

疼痛自动识别综述^①

支瑞聪, 周才霞

(北京科技大学 计算机与通信工程学院, 北京 100083)

通讯作者: 周才霞, E-mail: g20188844@xs.ustb.edu.cn



摘要: 疼痛是一种个人经历, 一般被分为急性疼痛和慢性疼痛, 可来源于受伤、疾病、手术或其他的健康问题。如果疼痛没有被及时治疗, 会对患者的身心健康带来极大的伤害。由于患者自身障碍等问题, 患者可能无法自我报告疼痛, 而专业人士评估无法保证连续性和客观性。因此, 对疼痛自动识别系统的需求日益加大, 近十年来, 很多研究人员在此领域取得突破性的成果, 本文对疼痛自动识别系统进行综述, 一方面从疼痛自动识别系统的结构组成方面进行描述, 主要包括数据获取、数据预处理、特征提取以及分类等; 另一方面从疼痛模态表征多角度进行技术汇总, 主要包括行为、语音、生理以及多模态融合等 4 个方面。本文论述疼痛自动识别系统中的关键技术, 并加以对比分析总结出该领域发展的一些挑战和方向。

关键词: 疼痛; 分类器; 数据库; 特征提取; 特征选择

引用格式: 支瑞聪, 周才霞. 疼痛自动识别综述. 计算机系统应用, 2020, 29(2): 9-27. <http://www.c-s-a.org.cn/1003-3254/7262.html>

Review of Automatic Pain Recognition

ZHI Rui-Cong, ZHOU Cai-Xia

(School of Computer and Communication Engineering, University of Science & Technology Beijing, Beijing 100083, China)

Abstract: Pain is a personal experience that is generally divided into acute pain and chronic pain, which can result from injury, illness, surgery, or other health problems. If pain is not treated in time, it will cause great harm to the patient's physical and psychological health. Not all patients, e.g. dementia, are able to self-report pain. The continuity and objectivity of the assessment cannot be guaranteed when medical care personnel assessing patient's pain. Therefore, the demand for automatic pain recognition system is increasing. In the past decade, many researchers have made breakthroughs in this area. This paper reviews the automatic pain recognition system. On the one hand, it describes the structural composition of the automatic pain recognition system, including data acquisition, data preprocessing, feature extraction, and classification. On the other hand, it summarizes a great number of techniques from pain modal representation, i.e. behavior, speech, physiology, and multi-modal fusion. This paper also discusses the key technologies in automatic pain recognition system, and analyzes several challenges and directions in the field.

Key words: pain; classifier; database; feature extraction; feature selection

国际疼痛研究协会将“疼痛”定义为“与实际或潜在的
组织损伤相关的令人不快的感觉和情绪体验, 或

对此类损伤的描述”。疼痛在医学上被当作一个重要指
标^[1], 正确地判断疼痛能够让治疗更加有效, 但目前没

① 基金项目: 国家自然科学基金面上项目 (61673052); 国家重点研发计划 (2017YFD0400100); 中央高校基本科研业务费专项资金 (2302018FRF-TP-18-014A2); 国家留学基金委 (201606465015)

Foundation item: General Program of National Natural Science Foundation of China (61673052); National Key Research and Development Program of China (2017YFD0400100); the Fundamental Research Funds for the Central Universities of China (2302018FRF-TP-18-014A2); Fund of China Scholarship Council (201606465015)

收稿时间: 2019-07-06; 修改时间: 2019-07-23; 采用时间: 2019-08-07; csa 在线出版时间: 2020-01-16

有直接高效检测疼痛的简单方法. 疼痛检测的黄金准则是病人的“自我报告”^[2-4], 然而这不是对所有病人都适用的一种方法, 例如对于婴儿、具有语言障碍的人^[5]来说, 这种方法就无法实施. 对于此类患者, 目前由其代理人, 如专业的医疗人士、婴儿的父母等, 来评估疼痛并进行治疗^[6]. 这种方法带来的另一个问题是评估的不连续性和不一致性, 而且需要代理人具有非常专业的医疗知识, 这对劳动力的需求非常高^[7]. 为了克服上述问题, 很多研究人员都开展了疼痛自动识别方法的探索中, 特别是 UNBC-McMaster (McMaster University and University of Northern British Columbia) 肩痛数据库^[8]和 BioVid 热痛数据库^[9]的公布, 推动了这一领域的发展.

疼痛发生时会在脸上显露出来以表达经历痛苦的感觉, 脸上显露的疼痛表情还可以提供有关疼痛强度的信息. 尽管自我报告的疼痛和面部疼痛表达之间的相关性存在争议^[10], 但许多文献发现这两者之间存在显著的关系^[11,12]. 然而从面部图像或视频自动评估疼痛水平是相当具有挑战性的. 这不仅是因为在疼痛与非疼痛面部框架之间缺乏足够的视觉差异来发现与疼痛相关的特征, 也因为存在外部因素导致疼痛表达的复杂性, 如“痛着微笑”现象、性别差异 (男性与女性的体验方式) 等对疼痛的影响^[13-15]. 因此除了利用面部图像的空间^[16]和时间信息^[17]外, 面部质量^[18-23]和面部捕获

模态^[24]、头部姿势^[25]、身体运动^[26-28]、语音^[29-31]和生理指标^[27,32]等许多其他因素也在疼痛自动评估中起了重要作用^[33].

本文内容安排如下: 第1节介绍了疼痛特征的提取方法, 并据此将现有的代表性研究分为基于行为的、基于语音的、基于生理的、基于多模态的, 这是本文的主体章节; 第2节中介绍了与疼痛研究相关常见的数据库, 以及该数据库的获取方式、规模; 第3节介绍了疼痛识别任务中使用到的常见分类器; 第4节说明了疼痛自动识别领域发展的挑战及方向; 最后在第5节中进行总结.

1 特征提取

一个疼痛自动识别模型的设计主要分为以下过程 (图1): 首先对选取的原始数据进行裁剪、旋转等预处理以减少干扰信息, 随后根据实验方案对预处理后数据提取特征, 然后将提取出的特征输入到预先选择或设计好的分类器中进行训练学习, 对分类器的输出结果进行各项指数计算来判断这个疼痛自动识别模型的好坏. 本文主要根据所提取的特征将近几年该领域的研究成果进行分类汇总. 我们通过查看医疗专业人员对成人、婴儿进行疼痛评估的量表 (如 NIPS、COMFORT 量表等^[34,35]), 将特征分成基于行为的、基于语音的、基于生理的以及多模态的来进行介绍 (图2).

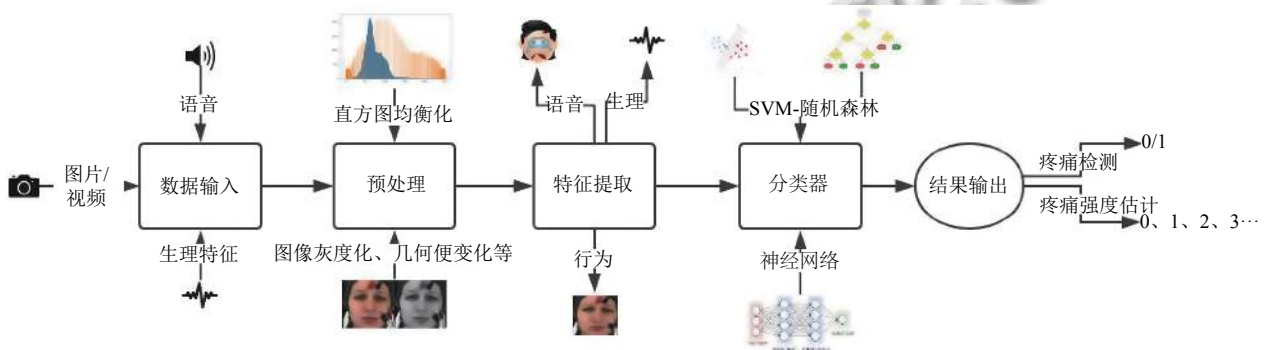


图1 疼痛自动识别模型的设计

1.1 基于行为的特征

与疼痛相关的评判标准中, 行为的变化是肉眼可见的明显变化之一, 且记录过程简单. 因此, 有大量文献基于这方面进行探讨. 我们将行为方面的特征分为基于面部表情的和基于身体运动的来进行介绍.

1.1.1 基于面部表情的特征

面部表情是疼痛的最常见的指标之一^[36], 所有目前公开发布的疼痛相关数据库面部表情都是必不可少的一部分. 疼痛的面部表情被定义为与疼痛刺激相关的面部肌肉的运动和扭曲, 面部运动可以由活动单元

(AU) 描述^[37]. 与成人疼痛相关的 AU^[38]包括眉毛下垂 (AU4), 脸颊抬高 (AU6), 眼睑收紧 (AU7), 鼻子皱纹 (AU9), 上唇抬高 (AU10), 斜唇抬起 (AU12), 水平唇拉伸 (AU20), 嘴唇分离 (AU25), 下颌下垂 (AU26), 嘴拉伸 (AU27) 和闭眼 (AU43). 与婴儿疼痛相关的 AU^[39]包括鼻唇沟的加深, 眉毛下垂, 眼睛变窄, 垂直和水平的嘴唇伸展, 唇部噘起, 唇部开口, 舌头突出, 绷紧的舌头和下巴颤动.

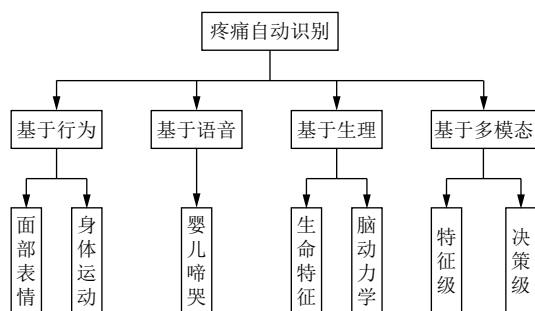


图2 疼痛自动识别方法框架

根据特征提取的对象差异, 将基于面部表情识别的方法分为基于静态图像的、基于序列的和基于面部动作编码系统 (Facial Action Coding System, FACS) 的. 基于静态图像的特征提取可能会丢失时间上的一些与疼痛相关的信息, 造成识别率低下; 基于序列的特征同时获得空间信息和时间信息, 但因此增加了运算复杂度; 基于静态图像和基于序列的都是对全脸进行分析的, 而基于 FACS 的是对人脸肌肉区域进行分析, 检测是否发生与疼痛相关的活动单元.

1.1.1.1 基于静态图像的特征

对静态图像进行分析时, 确定人脸位置后, 由于我们无法使用时间和空间上的信息, 因此我们利用图像标志点间的关系、像素点间的关系等来构造特征向量进行疼痛识别, 其中基于标志点之间关系的即为几何特征; 基于像素点间关系的即为外观特征. 基于几何特征的方法是早期的人脸识别算法^[40-42], 但由于疼痛表达的复杂性, 在静态图像上利用几何特征来进行疼痛识别十分困难, 相关文献几乎没有. 但是加上时间信息或者纹理特征的几何特征识别性比较理想, 会在后面进行介绍. 本节将基于外观的特征分为基于像素点间相对位置和基于像素点间数学变化来进行介绍.

1.1.1.1.1 基于像素点间相对位置的特征

记录像素点间相对位置的常用特征是局部二值模

式 (Local Binary Pattern, LBP), 它是一种用来描述图像局部纹理特征的算子, 计算简单, 所以 LBP 以及 LBP 的变体被广泛应用于人脸特征提取^[43-45]. 原始的 LBP 算子定义为在 3×3 的窗口内, 以窗口中心像素为阈值, 将相邻的 8 个像素的灰度值与其进行比较, 若周围像素值大于中心像素值, 则该像素点的位置被标记为 1, 否则为 0. 这样, 3×3 邻域内的 8 个点经比较可产生 8 位二进制数 (通常转换为十进制数即 LBP 码, 共 256 种), 即得到该窗口中心像素点的 LBP 值, 并用这个值来反映该区域的纹理信息. 这种 LBP 算子只覆盖了一个固定范围内的小区域, 可以通过将 3×3 邻域换为圆形区域来改进; 原始 LBP 算子具有灰度不变性不具有旋转不变性, 通过不断旋转圆形邻域得到一系列初始定义的 LBP 值, 取其最小值作为该邻域的 LBP 值来保证旋转不变性.

Aung 等^[29]为了检测 Emopain 数据库疼痛存在与否, 采用 LBP 特征和留一法交叉验证来训练 SVM, 识别精度相比 UNBC 肩部数据库较低, 这可以通过数据集性质之间的差异来解释: Emopain 数据库由长时间遭受慢性疼痛的参与者组成, 他们的许多表达被抑制, 这进一步使识别任务复杂化, 也体现了自然疼痛表情识别的难度很大. Nanni 等^[46]提出使用 LBP 纹理描述符及其变体 (局部三元模式 Local Ternary Pattern、细长二元模式 Elongated Binary Pattern 和细长三元模式 Elongated Ternary Pattern) 检测婴儿疼痛表达的方法. 在预处理阶段, 对图像进行调整大小、对齐、裁剪以获得精确的面部区域, 并将其划分为 25×25 维的块或单元格. 为了选择最有代表的单元格, 将 SFFS 特征选择算法应用于训练集, 结果表明将整个图像分成单元格可以改善性能. Li 等^[47]使用约束局部模型 (CLM) 进行婴儿不适检测, 将面部形状与检测到的面部区域相对应. 之后, 通过在相似性标准化外观 (SAPP) 上使用细长局部二值模式 (ELBP) 来提取表达特征.

由于数据收集和标记的困难, 在实际应用中特定人员数据非常稀少. 从稀疏数据中学习可能会严重过度拟合, 因此 Chen 等^[48]提出通过迁移学习来学习特定于人的面部表情模型. 通过传递来自其他人的信息知识, 它允许我们仅用少量某人的特定数据来学习针对新对象的准确人的特定模型. 实验在 UNBC 肩部数据库上测试, 圆形区域的 LBP 算子用作实验中的面部图像特征, 实验结果相比较 Lucey 等^[8]稍有提升, 而且

训练时间显著缩短。

1.1.1.1.2 基于像素点数学变化的特征

像素点数学变化中最常见的是利用主动外观模型 (Active Appearance Model, AAM) 和离散余弦变换 (Discrete Cosine Transform, DCT). AAM 是广泛应用于模式识别领域的一种特征点提取方法^[8,49]. 基于 AAM 的人脸特征定位方法在建立人脸模型过程中, 不但考虑局部特征信息, 而且综合考虑到全局形状和纹理信息, 通过对人脸形状特征和纹理特征进行统计分析, 建立人脸混合模型, 即为最终对应的 AAM 模型^[50]. DCT 变换是与傅里叶变换相关的一种变换, 它与离散傅里叶变换类似, 但是只使用实数。

Lukey 等^[8]利用 AAM 形状和外观来跟踪患者的面部, 使用跟踪到的信息从面部推导出相似的归一化形状 (SPTS) 和规范的归一化外观 (CAPP) 作为特征. Rupenga 等^[51]用 AAM 提取出特征, 并在 UNBC 数据库上评估 ELM(超限学习机) 和 SVM 两种分类器在疼痛事件检测中的有效性. Kaltwang 等^[52]提出基于面部表情对疼痛进行持续疼痛估计的模型中将离散余弦变换 (DCT) 应用于对齐的面部图像而获得外观特征. Agrawal 等^[53]首先提取面部区域, 再用 Gabor 过滤器提取特征. Fotiadou 等^[54]讨论了使用 AAM 检测急性疼痛过程中婴儿的疼痛表现, 从跟踪的面部提取 SPTS、SAPP(相似性标准化外观) 和 CAPP. 8 名婴儿的 15 个视频被用于构建自动化不适检测系统, 结果表明所提出的系统可以自动检测不适 (AUC=98%).

基于 AAM 模型的疼痛检测算法的限制是面部标注问题, 当已知面部的初始形状偏差太大、外观模型估计失败时, AAM 搜索不到一个很好的匹配来标注面部^[55], AAM 的另一个限制是其训练阶段计算复杂性较高。

此外, 进行逐帧标记的成本非常大, 因此 Ashraf 等^[49]研究序列级分类情况. 提供了一种利用 AAM 模型非刚性配准的方法来分离外观和形状分量, 利用 AAM 模型提取出 SPTS、CAPP 和 SAPP. 在 UNBC 肩痛数据库上利用 SVM 分类器进行分类, 实验结果为对于帧级疼痛检测, 正确率为 82%, 误报率为 30%; 序列水平标记的相应比率分别为 77% 和 44%. 结果表明, 学习帧级标签最适合于帧级检测疼痛. 然而, 即使在他们未被直接训练的任务上进行评估, 序列水平训练的分类也比几率好得多. 这个结果推动自动疼痛识别的研究人员在未来标记数据集时应该如何使用资源问题的探索. 继续在帧级标记确保学习的数据集的大小适中或

采用混合标记策略在帧中标记一些部分, 在序列级别标记某些部分, 以便在更大的数据集中学习。

基于帧的疼痛自动识别方法汇总见表 1.

1.1.1.2 基于序列的特征

基于静态图像的特征丢失了时间信息, 因此我们将静态图像加上时间信息组合成序列, 并从其中提取特征, 从而达到更好的识别效果. 在序列上提取特征的方法主要包括光流法、多示例学习、时空局部描述符以及深度学习等。

1.1.1.2.1 光流法提取的特征

光流是一种众所周知的运动估计方法, 其通过直接估计连续视频帧上的像素速度来工作. 它依据亮度守恒原理, 并提供密集的像素到像素的对应关系. 将光流法用于人脸提取相关特征的实验已有很多, Zhao 等在^[56]中应用了光流法和累积光流法取得很好的实验结果, 表明光流法适用于该领域。

Zamzmi 等^[57]提出光流法来检测视频中婴儿的疼痛表现. 在预处理阶段, 在每个帧中检测婴儿的面部并提取 68 个面部点. 然后将这些点用于对齐面部, 将其裁剪, 并将其分成 4 个区域. 为了提取与疼痛相关的特征, 在连续帧之间计算光学流动矢量并用于估计光学应变变量值, 以测量面部表情发生期间的面部形变. 然后, 将峰值检测器应用于应变曲线, 以找到对应于面部表情的最大应变变量值。

与疼痛相关的数据库在采集过程中保持了亮度恒定, 保证照明条件良好且不变, 因此数据比较适合采用光流法处理. 但是在现实情况下光流法的亮度恒定不变和小运动假设不容易满足, 时效性与精确性难以兼得. 在对于运动场景和目标无有效认知、运动模型难以预测、目标特征无法确定的情况下, 若使用光流法构建稠密光流场, 并且使用金字塔算法补充对于大幅度运动的检测精度, 可以有效提高检测精度. 但是, 这样的算法复杂度很高, 计算量巨大, 时效性较差。

1.1.1.2.2 多示例学习提取的特征

多示例学习 (MIL)^[58]可以被描述为: 假设训练数据集中的每个数据是一个包 (bag), 每个包都是一个示例 (instance) 的集合, 每个包都有一个训练标记, 如果包中存在一个以上正标记的示例, 则包被标为正标记; 而对于一个负标记的包, 其中所有的示例均为负标记. 通过定义可以看出, 与监督学习相比, 多示例学习数据集中的样本示例的标记并不都是已知的, 而监督学习的训

练样本集中, 每个示例都被标记; 与非监督学习相比, 多示例学习仅仅只有包的标记是已知的, 而非监督学

习样本所有示例均没有标记. 多示例学习是一种用于解决不同的弱监督问题的流行的建模框架.

表1 基于帧的疼痛自动识别方法汇总

人员	年份	使用特征	分类器	类别数	数据库	实验结果				
						准确率 (%)	AUC (%)	F1	Person 相关系数	均方误差 (MSE)
Lucey 等 ^[8]	2011	SPTS CAPP	SVM	2	UNBC	-	(单特征 CAPP) 80.9 (双特征融合) 83.9	-	-	-
Chen 等 ^[48]	2012	LBP	转移学习	2	UNBC	-	89.5	-	-	-
Min 等 ^[29]	2017	LBP	SVM	2	Emopain	-	64.5	0.386	-	-
Kaltwang 等 ^[52]	2012	PTS	RVR	17	UNBC	-	-	-	(单特征)	1.484
		LBP DCT							(三特征)	1.373
Agrawal 等 ^[53]	2015	Gabor	SVM	4	UNBC	82.43	-	-	-	-
Ashraf 等 ^[49]	2009	SPTS CAPP SAPP	SVM	2	UNBC	82	-	-	-	-
Li 等 ^[47]	2016	SPTS ELBP	SVM	2	11 名婴儿	81.5	-	-	-	-
Fotiadou 等 ^[54]	2014	SPTS CAPP SAPP	SVM	2	8 名婴儿 15 个视频	-	98	-	-	-
Nanni 等 ^[46]	2010	LBP LTB ELTP ELBP	SVM	2	COPE	-	ELTP 93	-	-	-

Sikka 等^[59]提出一种多段结合的多示例学习方法, 每个序列表示为包含多个段的包, 这些片段通过序列的多个聚类或运行多尺度时间扫描窗口生成, 并且使用词袋 (BoW) 来表示. 这种方法数据的表示不是作为单独的帧而是作为段来结合时间动态, 而且做到仅使用序列级正确标签来检测和定位疼痛的帧, 在试验过程中提取了多个段, 这适用于信号在视频中具有不确定的时间位置和持续时间的情况. Ruiz 等^[60]提出了新的多实例动态序数随机场 (MI-DORF) 模型, 这个模型将袋内的实例标签视为潜在的序数状态. MIL 假设通过将袋和实例标签相关的高阶基数潜力结合到能量函数中来建模, 这种方法充分考虑目标数据中的时间和序数两种类型的结构.

UNBC、EmoPain 数据库由人工标注时, 因数据量巨大, 难免会出现一些错误, 是典型的弱监督学习问题, 这类问题适合应用多示例学习来建模解决. 多示例学习中存在的问题就是示例标签的不稳定性, 也就是说, 如果训练集出现了轻微的变动 (例如数量略微增加和

减少), 那么输出的示例标签就会出现不稳定的情况.

1.1.1.2.3 时空描述符特征

仅通过面部表情的静态图像来准确判断疼痛强度十分困难, 可以利用静态图像的上下文信息^[61], 也就是在静态图像的时间序列、空间序列上提取相关特征.

Khan 等^[62]使用 Viola-Jones 算法^[63]检测视频中的人脸, 将检测到的面部分成两个相等的部分, 上面部分包括鼻子上部、眼睛和皱纹区域, 下部包括嘴部区域和鼻子下部, 从上下部分分别提取 PHOG(金字塔直方图) 和 PLBP(金字塔局部二进制模式) 特征, 并将它们连接起来, 形成最终的特征向量, PLBP 是 LBP 描述符的基于金字塔的空间表示.

Yang 等^[64]提出的方法包括 5 个主要步骤. 首先根据眼睛坐标和其他面部标注检测、分割和对齐人脸图像. 然后, 分别提取来自视频帧的空间纹理特征和来自视频序列的时空特征, 即 LBP, 局部相位量化 (LPQ) 和二值化统计图像特征 (BSIF), 它们都是使用 3 个正交平面 (TOP) 从视频中提取的. 使用早期融合将得到的

特征向量组合,作为SVM的输入,从而判别疼痛的表达是否存在于序列中. Werner等^[65]提出了一个新的特征集来描述面部动作及其动态,将其称为面部活动描述符. 这些描述符通过从时间序列及其第一和第二导数中提取特征,将特征信号(帧级特征的时间序列)浓缩成信号描述符. 新提出来的特征集的维度较低,而且运用了时间上的信息. Bourou等^[66]基于Openface跟踪面部和检测标注点并估计头部姿势,从感兴趣的区域提取若干距离及灰度、颜色强度值. 随后,在整个视频持续时间内,根据每个特征的时间序列计算统计指标,如均值、中位数、模式、最小值、最大值、标准差、方差、平均绝对偏差、偏度、峰度、熵、四分位数范围和相关系数. 鼻子和眉毛皱纹和鼻唇沟用于灰度强度值统计,以估计疼痛表达. 使用面颊来提取4个颜色通道中的值,以估计面部变红. 通过套索回归确定最具信息性的特征. 特征选择首次表明,在YCbCr颜色空间中编码的面部颜色确实是用于疼痛表达的高信息量. 与Werner等^[65]相比,该研究中鼻唇沟相关特征的相对重要性更高具有对称性. Chen等^[17]提出一种将帧级特征P-HOG、段级特征HOG-TOP,以及最大池化生成的序列级全局P-HOG和HOG-TOP结合. Zhi等^[67]提出了一种动态疼痛面部表情表示和融合方案,用于自动婴儿疼痛评估,结合时间外观面部特征和时间几何面部特征,构建不同的自动疼痛评估模型以更好地理解影响婴儿疼痛反应性的各种因素.

使用时空描述符表示的特征维度较低,但是特征集的特征数目较多,需要计算的特征较多,分类器需要学习的特征也比较多. 应用时空描述符方法的关键在于如何找到更具有特异性的特征,需要在实验中不断总结、尝试.

1.1.1.2.4 基于深度学习提取的特征

深度学习是对人脑处理信息时所采用方法的抽象总结和模拟. 未经处理的数据(图像,声音信息或者文字信息)被输入至输出层的“输入单元”;输入信息经过一定的映射输出至输出层的“输出节点”,映射的方法根据用户定义. 深度学习是一种通过多层神经网络对信息进行抽取和表示,并实现分类、检测等复杂任务的算法架构. 深度学习是近十年来人工智能领域取得的最重要的突破之一,它在人脸识别方面取得巨大成功^[68-72]. 深度学习架构已广泛用于人脸识别^[73-75]、面部表情识别^[76-78]和情感检测^[79-81],近几年很多研究人

员将深度学习应用到疼痛识别领域^[18,82-84].

Zhou等^[82]将UNBC视频序列裁剪出人脸,再拉平连接成帧向量序列传入到RCNN网络中预测帧的疼痛强度. Wang等^[83]使用现有的深度学习网络结构进行微调,将现有效果最佳的面部对准网络结构^[85]去除全连接层而取其卷积层,然后加入两个新的全连接层. 为了防止过拟合,将原有网络的全连接层的隐藏单元数从512改为50,另外加上center loss^[85]来正则化网络的回归损失函数.

Kharghanian等^[86]提出使用无监督学习特征的一种方法来实现连续的疼痛检测,使用卷积深度置信网络(CDBN)提取特征. 提取的特征包括疼痛图像的不同属性,例如头部运动、形状和外观信息. Martinez等^[87]提出了一种层次学习框架,采用双向长短期记忆递归神经网络(LSTM-RNN)来估计从面部图像中提取的面部标注点的PSPI(一种疼痛评估方式, $PSPI=AU4+\max(AU6, AU7)+\max(AU9, AU10)+AU43$). 然后,PSPI用作隐藏条件随机场(HCRF)的输入,以估计受试者的VAS(视觉类比量表). Egede等^[88]结合手工提取的特征和深度学习来进行疼痛自动评估. 将时间信息包括到学习过程中,时间 t 处的图像的特征表示包括来自前一帧和后一帧的信息,将提取出的图像区域和二进制掩码由CNN学习得到深度学习特征. Rodriguez等^[89]利用CNN-LSTM模型在数据库提供的面部标记点来裁剪和正面化面部图像,并作为CNN的输入来学习特征,然后对无痛数据进行欠采样来平衡数据的不均匀分布.

深度学习模型将特征提取和分类过程进行联动优化,并在疼痛自动识别中取得了有竞争力的结果. 然而,深度学习也存在一些不足. 首先,网络结构复杂致使调参过程复杂,且导致结果的可重复性低;其次,深度学习很容易造成过拟合现象,通过数据扩充、预训练模型微调、dropout、批标准化等操作可有效改善但不能完全解决;最后深度学习的网络需要大量数据才能获得良好的性能,而目前在疼痛识别方面公开的数据集数量非常有限,这使得网络模型在学习特征时候无法充分学习到它需要的信息,阻碍了它在自动疼痛识别中的应用. 基于序列的疼痛自动识别方法汇总见表2.

1.1.1.3 基于FACS的方法

面部表情可以表示疼痛存在和严重程度,是一个具有敏感性和特异性的行为指标. 上一章节我们主要

讨论了针对面部整体进行分析的疼痛自动分析,此外也可以通过 FACS 开展疼痛评估.面部动作编码系统通过使用称为 AU 的 46 个基于解剖学的组件运动来测量面部表情.面部表情自动识别的重大进展推动了其在基于 FACS 的情感计算问题中的众多应用^[90].

Sikka 等^[6]利用计算机表达识别工具箱 (CE RT),对阑尾切除术后儿童疼痛的面部表情进行自动分析.实验的对象是在过去 24 小时内接受腹腔镜阑尾切除术的 50 名年龄在 5 至 18 岁的青少年,利用计算机视觉机器学习 (CVML) 模型进行疼痛二元分析和疼痛评级. CVML 疼痛评估模型在检测临床显著疼痛和评估

术后疼痛严重性方面表现良好, CVML 模型与常用的代理性疼痛评估方法一样,表明在手术后监测儿科疼痛状态的功效.在成人方面, Bartlett 等在^[91]中利用 CERT 来检查真正疼痛和虚假疼痛,该研究成果进一步表明面部表情的动态变化是人类情感表达的重要区别特征.

基于 FACS 的方法的主要挑战是在每个视频帧中人工标记 AU 需要大量时间.据报道,人类专家需要大约 3 小时来编码 1 分钟的视频序列.降低标签成本的方法是开发一个可靠的系统来自动检测每个帧中的 AU 并将其用作标签.

表 2 基于序列的疼痛自动识别方法汇总

人员	年份	特征	分类器	类别数	交叉验证 协议	数据库	实验结果					
							准确率 (%)	MSE	CORR	ICC	MAE	AUC(%)
Sikka 等 ^[59]	2014	MIL	SVM	2	留一法	UNBC	83.7	-	-	-	-	-
Ruiz 等 ^[60]	2016	MIL	MILBo ost	6	留一法	UNBC	帧序列 52	-	0.40	0.40	0.19	-
Khan 等 ^[62]	2013	PLBP PHOG	KNN	2	10 折交叉验证	UNBC	96.9	-	-	-	-	-
Yang 等 ^[64]	2017	LBP-TOP	SVM	2	留一法	UNBC	83.42	-	-	-	-	-
		BioVid				65.17	-	-	-	-	-	
Werner 等 ^[65]	2017	面部活动描述符 信号描述符	SVM	2	5×2 折交叉验证	UNBC	91.7	-	-	-	-	-
			/RF	2		BioVid	72.4	-	-	-	-	-
			RF	5		BioVid	30.8	-	-	-	-	-
Bourou 等 ^[66]	2018	距离、灰度和 颜色强度值	RF	5	10 折交叉验证	BioVid	27.13±2.24	-	-	-	-	
Chen 等 ^[17]	2016	HOG HOG-TOP	SVM	2	留一法	UNBC	91.4	-	-	-	-	
Kharghanian 等 ^[86]	2016	CDBN 学习到的特 征	SVM	2	10 折交叉验证	UNBC	87.2	-	-	-	-	94.48
Egede 等 ^[88]	2017	HOG	RVR	16	留一法	UNBC	-	0.98	0.67	-	-	-
		手工提取几何特征 CNN 深度学习特征										
Zhou 等 ^[82]	2016	RCNN 学习的特征	RCNN	16	留一法	UNBC	-	1.54	-	-	-	-
Wang 等 ^[83]	2016	CNN 学习到的特征	CNN	16	25 折交叉验证	UNBC	-	1.72	-	-	0.99	-
Martinez 等 ^[87]	2017	RNN 学习的特征	HCRF	16PSPI	-	UNBC	-	-	-	0.30	0.94	-
			模型	11 VAS						0.34	2.46	-
Rodriguez 等 ^[89]	2017	CNN 深度学习的特 征	LSTM	6	留一法	UNBC	-	0.74	0.78	0.45	0.5	91.3
Zammi 等 ^[57]	2015	光学流动矢量	KNN	2	10 折交叉验证	10 名 婴儿	96	-	-	-	-	-
Zhi 等 ^[67]	2018	时间外观面部特征 时间几何面部特征	SVM	2	留一法	IPAD	89.33	-	-	-	-	-

1.1.2 基于身体运动的特征

在疼痛发生过程中,除了面部表情的表达,身体运动也是一个重要的参照指标,尤其是对慢性病患者和

婴儿来说,在他们的疼痛发生过程中,身体运动的变化会尤为明显.另外头部运动和姿势^[8]也可以作为衡量疼痛的标准, Haque 等^[21]在 BioVid 热痛数据库、UNBC

肩痛数据库和 BP4D 数据库上进行了头部姿势汇总、头部运动汇总、特定头部姿势发生、头部运动聚类的统计分析,发现疼痛伴随着头部姿势和运动,其倾向于向下或朝向疼痛部位。

在成人方面,大部分的疼痛自动识别的研究都是围绕面部表情或与生理特征相结合的多模态来开展的,这是因为在疼痛领域 UNBC 肩痛数据库和 BioVid 热痛数据库的促进作用。对于身体运动的疼痛评估研究相对较少。Olugbade 等^[92]提出在身体运动期间根据特征区分低慢性疼痛患者、高慢性疼痛患者和非慢性疼痛患者。实验通过两层 SVM 或随机森林分类器进行分类,在 EmoPain 数据库上进行验证。实验提取了完整躯干屈曲练习的特征集和坐到站练习的特征集,并加上抑郁分数,分别来区分这两个动作。在完整躯干屈曲中,最好的实验结果是在 SVM 模型下利用身体运动特征的优化,得到 94% 的准确率;在坐到站实验中,最好的实验结果是 80% 的准确率。Wang 等^[93]首次将诸如 LSTM 之类的递归神经网络用于检测保护行为,在 EmoPain 数据库上达到 81.5% 的准确率。

当婴儿经历疼痛时,婴儿往往会摇头,伸展手臂或腿,并伸展手指。Stahl 等^[94]提出了一种基于光流的算法来预测患有脑瘫 (CP) 疾病风险的婴儿。采用的数据库包括在 10–18 周龄范围内为 82 名婴儿 (15 名被诊断为 CP 和 67 名健康者) 记录的 136 个视频。对于每个视频,应用光流以生成运动轨迹。然后,将这些轨迹转移到时间相关的信号,并进一步分析以提取 3 种类型的特征:小波系数,绝对运动距离和相对频率特征。小波系数测量婴儿运动的种类,另外两个特征测量运动模式中的活动和发生的频率。对于婴儿分为受损或未受损的分类,线性 SVM 在使用相对频率特征,绝对运动距离和小波系数训练时分别达到 $93.7 \pm 2.1\%$, $91.7 \pm 2.2\%$ 和 $84.7 \pm 1.8\%$ 平均准确率。

一般来说,突发刺激、慢性疾病患者和婴儿的疼痛发生过程中更容易引发身体上的运动,而对其他人来说疼痛发生的过程中身体运动不一定存在,而且目前对于成人来说只在 EmoPain 数据集采集了参与者身体运动的信息,因此基于身体运动来进行疼痛分析的实验很少。此外,仅通过身体运动来判别疼痛不够全面,我们应将其他特征和身体运动相结合,相关内容将在

第 4 节进行介绍。

1.2 基于语音的特征

在疼痛表达过程中,除了行为特征外,语音特征也是一个重要因素,特别是对于婴儿来说,啼哭是非常重要的一个信号。对于成人来说,由于在疼痛过程中可能很少会发出语音信息,只有少部分的数据集记录与语音相关的信息,与成人相关的基于语音的疼痛识别的相研究还很少见,EmoPain、SenseEmotion 数据库中有相关语音记录,但还没被适当运用。Thiam 等^[95]使用 SenseEmotion 数据库上的音频特征进行二分类任务得到 $69 \pm 15\%$ 的准确率,效果不太理想。因此接下来重点介绍基于语音的婴儿疼痛检测。

婴儿啼哭是不适、饥饿或疼痛的常见信号^[96],它传达的信息可以帮助护理人员评估婴儿的情绪状态并做出适当的反应。我们将基于啼哭分析的特征提取方法分为时域方法、频域方法和倒谱域方法。

1.2.1 时域分析提取的特征

时域分析是关于时间的信号分析(信号的幅度随时间的变化)。线性预测编码(LPC)是用于分析声音的最常见的时域方法之一。LPC 主要原理是使用过去时域样本的线性组合来预测当前时域样本。用于婴儿声音分析的其他时域特征是能量、幅度和暂停持续时间。

Vempada 等^[30]提出了一种时域方法来检测与不适相关的哭泣。所提出的方法在数据集上评估,该数据集包括在疼痛(30个语料库)、饥饿(60个语料库)和湿尿布(30个语料库)期间收集的120个哭泣语料库。文中没有提供有关触发疼痛状态的刺激信息和数据收集程序的信息,婴儿的年龄范围为12–40周龄。使用索尼数字记录仪记录所有语料库,采样率为44.1 kHz。在特征提取阶段,计算两个特征:短时能量(STE),它是在合适的窗口样本值的平方的平均值;以及在哭泣段内暂停持续时间。部分样本用于构建SVM,其余样本用于评估其性能。痛苦哭泣、饥饿哭泣和湿尿布哭泣的识别表现分别为83.33%、27.78%和61.11%,平均识别率为57.41%。

1.2.2 频域分析提取的特征

频域分析显示信号在特定频率范围内的分布。基频(F0)是众所周知的频域特性,表示周期信号的最低频率。婴儿的哭声可以根据基本频率分类为:

(1) 带有平滑谐波结构的发出嘶嘶声(phonated cries),基本频率范围为400~500 Hz。

(2) 与发声的啼声 (phonated cries) 相比, 具有较少谐波结构的噪音。

(3) 声音突然向上移动 (高达 2000 Hz) 的超音速哭声 (Hyperphonated cries)。

Pai 等^[31]提出了一种光谱方法, 将婴儿的哭声分类为呜咽或有力。这项工作的数据库来自 27 名婴儿, 平均年龄为 36 孕周。在急性疼痛程序 (即, 足跟穿刺和免疫接种) 期间记录音频数据。记录两种类型的疼痛哭泣, 呜咽 (14 个样本) 和剧烈哭泣 (20 个样本)。记录样本的正确标签由训练有素的护士使用 NIPS 疼痛量表给出。为了获得每个样品的功率谱, Welch 的方法应用于 20 毫秒的窗口。获得光谱后, 从每个样本中提取线性预测系数 (LPC) 以及其他统计数据 (例如, 平均值和标准偏差) 并用于训练 KNN, 使用 10 倍交叉验证评估的分类器的平均准确率为 76.47%。

1.2.3 倒谱域分析提取的特征

通过采用信号频谱的对数的逆傅立叶变换 (IFT) 来生成信号的倒谱域。Mel 频率倒谱系数 (MFCC) 是一种常见的倒谱域方法, 用于从声音信号中提取有用且有代表性的特征集 (即系数), 并丢弃噪声和无用的特征。

Abdulaziz 和 Ahmad 等^[96]引入了一种倒谱域方法, 将婴儿哭泣分类为疼痛或无痛 (即饥饿和愤怒)。对于年龄范围从新生儿到 12 个月大的婴儿, 记录了一组 150 个疼痛样本和 30 个无痛样本。在当地医院的 NICU (新生儿重症监护病房) 中的常规免疫程序期间记录疼痛样本, 在婴儿家中记录无痛样本。在 180 个记录的样本中, 通过创建一个第二段获得 881 个样本。然后使用这些样本提取两组特征, 即 Mel 频率倒谱系数 (12 个 MFCC 系数) 和线性预测倒谱系数 (16 个 LPCC 系数), 并将提取的特征送到用缩放共轭梯度算法训练的神经网络。该方法分别对 LPCC 和 MFCC 的准确率分别达到 68.5% 和 76.2%。该结果表明, MFCC 在检测婴儿疼痛哭泣方面的表现优于 LPCC 特征。

1.3 基于生理指标的特征

基于生理测量的疼痛分析可以定义为从参与者的生理反应中提取疼痛相关特征的过程, 疼痛的发生同时会伴随生理指标的变化^[97]。与成人相关^[98]的生理反应有心电 (ECG)、脑电 (EEG)、肌电 (EMG)、皮肤电导水平 (SCL) 等生理信号。与婴儿疼痛相关的生理反应实例包括生命体征的变化和脑血流动力学活性 (脑

氧合和疼痛的变化), 适合床边检测的脑氧合变化测量方法为近红外光谱 (NIRS), 它使用附着在头部的小探针测量氧合血红蛋白 (HbO₂) 和脱氧血红蛋白 (HbH) 浓度的微小变化。另外, 体温大于 0.5 的变化也可能预示疾病或者疼痛^[28]。

在成人方面, Werner 等^[99]和 Walter 等^[100]在 Biovid 数据库上进行实验, 在生理特征方面提取电流皮肤反应 (GSR)、EMG、ECG 的幅度和变化特征用于检测疼痛。对单一生理特征来说, GSR 具有显著优势, 在疼痛强度 0 和 4 二分类上的正确率为 73.8%; 将 3 种生理特征结合会使 0 和 4 二分类正确率提高到 75.6%。Walter 等从生理信号的幅度、频率、平稳性、熵、线性和可变性的数学分组中提取 135 个特征并使用前向传播进行特征选择, 采用 SVM 进行二分类, 达到 77.05% 正确率。Kächele 等在文献^[101]中提取的基于生理信号的特征有 EMG、ECG、SCL, 实验结果表明最具有识别性的生理特征为 SCL, 在 Biovid 数据库上识别正确率为 81.9%。Panavaranan 等^[102]通过考虑单个通道的功率谱密度的 α 和 β 谱对 EEG 带来的影响进行疼痛分类。

在婴儿方面, Faye 等^[103]提出了一种分析 28 名患有慢性疼痛的婴儿 (年龄 > 34 孕周) 的心率变异性 (HRV) 的方法。为了研究慢性疼痛和心血管数据之间的关联, 使用心率 (HR), 呼吸率 (RR), 血氧饱和度 (SpO₂) 和高频变异指数 (HFV I) 的平均值进行线性回归分析, 两组之间的 HRV 显著下降, 而未发现 RR 和 SpO₂ 的显著变化。结果显示 HFVI (< 0.9 阈值) 能够评估疼痛, 实验结果为在 ROC 曲线下面积 (AUC) 为 0.81, 说明与婴儿疼痛发生最相关的生命体征为心率。Ranger 等^[104]提出了一种基于 NIRS 的方法, 通过分析大脑区域的血液动力学活动来评估婴儿的慢性疼痛。NIRS 数据来源于 40 名婴儿 (< 12 个月) 心脏手术后的胸腔排出移除过程 (T2)、脱掉衣服 (T1) 和基线 (T0) 3 个阶段的记录。为了验证 NIRS 数据和疼痛刺激之间的关联, 对提取的测量进行单变量线性回归, 结果显示疼痛期间 HbH 显著增加。

基于生理特征分析的疼痛检测在机理解释方面存在一些挑战^[28], 例如药物、病理状况、治疗、压力和对疼痛的恐惧可能直接影响这些特征并降低疼痛测量的可靠性。它还受婴幼儿的一般健康状况和年龄的影响, 所以仅通过生理信号来确定疼痛的发作是不合适

的,应该与其他指标结合使用可获得更好的结果。

1.4 基于多模态的特征

以上介绍的方法都是基于单个模态特征的,但是在疼痛发生的过程中,这些模态是同时存在的,多模态评估疼痛具有挑战性但是非常合理的^[105]。因此接下来介绍的方法是将面部表情、身体运动、语音以及生理等特征融合的多模态方法,并根据特征融合的时期将方法分为特征层级的融合和决策层级的融合。

1.4.1 特征层级融合得到的特征

特征级融合将所有模态的特征连接成单个高维特征向量,然后使用级联特征向量来训练单个分类器以进行分类。据我们所知,目前没有工作在特征层面结合不同的疼痛指标,以评估婴儿的疼痛。

在成人疼痛评估方面,Haque等^[24]利用深度学习方法将提取出来的颜色图像、热图像和深度图3种模态输入分别进行训练、特征级融合和决策级融合。早期融合为每个视频帧生成一个5维的矩阵做为CNN的新输入。晚期融合中,将各个分类器的输出集成为第二个堆叠分类器的输入特征向量,第二个分类器生成最终判定结果。在单特征中得到的最好结果是准确率为18.55%,融合再进行分类后得到的结果都有提升。Werner等^[99]提出了将视频和生物学信号结合起来用于成人疼痛评估。Thiam等^[95]提取了音频、几何描述符、头部姿势、LBP-TOP以及ECG、RSP、EMG、EDA等生理特征,用RF分别做特征级融合和决策级融合,在SenseEmotion数据库上做了二分类和多分类任务。

从理论上讲,特征级融合可以比决策级融合具有更高的性能,因为它包含更丰富的信息。但是,这种融合程度可能会在实践中引发一些问题,对这些问题的不当处理可能会降低性能。例如,将不同疼痛模态的特征连接成单个高维特征向量可能导致维度灾难等。特征级融合的另一个问题是由于未能记录特定模态或在特定时间无法获得数据而导致的数据缺失。

1.4.2 决策层级融合得到的特征

决策级融合旨在将多个分类器的决策或结果合并为一个单独的决策。换句话说,决策级方法考虑了多个分类的结果,即每个疼痛指标或模态的分类,以确定最终的决策或结果。Wagner等^[106]提出了几种不同方法来

将结果结合起来用于决策。

多数投票是将不同分类器结果融合起来的最常见方法之一。在多数投票方案中,每个指标贡献一票(即类别标签),并且投票得到的多数标签被选择作为最终决定或结果。Kächele等^[107]在BioVid热痛数据库上利用多模态进行疼痛自动评估的实验,提取了生理信号和面部表情、头部姿势;Kächele等^[101]利用生理信号EMG、ECG、SCL特征以及基于面部几何距离和基于外观的LBP-TOP特征,此外还介绍了一种自适应置信度学习来修改参数。Zamzmi等^[27]利用多数投票方法将不同的疼痛指标结合起来,以开发多模式疼痛评估系统。每种模态的特征(即,面部表情,身体运动和生理信号)被单独使用以构建分类器。Worley等^[108]结合EEG、EMG、ECG、NIRS和视频监控到的行为动作特征构建疼痛事件监测系统。

决策级融合的一个主要优点是易于实现,因为它依赖于组合不同的分类标签。这种融合水平可能导致信息丢失(即不同模态之间相关信息的丢失),因为它取决于组合指标是独立的假设。

由疼痛评估标准的量表可知,疼痛是由多种模态信息共同作用的,如身体运动、语音、生理特征等,因此,多模态的特征融合方法会带来更多有效信息,如何将这些信息进行有效结合利用是研究重点。

基于多模态的疼痛识别方法汇总如表3。

2 相关数据库

目前和疼痛相关的数据库并不是很多,而且规模一般偏小^[4],有些数据库的疼痛表达过程不够自然,而且记录的信息有限。下面从成人和婴儿两个方面介绍公开发布的疼痛数据库。

2.1 成人

与成人相关现在被用于研究最多的是Lucey等在发布的UNBC肩痛数据库^[8]和Walter等在发布的BioVid热痛数据库^[9]。近两年,Min等提出EmoPain多模态数据库^[29],Haque等提出MIntPAIN多模态数据库,Velana等提出的SenseEmotion多模态数据库^[109],还有一些未经公开的数据库其具体内容我们难以知晓,如文献^[16,110]中提到的自采数据库。与疼痛相关数据库的归纳如表4所示。

表3 基于多模态的疼痛自动识别方法汇总

研究者	年份	特征	分类器	类别数	数据库	融合策略	实验结果 (准确率 %)
Haque 等 ^[24]	2018	CNN 学习到的不同模态的特征 音频、几何描述符、头部姿势、	CNN	5	MIntPAIN	特征级	36.55
				5		决策级	25.40
Thiam 等 ^[95]	2016	LBP-TOP、ECG、RSP、 EMG、EDA 等生理特征	RF	2	SenseEmotion	特征级	85±12
						决策级	85±11
Werner 等 ^[99]	2014	几何和生理特征	RF	2	UNBC	特征级	80.6
Kächele 等 ^[101]	2017	EMG、ECG、SCL、面部 LBP-TOP	RF	2	BioVid	决策级	83.1
Zamzmi 等 ^[27]	2017	面部表情, 身体运动和生命体征	SVM	2	18 名婴儿	决策级	95

表4 疼痛相关数据库汇总

数据库	受试者人数	受试者类型	疼痛水平	疼痛类型	数据库大小	年份
UNBC 肩痛	129 (公布了 16 个)	自我识别为肩部疼痛 问题的参与者	帧: 0-16 (PSPI) 序列: 0-5 (OPI) 0-10 (VAS)	自发肩痛	200 个视频, 31 571 帧	2011
BioVid 热痛	90 (87 个可用)	健康参与者	1-4 (疼痛刺激)	热刺激	17 300 个视频, 帧速率为 25 fps/s	2013
EmoPain	50 (46 个可用)	22 名慢性腰背痛患者, 28 个健康参与者	0-1 (自行开发的操纵杆)	自发慢性疼痛	总帧数为 585 487	2017
MIntPAIN	20	健康参与者	0-4 (疼痛刺激)	电刺激	9366 个视频, 187 939 帧	2018
SenseEmotion	45	健康参与者	0-3 (疼痛刺激)	热刺激	每个受试者两段长约 30 分钟的视频	2016
COPE	26	健康参与者	4 种刺激	物理刺激	204 张静态图像	未公开
IPAD	31	急性疼痛治疗	3 种疼痛状态	治疗刺激	未知	未公开

2.1.1 UNBC 肩痛数据库

为了促进自动疼痛检测这项工作, UNBC 的研究人员招募了 129 名被自我识别为肩部疼痛问题的参与者 (63 名男性, 66 名女性). 参与者中 1/4 是学生, 其他人来自社区, 包括各种各样的职业. 参与者的肩痛的诊断类型各不相同, 超过一半的参与者报告使用药物治疗疼痛.

所有参与者参加了 8 项标准运动范围的主动和被动测试. 在测试期间, 两台相机记录了参与者的面部表情. 每次测试结束后, 参与者利用两个 Likert-type 刻度的卡片对疼痛进行口头评级. 实验结束后, 接受过大量训练的观察者独立评估视频的疼痛强度 (OPI), 范围从 0(无疼痛) 到 5(强烈疼痛), 通过 Pearson 相关系数评估了评分的高可靠性. 此外, 对视频中每一帧进 FACS 编码并计算 PSPI 分数, 其中 83.6% 的帧具有 0 的 PSPI 得分, 并且 16.4% 具有 PSPI 得分 ≥ 1 的帧.

2.1.2 BioVid 热痛数据库

共有 90 名受试者参加了该实验, 这些受试者是从 18-35 岁、36-50 岁、51-65 岁 3 个年龄组中招募, 每个年龄组各 30 名受试者, 其中男女性各 50%. 该实验

测量了包括 SCL、ECG、EMG 和 EEG 在内的生物电位, 另外还采集了视频信号.

视频录制过程使用了 3 台相机, 一台直接放在参与者面前, 另外两台放在两侧来保证参与者自由地移动头部. 相机以 25 Hz 的帧速同步触发, 并以 1388 \times 1038 的分辨率记录. 为了同步视频流和生物反馈数据, 使用 Nexus 设备记录了相机触发信号的分频版本以及生物反馈信号. 然后自动分析这个已知触发信号的记录, 以补偿生物反馈和相机视频记录之间的偏移和漂移. 此外, 在正面相机上方使用了 Kinect 传感器来记录深度图 (640 \times 480 像素, 大约 30 Hz), 彩色图像 (1280 \times 1024 像素, 大约 10 Hz) 以及 Kinect 提供的相关时间戳.

疼痛诱发通过右臂的热电极实现, 主要包括 6 步: 第 1 步进行阈值校准来确定每个人的特定疼痛水平温度; 第 2 步疼痛刺激, 用由第 1 步确定的 4 种个体特定疼痛水平随机刺激参与者. 每个水平被刺激 20 次, 总共 80 次刺激. 每个疼痛水平的最高温度保持 4 秒, 刺激之间随机暂停在 8-12 秒之间; 第 3 步参与者提出自己的每一个基本情绪; 第 4 步用图片来启发情感; 第

5步通过剪辑的视频来引发基本情绪;第6步用面部肌电图来重复第2步.该数据库的疼痛强度由刺激强度确定,是在视频序列上标记,没有帧上的标记.

2.1.3 多模态 EmoPain 数据集

该实验招募了50名参与者,其中22名慢性腰背痛患者(7名男性,15名女性,平均年龄50.5岁).虽然参与者数量很小,但该群体是典型的慢性疼痛寻求治疗的人.另外招募了28名没有慢性腰背痛病史的健康对照参与者(14名男性,14名女性,平均年龄37.1岁),用来记录在没有疼痛的情况下自然运动的各种方式.

数据采集过程将3个感觉系统连接到参与者:4个无线表面肌电探针,1个由18个基于微机电(MEMS)的惯性测量单元(IMU)组成的动作捕捉套装和1个头戴式麦克风.配置了8个相机让参与者在运动期间进行更多不受约束的指令,从而捕获自然运动.使用两个麦克风通道捕获音频信号,以48 kHz的速率和24位脉冲编码调制进行记录.

对记录到的数据让评估者进行面部表情标记,评分为0-1之间的任意一个值,将0.02作为阈值.专家观看视频并进行身体运动的标记,将身体行为运动分为6类.数据集中的总帧数为585 487,其中50 071(占总数的8.6%)被判断为疼痛表情.

2.1.4 多模式强度疼痛(MIntPAIN)数据库

MIntPAIN数据库除了面部表情以外,包含使用Kinect捕获的正面面部彩色图像和深度数据,以及由热像仪捕获的热数据.EMG从受试者左臂记录,EMG信号经过预放大和滤波(10-500 Hz)并存储为1 s记录.受试者的感知疼痛强度在10 cm电子VAS上评定,指定0作为感知阈值,5作为感受疼痛值,10作为可想象的最强烈疼痛.每个刺激由受试者评定并储存.

MIntPAIN数据库通过在20个健康受试者中给予5种不同水平(0级至4级,其中0表示无刺激,4表示最高刺激程度)的电刺激而获得多模式疼痛数据.每个受试者在数据捕获期间展示两个试验,且每个试验具有40次疼痛刺激扫描,最终获得9366个实验视频.

2.1.5 SenseEmotion 多模态数据库

SenseEmotion数据库通过热刺激前臂来诱发疼痛,热刺激的温度在32-55℃(分为T0-T3等级,T0 32℃为无疼痛).在记录数据之前,确定每个受试者的特定阈值温度(T1)和耐受温度(T3),通过T1和T3的平均值来计算中间温度(T2).每个温度的刺激随机施加

30次并保持4 s,在连续刺激之间暂停8-12 s.实验分为两个阶段,包括15分钟的校准时间、30分钟的实验时间和6分钟的评估时间,每个阶段结束后换另外一个手臂进行实验.此外,使用情感图片和声音刺激来诱发情感状态.

SenseEmotion数据库记录了SCL、ECG、EMG和RSP(呼吸)生理信号,视频信号以及语音信号.

2.2 婴幼儿

目前与婴幼儿相关的疼痛数据库基本尚未公开,下面描述的COPE和IPAD的数据集均是在疼痛自动识别相关文献实验数据库介绍部分得到的.

2.2.1 COPE 数据库

该数据集包括26名白人婴儿,其中一半为女孩,使用尼康D100数码相机拍摄的204张彩色图像.婴儿的年龄从18小时到3天不等,所有婴儿健康状况良好.在经历4种不同的刺激时拍摄婴儿的面部图像:脚跟穿刺期间的疼痛刺激,婴儿从一个婴儿床运送到另一个婴儿床期间的休息/哭泣刺激,空气刺激鼻子和摩擦刺激,包括用浸泡在酒精中的棉花接受脚跟外侧表面上的摩擦.尽管在COPE数据库上的研究已经取得了相对较好的疼痛识别效果,但COPE数据库仍然存在不足:它是一个相对较小的数据库,并且只有剧痛表情照片,没有任何有关疼痛程度的信息.此外,它是一个静态图像数据库,可提供的疼痛信息有限.

2.2.2 IPAD 数据库^[67]

该数据集共记录了31名婴儿,一半的婴儿是男性,婴儿的平均孕龄为36.4周,范围为30.4至40.6(标准差=2.7).采集在Tampa综合医院新生儿重症监护病房住院期间,接受常规疼痛手术(如足跟穿刺)刺激中婴儿的疼痛数据,持续时间约为5 s.妊娠37周前出生的婴儿称为早产,足月妊娠为37周至42周.

婴儿视频录像由GoPro Hero3+相机获取,捕捉他们的面部表情,身体动作和声音.摄像机设置在正常临床环境中,记录急性疼痛治疗期间婴儿的自发反应.新生儿婴儿疼痛量表(NIPS)是早产儿和足月儿的可靠且有效的基于指标的疼痛量表.涉及行为和生理指标,如面部表情,哭泣,呼吸模式,手臂运动,腿部运动和唤醒状态.婴儿的疼痛程度分为3组,由专业医护人员根据NIPS的总疼痛评分决定,即无疼痛(0-2),中度疼痛(3-4)或严重疼痛(>4).护士在疼痛过程中每隔1分钟评估婴儿疼痛指标的严重程度,并将总疼痛评分

作为婴儿疼痛评估的标签信息。

3 分类器

疼痛检测及评估的决策阶段,需要选择合适的分类器,对学习到的特征进行分类。在疼痛识别任务中,分类目标包括两大类,一种是二分类,即输出结果为疼痛或者无疼痛;另一种为疼痛强度的估计,输出结果为疼痛的等级,为0-4或1-4等。下面我们根据分类目标来介绍常用的一些分类器以及它们的原理。

3.1 疼痛检测

早期的研究往往是对受试者进行疼痛检测,目标是判断疼痛的存在与否,这是一个典型的二分类问题,最常用的分类器是 SVM。

SVM 可以分为线性和非线性两大类。其主要思想是寻找空间中的一个能够将所有数据样本划分开的超平面,并且使得所有数据到这个超平面的距离最短。给定训练样本集,在特征空间上找到一个分离超平面,将样本点分到不同的类。其中存在唯一的分类超平面,使得几何间距最大。其中,距离超平面最近的点为该超平面的支持向量。对于待测点,通过计算该点相对于超平面的位置进行分类。距离分离超平面的距离越大,表示分类预测的确信程度越高。在文献[8,29,46,47,49,54,59,62,64,67,86,94,111-113]均使用 SVM 作为分类器来输出结果,在二分类任务中,一般 SVM 是我们的首选,因为 SVM 非常适用于小样本集而且泛化能力比较好,目前的疼痛数据集数量都很小, SVM 成为疼痛检测任务中首先考虑的分类器。

另一种常见的二分类分类器为 KNN, KNN 算法的核心思想是如果一个样本在特征空间中的 k 个最相邻的样本中的大多数属于某一个类别,则该样本也属于这个类别,并具有这个类别上样本的特性。该方法在确定分类决策上只依据最邻近的一个或者几个样本的类别来决定待分样本所属的类别,在文献[32,57,62]中应用了 KNN 来进行二分类。KNN 算法非常简单有效,重复训练的代价比较低,但是该算法比较适用于样本容量比较大的类域的自动分类,而那些样本容量较小的类域采用这种算法比较容易产生误分。

每个分类算法都可用来进行二分类,但有些算法较复杂或者分类性能不如 SVM,故不常用。在实验过程中可多尝试几种分类算法来找到最好的分类结果,在结果相差不大的情况下应优先选择计算量简单的。

3.2 疼痛强度估计

疼痛强度估计为多分类任务,我们关心的结果不仅是疼痛的存在与否,更关心疼痛的强度,这项任务在现实生活中的应用价值更大,可根据疼痛强度来进行相关治疗。

在疼痛强度估计任务中,随机森林是最常用的分类算法。它通过大多数单个树的决策来预测给定测试模式的输出。每个树都是在一个训练集中随机选择的一个例子构建的。文献[65,66,101]使用 RF 来进行疼痛强度估计,RF 在训练过程中能够处理很高维度的数据,不用做特征选择,而且能预测各特征之间的互相影响,利用 RF 分类器能判断出各特征的重要性排名,这在实验中广泛应用。

多分类中另一个常用分类算法是神经网络及其变体,神经网络是将一组训练集送入模型,根据网络的实际输出与期望输出间的差别来调整权值。Martinez 等^[87]使用 HCRF 模型进行多分类,Egede 等^[88]应用 RVR, Rodriguez 等^[89]使用 LSTM。神经网络的自学习和自适应能力比较强,但是它的学习过程是一个黑盒模型,而且需要足够的数据才能学到相应信息。此外, SVM 也可被应用于多类分类任务,通过 1-1 或者 1-all 策略实现多类分类^[41,107,114,115]。

4 发展的挑战及方向

我们目前获得的实验结果与人为误差水平还有很大的差距,大部分性能差距不是由识别系统的技术弱点引起的,而是由于疼痛及其表达的复杂性所带来的挑战。在实验过程中,我们遇到的挑战主要有数据库问题、评判疼痛的标准问题、表现力差异等。

首先,公开发布的针对疼痛的数据库极大地推进了疼痛自动识别领域工作的推进,但每个数据库的参与者人数和数据量完全满足需求。现存的数据库也存在着一一些问题,比如不是自发情况下引起的疼痛等。数据的局限和数量的不足是制约自动识别系统算法开发和泛化应用的关键问题之一。针对此问题,目前可以采取的办法是根据现有的数据来进行数据增强或扩充,如裁剪、翻转、旋转、加入噪音点等,这些数据增强方式可以有有效的扩充训练样本数量。另外样本量较小时可以通过迁移学习来训练。

其次是疼痛评级的 PSPI 标准不足。如人实际上感到疼痛,但 PSPI 评分可能为零。此外,疼痛可能诱发的

不是 PSPI 评分中拥有模式 (AU4/6/7/9/10/43) 的面部反应. Kunz 等^[61]最近的研究表明有几个“痛苦的面孔”, 包含抬起眉毛 (AU1/2) 或张开嘴 (AU25/ 26/27), 这些都是 PSPI 不考虑的. 尽管观察到的受试者不感到疼痛, 但 PSPI 也可能不为零. 最明显的是, AU43(闭眼) 不是与疼痛特定相关, 例如它也发生在睡眠和放松期间. 此外, 情绪的几种面部表情与 PSPI 共享 AU, 例如, 厌恶 (AU9 或 10), 恐惧 (AU4), 悲伤 (AU4) 或幸福 (AU6). 需要更多的研究来找到更好的疼痛测量方法. 不同人由于生活经历文化背景等的差异, 对于同一疼痛刺激, 表现出不同的面部表情. 一些人对于疼痛刺激表现出的面部活动非常少, 另外有研究显示对于相同的刺激, 男性和女性表现出不同程度的疼痛表现. 因此, 从面部线索自动估计自我报告的疼痛强度的方法还应该考虑个体在疼痛的面部表现力方面的差异.

另外, 深度学习是潜力很大的智能算法, 其在疼痛自动识别中的应用价值亟待开发. 然而深度学习模型的巨大计算量使得资源消耗和计算时间增加很多. 近两年很多研究提出对网络结构进行轻量化, 通过改变卷积方式, 在不减少性能或稍微损失性能的前提下减少参数量, 从而减小内存和运算量、加快运算速度. 目前提出的网络结构有 SqueezeNet、MobileNet、ShuffleNet、Xception 等^[116-123]. 一般而言, 我们的计算资源都是非常有限的, 深度网络的轻量化可在节约计算资源的前提下获得有效的任务性能, 将会引起越来越多的关注.

5 结论

本文从疼痛自动识别系统的结构组成, 以及疼痛模态表征多角度方面进行描述, 对疼痛自动识别系统的流程环节和主要技术方法进行了系统性的梳理和总结. 疼痛检测包含多样化的评估指标, 例如面部表情、身体运动、生理信号等, 其中面部表情是疼痛最常见且最重要的评价指标, 语音信息也是具有潜在应用价值的表达模态. 由于疼痛表现形式复杂多样, 多模态特征融合技术的应用成为趋势, 模态特征信息提取、时空域特征融合以及多模态信息组合方式是重点研究的方向. 从众多文献的实验结果可以看出, 即使在可控条件下的疼痛数据库上, 有关疼痛检测任务的实验结果也不乐观, 距离实际应用的精度要求还有很大差距. 因此需要提取更合适、更具有特异性的模态特征, 并通过多特征、多模态的多角度综合分析, 促进疼痛自动

识别系统的性能.

参考文献

- 1 Prkachin KM. Assessing pain by facial expression: Facial expression as nexus. *Pain Research and Management*, 2009, 14(1): 53–58. [doi: [10.1155/2009/542964](https://doi.org/10.1155/2009/542964)]
- 2 Bahreini M, Jalili M, Moradi-Lakeh M. A comparison of three self-report pain scales in adults with acute pain. *The Journal of Emergency Medicine*, 2015, 48(1): 10–18. [doi: [10.1016/j.jemermed.2014.07.039](https://doi.org/10.1016/j.jemermed.2014.07.039)]
- 3 Hjermstad MJ, Fayers PM, Haugen DF, *et al.* Studies comparing numerical rating scales, verbal rating scales, and visual analogue scales for assessment of pain intensity in adults: A systematic literature review. *Journal of Pain and Symptom Management*, 2011, 41(6): 1073–1093. [doi: [10.1016/j.jpainsymman.2010.08.016](https://doi.org/10.1016/j.jpainsymman.2010.08.016)]
- 4 Chanques G, Viel E, Constantin JM, *et al.* The measurement of pain in intensive care unit: Comparison of 5 self-report intensity scales. *Pain*, 2010, 151(3): 711–721. [doi: [10.1016/j.pain.2010.08.039](https://doi.org/10.1016/j.pain.2010.08.039)]
- 5 Schnakers C, Chatelle C, Majerus S, *et al.* Assessment and detection of pain in noncommunicative severely brain-injured patients. *Expert Review of Neurotherapeutics*, 2010, 10(11): 1725–1731. [doi: [10.1586/ern.10.148](https://doi.org/10.1586/ern.10.148)]
- 6 Sikka K, Ahmed AA, Diaz D, *et al.* Automated assessment of children's postoperative pain using computer vision. *Pediatrics*, 2015, 136(1): e124–e131. [doi: [10.1542/peds.2015-0029](https://doi.org/10.1542/peds.2015-0029)]
- 7 Burns M, McIlfatrick S. Nurses' knowledge and attitudes towards pain assessment for people with dementia in a nursing home setting. *International Journal of Palliative Nursing*, 2015, 21(10): 479–487. [doi: [10.12968/ijpn.2015.21.10.479](https://doi.org/10.12968/ijpn.2015.21.10.479)]
- 8 Lucey P, Cohn JF, Prkachin KM, *et al.* Painful data: The UNBC-McMaster shoulder pain expression archive database. *Face and Gesture* 2011. Santa Barbara, CA, USA. 2011. 57–64.
- 9 Walter S, Gruss S, Ehleiter H, *et al.* The biovid heat pain database data for the advancement and systematic validation of an automated pain recognition system. *Proceedings of 2013 IEEE International Conference on Cybernetics (CYBCO)*. Lausanne, Switzerland. 2013. 128–131.
- 10 Haque MA, Nasrollahi K, Moeslund TB. Pain expression as a biometric: Why patients' self-reported pain doesn't match with the objectively measured pain? *Proceedings of 2017 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA)*. New Delhi, India. 2017. 1–8.
- 11 Hadjistavropoulos T, LaChapelle DL, MacLeod FK, *et al.*

- Measuring movement-exacerbated pain in cognitively impaired frail elders. *The Clinical Journal of Pain*, 2000, 16(1): 54–63. [doi: [10.1097/00002508-200003000-00009](https://doi.org/10.1097/00002508-200003000-00009)]
- 12 Prkachin KM, Schultz I, Berkowitz J, *et al.* Assessing pain behaviour of low-back pain patients in real time: Concurrent validity and examiner sensitivity. *Behaviour Research and Therapy*, 2002, 40(5): 595–607. [doi: [10.1016/S0005-7967\(01\)00075-4](https://doi.org/10.1016/S0005-7967(01)00075-4)]
- 13 Kunz M, Gruber A, Lautenbacher S. Sex differences in facial encoding of pain. *The Journal of Pain*, 2006, 7(12): 915–928. [doi: [10.1016/j.jpain.2006.04.012](https://doi.org/10.1016/j.jpain.2006.04.012)]
- 14 Kunz M, Prkachin K, Lautenbacher S. Smiling in pain: Explorations of its social motives. *Pain Research and Treatment*, 2013, 2013: 128093.
- 15 Keefe FJ, Lefebvre JC, Egert JR, *et al.* The relationship of gender to pain, pain behavior, and disability in osteoarthritis patients: The role of catastrophizing. *Pain*, 2000, 87(3): 325–334. [doi: [10.1016/S0304-3959\(00\)00296-7](https://doi.org/10.1016/S0304-3959(00)00296-7)]
- 16 Irani R, Nasrollahi K, Simon MO, *et al.* Spatiotemporal analysis of RGB-D-T facial images for multimodal pain level recognition. *Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops*. Boston, MA, USA. 2015. 88–95.
- 17 Chen JK, Chi ZR, Fu H. A new framework with multiple tasks for detecting and locating pain events in video. *Computer Vision and Image Understanding*, 2017, 155: 113–123. [doi: [10.1016/j.cviu.2016.11.003](https://doi.org/10.1016/j.cviu.2016.11.003)]
- 18 Bellantonio M, Haque MA, Rodriguez P, *et al.* Spatio-temporal pain recognition in CNN-based super-resolved facial images. *Proceedings of the 3rd International Workshop Video Analytics. Face and Facial Expression Recognition and Audience Measurement*. Cancun, Mexico. 2016. 151–162.
- 19 Haque MA, Irani R, Nasrollahi K, *et al.* Facial video-based detection of physical fatigue for maximal muscle activity. *IET Computer Vision*, 2016, 10(4): 323–329. [doi: [10.1049/iet-cvi.2015.0215](https://doi.org/10.1049/iet-cvi.2015.0215)]
- 20 Haque MA, Irani R, Nasrollahi K, *et al.* Heartbeat rate measurement from facial video. *IEEE Intelligent Systems*, 2016, 31(3): 40–48. [doi: [10.1109/MIS.2016.20](https://doi.org/10.1109/MIS.2016.20)]
- 21 Haque MA, Nasrollahi K, Moeslund TB. Real-time acquisition of high quality face sequences from an active pan-tilt-zoom camera. *Proceedings of the 10th IEEE International Conference on Advanced Video and Signal Based Surveillance*. Krakow, Poland. 2013. 443–448.
- 22 Haque MA, Nasrollahi K, Moeslund TB. Constructing facial expression log from video sequences using face quality assessment. *Proceedings of 2014 International Conference on Computer Vision Theory and Applications (VISAPP)*. Lisbon, Portugal. 2014. 517–525.
- 23 Haque MA, Nasrollahi K, Moeslund TB. Quality-aware estimation of facial landmarks in video sequences. *Proceedings of 2015 IEEE Winter Conference on Applications of Computer Vision*. Waikoloa, HI, USA. 2015. 678–685.
- 24 Haque MA, Bautista RB, Noroozi F, *et al.* Deep multimodal pain recognition: A database and comparison of spatio-temporal visual modalities. *Proceedings of the 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*. Xi'an, China. 2018. 250–257.
- 25 Werner P, Al-Hamadi A, Limbrecht-Ecklundt K, *et al.* Head movements and postures as pain behavior. *PLoS One*, 2018, 13(2): e0192767. [doi: [10.1371/journal.pone.0192767](https://doi.org/10.1371/journal.pone.0192767)]
- 26 Zamzmi G, Pai CY, Goldgof D, *et al.* Automated pain assessment in neonates. *Proceedings of the 20th Scandinavian Conference on Image Analysis*. Tromsø, Norway. 2017. 350–361.
- 27 Zamzmi G, Pai CY, Goldgof D, *et al.* An approach for automated multimodal analysis of infants' pain. *Proceedings of the 23rd International Conference on Pattern Recognition (ICPR)*. Cancun, Mexico. 2016. 4148–4153.
- 28 Mellor DJ, Stafford KJ. Physiological and behavioural assessment of pain in ruminants: Principles and caveats. *Alternatives to Laboratory Animals*, 2004, 32(S1): 267–271.
- 29 Aung MSH, Kaltwang S, Romera-Paredes B, *et al.* The automatic detection of chronic pain-related expression: Requirements, challenges and the multimodal EmoPain dataset. *IEEE Transactions on Affective Computing*, 2016, 7(4): 435–451. [doi: [10.1109/TAFFC.2015.2462830](https://doi.org/10.1109/TAFFC.2015.2462830)]
- 30 Vempada RR, Kumar BSA, Rao KS. Characterization of infant cries using spectral and prosodic features. *Proceedings of 2012 National Conference on Communications (NCC)*. Kharagpur, India. 2012. 1–5.
- 31 Pai CY. *Automatic Pain Assessment from Infants' Crying Sounds*[Master's thesis]. Tampa Bay: University of South Florida, 2016.
- 32 Kächele M, Thiam P, Amirian M, *et al.* Multimodal data fusion for person-independent, continuous estimation of pain intensity. *Proceedings of the 16th International Conference on Engineering Applications of Neural Networks*. Rhodes, Greece. 2015. 275–285.
- 33 Gruss S, Geiger M, Werner P, *et al.* Multi-modal signals for analyzing pain responses to thermal and electrical stimuli. *Journal of Visualized Experiments*, 2019, (146): e59057.
- 34 Hawker GA, Mian S, Kendzerska T, *et al.* Measures of adult pain: Visual analog scale for pain (VAS Pain),

- numeric rating scale for pain (NRS Pain), McGill pain questionnaire (MPQ), short - form McGill pain questionnaire (SF - MPQ), chronic pain grade scale (CPGS), short form - 36 bodily pain scale (SF - 36 BPS), and measure of intermittent and constant osteoarthritis pain (ICOAP). *Arthritis Care & Research*, 2011, 63(S11): S240-S252.
- 35 Taddio A, Hogan ME, Moyer P, *et al.* Evaluation of the reliability, validity and practicality of 3 measures of acute pain in infants undergoing immunization injections. *Vaccine*, 2011, 29(7): 1390-1394. [doi: [10.1016/j.vaccine.2010.12.051](https://doi.org/10.1016/j.vaccine.2010.12.051)]
- 36 Zamzmi G, Kasturi R, Goldgof D, *et al.* A review of automated pain assessment in infants: Features, classification tasks, and databases. *IEEE Reviews in Biomedical Engineering*, 2018, 11: 77-96. [doi: [10.1109/RBME.2017.2777907](https://doi.org/10.1109/RBME.2017.2777907)]
- 37 Ekman R, Rosenberg EL. *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)*. Oxford: Oxford University Press, 1997.
- 38 Prkachin KM, Solomon PE. The structure, reliability and validity of pain expression: Evidence from patients with shoulder pain. *Pain*, 2008, 139(2): 267-274. [doi: [10.1016/j.pain.2008.04.010](https://doi.org/10.1016/j.pain.2008.04.010)]
- 39 Grunau RVE, Craig KD. Pain expression in neonates: Facial action and cry. *Pain*, 1987, 28(3): 395-410. [doi: [10.1016/0304-3959\(87\)90073-X](https://doi.org/10.1016/0304-3959(87)90073-X)]
- 40 Ruder S. An overview of multi-task learning in deep neural networks. arXiv: 1706.05098, 2017.
- 41 Ghimire D, Lee J. Geometric feature-based facial expression recognition in image sequences using multi-class adaboost and support vector machines. *Sensors*, 2013, 13(6): 7714-7734. [doi: [10.3390/s130607714](https://doi.org/10.3390/s130607714)]
- 42 Li XL, Ruan QQ, Ming Y. 3D facial expression recognition based on basic geometric features. *Proceedings of the IEEE 10th International Conference on Signal Processing*. Beijing, China. 1366-1369.
- 43 Chao WL, Ding JJ, Liu JZ. Facial expression recognition based on improved local binary pattern and class-regularized locality preserving projection. *Signal Processing*, 2015, 117: 1-10. [doi: [10.1016/j.sigpro.2015.04.007](https://doi.org/10.1016/j.sigpro.2015.04.007)]
- 44 Huang XH, Wang SJ, Zhao GY, *et al.* Facial micro-expression recognition using spatiotemporal local binary pattern with integral projection. *Proceedings of 2015 IEEE International Conference on Computer Vision Workshop*. Santiago, Chile. 2015: 1-9.
- 45 Shyam R, Singh YN. Face recognition using augmented local binary pattern and Bray Curtis dissimilarity metric. *Proceedings of the 2nd International Conference on Signal Processing and Integrated Networks (SPIN)*. Noida, India. 2015. 779-784.
- 46 Nanni L, Brahnam S, Lumini A. A local approach based on a Local Binary Patterns variant texture descriptor for classifying pain states. *Expert Systems with Applications*, 2010, 37(12): 7888-7894. [doi: [10.1016/j.eswa.2010.04.048](https://doi.org/10.1016/j.eswa.2010.04.048)]
- 47 Li C, Zinger S, Ten WETA, *et al.* Video-based discomfort detection for infants using a constrained local model. *Proceedings of 2016 International Conference on Systems, Signals and Image Processing (IWSSIP)*. Bratislava, Slovakia. 2016. 1-4.
- 48 Chen JX, Liu XM, Tu P, *et al.* Person-specific expression recognition with transfer learning. *Proceedings of the 19th IEEE International Conference on Image Processing*. Orlando, FL, USA. 2012. 2621-2624.
- 49 Ashraf AB, Lucey S, Cohn JF, *et al.* The painful face-pain expression recognition using active appearance models. *Image and Vision Computing*, 2009, 27(12): 1788-1796. [doi: [10.1016/j.imavis.2009.05.007](https://doi.org/10.1016/j.imavis.2009.05.007)]
- 50 Edwards GJ, Cootes TF, Taylor CJ. Face recognition using active appearance models. *Proceedings of the 5th European Conference on Computer Vision*. Germany. 1998. 581-595.
- 51 Rupenga M, Vadapalli HB. Automatic spontaneous pain recognition using supervised classification learning algorithms. *Proceedings of 2016 Pattern Recognition Association of South Africa and Robotics and Mechatronics International Conference (PRASA-RobMech)*. Stellenbosch, South Africa. 2016. 1-6.
- 52 Kaltwang S, Rudovic O, Pantic M. Continuous pain intensity estimation from facial expressions. *Proceedings of the 8th International Symposium on Advances in Visual Computing*. Crete, Greece. 2012. 368-377.
- 53 Agrawal S, Khatri P. Facial expression detection techniques: Based on Viola and Jones algorithm and principal component analysis. *Proceedings of the 5th International Conference on Advanced Computing & Communication Technologies (ACCT)*. Haryana, India. 2015. 108-112.
- 54 Fotiadou E, Zinger S, Ten WETA, *et al.* Video-based facial discomfort analysis for infants. *Proceedings of SPIE 9029, Visual Information Processing and Communication V*. San Francisco, CA, USA. 2014. 90290F.
- 55 Lucey P, Cohn JF, Matthews I, *et al.* Automatically detecting pain in video through facial action units. *IEEE*

- Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 2011, 41(3): 664–674. [doi: [10.1109/TSMCB.2010.2082525](https://doi.org/10.1109/TSMCB.2010.2082525)]
- 56 Zhao JF, Mao X, Zhang J. Learning deep facial expression features from image and optical flow sequences using 3D CNN. *The Visual Computer*, 2018, 34(10): 1461–1475. [doi: [10.1007/s00371-018-1477-y](https://doi.org/10.1007/s00371-018-1477-y)]
- 57 Zamzami G, Ruiz G, Goldgof D, *et al.* Pain assessment in infants: Towards spotting pain expression based on infants' facial strain. *Proceedings of the 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*. Ljubljana, Slovenia. 2015. 1–5.
- 58 Wei XS, Zhou ZH. An empirical study on image bag generators for multi-instance learning. *Machine Learning*, 2016, 105(2): 155–198. [doi: [10.1007/s10994-016-5560-1](https://doi.org/10.1007/s10994-016-5560-1)]
- 59 Sikka K, Dhall A, Bartlett MS. Classification and weakly supervised pain localization using multiple segment representation. *Image and Vision Computing*, 2014, 32(10): 659–670. [doi: [10.1016/j.imavis.2014.02.008](https://doi.org/10.1016/j.imavis.2014.02.008)]
- 60 Ruiz A, Rudovic O, Binefa X, *et al.* Multi-instance dynamic ordinal random fields for weakly-supervised pain intensity estimation. *Proceedings of the 13th Asian Conference on Computer Vision*. Taipei, China. 2016. 171–186.
- 61 Hammal Z, Kunz M. Pain monitoring: A dynamic and context-sensitive system. *Pattern Recognition*, 2012, 45(4): 1265–1280. [doi: [10.1016/j.patcog.2011.09.014](https://doi.org/10.1016/j.patcog.2011.09.014)]
- 62 Khan RA, Meyer A, Konik H, *et al.* Pain detection through shape and appearance features. *Proceedings of 2013 IEEE International Conference on Multimedia and Expo (ICME)*. San Jose, CA, USA. 2013. 1–6.
- 63 Roy SD, Bhowmik MK, Saha P, *et al.* An approach for automatic pain detection through facial expression. *Procedia Computer Science*, 2016, 84: 99–106. [doi: [10.1016/j.procs.2016.04.072](https://doi.org/10.1016/j.procs.2016.04.072)]
- 64 Yang RJ, Tong SJ, Bordallo M, *et al.* On pain assessment from facial videos using spatio-temporal local descriptors. *Proceedings of the 6th International Conference on Image Processing Theory, Tools and Applications (IPTA)*. Oulu, Finland. 2016. 1–6.
- 65 Werner P, Al-Hamadi A, Limbrecht-Ecklundt K, *et al.* Automatic pain assessment with facial activity descriptors. *IEEE Transactions on Affective Computing*, 2017, 8(3): 286–299. [doi: [10.1109/TAFFC.2016.2537327](https://doi.org/10.1109/TAFFC.2016.2537327)]
- 66 Bourou D, Pampouchidou A, Tsiknakis M, *et al.* Video-based pain level assessment: Feature selection and inter-subject variability modeling. *Proceedings of the 41st International Conference on Telecommunications and Signal Processing (TSP)*. Athens, Greece. 2018. 1–6.
- 67 Zhi RC, Zamzami GZD, Goldgof D, *et al.* Automatic infants' pain assessment by dynamic facial representation: Effects of profile view, gestational age, gender, and race. *Journal of Clinical Medicine*, 2018, 7(7): 173. [doi: [10.3390/jcm7070173](https://doi.org/10.3390/jcm7070173)]
- 68 Schroff F, Kalenichenko D, Philbin J. FaceNet: A unified embedding for face recognition and clustering. *Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Boston, MA, USA. 2015. 815–823.
- 69 Parkhi OM, Vedaldi A, Zisserman A. Deep face recognition. *Proceedings of 2015 British Machine Vision Conference (BMVC)*. Swansea, UK. 2015. 6.
- 70 Zhang KP, Zhang ZP, Li ZF, *et al.* Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, 2016, 23(10): 1499–1503. [doi: [10.1109/LSP.2016.2603342](https://doi.org/10.1109/LSP.2016.2603342)]
- 71 Hermans A, Beyer L, Leibe B. In defense of the triplet loss for person re-identification. *arXiv: 1703.07737*, 2017.
- 72 Liu WY, Wen YD, Yu ZD, *et al.* SphereFace: Deep hypersphere embedding for face recognition. *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, HI, USA. 2017. 212–220.
- 73 Yang JL, Ren PR, Zhang DQ, *et al.* Neural aggregation network for video face recognition. *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, HI, USA. 2017: 4362–4371.
- 74 Huang ZW, Wang RP, Shan S, *et al.* Face recognition on large-scale video in the wild with hybrid Euclidean-and-Riemannian metric learning. *Pattern Recognition*, 2015, 48(10): 3113–3124. [doi: [10.1016/j.patcog.2015.03.011](https://doi.org/10.1016/j.patcog.2015.03.011)]
- 75 Li HX, Hua G. Hierarchical-PEP model for real-world face recognition. *Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Boston, MA, USA. 2015. 4055–4064.
- 76 Ofodile I, Kulkarni K, Corneanu CA, *et al.* Automatic recognition of deceptive facial expressions of emotion. *arXiv: 1707.04061*, 2017.
- 77 Yu ZD, Zhang C. Image based static facial expression recognition with multiple deep network learning. *Proceedings of 2015 ACM on International Conference on Multimodal Interaction*. Seattle, WA, USA. 2015. 435–442.
- 78 Kim BK, Roh J, Dong SY, *et al.* Hierarchical committee of deep convolutional neural networks for robust facial expression recognition. *Journal on Multimodal User Interfaces*, 2016, 10(2): 173–189. [doi: [10.1007/s12193-015-0209-0](https://doi.org/10.1007/s12193-015-0209-0)]
- 79 Zhang C, Zhang ZY. Improving multiview face detection with multi-task deep convolutional neural networks. *IEEE*

- Winter Conference on Applications of Computer Vision. Steamboat Springs, CO, USA. 2014. 1036–1041.
- 80 Sun B, Xu QH, He J, *et al.* Audio-video based multimodal emotion recognition using SVMs and deep learning. Proceedings of the 7th Chinese Conference on Pattern Recognition. Chengdu, China. 2016. 621–631.
- 81 Noroozi F, Marjanovic M, Njegus A, *et al.* Audio-visual emotion recognition in video clips. IEEE Transactions on Affective Computing, 2019, 10(1): 60–75. [doi: [10.1109/TAFFC.2017.2713783](https://doi.org/10.1109/TAFFC.2017.2713783)]
- 82 Zhou J, Hong XP, Su F, *et al.* Recurrent convolutional neural network regression for continuous pain intensity estimation in video. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Las Vegas, NV, USA. 2016. 84–92.
- 83 Wang F, Xiang X, Liu C, *et al.* Regularizing face verification nets for pain intensity regression. arXiv: 1702.06925, 2017.
- 84 Soar J, Bargshady G, Zhou XJ, *et al.* Deep learning model for detection of pain intensity from facial expression. Proceedings of the 16th International Conference on Smart Homes and Health Telematics, Designing a Better Future: Urban Assisted Living. Singapore. 2018. 249–254.
- 85 Wen YD, Zhang KP, Li ZF, *et al.* A discriminative feature learning approach for deep face recognition. Proceedings of the 14th European Conference on Computer Vision. Amsterdam, The Netherlands. 2016. 499–515.
- 86 Kharghanian R, Peiravi A, Moradi F. Pain detection from facial images using unsupervised feature learning approach. Proceedings of the 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Orlando, FL, USA. 2016. 419–422.
- 87 Martinez DL, Rudovic O, Picard R. Personalized automatic estimation of self-reported pain intensity from facial expressions. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Honolulu, HI, USA. 2017. 70–79.
- 88 Egede J, Valstar M, Martinez B. Fusing deep learned and hand-crafted features of appearance, shape, and dynamics for automatic pain estimation. Proceedings of the 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017). Washington, DC, USA. 2017. 689–696.
- 89 Rodriguez P, Cucurull G, González J, *et al.* Deep pain: Exploiting long short-term memory networks for facial expression classification. IEEE Transactions on Cybernetics, 2017. [doi: [10.1109/TCYB.2017.2662199](https://doi.org/10.1109/TCYB.2017.2662199)]
- 90 Chen ZL, Ansari R, Wilkie D. Automated pain detection from facial expressions using FACS: A review. arXiv: 1811.07988, 2018.
- 91 Bartlett MS, Littlewort GC, Frank MG, *et al.* Automatic decoding of facial movements reveals deceptive pain expressions. Current Biology, 2014, 24(7): 738–743. [doi: [10.1016/j.cub.2014.02.009](https://doi.org/10.1016/j.cub.2014.02.009)]
- 92 Olugbade TA, Bianchi-Berthouze N, Marquardt N, *et al.* Pain level recognition using kinematics and muscle activity for physical rehabilitation in chronic pain. Proceedings of 2015 International Conference on Affective Computing and Intelligent Interaction (ACII). Xi'an, China. 2015. 243–249.
- 93 Wang CY, Olugbade TA, Mathur A, *et al.* Automatic detection of protective behavior in chronic pain physical rehabilitation: A recurrent neural network approach. arXiv: 1902.08990, 2019.
- 94 Stahl A, Schellewald C, Stavadahl Ø, *et al.* An optical flow-based method to predict infantile cerebral palsy. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2012, 20(4): 605–614. [doi: [10.1109/TNSRE.2012.2195030](https://doi.org/10.1109/TNSRE.2012.2195030)]
- 95 Thiam P, Kessler V, Amirian M, *et al.* Multi-modal pain intensity recognition based on the SenseEmotion database. IEEE Transactions on Affective Computing, 2019. [doi: [10.1109/TAFFC.2019.2892090](https://doi.org/10.1109/TAFFC.2019.2892090)]
- 96 Abdulaziz Y, Ahmad SMS. Infant cry recognition system: A comparison of system performance based on mel frequency and linear prediction cepstral coefficients. Proceedings of 2010 International Conference on Information Retrieval & Knowledge Management (CAMP). Shah Alam, Selangor, Malaysia. 2010. 260–263.
- 97 Franck LS, Ridout D, Howard R, *et al.* A comparison of pain measures in newborn infants after cardiac surgery. Pain, 2011, 152(8): 1758–1765. [doi: [10.1016/j.pain.2011.03.017](https://doi.org/10.1016/j.pain.2011.03.017)]
- 98 Werner P, Al-Hamadi A, Walter S, *et al.* Automatic heart rate estimation from painful faces. Proceedings of 2014 IEEE International Conference on Image Processing (ICIP 2014). Paris, France. 2014. 1947–1951.
- 99 Werner P, Al-Hamadi A, Niese R, *et al.* Automatic pain recognition from video and biomedical signals. Proceedings of the 22nd International Conference on Pattern Recognition. Stockholm, Sweden. 2014. 4582–4587.
- 100 Walter S, Gruss S, Limbrecht-Ecklundt K, *et al.* Automatic pain quantification using autonomic parameters. Psychology & Neuroscience, 2014, 7(3): 363–380.
- 101 Kächele M, Amirian M, Thiam P, *et al.* Adaptive confidence learning for the personalization of pain intensity estimation systems. Evolving Systems, 2017, 8(1): 71–83.

- [doi: [10.1007/s12530-016-9158-4](https://doi.org/10.1007/s12530-016-9158-4)]
- 102 Panavaranan P, Wongsawat Y. EEG-based pain estimation via fuzzy logic and polynomial kernel support vector machine. The 6th 2013 Biomedical Engineering International Conference. Amphur Muang, Thailand. 2013. 13974380. [doi: [10.1109/BMEiCon.2013.6687668](https://doi.org/10.1109/BMEiCon.2013.6687668)]
- 103 Faye PM, De Jonckheere J, Logier R, *et al.* Newborn infant pain assessment using heart rate variability analysis. The Clinical Journal of Pain, 2010, 26(9): 777–782. [doi: [10.1097/AJP.0b013e3181ed1058](https://doi.org/10.1097/AJP.0b013e3181ed1058)]
- 104 Ranger M, Johnston CC, Rennick JE, *et al.* A multidimensional approach to pain assessment in critically ill infants during a painful procedure. The Clinical Journal of Pain, 2013, 29(7): 613–620. [doi: [10.1097/AJP.0b013e31826dfb13](https://doi.org/10.1097/AJP.0b013e31826dfb13)]
- 105 Wideman TH, Edwards RR, Walton DM, *et al.* The multimodal assessment model of pain: A novel framework for further integrating the subjective pain experience within research and practice. The Clinical Journal of Pain, 2019, 35(3): 212–221. [doi: [10.1097/AJP.0000000000000670](https://doi.org/10.1097/AJP.0000000000000670)]
- 106 Wagner J, Andre E, Lingensfelder F, *et al.* Exploring fusion methods for multimodal emotion recognition with missing data. IEEE Transactions on Affective Computing, 2011, 2(4): 206–218. [doi: [10.1109/T-AFFC.2011.12](https://doi.org/10.1109/T-AFFC.2011.12)]
- 107 Kächele M, Werner P, Al-Hamadi A, *et al.* Bio-visual fusion for person-independent recognition of pain intensity. Proceedings of the 12th International Workshop on Multiple Classifier Systems. Günzburg, Germany. 2015. 220–230.
- 108 Worley A, Fabrizi L, Boyd S, *et al.* Multi-modal pain measurements in infants. Journal of Neuroscience Methods, 2012, 205(2): 252–257. [doi: [10.1016/j.jneumeth.2012.01.009](https://doi.org/10.1016/j.jneumeth.2012.01.009)]
- 109 Velana M, Gruss S, Layher G, *et al.* The senseEmotion database: A multimodal database for the development and systematic validation of an automatic pain-and emotion-recognition system. Proceedings of the 4th IAPR TC 9 Workshop on Multimodal Pattern Recognition of Social Signals in Human-Computer-Interaction. Cancun, Mexico. 2016. 127–139.
- 110 Singh SK, Tiwari S, Abidi AI, *et al.* Prediction of pain intensity using multimedia data. Multimedia Tools and Applications, 2017, 76(18): 19317–19342. [doi: [10.1007/s11042-017-4718-6](https://doi.org/10.1007/s11042-017-4718-6)]
- 111 Rathee N, Ganotra D. A novel approach for pain intensity detection based on facial feature deformations. Journal of Visual Communication and Image Representation, 2015, 33: 247–254. [doi: [10.1016/j.jvcir.2015.09.007](https://doi.org/10.1016/j.jvcir.2015.09.007)]
- 112 Adibuzzaman M, Ostberg C, Ahamed S, *et al.* Assessment of pain using facial pictures taken with a smartphone. Proceedings of the IEEE 39th Annual Computer Software and Applications Conference. Taichung, China. 2015. 726–731.
- 113 Florea C, Florea L, Vertan C. Learning pain from emotion: Transferred hot data representation for pain intensity estimation. European Conference on Computer Vision. Zurich, Switzerland. 2014. 778–790.
- 114 Shah JH, Sharif M, Yasmin M, *et al.* Facial expressions classification and false label reduction using LDA and threefold SVM. Pattern Recognition Letters, 2017. [doi: [10.1016/j.patrec.2017.06.021](https://doi.org/10.1016/j.patrec.2017.06.021)]
- 115 Werner P, Al-Hamadi A, Niese R. Pain recognition and intensity rating based on comparative learning. Proceedings of the 19th IEEE International Conference on Image Processing. Orlando, FL, USA. 2012. 2313–2316.
- 116 Iandola FN, Han S, Moskewicz MW, *et al.* SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size. arXiv: 1602.07360, 2016.
- 117 Howard AG, Zhu ML, Chen B, *et al.* Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv: 1704.04861, 2017.
- 118 Sandler M, Howard A, Zhu MH, *et al.* Mobilenetv2: Inverted residuals and linear bottlenecks. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. 2018. 4510–4520.
- 119 Zhang XY, Zhou XY, Lin MX, *et al.* Shufflenet: An extremely efficient convolutional neural network for mobile devices. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. 2018. 6848–6856.
- 120 Ma NN, Zhang XY, Zheng HT, *et al.* ShuffleNet v2: Practical guidelines for efficient CNN architecture design. Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich, Germany. 2018. 116–131.
- 121 Chollet F. Xception: Deep learning with depthwise separable convolutions. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA. 2017. 1251–1258.
- 122 Zhang T, Qi GJ, Xiao B, *et al.* Interleaved group convolutions. Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy. 2017. 4373–4382.
- 123 Sun K, Li MJ, Liu D, *et al.* Igcv3: Interleaved low-rank group convolutions for efficient deep neural networks. arXiv: 1806.00178, 2018.