Review

A Review of Automatic Pain Assessment from Facial Information Using Machine Learning

Najib Ben Aoun [1,2]

1 Department of Information Technology, Faculty of Computing and Information, Al-Baha University, Al-Baha 65799, Saudi Arabia; nbenaoun@bu.edu.sa
2 REGIM-Lab: Research Groups in Intelligent Machines, National School of Engineers of Sfax (ENIS), University of Sfax, Sfax 3038, Tunisia

Abstract: Pain assessment has become an important component in modern healthcare systems. It aids medical professionals in patient diagnosis and providing the appropriate care and therapy. Conventionally, patients are asked to provide their pain level verbally. However, this subjective method is generally inaccurate, not possible for non-communicative people, can be affected by physiological and environmental factors and is time-consuming, which renders it inefficient in healthcare settings. So, there has been a growing need to build objective, reliable and automatic pain assessment alternatives. In fact, due to the efficiency of facial expressions as pain biomarkers that accurately expand the pain intensity and the power of machine learning methods to effectively learn the subtle nuances of pain expressions and accurately predict pain intensity, automatic pain assessment methods have evolved rapidly. This paper reviews recent spatial facial expressions and machine learning-based pain assessment methods. Moreover, we highlight the pain intensity scales, datasets and method performance evaluation criteria. In addition, these methods' contributions, strengths and limitations will be reported and discussed. Additionally, the review lays the groundwork for further study and improvement for more accurate automatic pain assessment.

Keywords: automatic pain assessment; pain intensity estimation; facial information; facial expressions; machine learning; deep earning

1. Introduction

Acute and chronic pain pose serious healthcare concerns, affecting millions of people worldwide and having an effect on quality of life. Indeed, for pain management and treatment to be successful, accurate pain evaluation is essential. Conventional techniques depend on self-reporting, which is subjective, can be biased by other environmental and psychological factors and is not possible for non-communicative patients, such as infants or people with cognitive disabilities. So, it is crucial to build an automatic pain assessment method that will help healthcare providers to precisely measure and monitor several types of pain including chronic and postoperative. This will also aid in providing the correct therapy and care and monitoring patient reaction to a medical treatment.

In recent years, numerous automatic pain estimation methods have been proposed [1–21]. They aim at recognizing the pain level using different modalities such as facial expressions [11–15,19], voice [22], human behavior (e.g., human activity, body movement, coordination and speed) [22,23] and physiological signals (e.g., ECG brain signal and heart rate) [24–26]. Nonetheless, the study of facial expressions is the most often used data source to forecast pain assessment. Face expressions reveal important information about an individual's level of discomfort because they are a normal and frequently unconscious reaction to it. In addition, the only specialized equipment needed for this non-invasive technique is a camera, which is a feature of every smartphone. Thus, many algorithms have been created [1–21] to precisely extract facial traits linked to pain, like biting of the lip, brow furrowing and clenching of the jaw.

In the past, techniques for estimating pain intensity depended on manually extracting pain-related information with the assistance of medical professionals [27–29]. However, with the emergence of machine learning techniques and their success in the computer vision and image processing fields [30–34], many machine learning-based methods have tackled the task of pain assessment from facial expressions [11–15,19,22]. These methods have achieved impressive results.

This review delves into the newly emerging field of machine learning and facial expressions for automatic pain assessment. Facial expressions are a natural and frequently unconscious reaction to pain; they can reveal important information about a person's level of pain. However, machine learning offers a viable path toward creating automated systems that can precisely identify and gauge the degree of pain based on facial features because of its capacity to understand intricate patterns.

In comparison with recent pain assessment review papers [35–43], this paper principally focuses on the effective machine learning and spatial facial expression-based pain assessment methods, reports more novel works (March 2018–May 2024) and categorizes the pain assessment works from the learning model perspective, which highlights the main contribution of each work. In addition, it offers a meta-analysis comparing the latest methods while evaluating them on the widely used pain assessment datasets employing the common performance evaluation criteria.

By providing a thorough review of the spatial facial expressions and machine learning-based pain assessment research landscape, this study aims to achieve the following:

- Highlight the limitations of self-reporting pain levels and emphasize the power of automated pain detection through spatial facial recognition and machine learning in healthcare settings.
- Present a background about the pain intensity scales, pain datasets and method performance evaluation criteria used in automatic pain assessment.
- Analyze the state-of-the-art spatial facial information and machine learning-based pain assessment methods to determine the areas in which their accuracies and resilience can be enhanced.
- Encourage the application of this cutting-edge technology in clinical settings for better pain management.

For the rest of this paper, in Section 2, we will provide the review methodology followed to collect the most relevant papers. In Section 3, we will give some background information on the pain datasets, pain intensity scales, and method performance evaluation criteria applied to automatic pain detection. Then, Section 4 presents a systematic overview of spatial facial expressions and machine learning-based pain assessment methods while highlighting their contributions, strengths and limitations. Furthermore, the results will be analyzed and interpreted, the automatic pain assessment challenges will be addressed and the limitations of this review will be presented in the discussion in Section 5. Finally, the conclusion, Section 6, will wrap up the paper with the pivotal findings along with recommendations for further pain assessment studies.

2. Review Methodology

2.1. Search Strategy

To conduct a deep and efficient review of the most recent and relevant research papers investigating the automatic pain assessment task from facial information using machine learning methods, a list of related keywords was built and used to collect high-level journal articles, book chapters and conference papers published in trustworthy databases such as IEEE Xplore, ACM Digital Library, Scopus, PubMed and Web of Science (WoS).

Indeed, pain assessment, facial information and machine learning-related keywords were used, mainly the following:

- For pain assessment, we used "pain assessment", "pain intensity", "pain estimation", "pain diagnosis" and "pain recognition" keywords.

- For facial information, we used "facial information", "facial expressions", "facial traits" and "faces" keywords.
- For machine learning, we used "machine learning", "deep learning" and "artificial intelligence" keywords.

After collecting the papers, they were scanned again to guarantee they accurately met this review's requirements.

2.2. Inclusion/Exclusion Criteria

This paper aims to investigate automatic pain assessment methods using static facial information and machine learning. So, only the recent works using static facial information as pain sources were selected while excluding the physiological, dynamic facial expressions, speech, self-reporting, behavioral/movement modalities and other indicators. In addition, the focus was only on the machine learning-based methods.

Additionally, to present a recent paper review, it was decided to preserve the original research and review articles published within the last six years (March 2018–May 2024). In addition, only the papers studying the pain of adults were collected while excluding the ones working on neonatal or infant pains. Moreover, mainly the works evaluated on the most commonly used and publicly available pain assessment datasets (UNBC–McMaster [44], BioVid database [45] and MIntPain database [46]) were retained (see Section 3.2).

2.3. Categorization Method

After filtering only the relevant papers presenting effective pain assessment methods, they were scanned to find an adequate categorization that would help the reader to obtain a clear understanding of their contributions and differences. So, based on the employed model, three main categories were identified : classical machine learning-based methods excluding the deep learning ones (referred to, in this paper, as "machine learning methods"), deep learning-based methods and hybrid model-based methods combining machine and/or deep learning methods (see Figure 1).
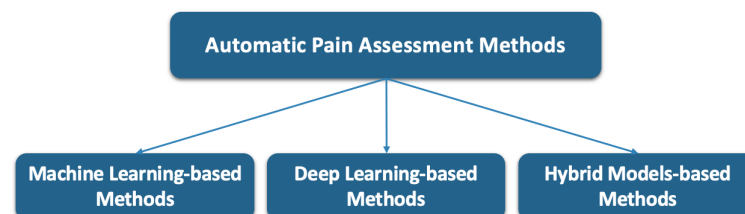

Figure 1. Automatic pain assessment methods.

The main contributions, results, advantages, limitations and suggested future directions were extracted from these research papers and used in developing and enriching our study. Furthermore, other review paper investigations [3,26,35–37,39,41] were leveraged to incorporate their recommended and relevant research extensions.

3. Background

A facial information-based pain assessment system is composed of three main phases, as shown in Figure 2.

The first phase consists of applying several pre-processing techniques to the input image, such as face detection, cropping and alignment, which leads to focusing only on the face while subtracting the background. Additionally, other image quality enhancement and denoising techniques can be applied, which will help in accurately extracting the facial features. In addition, image segmentation can be used to extract facial parts, which can help in extracting local facial features. Moreover, due to the need for large image data, especially for deep learning models, data augmentation can be employed where the image data are

limited. Furthermore, in cases where the dataset is imbalanced, it is crucial to conduct data balancing across all the classes to effectively train the learning model.
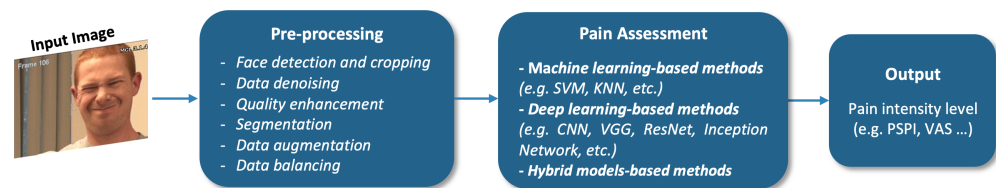

Figure 2. Facial information-based pain assessment flowchart.

Then, after pre-processing the image data, pain assessment models are applied. In this phase, facial features are extracted and used to learn the model to effectively characterize and classify the pain level. Indeed, three machine learning-based methods can be applied. The conventional machine learning-based methods [1–4] manually extract handcrafted facial features, then classify them using conventional machine learning algorithms such as Support Vector Machines (SVMs) [47] and K-Nearest Neighbors (KNNs) [48]. However, deep learning-based methods [5–13] use deep learning models (e.g., Convolutional Neural Networks (CNNs) [49], Residual neural network (ResNet) [50], VGG [51] and Inception Network [52]) for both feature extraction and model-learning tasks. Recently, hybrid model-based methods [14–21], which combine machine learning and/or deep learning have been proposed for the pain assessment task. These methods follow different ensemble learning strategies to leverage the advantages of different learning models.

Finally, once the facial features are extracted and classified into several pain classes, the pain intensity level is determined. Pain can be categorized into several pain levels based on its severity. That is why different pain scales have been suggested to accurately identify the pain level is person is experiencing.

3.1. Pain Intensity Scales

Pain is perceptible and may be precisely measured. Different subjective pain intensity scales have been introduced (see Figure 3) to help in communicating the pain degree between the patient and the healthcare providers, which have helped them to better understand a patient's pain and to develop an appropriate treatment plan. These subjective pain measurement tools have been a major supporter of the development of automatic pain assessment methods. The most commonly used scales by the recent automatic pain assessment methods are the Visual Analog Scale (VAS) [53] and the Prkachin and Solomon Pain Intensity (PSPI) score [54].
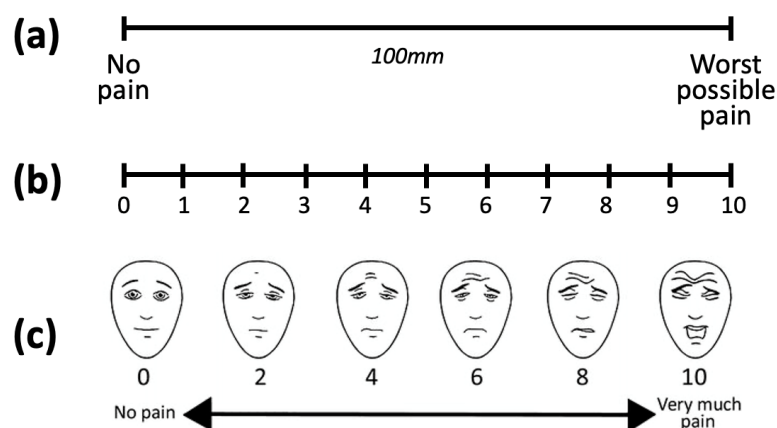

Figure 3. Pain intensity Scales. (**a**) VAS (**b**) NRS (**c**) FPS-R.

The VAS uses a 10-unit scale to rate the pain degree. To rate, a handwritten mark must be made along a line of 100 mm, which represents a continuum going from left to

right between "no pain" (0 mm) and "the worst pain imaginable" (100 mm). However, for an appropriate pain intensity measurement, it was advised to standardize the scale into 10 levels (equally spaced by 10 mm) where a patient chooses from level 0 (no pain) to level 10 (unbearable pain). Additionally, this scale can be scaled up to four sections: no pain (0–4 mm), mild pain (5–44 mm), moderate pain (45–74 mm) and severe pain (75–100 mm).

Another common metric for estimating pain intensity at the frame-level ground truth is the PSPI score. In fact, it uses the Facial Action Coding System (FACS) [55] that consists of 44 facial Action Units (AUs), which are the smallest observable contractions of facial muscles. However, PSPI focuses on six specific actions associated with pain (see Figure 4): *AU*4 (Brow Lowerer), *AU*6 (Cheek Raiser), *AU*7 (Lid Tightener), *AU*9 (Nose Wrinkler), *AU*10 (Upper Lip Raiser) and *AU*43 (Eyes Closed). After measuring these six AUs on a scale from 0 to 5 (except *AU*43, which is measured as 0 when the eye is opened or 1 when the eye is closed), PSPI performs a linear combination between their intensities following Equation (1). Originally, PSPI scores were set on a scale ranging from 0 to 16 [54]; although, in several research studies, they are often standardized to only four levels (0, 1, 2 and $\geq$3) [56].

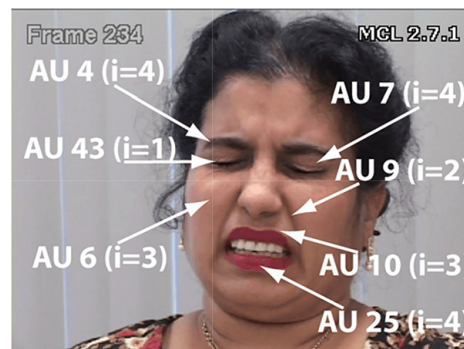$$\text{PSPI pain} = AU4 + \max(AU6, AU7) + \max(AU9, AU10) + AU43 \qquad (1)$$

Figure 4. Example of a painful face from the UNBC–McMaster database with the PSPI-related AUs intensities [56]. Here, PSPI = 4 + Max (3,4) + Max (2,3) + 1 = 12.

More other less-used pain intensity scales have been provided, such as the Numeric Rating Scale (NRS) [57], which is a simple 0-to-10 scale, where 0 means "no pain" and 10 means "the worst pain imaginable", as well as the Faces Pain Scale-Revised (FPS-R) [58], which help the patient to define their pain by choosing between six cartoon faces that range from happy (no pain) to crying (worst pain). This scale is often used for children or people having difficulty using a numeric scale.

3.2. Publicly Accessible Pain Assessment Datasets

The availability of many face images and/or video datasets for pain assessment has driven recent advances in the field of automatic pain assessment. The UNBC–McMaster Shoulder Pain Expression Archive Database (UNBC–McMaster) [44] is the one of the most widely utilized of these datasets [1–17,17–21]. This dataset was gathered from 25 adult participants suffering from shoulder pain, which form 48,398 RGB frames issued from 200 variable-length videos (see details in Table 1). In this database, images are mainly labeled into 17 PSPI levels (0–16) and 11 VAS levels (0–10). However, for some studies, the PSPI level is normalized into four to six levels [13,19] and four VAS levels. Sample images from the UNBC–McMaster database with the corresponding PSPI levels can be seen in Figure 5. Indeed, the UNBC–McMaster dataset, like many other image datasets, suffers from the imbalanced data problem where more than 80% of the dataset has a PSPI score of zero (meaning "no pain") [3]. So, data balancing has been conducted in many works [3,16,19] by applying an under-resampling technique to decrease the no-pain

class. As a result, the UNBC–McMaster dataset size was reduced from 48,398 frames to 10,783 frames.

Table 1. Pain assessment datasets details.

| Attribute	UNBC–McMaster (2011) [44]	BioVid Database (2013) [45]	MIntPain Database (2018) [46]
Number of subjects	25 adults with shoulder pain	90 subjects (87 are available)	20 subjects
Subject type	Self-identified pain patient	Healthy volunteers	Healthy volunteers
Pain nature	Natural shoulder pain	Inducted heat pain	Stimulated electrical pain
Pain levels	0–16 (PSPI) and 0–10 (VAS)	1–4 (Stimuli)	0–4 (Stimuli)
Modalities	RGB	RGB	RGB, Depth and Thermal
Dataset size	48,398 frames issued from 200 variable-length videos	17,300 5 s videos (25 fps)	9366 variable-length videos with 187,939 frames
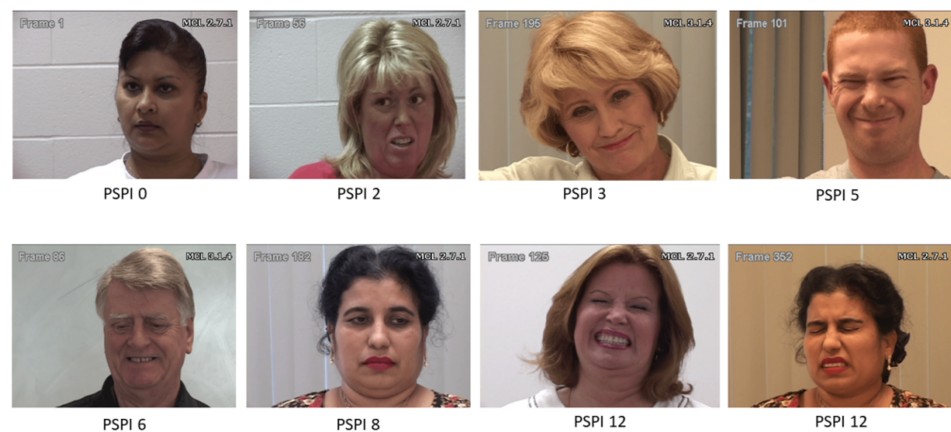
Figure 5. Sample images from the UNBC–McMaster database with their corresponding PSPI levels [3].

Furthermore, another well-used dataset [6,8,20,45,59] is the BioVid Heat Pain dataset [45,59], which was formed by collecting 17300 RGB videos from 87 subjects of 5 s each with a frame rate of 25 fps. An inducted heat pain was the cause of the pain represented in these videos. Additionally, four pain levels are present in this dataset (from 1 to 4).

In addition, another less-used pain assessment database is the Multimodal Intensity Pain (MIntPAIN) database [16,46], which contains 9366 variable-length videos having 187,939 frames issued from 20 subjects. These videos are of three modalities: RGB, depth and thermal of the same video sequences. The pain elicited using controlled electrical stimulation is labeled into five pain levels (from 0 to 4).

More pain assessment datasets have been developed such as X-ITE [60] and EmoPain [61]. Yet, their use to evaluate pain assessment methods is still limited and many of them lack diversity with regard to subject age, gender as well as poses, occlusions and lighting conditions. Data balancing is also required for some datasets.

3.3. Criteria for Performance Evaluation

To efficiently assess and compare the different automatic pain assessment methods, it is crucial to evaluate them on the same benchmarks while adopting common performance evaluation criteria based on accurate pain intensity levels. In fact, the two main performance evaluation criteria used for pain assessment methods are classification accuracy [11] and Mean Square Error (MSE) [3], which measures how accurate a model prediction is. A learning model indeed strives for both high accuracy and low MSE.

Following Equation (2), the accuracy metric compares the pain intensities' ground truth (i.e., PSPI, VAS scores, etc.) of the test samples to their pain intensity predicted by the method:

$$\text{Accuracy} = \frac{\text{\# of predicted test samples with correct pain intensities}}{\text{Total number of test samples}} \quad (2)$$

However, as it can be seen from Equation (3), the MSE measures the average squared difference between the predicted values ($\hat{y}_i$) and the actual values (y_i):

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \quad (3)$$

4. Facial Information-Based Pain Assessment Methods

One of the most trustworthy data sources for techniques estimating pain severity is facial expressions. In fact, brow furrowing, jaw clenching, lip biting, the degree of eye closure and other facial expressions are often used to convey a patient's level of pain. Thus, it is promising to develop a system that can precisely identify the pain intensity level by extracting the most relevant pain-related face information. Facial expressions can be either dynamic (derived from the video's temporal dimension) or static (derived from the face image). However, the main focus of this review will be on recent studies that estimated patient pain using static facial expressions.

Indeed, with the success of machine learning algorithms in learning and predicting the pain intensity degree, static facial expression-based methods have largely deployed them in recent years. So, as mentioned in the methodology in Section 2 and based on the learning methodology, these methods are categorized into three categories (see Figure 1): machine learning, deep learning and hybrid model-based methods.

In general, most of the methods have followed the flowchart illustrated in Figure 2 with different inputs, pre-processing techniques and model architectures.

4.1. Machine-Learning-Based Methods

The baseline of classical machine learning models (e.g., SVM, KNN, etc.) is to classify pre-extracted hand-crafted features. The main limitation of these models is that their performances are highly dependent on the quality and the pertinence of the extracted features, which require domain expertise. However, they have allowed the automation of image classification and computer vision tasks [39]. For the spatial facial expression-based pain assessment task, machine learning-based methods [44,56] have been widely used during the last fifteen years due to their encouraging results (see Table 2), whereas, with the emergence of more effective models like deep learning and hybrid models, their use for this task has been recently limited [1–4].

For instance, in [1], a relatively shallow CNN architecture with three convolutional layers was proposed. This computationally efficient network with few parameters has obtained an accuracy of 93.34% when evaluated on the UNBC–McMaster dataset.

Additionally, in [2], a hierarchical network architecture was proposed, where two per-frame feature modules are employed. The first module extracts low-level features from the image patches and assembles them using second-order pooling. The second module extracts deep learning features from the image using a deep CNN. Then, the output face representations of the two modules are weighted and combined to form a holistic representation that boosts the pain estimation process. The resulting feature is afterward classified with a linear L2-regularized L2-loss Support Vector Regressor (SVR) to predict the pain intensity. An MSE of 1.45 was obtained when evaluated on the UNBC–McMaster dataset.

Furthermore, transfer learning is increasingly adopted for image classification works. The main idea is to pre-train a model for a specific task and acquire knowledge from it. Then, this model is re-used for another task where it is fine-tuned according to its data, which will lead to improving the model performance and decreasing the training time. For instance, in [3], a pre-trained DenseNet-161 model [62] was retrained on the UNBC–McMaster dataset. Then, the features were extracted from ten middle layers of the

fine-tuned network and used as inputs to a Support Vector Regression (SVR) classifier. The evaluation of this model on the UNBC–McMaster dataset obained an MSE of 0.34.

Table 2. Summary of machine learning-based pain assessment methods using spatial facial expressions and their performances on UNBC–McMaster dataset. "ACC" refers to accuracy while "MSE" refers to Mean Square Error.

Method	Feature Extraction	Machine Learning Model	Main Contribution	Performance on UNBC–McMaster
Semwal and Londhe [1]	CNN-based deep features	3-layer CNN	Optimize the model parameters	ACC: 93.34%
Yang et al. [2]	Low-level image patches features + CNN-based deep learning features	SVR	A hierarchical network architecture with two feature modules to extract low-level and deep learning features.	MSE: 1.45
Morabit et al. [3]	Deep features extracted with a fine-tuned DenseNet-161 model	SVR	Adopting the transfer learning technique to fine-tune a pre-trained DenseNet-161 model for feature extraction	MSE: 0.34
Barua et al. [4]	Deep features extracted with a fine-tuned DarkNet19 and selected with INCA	KNN	Using a pre-trained DarkNet19 model to generate deep features that were optimized by the INCA algorithm	ACC: 95.57%

Moreover, in [4], a KNN-based pain assessment method was proposed. This method extracts facial features from face patches using a pre-trained DarkNet19 [63] model and selects the most informative features with the iterative neighborhood component analysis (INCA) technique. Then, the resulting features are classified with the KNN algorithm to efficiently predict the pain intensities. This method has achieved a pain intensity estimation accuracy of 95.57% on the UNBC–McMaster dataset.

However, despite the promising pain assessment results gained by using the machine learning-based methods [1–4], most of them still need more improvement to reach satisfactory results. A very encouraging alternative is the use of deep learning models which have outperformed the machine learning models in pain assessment tasks [5–13].

4.2. Deep Learning-Based Methods

According to the collected research papers and as reported in [40], deep learning models have rapidly taken the lead for automatic pain assessment since 2018. This is principally due to the success of deep learning models for data classification and the availability of large pain datasets. Most deep learning-based pain assessment studies [7–10,12] have utilized a variation of the successful Convolutional Neural Network (CNN) [49] (see Table 3). Many of them have exploited more sophisticated recent deep learning models (e.g., ResNet, DenseNet and InceptionV3) [5,6,11,13].

In [12], an improved version of [1] was proposed, where a compact shallow CNN model for pain severity assessment (SPANET) including a false positive reduction method were presented. In an unrestricted hospital setting and on the UNBC–McMaster dataset, the proposed SPANET method demonstrated strong performance, with a pain intensity estimation accuracy of 97.48% on UNBC–McMaster .

In addition, to overcome the possible limitations in the original pain ground truth according to the subjects themselves or data annotation experts, [5] employed seven experts to re-annotate the UNBC–McMaster dataset. Then, they translated the frames to illumination-invariant 3D space using the multidimensional scaling method in order to feed a pre-trained AlexNet [49] model. This method obtained an accuracy of 80% on the UNBC–McMaster dataset.

Table 3. Summary of deep learning-based pain assessment methods using spatial facial expressions and their performances on UNBC–McMaster and BioVid pain datasets. "ACC" refers to accuracy while "MSE" refers to Mean Square Error.

Method	Deep Learning Model	Main Contribution	Performance	
			UNBC–McMaster	BioVid
Casti et al. [5]	Pre-trained AlexNet	Re-annotating the dataset with seven experts to improve the ground truth and translating the frames to illumination-invariant 3D space using the multidimensional scaling	ACC: 80%	-
Dragomir et al. [6]	ResNet	Hyper-parameters optimization + data augmentation	-	ACC: 36.6%
Xi et al. [7]	9-layer CNN	Including an attention mechanism to focus on the region-related pain	ACC: 51.1%	-
Xi et al. [8]	CNN	Incorporating locality-aware and identity-aware modules.	ACC: 89.17%	ACC: 40.4%
Rathee et al. [9]	CNN	Efficiently initializing CNN parameters with an MAML++ module.	ACC: 90.3%	-
Cui et al. [10]	Multi-scale regional attention network (MSRAN)	Integrating self-attention and relation attention modules to emphasize the pain-related face parts and their inter-relationships	ACC: 91.13%	-
Karamitsos et al. [11]	Customized and Deeper VGG16	Feeding the model with effectively pre-processed images (e.g., gray-scaling, histogram equalization, face detection, image cropping, mean filtering and normalization).	ACC: 92.5%	-
Semwal and Londhe [12]	SPANET	Shallow CNN including a false positive reduction strategy.	ACC: 97.48%	-
Alghamdi and Alaghband [13]	InceptionV3	The convolutional block is frozen and a shallow CNN is used instead of the prediction layer.	ACC: 99.1%	-

Moreover, in [7], to focus on the pain-related face regions, an attention mechanism was incorporated in a nine-layer CNN with a Softmax loss function to assign different weights for each region according to its pain expressiveness. The integration of this attention mechanism has improved the prediction accuracy to reach 51.1%. Afterward, a more powerful multi-task pain assessment architecture was proposed in [8], which consists of a locality and identity-aware network (LIAN). This network, first, conducts a dual-branch locality-aware module to highlight the information about the facial regions connected to pain. Then, the identity-aware module (IAM) is employed to decouple the pain assessment and the identity recognition tasks, which could achieve identity-invariant pain assessment. As it can be seen in Table 3, the outcomes demonstrate that this approach delivers a good performance on UNBC–McMaster (an accuracy of 89.17%). Furthermore, a similar pain-locality-based method was suggested by Cui et al. [10]. They created a multi-scale regional attention network (MSRAN) that uses adaptive learning to identify the degree of pain by capturing information about the facial pain regions. To highlight the facial areas associated with pain and their relationship, a self-attention and a relation attention module were incorporated. With this method, an accuracy of 91.13% in pain estimation was obtained on the UNBC–McMaster dataset.

Additionally, evaluating the UNBC–McMaster dataset, Rathee et al. [9] have suggested a CNN-based pain assessment approach with an estimation accuracy of 90.3%. This technique uses an improved version of the Model Agnostic Meta-Learning (MAML++) module, which aids in efficiently initializing the CNN model parameter so that it can rapidly converge to the optimal performance with the fewest number of images and the least amount of time. When determining the degree of pain in fresh subjects, this method has shown good results. In addition, a customized and deeper CNN model based on the VGG16 architecture was proposed in [11]. Using the UNBC–McMaster dataset, this modified VGG16 model yields a 92.5% accuracy in estimating pain intensity. To reach this compelling result, this model was fed with pre-processed face images that included gray-scaling, histogram equalization, face detection, image cropping, mean filtering and normalization.

Recently, on the UNBC–McMaster dataset, the method described in [13] produced the best-reported pain intensity estimation accuracy (99.1%). In this method, two concurrent deep CNNs (InceptionV3 models [64] with an SGD optimizer) are used to extract the relevant features while freezing all convolutional blocks and replacing the classifier layer with a shallow CNN. The resulting outputs are concatenated and sent to a dense layer, then to a fully connected layer to classify the data. This makes this architecture a deep learning architecture.

Compared to the spatial facial expression-based pain assessment works evaluated on the UNBC–McMaster dataset, fewer methods have been examined on the BioVid dataset [6,8] (see Table 3). Apart from the previously reported accuracy result (40.4%) reached in [8], Dragomir et al. [6] presented a ResNet-based pain intensity estimation model, optimizing the model hyper-parameters and trying several strategies for data augmentation. On the BioVid dataset, the model's accuracy in estimating pain intensity was found to be 36.6%.

Based on these reported methods, it can be concluded that deep learning models have performed exceptionally well when fed with static facial expressions for the task of estimating pain severity (see Table 3). However, more recent deep learning models have to be evaluated for pain assessment and more model validation has to be conducted on more challenging pain datasets. In addition, it should be mentioned that several model fusion attempts have been proposed to get the most benefit from the combined models [14–21].

4.3. Hybrid Model Methods

Due to the success of machine learning and deep learning models in the automatic pain assessment task, several ensemble learning methods [14–21] have been developed in an attempt to combine the efficiency of many models (see Table 4).

Table 4. Summary of the hybrid model-based pain assessment methods using spatial facial expressions and their performances on UNBC–McMaster , BioVid and MIntPain pain datasets. "ACC" refers to accuracy while "MSE" refers to Mean Square Error.

Method	Hybrid Model	Main Contribution	Performance		
			UNBC–McMaster	BioVid	MIntPain
Bargshady et al. [16]	EDLM: three independent CNN-RNN deep learners with different weights	Extract face features with a fine-tuned VGGFace and optimize them with the PCA algorithm.	ACC: 86%	-	ACC: 92.26%
Huang et al. [17]	Pain-awareness multi-stream convolutional neural network with 4 CNNs	4 pain related-face region features are extracted with 4 CNNs, weighted according to their contribution to the pain expression and classified to estimate the pain intensity	ACC: 88.19%	-	-
Huang et al. [18]	Hierarchical Deep Network (HDN)	Two scale branches are performed to extract local facial patches features with a multi-stream CNN and their inter-dependencies with a multi-task learning technique	MSE: 0.49	-	-
Alghamdi and Alaghband [20]	SAFEPA: 2 fine-tuned concurrent and coupled CNNs	Placing greater emphasis on the upper part of the face to extract pose-invariant face features which are coupled with global face feature leading to better pain assessment.	ACC: 89.93%	ACC: 33.28%	-
Semwal and Londhe [14]	ECCNET = VGG-16 + MobileNet + GoogleNet	Averaging the prediction of three deep learning models	ACC: 93.87%	-	-
Ye et al. [19]	Parallel CNN framework with regional attention: VGGNet + ResNet	Used VGGNet + ResNet to extract the face features with regional attention and the SogtMax algorithm to classify them.	ACC: 95.11%	-	-
Semwal and Londhe [15]	3 CNNs: VGG-TL + ETNet + DSCNN	Used multiple neural networks with high-level spatial features with local and global geometric cues.	ACC: 96%	-	-
Sait and Dutta [21]	Stacking XGBoost and CatBoost models with an SVM as meta-learner	Relevant spatial features are extracted with fine-tuned ShuffleNet V2 model + Application of class activation map and fusion feature approaches	ACC: 98.7%	-	-

Semwal and Londhe [14] proposed an Ensemble of Compact Convolutional Neural Networks (ECCNET) model that combines three CNNs (VGG-16 [51], MobileNet [65] and GoogleNet [66]) while aggregating their predictions using the average ensemble rule. The experiments demonstrate that merging the CNNs leads to better classification performance than using them individually. As a result, an accuracy of 93.87% on the UNBC–McMaster dataset was reached. Afterward, the CNN models fusion technique was re-used in a follow-up study by the same authors [15] to combine three CNNs into a single model: a cross-dataset Transfer Learning VGG-based model (VGG-TL), Entropy Texture Network (ETNet) and Dual Stream CNN (DSCNN). In fact, three distinct image features (RGB features, an entropy-based texture feature and a complimentary feature that was learned from both of them) were combined and fed into the suggested pain assessment model to improve its generalization. Furthermore, a range of data augmentation methodologies were applied to the dataset in an effort to mitigate the issue of overfitting in the model. As a result, using the same UNBC–McMaster dataset, pain level detection accuracy increased by 2.13% to reach 96% compared to [14].

Furthermore, in [16], an ensemble deep learning model (EDLM) for pain assessment was proposed. This model uses a fine-tuned VGGFace to extract facial features followed by the Principal Component Analysis (PCA) algorithm to reduce the feature dimension while retaining the most informative of them, which helps in decreasing the training time. Then, three independent CNN-RNN deep learners with different weights are used to classify the extracted facial features. This model achieved a satisfactory accuracy of 86% on the UNBC–McMaster dataset and 92.26% on the MIntPain dataset.

Several methods have focused on the face regions that are mostly affected by pain while neglecting image background information, which disturbs pain intensity detection. So, after detecting the face regions (left eye, right eye, nose and mouth) Huang et al. [17] applied a multi-stream CNN of four sub-CNNs to extract the features from these four face regions. Then, these features were classified to estimate the pain intensity while assigning a learned weight for each of them, proportionally to their contribution to the pain expression. This method was evaluated on the UNBC–McMaster dataset and given an accuracy of 88.19%. Afterwards, in a further study [18], a hierarchical deep network (HDN) architecture was described. Within HDN, two scale branches are implemented where a region-wise branch is intended to extract characteristics from face image regions related to pain while a global-wise branch investigates the inter-dependencies of pain-associated areas. Indeed, in the region-wise branch, a multi-stream CNN is used to extract local features while, in the global-wise branch, a pre-trained CNN is used to extract holistic features from the face. Additionally, a multi-task learning technique is used in the global-wise branch to identify action units and estimate pain intensity. Ultimately, a decision level fuses the outputs of two branches' pain estimation. In fact, it is empirically demonstrated that the proposed HDN performs satisfactorily and provides an MSE of 0.49 when evaluated on the UNBC–McMaster dataset.

Likewise, Ye et al. [19] proposed a parallel CNN framework with regional attention focusing on the most important pain-sensitive face regions. Their method merged a VGGNet [51] and a ResNet [50] model to extract the facial feature and a SoftMax classifier to classify them. The method has obtained an accuracy of 95.11% on the UNBC–McMaster dataset. In addition, in [20], a model that overcomes the challenges posed by full left and right profiles was proposed. This model utilizes Sparse Autoencoders (SAEs) to reconstruct the pain-affected upper-face part from the input image. Then, two pre-trained concurrent and coupled CNNs are fed the constructed upper face part as well as the original image. Indeed, this Sparse Autoencoders for Facial Expressions-based Pain Assessment (SAFEPA) approach produces better identification performance by placing greater emphasis on the upper part of the face. Furthermore, SAFEPA's architecture makes use of CNNs' advantages while also taking into account differences in head positions, which removes the requirement for pre-processing steps necessary for face detection and upper-face extraction in other models. Using the widely established UNBC–McMaster dataset, SAFEPA achieves a good accuracy of 89.93% while reaching an accuracy of 33.28% on the BioVid dataset (see Table 4).

More recently, Sait and Dutta [21] have proposed an ensemble learning model with a ShuffleNet V2 model [67], which is fine-tuned for feature extraction through the application of class activation map and fusion feature approaches. Then, following a stacking ensemble learning strategy, XGBoost and CatBoost are used as base models followed by an SVM as a meta-learner to predict the pain intensities. As a result, an accuracy of 98.7% on the UNBC–McMaster dataset proves the reliability of the proposed method and the possibility that it can be deployed in healthcare centers.

Table 4 illustrates how, by aggregating the model performances, hybrid model-based techniques are effectively challenging deep learning-based techniques. More work is yet required to effectively combine the model's advantages in order to gain a performance advantage without increasing method complexity or calculation time.

5. Discussion

Based on our search for spatial facial expression-based pain assessment methods and after meticulous paper scanning and filtering, twenty-one papers were selected: three machine learning-based methods, ten deep learning-based methods and eight hybrid model-based methods. As previously mentioned in Section 1, different modalities have been used for pain assessment, while the facial expressions one is the most efficient of them. That is what led us to focus on facial expression-based methods and, more precisely, on the approaches using spatial face information.

5.1. Result Analysis

After studying the three learning approaches, we can conclude that deep learning models [5–13] are more effective (see Table 3) in comparison with the classical machine learning models [1–4] (see Table 2). This can be explained by the power of deep learning models to extract the most relevant features, classify them and coordinate between the feature extraction and classification parts through the backpropagation process. In addition, deep learning models have leveraged the availability of large pain datasets and efficient computational capacities. As well, the hybrid model performances [14–21] have proven the success of merging several deep and machine learning techniques for pain assessment (see Table 4). The promising results show that the hybrid models seriously compete with the DL-based models.

In addition, it was noticed that several strategies and techniques have significantly improved the methods' efficiencies. Indeed, various methods [2,7,8,10,17–20] have focused on the pain-related face parts to extract the most pertinent pain features. However, many methods [3–5,15,16,18,20,21] got the benefit of the transfer learning strategy and utilized pre-trained models to speed up the training process and obtain better model performances. Moreover, a local feature relations attention module was incorporated in a variety of methods [10,18] to include the feature relationship information into the model account. Furthermore, for some methods, data augmentation was employed to enlarge the data size [6,15] and the dataset pain ground truth was enhanced by means of experts [5], which led to the increase in the accuracy results. Additionally, pre-processed images were supplied to the model in [11] to aid in the extraction of the most valuable pain features.

Additionally, several model-boosting techniques were incorporated to enhance the model performance such as the MAML++ algorithm for efficient model weight initialization [9], the false positive reduction technique [12] as well as the INCA [4] and the PCA [16] algorithms, which are used for feature selection.

5.2. Automatic Pain Assessment Challenges

Actually, several non-related pain factors can affect a patient's facial expressions, which lead to a mistaken pain intensity measurement, such as the external environment (pain distraction factors, weather, etc.), psychological factors, ethnicity, gender, region, patient sensitivity to pain (central sensitization), wondering/astonishment, previous experiences or even painkillers/drugs in medical settings or intensive care. This will make the

development of a generalizable pain intensity estimation system a difficult task. So, it is essential to consider the specific context in which the automated pain assessment systems will be used to ensure high accuracy and avoid these confounding effects.

In addition, for the same pain intensity and depending on the pain types (postoperative, acute, chronic, etc.), the patient's facial expressions can be different. Indeed, in chronic pain, which is more challenging than acute pain, the patient gets used to pain so their facial expressions are less intense. So, a pain type-customized system may be a promising solution.

Furthermore, more studies have to be done to encourage the use of automatic pain assessment technology in clinical settings and analyze its cost-effectiveness.

5.3. Limitations of This Review

Through our paper search methodology, we tried to collect all recent and relevant papers from many trustworthy research papers databases. Then, we did our best to exhibit the current state of research in automated pain assessment from spatial facial expressions using machine learning methods, identifying trends, capabilities, limitations, potential healthcare applications, and knowledge gaps. However, we have not succeeded in accessing five relevant papers because of university subscription limitations.

Moreover, it was not easy to conduct a deep comparative study between the reviewed studies since they were evaluated on different pain datasets (or a subset of a dataset) with varying pain intensity levels and used several performance evaluation criteria and different cross-validation techniques. In addition, most of them did not provide sufficient details such as the use of data augmentation, validation strategies (cross-validation, K-fold, etc.) and model parameter optimization techniques.

6. Conclusions

This study was designed to review the recent spatial facial expression-based pain assessment methods using machine learning models in their broad meaning. It is an attempt to support the ongoing efforts in leveraging machine learning technologies in the healthcare field, mainly for pain assessment and management. Indeed, several research works were reviewed and analyzed. Their very promising capabilities have confirmed that the automation of the pain assessment task is essential and that they can be employed for broader real-time applications in medical diagnosis and health informatics areas.

However, despite the power of automatic deep learning and hybrid model-based pain assessment methods, they still need more effort to improve model efficiency. Indeed, since model fusion has given promoting results, it would be beneficial to combine several well-performing models (e.g., the visual transformer) following a hierarchical strategy while keeping the architecture simple and the training time acceptable and validate it on larger facial image datasets, accurately labeled and collected from more subjects of diverse ethnicities. In addition, it is advised to effectively leverage the already used method enhancement techniques (pain-related features extraction, incorporating feature interdependencies, feature selection, image pre-processing, transfer learning, data augmentation, model parameter optimization, etc.).

Additionally, another future research direction is to exploit many pain modalities (facial, voice, physiological, behavioral, etc.), which may allow other pain-related tasks such as pain location and cause recognition. Furthermore, for long-term pain scenarios (e.g., chronic pain), it is encouraging to include the temporal dimension for the facial expressions to extract the facial dynamics, since spatio-temporal facial features are more expressive for pain. One instance is that pain-related face landmarks can be used to form a spatio-temporal graph or several graphs for each pain-related face region. So, the graph neural network (GNN) model [68] may be used and trained with these graphs for an accurate pain intensity estimation.

Furthermore, it is anticipated that explainable AI systems will be created to aid in decision-making so that medical professionals can better interpret and control pain.

References

1. Semwal, A.; Londhe, N.D. Automated Pain Severity Detection Using Convolutional Neural Network. In Proceedings of the 2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS), Belagavi, India, 21–22 December 2018; pp. 66–70. [CrossRef]
2. Yang, R.; Hong, X.; Peng, J.; Feng, X.; Zhao, G. Incorporating high-level and low-level cues for pain intensity estimation. In Proceedings of the 2018 24th International Conference on Pattern Recognition (ICPR), Beijing, China, 20–24 August 2018; pp. 3495–3500. [CrossRef]
3. El Morabit, S.; Rivenq, A.; Zighem, M.E.N.; Hadid, A.; Ouahabi, A.; Taleb-Ahmed, A. Automatic Pain Estimation from Facial Expressions: A Comparative Analysis Using Off-the-Shelf CNN Architectures. *Electronics* **2021**, *10*, 1926. [CrossRef]
4. Barua, P.; Baygin, N.; Dogan, S.; Baygin, M.; Arunkumar, N.; Fujita, H.; Tuncer, T.; Tan, R.; Palmer, E.; Azizan, M.; et al. Automated detection of pain levels using deep feature extraction from shutter blinds-based dynamic-sized horizontal patches with facial images. *Sci. Rep.* **1022**, *12*, 17297. [CrossRef] [PubMed]
5. Casti, P.; Mencattini, A.; Comes, M.C.; Callari, G.; Di Giuseppe, D.; Natoli, S.; Dauri, M.; Daprati, E.; Martinelli, E. Calibration of Vision-Based Measurement of Pain Intensity with Multiple Expert Observers. *IEEE Trans. Instrum. Meas.* **2019**, *68*, 2442–2450. [CrossRef]
6. Dragomir, M.C.; Florea, C.; Pupezescu, V. Automatic Subject Independent Pain Intensity Estimation using a Deep Learning Approach. In Proceedings of the International Conference on e-Health and Bioengineering (EHB), Iasi, Romania, 29–30 October 2020; pp. 1–4. [CrossRef]
7. Xin, X.; Lin, X.; Yang, S.; Zheng, X. Pain intensity estimation based on a spatial transformation and attention CNN. *PLoS ONE* **2020**, *15*, e0232412. [CrossRef]
8. Xin, X.; Li, X.; Yang, S.; Lin, X.; Zheng, X. Pain expression assessment based on a locality and identity aware network. *IET Image Process.* **2021**, *15*, 2948–2958. [CrossRef]
9. Rathee, N.; Pahal, S.; Sheoran, P. Pain detection from facial expressions using domain adaptation technique. *Pattern Anal. Appl.* **2021**, *25*, 567–574. [CrossRef]
10. Cui, S.; Huang, D.; Ni, Y.; Feng, X. Multi-Scale Regional Attention Networks for Pain Estimation. In Proceedings of the International Conference on Bioinformatics and Biomedical Technology (ICBBT), Xi'an, China, 21–23 May 2021; pp. 1–8. [CrossRef]
11. Karamitsos, I.; Seladji, I.; Modak, S. A Modified CNN Network for Automatic Pain Identification Using Facial Expressions. *J. Softw. Eng. Appl.* **2021**, *14*, 400–417. [CrossRef]
12. Semwal, A.; Londhe, N. S-PANET: A Shallow Convolutional Neural Network for Pain Severity Assessment in Uncontrolled Environment. In Proceedings of the IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC), Virtual, 27–30 January 2021; pp. 800–806. [CrossRef]
13. Alghamdi, T.; Alaghband, G. Facial Expressions Based Automatic Pain Assessment System. *Appl. Sci.* **2022**, *12*, 6423. [CrossRef]
14. Semwal, A.; Londhe, N. ECCNet: An Ensemble of Compact Convolution Neural Network for Pain Severity Assessment from Face images. In Proceedings of the International Conference on Cloud Computing, Data Science and Engineering (Confluence), Noida, India, 28–29 January 2021; pp. 761–766. [CrossRef]
15. Semwal, A.; Londhe, N. MVFNet: A multi-view fusion network for pain intensity assessment in unconstrained environment. *Biomed. Signal Process. Control.* **2021**, *67*, 102537. [CrossRef]
16. 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]
17. Huang, D.; Xia, Z.; Li, L.; Wang, K.; Feng, X. Pain-awareness multistream convolutional neural network for pain estimation. *J. Electron. Imaging* **2019**, *28*, 043008. [CrossRef]
18. Huang, D.; Xia, Z.; Li, L.; Ma, Y. Pain estimation with integrating global-wise and region-wise convolutional networks. *IET Image Process.* **2022**, *17*, 637–648. [CrossRef]
19. Ye, X.; Liang, X.; Hu, J.; Xie, Y. Image-Based Pain Intensity Estimation Using Parallel CNNs with Regional Attention. *Bioengineering* **2022**, *9*, 804. [CrossRef]
20. Alghamdi, T.; Alaghband, G. SAFEPA: An Expandable Multi-Pose Facial Expressions Pain Assessment Method. *Appl. Sci.* **2023**, *13*, 7206. [CrossRef]
21. Sait, A.; Dutta, A. Ensemble Learning-Based Pain Intensity Identification Model Using Facial Expressions. *J. Disabil. Res.* **2024**, *3*, e20240029. [CrossRef]
22. Salekin, M.; Zamzmi, G.; Goldgof, D.; Kasturi, R.; Ho, T.; Sun, Y. Multimodal spatio-temporal deep learning approach for neonatal postoperative pain assessment. *Comput. Biol. Med.* **2021**, *129*, 104150. [CrossRef]

23. Szczapa, B.; Daoudi, M.; Berretti, S.; Pala, P.; Del Bimbo, A.; Hammal, Z. Automatic Estimation of Self-Reported Pain by Trajectory Analysis in the Manifold of Fixed Rank Positive Semi-Definite Matrices. *IEEE Trans. Affect. Comput.* **2022**, *13*, 1813–1826. [CrossRef]
24. Thiam, P.; Hihn, H.; Braun, D.; Kestler, H.; Schwenker, F. Multi-Modal Pain Intensity Assessment Based on Physiological Signals: A Deep Learning Perspective. *Front. Physiol.* **2021**, *12*, 720464. [CrossRef]
25. Phan, K.N.; Iyortsuun, N.K.; Pant, S.; Yang, H.J.; Kim, S.H. Pain Recognition With Physiological Signals Using Multi-Level Context Information. *IEEE Access* **2023**, *11*, 20114–20127. [CrossRef]
26. Gkikas, S.; Chatzaki, C.; Pavlidou, E.; Verigou, F.; Kalkanis, K.; Tsiknakis, M. Automatic Pain Intensity Estimation based on Electrocardiogram and Demographic Factors. In Proceedings of the International Conference on Information and Communication Technologies for Ageing Well and e-Health, Virtual, 23–25 April 2022; pp. 155–162.
27. Mieronkoski, R.; Syrjälä, E.; Jiang, M.; Rahmani, A.; Pahikkala, T.; Liljeberg, P.; Salanterä, S. Developing a pain intensity prediction model using facial expression: A feasibility study with electromyography. *PLoS ONE* **2020**, *15*, 0235545. [CrossRef]
28. Khan, R.A; Meyer, A.; Konik, H.; Bouakaz, S. Pain detection through shape and appearance features. In Proceedings of the 2013 IEEE International Conference on Multimedia and Expo (ICME), San Jose, CA, USA, 15–19 July 2013; pp. 1–6. [CrossRef]
29. Zafar, Z.; Khan, N. Pain Intensity Evaluation through Facial Action Units. In Proceedings of the 2014 22nd International Conference on Pattern Recognition, Stockholm, Sweden, 24–28 August 2014; pp. 4696–4701. [CrossRef]
30. Brahimi, S.; Ben Aoun, N.; Ben Amar, C.; Benoit, A.; Lambert, P. Multiscale Fully Convolutional DenseNet for Semantic Segmentation. *J. WSCG* **2018**, *26*, 104–111. [CrossRef]
31. Ben Aoun, N.; Mejdoub, M.; Ben Amar, C. Bag of sub-graphs for video event recognition. In Proceedings of the 39th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'14), Florence, Italy, 4–9 May 2014; pp. 1566–1570. [CrossRef]
32. Brahimi, S.; Ben Aoun, N.; Ben Amar, C. Very Deep Recurrent Convolutional Neural Network for Object Recognition. In Proceedings of the International Conference on Machine Vision (ICMV'2016), Nice, France, 18–20 November 2016; Volume 10341, p. 1034107. [CrossRef]
33. Nhidi, W.; Ben Aoun, N.; Ejbali, R. Deep Learning-Based Parasitic Egg Identification from a Slender-Billed Gull's Nest. *IEEE Access* **2023**, *11*, 37194–37202. [CrossRef]
34. Nhidi, W.; Ben Aoun, N.; Ejbali, R. Ensemble Machine Learning-Based Egg Parasitism Identification for Endangered Bird Conservation. In Proceedings of the 15th International Conference on Advances in Computational Collective Intelligence (ICCCI'2023), Communications in Computer and Information Science, Budapest, Hungary, 27–29 September 2023; Volume 1864, pp. 364–375.
35. Chen, Z.; Ansari, R.; Wilkie, D.J. Automated Pain Detection from Facial Expressions using FACS: A Review. *arXiv* **2018**, arXiv:1811.07988.
36. Al-Eidan, R.; Al-Khalifa, H.; Al-Salman, A. Deep-Learning-Based Models for Pain Recognition: A Systematic Review. *Appl. Sci.* **2020**, *10*, 5984. [CrossRef]
37. Hassan, T.; Seus, D.; Wollenberg, J.; Weitz, K.; Kunz, M.; Lautenbacher, S.; Garbas, J.; Schmid, U. Automatic Detection of Pain from Facial Expressions: A Survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *43*, 1815–1831. [CrossRef]
38. 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]
39. Matsangidou, M.; Liampas, A.; Pittara, M.; Pattichi, C.; Zis, P. Machine Learning in Pain Medicine: An Up-To-Date Systematic Review. *Pain Ther.* **2021**, *10*, 1067–1084. [CrossRef]
40. Gkikas, S.; Tsiknakis, M. Automatic assessment of pain based on deep learning methods: A systematic review. *Comput. Methods Programs Biomed.* **2023**, *231*, 107365. [CrossRef]
41. De Sario, G.; Haider, C.; Maita, K.; Torres-Guzman, R.; Emam, O.; Avila, F.; Garcia, J.; Borna, S.; McLeod, C.; Bruce, C.; et al. Using AI to Detect Pain through Facial Expressions: A Review. *Bioengineering* **2023**, *10*, 548. [CrossRef]
42. Cascella, M.; Schiavo, D.; Cuomo, A.; Ottaiano, A.; Perri, F.; Patrone, R.; Migliarelli, S.; Bignami, E.G.; Vittori, A.; Cutugno, F. Artificial Intelligence for Automatic Pain Assessment: Research Methods and Perspectives. *Pain Res. Manag.* **2023**, *2023*, 6018736. [CrossRef]
43. Huo, J.; Yu, Y.; Lin, W.; Hu, A.; Wu, C. Application of AI in Multilevel Pain Assessment Using Facial Images: Systematic Review and Meta-Analysis. *J. Med. Internet Res.* **2024**, *26*, e51250. [CrossRef]
44. Lucey, P.; Cohn, J.; Prkachin, K.; Solomon, P.; Matthews, I. Painful data: The UNBC-McMaster shoulder pain expression archive database. In Proceedings of the 2011 IEEE International Conference on Automatic Face and Gesture Recognition (FG), Santa Barbara, CA, USA, 21–23 March 2011; pp. 57–64. [CrossRef]
45. Kächele, M.; Werner, P.; Al-Hamadi, A.; Palm, G.; Walter, S.; Schwenker, F. Bio-Visual Fusion for Person-Independent Recognition of Pain Intensity. In Proceedings of the International Workshop on Multiple Classifier Systems (MCS), Nanjing, China, 15–17 May 2015; Volume 132, pp. 220–230. [CrossRef]
46. Haque, M.; Bautista, R.; Noroozi, F.; Kulkarni, K.; Laursen, C.; Irani, R.; Bellantonio, M.; Escalera, S.; Anbarjafari, G.; Nasrollahi, K.; et al. Deep Multimodal Pain Recognition: A Database and Comparison of Spatio-Temporal Visual Modalities. In Proceedings of the 2018 13th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2018), Xi'an, China, 15–19 May 2018; pp. 250–257. [CrossRef]

47. Chandra, M.A.; Bedi, S. Survey on SVM and their application in image classification. *Int. J. Inf. Technol.* **2021**, *13*, 1–11. [CrossRef]

48. Taunk, K.; De, S.; Verma, S.; Swetapadma, A. A Brief Review of Nearest Neighbor Algorithm for Learning and Classification. In Proceedings of the 2019 International Conference on Intelligent Computing and Control Systems (ICCS), Madurai, India, 15–17 May 2019; pp. 1255–1260. [CrossRef]

49. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 1097–1105. [CrossRef]

50. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

51. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. In Proceedings of the International Conference on Learning Representations (ICLR), San Diego, CA, USA, 7–9 May 2015.

52. Szegedy, C.; Ioffe, S.; Vanhoucke, V. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. *arXiv* **2016**, arXiv:1602.07261.

53. Karcioglu, 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]

54. Prkachin, K.; Solomon, P.E. The structure, reliability and validity of pain expression: Evidence from patients with shoulder pain. *Pain* **2008**, *139*, 267–274. [CrossRef]

55. Ekman, P.; Friesen, W. *Facial Action Coding System: A Technique for the Measurement of Facial Movements*; Consulting Psychologists Press: Palo Alto, CA, USA, 1978.

56. Hammal, Z.; Cohn, J. Automatic detection of pain intensity. In Proceedings of the ACM International Conference on Multimodal Interaction (ICMI), Santa Monica, CA, USA, 22–26 October 2012; pp. 47–52.

57. Breivik, H.; Borchgrevink, P.C.; Allen, S.M.; Rossel, L.A.; Romundstad, L.; Breivik Hals, E.K.; Kvarstein, G.; Stubhaug, A. Assessment of pain. *Br. J. Anaesth.* **2008**, *101*, 17–24. [CrossRef]

58. Hicks, C.L.; von Baeyer, C.L.; Spafford, P.A.; van Korlaar, I.; Goodenough, B. The Faces Pain Scale—Revised: Toward a common metric in pediatric pain measurement. *Pain* **2001**, *93*, 173–183. [CrossRef]

59. Walter, S.; Gruss, S.; Ehleiter, H.; Tan, J.; Traue, H.; Werner, P.; Al-Hamadi, A.; Crawcour, S.; Andrade, A.; Moreira da Silva, G. 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. [CrossRef]

60. Werner, P.; Al-Hamadi, A.; Gruss, S.; Walter, S. Twofold-Multimodal Pain Recognition with the X-ITE Pain Database. In Proceedings of the 2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW), Cambridge, UK, 3–6 September 2019; pp. 290–296. [CrossRef]

61. Aung, M.S.H.; Kaltwang, S.; Romera-Paredes, B.; Martinez, B.; Singh, A.; Cella, M.; Valstar, M.; Meng, H.; Kemp, A.; Shafizadeh, M.; et al. The Automatic Detection of Chronic Pain-Related Expression: Requirements, Challenges and the Multimodal EmoPain Dataset. *IEEE Trans. Affect. Comput.* **2016**, *7*, 435–451. [CrossRef]

62. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.

63. Redmon, J.; Farhadi, A. YOLO9000: Better, Faster, Stronger. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 6517–6525. [CrossRef]

64. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the Inception Architecture for Computer Vision. In Proceedings of the 33rd International Conference on Machine Learning, New York, NY, USA, 20–22 June 2016; pp. 2818–2826.

65. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1382–1391.

66. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.

67. Ma, N.; Zhang, X.; Zheng, H.; Sun, J. ShuffleNet V2: Practical Guidelines for Efficient CNN Architecture Design. In Proceedings of the European Conference on Computer Vision, Munich, Germany, 8–14 September 2018; pp. 122–138.

68. Zhou, J.; Cui, G.; Zhang, Z.; Yang, C.; Liu, Z.; Sun, M. Graph Neural Networks: A Review of Methods and Applications. *arXiv* **2018**, arXiv:1812.08434.