



# Using AI to Detect Pain through Facial Expressions: A Review

Gioacchino D. De Sario<sup>1</sup>, Clifton R. Haider<sup>2</sup>, Karla C. Maita<sup>1</sup>, Ricardo A. Torres-Guzman<sup>1</sup> , Omar S. Emam<sup>3</sup>, Francisco R. Avila<sup>1</sup>, John P. Garcia<sup>1</sup>, Sahar Borna<sup>1</sup>, Christopher J. McLeod<sup>4</sup>, Charles J. Bruce<sup>4</sup>, Rickey E. Carter<sup>5</sup> and Antonio J. Forte<sup>1,\*</sup> 

<sup>1</sup> Division of Plastic Surgery, Mayo Clinic, Jacksonville, FL 32224, USA

<sup>2</sup> Department of Physiology and Biomedical Engineering, Mayo Clinic, Rochester, MN 55902, USA

<sup>3</sup> Division of AI in Health Sciences, University of Louisville, Louisville, KY 40292, USA

<sup>4</sup> Department of Cardiovascular Medicine, Mayo Clinic, Jacksonville, FL 32224, USA

<sup>5</sup> Department of Health Sciences Research, Mayo Clinic, Jacksonville, FL 32224, USA

\* Correspondence: [ajvforte@yahoo.com.br](mailto:ajvforte@yahoo.com.br)

**Abstract:** Pain assessment is a complex task largely dependent on the patient's self-report. Artificial intelligence (AI) has emerged as a promising tool for automating and objectifying pain assessment through the identification of pain-related facial expressions. However, the capabilities and potential of AI in clinical settings are still largely unknown to many medical professionals. In this literature review, we present a conceptual understanding of the application of AI to detect pain through facial expressions. We provide an overview of the current state of the art as well as the technical foundations of AI/ML techniques used in pain detection. We highlight the ethical challenges and the limitations associated with the use of AI in pain detection, such as the scarcity of databases, confounding factors, and medical conditions that affect the shape and mobility of the face. The review also highlights the potential impact of AI on pain assessment in clinical practice and lays the groundwork for further study in this area.

**Keywords:** artificial intelligence; AI; machine learning; pain; facial expression



**Citation:** De Sario, G.D.; Haider, C.R.; Maita, K.C.; Torres-Guzman, R.A.; Emam, O.S.; Avila, F.R.; Garcia, J.P.; Borna, S.; McLeod, C.J.; Bruce, C.J.; et al. Using AI to Detect Pain through Facial Expressions: A Review. *Bioengineering* **2023**, *10*, 548. <https://doi.org/10.3390/bioengineering10050548>

Academic Editor: Chien-Hung Liao

Received: 27 March 2023

Revised: 24 April 2023

Accepted: 27 April 2023

Published: 2 May 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Pain is an unpleasant subjective experience caused by actual or potential tissue damage associated with complex neurological and psychosocial components [1,2]. Self-reporting is the primary method of assessing pain, as it is highly individualized and dependent on the individual's perception [3,4].

The medical literature provides several pain scoring systems for pain assessment, including the 100 mm visual analog scale (VAS), the numeric rating scale (NRS), and the color analog scale [5–7]. Studies have shown that the VAS is a highly reliable and valid measure of pain and is the most responsive to treatment effects based on substantial evidence [8].

Despite its value, the VAS is beset by several shortcomings. For instance, it is not feasible to employ it in situations where the individuals are either unconscious, cognitively impaired, or unable to articulate themselves verbally [9].

Observational scales have been developed and validated for use in different clinical settings and with specific patient populations to address patients' inability to communicate their pain. These scales, such as the Behavioral Pain Scale, Nociception Coma Scale, and Children's Revised Impact of Event Scale [10–12], offer an alternative method for assessing pain but are limited by the observer's previous training and ability to interpret the pain responses accurately.

Additionally, studies have found that observer biases can affect the results of these scales [13–15]. Therefore, there is a need for a genuinely objective pain assessment method that is also time-sensitive to detect changes in the patient's pain experience.

Artificial intelligence (AI) has the potential to transform the healthcare system by making the analysis of facial expressions during pain more efficient and lessening the workload of human professionals. In particular, AI can automate feature extraction and perform repetitive and time-consuming tasks requiring much human effort by utilizing machine learning (ML) algorithms and data analysis techniques; this may result in better patient outcomes, better use of resources, and lower operating costs [16,17].

The potential of AI in medical imaging analysis has recently come to light in recent studies. Large datasets of medical images, including computed tomography scans and X-rays, can be used to train AI algorithms to recognize abnormalities that point to the presence of a disease. These algorithms have been shown to perform better than human radiologists in some diagnostic tasks [18,19], highlighting the potential of AI to increase the accuracy and efficiency of medical imaging analysis.

#### *Research Question and Objectives*

This study aims to explore the current state of AI-assisted pain detection using facial expressions. The specific objectives of this study are as follows:

1. Summarize the current state of research in this field.
2. Identify and discuss the potential implications and challenges of deploying this technology in the healthcare system.
3. Determine research gaps and propose areas for future work.

## **2. Materials and Methods**

For the literature review, we conducted a search on 23 January 2023 using keywords in 4 databases: PubMed, EMBASE/MEDLINE, Google Scholar, Cumulative Index of Nursing and Allied Health Literature (CINAHL), and Web of Science, to identify relevant literature and evidence on the use of AI and ML to detect pain through facial expressions. Posteriorly, we conducted a narrative synthesis to provide a comprehensive overview of the current state of the art, the potential for clinical use, challenges, limitations, ethical concerns, and knowledge gaps for future research.

## **3. Objective Pain Measurement and AI**

There has been considerable research on pain responses to develop a more “objective” way of assessing pain. Pain responses include changes in physiological parameters such as galvanic skin response, pupil reflexes, blood pressure, heart rate variability, and hormonal and biochemical markers [20–24]. Additionally, behavioral pain responses can be verbal, such as describing or vocalizing pain, and nonverbal, such as withdrawal behavior, body posture, and facial expressions [25–27].

Most attempts to recognize facial expressions have focused on the identification of action units (AUs), defined in the Facial Action Coding System (FACS) [28]. Numerous AUs in the FACS have been linked to pain. However, according to Prkachin (1992), the ones that convey the most information regarding pain are brow lowering, eye closure, orbit tightening, and levator muscle contraction [29]. These four “core” factors also contribute to the majority of the heterogeneity in pain expression [30].

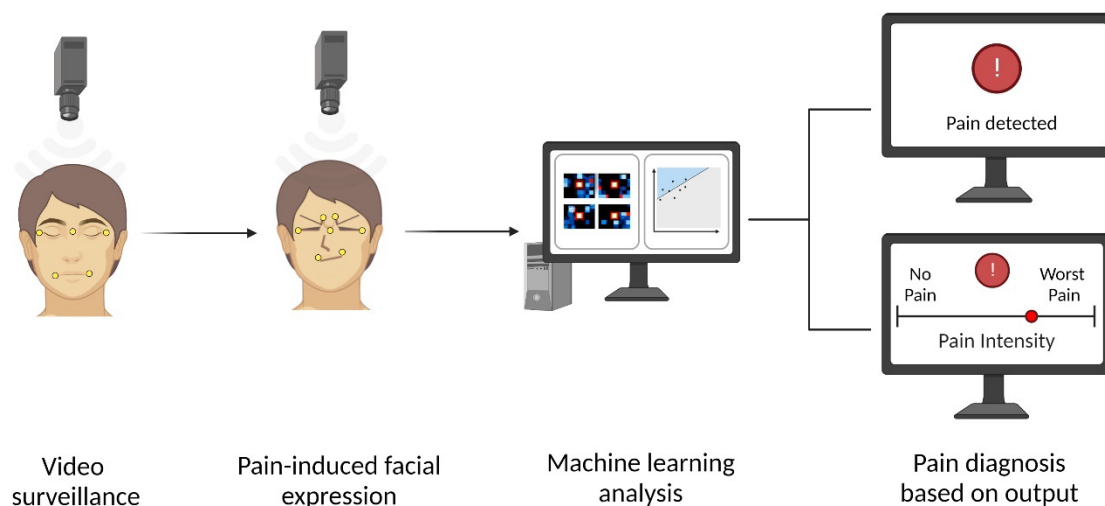
However, the facial indicators of pain that have been validated in the past are impractical for clinical settings due to their reliance on highly skilled observers to label facial AUs, a time-consuming task unsuitable for real-time pain assessment [31,32]. Nevertheless, facial expressions are advantageous in AI/ML because they can provide relevant data in each video frame and changes over time, and computer vision systems could perform this operation automatically through the training of a classifier to recognize the facial expressions connected to pain [33].

#### *Models Using AI/ML for Pain Detection through Facial Expressions*

The first step for the automated detection of pain tasks is to develop a pre-trained ML system. For supervised ML models, this step involves training with large datasets labeled

with the correct output, processed by algorithms and mathematical models to recognize patterns associated with the output. Afterward, the inferential phase is started, where the ML model is loaded with new data to generate categorizations. Typically, a camera records video data of a subject’s face. The facial features are then extracted from the video data using computer vision techniques to identify pain-related patterns. These facial features found in frames or video sequences are later processed by the pre-trained ML models, providing their estimation of the subject’s pain experience [34,35].

Figure 1 depicts a standard proposed scenario for detecting pain through video surveillance of patient faces using computer vision and ML techniques.



**Figure 1.** Automated pain detection using AI. This image depicts video surveillance being used to capture facial expressions associated with pain, which are then analyzed by a computer system using machine learning to provide an accurate output of pain detection or intensity estimation. Created with [BioRender.com](https://www.biorender.com) (accessed on 14 March 2023).

#### 4. Current Evidence of AI-Based Pain Detection through Facial Expressions

Several studies have found promising findings on the precision of AI-based pain detection using facial expressions. Table 1 summarizes the results of 15 experimental studies that used AI/ML to detect pain using facial expressions.

**Table 1.** Summary of studies assessing the use of AI to detect pain through facial expressions.

Author and Date	Population	Pain Setting	Ground Truth	ML Classifiers	Outcomes	Performance
Fontaine et al. (2022) [36]	Adult patients from a single university hospital	Postoperative pain	NRS	CNN	Pain intensity estimation	Estimation of pain intensity <ul style="list-style-type: none"> <li>• Accuracy = 53%</li> <li>• Mean error = 2.4 points</li> </ul> Detection of pain (NRS $\geq 4/10$ ) <ul style="list-style-type: none"> <li>• Sensitivity = 89.7%</li> <li>• Specificity = 61.5%</li> </ul> Detection of severe pain (NRS $\geq 7/10$ ) <ul style="list-style-type: none"> <li>• Sensitivity = 77.5%</li> <li>• Specificity = 45%</li> </ul>
Bargshady et al. (2020) [37]	UNBC-McMaster database MIntPAIN database	UNBC-McMaster database: self-identified shoulder pain MIntPAIN database: electrical-induced pain	UNBC-McMaster database: PSPI MIntPAIN database: stimuli-based pain levels (0–4)	CNN-RNN	Pain intensity estimation <ul style="list-style-type: none"> <li>• UNBC-McMaster: categorized into five levels (PSPI 0, 1, 2–3, 4–5, and <math>\geq 6</math>)</li> <li>• MINT: categorized into five levels (0–4)</li> </ul>	UNBC-McMaster <ul style="list-style-type: none"> <li>• Accuracy = 86%</li> <li>• AUC = 90.5%</li> <li>• MSE = 0.081</li> <li>• MAE = 0.103</li> </ul> MIntPAIN <ul style="list-style-type: none"> <li>• Accuracy = 92.26%</li> <li>• AUC = 93.67%</li> <li>• MSE = 0.0245</li> <li>• MAE = 0.0341</li> </ul>

Author and Date	Population	Pain Setting	Ground Truth	ML Classifiers	Outcomes	Performance
Bartlett et al. (2014) [38]	Healthy subjects	Cold pressor-induced pain	Pain stimuli-dependent assessments	SVM	Detection of genuine vs. faked pain	<ul style="list-style-type: none"> <li>AUC = 0.91%</li> <li>Accuracy = 85%</li> </ul>
Othman et al. (2021) [39]	X-ITE Pain Database	Heat-induced and electrical-induced pain	NRS categorized into 4 pain intensities (no pain, low, medium, and severe)	Two-CNN with sample weighting	Pain intensity detection for electrical and thermal stimuli using two groupings of pain levels: none/low/severe and none/moderate/severe	Mean accuracy = 51.7%
Rodriguez et al. (2022) [40]	UNBC-McMaster database	Self-identified shoulder pain	PSPI	CNN-LSTM	Pain detection Estimation of pain intensity, categorized into 6 levels: PSPI 0, 1, 2, 3, 4–5, and $\geq 6$	Pain detection <ul style="list-style-type: none"> <li>Accuracy = 83.1%</li> <li>AUC = 93.3%</li> </ul> Pain intensity estimation <ul style="list-style-type: none"> <li>MSE = 0.74</li> <li>MAE = 0.5</li> </ul>
Rathee et al. (2015) [41]	UNBC-McMaster database	Self-identified shoulder pain	PSPI	DML combined with SVM	Detection of pain intensity by PSPI score (16 levels)	Accuracy = 96%
Lucey et al. (2011) [35]	UNBC-McMaster database	Self-identified shoulder pain	PSPI	SVM	Pain detection	<ul style="list-style-type: none"> <li>Accuracy = 80.9%</li> <li>AUC = 84.7%</li> </ul>
Bargshady et al. (2020) [42]	UNBC-McMaster database	Self-identified shoulder pain	PSPI	Hybrid CNN-bidirectional LSTM	Estimation of pain intensity, categorized into four levels: PSPI 0, 1, 2–3, and $\geq 4$	<ul style="list-style-type: none"> <li>Accuracy = 85% †</li> <li>AUC = 88.7% †</li> <li>MSE = 0.21 †</li> <li>MAE = 0.18 †</li> <li>F-measure = 78.2%</li> </ul>
Littlewort et al. (2009) [43]	University students	Cold pressor-induced pain	Pain stimuli-dependent assessments	Gaussian SVM	Detection of genuine vs. faked pain	Accuracy = 88%
Barua et al. (2022) [44]	UNBC-McMaster database	Self-identified shoulder pain	PSPI	K-Nearest Neighbor	Estimation of pain intensity, categorized into four levels: PSPI 0, 1, 2–3, and $\geq 4$	<ul style="list-style-type: none"> <li>Accuracy = 95.57%</li> <li>Average F1 = 95.67%</li> </ul>
Bargshady et al. (2020) [45]	UNBC-McMaster database MIntPAIN database	UNBC-McMaster database: self-identified shoulder pain MIntPAIN database: electrical-induced pain	UNBC-McMaster database: PSPI MIntPAIN database: stimuli-based pain levels (0–4)	Temporal Convolutional Network	Estimation of pain intensity <ul style="list-style-type: none"> <li>UNBC-McMaster: categorized into four levels: PSPI 0, 1, 2–3, and <math>\geq 4</math></li> <li>MINT: categorized into five levels (0–4)</li> </ul>	UNBC-McMaster <ul style="list-style-type: none"> <li>Accuracy = 94.14%</li> <li>AUC = 91.3%</li> <li>MSE = 0.186</li> <li>MAE = 0.234</li> </ul> MIntPAIN <ul style="list-style-type: none"> <li>Accuracy = 89%</li> <li>AUC = 92%</li> <li>MSE = 0.22</li> <li>MAE = 0.26</li> </ul>
Rathee et al. (2016) [46]	UNBC-McMaster database	Self-identified shoulder pain	PSPI	SVM	Pain detection Estimation of pain intensity, categorized into four levels: PSPI 0, 1, 2, and $\geq 3$	Pain detection <ul style="list-style-type: none"> <li>Accuracy = 89.59%</li> </ul> Pain intensity estimation <ul style="list-style-type: none"> <li>Accuracy = 75%</li> </ul>
Casti et al. (2021) [47]	UNBC-McMaster database	Self-identified shoulder pain	VAS	Linear discriminant analysis	Pain detection (VAS $\geq 0$ ) Pain intensity (VAS) estimation	Pain detection <ul style="list-style-type: none"> <li>AUC = 0.87</li> </ul> Pain intensity estimation <ul style="list-style-type: none"> <li>MAE = 2.44</li> </ul>
Tavakolian et al. (2020) [48]	UNBC-McMaster database BioVid database (part A)	UNBC-McMaster database: self-identified shoulder pain BioVid database: heat-induced pain	UNBC-McMaster database: PSPI BioVid database: stimuli-based pain (5 levels)	CNNs	Estimation of pain intensity UNBC-McMaster: 16 pain levels BioVid: 5 pain levels	Training with BioVid and testing on UNBC-McMaster <ul style="list-style-type: none"> <li>Self-supervised model: AUC = 69.2%</li> <li>Supervised model: AUC = 80.1%</li> </ul> Training with UNBC-McMaster and testing on BioVid <ul style="list-style-type: none"> <li>Self-supervised model: AUC = 65.5%</li> <li>Supervised model: AUC = 75.5%</li> </ul>

**Table 1.** *Cont.*

Author and Date	Population	Pain Setting	Ground Truth	ML Classifiers	Outcomes	Performance
Sikka et al. (2015) [49]	Pediatric patients from a tertiary care center	Postoperative pain	NRS	Logistic regression and linear regression models	Detection of clinically significant pain (NRS ≥ 4)  Pain intensity (NRS) estimation	Clinically significant pain detection <ul style="list-style-type: none"> <li>• Baseline pain: AUC = 0.84</li> <li>• Transient pain: AUC = 0.91</li> </ul> Pain intensity estimation <ul style="list-style-type: none"> <li>• Baseline pain: r = 0.47; z = 4.4 *</li> <li>• Transient pain: r = 0.47; z = 6.0 *</li> </ul>

† Performance using leave-one-subject-out cross validation. \*  $p < 0.0001$ . Abbreviations: UNBC-McMaster (UNBC-McMaster Shoulder Pain Archive), NRS (numeric rating scale), PSPI (Prkachin and Solomon Pain Intensity), VAS (visual analog scale), CERT (Computer Expression Recognition Toolbox), DML (Distance Metric Learning), AAMs (Active Appearance Models), CNN (Convolutional Neural Network), RNN (Recurrent Neural Network), LSTM (Long Short-Term Memory), SVM (Support Vector Machine), AUC (area under the curve), MSE (mean square error), MAE (mean absolute error).

Overall, the studies showed varying levels of accuracy in pain intensity estimation and detection of pain, with some models performing better than others.

The principal outcomes differed among studies. For instance, one study focused only on the detection of pain [35], eight studies only on the estimation of multilevel pain intensity [36,37,39,41,42,44,45,48], and four studied both the detection of pain and the assessment of multilevel pain intensity [40,46,47,49]. Additionally, two studies proposed their automated detection model to differentiate between genuine and faked facial expressions of pain [38,43].

All the presented studies included videos featuring patients’ faces experiencing varied pain levels, including the absence of pain. AI/ML models were trained and tested on these videos to evaluate their performance in detecting pain through facial expressions.

Four studies applied their automated pain detection systems to videos from their recruited patients [36,38,43,49], and eleven used them on at least one public database of pre-recorded patients experiencing pain [35,37,39–42,44–48].

From the 11 studies using public databases, 7 used only one database [35,39–42,46,47], while 3 used a second database to validate further their AI/ML model [37,45,48]. The most used was the UNBC-McMaster Shoulder Pain Archive database, utilized in 10 studies; this consisted of videos of 25 subjects with unilateral shoulder injuries whose pain was elicited by passive and active arm movements [35,37,40–42,44–48]. Two studies used the MintPAIN database, consisting of videos of 20 participants with induced pain from electrical stimulation.

One study used the BioVid database (part A), involving 87 subjects experiencing induced painful heat stimuli [48]. Lastly, one study used the X-ITE Pain database, consisting of 127 individuals whose pain was caused by heat and electrical stimulation [39].

Of the four studies that recruited patients for AI/ML model assessment, one consisted of 1189 patients undergoing different surgeries in a single healthcare center [36]. In addition, two studies assessed pain induced through cold pressor methods in 26 healthy university students [43] and healthy volunteers [38]. Lastly, one study consisted of 50 children who underwent laparoscopic appendectomies, assessing their baseline and palpation-induced pain during the preoperative stage and 3 days post operation [49].

## 5. Discussion

### 5.1. The Ground Truth for Pain Assessment

In the context of pain recognition, ground truth refers to the labels that are used to train and evaluate pain recognition systems. There are three types of ground truth: self-report, observer assessment, and study design [50]. Self-report scales are widely considered the gold standard for measuring pain intensity [51,52]. Observer assessment can be conducted with subjective or validated systematic observation scales, and despite being advantageous

in particular populations unable to report pain, it might have limited accuracy, especially in untrained observers [53,54]. Study design ground truth is based on prior knowledge about the circumstances in which pain is likely to be felt, such as the effects of certain procedures [55].

In the studies presented in Table 1, the ground truth for the pain assessment varied among studies. The validated Prkachin and Solomon Pain Intensity (PSPI) scale was the most frequently used ground truth scale, used in nine studies [35,37,40–42,44–46,48]. In addition, four studies relied on self-reported pain on different scales [36,39,47,49]. Finally, five studies relied on study design ground truth; of these, three used the intensity of the applied stimuli (i.e., study design ground truth), which was previously calibrated to cause different levels of pain in the participants [37,45,48], and two used circumstantial knowledge of painful stimulation [38,43].

### 5.2. Is PSPI Suitable for Estimating Pain?

Recent advances in automatic pain estimation have focused on recognizing AUs as defined in the FACS [56]. PSPI is a scale based on frame-level ground truth calculated by assessing AUs [30].

However, the main strength of the PSPI score is its simplicity, as it condenses facial expressions into one number, making it easy to analyze with regression and classification algorithms, thereby leading to its wide acceptance as a tool for measuring pain [35].

Some weaknesses of the PSPI scale are that it does not reflect the experienced pain severity in all cases. There may be instances where a person experiencing pain may have a low PSPI score despite the presence of significant pain or vice versa. For example, some observers may underestimate a patient's pain experience, and some patients, especially those with motor disorders such as Parkinson's disease, may not exhibit the facial changes assessed in the PSPI scale [57]. Furthermore, it measures the facial expression of pain but does not provide a comprehensive understanding of the experience of pain, which can be influenced by various factors, including psychological and cultural factors [58].

Regarding the validity of the PSPI scale, research has yielded mixed results regarding the correlation between self-reported pain and facial expressions of pain; however, many studies have demonstrated a significant relationship between both [30,59–61].

### 5.3. Performance of AI for Pain Detection through Facial Expressions

In the studies presented in Table 1, the reported accuracy for pain detection ranged from 80.9% to 89.59%, while the AUC ranged from 84% to 93.3%. In pain intensity estimation, the accuracy range was between 51.7% and 96%, while the AUC ranged from 65.5% to 93.67%. Finally, the accuracy range was between 85% and 88% for distinguishing between real and faked pain, with an AUC of 91%.

Most research analyzing facial expressions has examined responses to experimental short-term pain anticipated by subjects. However, it could be possible that facial expressions induced by longer-term pain, such as in cancer pain, may differ from acute pain due to a lack of surprises or expectations. Indeed, this variance may explain the difficulty in creating reliable digital tools to evaluate pain through facial expression analysis for clinical use [48,49].

### 5.4. AI/ML Characteristics and Differences

There are variations among studies in the employed feature extraction tools, ML algorithms, data processing techniques, video or image quality, cross-validation techniques, and other factors that can significantly impact the performance of each model [62].

It is notable that studies utilizing varying techniques on the same populations achieved different degrees of performance (Table 1). Furthermore, the feature extraction tools can significantly impact the accuracy of the models, as demonstrated by some studies where different tools were employed using the same classifiers, resulting in varying levels of accuracy [35,41].

Moreover, as shown in Table 1, pain identification and quantification performance varied even within studies that utilized the same video database.

Although the accuracy variations could be mainly attributed to the feature extraction tools and AI/ML algorithms, further research is necessary to assess the impact of other potential factors.

#### 5.5. Combining Facial Expressions with Other Physiological Data as Input

AI/ML has also been applied to assess pain by fusing the information from facial expressions and other physiological and demographic data. Similar to Sikka et al. (2015) [49], other authors also employed their automated pain detection algorithm on children undergoing laparoscopic appendectomies, demonstrating higher accuracy in detecting clinically significant pain when fusing facial expressions and electrodermal activity as input [63]. Furthermore, other studies have demonstrated that combining facial expression data with demographic and bio-physiological features such as electrocardiograms, electromyography, and skin conductivity can increase the accuracy of pain detection [64–67].

#### 5.6. Machine Learning vs. Human Observers for Pain Estimation

In addition to assessing the performance of the automated detection and quantification of pain, five studies compared the accuracy of human observers to their proposed ML model [36,38,43,49].

Two studies specifically assessed the capability of humans to discriminate genuine vs. faked facial expressions of pain. In the study conducted by Bartlett et al. (2014), trained human observers accurately detected pain in 54.6% of the cases [38]. Moreover, Littlewort et al. (2009) tested human observers and achieved accuracy of 49.1% [43]. In both studies, the authors compared trained and tested ML models, which performed better than human observers (see Table 1), even after training.

Two studies evaluated nurses' capacity to detect pain in postoperative patients. Fontaine et al. (2022) [36] reported on 33 skilled nurses who estimated pain intensity by looking at facial expressions, with 14.9% accuracy and a mean absolute error of 3.04. Their sensitivity and specificity in the detection of pain ( $\text{NRS} \geq 4/10$ ) was 44.9% and 68.4%, while for severe pain ( $\text{NRS} \geq 7/10$ ) the values were 17.0% and 41.1%, respectively. However, the study showed that their AI/ML model outperformed nurses in detecting pain and estimating pain levels, as demonstrated in Table 1 [36]. On the other hand, the results of the study conducted by Sikka et al. (2015) [49] showed that AI/ML performed similarly to nurses estimating baseline postoperative pain and performed better in palpation-induced transient pain. Compared to their ML model's performance (Table 1), the mean AUC achieved by nurses for pain detection was 0.86 and 0.93 for ongoing and transient pain, respectively; for the pain intensity assessment, nurses estimated ongoing and transient pain intensity with a correlation coefficient of  $r = 0.53$  and  $r = 0.54$ , respectively [49]. Moreover, results for automated detection were not impacted by demographic differences, suggesting its advantage against human observers as it does not pose the risk of observer bias [49,68,69].

Lastly, Othman et al. (2021) evaluated the performance of human observers in detecting pain categorized into seven classes, which included three intensities each of heat and electrical pain stimuli and a seventh class for no stimulation. The reported accuracy in the seven-class classification of pain was 21.1%, while for the Convolutional Neural Network classifier accuracy was 27.8% [39].

#### 5.7. Potential Applications

The application of AI/ML techniques in the detection of pain through facial expressions presents a plethora of potential advantages. Firstly, it can provide objective and accurate measurements of pain intensity, which can be used to provide more accurate diagnoses and treatments. Additionally, it can be helpful for the detection of pain in situations where it is difficult to assess, such as in patients unable to communicate verbally, critically ill patients, and during the perioperative period [36,49,70–73].

Inadequate pain management after surgery can have serious consequences, including increased morbidity and mortality, longer recovery times, unexpected hospital readmissions, and chronic persistent pain [74]. Overcoming obstacles to effective pain management, including those related to healthcare providers, is crucial for achieving optimal pain relief after surgery. For example, Sikka et al. (2015) and several other authors have determined that healthcare personnel tend to underestimate children's self-reported pain [49,75,76], which could be translated to a relevant advantage of AI/ML in assisting healthcare personnel in the effective management of postoperative pain.

By utilizing AI/ML technologies, healthcare providers can analyze and interpret patients' facial expressions that coincide with pain, ultimately enabling them to customize treatments and dosages based on individual needs. Moreover, an objective and continuous method for monitoring postoperative pain intensity would be highly advantageous, potentially enabling reliable and cost-effective evaluation of pain intensity.

The results of some studies suggest that AI/ML performs better than human observers at differentiating genuine vs. faked pain [38,43]. The practical implications of this capability are broad, including the detection of malingering, which has been reported to be important in patients seeking compensation [77–79]. Additionally, it could help prevent insurance fraud and unnecessary narcotics prescriptions, reduce healthcare costs, and ultimately improve the quality of care [36].

#### 5.8. Confounding Effect

Evidence suggests that facial expressions of pain are sensitive and specific to pain, and that these expressions can be distinguished from facial expressions associated with basic emotions [80,81]. However, some studies have found that ML algorithms are prone to misinterpreting unpleasant disgust as pain in facial expressions [82]. For instance, Barua et al. (2022) tested their predesigned AI/ML algorithm on the Denver Intensity of Spontaneous Facial Action database, which comprised a set of video frames of the facial expressions of spontaneous emotional expressions. They reported that the proposed pain intensity classification model achieved greater than 95% accuracy in pain detection [44]. Although this database was not designed to study actual pain, AUs associated with pain response are identifiable in video frames, allowing them to be coded using the FACS and the corresponding PSPI scores. Hence, it is essential to consider the specific context in which the automated systems will be used to ensure high accuracy and avoid this confounding effect.

#### 5.9. Ethical Concerns

Using AI/ML algorithms to detect pain through facial expressions raises ethical concerns that must be addressed. For instance, it is essential to consider the potential for errors and inaccuracies in pain detection models. Relying only on inaccurate models could lead to dangerous or inappropriate decisions, such as misdiagnosis, inappropriate treatment, or even legal actions [83].

For instance, misdiagnosing certain conditions based on inaccurate pain detection models may lead to low-quality or no care, or prompt unnecessary surgery or medication; this could lead to an erosion of trust between patients and healthcare providers, with the potential for significant legal and financial implications [84].

Additionally, concerns are being raised regarding patient privacy and autonomy. For example, patients should provide informed consent beforehand as they may refuse facial analysis [85,86]. Furthermore, algorithms might be trained for particular demographics, further marginalizing already vulnerable groups [87,88].

## 6. Challenges and Limitations

Automatic pain detection is challenging because it is complex, subjective, and subject to a variety of factors, such as an individual's personality, social context, and past experiences [89].



Despite the promising results of using AI/ML algorithms to detect pain through facial expressions, they face several limitations. For example, the presence of head motion and rotation, part of typical human behavior in real clinical scenarios, can significantly reduce the accuracy of the AI model's ability to detect AUs [90,91]. Additionally, its utility may be limited by medical conditions affecting facial shape and mobility, such as Parkinson's, stroke, facial injury, or deformity [92–96].

The scarcity of diverse databases further limits the development of a reliable and widely generalizable system for recognizing pain through facial expressions [97]. Additionally, differences between sex, age, and pain setting require validation across large pools of data, prompting the debate over whether to adopt a universal approach or create tailored models for each target population [97].

The Hawthorne effect can be considered a potential limitation of the included studies, whereby the participants' awareness of being observed or filmed may have led to changes in their behavior [98].

Additionally, the application of ML is regarded as a “black-box” method of reasoning, making it challenging to communicate the rationale behind classification choices in a way humans can comprehend [99]. This can be a significant issue as healthcare providers need to understand and interpret the reasoning behind an algorithm's classification decisions in order to make informed decisions about patient care. Therefore, additional research is required to investigate how to improve the clarity and understanding of the reasoning process.

#### *Limitations of This Review*

Most studies concentrated mainly on the technical elements of automated pain identification, with limited exploration of consequences in healthcare as a whole. It is necessary to consider how these innovations may affect patient care and clinical decision making, even if the technical components of this sector are unquestionably crucial. A more comprehensive strategy that considers both technology and healthcare viewpoints might be advantageous for future research.

Although automated pain recognition could be a particularly valuable tool for specific populations limited to self-reported pain, such as individuals with dementia, newborns, patients under anesthesia, and unconscious patients, these groups remained out of the scope of this review.

Given the multiple factors and confounders that could have altered the accuracy of the AI/ML technologies in detecting pain through facial expressions, we could not establish the most dependable and precise methodology. However, we have exhibited the current state of research in automated pain recognition, identifying trends, capabilities, limitations, potential healthcare applications, and knowledge gaps.

## **7. Conclusions**

This review confirms that AI/ML technologies have been used to detect pain through facial expressions to demonstrate their potential to assist during clinical practice. Furthermore, the results indicate that AI/ML can accurately detect and quantify pain through facial expressions, outperforming human observers in pain assessment and detecting deceptive facial expressions of pain. Thus, AI/ML could be a helpful tool in providing objective and accurate measurements of pain intensity, enabling clinicians to make more informed decisions regarding the diagnosis and treatment of pain.

However, it would be wise to encourage the sharing of more diverse and complex publicly available data with the appropriate ethical considerations and proper permissions to allow AI experts to develop reliable and robust methods of facial expression analysis for use in clinical practice. Likewise, well-designed randomized control trials are needed to determine the reliability and generalizability of automated pain detection in real clinical scenarios across medical conditions affecting facial shape and mobility.

Further research is needed to expand the capabilities of AI/ML and test its performance in different pain settings, such as those pertaining to chronic pain conditions, to assess its full potential for use in clinical practice. Additionally, patient satisfaction and preferences regarding the usage and acceptance of AI/ML systems should be explored. Finally, ethical considerations around privacy and algorithm biases are complex and must be addressed.

**Author Contributions:** Conceptualization, G.D.D.S., O.S.E. and R.A.T.-G.; methodology, G.D.D.S. and K.C.M.; validation, G.D.D.S., K.C.M. and R.A.T.-G.; formal analysis, C.R.H.; investigation, F.R.A. and S.B.; data curation, J.P.G. and S.B.; writing—original draft preparation, G.D.D.S.; writing—review and editing, C.R.H., C.J.M., C.J.B. and R.E.C.; visualization, J.P.G.; supervision, A.J.F.; project administration, A.J.F. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The databases generated for drafting this manuscript can be solicited from the author upon reasonable request.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Ghazisaeidi, S.; Muley, M.M.; Salter, M.W. Neuropathic Pain: Mechanisms, Sex Differences, and Potential Therapies for a Global Problem. *Annu. Rev. Pharmacol. Toxicol.* **2023**, *63*, 565–583. [[CrossRef](#)] [[PubMed](#)]
2. Witte, W. Pain and anesthesiology: Aspects of the development of modern pain therapy in the twentieth century. *Der Anaesthesist* **2011**, *60*, 555–566. [[CrossRef](#)] [[PubMed](#)]
3. Katz, J.; Melzack, R. Measurement of pain. *Surg. Clin. N. Am.* **1999**, *79*, 231–252. [[CrossRef](#)] [[PubMed](#)]
4. Melzack, R.; Katz, J. Pain assessment in adult patients. In *Wall and Melzack's Textbook of Pain*; Elsevier: Amsterdam, The Netherlands, 2006.
5. Bulloch, B.; Garcia-Filion, P.; Notricia, D.; Bryson, M.; McConahay, T. Reliability of the color analog scale: Repeatability of scores in traumatic and nontraumatic injuries. *Acad. Emerg. Med.* **2009**, *16*, 465–469. [[CrossRef](#)]
6. Bahreini, M.; Jalili, M.; Moradi-Lakeh, M. A comparison of three self-report pain scales in adults with acute pain. *J. Emerg. Med.* **2015**, *48*, 10–18. [[CrossRef](#)] [[PubMed](#)]
7. Karcioğlu, O.; Topacoglu, H.; Dikme, O.; Dikme, O. A systematic review of the pain scales in adults: Which to use? *Am. J. Emerg. Med.* **2018**, *36*, 707–714. [[CrossRef](#)]
8. Benzon, H.; Raja, S.N.; Fishman, S.E.; Liu, S.S.; Cohen, S.P. *Essentials of Pain Medicine E-Book*; Elsevier Health Sciences: Amsterdam, The Netherlands, 2011.
9. Severgnini, P.; Pelosi, P.; Contino, E.; Serafinelli, E.; Novario, R.; Chiaranda, M. Accuracy of Critical Care Pain Observation Tool and Behavioral Pain Scale to assess pain in critically ill conscious and unconscious patients: Prospective, observational study. *J. Intensive Care* **2016**, *4*, 68. [[CrossRef](#)]
10. Payen, J.F.; Bru, O.; Bosson, J.L.; Lagrasta, A.; Novel, E.; Deschaux, I.; Lavagne, P.; Jacquot, C. Assessing pain in critically ill sedated patients by using a behavioral pain scale. *Crit. Care Med.* **2001**, *29*, 2258–2263. [[CrossRef](#)]
11. Ahn, Y.; Kang, H.; Shin, E. Pain assessment using CRIES, FLACC and PIPP in high-risk infants. *J. Korean Acad. Nurs.* **2005**, *35*, 1401–1409. [[CrossRef](#)]
12. Vink, P.; Lucas, C.; Maaskant, J.M.; van Erp, W.S.; Lindeboom, R.; Vermeulen, H. Clinimetric properties of the Nociception Coma Scale (-Revised): A systematic review. *Eur. J. Pain* **2017**, *21*, 1463–1474. [[CrossRef](#)]
13. Robinson, M.E.; Wise, E.A. Gender bias in the observation of experimental pain. *Pain* **2003**, *104*, 259–264. [[CrossRef](#)]
14. Contreras-Huerta, L.S.; Baker, K.S.; Reynolds, K.J.; Batalha, L.; Cunningham, R. Racial Bias in Neural Empathic Responses to Pain. *PLoS ONE* **2013**, *8*, e84001. [[CrossRef](#)]
15. Khatibi, A.; Mazidi, M. Observers' impression of the person in pain influences their pain estimation and tendency to help. *Eur. J. Pain* **2019**, *23*, 936–944. [[CrossRef](#)] [[PubMed](#)]
16. Davenport, T.; Kalakota, R. The potential for artificial intelligence in healthcare. *Future Healthc. J.* **2019**, *6*, 94–98. [[CrossRef](#)] [[PubMed](#)]
17. Bajwa, J.; Munir, U.; Nori, A.; Williams, B. Artificial intelligence in healthcare: Transforming the practice of medicine. *Future Healthc. J.* **2021**, *8*, e188–e194. [[CrossRef](#)] [[PubMed](#)]
18. Ardila, D.; Kiraly, A.P.; Bharadwaj, S.; Choi, B.; Reicher, J.J.; Peng, L.; Tse, D.; Etemadi, M.; Ye, W.; Corrado, G.; et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat. Med.* **2019**, *25*, 954–961. [[CrossRef](#)]

19. McKinney, S.M.; Sieniek, M.; Godbole, V.; Godwin, J.; Antropova, N.; Ashrafian, H.; Back, T.; Chesus, M.; Corrado, G.S.; Darzi, A.; et al. International evaluation of an AI system for breast cancer screening. *Nature* **2020**, *577*, 89–94. [[CrossRef](#)] [[PubMed](#)]
20. Forte, G.; Troisi, G.; Pazzaglia, M.; Pascalis, V.D.; Casagrande, M. Heart Rate Variability and Pain: A Systematic Review. *Brain Sci.* **2022**, *12*, 153. [[CrossRef](#)]
21. Saccò, M.; Meschi, M.; Regolisti, G.; Detrenis, S.; Bianchi, L.; Bertorelli, M.; Pioli, S.; Magnano, A.; Spagnoli, F.; Giuri, P.G.; et al. The Relationship Between Blood Pressure and Pain. *J. Clin. Hypertens.* **2013**, *15*, 600–605. [[CrossRef](#)]
22. Jha, R.K.; Amatya, M.S.; Nepal, O.; Bade, M.M.; Jha, M.M.K. Effect of cold stimulation induced pain on galvanic skin response in medical undergraduates of Kathmandu University School of medical sciences (KUSMS). *Int. J. Sci. Healthc. Res.* **2017**, *2*, 55–58.
23. López de Audicana-Jimenez de Aberasturi, Y.; Vallejo-De la Cueva, A.; Aretxabala-Cortajarena, N.; Quintano-Rodero, A.; Rodriguez-Nuñez, C.; Pelegrin-Gaspar, P.M.; Gil-Garcia, Z.I.; Margüello-Fernandez, A.A.; Aparicio-Cilla, L.; Parraza-Diez, N. Pupillary dilation reflex and behavioural pain scale: Study of diagnostic test. *Intensive Crit. Care Nurs.* **2023**, *74*, 103332. [[CrossRef](#)]
24. Cowen, R.; Stasiowska, M.K.; Laycock, H.; Bantel, C. Assessing pain objectively: The use of physiological markers. *Anaesthesia* **2015**, *70*, 828–847. [[CrossRef](#)] [[PubMed](#)]
25. Fordyce, W.E. *Behavioral Methods for Chronic Pain and Illness*; Mosby: Maryland Heights, MO, USA, 1976.
26. Keefe, F.J.; Pryor, R.W. Assessment of Pain Behaviors. In *Encyclopedia of Pain*; Schmidt, R.F., Willis, W.D., Eds.; Springer: Berlin/Heidelberg, Germany, 2007; pp. 136–138.
27. LeResche, L.; Dworkin, S.F. Facial expressions of pain and emotions in chronic TMD patients. *Pain* **1988**, *35*, 71–78. [[CrossRef](#)] [[PubMed](#)]
28. Ekman, P.; Friesen, W.V. Facial Action Coding System. *Environ. Psychol. Nonverbal Behav.* **1978**. Available online: <https://psycnet.apa.org/doiLanding?doi=10.1037%2F027734-000> (accessed on 26 March 2023).
29. Prkachin, K.M. The consistency of facial expressions of pain: A comparison across modalities. *Pain* **1992**, *51*, 297–306. [[CrossRef](#)] [[PubMed](#)]
30. Prkachin, K.M.; Solomon, P.E. The structure, reliability and validity of pain expression: Evidence from patients with shoulder pain. *Pain* **2008**, *139*, 267–274. [[CrossRef](#)]
31. Coan, J.A.; Allen, J.J. *Handbook of Emotion Elicitation and Assessment*; Oxford University Press: Oxford, UK, 2007.
32. Craig, K.; Prkachin, K.; Grunau, R.; Turk, D.; Melzack, R. The facial expression of pain. In *Handbook of Pain Assessment*; The Guilford Press: New York, NY, USA, 2001.
33. Lucey, P.; Cohn, J.; Lucey, S.; Matthews, I.; Sridharan, S.; Prkachin, K.M. Automatically Detecting Pain Using Facial Actions. In Proceedings of the 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, Amsterdam, The Netherlands, 10–12 September 2009; pp. 1–8.
34. Feighelstein, M.; Shimshoni, I.; Finka, L.R.; Luna, S.P.L.; Mills, D.S.; Zamansky, A. Automated recognition of pain in cats. *Sci. Rep.* **2022**, *12*, 9575. [[CrossRef](#)]
35. Lucey, P.; Cohn, J.F.; Matthews, I.; Lucey, S.; Sridharan, S.; Howlett, J.; Prkachin, K.M. Automatically detecting pain in video through facial action units. *IEEE Trans. Syst. Man Cybern. B Cybern.* **2011**, *41*, 664–674. [[CrossRef](#)]
36. Fontaine, D.; Vielzeuf, V.; Genestier, P.; Limeux, P.; Santucci-Sivilotto, S.; Mory, E.; Darmon, N.; Lanteri-Minet, M.; Mokhtar, M.; Laine, M.; et al. Artificial intelligence to evaluate postoperative pain based on facial expression recognition. *Eur. J. Pain* **2022**, *26*, 1282–1291. [[CrossRef](#)]
37. Bargshady, G.; Zhou, X.; Deo, R.C.; Soar, J.; Whittaker, F.; Wang, H. Ensemble neural network approach detecting pain intensity from facial expressions. *Artif. Intell. Med.* **2020**, *109*, 101954. [[CrossRef](#)]
38. Bartlett, M.S.; Littlewort, G.C.; Frank, M.G.; Lee, K. Automatic decoding of facial movements reveals deceptive pain expressions. *Curr. Biol.* **2014**, *24*, 738–743. [[CrossRef](#)]
39. Othman, E.; Werner, P.; Saxen, F.; Al-Hamadi, A.; Gruss, S.; Walter, S. Automatic vs. Human Recognition of Pain Intensity from Facial Expression on the X-ITE Pain Database. *Sensors* **2021**, *21*, 3273. [[CrossRef](#)] [[PubMed](#)]
40. Rodriguez, P.; Cucurull, G.; Gonzalez, J.; Gonfaus, J.M.; Nasrollahi, K.; Moeslund, T.B.; Roca, F.X. Deep Pain: Exploiting Long Short-Term Memory Networks for Facial Expression Classification. *IEEE Trans. Cybern.* **2022**, *52*, 3314–3324. [[CrossRef](#)] [[PubMed](#)]
41. Rathee, N.; Ganotra, D. A novel approach for pain intensity detection based on facial feature deformations. *J. Vis. Commun. Image Represent.* **2015**, *33*, 247–254. [[CrossRef](#)]
42. Bargshady, G.; Zhou, X.J.; Deo, R.C.; Soar, J.; Whittaker, F.; Wang, H. Enhanced deep learning algorithm development to detect pain intensity from facial expression images. *Expert. Syst. Appl.* **2020**, *149*, 10. [[CrossRef](#)]
43. Littlewort, G.C.; Bartlett, M.S.; Lee, K. Automatic coding of facial expressions displayed during posed and genuine pain. *Image Vis. Comput.* **2009**, *27*, 1797–1803. [[CrossRef](#)]
44. Barua, P.D.; Baygin, N.; Dogan, S.; Baygin, M.; Arunkumar, N.; Fujita, H.; Tuncer, T.; Tan, R.S.; Palmer, E.; Azizan, M.M.B.; et al. Automated detection of pain levels using deep feature extraction from shutter blinds-based dynamic-sized horizontal patches with facial images. *Sci. Rep.* **2022**, *12*, 17297. [[CrossRef](#)]
45. Bargshady, G.; Zhou, X.; Deo, R.C.; Soar, J.; Whittaker, F.; Wang, H. The modeling of human facial pain intensity based on Temporal Convolutional Networks trained with video frames in HSV color space. *Appl. Soft Comput. J.* **2020**, *97*, 106805. [[CrossRef](#)]
46. Rathee, N.; Ganotra, D. Multiview Distance Metric Learning on facial feature descriptors for automatic pain intensity detection. *Comput. Vis. Image Underst.* **2016**, *147*, 77–86. [[CrossRef](#)]

47. Casti, P.; Mencattini, A.; Filippi, J.; D'Orazio, M.; Comes, M.C.; Di Giuseppe, D.; Martinelli, E. Metrological Characterization of a Pain Detection System Based on Transfer Entropy of Facial Landmarks. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 8. [[CrossRef](#)]
48. Tavakolian, M.; Lopez, M.B.; Liu, L. Self-supervised pain intensity estimation from facial videos via statistical spatiotemporal distillation. *Pattern Recogn. Lett.* **2020**, *140*, 26–33. [[CrossRef](#)]
49. Sikka, K.; Ahmed, A.A.; Diaz, D.; Goodwin, M.S.; Craig, K.D.; Bartlett, M.S.; Huang, J.S. Automated Assessment of Children's Postoperative Pain Using Computer Vision. *Pediatrics* **2015**, *136*, E124–E131. [[CrossRef](#)] [[PubMed](#)]
50. Werner, P.; Lopez-Martinez, D.; Walter, S.; Al-Hamadi, A.; Gruss, S.; Picard, R.W. Automatic Recognition Methods Supporting Pain Assessment: A Survey. *IEEE Trans. Affect. Comput.* **2022**, *13*, 530–552. [[CrossRef](#)]
51. Craig, K.D. The social communication model of pain. *Can. Psychol. Psychol. Can.* **2009**, *50*, 22. [[CrossRef](#)]
52. Herr, K.; Coyne, P.J.; McCaffery, M.; Manworren, R.; Merkel, S. Pain assessment in the patient unable to self-report: Position statement with clinical practice recommendations. *Pain Manag. Nurs.* **2011**, *12*, 230–250. [[CrossRef](#)]
53. Gregory, J. Use of pain scales and observational pain assessment tools in hospital settings. *Nurs. Stand.* **2019**, *34*, 70–74. [[CrossRef](#)] [[PubMed](#)]
54. Lukas, A.; Barber, J.B.; Johnson, P.; Gibson, S.J. Observer-rated pain assessment instruments improve both the detection of pain and the evaluation of pain intensity in people with dementia. *Eur. J. Pain* **2013**, *17*, 1558–1568. [[CrossRef](#)]
55. Walter, S.; Gruss, S.; Ehleiter, H.; Junwen, T.; Traue, H.C.; Werner, P.; Al-Hamadi, A.; Crawcour, S.; Andrade, A.O.; da Silva, G.M. The biovid heat pain database data for the advancement and systematic validation of an automated pain recognition system. In Proceedings of the 2013 IEEE International Conference on Cybernetics (CYBCO), Lausanne, Switzerland, 13–15 June 2013; pp. 128–131.
56. Kunz, M.; Meixner, D.; Lautenbacher, S. Facial muscle movements encoding pain—A systematic review. *Pain* **2019**, *160*, 535–549. [[CrossRef](#)]
57. Prkachin, K.M.; Craig, K.D. Expressing pain: The communication and interpretation of facial pain signals. *J. Nonverbal Behav.* **1995**, *19*, 191–205. [[CrossRef](#)]
58. Kunz, M.; Mylius, V.; Schepelmann, K.; Lautenbacher, S. On the relationship between self-report and facial expression of pain. *J. Pain* **2004**, *5*, 368–376. [[CrossRef](#)]
59. Hadjistavropoulos, T.; LaChapelle, D.L.; MacLeod, F.K.; Snider, B.; Craig, K.D. Measuring movement-exacerbated pain in cognitively impaired frail elders. *Clin. J. Pain* **2000**, *16*, 54–63. [[CrossRef](#)]
60. Prkachin, K.M.; Schultz, I.; Berkowitz, J.; Hughes, E.; Hunt, D. Assessing pain behaviour of low-back pain patients in real time: Concurrent validity and examiner sensitivity. *Behav. Res. Ther.* **2002**, *40*, 595–607. [[CrossRef](#)] [[PubMed](#)]
61. Werner, P.; Al-Hamadi, A.; Limbrecht-Ecklundt, K.; Walter, S.; Gruss, S.; Traue, H. Supplemental Material (Automatic Pain Assessment with Facial Activity Descriptors). *IEEE Trans. Affect. Comput.* **2016**, *8*, 99.
62. Badi Mame, A.; Tapamo, J.-R. A Comparative Study of Local Descriptors and Classifiers for Facial Expression Recognition. *Appl. Sci.* **2022**, *12*, 12156. [[CrossRef](#)]
63. Susam, B.; Riek, N.; Akcakaya, M.; Xu, X.; de Sa, V.; Nezamfar, H.; Diaz, D.; Craig, K.; Goodwin, M.; Huang, J. Automated Pain Assessment in Children Using Electrodermal Activity and Video Data Fusion via Machine Learning. *IEEE Trans. Biomed. Eng.* **2022**, *69*, 422–431. [[CrossRef](#)]
64. Kächele, M.; Thiam, P.; Amirian, M.; Werner, P.; Walter, S.; Schwenker, F.; Palm, G. Multimodal data fusion for person-independent, continuous estimation of pain intensity. In *Communications in Computer and Information Science*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 275–285.
65. Kächele, M.; Werner, P.; Al-Hamadi, A.; Palm, G.; Walter, S.; Schwenker, F. *Bio-Visual Fusion for Person-Independent Recognition of Pain Intensity*; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2015; pp. 220–230.
66. Xu, X.; Susam, B.T.; Nezamfar, H.; Diaz, D.; Craig, K.D.; Goodwin, M.S.; Akcakaya, M.; Huang, J.S.; Virginia, R.S. Towards Automated Pain Detection in Children using Facial and Electrodermal Activity. In Proceedings of the International Workshop on Artificial Intelligence in Health, Stockholm, Sweden, 13–14 July 2018; Volume 2142, pp. 208–211.
67. Liu, D.; Peng, F.; Shea, A.; Ognjen; Rudovic, O.; Picard, R. DeepFaceLIFT: Interpretable Personalized Models for Automatic Estimation of Self-Reported Pain. In Proceedings of the IJCAI 2017 Workshop on Artificial Intelligence in Affective Computing, Melbourne, Australia, 19–25 August 2017.
68. Craig, K.D. Social communication model of pain. *Pain* **2015**, *156*, 1198–1199. [[CrossRef](#)]
69. Kaseweter, K.A.; Drwecki, B.B.; Prkachin, K.M. Racial Differences in Pain Treatment and Empathy in a Canadian Sample. *Pain Res. Manag.* **2012**, *17*, 803474. [[CrossRef](#)]
70. Al Darwish, Z.Q.; Hamdi, R.; Fallatah, S. Evaluation of Pain Assessment Tools in Patients Receiving Mechanical Ventilation. *AACN Adv. Crit. Care* **2016**, *27*, 162–172. [[CrossRef](#)]
71. Rahu, M.A.; Grap, M.J.; Cohn, J.F.; Munro, C.L.; Lyon, D.E.; Sessler, C.N. Facial Expression as an Indicator of Pain in Critically Ill Intubated Adults during Endotracheal Suctioning. *Am. J. Crit. Care* **2013**, *22*, 412–422. [[CrossRef](#)]
72. Wu, C.L.; Liu, S.F.; Yu, T.L.; Shih, S.J.; Chang, C.H.; Yang Mao, S.F.; Li, Y.S.; Chen, H.J.; Chen, C.C.; Chao, W.C. Deep Learning-Based Pain Classifier Based on the Facial Expression in Critically Ill Patients. *Front. Med.* **2022**, *9*, 851690. [[CrossRef](#)]
73. Jiménez-Moreno, C.; Aristizábal-Nieto, J.K.; Giraldo-Salazar, O.L. Classification of Facial Expression of Post-Surgical Pain in Children: Evaluation of Convolutional Neural Networks. *Vis. Electrón.* **2021**, *15*, 7–16.

74. Baratta, J.L.; Schwenk, E.S.; Viscusi, E.R. Clinical consequences of inadequate pain relief: Barriers to optimal pain management. *Plast. Reconstr. Surg.* **2014**, *134*, 15s–21s. [[CrossRef](#)] [[PubMed](#)]
75. Birnie, K.A.; Chambers, C.T.; Fernandez, C.V.; Forgeron, P.A.; Latimer, M.A.; McGrath, P.J.; Cummings, E.A.; Finley, G.A. Hospitalized children continue to report undertreated and preventable pain. *Pain Res. Manag.* **2014**, *19*, 198–204. [[CrossRef](#)] [[PubMed](#)]
76. Kang, M.S.; Park, J.; Kim, J. Agreement of Postoperative Pain Assessment by Parents and Clinicians in Children Undergoing Orthopedic Surgery. *J. Trauma Nurs.* **2020**, *27*, 302–309. [[CrossRef](#)] [[PubMed](#)]
77. Schmand, B.; Lindeboom, J.; Schagen, S.; Heijt, R.; Koene, T.; Hamburger, H.L. Cognitive complaints in patients after whiplash injury: The impact of malingering. *J. Neurol. Neurosurg. Psychiatry* **1998**, *64*, 339. [[CrossRef](#)]
78. Greve, K.W.; Ord, J.S.; Bianchini, K.J.; Curtis, K.L. Prevalence of malingering in patients with chronic pain referred for psychologic evaluation in a medico-legal context. *Arch. Phys. Med. Rehabil.* **2009**, *90*, 1117–1126. [[CrossRef](#)]
79. Mittenberg, W.; Patton, C.; Canyock, E.M.; Condit, D.C. Base Rates of Malingering and Symptom Exaggeration. *J. Clin. Exp. Neuropsychol.* **2002**, *24*, 1094–1102. [[CrossRef](#)]
80. Simon, D.; Craig, K.D.; Gosselin, F.; Belin, P.; Rainville, P. Recognition and discrimination of prototypical dynamic expressions of pain and emotions. *Pain* **2008**, *135*, 55–64. [[CrossRef](#)]
81. Williams, A. Facial expression of pain: An evolutionary account. *Behav. Brain Sci.* **2002**, *25*, 439–455, discussion 455. [[CrossRef](#)]
82. Dirupo, G.; Garlasco, P.; Chappuis, C.; Sharvit, G.; Corradi-Dell'Acqua, C. State-Specific and Supraordinal Components of Facial Response to Pain. *IEEE Trans. Affect. Comput.* **2022**, *13*, 793–804. [[CrossRef](#)]
83. Prkachin, K.M.; Hammal, Z. Computer mediated automatic detection of pain-related behavior: Prospect, progress, perils. *Front. Pain Res.* **2021**, *2*, 788606. [[CrossRef](#)]
84. Price, W.N., II; Gerke, S.; Cohen, I.G. Potential Liability for Physicians Using Artificial Intelligence. *JAMA* **2019**, *322*, 1765–1766. [[CrossRef](#)] [[PubMed](#)]
85. Balthazar, P.; Harri, P.; Prater, A.; Safdar, N.M. Protecting Your Patients' Interests in the Era of Big Data, Artificial Intelligence, and Predictive Analytics. *J. Am. Coll. Radiol.* **2018**, *15*, 580–586. [[CrossRef](#)] [[PubMed](#)]
86. Martinez-Martin, N. What Are Important Ethical Implications of Using Facial Recognition Technology in Health Care? *AMA J. Ethics* **2019**, *21*, E180–E187. [[CrossRef](#)] [[PubMed](#)]
87. Fillingim, R.B.; King, C.D.; Ribeiro-Dasilva, M.C.; Rahim-Williams, B.; Riley, J.L., 3rd. Sex, gender, and pain: A review of recent clinical and experimental findings. *J. Pain* **2009**, *10*, 447–485. [[CrossRef](#)]
88. Wandner, L.D.; Scipio, C.D.; Hirsh, A.T.; Torres, C.A.; Robinson, M.E. The perception of pain in others: How gender, race, and age influence pain expectations. *J. Pain* **2012**, *13*, 220–227. [[CrossRef](#)]
89. Grouper, H.; Eisenberg, E.; Pud, D. More Insight on the Role of Personality Traits and Sensitivity to Experimental Pain. *J. Pain Res.* **2021**, *14*, 1837–1844. [[CrossRef](#)]
90. Littlewort, G.; Frank, M.; Lainscsek, C.; Fasel, I.; Movellan, J. Automatic Recognition of Facial Actions in Spontaneous Expressions. *J. Multimed.* **2006**, *1*, 22–35. [[CrossRef](#)]
91. Rudovic, O.; Pavlovic, V.; Pantic, M. Context-Sensitive Dynamic Ordinal Regression for Intensity Estimation of Facial Action Units. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *37*, 944–958. [[CrossRef](#)]
92. Kunz, M.; Seuss, D.; Hassan, T.; Garbas, J.U.; Siebers, M.; Schmid, U.; Schöberl, M.; Lautenbacher, S. Problems of video-based pain detection in patients with dementia: A road map to an interdisciplinary solution. *BMC Geriatr.* **2017**, *17*, 33. [[CrossRef](#)]
93. Taati, B.; Zhao, S.; Ashraf, A.B.; Asgarian, A.; Browne, M.E.; Prkachin, K.M.; Mihailidis, A.; Hadjistavropoulos, T. Algorithmic Bias in Clinical Populations—Evaluating and Improving Facial Analysis Technology in Older Adults with Dementia. *IEEE Access* **2019**, *7*, 25527–25534. [[CrossRef](#)]
94. Priebe, J.A.; Kunz, M.; Morcinek, C.; Rieckmann, P.; Lautenbacher, S. Does Parkinson's disease lead to alterations in the facial expression of pain? *J. Neurol. Sci.* **2015**, *359*, 226–235. [[CrossRef](#)] [[PubMed](#)]
95. Maza, A.; Moliner, B.; Ferri, J.; Llorens, R. Visual Behavior, Pupil Dilation, and Ability to Identify Emotions from Facial Expressions After Stroke. *Front. Neurol.* **2020**, *10*, 1415. [[CrossRef](#)] [[PubMed](#)]
96. Delor, B.; D'Hondt, F.; Philippot, P. The Influence of Facial Asymmetry on Genuineness Judgment. *Front. Psychol.* **2021**, *12*, 727446. [[CrossRef](#)] [[PubMed](#)]
97. Prkachin, K.; Hammal, Z. Automated Assessment of Pain: Prospects, Progress, and a Path Forward. In Proceedings of the 2021 International Conference on Multimodal Interaction, Montreal, QC, Canada, 18–22 October 2021; pp. 54–57.
98. Sedgwick, P.; Greenwood, N. Understanding the Hawthorne effect. *BMJ* **2015**, *351*, h4672. [[CrossRef](#)]
99. London, A.J. Artificial Intelligence and Black-Box Medical Decisions: Accuracy versus Explainability. *Hastings Cent. Rep.* **2019**, *49*, 15–21. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.