtle movements[J]. *Science Technology and Engineering*, 2016, 16(1): 114-118(in Chinese)
(王霞, 付晓静, 王楠, 等. 预决策金字塔层数的欧拉视频微弱运动放大算法[J]. *科学技术与工程*, 2016, 16(1): 114-118)