This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2019.2958341, IEEE Transactions on Pattern Analysis and Machine Intelligence

Automatic Detection of Pain from Facial Expressions: A Survey

Teena Hassan, Dominik Seuß, Johannes Wollenberg, Katharina Weitz, Miriam Kunz, Stefan Lautenbacher, Jens-Uwe Garbas, Ute Schmid

Abstract—Pain sensation is essential for survival, since it draws attention to physical threat to the body. Pain assessment is usually done through self-reports. However, self-assessment of pain is not available in the case of noncommunicative patients, and therefore, observer reports should be relied upon. Observer reports of pain could be prone to errors due to subjective biases of observers. Moreover, continuous monitoring by humans is impractical. Therefore, automatic pain detection technology could be deployed to assist human caregivers and complement their service, thereby improving the quality of pain management, especially for noncommunicative patients. Facial expressions are a reliable indicator of pain, and are used in all observer-based pain assessment tools. Following the advancements in automatic facial expression analysis, computer vision researchers have tried to use this technology for developing approaches for automatically detecting pain from facial expressions. This paper surveys the literature published in this field over the past decade, categorizes it, and identifies future research directions. The survey covers the pain datasets used in the reviewed literature, the learning tasks targeted by the approaches, the features extracted from images and image sequences to represent pain-related information, and finally, the machine learning methods used.

Index Terms—automatic pain detection, facial expressions of pain, pain datasets, pain feature representation, facial expression analysis, machine learning, survey.

1 Introduction

The International Association for the Study of Pain (IASP) [1, p. 209] defines pain as "an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage." Pain has the function of increasing attention, and initializing and maintaining mechanisms such as self-protection, recovery, and healing [2]. Without pain, human life would be significantly shorter [3]. The expression of pain triggers social reactions such as empathy, care, and nursing [2]. However, untreated pain is known to be a major contributor to reduced quality of life [4], to a progressive decline of functional and mental capacity [5], loss of appetite [6], reduced sleep [7], and behavioral disturbances including agitation, depression, and anxiety [8]. Therefore, timely detection and adequate treatment of pain is important.

Reliable assessment of pain is necessary for determining appropriate analgesics (pain-relieving medication) and their dosage. Self-reports or observational scales are used to assess pain. Self-reporting methods include rating scales (e.g., Visual Analogue Scale for Pain (VAS) [9], Numeric Rating Scale (NRS) [10]), pain diaries [11], or verbal descriptions (e.g., [12]). Patients who are noncommunicative due to a critical illness, narcotic medication, cognitive impairment, or infancy, cannot use self-reporting methods to communicate the pain they are experiencing. Therefore, assessment by other people, especially caregivers and nursing staff, is necessary. For this, different observational pain scales such as Behavioral Pain Scale (BPS) [13], Pain Assessment in Advanced Dementia (PAINAD) [14], or Neonatal Infant Pain Scale (NIPS) [15], are used in clinical settings. Facial expressions, body movements, and vocalizations are part of such observational pain scales. In research settings, other assessment tools are also used to study these observable dimensions of pain expression in greater detail. For example, the Prkachin-Solomon-Pain-Intensity (PSPI) scale [16] is used quite frequently by human coders (cf. [17]) as well as by computer scientists (cf. [18]) to annotate intensities of facial expressions of pain.

Pain assessment through observation is very challenging, and is affected by the subjective biases and errors in beliefs of the observer [19]. Studies such as [20] and [21] have found that pain is underestimated by nursing

- *T. Hassan is currently a Research Associate at Bielefeld University, and an external researcher at the Electronic Imaging Department, Fraunhofer IIS, 91058 Erlangen, Germany. The contributions in this paper are part of her research at Fraunhofer IIS.*
E-mail: teena.c.hassan@gmail.com
- *D. Seuß and J. Garbas are with the Electronic Imaging Department, Fraunhofer IIS, 91058 Erlangen, Germany.*
- *J. Wollenberg was with the Electronic Imaging Department, Fraunhofer IIS, 91058 Erlangen, Germany, at the time his contributions to the paper were created.*
- *K. Weitz is currently with the Department of Computer Science, University of Augsburg, Germany. She was with the Electronic Imaging Department, Fraunhofer IIS, 91058 Erlangen, Germany, at the time her contributions to the paper were created.*
- *M. Kunz is with the Faculty of Medical Sciences, University of Augsburg, Germany.*
- *S. Lautenbacher is with the Faculty of Human Sciences and Education, University of Bamberg, Germany.*
- *U. Schmid is with the Faculty of Information Systems and Applied Computer Science, University of Bamberg, Germany.*
E-mail: ute.schmid@uni-bamberg.de

Manuscript received Month dd, yyyy.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2019.2958341, IEEE Transactions on Pattern Analysis and Machine Intelligence

staff. In addition, it is not possible for human caregivers to continuously monitor a patient. All these factors lead to inappropriate pain management. Undertreatment of pain can be life-threatening for critically-ill patients [22]. Therefore, technical solutions to support caregivers and nursing staff to ensure continuous pain monitoring could promote better pain management. Automatic recognition of pain would enable individualised, patient-centered care, and help caregivers to provide timely and appropriate care to the pain felt by patients. By adding diagnostic decision explanation capabilities to such technical solutions (e.g. [23]), they could be used to train caregivers, medical practitioners, and nursing staff to improve their ability to correctly assess pain.

Facial expression is one of the valid indicators of pain [24] [25], and it appears in the observational scales for pain assessment. The advancements in the field of automatic facial image analysis inspired computer vision researchers to apply these techniques to detect pain from facial expressions. In [26], we discuss the challenges and present an interdisciplinary roadmap for developing a practically useful facial video-based pain detection system.

Apart from facial expressions, attempts have been made to use other modalities either individually or in combination for automatic pain detection. For example, Aung et al. [27] examined the use of body posture, body motion, and muscle activity for detecting patterns in body movements that could be indicative of pain during physical exercise; Tsai et al. [28] combined facial expressions and acoustic features to detect pain intensities in emergency cases; Werner et al. [29] investigated the use of facial expressions, electromyogram (EMG) recorded from trapezius muscles, and autonomic signals such as skin conductance and electrocardiogram (ECG) to automatically detect the different heat pain stimulus levels. Attempts have also been made to investigate the use of brain activation–acquired via either electroencephalography (EEG) [30] [31] or functional imaging [32]–for automatic pain assessment. However, brain activation based methods are often limited to experimental pain conditions, are very expensive (especially functional imaging), and require long and careful preparations (especially for EEG). Additionally, the amount of explained variance is often quite low [30].

The recording of brain activation or other physiological signals such as EMG, ECG, and skin conductance mostly requires sensors that are in contact with the body/skin[1], and this could become an additional cause of distress. In contrast, facial expressions of pain can be recorded in a contactless and nonintrusive manner. Facial expressions do not differ fundamentally between clinical and experimental pain [35]. The coding of facial expressions is well defined within the Facial Action Coding System (FACS) [36]. A similar, standard or widely used coding framework is not yet available for other modalities such as body movements[2]. Given the prominence of facial expressions in the assessment of pain across different age groups and health conditions (cf. [35], [15], [13], [14]), this survey focuses on automatic detection of pain from facial expressions.

1. Video-based, contactless measurement of physiological signals (cf. [33], [34]) could offer a promising alternative in the future.

2. Some efforts have been made to develop a coding system for body movements, posture, and muscle activity (cf. [37], [38], [27]).

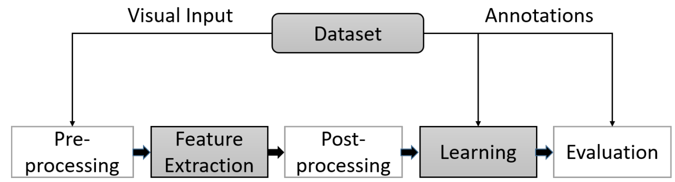

Fig. 1. General steps involved in developing an automatic pain detection system based on facial expressions. The boxes marked in gray highlight the elements that are covered in detail in this survey.

In this survey, we aim to review, consolidate, and structure the extensive work that has been done over the last decade to automatically detect pain from facial expressions, to identify the challenges, and define future research directions. In this work, the term "detection" is used more generally to cover both the detection of presence of pain and estimation of its intensity. This paper first surveys the *pain datasets* and then the automatic pain detection approaches. The general steps involved in automatic detection of pain from facial expression images or videos are shown in Figure 1. In this survey, we focus on the key elements that have close semantic relevance to pain, namely the *learning task*, *representation* (i.e. extracted features), and *learning method*. A survey of the methods used for input preprocessing and feature postprocessing (e.g. feature selection for dimensionality reduction) is not included. Due to the heterogeneity in the datasets, performance metrics, crossvalidation schemes, and training-validation-test splits used for performance evaluation, the results reported in the reviewed literature cannot be compared with each other. Therefore, a summary of the performance of the approaches is excluded from this survey.

The rest of the paper is organized as follows: Section 2 describes the methodology followed to collect and review the literature; Section 3 summarizes the findings from psychological studies on facial expressions associated with pain; Section 4 presents the results of the survey of datasets containing facial expressions of pain that have been used in the reviewed literature; Section 5 presents the results of the survey of automatic pain detection approaches based on facial expressions, with a special focus on the learning tasks, extracted features, and machine learning methods used; Section 6 discusses the open challenges and identifies future research directions necessary to address the challenges; Section 7 concludes the paper.

2 Review Methodology

Peer-reviewed papers were collected mainly by searching online digital libraries such as IEEE Xplore[3], ACM[4], and ScienceDirect[5]. Additionally, publication lists on the webpages of research groups known to be working on automatic pain/facial expression analysis were examined. Research projects related to automatic pain detection also served as a source for finding relevant papers. The Google Scholar search engine[6] was also used to search for relevant papers.

3. https://ieeexplore.ieee.org

4. https://www.acm.org

5. https://www.sciencedirect.com

6. https://scholar.google.de

The search keywords used included: 'automatic pain recognition', 'facial expressions of pain', 'automatic pain detection', and 'pain intensity estimation'. Only papers that used video or image-based facial expression information were considered[7]. In cases where the facial expression modality was combined with other modalities, only the processing of the facial expression modality was studied in detail.

The papers that were collected for review[8] appeared during the period from 2006 to 2018. The field of automatic pain detection from facial expressions started receiving attention since 2006, closely following the success of face detection [39] and automatic facial image analysis [40] [41]. Since 2015, the focus is shifting towards the use of deep learning methods for automatic pain detection. This has been prompted by the recent success of Convolutional Neural Networks (CNN) in image classification tasks [42] [43] in general, and in face analysis tasks [44] in specific.

The following items were focused on during the literature review:

- Objective of the paper (learning task);
- Dataset(s) used, along with the pain induction method, the demographics and health condition of participants, the size of the material, and the provided annotations;
- Visual input type (single images or image sequences);
- Feature representations extracted directly or indirectly from the visual input;
- Learning strategies (supervised, semi-supervised, weakly supervised, or unsupervised), and learning methods used.

3 Facial Expressions of Pain

Facial expression during the experience of pain is not unspecific grimacing, but conveys pain specific information. Studies investigating facial expressions of pain have most often used FACS [36], the gold-standard for facial expression research. FACS is a fine-grained, objective, and anatomically-based coding system that differentiates between 44 facial movements known as Action Units (AU). Coders are trained to apply specific operational criteria to determine the onset and offset as well as the intensity of the AUs. Using FACS, it was shown that facial expressions of pain are composed of a small subset of facial activities, namely lowering the brows (AU4), cheek raise/lid tightening (AUs 6_7), nose wrinkling/raising the upper lip (AUs 9_10), opening the mouth (AUs 25_26_27), and eye closure longer than 0.5s (AU 43) [35] [45] [16]. These facial activities are displayed during the experience of experimental pain as well as in clinical pain conditions, and seem to be largely inborn [46] [2]. Nevertheless, this does not mean that there is only one uniform facial expression of pain that can be observed at all times and in each individual [47]. Rather, individuals often display only parts of this subset or combine this subset of facial activities differently. Using cluster analyses, it was shown that facial expressions of pain can indeed be clustered into four distinct facial activity patterns of pain [48]. Besides a stoic expression, the most stable patterns are: "narrowed eyes" combined with either (I) "raising the upper lip/nose wrinkling" and "furrowed brows"; (II) "furrowed brows" or (III) "opening of the mouth". The fourth cluster "raised eyebrows" was less stable and less frequent, and in a recent review article [35] on facial expressions of pain, it was found that this type of response occurs more frequently in response to experimentally induced pain and could reflect a kind of novelty/surprise response. These different facial activity patterns seem to represent behavioral synonyms for the internal state "pain". Training individuals to recognize these distinct facial activity patterns of pain was shown to improve recognition of pain significantly compared to only focusing on one prototypical expression of pain [49]. Thus, embracing the idea of some variability in facial expressions of pain holds the potential to improve the communication of pain.

4 Datasets of Facial Expressions of Pain

Data forms the basis for developing machine learning models for automatic pain detection. Tables 1 and 2 summarize the datasets that were used in the reviewed literature. The methods used to induce pain, the demographics of the participants, the available annotations, and the size of the available video or image material, were examined closely.

Most of the works on automatic pain detection used one or more of the publicly available datasets listed in Table 1. In the remaining works, the researchers collected their own pain datasets. These are listed in Table 2. The availability of these datasets for research purposes is not clearly known. It can be noted from both these tables that the datasets contain the visual material for facial expressions of pain in the form of either *single images* (e.g. [50], [51]), *videos* (e.g. [52], [53]), or *sequence of images* extracted from videos or video clips (e.g. [18], [54]). The UNBC McMaster Shoulder Pain Archive Database [18] is the most widely used publicly available dataset for automatic pain detection from facial expressions. The BioVid[9] Heat Pain Database [52] is yet another large publicly available database consisting of facial expressions and physiological signals recorded during administration of painful heat stimuli. The Infant COPE dataset [50] [55] is a relatively small dataset containing images of facial expressions of neonates experiencing pain during heel lancing.

It is clear from Tables 1 and 2 that a variety of pain inducing methods have been used for creating pain datasets by different researchers. Under laboratory settings, acute pain was usually induced using cold stimuli (cf. [56]), heat stimuli (cf. [52], [57]), or mechanical pressure (cf. [53]). In datasets where participants were already suffering from pain, physical movements (cf. [18], [27]), activities of daily living (cf. [58]), or manual pressure (cf. [59], [60]) were used to induce acute pain.

It can also be seen from Tables 1 and 2 that different annotation methods were used in the different datasets for describing the facial expressions and the experienced pain or emotion. In general, self-reports, observer reports, stimulus type, stimulus level, or AU based scores were used to annotate pain. Annotations were done either at *frame level*,

7. The use of facial EMG information is not reviewed in this survey.
8. We have included all relevant papers, to the best of our knowledge.
9. http://www.iikt.ovgu.de/BioVid.print

sequence level, or *segment level*. Sequence-level annotation refers to annotation given to the entire video. Segment-level annotation refers to annotation of a chosen snippet or session within the video. It is noted that self-reports were provided at sequence or segment level. AU-based annotation of facial expressions was provided at frame level or segment level. Some datasets provided multiple forms of annotations. For example, the UNBC McMaster Shoulder Pain Archive Database [18] provides frame-level AU annotations as well as sequence-level self and observer reports.

Pain management is challenging, especially in non-communicative subjects [61] [62]. The ageing population [63] and forecasts of increasing incidence of dementia in the coming years [64], raise the need for investigating pain in older adults, and developing and testing automatic pain detection systems specifically for the older old. Facial expression analysis systems that are developed for a young age-group would not generalise well to older age-groups [65]. However, very few datasets [53] [66] have included participants above 67 years of age.

Based on the analysis of the existing datasets and interdisciplinary consultations between psychologists and computer scientists, the following recommendations are made:

- We need datasets covering a larger age-range, including also the oldest old.
- Since pain is often confused with negative emotions [67], we need more datasets with genuine pain and emotions as control condition in order to develop reliable pain detection systems.
- Since pain is experienced not only during rest, more datasets that are ecologically more valid are needed. That is, participants should also be filmed while in motion and not only while sitting or lying down.
- In order to provide a good comparison between manual and automatic pain detection, the datasets should always provide manual FACS codes as well as self-report or observer-report (especially in cases where self-report is not possible).
- For the development, validation, and benchmarking of different automatic pain detection approaches, it is important to have annotations at a granularity finer than the sequence level, along with a precise temporal alignment of these annotations with the video.

5 Automatic Pain Detection

This section summarizes the results of the survey of automatic pain detection approaches based on facial expressions. It is organized into three subsections. Subsection 5.1 provides an overview of the approaches by categorizing them based on the learning task. Subsection 5.2 consolidates the features used for learning, and Subsection 5.3 lists and categorizes the different machine learning strategies and methods used.

5.1 Overview of Approaches

We categorize the approaches adopted by the research community for automatic pain recognition from facial expressions into *one-step* and *two-step* approaches. The one-step approaches predict pain or pain intensity based on geometric, textural and/or temporal features extracted directly from the input image or image sequence. The two-step approaches use or require an intermediate learning stage for describing the facial expression in terms of AUs or AU intensities. While the one-step approaches correspond to the classical way of learning the target from the input features, the two-step approaches are motivated by the way in which human observers detect and code pain [74]–on the basis of specific facial expression elements (cf. [13]).

Table 3 summarizes the approaches used in the literature reviewed in this paper. As can be seen, (i) most works employed a one-step approach; (ii) the predominant learning task was the detection of presence or absence of pain in single images; (iii) other learning tasks include distinguishing pain from other emotions or states, detecting discrete pain intensity levels, estimating continuous-valued pain intensities, distinguishing genuine pain from posed expression of pain, and detecting/localizing pain events in image sequences; (iv) very few works have so far investigated the task of distinguishing pain from other emotions.

It can also be seen that many of the approaches were developed for single images, and did not consider temporal information about pain. In Table 3, temporal information is considered to be included when the temporal dimension is considered for feature extraction (cf. [75]) or when dynamic models, for example the latent-dynamic conditional random field method in [76], are used for learning. Temporal information is also considered to be included when frame-level features are aggregated–for example, the statistical features in [59]–or when sequence-level events are considered for the corresponding learning task(s)–for example, the sequence-level AU events in [77] and [78].

5.2 Feature Extraction

Features are extracted from facial images and image sequences to describe the facial shape and appearance, or their changes, that are caused by facial expressions. The terminology used in [132] for categorizing feature representations for facial affect analysis has been generally adopted in this survey. Tables 4 and 5 summarize the features that have been extracted directly from single images or image sequences for automatic detection of pain in the reviewed literature[10]. These features are mainly categorised into *spatial* and *spatiotemporal* features. Spatial features provide a static description of what is visible in an image: facial shape and facial texture. Spatiotemporal features encode the changes in facial shape and appearance that are visible over time in a sequence of images. In one-step approaches, spatial or spatiotemporal features were used for pain detection. In two-step approaches, these features were used for AU detection. It can be noted from Tables 4 and 5 that spatial features were most widely used for pain or AU detection. A good number of works extracted spatiotemporal features from the visual input, and a few works (e.g. [75], [116]) used a *mixture* of spatial and spatiotemporal features for the detection task.

10. References, in which the extracted features are not stated clearly, have been excluded from Tables 4 and 5.

TABLE 1
Summary of datasets containing facial expressions of pain that are available upon request via email to first author or through a website.

Reference	Diagnostic Status	Pain Stimulus	Demographics	Sample Size	Annotation Granularity	Annotation (Labels)
Infant COPE [50] [55]	healthy	heel lancing	26 neonates (18–72 hours); 13 male, 13 female; Caucasian	204 facial images (pain images: 60)	frame-level	pain, crying, heel friction, nasal air stimulus, rest
UNBC-McMaster Shoulder Pain Expression Archive Database [18]	shoulder pain	range of motion tests on shoulders	129 adults; 63 male, 66 female	200 image sequences (total frames: 48,398; pain frames: 8369)	frame-level	12 AUs and their intensities (A–E), 66 facial landmarks, PSPI score
					sequence-level	self-report via VAS, sensory scale, affective-motivational scale; observer report via Observer Rated Pain Intensity (OPI)
BioVid Heat Pain Database [52]	healthy	heat	Part A: 87 adults (18–65 years); 44 male, 43 female	8700 videos (pain videos: 6960)	sequence-level	baseline (no pain), 4 pain stimulus intensity levels
			Part B: 86 adults (18–65 years); 42 male, 44 female	8600 videos (pain videos: 6880; partial facial occlusion due to facial EMG electrodes)	sequence-level	baseline (no pain), 4 pain stimulus intensity levels
			Part C: 87 adults (18–65 years); 44 male, 43 female	87 videos (long version of Part A with one video per subject)	segment-level	pain stimulus
		case vignette	Part D: 90 adults (18–65 years); 45 male, 45 female	630 videos (posed pain videos: 90)	sequence-level	7 posed expressions: neutral, pain, anger, disgust, fear, happy, sad
Hi4D-ADSIP [68]	healthy	none	80 adults (18–60 years); 32 male, 48 female; diverse ethnicities	3360 3D sequences (pain sequences: 240)	sequence-level	posed pain, 6 posed emotions (anger, disgust, surprise, fear, sadness, happiness), and 7 other facial articulations at 3 intensity levels each: mild, normal, extreme
BP4D-Spontaneous [69]	healthy	cold	41 adults (18–29 years); 18 male, 23 female; 20 Euro-American, 11 Asian, 6 African-American, 4 Hispanic	328 2D and 328 3D videos (pain videos: 41 2D, 41 3D)	frame-level	27 AUs (max. 20 sec segments), 3D head rotation angles, facial landmarks (83 for 3D, 49 for 2D)
					sequence-level	pain, anger, startle, fear, sadness, disgust, embarrassment, happiness

Spatial features consist of *geometric* or *textural* features. Geometric features describe the shape of the face in terms of point-based shape description schemes. These define point placements on facial features such as eyes, eyebrows, cheek, nose, lips, chin, and/or facial boundary. The locations of these facial feature points or higher-order features such as distances and angles between the facial feature points, are used as geometric features. Textural features describe the appearance of the face and facial features. Textural features include a description of the edges of facial features, and the wrinkles or folds that appear on or around them. Textural features used in literature range from raw pixel intensities to hand-crafted or self-learned features. The commonly used hand-crafted textural feature descriptors are Gabor filters [133], Local Binary Patterns (LBP) [134] [135], and Histogram of Oriented Gradients (HOG) [136]. Geometric features were rarely used alone. Textural features, either alone or in combination with geometric features, are the most widely used features in automatic pain detection. The combination of geometric and textural features are denoted in Tables 4 and 5 as *hybrid* features.

Spatiotemporal features describe changes in spatial features over time, and can be categorized in a similar fashion as spatial features into geometric, textural, and hybrid categories. In other words, spatiotemporal geometric and textural features were used either independently or in combination (hybrid). Geometric features extracted from a sequence of images were summarized using mathematical and statistical operators. Spatiotemporal textural features such as LBP-TOP [137] and HOG-TOP [138] were extracted from Three Orthogonal Planes (TOP), one of which covers the temporal dimension that spans a temporally ordered sequence of images. Yang et al. [104] compared the performance of several spatiotemporal textural features, such as LBP-TOP, LPQ-TOP[11], BSIF-TOP[12], and their combinations.

In cases where hybrid, mixed, or multiple features of the same type were used, the fusion of features was performed either before the learning step (cf. [75], [123]), or by fusing the decisions of classifiers trained separately for each feature (cf. [107], [116]). The former is commonly referred to as "early fusion", and the latter is commonly referred to as "late fusion".

In two-step approaches, pain detection is done based on an intermediate representation of the face in terms of

11. LPQ-TOP is the spatiotemporal variant of Local Phase Quantization (LPQ) [139]
12. BSIF-TOP is the spatiotemporal variant of Binarized Statistical Image Features (BSIF) [140]

TABLE 2
Summary of datasets whose availability is unknown at the time of writing this survey. The authors of the respective papers might be contacted for potential access. Note: 'elderly' denotes adults aged over 65 years.

Reference	**Diagnostic Status**	**Pain Stimulus**	**Demographics**	**Sample Size**	**Annotation Granularity**	**Annotation (Labels)**
Wilkie [58]	lung cancer	Activities of Daily Living (ADL)	43 adults; 27 male, 16 female; 31 Caucasian, 12 others	43 videos	segment-level	9 AUs (20 sec segments)
					sequence-level	self-report via VAS and State-Trait Anxiety Inventory (STAI)
Roy et al. [70]	healthy	none	34 adults	1088 videos (pain videos: 136)	sequence-level	pain, neutral, 6 basic emotions (anger, sadness, happiness, disgust, fear, surprise)
Kunz et al. [53]	demented	mechanical pressure	42 elderly (mean: 76.7 years); 20 male, 22 female	42 videos (pain stimulus sessions: 840)	segment-level	for each 5 sec stimulus session: 44 AUs and their intensities (A–E), self-report of pain level via verbal category scale
	healthy		54 elderly (mean: 74.2 years); 11 male, 43 female	54 videos (pain stimulus sessions: 1080)		
Lu et al. [51]	healthy	heel lancing	57 neonates; 30 male, 27 female	510 images (pain images: 160)	frame-level	pain, cry, calm
Hammal et al. [71]	healthy	heat	20 adults	20 videos (pain stimulus sessions: 40)	segment-level	for each 5 sec stimulus session: 44 AUs and their intensities (A–E)
Kunz et al. [57]	healthy	heat	44 young adults (18–30 years); 22 male, 22 female	44 videos (pain stimulus sessions: 352)	segment-level	for each 5 sec segment after stimulus reached peak: 44 AUs and their intensities (A–E), self-report via VAS
Littlewort et al. [56] [72]	healthy	cold	26 adults; 6 male, 20 female	78 one-minute videos (real pain videos: 26; faked pain videos: 26)	sequence-level	baseline (no pain), real pain, faked pain
Niese et al. [54]	healthy	hand movements with tourniquet attached	21 adults (20–30 years); 10 male, 11 female	21 image sequences (total frames: 966000; pain frames: 31500)	segment-level	self-report of pain intensity via NRS
EmoPain [27]	chronic lower back pain	physical exercises	22 adults (19–67 years); 7 male, 15 female; 18 Caucasian, 4 others	44 videos (total frames: 585,487; pain frames: 50,071)	frame-level	pain, no pain
					segment-level	self-report of pain and anxiety on 1–10 scale
	healthy		28 adults (mean age: 37.1 years); 14 male, 14 female; 26 Caucasian, 2 Asian	–	–	no pain
Irani et al. [66]	healthy	mechanical pressure	12 elderly females (66–90 years)	96 videos (total frames: 2388; pain frames: 1631)	sequence-level	self-report of pain intensity via NRS
Pediatric Pain Dataset [59] [73]	after appendectomy	endogenous and exogenous (manual pressure at surgical site)	50 youth (5–18 years); 27 male, 23 female; 35 Hispanic, 9 non-Hispanic white, 5 Asian, 1 Native American	300 videos (endogenous pain: 150 exogenous pain: 150)	sequence-level	self and observer reports of pain intensity via NRS
Singh [60]	back/neck/knee pain	manual pressure on affected area	21 adults; 12 male, 9 female	21 image sequences (total frames: 336)	frame-level	7 AUs and their intensities
Tsai et al. [28]	emergency cases with pain or headache	endogenous	117 adults	205 videos	sequence-level	self-report via NRS

TABLE 3
Summary of the learning approaches that have been developed and tested for automatic pain detection from facial expressions.

Learning Task	Temporal Information	References
One-Step Approaches		
pain and no-pain	no	Brahnam et al. [79], Monwar and Rezaei [80], Brahnam et al. [81], Lu et al. [51], Ashraf et al. [82], Lucey et al. [83], Siebers et al. [84], Nanni et al. [85], Gholami et al. [86], Monwar and Rezaei [87], Wei and Li-min [88], Lucey et al. [18], Lucey et al. [89], Werner et al. [90], Chen et al. [91], Khan et al. [92], Pedersen [93], Neshov and Manolova [94], Rathee and Ganotra [95], Aung et al. [27], Kharghanian et al. [96], Roy et al. [97], Rupenga and Vadapalli [98], Meawad et al. [99], Alphonse and Dharma [100]
	yes	Werner et al. [101], Meng and Bianchi-Berthouze [102], Werner et al. [29], Kächele et al. [103], Yang et al. [104]
pain and emotions	no	Niese et al. [54]
	yes	Hammal et al. [71], Hammal and Kunz [105]
pain and states (crying, calm/rest)	no	Brahnam et al. [79], Lu et al. [51], Yuan et al. [106]
pain and distress (via heel friction or air stimulus on nose)	no	Brahnam et al. [79]
pain intensity (continuous)	no	Werner et al. [90], Kaltwang et al. [107], Romera-Paredes et al. [108], Neshov and Manolova [94], Wang et al. [109], Liu et al. [110]
	yes	Kächele et al. [103], Florea et al. [111], Zhou et al. [112], Kaltwang et al. [113], Zhao et al. [114], Rodriguez et al. [115], Egede et al. [116], Egede and Valstar [117], Lopez-Martinez et al. [118], Tavakolian and Hadid [119]
pain intensity (discrete)	no	Gholami et al. [86], Lucey et al. [89], Hammal and Cohn [120], Singh [60], Rathee and Ganotra [95], Roy et al. [97], Alphonse and Dharma [100]
	yes	Rudovic et al. [121], Irani et al. [122], Irani et al. [66], Werner et al. [123], Tsai et al. [28], Lopez-Martinez et al. [118]
pain event in sequence	yes	with localization: Sikka et al. [124], Sikka et al. [125], Lo Presti and La Cascia [126], Lo Presti and La Cascia [127]
		without localization: Chen et al. [75]
Two-Step Approaches		
pain and no-pain	no	Lucey et al. [83], Lucey et al. [128], Zafar and Khan [129]
	yes	Schmid et al. [77], Sikka et al. [59], Siebers et al. [78]
pain intensity (continuous)	yes	Sikka [73], Sikka et al. [59], Zhang et al. [76], Lopez-Martinez et al. [130]
pain intensity (discrete)	no	Zafar and Khan [129]
	yes	Ghasemi et al. [74]
posed and genuine pain	yes	Littlewort et al. [72], Littlewort et al. [56], Bartlett et al. [131]

AUs. Features used for learning pain-related targets were therefore extracted from the AU labels or AU scores[13] provided by the first step. We categorize these features that are indirect representations of the input image or image sequence into *non-temporal* and *temporal* features. Table 6 provides an overview of the indirect features that have been used for automatic pain detection. Non-temporal features refer to the AU representations for a single image or a *single timestep* in an image sequence. In this case, the AU labels or scores for the image are used as features for pain detection (e.g. [77], [78], [83]). Temporal features refer to AU representations for a sequence of images spanning *multiple timesteps*. In this case, AU scores provided by the first learning stage are aggregated using statistical operators (cf. [59]) or dynamic features are extracted using temporal filters (cf. [131]). Note that the categorization into non-temporal and temporal features is based purely on whether the pain detection in the second step used AU detection outputs for a single image/timestep or for multiple timesteps. It does not take into account whether temporal information was used in the first step for AU detection. It was noted that the two-step approach followed by Lopez-Martinez et al. [130] used a combination of direct and indirect features for continuous pain intensity estimation (see Table 6).

The extracted features are often post-processed to increase their discriminative power or to extract the most important information. Principal Component Analysis (PCA) is a commonly used method to select the most important feature dimensions and thereby transform the features into a lower-dimensional space (cf. [76], [79], [114]). Rathee and Ganotra [95] proposed multiview distance metric learning to fuse LBP, HOG, and Gabor features, and to increase the discriminative power of the new set of features. Florea et al. [111] used a semi-supervised transfer learning method based on spectral regression to learn the most discriminative feature dimensions of the extracted Histogram of Topological (HoT) features and to reduce the dimensionality of the feature space. An exhaustive survey of the feature post-processing methods is outside the scope of this paper. The reader is advised to refer to other surveys on facial expression analysis (e.g. [132]) to obtain an overview about

13. The term “scores” is used in this paper to broadly refer to scores/probabilities/intensities of AUs.

TABLE 4
Summary of spatial representations extracted directly from facial images for automatic pain detection.

Feature Sub-Type	Features	References
geometric	facial landmark positions	Meng and Bianchi-Berthouze [102], Ghasemi et al. [74], Aung et al. [27], Rupenga and Vadapalli [98], Liu et al. [110], Lopez-Martinez et al. [118]
	facial landmark distances	Romera-Paredes et al. [108], Meawad et al. [99]
	facial landmark distances and angles	Niese et al. [54], Siebers et al. [84]
	facial landmark positions, distances, angles	Zafar and Khan [129]
textural	pixel intensities	Brahnam et al. [79], Gholami et al. [86], Ghasemi et al. [74]
	Gabor filters	Littlewort et al. [72], Lu et al. [51], Yuan et al. [106], Littlewort et al. [56], Sikka [73], Bartlett et al. [131], Sikka et al. [59], Roy et al. [97]
	Discrete Cosine Transform (DCT)	Brahnam et al. [81], Aung et al. [27]
	Local Binary Pattern (LBP) or its variant	Nanni et al. [85], Chen et al. [91], Rudovic et al. [121], Aung et al. [27]
	Local Ternary Pattern (LTP) or its variant	Nanni et al. [85]
	histogram of quantized edge directions	Monwar and Rezaei [80]
	Histogram of Oriented Gradients (HOG) around facial landmarks	Chen et al. [75]
	Histogram of Topological (HoT) features	Florea et al. [111]
	variants of Local Directional Pattern (LDP)	Alphonse and Dharma [100]
	Scale-Invariant Feature Transform (SIFT)	Neshov and Manolova [94], Singh [60]
	Speeded-Up Robust Features (SURF)	Singh [60]
	pyramid HOG and pyramid LBP	Khan et al. [92]
	supervised locality preserving projection	Wei and Li-min [88]
	log-normal filters	Hammal and Cohn [120]
	Gabor filters, HOG, and LBP	Rathee and Ganotra [95]
	3D binary edges	Zhang et al. [76]
	features learned by semi-supervised autoencoder	Pedersen [93]
	deep learned features	Kharghanian et al. [96], Rodriguez et al. [115], Wang et al. [109]
hybrid (geometric + textural)	facial landmark distances, nasal root wrinkles, context variable	Hammal et al. [71], Hammal and Kunz [105]
	facial landmark distances, histogram of quantized edge directions	Monwar and Rezaei [87]
	facial landmark positions, LBP, Gabor filters	Zhao et al. [114]
	facial landmark distances, mean gradient magnitude in facial regions	Werner et al. [90]
	facial landmark positions, DCT, LBP	Kaltwang et al. [107]
	facial landmark positions, pixel intensities	Ashraf et al. [82], Lucey et al. [128], Lucey et al. [18], Lucey et al. [89]
	facial landmark positions, DCT	Lucey et al. [83]
	facial landmark positions, distances, angles, HOG	Egede et al. [116]

commonly used feature post-processing or feature selection methods.

5.3 Learning Methods

The different learning tasks examined by the existing automatic pain detection approaches were discussed in Section 5.1 and listed in Table 3. The majority of the learning tasks were either binary or multiclass classification tasks such as pain versus no-pain, genuine versus posed pain, discrete pain intensity levels, and pain versus emotions. The other type of learning tasks were regression tasks for estimating pain intensity as a continuous-valued function. Table 7 lists the machine learning methods that have been used in the reviewed literature for pain-related classification and regression tasks. Two-step approaches involve two learning tasks: AU detection and pain detection. AU detection tasks include binary or multiclass classification for detecting the presence of different AUs, and regression for estimating intensities of different AUs. The machine learning methods used for AU detection in two-step approaches are also listed in Table 7. Certain one-step approaches did not use machine learning methods for pain detection. Irani et al. [122] [66] used experimentally determined thresholds on spatiotemporal features to determine three discrete levels of pain. Meawad et al [99] defined a mapping between facial landmark distances and pain-related AUs. Based on this mapping, the PSPI scale was modified. Sequence-level pain detection was then performed by checking whether a predefined number of consecutive frames showed the presence of pain according to the modified PSPI scale. Certain two-step approaches did not use machine learning methods in the second step. Zafar and Khan [129] applied the PSPI scale on the discrete AU intensities predicted by a set of k-nearest neighbor classifiers. Zhang et al. [76] averaged the probabilities of selected pain-related AUs to calculate the pain intensity estimate.

TABLE 5
Summary of spatiotemporal and mixed representations extracted directly from facial images for automatic pain detection.

Feature Sub-Type	Features	References
Spatiotemporal Features		
geometric	Hankel matrices based on facial landmark positions and/or distances	Lo Presti and La Cascia [126]
	statistical features from sequence of facial landmark distances and quadratic polynomial coefficients of mouth shape	Tsai et al. [28]
	bag of words from k-means based clusters of sequence of geometric features (facial landmark distances and quadratic polynomial coefficients of mouth shape)	Tsai et al. [28]
textural	HOG from Three Orthogonal Planes (HOG-TOP)	Chen et al. [75]
	LBP from Three Orthogonal Planes (LBP-TOP)	Kaltwang et al. [113]
	combinations of LBP-TOP, LPQ-TOP, BSIF-TOP	Yang et al. [104]
	energy from optical flow	Ghasemi et al. [74]
	time-integral of histogram of oriented energies	Irani et al. [122], Irani et al. [66]
	Hankel matrices from time series of Haar and/or Gabor features	Lo Presti and La Cascia [127]
	deep learned spatiotemporal features	Zhou et al. [112], Egede et al. [116], Tavakolian and Hadid [119]
	max temporal pooling on sequence of SIFT based features	Sikka et al. [124], Sikka et al. [125]
hybrid (geometric + textural)	statistical features from sequence of head pose, facial landmark distances, mean gradient magnitudes	Werner et al. [101], Werner et al. [29]
	statistical and time features from sequence of head pose, facial landmark distances, mean gradient magnitudes	Werner et al. [123]
	LBP-TOP, statistical features from facial distances	Kächele et al. [103]
Mixed (Spatial + Spatiotemporal) Features		
texture	HOG around facial landmarks, HOG-TOP	Chen et al. [75]
hybrid (geometric + textural)	facial landmark positions, distances, angles, HOG, deep learned features from image sequence	Egede et al. [116], Egede and Valstar [117]

TABLE 6
Summary of indirect representations of facial images and image sequences used in two-step automatic pain detection approaches either alone or in combination with direct representations.

Feature Category	Features	References
Indirect Features		
non-temporal or single timestep	AU scores	Lucey et al. [83], Lucey et al. [128], Zafar and Khan [129], Zhang et al. [76]
	AU labels	Schmid et al. [77], Siebers et al. [78]
temporal or multiple timesteps	statistical features from AU scores	Littlewort et al. [72], Sikka [73], Sikka et al. [59]
	temporal filters on AU scores	Littlewort et al. [56], Bartlett et al. [131]
	histogram of AUs	Ghasemi et al. [74]
Combined (Direct + Indirect) Features		
temporal or multiple timesteps	statistical features from sequence of AU scores, facial landmark distances, eye gaze coordinates, and head pose	Lopez-Martinez et al. [130]

Almost all classification and regression tasks were supervised. Ground truth in the form of pain or AU labels, and discrete or continuous-valued pain or AU intensities, were used to train the machine learning models. Support Vector Machines (SVM) and its variants such as multiple kernel SVM and Support Vector Regressors (SVR), are the most widely used supervised machine learning methods. Probabilistic methods such as Relevance Vector Machines (RVM) and different variants of conditional random fields, have been used for supervised classification tasks. Random forests have been used for both supervised classification and supervised regression tasks in automatic pain detection. Less commonly used methods for supervised learning include decision trees and classical regression methods such as linear and logistic regression. More recently, deep learning methods such as Convolutional Neural Networks (CNN), Recurrent CNN, and Long Short-Term Memory (LSTM) recurrent neural networks, are increasingly being used for end-to-end learning of pain intensities from single images (e.g. [109]) or image sequences (e.g. [115], [112]).

Very few works explored machine learning strategies other than supervised learning. For example, an unsupervised comparative learning method was used by Werner et al. in [90] for estimating continuous-valued pain intensity; Sikka et al. [124], [125] employed weakly supervised learning with the help of a multiple-instance variant of boosting algorithms for pain event localization in an image sequence. Semi-supervised learning strategies have not yet been explored in the context of automatic pain detection from facial expressions.

The metrics used to quantify the performance of automatic pain detection from facial expressions depend on the learning task. For classification tasks, metrics such as accuracy, F1 score, and area under Receiver Operating

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2019.2958341, IEEE Transactions on Pattern Analysis and Machine Intelligence

TABLE 7
Summary of machine learning methods used in the automatic pain detection approaches.

Prediction Task	Approach	Machine Learning Method	References
Supervised Methods			
classification	one-step	Support Vector Machine (SVM)	Brahnam et al. [79], Brahnam et al. [81], Lu et al. [51], Monwar and Rezaei [87], Lucey et al. [83], Ashraf et al. [82], Niese et al. [54], Siebers et al. [84], Nanni et al. [85], Gholami et al. [86], Lucey et al. [18], Lucey et al. [89], Hammal and Cohn [120], Werner et al. [90], Werner et al. [101], Khan et al. [92], Pedersen [93], Neshov and Manolova [94], Lo Presti and La Cascia [126], Singh [60], Chen et al. [75], Aung et al. [27], Rathee and Ganotra [95], Kharghanian et al. [96], Roy et al. [97], Werner et al. [123], Yang et al. [104], Rupenga and Vadapalli [98], Tsai et al. [28]
		Relevance Vector Machine (RVM)	Gholami et al. [86]
		random forest	Khan et al. [92], Werner et al. [29], Kächele et al. [103], Werner et al. [123]
		multiple kernel SVM	Wei and Li-Min [88], Chen et al. [75]
		Neural Network (NN)	Monwar and Rezaei [80]
		Neural Network Simultaneous Optimization Algorithm (NNSOA)	Brahnam et al. [81]
		extreme learning machine	Rupenga and Vadapalli [98], Alphonse and Dharma [100]
		decision tree	Khan et al. [92]
		k-nearest neighbors	Siebers et al. [84], Khan et al. [92], Lo Presti and La Cascia [126]
		k-nearest neighbors + hidden Markov model	Meng and Bianchi-Berthouze [102]
		Adaboost or its variants	Yuan et al. [106], Chen et al. [91], Lo Presti and La Cascia [127]
		transferable belief model	Hammal et al. [71], Hammal and Kunz [105]
		heteroscedastic conditional ordinal random field	Rudovic et al. [121]
		hidden conditional random field	Lopez-Martinez et al. [118]
		regularized multi-task learning	Romera-Paredes et al. [108]
	two-step	SVM	step1-AU: Lucey et al. [83], Lucey et al. [128] step2-pain: Bartlett et al. [131] both steps: Littlewort et al. [72], Littlewort et al. [56], Ghasemi et al. [74]
		logistical linear regression	step2-pain: Lucey et al. [83], Lucey et al. [128]
		k-nearest neighbors	step1-AU: Zafar and Khan [129]
		logistic regression	step2-pain: Sikka et al. [59]
		alignment-based learning	step2-pain: Schmid et al. [77], Siebers et al. [78]
		hidden conditional random field	step2-pain: Ghasemi et al. [74]
		latent-dynamic conditional random field	step1-AU: Zhang et al. [76]
regression	one-step	support vector regression	Florea et al. [111], Lopez-Martinez et al. [118]
		ordinal support vector regression	Zhao et al. [114]
		relevance vector regression or its variants	Kaltwang et al. [107], Kaltwang et al. [113], Egede et al. [116], Egede and Valstar [117]
		random forest	Kächele et al. [103]
		linear regression	Neshov and Manolova [94]
		ordinal support vector regression	Zhao et al. [114]
		NN	Lopez-Martinez et al. [118]
		Convolutional Neural Network (CNN)	Wang et al. [109]
		3D CNN with kernels of varying temporal lengths	Tavakolian and Hadid [119]
		recurrent CNN	Zhou et al. [112]
		LSTM recurrent neural network	Rodriguez et al. [115], Lopez-Martinez et al. [118]
	two-step	support vector regression	step1-AU: Bartlett et al. [131], Sikka et al. [59] both steps: Sikka [73]
		linear regression	step2-pain: Sikka et al. [59]
		multi-task NN	step2-pain: Lopez-Martinez et al. [130]
Weakly-Supervised Methods			
classification	one-step	multiple instance learning-Boost	Sikka et al. [124], Sikka et al. [125]
regression	one-step	ordinal support vector regression	Zhao et al. [114]
		NN + Gaussian process regression model	Liu et al. [110]
Unsupervised Methods			
regression	one-step	comparative learning	Werner et al. [90]
		ordinal support vector regression	Zhao et al. [114]

Characteristic (ROC) curves have been used to report the performance. For regression tasks, correlation and mean absolute error have been commonly used. Leave-one-subject-out (LOSO) or person-independent 10-fold crossvalidation is normally performed to evaluate the learning performance. A detailed survey of the performance metrics and evaluation strategies used is out of scope of this paper.

6 Discussion

The survey of papers on automatic pain detection from facial expressions–covered in Sections 4 and 5–shows significant progress in the field since the first methods appeared in 2006. The approaches followed two paradigms: learning pain targets directly from input features (one-step approaches); and learning an intermediate representation of facial expressions in terms of AUs, based on which pain detection was performed (two-step approaches). The approaches performed pain detection at frame-level, or at sequence-level. Dynamic information about facial expressions of pain was considered in some of the approaches either by extracting spatiotemporal features or by using dynamic learning methods. Spatial and spatiotemporal information have been used for representing the shape and appearance of facial expressions of pain, and its changes over time, respectively. Shape information in the form of geometric features has been rarely used alone for automatic pain detection. Appearance information in the form of texture descriptors has been used more successfully, sometimes in combination with geometric features. A wide variety of machine learning methods have been explored (see Table 7) for classification as well as regression tasks. Supervised machine learning methods dominate the field. Classification tasks (such as distinguishing pain from non-pain condition, other emotions or states, and detection of discrete pain intensity levels) have received more focus than the regression task of continuous pain intensity estimation. The UNBC McMaster Shoulder Pain Archive Database [18] is the most widely used dataset by researchers in this field. Other datasets have appeared to address other diagnostic conditions, to investigate other acute pain stimulation methods, and to provide multimodal pain information (see Section 4 and Tables 1 and 2). However, there are a number of challenges that should be addressed in order to create robust, real-time systems for inferring pain by analyzing facial expressions. These challenges pertain to data as well as approaches, and are interwined with each other. Future research needs to address these challenges in order to bring automatic pain detection systems closer to practical usability. The following paragraphs highlight some of these challenges, and propose future research directions.

One of the main challenges is the acquisition of appropriate training and testing data. When it comes to automatic pain monitoring in infants, critically ill, elderly, or cognitively impaired groups, data acquisition becomes even harder. Due to the ethical challenges involved, the common practice is to record videos of cognitively healthy individuals, mostly young adults, by experimentally inducing acute pain and other distress states in controlled laboratory settings. This is evident from the Tables 1 and 2, where the diagnostic condition is mostly listed as 'healthy', and the demographic information show less coverage of the older old and infants. It is known that systems trained only on young faces do not generalise well to older faces [65] due to the textural differences caused by skin ageing, and the variations in facial muscle elasticity and facial motion dynamics. In addition, systems trained only on cognitively healthy cannot capture well the differences in pain expressions and their underlying nuances that could possibly be characteristic of persons with cognitive impairments. Furthermore, a system trained only on facial expressions of experimentally induced acute pain, cannot perform well in the context of clinical acute pain. Although the facial expressions of experimental pain and clinical pain are very similar, a significant difference lies in the fact that experimental pain induction is usually performed while a person is sitting still (not moving), whereas clinical pain becomes more visible in a person who is in motion (e.g. during morning care). Moreover, the temporal characteristics vary a lot between short experimental stimulation and long lasting clinical pain. To address the above-mentioned challenge, there is a need to develop automatic pain detection methods that can first be trained and tested on acute pain expression data of healthy adults which is relatively easier to collect, and can later be tuned with limited amout of clinical or experimental data collected from different cohorts belonging to the vulnerable categories.

The pain datasets have mostly been created under controlled settings, and therefore, do not sufficiently cover the variations that could occur due to environmental influences in real application settings. For example, in an uncontrolled monitoring situation, abrupt changes in head pose and facial expressions could be caused by events such as a person entering the patient's room or an object falling down. Moreover, head and body movements are part of daily life activities. Another set of variations that can occur in real-life settings is the variations in illumination conditions. Such variations have not been systematically covered in the existing pain datasets. In order to develop pain detection approaches that are robust to such environmental influences, there is a need for building datasets that cover such variations in good proportions. Future research on automatic pain detection needs to focus on developing solutions that combine multiple tasks in order to identify and isolate contextual and environmental influences. Head and body pose detection, as well as motion tracking should be integrated in automatic pain detection systems to improve their robustness to sudden head and body movements that may be unrelated to pain.

Pain can be experienced and expressed during rest as well as during movement. However, most of the datasets listed in Tables 1 and 2 have been gathered under the condition that the person is sitting on a chair. In the Pediatric Pain Dataset [59] [73], the monitored patients were lying on a bed with raised head, and the camera was positioned such that an almost frontal view of the face was obtained. More pain datasets that systematically cover different mobility modes such as sitting, standing, walking, and lying need to be built. Future research should explore automatic pain detection during these modes. This could require integration of activity recognition and activity-specific pain detection models.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2019.2958341, IEEE Transactions on Pattern Analysis and Machine Intelligence

Another challenge is to obtain good quality annotations for the collected data. Facial expressions of pain involve high variability [48] and closely resemble facial expressions associated with distress states [141]. Therefore, objective coding standards like FACS need to be applied in order to study the differences between different states. However, FACS coding is time-consuming and requires trained coders. Self-reports and observer-reports are also used to label pain sequences. In the case of noncommunicative patients, self-report of pain cannot be obtained. Observer reports may have biases or errors, especially if the observer is not a trained expert. The reviewed literature used different types of information as the ground truth for training and testing the automatic pain detection approaches. For example, Werner et al. [101] used pain stimulus level as the segment-level ground truth; Liu et al. [110] used self-report and observer report as the sequence-level ground truth; Zhou et al. [112] used AU-based PSPI [16] as the frame-level ground truth. The use of semi-supervised and weakly supervised machine learning methods could reduce the need for labeling of entire data at frame and sequence levels by human experts. Instead, the experts could focus on labeling the exemplary instances of different pain expression patterns and on the challenging or ambiguous instances that are present in the dataset. This would reduce the time and costs involved in annotation of large pain datasets. However, so far, semi-supervised methods have not been explored for automatic pain detection, and very few weakly supervised approaches have been reported. Future research could focus more on semi-supervised and weakly supervised methods to overcome the practical challenge of annotating large amounts of data.

Automatic pain detection systems based on facial expressions need to be capable of adapting itself to the person-specific facial morphology, facial texture, and pain expression. This adaption should preferably occur online, without the need for manual intervention or cooperation from monitored user. So far, offline methods have been explored for person-specific adaption of learned pain models. Werner et al. [101] explored person-specific models, where a separate, customized model was developed for each subject. This approach would require sufficiently large amount of data from each person, and would not scale as a generic pain detection solution that can be deployed widely. It would be more advantageous to learn models for different categories of users. Liu et al. [110] explored the use of personal information such as age, gender, and complexion as additional personalised features in order to take different cohort-specific variations into consideration. Most approaches have tried to learn generic, person-independent pain expression models, and have used LOSO or person-independent crossvalidation methods to demonstrate the generalisation capability of these approaches to unseen faces. In order for these approaches to adapt well to new users, the pain datasets should include a large number of subjects to cover as much as possible, the identity-related variance in facial morphology, facial texture, and facial expression of pain. There is a need to develop validated benchmark datasets that would enable the comparison of different automatic pain detection approaches, especially their ability to generalize or dynamically adapt to unseen faces. This benchmark dataset could also be created by combining existing or new datasets, after validation by psychologists. Evaluation on a benchmark dataset should be promoted, since it is needed to identify which approaches are most promising.

The features extracted from images and image sequences were either simple features like facial landmark positions or complex features like those learned by CNNs (see Tables 4 and 5). The feature representations learned by machine learning methods are often not interpretable. In addition, it is difficult for humans to interpret which features and feature combinations affect the decision-making in which way. Explainable Artificial Intelligence (AI) methods such as Layer-wise Relevance Propagation [142] and Local Interpretable Model-Agnostic Explanations (LIME) [143] could be used to make decision-making transparent and comprehensible to humans [23]. Some approaches (e.g. [76]) used an intermediate representation of facial expressions based on FACS and used simple statistical operations to detect pain. Such methods provide better interpretability of predictions. Human diagnosis of pain using observational pain assessment tools are based on rules. For human comprehensibility and deployment in the area of care giving, it would be beneficial to have a rule-based system that can explain the decisions. Such rules might be learned with interpretable machine learning methods such as grammar inference [78] or inductive logic programming [144]. Experienced medical staff can refer to explanations of the system's decisions to control the quality of automated decision making. Inexperienced persons can profit from explanations by gaining deeper insight into facial cues for pain [145].

The decision boundaries/thresholds/strategies that are learned for pain detection are sensitive to noise in observed features. Probabilistic methods can be used to model the observation-based and model-based uncertainties. Liu et al. [110] used a Gaussian process model that predicts sequence-level pain intensity rating and an associated uncertainty estimate. Hammal and Kunz [105] used Transferable Belief Model to model uncertainty in the features and to adjust the certainty of its predictions accordingly. Future research should focus more on integrating machine learning with different AI methods such as symbolic inference, logic, and reasoning with uncertainty, in order to leverage the strengths of different methods.

As revealed by Table 3, there has been very limited focus on the problem of distinguishing pain from emotions, especially negative emotions. However, it is crucial in clinical applications to differentiate the different distress states such as pain, anger, fear, and disgust [26]. Future research should investigate this problem more intensely. To address the need for sufficient data on pain and emotions, combination of existing pain datasets and emotion datasets could be considered.

It would be interesting to study the long-term pain expression dynamics and include this information in the development of a continuous pain monitoring system. This information could be used to choose or adapt the desired temporal granularity for pain predictions, and to determine how this granularity changes over time. Adapting the temporal granularity could help in making specific facial expression patterns related to or involved in pain to be more/less likely to be indicative of other distress states rather than pain itself. This could help in reducing the occurrence of false alarms and missed detections. For example, information about the duration for which pain medication would be effective could be useful in adjusting the pain detection thresholds as well as the window of time that is examined for pain events, immediately after the administration of the medication. It is highly likely, based on earlier observations, that continuous subjective pain does not result in a continuous facial expression of pain. Rather, there are fluctuations in the form of a sporadic waxing and waning of facial expression. Learning more about the timing of these episodes would also help to improve the automatic pain detection models.

The use of AUs defined in FACS for automatic facial expression analysis allows objective, reliable, and anatomically-based coding of facial responses. They represent a common language that enables results to be easily compared between studies. However, the coding of AUs is limited by human observation capacities. That is, only visible movements are coded. Consequently, very subtle facial expressions cannot be assessed. In order to assess visible as well as previsible facial activity, the use of a sufficiently sensitive EMG is necessary. However, although EMG can pick up even subtle muscle activities, only a limited number of facial muscles can be assessed simultaneously. Since EMG electrodes have to be attached to the skin of the face, a placement of more than 3 or 4 electrodes is not advisable, because it would otherwise become too obtrusive and would interfere with the facial expression [146]. Moreover, the ability of surface EMG to isolate a single facial muscle is much poorer compared to FACS based analysis (due to EMG crosstalk amongst neighboring muscles) [146]. In addition, facial EMG electrodes would introduce facial occlusions that would make a simultaneous video-based analysis of facial expressions difficult. Therefore, within the realm of pain assessment from the face, FACS based facial expression analysis is the more practical option for continuous pain monitoring.

In this survey, we focused on automatic detection of pain from facial expressions. However, there have been efforts to apply machine learning methods to combine facial activity with other modalities such as vocalizations [28] or ECG, EMG, and skin conductance [29]. Improvements in pain recognition rates were reported when information from multiple modalities were fused. Combining facial expressions, vocalizations, body movements, brain activation, and autonomic responses, would be important as well as useful in order to get a holistic picture of the pain experienced, as well as to improve the performance of automatic pain detection systems through complementary and/or supplementary evidence. Future research should investigate the potential as well as the limits of such multimodal systems for robust detection of pain in clinical settings. The development of reliable, contactless or minimally obtrusive methods for assessing physiological signals would also be necessary for improving the practical usability of such multimodal methods.

7 Conclusion

This paper surveys literature on automatic pain detection from facial expressions. The reviewed literature is structured and categorized based on the learning tasks, and the features and machine learning methods used. Two main types of approaches were identified. One detects pain directly from the visual input, and the other learns an intermediate representation of facial expression. The datasets consisting of facial expressions of pain that have been used by the reviewed literature have also been summarized, while highlighting important details such as the pain induction methods, demographic information, and available annotations. Several challenges related to data acquisition and development of learning methods have been discussed, and future research directions have been identified. More datasets fulfilling additional criteria are required, and there is a need to build and designate a dataset or a combination of datasets for benchmarking automatic pain detection approaches. Semi-supervised and weakly supervised approaches should be explored for automatic pain detection to reduce the heavy dependence on labeled data. Future methods should focus more on the task of distinguishing pain from other emotions. Interpretability of features learned by AI methods and the explainability of the decisions made by them is crucial for their applicability and acceptance in clinical practice, as supporting systems for pain diagnosis and pain monitoring. Using multimodal information about pain, and integrating other tasks such as head pose, motion, and facial occlusion detection could help in improving robustness and performance of automatic pain detection systems. Probabilistic information is also essential for ensuring robustness and reliability in highly dynamic real-life conditions. Interdisciplinary research is needed to solve several challenges involved in developing a robust, automatic pain detection system. The temporal characteristics of pain episodes, and the effect of pain medication could be useful for reducing the occurrence of false alarms. The ability of automatic pain detection systems to generalize across cohorts with different diagnostic status should be investigated in the future.

Acknowledgments

This work has been partially funded by the grant received by the first author under the TALENTA *start* programme of the Fraunhofer Society during the period from February 2015 to January 2017. This work has also been partially supported by DFG grants GA 2485/3-1, LA 685/18, Schm 1239/15 , "Video-based automated pain detection exploiting compositional and temporal characteristics of action unit."

References

[1] H. Merskey and N. Bogduk, Eds., *PART III pain terms, a current list with definitions and notes on usage*. IASP Press, 2012.

[2] A. C. d. C. Williams, "Facial expression of pain: an evolutionary account," *Behavioral and Brain Sciences*, vol. 25, no. 4, pp. 439–455, 2002.

[3] P. D. Wall, "Introduction to the fourth edition," in *The Textbook of Pain*, P. D. Wall and R. Melzack, Eds. Harcourt Publishers, 1999, pp. 1–8.

[4] D. Cipher and R. Clifford, "Dementia, pain, depression, behavioral disturbances, and ADLs: toward a comprehensive conceptualization of quality of life in long-term care," *International Journal of Geriatric Psychiatry*, vol. 19, no. 8, pp. 741–748, 08 2004.

[5] O. Moriarty, B. McGuire, and D. Finn, "The effect of pain on cognitive function: a review of clinical and preclinical research," *Progress in Neurobiology*, vol. 93, no. 3, pp. 385–404, 03 2011.

[6] B. Bosley, D. Weiner, T. Rudy, and E. Granieri, "Is chronic nonmalignant pain associated with decreased appetite in older adults? preliminary evidence," *Journal of the American Geriatrics Society*, vol. 52, no. 2, pp. 247–251, 02 2004.

[7] M. Giron, Y. Forsell, C. Bernsten, M. Thorslund, B. Winblad, and J. Fastbom, "Sleep problems in a very old population: drug use and clinical correlates," *Journals of Gerontology A-Biological Science and Medical Science*, vol. 57, no. 4, pp. 236–240, 04 2002.

[8] B. Husebo, C. Ballard, R. Sandvik, O. Nilsen, and D. Aarsland, "Efficacy of treating pain to reduce behavioural disturbances in residents of nursing homes with dementia: cluster randomised clinical trial," *British Medical Journal*, vol. 343, pp. 1–10, 07 2011.

[9] H. M. McCormack, J. d. L. David, and S. Sheather, "Clinical applications of visual analogue scales: a critical review," *Psychological Medicine*, vol. 18, no. 4, pp. 1007–1019, 1988.

[10] W. Downie, P. Leatham, V. Rhind, V. Wright, J. Branco, and J. Anderson, "Studies with pain rating scales," *Annals of the Rheumatic Diseases*, vol. 37, no. 4, pp. 378–381, 1978.

[11] R. de Wit, F. van Dam, M. Hanneman, L. Zandbelt, A. van Buuren, K. van der Heijden, G. Leenhouts, S. Loonstra, and H. H. Abu-Saad, "Evaluation of the use of a pain diary in chronic cancer pain patients at home," *Pain*, vol. 79, no. 1, pp. 89–99, 1999.

[12] R. H. Gracely, P. McGrath, and R. Dubner, "Ratio scales of sensory and affective verbal pain descriptors," *Pain*, vol. 5, no. 1, pp. 5–18, 1978.

[13] H. Dehghani, H. Tavangar, and A. Ghandehari, "Validity and reliability of behavioral pain scale in patients with low level of consciousness due to head trauma hospitalized in intensive care unit," *Archives of Trauma Research*, vol. 3, no. 1, 2014.

[14] V. Warden, A. Hurley, and L. Volicer, "Development and psychometric evaluation of the Pain Assessment in Advanced Dementia (painad) scale," *Journal of the American Medical Directors Association*, vol. 4, no. 1, pp. 9–15, 2003.

[15] J. Lawrence, D. Alcock, P. McGrath, J. Kay, S. MacMurray, and C. Dulberg, "The development of a tool to assess neonatal pain," *Neonatal Network : NN*, vol. 12, no. 6, pp. 59–66, September 1993.

[16] K. M. Prkachin and P. E. Solomon, "The structure, reliability and validity of pain expression: evidence from patients with shoulder pain," *PAIN*, vol. 139, no. 2, pp. 267–274, 2008.

[17] T. Hadjistavropoulos, M. E. Browne, K. M. Prkachin, B. Taati, A. Ashraf, and A. Mihailidis, "Pain in severe dementia: a comparison of a fine-grained assessment approach to an observational checklist designed for clinical settings," *European Journal of Pain*, vol. 22, no. 5, pp. 915–925, May 2018.

[18] P. Lucey, J. F. Cohn, K. M. Prkachin, P. E. Solomon, and I. Matthews, "Painful data: the UNBC-McMaster shoulder pain expression archive database," in *Face and Gesture 2011*, March 2011, pp. 57–64.

[19] J. Kappesser and A. C. de C. Williams, "Pain estimation: asking the right questions," *Pain*, vol. 148, no. 2, pp. 184–187, 2010.

[20] V. Engle, M. Graney, and A. Chan, "Accuracy and bias of licensed practical nurse and nursing assistant ratings of nursing home residents' pain," *The Journals of Gerontology: Series A*, vol. 56, no. 7, pp. M405–M411, 2001.

[21] W. Achterberg, M. Pieper, A. van Dalen-Kok, M. de Waal, B. Husebo, S. Lautenbacher, M. Kunz, E. Scherder, and A. Corbett, "Pain management in patients with dementia," *Clinical Interventions in Aging*, vol. 8, pp. 1471–1482, 2013.

[22] K. Shannon and T. Bucknall, "Pain assessment in critical care: what have we learnt from research," *Intensive and Critical Care Nursing*, vol. 19, no. 3, pp. 154–162, 2003.

[23] K. Weitz, T. Hassan, U. Schmid, and J.-U. Garbas, "Deep-learned faces of pain and emotions: elucidating the differences of facial expressions with the help of explainable ai methods," *tm-Technisches Messen*, vol. 86, no. 7-8, pp. 404–412, 2019.

[24] E. Sheu, J. Versloot, R. Nader, D. Kerr, and C. K.D., "Pain in the elderly: validity of facial expression components of observational measures," *The Clinical Journal of Pain*, vol. 27, no. 7, pp. 593–601, 2011.

[25] M. Kunz, V. Mylius, S. Scharmann, K. Schepelman, and S. Lautenbacher, "Influence of dementia on multiple components of pain," *European Journal of Pain*, vol. 13, no. 3, pp. 317–325, 2009.

[26] M. Kunz, D. Seuss, T. Hassan, J. Garbas, M. Siebers, U. Schmid, M. Schöberl, and S. Lautenbacher, "Problems of video-based pain detection in patients with dementia: a road map to an interdisciplinary solution," *BMC Geriatrics*, vol. 17, no. 33, 2017.

[27] M. S. H. Aung, S. Kaltwang, B. Romera-Paredes, B. Martinez, A. Singh, M. Cella, M. Valstar, H. Meng, A. Kemp, M. Shafizadeh, A. C. Elkins, N. Kanakam, A. de Rothschild, N. Tyler, P. J. Watson, A. C. d. C. Williams, M. Pantic, and N. Bianchi-Berthouze, "The automatic detection of chronic pain-related expression: requirements, challenges and the multimodal EmoPain dataset," *IEEE Transactions on Affective Computing*, vol. 7, no. 4, pp. 435–451, Oct 2016.

[28] F. Tsai, Y. Hsu, W. Chen, Y. Weng, C. Ng, and C. Lee, "Toward development and evaluation of pain level-rating scale for emergency triage based on vocal characteristics and facial expressions," in *Interspeech 2016*, 2016, pp. 92–96.

[29] P. Werner, A. Al-Hamadi, R. Niese, S. Walter, S. Gruss, and H. C. Traue, "Automatic pain recognition from video and biomedical signals," in *2014 22nd International Conference on Pattern Recognition*, Aug 2014, pp. 4582–4587.

[30] S. F. Bunk, S. Lautenbacher, J. Rüsseler, K. Müller, J. Schultz, and M. Kunz, "Does EEG activity during painful stimulation mirror more closely the noxious stimulus intensity or the subjective pain sensation?" *Somatosensory & Motor Research*, vol. 35, no. 3-4, pp. 192–198, 2018.

[31] M. M. Nickel, E. S. May, L. Tiemann, P. Schmidt, M. Postorino, S. T. Dinh, J. Gross, and M. Ploner, "Brain oscillations differentially encode noxious stimulus intensity and pain intensity," *NeuroImage*, vol. 148, pp. 141–147, 2017.

[32] T. D. Wager, L. Y. Atlas, M. A. Lindquist, M. Roy, C.-W. Woo, and E. Kross, "An fMRI-based neurologic signature of physical pain," *New England Journal of Medicine*, vol. 368, no. 15, pp. 1388–1397, 2013.

[33] C. Wang, T. Pun, and G. Chanel, "A comparative survey of methods for remote heart rate detection from frontal face videos," *Frontiers in Bioengineering and Biotechnology*, vol. 6, 2018.

[34] L. Tarassenko, M. Villarroel, A. Guazzi, J. Jorge, D. A. Clifton, and C. Pugh, "Non-contact video-based vital sign monitoring using ambient light and auto-regressive models," *Physiological Measurement*, vol. 35, no. 5, pp. 807–831, March 2014.

[35] M. Kunz, D. Meixner, and S. Lautenbacher, "Facial muscle movements encoding pain–a systematic review," *Pain*, vol. 160, no. 3, pp. 535–549, March 2019.

[36] P. Ekman and W. V. Friesen, *Facial Action Coding System*. Palo Alto, CA: Consulting Psychologists Press, 1978.

[37] N. Dael, M. Mortillaro, and K. R. Scherer, "The Body Action and Posture coding system (BAP): development and reliability," *Journal of Nonverbal Behavior*, vol. 36, no. 2, pp. 97–121, June 2012.

[38] E. M. J. Huis In 't Veld, G. J. M. van Boxtel, and B. de Gelder, "The Body Action Coding System II: muscle activations during the perception and expression of emotion," *Frontiers in Behavioral Neuroscience*, vol. 8, p. 330, 2014.

[39] P. Viola and M. J. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, May 2004.

[40] B. Fasel and J. Luettin, "Automatic facial expression analysis: a survey," *Pattern Recognition*, vol. 36, no. 1, pp. 259–275, 2003.

[41] D. Huang, C. Shan, M. Ardabilian, Y. Wang, and L. Chen, "Local binary patterns and its application to facial image analysis: a survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 41, no. 6, pp. 765–781, Nov 2011.

[42] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105.

[43] P. N. Druzhkov and V. D. Kustikova, "A survey of deep learning methods and software tools for image classification and object detection," *Pattern Recognition and Image Analysis*, vol. 26, no. 1, pp. 9–15, Jan 2016.

[44] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," in *British Machine Vision Conference*, 2015.

[45] K. M. Prkachin, "The consistency of facial expressions of pain: a comparison across modalities," *Pain*, vol. 51, no. 3, pp. 297–306, 1992.

[46] M. Kunz, N. Faltermeier, and S. Lautenbacher, "Impact of visual learning on facial expressions of physical distress: a study on voluntary and evoked expressions of pain in congenitally blind

and sighted individuals," *Biological Psychology*, vol. 89, no. 2, pp. 467–476, 2012.

[47] K. D. Craig, K. M. Prkachin, and R. E. Grunau, "The facial expression of pain," in *Handbook of Pain Assessment*, D. C. Turk and R. Melzack, Eds. New York, NY, US: Guilford Press, 2001, pp. 153–169.

[48] M. Kunz and S. Lautenbacher, "The faces of pain: a cluster analysis of individual differences in facial activity patterns of pain." *European Journal of Pain*, vol. 18, no. 6, pp. 813–823, July 2014.

[49] ——, "Improving recognition of pain by calling attention to its various faces," *European Journal of Pain*, vol. 19, no. 9, pp. 1350–1361, 2015.

[50] S. Brahnam, C.-F. Chuang, F. Y. Shih, and M. R. Slack, "Machine recognition and representation of neonatal facial displays of acute pain," *Artificial Intelligence in Medicine*, vol. 36, no. 3, pp. 211–222, 2006.

[51] G. Lu, X. Li, and H. Li, "Facial expression recognition for neonatal pain assessment," in *2008 International Conference on Neural Networks and Signal Processing*, June 2008, pp. 456–460.

[52] S. Walter, S. Gruss, H. Ehleiter, , H. C. Traue, P. Werner, A. Al-Hamadi, S. Crawcour, A. O. Andrade, and G. Moreira da Silva, "The BioVid heat pain database data for the advancement and systematic validation of an automated pain recognition system," in *2013 IEEE International Conference on Cybernetics (CYBCO)*, June 2013, pp. 128–131.

[53] M. Kunz, S. Scharmann, U. Hemmeter, K. Schepelmann, and S. Lautenbacher, "The facial expression of pain in patients with dementia," *PAIN*, vol. 133, no. 1, pp. 221 – 228, 2007.

[54] R. Niese, A. Al-Hamadi, A. Panning, D. Brammen, U. Ebmeyer, and B. Michaelis, "Towards pain recognition in post-operative phases using 3D-based features from video and support vector machines," *International Journal of Digital Content Technology and its Applications*, vol. 3, no. 4, Dec. 2009.

[55] S. Brahnam, C.-F. Chuang, R. S. Sexton, and F. Y. Shih, "Machine assessment of neonatal facial expressions of acute pain," *Decision Support Systems*, vol. 43, no. 4, pp. 1242–1254, 2007, special Issue Clusters.

[56] G. C. Littlewort, M. S. Bartlett, and K. Lee, "Automatic coding of facial expressions displayed during posed and genuine pain," *Image and Vision Computing*, vol. 27, no. 12, pp. 1797–1803, 2009.

[57] M. Kunz, C. Chatelle, S. Lautenbacher, and P. Rainville, "The relation between catastrophizing and facial responsiveness to pain," *PAIN*, vol. 140, no. 1, pp. 127–134, 2008.

[58] D. J. Wilkie, "Facial expressions of pain in lung cancer," *Analgesia*, vol. 1, no. 2, pp. 91–99, 1995.

[59] K. Sikka, A. A. Ahmed, D. Diaz, M. S. Goodwin, K. D. Craig, M. S. Bartlett, and J. S. Huang, "Automated assessment of children's postoperative pain using computer vision," *Pediatrics*, vol. 136, no. 1, pp. e124–e131, 2015.

[60] S. K. Singh and S. K. Singh, "SIFT and SURF performance evaluation for pain assessment using facial expressions," *Journal of Biological Engineering Research and Review*, vol. 2, no. 1, pp. 06–14, 2015.

[61] K. Herr, K. Bjoro, and S. Decker, "Tools for assessment of pain in nonverbal older adults with dementia: a state-of-the-science review," *Journal of Pain and Symptom Management*, vol. 31, no. 2, pp. 170–192, 2006.

[62] C. Regnard, D. Mathews, L. Gibson, and C. Clarke, "Difficulties in identifying distress and its causes in people with severe communication problems," *International Journal of Palliative Nursing*, vol. 9, no. 4, 2013.

[63] "World population prospects: the 2017 revision, key findings & advance tables," United Nations, Department of Economic and Social Affairs, Population Division, Tech. Rep. ESA/P/WP/248, 2017.

[64] M. Prince, R. Bryce, E. Albanese, A. Wimo, W. Ribeiro, and C. P. Ferri, "The global prevalence of dementia: a systematic review and metaanalysis," *Alzheimer's & Dementia*, vol. 9, no. 1, 2013.

[65] T. C. Hassan, "Recognizing emotions conveyed through facial expressions," Department of Computer Science, Hochschule Bonn-Rhein-Sieg - University of Applied Sciences, Sankt Augustin, Germany, Tech. Rep. 04-2017, 2017.

[66] R. Irani, K. Nasrollahi, M. O. Simon, C. A. Corneanu, S. Escalera, C. Bahnsen, D. H. Lundtoft, T. B. Moeslund, T. L. Pedersen, M. Klitgaard, and L. Petrini, "Spatiotemporal analysis of RGB-D-T facial images for multimodal pain level recognition," in *2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, June 2015, pp. 88–95.

[67] J. Kappesser and A. C. C. Williams, "Pain and negative emotions in the face: judgements by health care professionals," *PAIN*, vol. 99, no. 1, pp. 197–206, 2002.

[68] B. J. Matuszewski, W. Quan, and L. K. Shark, "High-resolution comprehensive 3-D dynamic database for facial articulation analysis," in *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, Nov 2011, pp. 2128–2135.

[69] X. Zhang, L. Yin, J. F. Cohn, S. Canavan, M. Reale, A. Horowitz, P. Liu, and J. M. Girard, "BP4D-Spontaneous: a high-resolution spontaneous 3D dynamic facial expression database," *Image and Vision Computing*, vol. 32, no. 10, pp. 692–706, 2014, best of Automatic Face and Gesture Recognition 2013.

[70] S. Roy, C. Roy, I. Fortin, C. Either-Majcher, P. Belin, and F. Gosselin, "A dynamic facial expression database [abstract]," *Journal of Vision*, vol. 7, no. 9, June 2007.

[71] Z. Hammal, M. Kunz, M. Arguin, and F. Gosselin, "Spontaneous pain expression recognition in video sequences," in *Proceedings of the 2008 International Conference on Visions of Computer Science: BCS International Academic Conference*, ser. VoCS'08. Swindon, UK: BCS Learning & Development Ltd., 2008, pp. 191–210.

[72] G. C. Littlewort, M. S. Bartlett, and K. Lee, "Faces of pain: automated measurement of spontaneous facial expressions of genuine and posed pain," in *Proceedings of the 9th International Conference on Multimodal Interfaces*, ser. ICMI '07. New York, NY, USA: ACM, 2007, pp. 15–21.

[73] K. Sikka, "Facial expression analysis for estimating pain in clinical settings," in *Proceedings of the 16th International Conference on Multimodal Interaction*, ser. ICMI '14. New York, NY, USA: ACM, 2014, pp. 349–353.

[74] A. Ghasemi, X. Wei, P. Lucey, S. Sridharan, and C. Fookes, "Social signal processing for pain monitoring using a hidden conditional random field," in *2014 IEEE Workshop on Statistical Signal Processing (SSP)*, June 2014, pp. 61–64.

[75] J. Chen, Z. Chi, and H. Fu, "A new approach for pain event detection in video," in *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*, Sep. 2015, pp. 250–254.

[76] X. Zhang, L. Yin, and J. F. Cohn, "Three dimensional binary edge feature representation for pain expression analysis," in *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, vol. 1, May 2015, pp. 1–7.

[77] U. Schmid, M. Siebers, D. Seuss, M. Kunz, and S. Lautenbacher, "Applying grammar inference to identify generalized patterns of facial expressions of pain," in *Heinz, Jeffrey; de la Higuera, Colin; Oates, Tim (Hrsg.): Proceedings of the Eleventh International Conference on Grammatical Inference, PMLR*, vol. 21. Heidelberg: Springer, 2012, pp. 183–188.

[78] M. Siebers, U. Schmid, D. Seuß, M. Kunz, and S. Lautenbacher, "Characterizing facial expressions by grammars of action unit sequences - a first investigation using ABL," *Information Sciences*, vol. 329, pp. 866–875, 2016, special issue on Discovery Science.

[79] S. Brahnam, C.-F. Chuang, F. Y. Shih, and M. R. Slack, "SVM classification of neonatal facial images of pain," in *Fuzzy Logic and Applications*, I. Bloch, A. Petrosino, and A. G. B. Tettamanzi, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 121–128.

[80] M. M. Monwar and S. Rezaei, "Pain recognition using artificial neural network," in *2006 IEEE International Symposium on Signal Processing and Information Technology*, Aug 2006, pp. 28–33.

[81] S. Brahnam, L. Nanni, and S. Randall, "Neonatal facial pain detection using NNSOA and LSVM," in *The 2008 International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV'08), Las Vegas*, 2008.

[82] A. B. Ashraf, S. Lucey, J. F. Cohn, Tsuhan Chen, Z. Ambadar, K. M. Prkachin, and P. E. Solomon, "The painful face – Pain expression recognition using active appearance models," *Image and Vision Computing*, vol. 27, no. 12, pp. 1788–1796, 2009.

[83] P. Lucey, J. Cohn, S. Lucey, I. Matthews, S. Sridharan, and K. M. Prkachin, "Automatically detecting pain using facial actions," in *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, Sep. 2009, pp. 1–8.

[84] M. Siebers, M. Kunz, S. Lautenbacher, and U. Schmid, "Classifying facial pain expressions: individual classifiers vs. global classifiers," in *Reichardt, Dirk (Hrsg.): Proceedings of the 4th Workshop on

Emotion and Computing - Current Research and Future Impact, Sep. 2009.

[85] L. Nanni, S. Brahnam, and A. Lumini, "A local approach based on a local binary patterns variant texture descriptor for classifying pain states," *Expert Systems with Applications*, vol. 37, no. 12, pp. 7888–7894, 2010.

[86] B. Gholami, W. M. Haddad, and A. R. Tannenbaum, "Relevance vector machine learning for neonate pain intensity assessment using digital imaging," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 6, pp. 1457–1466, June 2010.

[87] M. M. Monwar and S. Rezaei, "Support vector machine for automatic pain recognition," in *Proc. SPIE 7246, Computational Imaging VII*, vol. 7246, 2009.

[88] Z. Wei and X. Li-min, "Pain expression recognition based on SLPP and MKSVM," *International Journal of Engineering and Manufacturing*, vol. 1, no. 3, Jun. 2011.

[89] P. Lucey, J. F. Cohn, K. M. Prkachin, P. E. Solomon, S. Chew, and I. Matthews, "Painful monitoring: automatic pain monitoring using the UNBC-McMaster shoulder pain expression archive database," *Image and Vision Computing*, vol. 30, no. 3, pp. 197–205, 2012, best of Automatic Face and Gesture Recognition 2011.

[90] P. Werner, A. Al-Hamadi, and R. Niese, "Pain recognition and intensity rating based on comparative learning," in *2012 19th IEEE International Conference on Image Processing*, Sep. 2012, pp. 2313–2316.

[91] J. Chen, X. Liu, P. Tu, and A. Aragones, "Person-specific expression recognition with transfer learning," in *2012 19th IEEE International Conference on Image Processing*, Sept 2012, pp. 2621–2624.

[92] R. A. Khan, A. Meyer, H. Konik, and S. Bouakaz, "Pain detection through shape and appearance features," in *2013 IEEE International Conference on Multimedia and Expo (ICME)*, July 2013, pp. 1–6.

[93] H. Pedersen, "Learning appearance features for pain detection using the UNBC-McMaster shoulder pain expression archive database," in *Computer Vision Systems*, L. Nalpantidis, V. Krüger, J.-O. Eklundh, and A. Gasteratos, Eds. Cham: Springer International Publishing, 2015, pp. 128–136.

[94] N. Neshov and A. Manolova, "Pain detection from facial characteristics using supervised descent method," in *2015 IEEE 8th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, vol. 1, Sep. 2015, pp. 251–256.

[95] N. Rathee and D. Ganotra, "Multiview distance metric learning on facial feature descriptors for automatic pain intensity detection," *Computer Vision and Image Understanding*, vol. 147, no. C, pp. 77–86, Jun. 2016.

[96] R. Kharghanian, A. Peiravi, and F. Moradi, "Pain detection from facial images using unsupervised feature learning approach," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug 2016, pp. 419–422.

[97] S. D. Roy, M. K. Bhowmik, P. Saha, and A. K. Ghosh, "An approach for automatic pain detection through facial expression," *Procedia Computer Science, Proceeding of the Seventh International Conference on Intelligent Human Computer Interaction (IHCI 2015)*, vol. 84, pp. 99–106, 2016, proceeding of the Seventh International Conference on Intelligent Human Computer Interaction (IHCI 2015).

[98] M. Rupenga and H. B. Vadapalli, "Automatic spontaneous pain recognition using supervised classification learning algorithms," in *2016 Pattern Recognition Association of South Africa and Robotics and Mechatronics International Conference (PRASA-RobMech)*, Nov 2016, pp. 1–6.

[99] F. Meawad, S.-Y. Yang, and F. L. Loy, "Automatic detection of pain from spontaneous facial expressions," in *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, ser. ICMI '17. New York, NY, USA: ACM, 2017, pp. 397–401.

[100] A. S. Alphonse and D. Dharma, "Chapter 4 - novel computational intelligence techniques for automatic pain detection and pain intensity level estimation from facial expressions using distributed computing for big data," in *Computational Intelligence for Multimedia Big Data on the Cloud with Engineering Applications*, ser. Intelligent Data-Centric Systems, A. K. Sangaiah, M. Sheng, and Z. Zhang, Eds. Academic Press, 2018, pp. 91–121.

[101] P. Werner, A. Al-Hamadi, R. Niese, S. Walter, S. Gruss, and H. Traue, "Towards pain monitoring: facial expression, head pose, a new database, an automatic system and remaining challenges," in *Proceedings of the British Machine Vision Conference*. BMVA Press, 2013.

[102] H. Meng and N. Bianchi-Berthouze, "Affective state level recognition in naturalistic facial and vocal expressions," *IEEE Transactions on Cybernetics*, vol. 44, no. 3, pp. 315–328, March 2014.

[103] M. Kächele, P. Thiam, M. Amirian, P. Werner, S. Walter, F. Schwenker, and G. Palm, "Multimodal data fusion for person-independent, continuous estimation of pain intensity," in *Engineering Applications of Neural Networks*, L. Iliadis and C. Jayne, Eds. Cham: Springer International Publishing, 2015, pp. 275–285.

[104] R. Yang, S. Tong, M. Bordallo, E. Boutellaa, J. Peng, X. Feng, and A. Hadid, "On pain assessment from facial videos using spatio-temporal local descriptors," in *2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, Dec 2016, pp. 1–6.

[105] Z. Hammal and M. Kunz, "Pain monitoring: a dynamic and context-sensitive system," *Pattern Recognition*, vol. 45, no. 4, pp. 1265–1280, 2012.

[106] L. Yuan, F. S. Bao, and G. Lu, "Recognition of neonatal facial expressions of acute pain using boosted Gabor features," in *2008 20th IEEE International Conference on Tools with Artificial Intelligence*, vol. 2, Nov 2008, pp. 473–476.

[107] S. Kaltwang, O. Rudovic, and M. Pantic, "Continuous pain intensity estimation from facial expressions," in *Advances in Visual Computing*, G. Bebis, R. Boyle, B. Parvin, D. Koracin, C. Fowlkes, S. Wang, M.-H. Choi, S. Mantler, J. Schulze, D. Acevedo, K. Mueller, and M. Papka, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 368–377.

[108] B. Romera-Paredes, M. S. H. Aung, M. Pontil, N. Bianchi-Berthouze, A. C. d. C. Williams, and P. Watson, "Transfer learning to account for idiosyncrasy in face and body expressions," in *2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, April 2013, pp. 1–6.

[109] F. Wang, X. Xiang, C. Liu, T. Tran, A. Reiter, G. Hager, H. Quon, J. Cheng, and A. Yuille, "Regularizing face verification nets for pain intensity regression," in *2017 IEEE International Conference on Image Processing, ICIP 2017 - Proceedings*, vol. 2017-September. IEEE Computer Society, 2 2018, pp. 1087–1091.

[110] D. Liu, F. Peng, A. Shea, O. Rudovic, and R. W. Picard, "DeepFaceLIFT: interpretable personalized models for automatic estimation of self-reported pain," in *Proceedings of Machine Learning Research*, vol. 66, 2017, pp. 1–16.

[111] C. Florea, L. Florea, R. Butnaru, A. Bandrabur, and C. Vertan, "Pain intensity estimation by a self-taught selection of histograms of topographical features," *Image and Vision Computing*, vol. 56, no. C, pp. 13–27, Dec. 2016.

[112] J. Zhou, X. Hong, F. Su, and G. Zhao, "Recurrent convolutional neural network regression for continuous pain intensity estimation in video," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2016.

[113] S. Kaltwang, S. Todorovic, and M. Pantic, "Doubly sparse relevance vector machine for continuous facial behavior estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 9, pp. 1748–1761, Sep. 2016.

[114] R. Zhao, Q. Gan, S. Wang, and Q. Ji, "Facial expression intensity estimation using ordinal information," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016, pp. 3466–3474.

[115] P. Rodriguez, G. Cucurull, J. Gonzlez, J. M. Gonfaus, K. Nasrollahi, T. B. Moeslund, and F. X. Roca, "Deep pain: exploiting long short-term memory networks for facial expression classification," *IEEE Transactions on Cybernetics*, pp. 1–11, 2017.

[116] J. Egede, M. Valstar, and B. Martinez, "Fusing deep learned and hand-crafted features of appearance, shape, and dynamics for automatic pain estimation," in *2017 12th IEEE International Conference on Automatic Face Gesture Recognition (FG 2017)*, May 2017, pp. 689–696.

[117] J. O. Egede and M. Valstar, "Cumulative attributes for pain intensity estimation," in *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, ser. ICMI '17. New York, NY, USA: ACM, 2017, pp. 146–153.

[118] D. Lopez-Martinez, O. Rudovic, and R. Picard, "Personalized automatic estimation of self-reported pain intensity from facial expressions," in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, July 2017, pp. 2318–2327.

[119] M. Tavakolian and A. Hadid, "Deep spatiotemporal representation of the face for automatic pain intensity estimation," in *2018 24th International Conference on Pattern Recognition (ICPR)*, Aug 2018, pp. 350–354.

[120] Z. Hammal and J. F. Cohn, "Automatic detection of pain intensity," in *Proceedings of the 14th ACM International Conference on Multimodal Interaction*, ser. ICMI '12. New York, NY, USA: ACM, 2012, pp. 47–52.

[121] O. Rudovic, V. Pavlovic, and M. Pantic, "Automatic pain intensity estimation with heteroscedastic conditional ordinal random fields," in *Advances in Visual Computing*, G. Bebis, R. Boyle, B. Parvin, D. Koracin, B. Li, F. Porikli, V. Zordan, J. Klosowski, S. Coquillart, X. Luo, M. Chen, and D. Gotz, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 234–243.

[122] R. Irani, K. Nasrollahi, and T. B. Moeslund, "Pain recognition using spatiotemporal oriented energy of facial muscles," in *2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, June 2015, pp. 80–87.

[123] P. Werner, A. Al-Hamadi, K. Limbrecht-Ecklundt, S. Walter, S. Gruss, and H. C. Traue, "Automatic pain assessment with facial activity descriptors," *IEEE Transactions on Affective Computing*, vol. 8, no. 3, pp. 286–299, July 2017.

[124] K. Sikka, A. Dhall, and M. Bartlett, "Weakly supervised pain localization using multiple instance learning," in *2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, April 2013, pp. 1–8.

[125] K. Sikka, A. Dhall, and M. S. Bartlett, "Classification and weakly supervised pain localization using multiple segment representation," *Image and Vision Computing*, vol. 32, no. 10, pp. 659–670, 2014, best of Automatic Face and Gesture Recognition 2013.

[126] L. Lo Presti and M. La Cascia, "Using hankel matrices for dynamics-based facial emotion recognition and pain detection," in *2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, June 2015, pp. 26–33.

[127] L. L. Presti and M. L. Cascia, "Boosting hankel matrices for face emotion recognition and pain detection," *Computer Vision and Image Understanding*, vol. 156, pp. 19–33, 2017, image and Video Understanding in Big Data.

[128] P. Lucey, J. F. Cohn, I. Matthews, S. Lucey, S. Sridharan, J. Howlett, and K. M. Prkachin, "Automatically detecting pain in video through facial action units," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 41, no. 3, pp. 664–674, June 2011.

[129] Z. Zafar and N. A. Khan, "Pain intensity evaluation through facial action units," in *2014 22nd International Conference on Pattern Recognition*, Aug 2014, pp. 4696–4701.

[130] D. Lopez-Martinez, O. Rudovic, and R. Picard, "Physiological and behavioral profiling for nociceptive pain estimation using personalized multitask learning," in *Neural Information Processing Systems (NIPS) Workshop on Machine Learning for Health*, Long Beach, USA, 2017.

[131] M. Bartlett, G. Littlewort, M. Frank, and K. Lee, "Automatic decoding of facial movements reveals deceptive pain expressions," *Current Biology*, vol. 24, no. 7, pp. 738 – 743, 2014.

[132] E. Sariyanidi, H. Gunes, and A. Cavallaro, "Automatic analysis of facial affect: a survey of registration, representation, and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 6, pp. 1113–1133, June 2015.

[133] D. Gabor, "Theory of communication. part 1: the analysis of information," *Journal of the Institution of Electrical Engineers - Part III: Radio and Communication Engineering*, vol. 93, pp. 429–441, November 1946.

[134] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 07 2002.

[135] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognition*, vol. 29, no. 1, pp. 51–59, 1996.

[136] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," in *International Conference on Computer Vision & Pattern Recognition (CVPR '05)*, C. Schmid, S. Soatto, and C. Tomasi, Eds., vol. 1. San Diego, United States: IEEE Computer Society, Jun. 2005, pp. 886–893.

[137] G. Zhao and M. Pietikainen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 29, no. 6, pp. 915–928, June 2007.

[138] J. Chen, Z. Chen, Z. Chi, and H. Fu, "Emotion recognition in the wild with feature fusion and multiple kernel learning," in *Proceedings of the 16th International Conference on Multimodal Interaction*, ser. ICMI '14. New York, NY, USA: ACM, 2014, pp. 508–513.

[139] V. Ojansivu and J. Heikkilä, "Blur insensitive texture classification using local phase quantization," in *Image and Signal Processing*, A. Elmoataz, O. Lezoray, F. Nouboud, and D. Mammass, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 236–243.

[140] J. Kannala and E. Rahtu, "BSIF: binarized statistical image features," in *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*, Nov 2012, pp. 1363–1366.

[141] M. Kunz, J. Peter, S. Huster, and S. Lautenbacher, "Pain and disgust: The facial signaling of two aversive bodily experiences," *PLOS ONE*, vol. 8, no. 12, 12 2013.

[142] S. Bach, A. Binder, G. Montavon, F. Klauschen, K.-R. Mller, and W. Samek, "On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation," *PLOS ONE*, vol. 10, no. 7, pp. 1–46, 07 2015.

[143] M. T. Ribeiro, S. Singh, and C. Guestrin, """Why should I trust you?": explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '16. New York, NY, USA: ACM, 2016, pp. 1135–1144.

[144] S. H. Muggleton, U. Schmid, C. Zeller, A. Tamaddoni-Nezhad, and T. Besold, "Ultra-strong machine learning: comprehensibility of programs learned with ilp," *Machine Learning*, vol. 107, no. 7, pp. 1119–1140, Jul 2018.

[145] U. Schmid, "Inductive programming as approach to comprehensible machine learning," in *KI 2018 Workshop on Formal and Cognitive Reasoning*, 2018, invited Paper.

[146] U. Hess, "Facial EMG," in *Methods in social neuroscience*, E. Harmon-Jones and J. S. Beer, Eds. New York, NY, US: Guilford Press, 2009, pp. 70–91.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2019.2958341, IEEE Transactions on Pattern Analysis and Machine Intelligence

Teena Hassan holds a Bachelor of Technology degree in Computer Science and Engineering from Cochin University of Science and Technology in Kerala, India. She received her Master of Science degree in Autonomous Systems from the Bonn-Rhein-Sieg University of Applied Sciences, Sankt Augustin, Germany, in 2014. After graduation, she continued her research in the field of automatic facial expression analysis at Fraunhofer IIS, Erlangen, Germany. The focus of her research was the modeling of facial muscle movements, fusing multiple sources of facial expression information, and modeling uncertainty in observations. Currently, she is a Research Associate at Bielefeld University, and is working on interaction architectures for social robots. She is an external PhD candidate at University of Bamberg.

Dominik Seuß received a Master of Science degree in Applied Computer Science from the University of Bamberg, Germany. After graduation, he joined Fraunhofer IIS, Erlangen, Germany, where he continued his research on automatic facial expression analysis. He is interested in machine learning topics in the field of deep learning in computer vision and generative networks. He is an external PhD candidate at University of Bamberg.

Johannes Wollenberg received his Diploma degree in Computer Science from University of Bremen, Germany, in 2013. After graduation, he worked as Research Associate at Fraunhofer IIS, Erlangen, Germany. His research interests include content-based video understanding, spatiotemporal feature extraction, and machine learning. His current work focuses on machine learning projects in the life sciences industry.

Katharina Weitz received a Master of Science in Psychology and a Master of Science in Computing in the Humanities (Applied Computer Science) at the University of Bamberg, Germany. She is currently working at the University of Augsburg at the chair for Human-Centered Multimedia. She is interested in machine learning topics in the field of social robotics and virtual agents. The influence of explainability and transparency of intelligent systems on people's trust is a central point of her research activities. She supports a human-centered usage of artificial intelligence and delves into ethical issues. In addition to her research activities, the communication of research knowledge to the general public in the form of lectures, workshops, and exhibitions is an important concern to her.

Miriam Kunz was born near Duesseldorf, Germany in 1977. She studied Psychology in Berlin (1997 2002) and received her PhD from the University of Bamberg (2006). She conducted her postdoctoral research at the University of Montreal (Prof. Pierre Rainville) and the University of Bamberg (Prof. Lautenbacher). In 2015, she moved to the Netherlands to start a Tenure Track position at the University of Groningen at the Department of General Practice and Elderly Care Medicine. In 2019, she moved to Augsburg (Germany) to start as a full professor for Medical Psychology and Sociology at the University of Augsburg. She is an expert in the field of facial expression of pain and in the field of pain in dementia, where her research bridges from experimental basic research to clinical applications.

Stefan Lautenbacher from Munich received his Diploma in Psychology 1983 in Munich and his PhD 1990 in Bamberg. He finished 1997 his habilitation thesis in Bamberg. He worked as fellow and assistant lecturer at the Max Planck Institute in Munich from 1983 to 1994 with an interruption of one year (1991-1992) for a postdoctoral stay in London (Canada). After 7 years as managing psychologist at the University Hospital in Marburg starting 1994, he was appointed 2001 to become full professor for Physiological Psychology in Bamberg. His expertise in the biopsychology of pain allowed him supervising many PhD theses and publishing several book and close to 200 journal articles.

Jens-Uwe Garbas received the diploma and PhD degrees in electrical engineering from Friedrich-Alexander University Erlangen-Nuremberg, Germany, in 2004 and 2010, respectively. In 2010, he joined Fraunhofer Institute for Integrated Circuits IIS, where he was appointed Head of the group Intelligent Systems in 2011 and Deputy Head of Department Electronic Imaging in 2012. He is responsible for industrial and public research projects as well as software licensing in the areas of real-time computer vision, affective computing, and facial analysis.

Ute Schmid holds a diploma in psychology and a diploma in computer science, both from Technical University Berlin (TUB), Germany. She received her doctoral degree (Dr. rer.nat.) in computer science from TUB in 1994 and her habilitation in computer science in 2002. She was assistant professor at the AI/Machine Learning group of TUB and lecturer at University Osnabrück. Since 2004 she holds a professorship of Applied Computer Science/Cognitive Systems at the University of Bamberg. Research interests of Ute Schmid are mainly in the domain of comprehensible machine learning, explainable AI, and high-level learning on relational data, especially inductive programming. Ute Schmid dedicates a significant amount of her time to measures supporting women in computer science and to promote computer science as a topic in elementary, primary, and secondary education.