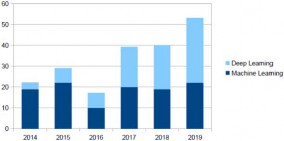
The current challenges of automatic recognition of facial expressions：A systematic review

1. **论文摘要**
2. 尽管人脸表情识别已经取得了重大进展，但该领域仍面临许多理论争论和技术挑战。
3. 通过对13个文献数据库的搜索，在2014-2019年期间共发现了220个参考文献。在全面介绍了当前系统及其性能之后，根据在自动面部表情识别领域遇到的主要问题对所选文章进行了分组和分析。
4. 强调了该领域未来研究的优势、局限性和主要方向。
5. **知识点**
6. **情感计算的应用领域**
   1. 大众情绪分析（人群的总体情绪、娱乐或旅游结构产生的幸福水平等）；
   2. 安全（体育场内或公共交通工具上的攻击风险、疲劳驾驶）
   3. 营销和娱乐（对产品、广告或电影的情感反应）；
   4. 健康（疼痛检测、对沟通障碍患者的支持，以及医疗和心理病理诊断方面的协助）。
7. **自动面部表情识别（AFER）**

事实上，AFER的大多数系统都是基于情绪的线性模型

1. **该论文从其它论文中提取出来的数据**
2. 完整参考资料和出版年份
3. 研究问题或主要主题
4. 使用的技术（分类器、培训和验证、数据集）
5. 标签的数量和类型（基本、复合情绪、维度、行动单位）
6. 绩效指标（准确性、F1分数）
7. 研究的优势和局限性
8. 将数据组织在表格中，以显示基本数据每项研究的信息。方差分析（ANOVA）也用于比较性能根据以下内容对系统进行分类：
9. 算法的一般类型（机器学习）/（深度学习）
10. 标签类型（基本/复杂/尺寸/情绪）
11. 使用的数据类型（姿势/自发表达）
12. **系统使用的算法**

****

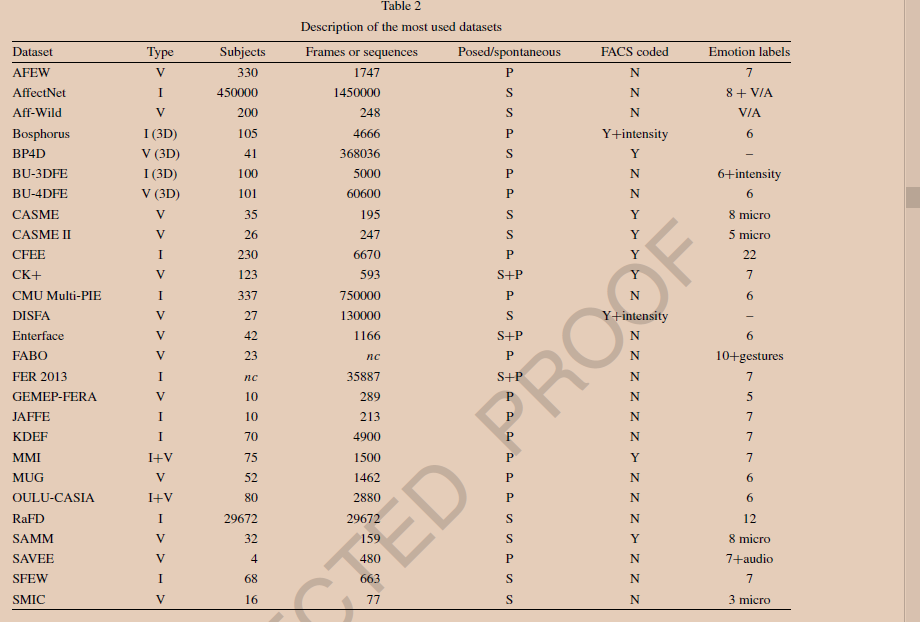
1. **机器学习**
2. 机器学习技术主要基于统计数据，以便对函数进行建模并得出预测，从而对输入数据进行分类。
3. 有各种各样的机器学习算法（回归、k近邻、AdaBoost、朴素贝叶斯、决策树、随机森林、SVM等）
4. 它们依赖于不同的统计方法，但按照相同的主要步骤操作：提取由数据分析人员手工选择的特征，进行模型估计，然后进行分类。
5. **特征提取**

使用机器学习算法的研究报告了各种各样的特征提取技术，其中Gabor滤波器和局部模式（LBP，LDP）应用最为广泛。

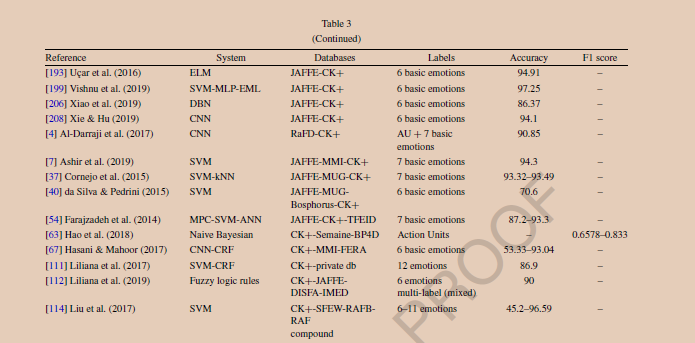
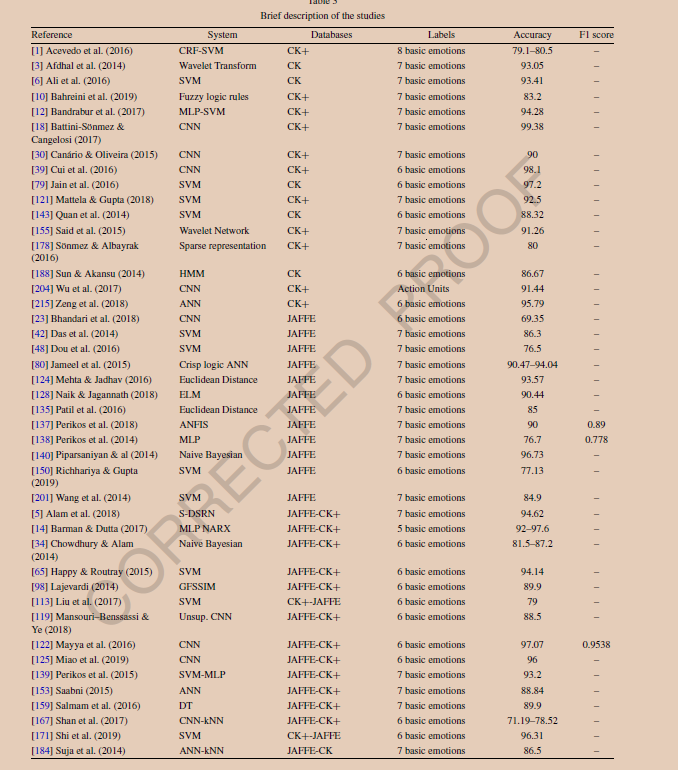
1. **深度学习**
2. 优缺点：不同于传统机器学习算法，在学习之前几乎不需要人工干预，但是需要大量的数据集
3. 主要用到的两种深度学习算法
4. CNN
5. RNN
6. **目前存在的技术挑战**
7. 大多数系统实时运行并处理大多数问题，如遮挡、不同的头部方向或亮度差异。然而，目前仍然存在一些技术挑战：面部表情的时间动力学问题。时间维度是面部表情的一个重要元素，情绪及其表达会随着时间的推移而波动。
8. 基于卷积神经网络的方法无法解释特征表达的时间动力学，一些学者使用了其它网络例如递归神经网络（RNN）和长短期记忆（LSTM）网络（一种特殊类型的RNN，能够学习和连接多个输入之间的时间依赖关系）。
9. 即使它带来了一些技术问题（特别是连续序列的分割问题），时间动力学也可以提供丰富的信息，并可以极大地提高面部表情识别系统的性能。
10. **静态识别与动态识别**
11. 静态识别与动态识别相比，平均准确率从74.5%增加到了80.5%.
12. 动态表情识别允许从变化的角度观察面部表情，能够促进微表情的检测和分类
13. 通过增大面部运动的幅度能够改善微表情的检测，表情强度的时间演变可以区分摆姿势的微笑和自然的微笑
14. **微表情难以检测的原因**
15. 微表情振幅低，需要非常精确的外观描述
16. 持续时间短（持续时间为1/25到1/15s）
17. 它们只能出现在特定的情况下
18. 不同的性格可能会有不同的微表情表现
19. **微表情的检测**
20. 一些研究[202,205]使用了运动放大技术，通过增加面部运动的幅度来改进微表情的检测。以下是其常用的定位技术
21. 阈值技术
22. 峰值技术
23. 还有一些研究[26]使用了高速采集系统以及基于当前帧和前一帧之间的绝对图像差的运动描述符
24. [133]提到了短时间的非情绪运动也可能被检测为微表情峰值，因此为解决这个问题，[134]开发了一种阈值技术来消除头部移动、眨眼、凝视的变化
25. **面部表情强度**
26. 面部表情的强度有助于解释复杂或者模糊的行为，[109]和[200]开发了用于估算行动单元强度（AU）的系统，其中考虑了AU与其强度之间的结构相关性和时间动力学。
27. [152]使用一个条件随机场来估计强度，该随机场考虑了他们称之为上下文的几个因素：观察到的受试者（谁）、面部表情的变化（如何）以及不同AU强度水平之间的时间关系（何时）。
28. 一些研究探索了在离散水平上提供强度估计的技术，主要是通过在多个数据库（Bosphorus、DISFA、GFT、UNBC）中提供5级强度的行动单元。
29. [11] 使用局部-全局排序技术，比较同一个人的两张图像之间动作单位强度的差异，以避免个体之间的形态差异问题。
30. [17] 接受多任务CNN培训，同时检测所有动作单元的发生和强度，同时考虑不同的头部方向。
31. [74]建议使用分离为多个ROI的特征提取系统检测AU的存在和强度，以减少AU共现问题，并避免因最终输出中要反映的人脸某些部分的遮挡而产生的错误。
32. **情绪理论模型**
33. 目前存在很多人提出了不同的理论模型，而在面部表情识别领域，大多数研究都假设情绪于面部表情之间存在直接联系，以下是各种研究中考虑的情绪类型：

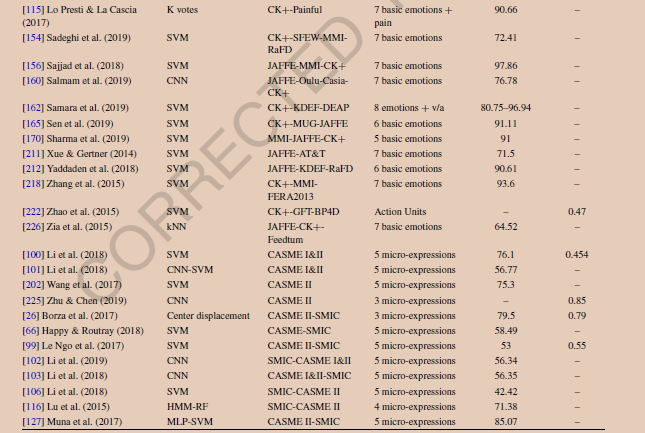


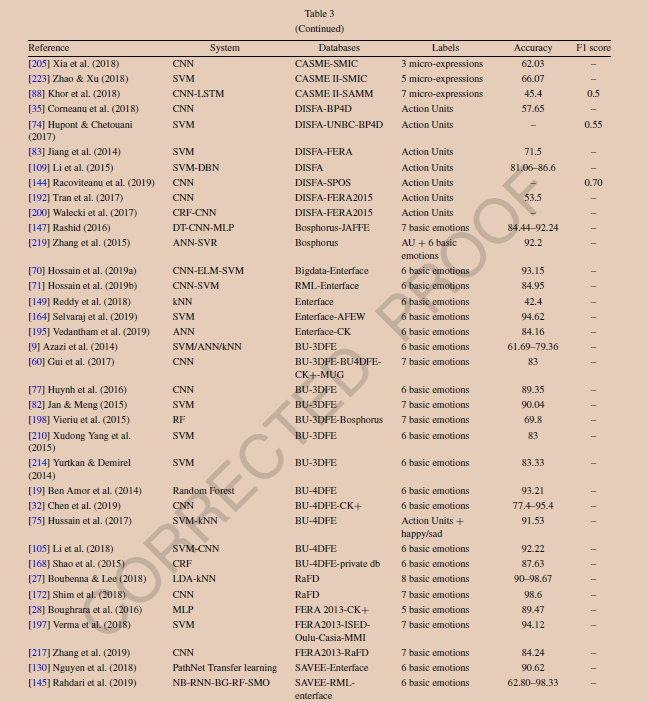
1. 目前基本表情一共分为7种，绝大多数研究（94.2%）都集中关注基本的离散情绪， 而复合情绪来自于两种或两种以上的基本情绪的组合，但是对于复合情绪的研究很少，主要是因为对于这种情绪的特定面部表情的存在缺乏理论共识。
2. [180]使用Plutchik的理论模型开发了一个系统，该系统经过训练可检测动作单元（AU），然后实现了构成6种基本情绪表达的AU组合，从而允许对应于次要情绪的组合（尤其是敬畏，恐惧和惊讶的组合）。
3. [111,112] 通过应用称为“模糊推理机”的系统，重新调查了以基本情绪（尤其是JAFFE、CK+、DISFA和IMED）分类的数据集，给出了多标签分类，以表示复合情绪。
4. 每一种情绪可以用二维或三维空间上的一个点来表示，即效价/唤醒度。[85,126,207]使用回归和效价/唤醒标记的数据库（RECOLA, Amigos individual and groupDB, ffectNet）来表示二维空间中的三维面部表情。
5. **算法试图识别什么：理论模型的问题**
6. 一些研究[33,43,57,118]只关注于微笑的检测，而不是对情绪进行分类
7. 自2014年以来，只有8项研究试图探索面部表情出现的背景。
8. 对于自动识别面部表情系统而言，它需要关注文化因素或者语境因素，因为同一表情在不同语境下可能表达不同的含义。[22]使用了一种基于事件的方法，他们开发了后处理分类器来识别一个时间窗口内的情绪背景。 这些算法使他们能够通过事件检测来预测情绪表达的突然或渐进变化。
9. [158]开发了一个基于背景信息（社会同伴的存在和类型、活动、温度、生理状态、位置和一天中的时间））的**情绪识别系统**。研究者表明，上下文信息是相关的，尤其是在个人层面，并提高了检测性能。
10. **自动面部表情识别系统的精度**
11. 深度学习的系统在平均识别精度上并不优于传统方法
12. **识别精度：基本情绪模型>效价/唤醒度模型>复合情绪模型**，情绪的类型对准确性有显着影响，例如在通过摆姿势获得的数据集上得到的准确度比自然采集的数据集准确度高。

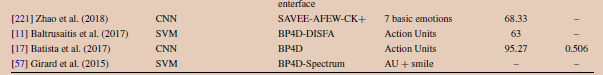


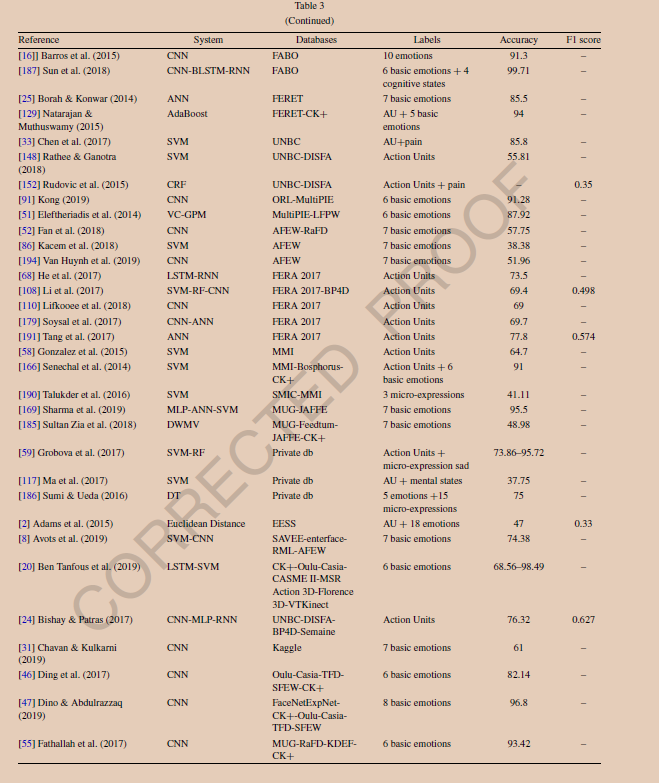
1. 平均准确率并不是评估算法的充分指标：根据情绪类别了解系统的性能也很重要。
2. 下面几张图片是不同研究者使用不同算法、数据集得到的识别精度，包括该数据集的类型

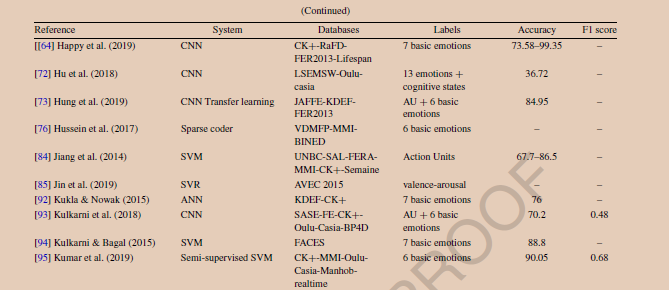


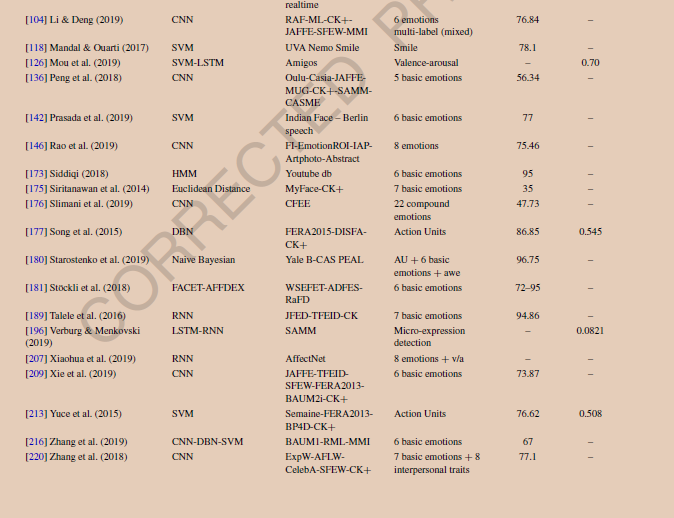












1. **总结**
2. 虽然深度学习被广泛使用，特别是在最近的研究中，但传统的机器学习技术仍然占大多数，并且涉及广泛的算法（SVM、随机森林） 、朴素贝叶斯网络、决策树等）。
3. 当系统面临更自然的数据时，性能会明显下降，即在产生情绪的现实生活条件（日常情况、自然对话）中获得的数据。这是因为情绪是一种多方面的现象，它不仅限于少数典型的面部表情，而且仍然受到许多理论争议的影响，因此在其复杂性方面仍然难以建模。
4. 利用多模态（语音、生物信号等）来进行情感识别将更加全面
5. 大量收集自然条件下的面部数据集有必要的，但是标注数据集非常困难，因此非注释的数据研究具有重要意义。
6. **备注**

这篇论文主要讲的是连续的动态表情识别、情绪理论、以及在面部表情识别系统用于实际应用（不同领域）时，它所需要考虑到的语境因素。这篇文章总结了220篇研究所使用的方法、数据集、数据集类型以及获得的精度，考虑了基本情绪模型、复合情绪模型、微表情、动作单元、多种情绪、多标签情绪以及上述几种的结合。

整理人：唐舒琪

整理时间：2021年11月17日