Survey on Pain Detection using Machine Learning Models

Ruijie Fang, Elahe Hosseini, Ruoyu Zhang, Chongzhou Fang, Setareh Rafatirad, Houman Homayoun

Submitted to: JMIR AI
on: September 22, 2023

Table of Contents

Survey on Pain Detection using Machine Learning Models

Ruijie Fang[1] BEng; Elahe Hosseini[1] MS; Ruoyu Zhang[1] MS; Chongzhou Fang[1] BE; Setareh Rafatirad[2] PhD; Houman Homayoun[1] PhD

[1]Department of Electrical and Computer Engineering University of California Davis US
[2]Department of Computer Science University of California Davis US

Corresponding Author:
Ruijie Fang BEng
Department of Electrical and Computer Engineering
University of California
One Shields Avenue
Davis
US

Abstract

Background: Approximately 20% of adults in the United States suffer from chronic pain, making it the most common reason for adults seeking medical care. This widespread issue has far-reaching consequences for society, leading to an estimated annual cost of $560 million in terms of medical expenses, lost productivity, and disability. These negative effects underscore the ongoing public health concern posed by chronic pain. On an individual level, the repercussions of inadequate pain management are profound, encompassing physical, psychological, social, and financial hardships for patients. As the initial step in the journey of pain management, pain assessment assumes a critical role. Traditionally, pain assessment relies on self-reports and observational scales, both of which are inherently subjective, time-consuming, and resource-intensive. Consequently, researchers have pioneered the development of automated pain assessment approaches harnessing the power of machine learning to mitigate these challenges.

Objective: In pursuit of a comprehensive understanding of the current landscape in automated pain assessment, this paper endeavors to conduct a comprehensive survey and review of the recent strides made in this field. The assessment encompasses critical aspects such as datasets, modalities, and machine learning models, which have all played pivotal roles in shaping the trajectory of automated pain assessment. By delving into these key elements, we aim to shed light on the evolving methodologies and approaches that have emerged in recent years.

Methods: In this comprehensive survey of recent developments within the field of automated pain assessment, we embark on a multifaceted journey. Our survey begins by examining the datasets. We delve into the meta-information and distinctions inherent to various datasets, gaining valuable insights into their applicability and scope. Our analysis extends to the diverse modalities employed within the field, including facial expression analysis, physiological signals, audio data, pupil size variations, and the multimodal inputs. These modalities offer a rich tapestry of sources for pain assessment. Moreover, we dissect the machine learning models that have become pillars in this domain, review their functionalities, strengths, and limitations, thereby providing a holistic overview of the current state of automated pain assessment methodologies.

Results: The advancements of the automated pain assessment field has been comprehensively surveyed. These advances have been critically assessed and analyzed from various angles, including datasets, modalities, and machine learning models.

Conclusions: This paper offers a comprehensive survey of the current state-of-the-art in the domain of machine learning-driven automated pain assessment. It initiates by providing a succinct overview of pain's fundamental mechanisms and the accompanying physiological and behavioral responses. Furthermore, a selection of publicly available datasets frequently employed in this field is introduced, including UNBC-McMaster, BioVid, BP4D-Spontaneous, BP4D+, COPE, the YouTube dataset, X-ITE, EmoPain, and SenseEmotion. Subsequently, the paper delves into the exploration of commonly used machine learning techniques for automated pain assessment, with a particular focus on aspects like modality selection, measurement devices, feature extraction, and classification models. Additionally, recent studies concentrating on audio-based and pupil size-based automated pain assessment are also summarized.

(JMIR Preprints 22/09/2023:53026)
DOI: https://doi.org/10.2196/preprints.53026

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Original Paper

Survey on Pain Detection using Machine Learning Models

Abstract

Pain, as a primary reason driving individuals towards seeking medical attention, has become a social problem. In recent decades, great advancements in automated pain assessment have been achieved. In this comprehensive survey, we commence by offering an overview of pain and its underlying mechanisms. Subsequently, we examine existing literature encompassing various modalities proposed for automated pain recognition. These modalities encompass facial expressions, physiological signals, audio, and pupil dilation. Concluding our survey, we delve into the prevalent challenges and propose directions for the progressive advancement of this field.

Keywords: pain; pain assessment; machine learning; survey

Introduction

Pain is “an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage”, according to The International Association for the Study of Pain (IASP) [1]. However, the discussion on the most precise definition of pain is still ongoing and the advances in the understanding of pain instantiate the biopsychosocial perspective on pain to capture evidence-based understanding and the evolution of pain [2]. Based on the pain origin, it is categorized as Nociceptive (due to stimulation of sensory nerve fibers), Neuropathic (due to impaired somatosensory nervous system), or Psychogenic pain (caused, increased, or prolonged by mental, emotional, or behavioral factors). Based on the time duration of the pain, it may be categorized as Acute (short duration) and Chronic (long duration, may last longer than 3 months).

Approximately 20% of adults suffer from chronic pain in the United States and chronic pain is the most common reason adults seeking medical care. For society, chronic pain contributes to an estimated 560$ Million each year of medical cost, loss of productivity and disability [3], [4]. These negative effects make it remain a public health concern. For individuals, inappropriate pain management has very deleterious physical, psychological, social and financial consequences on patients. Untreated pain can lead to chronic pain syndrome, which is often accompanied by decreased mobility, impaired immunity, decreased concentration, anorexia, and sleep disturbances. More importantly, the use of prescription opioids for the treatment of chronic non-cancer pain is associated with a substantial risk for abuse, dependence and overdose [5].

As the first step of pain management, pain assessment holds an essential role [6]. Unprecise pain assessment can lead to severe consequences. Undertreatment of pain not only causes psychological suffering, but also physiological consequences, e.g. increased blood pressure and heart rate. On the other hand, Overtreatment of pain may result in nausea, vomiting or constipation immediately and drug addiction in long-term. Traditionally, pain assessment is conducted through self- reports or observational scales. Self-report refers to the conscious communication of pain-related information by the person in pain, typically using spoken or written language or gestures. Various pain rating scales have been developed to capture patients’ self-report of pain intensity. Traditional approaches used to play an important role in pain assessment, including the Verbal Rating Scale (VRS) [7], the Visual Analog Scale (VAS) [8], the Numerical Rating Scale (NRS) [9], the Wong-Baker FACES Scale (WBS) [10].

However, such scoring methods are not feasible on certain patients like unconscious patients. For this, different observational pain scales such as Behavioral Pain Scale (BPS) [11], Pain Assessment in Advanced Dementia (PAINAD) [12], or Neonatal Infant Pain Scale (NIPS) [13], are used in clinical settings. Most scales consider facial expressions, vocalizations, and body language, while some include vital parameters. It is hard to assess and compare the validity of the various scales, because studies differ a lot in design, methodology, subjects, and conceptualization of the pain phenomenon. Pain assessment through observation is very challenging and is affected by the subjective biases and errors in beliefs of the observer [14].

To solve these challenges, it is necessary to develop an objective, accurate, continuous pain assessment method as shown in Fig. 1. In the last decades, multiple studies have been conducted to evaluate the feasibility of automated pain assessment using multi-modality and machine learning techniques. This paper surveys and reviews the recent advances in the field in terms of datasets, modalities, machine learning models. Lastly, we present the challenges remaining in the field and propose future directions.

Pain Mechanism

The pain mechanism is not fully understood because of its complexity and diversity [15]. Pain, created by the brain, is a psychological state rather than a physical one[16]. Different from pain, nociception is related to the peripheral nervous system (PNS) and central nervous system (CNS) reacting to the stimulation from the internal or external environment, as generated by the activation of nociceptors [17]. The noxious stimulus damages the tissue or potentially activates the nociceptors in the peripheral structure. Then, the information will be transmitted to the spinal cord dorsal horn or the nucleus caudalis. From there, the information will continue to the cerebral cortex via the brainstem in the brain, and the perception of pain is generated. Thus, no brain, no pain [18]. Fig. 2 presents the pain mechanism.

Usually, pain is regarded as chronic pain or acute pain according to the duration of the pain. Acute pain is a kind of sudden pain. The mechanism of momentary pain is well understood [19]. The nociceptors generate the nociception, and the information is transmitted to the brain, where the perception of pain is caused. There are two major types of nociceptors responding to different stimuli: C-fibers, associated with unmyelinated axons, and A-delta fibers, associated with thinly myelinated axons [20]. C-fibers generate slow, diffuse pain, while A-delta fibers are related to sharp, pricking pain. Silent nociceptors typically respond to endogenous chemical mediators related to tissue injury [19].

Chronic pain, lasting more than three months, does not have a useful biological function and is challenging to treat due to its varied etiologies [21], [22], [23]. According to the International Classification of Diseases (ICD) -11, chronic pain can be categorized into musculoskeletal, neuropathic, visceral, and cancer pain [21].

Psychological distress refers to a diffuse subjective experience as an internal response to noxious stimuli. Many patients argue that psychological pain is more severe than intense physical pain [24]. Chronic pain can lead to psychological pain and depression, while depression can exacerbate chronic pain [25], [26]. Psychogenic pain is physical pain caused or increased by mental and emotional factors [27]. Treatments like transcutaneous electrical nerve stimulation (TENS) or psychotherapy are often more effective for reducing psychogenic pain compared to traditional painkillers [28], [29].

The body responds to pain via multiple physiological processes: the sympathetic nervous system (SNS), neuroendocrine system, immune system, as well as emotions [30]. The SNS, known for the fight or flight response, increases heart rate and blood pressure via hormones like catecholamines, epinephrine, and norepinephrine when activated [31]. The SNS also activates sweat glands via acetylcholine, reflecting the active level of SNS through the volume of secreted sweat within a time range [32].

Pain Datasets

Data that is representative is crucial in the creation of a pain recognition system and the demonstration of its efficacy. Crucially, the system should perform optimally within the intended medical context, a fact that must be validated through clinical studies involving patients. Nevertheless, preliminary phases of development could leverage experimental pain research with healthy volunteers, characterized by strictly managed conditions and possibly large participant pools, along with repeated application of the pain stimulus. This data is foundational to the development of machine learning models for automated pain detection.

For studying pain in healthy adults, an external stimulus is needed. Common methods include heat applied via contact (e.g., heated objects, electrical heaters) or radiant sources (e.g., infrared light). Tables. 1 summarize the publicly available datasets that were used for pain recognition research. The UNBC-McMaster Shoulder Pain Expression Archive Database [33] includes 200 video sequences that capture the facial expressions of 25 participants experiencing shoulder pain. Each video sequence includes individuals performing a series of active and passive range-of-motion tests to provoke visible responses to pain, providing a unique dataset rich in both the variety and volume of pain expressions. The dataset includes self-reported and observer assessments of pain intensity at the video level, along with FACS coding at the frame level. The BioVid Heat Pain Database [34] is a collection of physiological data and videos from 90 healthy adults subjected to controlled heat stimuli. BioVid is comprised of several sections: A, B, and C, which focus on pain stimulation, along with sections D and E, which are dedicated to posed expressions and emotion elicitation, respectively. The MIntPAIN database [35] collected color, depth, and thermal videos from 20 healthy adults who were subjected to approximately 1,600 instances of electrical pain stimuli at four different intensity levels. EmoPain [36], SenseEmotion [37], X-ITE Pain [38], BP4D-Spontaneous [39], and BP4D+ [40] datasets are significant resources for pain and emotion studies. EmoPain contains video, audio, motion, and sEMG for lower back pain. SenseEmotion and X-ITE Pain include audio and physiological data from healthy adults subjected to experimental pain stimuli, while X-ITE provides thermal videos, body movement, and EMG. BP4D-Spontaneous and BP4D+ offer facial video recordings from individuals undergoing the cold presser task, with BP4D+ further providing 3D and thermal videos, along with physiological signals.

In the field of infant pain research, the iCOPE [41], YouTube [42], APN-db [43], iCOPEvid [44], and USF-MNPAD-I [45] database are the publicly available datasets. The iCOPE consists of 204 static photographs that capture 26 neonates during various procedures. The images provide valuable insights into the facial expressions associated with infant pain experiences. The YouTube dataset offers 142 videos accompanied by audio, showcasing the reactions of different infants undergoing immunizations. The APN-db is a dataset that includes over 200 videos of infants undergoing various procedures, and it features unique annotations such as Neonatal Face and Limb Acute Pain (NFLAP) intensity. The USF-MNPAD-I dataset collects video, audio, and physiological data from 58 neonates during their hospitalization in the neonatal intensive care unit and is annotated by using NIPS and N-

PASS scales.

Post-operative pain: Although automated pain assessment in controlled settings is well-studied, post-operative pain has not been extensively researched due to the difficulty of data collection. Postoperative pain results from tissue injury following surgery and is critical to manage, as inadequate treatment can lead to serious physiological and psychological outcomes. Post-operative pain datasets often exhibit imbalanced distributions and may contain missing labels due to variability in patient experiences and clinical settings, further complicating accurate and comprehensive pain assessment. The NPAD-IA database [46] captures video, audio, and physiological data from 40 infants undergoing procedural (heel lancing and immunization) and postoperative (Gastrostomy tube) pain. Notably, it includes postoperative pain data, addressing the complexity and variability of pain levels in real-world clinical settings, thereby enhancing the ecological validity of the assessment. Salekin et al. [47] presents a novel fully automated deep learning framework to assess neonatal postoperative pain. It employs a Bilinear Convolutional Neural Network (B-CNN) to extract facial features and a Recurrent Neural Network (RNN) to model the temporal patterns of postoperative pain. The study uses a dataset of over 600 minutes of visual, vocal, and physiological data from neonates, demonstrating the feasibility and efficiency of combining B-CNN and RNN for continuous and accurate assessment of post-operative pain intensity in clinical settings. Salekin et al. [45] introduces an automated system for assessing neonatal post-operative pain by integrating visual, vocal, and physiological data. The study also employs a Bilinear Convolutional Neural Network (B-CNN) for spatial feature extraction but use a Long Short-Term Memory (LSTM) network for capturing temporal patterns, demonstrating that the multimodal spatial-temporal approach significantly outperforms unimodal methods, achieving an AUC of 0.87 and accuracy of 79%. Automated postoperative pain assessment is still in its nascent stages, primarily hindered by a lack of comprehensive datasets and consistent research efforts. The current methods, often unimodal and focused on short-term procedural pain, fail to capture the complex and prolonged nature of postoperative pain. There is a pressing need for more extensive and diverse datasets to improve the accuracy and reliability of these systems. Despite these challenges, the potential benefits of automated pain assessment are immense, offering more consistent and objective pain management that can significantly enhance patient outcomes and reduce the burden on healthcare providers.

Table. 1 Pain databases.

Database	Subjects	Modalities	Annotation
UNBC-McMaster [33]	25 adults with shoulder pain	Video of face (RGB)	FACS, VAS, OPI
BioVid [34]	87 healthy adults	Video of face (RGB), EDA, ECG, EMG	Stimulus (calibrated per person)
MIntPAIN [35]	20 healthy adults	Video of face (RGB, depth, thermal)	Stimulus (calibrated per person)
EmoPain [36]	22 adults with chronic back pain	Video, Audio, EMG, Motion capture	Self-report, naïve OPI
SenseEmotion [37]	45 healthy adults	Video of face, Audio, EDA, ECG, EMG	Stimulus (calibrated per person)
X-ITE [38]	134 healthy adults	Video of face, Video of body,	Stimulus (calibrated per

		Audio, EDA, ECG, EMG	person)
BP4D-Spontaneous [39]	41 healthy adults	Video of face (RGB, 3D)	Stimulus, FACS
BP4D+ [40]	140 healthy adults	Video of face (RGB, 3D, thermal), Heart rate, Respiration rate, Blood pressure, EDA	Stimulus, FACS
Database with neonates			
iCOPE [41]	26 healthy neonates	204 RGB photographs of face	Category (pain, rest, cry, air puff, friction)
YouTube [42]	142 infants	Video, Audio	FLACC
APN-db [43]	112 healthy neonates	Video of face (RGB)	NFLAPS, NIPS, NFCS
NPAD-ID [46]	36 healthy neonates, 9neonates with surgery	Video of face and body (RGB)	NIPS, N-PASS
iCOPEvid [44]	49 neonates	Video of face (Grayscale)	Category (pain, no pain)
USF-MNPAD-I [45]	36 neonates	Video of face (RGB), Audio, Heart rate, Blood pressure, SpO2, Deoxyhemoglobin (HbH), oxyhemoglobin (HbO2)	NIPS, N-PASS

Automatic Pain Assessment

Automated tools for pain assessment have great promise. Because pain results in different physiological and behavioral responses, signals that capture these may be used to detect the presence of pain. However, prior research work has been limited, and automated approaches have not yet become widely used in clinical practice. In this section, we briefly outline different approaches relevant to the development of automated pain assessment methods described in the research literature. Specifically, we review their system architecture (inputs and outputs) and describe the data sources available for the research and development of ML-based automated pain assessment tools, together with an overview of system validation challenges. This section summarizes the results of the survey of automatic pain detection approaches.

The use of modalities

The selection of sensors is a critical aspect of automated pain assessment, as different sensors can convey varying levels of information and have different discriminative abilities. Modalities commonly used in this field can be broadly classified into four categories: video, audio,

physiological signals as shown in Table. 2, fMRI was found to be the most prevalent sensor in pain studies, with a prevalence score of 95.9. EEG and ECG were also frequently used, with prevalence scores of 69.6 and 39.1, respectively. In contrast, fNIRS and PPG had much lower prevalence scores of less than 10. Besides, Appendix. 1 also includes information on modalities used in studies. (including brain activity, cardiovascular activity, electrodermal activity, and respiration activity), pupil size. In terms of physiological signals, brain activity can be measured using electroencephalogram (EEG), functional magnetic resonance imaging (fMRI), and functional near-infrared spectroscopy (fNIRS). Cardiovascular activity can be measured using electrocardiogram (ECG/EKG) or photoplethysmography (PPG), while electrodermal activity is often measured by skin conductance level (SCL) or surface electromyogram (sEMG). To gain insight into the prevalence of each modality, we conducted a search for "Modality AND Pain AND Machine learning" (e.g., "EEG AND Pain AND Machine learning") on PubMed and Scopus, limited to the period from Jan 1st, 2010, to August 1st, 2023. We then recorded the number of results and normalized them to the range of (0,100) for each database. The prevalence scores were then calculated as the average of the normalized results from PubMed and Scopus.

As shown in Table. 2, video was found to be the most prevalent sensor in pain studies, with a prevalence score of 100. FMRI, EEG and ECG were also frequently used, with prevalence scores of 95.9, 69.6 and 39.1, respectively. In contrast, fNIRS and PPG had much lower prevalence scores of less than 10.

Convenience and feasibility should also be considered when selecting sensors. For example, some sensors like EEG and fMRI are non-wearable and can be invasive, which may limit their utility in certain settings. Moreover, complex signals require more sophisticated processing techniques and computing resources, which may not be practical in some situations, such as those involving microprocessors.

Facial Expression

Facial expression during the experience of pain is not unspecific grimacing but conveys pain-specific information. Studies investigating facial expressions of pain have most often used Facial Action Coding System (FACS) [57], the gold standard for facial expression research. FACS is a fine grained, objective, and anatomically based coding system that differentiates between 44 facial movements known as Action Units (AU). Coders are trained to apply specific operational criteria to determine the onset and offset as well as the intensity of the AUs. Using FACS, it was shown that facial expressions of pain are composed of a small subset of facial activities, namely lowering the brows (AU4), cheek raise/lid tightening (AUs 6, 7), nose wrinkling/raising the upper lip (AUs 9, 10), and eye closure longer than 0.5s (AU 43). Prkachin et al. [58] developed the Prkachin and Solomon Pain Intensity (PSPI) metric based on this observation, which is a 16-level scale based on the contribution of the individual intensity of pain-related AUs and is defined as:

$$Pain = AU\,4 + (AU\,6, AU\,7) + (AU\,9, AU\,10) + AU\,43$$

The list of pain-related AUs has been further expanded in more extensive research [59] to include lip corner puller (AU12), lip stretch (AU20), lips part (AU25), jaw drop (AU26) and mouth stretch (AU27).

Facial activities during experimental and clinical pain are largely inborn but not uniform across individuals. People display different parts or combinations of facial activities. Cluster analyses identified four distinct facial activity patterns: (1) narrowed eyes with raised upper lip/nose wrinkling and furrowed brows, (2) narrowed eyes with furrowed brows, (3) narrowed eyes with mouth opening, and (4) raised eyebrows, which is less frequent and stable, often indicating novelty or

surprise in response to pain. Recognizing these patterns improves pain detection more than focusing on a single expression. Thus, acknowledging variability in facial expressions can enhance pain communication.

Facial expression analysis uses spatial and spatiotemporal features. Spatial features capture static details of the face, such as the geometric and textural characteristics of eyes, eyebrows, nose, lips, and facial contours, using techniques like facial landmark detection, geometric feature extraction, Gabor filters, LBP, and HOG. Spatiotemporal features capture dynamic changes in expressions over time, using techniques like optical flow or differences between consecutive frames. Advanced methods may involve 3D facial modeling or LSTM networks to identify temporal dependencies. Combining spatial and spatiotemporal features provides a comprehensive analysis of facial expressions.

Vision-based: Spatial Features

In the research conducted by Ashraf et al. [60] and Lucey et al. [61], features derived from the Active Appearance Model (AAM) were input into Support Vector Machine (SVM) classifiers for the purpose of frame-level pain recognition. In addition, they implemented pain detection at the sequence level by averaging the frame-level predictions. Gholami et al. [62] used a Bayesian extension of SVM, known as the Relevance Vector Machine (RVM), to differentiate between instances of pain and no pain in neonates. They also used this methodology to assess varying pain intensity levels. Meanwhile, Hammal et al. [63] identified four levels of pain intensity through the use of Log-Normal filter-based features and an SVM classifier. Kaltwang et al. [64] conducted a comparative study involving three separate methodologies. They utilized facial landmarks, Discrete Cosine Transform (DCT), and LBP features to train three distinct Relevance Vector Regression (RVR) models for estimating PSPI. The best results were achieved by training an additional RVR model that consolidated the predictions from the three previously trained RVR models. The system [65] utilizes a pyramid histogram of orientation gradients (PHOG) for shape information and a pyramid local binary pattern (PLBP) for appearance information, offering a more automated and objective approach to pain monitoring.

Pedersen [66] implementation used a 4-layer contractive autoencoder, along with SVM, which resulted in an effective pain detection system at the frame level. Egede et al. [67] extracted features using both deep learning models and handcrafted methodologies. Facial landmarks, HOG, and deep vectors drawn from pre-trained VGG-16 [68] and ResNet-50 [69] models were employed. Rudovic et al. [70] introduced a Personalized Federated Deep Learning (PFDL) technique for pain estimation derived from facial images. This approach involved using a compact CNN architecture across various clients without the need to share their facial images. Contrary to the full sharing of model parameters, the PFDL technique keeps the last layer localized. Hosseini et al. [71] utilized pre-trained ResNet-18 on a large emotion recognition dataset FER+ [72] and transfer learning techniques to improve accuracy and performance. Huang et al. [73] proposed a pain-awareness multi-stream CNN approach for feature extraction, focusing on specific regions most relevant to pain expression instead of entire face images. [74] proposed an Ensemble of Compact Convolutional Neural Networks (ECCNet) utilizing three compact CNNs (variants of VGG, MobileNet, and GoogleNet) and integrating their predictions using the average ensemble rule. Kharghanian et al. [75], [76] developed a four-layer Convolutional Deep Belief Network (CDBN), trained as Convolutional Restricted Boltzmann Machines (CRBM) to extract features. The addition of batch normalization units and utilization of a linear SVM for frame classification into pain or no-pain classes. Semwal et al. [77] introduces a novel fusion method for pain severity assessment in unconstrained environments, utilizing a decision-level fusion of three distinct features: data-driven RGB features,

entropy-based texture features, and complementary features from both RGB and texture data. Employing three CNNs (VGG-TL, ETNet, and DSCNN) with Transfer Learning, Entropy Texture Network, and Dual Stream CNN—the model and various data augmentation techniques avoid overfitting and improve performance. The system demonstrates a 94.0% F1-score on a self-generated dataset from an unconstrained hospital setting.

Alghamdi and Alaghband [78] presented a facial expressions-based automatic pain assessment system (FEAPAS) utilizing two concurrent subsystems that analyze both the full face and upper half of the face through pre-trained convolutional neural networks (CNNs) such as VGG16, InceptionV3, ResNet50, or ResNeXt50. Dai et al. [79] develops a real-time pain detection system by mixing pain and emotion datasets for optimal real-time performance and conducting a cross-corpus test. The study experiments with both AU-based and non-AU-based methods, ultimately implementing the method on a robot for frozen shoulder therapy, thus emphasizing the need for balanced and ecologically valid pain datasets, and the importance of real-world application and testing. Karamitsos et al. [80] utilizes Haarcascade Frontal Face Detector by OpenCV for face detection, then faces undergo gray scaling, histogram equalization, cropping, mean filtering, and normalization. The CNN is built upon a modified VGG16 architecture, achieving an impressive 92.5% accuracy. Barua et al. [81] utilized a shutter blinds-based model inspired by spontaneous facial expressions and patch-based learning to achieve over 95% accuracy in pain detection from facial images, leveraging transfer learning for efficient deep feature extraction. The model uniquely uses horizontal dynamic-sized patches, or "shutter blinds," to mine hidden facial signatures. Semwal et al. [77] assesses pain severity in unconstrained hospital environments by using a decision-level fusion of three distinct types of features: data-driven RGB, entropy-based texture, and complementary features. Employing three CNNs—VGG-CNN with Transfer Learning, Entropy Texture Network, Dual Stream CNN, and various data augmentation techniques to avoid overfitting. The system demonstrates a 94.0% F1-score on a self-generated dataset from an unconstrained hospital setting.

Li et al. [82] introduced a video-based infant monitoring system to analyze infant pain using three databases: Train-Data, Data-Clinic, and Data-YouTube. Utilizing Fast R-CNN with object tracking and a Hidden Markov Model, the system precisely detects infant expressions and states. With a significant dataset from varied sources, including over 16,000 images and real-world clinical videos, the approach offers enhanced accuracy and reliability in infant pain detection. Zamzmi et al. [83] introduced a Neonatal Convolutional Neural Network (N-CNN) that employs a cascaded architecture with three convolutional branches. This design merges image-specific and general information for pain detection. The N-CNN demonstrated 91% accuracy and 0.93 AUC on the Neonatal Pain Assessment Dataset and 84.5% on the Infant Classification of Pain Expression dataset. Witherow et al. [84] developed FACE-BE-SELF, a novel deep adaptive method for adult-child facial expression classification. It fuses facial landmark data with deep feature representations, achieving domain-invariant classification. Using a unique mixture of Beta distributions, facial features are selected based on expression, domain, and identity correlations. The FACE-BE-SELF method stands out by concurrently adapting adult-child domains, providing a unified expression representation for both groups. Compared to standard approaches, it surpasses in aligning latent representations of expressions across age groups.

Vision-based: Spatio-temporal Features

Bargshady et al. [85] presents an Ensemble Deep Learning Model (EDLM) that combines a three-stream hybrid neural network with CNNs to extract facial features and classify pain levels. The VGGFace, integrated with PCA, is utilized for early feature extraction, while a three-layer hybrid of CNN and Bi-LSTM is developed for late fusion classification. This approach, tested on multiple pain

databases, surpasses competing models with an accuracy of over 89%. SAFEPA (Sparse Autoencoders for Facial Expressions-based Pain Assessment) [86] reconstructs the upper part of the face from input images and then feeds both the original and reconstructed images into two concurrent and coupled InceptionV3 by utilizing Sparse Autoencoders (SAE). This dual-input approach emphasizes the upper facial features, essential for pain detection. By eliminating the need for conventional preprocessing steps like face detection, and adeptly handling varying head poses, SAFEPA offers enhanced performance and accuracy across multiple datasets, even in challenging profile views. [80] modified Temporal Convolutional Network (TCN) algorithm and processes facial features extracted from fine-tuned VGG-Face and PCA combined with HSV color spaces. The TCN-based approach showcases faster performance and higher efficiency, achieving an accuracy of 92.44% and an AUC of 85%. Bargshady et al. [88] proposes an Enhanced Joint Hybrid CNN-BiLSTM (EJH-CNN-BiLSTM) model by leveraging a fine-tuned VGG-Face for feature extraction and applies PCA to focus on the most significant features, improving computational efficiency. These features are then classified by a CNN-BiLSTM hybrid network into four levels of pain intensity.

3D CNNs have gained attention in several studies. Tavakolian and Hadid [110], [111] created a 3D CNN that captures dynamic facial representations from videos and emphasizes the typical use of a fixed temporal kernel depth in research, which often misses capturing different time ranges. In [112], a hybrid network (HybNet) by combining 3D, 2D, and 1D CNNs has been introduced to extract spatiotemporal, spatial, and geometric features from image sequences. Wang et al. [113] utilized the Convolutional 3D (C3D) network for pain expression recognition, which primarily employs a 3×3×3 convolutional layer. However, this method often fails to capture the full spectrum of facial expression variations. To address this, they combined three distinct features: 3D CNN, histogram of oriented gradients, and geometric features using Support Vector Regression (SVR) for pain estimation. They integrated the C3D network for spatiotemporal facial feature extraction and employed the HOG in 2D images for geometric information to discern pain levels in facial expressions. De et al. [114] present a deep learning architecture, the Decomposed Multiscale Spatiotemporal Network (DMSN). It employs three innovative blocks, DMSN-A, DMSN-B, and DMSN-C, to efficiently capture varied facial dynamics across conditions like depression and pain. DMSN-A block focuses on pain which might vary rapidly. It uses a sequence of 3×1×1 temporal convolutions, capturing short to long temporal ranges. The studies in [115] and [116] implemented weak-supervised domain adaptation, focusing on a shift from general affective expressions to specific pain expressions. Their framework utilized an inflated 3D-CNN (I3D) [117] with three convolutional layers and three inception modules, extracting both spatial and temporal data from videos.

Physiological Signals

While facial expressions are commonly used to identify pain, physiological signals are also a valuable modality for automatic pain detection. As detailed in Section II, pain triggers changes in physiological signals, such as increased heart rate and skin conductivity, due to the activation of the SNS and PNS [118]. Conversely, changes in physiological signals can indicate the presence of pain. However, extracting discriminative information from physiological signals is challenging. On the other hand, they are objective indicators of pain because they cannot be artificially controlled [119] while exterior signals, such as facial expressions and gestures, may be unreliable, as individuals can deliberately disguise their behaviors. It makes physiological signals more reliable than exterior signals. In addition, physiological signals can be measured during daily life while video, hand gesture can only be measured in lab settings. Thus, researchers have spent a great amount of work exploring the feasibility of physiological signal-based pain assessment. Recent advances in sensor technology, signal processing, feature extraction, and machine learning algorithms are essential to

the success of physiological signal-based automatic pain assessment.

This section provides a comprehensive review of the latest developments in pain detection approaches based on physiological signals. Four key components are exploited: (1) the use of modalities, (2) measurement devices, (3) feature extraction methods, and (4) machine learning models. The use of modalities refers to the type of physiological signals used for pain detection, including EEG, fMRI, ECG, EDA. Measurement devices contain wearable devices, non-wearable devices covering cardiac, skin conductivity, temperature, accelerometer devices, etc. Feature extraction methods are techniques used to extract informative features from physiological signals, such as time-domain features, frequency-domain features, and time-frequency features. Finally, machine learning models, such as support vector machines (SVM), artificial neural networks (ANN), and random forest (RF), are employed to classify pain based on the extracted features.

Table. 2 Summary of commonly used modalities

Category	Abbreviation	Name	Description	Prevalence	References
Video	-	Video analysis	Analyzes facial expressions and body movements to assess pain levels [57].	100	[35] [33]
Audio	-	Audio analysis	Analyzes vocal characteristics and speech patterns to assess pain [89].	48.2	[89]
Pupil Size	-	Pupil size measurement	Measures changes in pupil diameter as an indicator of pain [90].	12.7	[91] [92]
Brain activity	EEG	Electroencephalogram	EEG is a test that detects tiny electrical charges that result from the activity of brain cells [82].	69.6	[94] [52] [95]
	fMRI	Functional magnetic resonance imaging	fMRI use magnetic resonance imaging to measure the changes in hemodynamics caused by neuronal activity [86].	95.9	[97] [98] [99]
	fNIRS	Functional Near-Infrared Spectroscopy	fNIRS Utilize the good scattering of the main components of blood to 600-900nm near-infrared light, so as to obtain the changes of oxyhemoglobin and deoxyhemoglobin during brain activity [90].	7.9	[101] [102]
Cardiovascular activity	ECG	Electrocardiogram	ECG is a test that measures the electrical activity of the heartbeat [93].	39.1	[104] [55] [49]
	PPG	Photoplethysmograph	PPG is an optical technique that can be used to detect blood volume changes in the microvascular bed of tissue [97].	9.4	[55] [54]
Electrodermal	SCL	Skin conductanc	SCL is measurement of the electrical conductivity of the skin	25.9	[104] [55]

activity		e level	[99].		[50]
	sEMG	surface electromyogram	sEMG is a technique to measure muscle activity noninvasively using surface electrodes placed on the skin overlying the muscle [101].	25.6	[108] [104] [48]
Respiration	RSP	respiration	Respiration refers to a person's breathing and the movement of air into and out of the lungs [104].	17.5	[53] [48]

[a]Prevalence Prevalence is measured by weighted searching results showed in Scopus and PubMed from 2010 to 2023 with key word "Name" AND "Pain" AND "Machine learning" on Aug 1st, 2023, and standardized in range of (0,100).

EEG

Electroencephalography (EEG) is a non-invasive technique widely used in the automatic detection of pain. The electrodes detect electrical activity and amplify it, producing a graphical representation of the brain activity over time. EEG recordings typically show a series of waveforms or oscillations that are grouped into different frequency bands, such as delta, theta, alpha, beta, and gamma. These frequency bands have been associated with different mental states and cognitive functions. Various studies have shown the potential of EEG-based pain detection, and different approaches have been proposed to extract discriminative features from EEG signals for pain classification. For instance, Panavaranan et al. [120] extracted power spectral density of EEG using Fast Fourier Transform and used Support Vector Machines (SVM) to classify thermal pain. Hadjileontiadis et al. [52] proposed a novel approach that analyzes wavelet higher-order spectral (WHOS) features of EEG to predict tonic cold pain. Vijayakumar et al. [121] extracted time-frequency wavelet representations of independent components from EEG data and trained a random forest model to classify pain levels, achieving an intra-subject accuracy of 93.26%.

The use of EEG techniques for pain detection has great potential to provide objective measures of pain, as these methods directly measure brain activity related to pain perception. However, these techniques also have limitations, including high cost, limited availability, and the need for specialized expertise for data analysis.

fMRI

Functional magnetic resonance imaging (fMRI) is a powerful neuroimaging tool that measures changes in blood flow within the brain as a proxy for neural activity. By measuring changes in the blood oxygen level-dependent (BOLD) signal, fMRI can indirectly map changes in neural activity in response to a specific stimulus, such as a painful stimulus.

The fMRI technique has been widely used in pain research, revealing a network of brain regions that are activated by painful stimuli. These regions include the primary and secondary somatosensory cortex, thalamus, insular cortex, and anterior cingulate cortex, among others. The activation of these regions is believed to be involved in the sensory and affective components of pain processing.

Activation of these regions is thought to be involved in the sensory discrimination aspects of pain processing. Thus, neuroimaging techniques allow us to visualize and quantify brain activities and then quantify pain. It is frequently employed in the research of automatic pain assessment. Wager et al. [122] used LASSO-PCR machine learning regression algorithm to recognize induced heat pain by

assessing the fMRI activity patterns. Shen et al. [99] derived primary, dorsal and ventral visual networks from BOLD-fMRI scans by using independent component analysis and employed a machine learning algorithm support vector machine (SVM) to discriminate chronic low back pain patients with healthy volunteers and achieved an accuracy of 79.3%. Tu et al. [98] proposed a novel sliced inverse regression-based fMRI decoding method to reduce the fMRI data dimension and showed overperformance than traditional regularization-based decoding analyses (PCA-DA, PLS-DA, LASSO-PCR). Robinson et al. [97] scanned fMRI and applied machine learning algorithms to classify fibromyalgia patients and healthy volunteers.

ECG

Electrocardiogram (ECG) is a widely used technique to measure the electrical activity of the heart and its changes during each cardiac cycle. The ECG waveform consists of several characteristic waves and intervals that correspond to the different phases of the cardiac cycle, including the P wave, QRS complex, and T wave. By analyzing the size, shape, and timing of these waves and intervals, a wide range of cardiac conditions, such as arrhythmias, heart attacks, and heart failure can be diagnosed. The use of ECG in pain detection assumes that pain can cause a physiological stress response, leading to cardiovascular changes that are related to the pain stimuli. The autonomic nervous system responds to pain by increasing sympathetic tension and decreasing parasympathetic tension, leading to an increase in heart rate and blood pressure. By analyzing the ECG signal, features that reflect the autonomic nervous system status, such as heart rate variability (HRV), can be extracted and used to detect pain.

Several studies have shown the potential of ECG for pain detection. Walter et al. [34] collected ECG data from 90 subjects using heat as pain stimuli and created the BioVid dataset, which also included skin conductance level (SCL), surface electromyography (sEMG), and video data. Adjei et al. [95] performed spectral analysis on ECG data and extracted HRV features, such as the low-frequency component (LF) and high-frequency component (HF), which were significantly correlated with pain level. Jiang et al. [49] extracted time-domain and frequency-domain HRV features from ECG data to classify pain level and obtained an area under the curve (AUC) of 0.82 in the receiver operating characteristic (ROC) curve.

However, there are also studies that suggest a lack of correlation between HRV and pain level. Meeuse et al. [123] found no significant correlation between HRV features and heat pain level in their study. It is important to note that ECG alone may not be sufficient to accurately detect pain, and other physiological signals, such as skin conductance and electromyography, may need to be considered as well. Furthermore, individual differences in pain perception and the variability of pain stimuli may affect the reliability of pain detection using ECG.

EDA

Electrodermal activity (EDA), also referred to as galvanic skin response (GSR), is a physiological gauge of the skin's electrical conductance. This conductance changes according to the functioning of sweat glands within the skin [124]. The measurement of EDA is a non-invasive process involving the placement of two electrodes, often on the fingers or palms. Activation of the sympathetic nervous system, triggered by situations like stress or pain, leads to increased sweat gland activity, causing a rise in skin's electrical conductance.

Within the context of automated pain recognition, EDA serves as a valuable indicator due to its reflection of sympathetic nervous system activity [125], closely tied to the body's response to pain. Numerous research studies have highlighted EDA's potential in pain detection. For instance, in the

BioVid dataset developed by Walter et al. [34], EDA was employed as one of the methods, revealing a correlation between EDA features and the intensity of pain.

Surface electromyogram (sEMG) is another important tool for measuring EDA in automatic pain detection. sEMG can measure the electrical activity of muscles and has been used to measure facial expression [126] or muscle movement of specific body parts, such as the back muscles [127]. These measures can provide additional information about the pain experience and may be used in combination with other modalities for better pain detection accuracy [128].

Devices

Data collection is indeed crucial in research, especially in statistical and machine learning-based studies. It is essential to ensure that the data collected is accurate, informative, and clean. However, data school, selecting the right measurement devices is critical to obtain high-quality data.

Table. 3 is a summary of previously used measurement devices in pain assessment studies. Fig. 4 presents three typical types of devices used in physiological signal-based pain assessment: wristband, headset and chest band. The importance of wearable devices in this context cannot be overstated; they enable ubiquitous, real-time data collection[129], [130], especially with the rise of body sensor networks. This technological advancement allows for extensive data gathering in wearable and remote settings, making continuous monitoring both feasible and affordable.

Table. 3 Physiological signals measurement devices used in pain assessment studies

Device	Physiological signals	Connectivity	Type	FDA-cleared[a]	Reference
Bioharness 3	ECG[b]	Bluetooth	Chest band	Y	[48][49]
Affectiva Q sensor	EDA[c]	Bluetooth	Wrist band	Y	[50]
Procomp+	EDA, HR[d]	Wired	Measurement hub	Y	[51]
Emotiv EPOC 14-channel EEG wireless recording headset	EEG[e]	Bluetooth	Headset	N	[52]
RespiBan	RSP[f]	Bluetooth	Chest band	N	[53]
Empatica E4	EDA, BVP[g], RSP	Wired	Wired sensor	Y	[53]
Infiniti 3000A platform with Flex/Pro sensors	BVP, ECG, EDA	Wired	Sensorhub	Y	[54][55]
Polar RS800CX	HRV[h]	Wired	Watch	N	[56]

[a]Y: FDA-cleared, N: not FDA-cleared, [b]eletrocardiacdiagram, [c]electrodermal activity, [d]heart rate, [e]electroencephalogram, [f]repiration rate, [g]blood volumn pulse, [h]heart rate variability

There are several studies that have evaluated the usability and reliability of different measurement devices. Researchers can use these studies as references when selecting measurement devices for

their studies. Ajayi et al. [134] evaluated the Empatica E4 by comparing the results with nurse-recorded data and pooling questionnaires from participants. Nazari et al. [135] tested the reliability of Bioharness and Fitbit measures of heart rate and activity at rest status. Rawstorn et al. [136] tested Bioharness under simulated daily activities, low, moderate, and high-intensity exercises on volunteers with sinus rhythm and atrial fibrillation. Loberg et al. [137] evaluated four different respiratory effort sensors and compared them with a RIP sensor from NOX Medical as the golden reference device.

Feature extraction

In the field of machine learning, pattern recognition, and image processing, feature extraction is a crucial step that involves transforming raw data into informative and non-redundant features to facilitate subsequent learning and generalization. Physiological signals typically carry implicit information that needs to be revealed through appropriate feature extraction techniques. While deep learning methods often generate features automatically, traditional machine learning methods require manual feature extraction.

In the case of physiological signals, time-window segmentation is commonly used to extract features. This involves segmenting the signals into chunks of equal time intervals and generating a row vector for each segment with one feature value for each feature, e.g., the mean value of the segmentation. Physiological signal features can be classified into four categories: time-domain, frequency-domain, time-frequency-domain, and space-domain features.

Time-domain features describe the statistical and morphological properties of physiological signals, such as maximum value, standard deviation, entropy, and mean R-R interval in ECG signals. Frequency-domain features characterize the spectral properties of signals, such as low-frequency band power and low-high frequency ratio. Time-frequency-domain features consider both time-domain and frequency-domain properties simultaneously to account for the short duration and changing nature of physiological signals. Space-domain features, such as multi-spectral imaging and topography, are used to represent topographic characteristics of brain activity features, including EEG, fMRI, and fNIRS.

The complexity of physiological signals can guide feature selection. Signals with high stochastic stationarity and low signal-to-noise ratio, such as PPG and EDA, are considered low in complexity and can be represented by one or two feature domains. Signals with low stochastic stationarity and high signal-to-noise ratio, such as ECG, EEG, and fMRI, are high in complexity and require three to four feature domains to capture all relevant information. Nowadays, numerous Python libraries are available that facilitate the rapid extraction of features in physiological signals [138], [139], EEG [140], video[141], and audio[142] domains. A summary of commonly used features is presented in Table. 4.

Brain-activity features

Physiological signals, including EEG, fMRI, and fNIRS, have unique characteristics that require specific feature extraction techniques. EEG signals, for example, have high topological complexity as multiple channels are measuring simultaneously. They can be divided into different frequency bands such as Delta, Theta, Alpha1, Alpha2, Beta1, Beta2, Gamma1, and Gamma2. To assess pain, Panavaranan et al. [120] used Power Spectral Density (PSD) features calculated using Fast Fourier Transform (FFT). Hadjileontiadis et al. [52] combined continuous wavelet transform (CWT) with higher-order statistics/spectra (HOS) to create a new feature space for EEG. Rissacher et al. [94]

found temporal parietal alpha of EEG to be a useful feature for pain assessment.

In fMRI, Tu et al. [98] proposed a novel dimension reduction method by incorporating singular value decomposition (SVD) into sliced inverse regression (SIR) to overcome the limitations of SIR when dealing with high-dimensional data. This method was used to assess pain, achieving 77.61% binary classification accuracy.

There are various feature extraction approaches for EEG signals, as summarized by Behzadfar et al. [143]. For brain activity signals in general, van der Miesen et al. [144] outlined the state and progress in pain detection using these signals.

ECG features

Unlike general statistical feature extraction methods, ECG feature extraction involves more human experience on ECG and is more interpretable. Shaffer et al. [145] overviewed heart rate variability features from time-domain, frequency-domain and non-linear measure features. Time-domain and frequency-domain features are widely used in pain assessment studies. On BioVid dataset, Werner et al. [146] derived MNRR, RMSSD, slopeRR from ECG signal. Gruss et al. [147], Campbell et al. [148] and Kachele et al. [149] used the same three features in their studies. Kachele et al. [149] also applied four level wavelet decomposition on detected R peaks to extract the mean α1 coefficients. Jiang et al. [49] extracted time-domain features AVNN, SDNN, RMSSD, pNN20 and frequency-domain features LF, HF, LF/HF from ECG and attained AUC = 0.82 for induced electrical pain and AUC = 0.75 for induced thermal pain.

Apart from heart rate variability, other features have been used for various purposes. For instance, some studies have used morphological features such as QRS complex duration and amplitude, T-wave amplitude, and ST-segment changes for diagnosing cardiac abnormalities [149].

EDA/EMG features

Electrodermal activity (EDA) and electromyography (EMG) are critical tools in pain detection because they measure physiological responses that are directly linked to the autonomic nervous system's reactions, which vary significantly with pain perception [150], [124]. Walter et al. [151] systematically gathered and summarized feature extraction methods for EDA/EMG signals from previous research and categorized into mathematical groups of (1) amplitude, (2) frequency [152], (3) stationarity [153], (4) entropy [154], (5) linearity [155] and (6) variability. In total, 33 different features were listed and their efficiency in pain assessment on BioVid dataset was proved. Then, Gruss et al. [147] deployed the feature table and derived it to 39 features. Campbell et al. [148] also developed a feature list based on Walter et al. [151]'s work. They also proposed a machine learning-based feature selection approach which deploys univariate feature selection and sequential forward selection for 100 epochs with cross-validation as the metric to explore the optimal feature set. From their results, a relationship table between features and pain was displayed illustrating the discriminative strength of features. In addition, amplitude, power, and unique functional features of EMG signal are noted as useful in all different feature sets. Table.4 summarized the features used in previous studies.

Table. 4 Summary of commonly used physiological signal features in pain assessment studies.

Category	Feature	Description	Reference
HRV time-domain measures	SDNN	Standard deviation of NN intervals	[145]

	SDRR	Standard deviation of RR intervals	[145]
	SDANN	STD of the average NN intervals for each 5min segment of a 24h HRV recording &	[145]
	SDNNI	Mean of the STD of all the NN intervals for 5min segment of a 24h HRV rec	[145]
	pNN50	Percentage of successive RR intervals that differ by more than 50 ms	[145]
	HR Range	Average difference between the highest and lowest heart rates during each respiratory cycle	[145]
	RMSSD	Root mean square of successive RR interval differences	[145]
	HRV triangle index	Integral of the density of the RR interval histogram divided by its height	[145]
	TINN	Baseline width of the RR interval histogram	[145]
HRV frequency-domain measures	ULF power	Absolute power of the ultra-low-frequency band (≤0.003 Hz)	[145]
	VLF power	Absolute power of the very-low-frequency band (0.0033–0.04 Hz)	[145]
	LF peak	Peak frequency of the low-frequency band (0.04–0.15 Hz)	[145]
	LF power (ms^2)	Absolute power of the low-frequency band (0.04–0.15 Hz)	[145]
	LF power (nu)	Relative power of the low-frequency band (0.04–0.15 Hz) in normal units	[145]
	LF power (%)	Relative power of the low-frequency band (0.04–0.15 Hz)	[145]
	HF peak	Peak frequency of the high-frequency band (0.15–0.4 Hz)	[145]
	HF power (ms^2)	Absolute power of the high-frequency band (0.15–0.4 Hz)	[145]
	HF power (nu)	Relative power of the high-frequency band (0.15–0.4 Hz) in normal units	[145]
	HF power (%)	Relative power of the high-frequency band (0.15–0.4 Hz)	[145]
	LF/HF	Ratio of LF-to-HF power	[145]
HRV non-linear measures	S	Area of the ellipse which represents total HRV	[145]
	SD1	Poincare plot standard deviation perpendicular the line of identity	[145]

	SD2	Poincare plot standard deviation along the line of identity	[145]
	SD1/SD2	Ratio of SD1-to-SD2	[145]
	DFA α1	Detrended fluctuation analysis, which describes short-term fluctuations	[145]
	DFA α2	Detrended fluctuation analysis, which describes long-term fluctuations	[145]
	D_2	Correlation dimension, which estimates the minimum number of variables required to construct a model of system dynamics	[145]
Amplitude	Peak	Peak Amplitude	[151]
	P2P	Peak to Peak Amplitude	[151]
	RMS	Root Mean Square	[197]
	MAV	Mean Absolute Value	[197]
	TMNP	Mean Relative Time of the Peaks	[198]
	TMNV	Mean Relative Time of the Valleys	[198]
Variability	IQR	Interquartile Range	[198]
	R	Range	[151]
	SD	Standard Deviation	[151]
	VAR	Variance	[197]
	MNRR	Mean Resting Rate	[145]
	SlopeRR	Slope Resting Rate	[145]
Stationarity	IDS	Integral Degree of Stationarity	[153]
	MIDS	Modified Integral Degree of Stationarity	[153]
	MMNDS	Modified Mean Degree of Stationarity	[153]
	MD	Median	[151]
	SDSD	Standard Deviation of Standard Deviation Vector	[151]
Entropy	ApEn	Approximate Entropy	[199]
	FuzzyEn	Fuzzy Entropy	[200]
	SampEn	Sample Entropy	[201]
	ShannonEn	Shannon Entropy	[202]
	SpectralEn	Spectral Entropy	[203]
Linearity	LDF	Lag Dependence Function	[151]
	PLDF	Population Lag Dependence Function	[151]
Similarity	CC	Correlation Coefficient	[204]
	MDCOH	Median Coherence	[205]
	MNCOH	Mean Coherence	[205]
	MMNCOH	Modified Mean Coherence	[205]
	MICOH	Modified Integral of Coherence	[205]
	MI	Mutual Information	[155]
Frequency	BW	Bandwidth	[151]

	CF	Center Frequency	[151]
	MDF	Median Frequency	[197]
	MNF	Mean Frequency	[197]
	MOF	Mode Frequency	[151]
	ZC	Zero Crossings	[197]

Models

In the field of machine learning, “no free lunch” theorem has been referred often when talking about model selection [161]. This theorem illustrates that “any two optimization algorithms are equivalent when their performance is averaged across all possible problems” which implies there’s no single algorithm always has the best performance for all machine learning tasks. Thus, appropriate model selection is necessary for the success of machine learning-based pain assessment. In this section, we compare different machine learning algorithms by illustrating their advantages and disadvantages and their applicable scenarios. Table.5 gives a summary of prevalent machine learning algorithms used in pain assessment.

Table. 5 Summary of prevalent machine learning algorithms in pain assessment studies.

Abbrev.	Model	Advantage	Disadvantage	Reference
SVM	Support vector machine	● Suitable for small dataset ● Take advantages of kernel functions	● Low performance in multi-class tasks	[49][53]
DT	Decision tree	● Easily interpretable ● Computation friendly	● High risk of overfitting ● Discard of correlations between features	[156]
RF	Random forest	● Applicable on large dataset ● Fix the overfitting problem of DT ● Easy to parallelize	● Low performance on low-dimensional dataset ● Time consuming	[157][158]
NN	Neural networks	● High performance with large amount of data ● Flexible with layers configurations	● Uninterpretable ● Computation consuming	[159][160]

Support vector machine

The first commonly used machine learning model in physiological-signal-based automatic pain detection is Support Vector Machine (SVM) [53], [49]. SVM is a type of generalized linear classifier that classifies data in a supervised learning way [162]. Its decision boundary is the maximum margin hyperplane for learning samples. SVM also includes kernel tricks, which makes it a substantially nonlinear classifier. The final decision of SVM only depends on the support vectors, which makes it suitable for small sample learning. By contrary, SVM lacks the ability to provide restoration of variables to the formation of derived predictors [163] which is important in some areas such as financial prediction and health applications. In addition, SVM requires delicate preprocessing and

tuning to acquire the best performance. Panavaranan et al. [120] applied polynomial kernel SVM on EEG data and obtained accuracy of 96.97%. Gruss et al. [147] used SVM on BioVid dataset and gained 90.94% accuracy on pain tolerance classification. In addition, Jiang et al. [49] obtained AUC of 0.82 with the use of SVM. More recently, Badura et al. [53] achieved 94% accuracy by using Gaussian kernel SVM.

Decision tree

Unlike SVM, Decision Tree is known for its interpretable characteristic. The decision tree algorithm is a method of approximating the value of a discrete function [164], [165]. It is a typical classification method which uses an induction algorithm to generate readable rules and decision trees, and then uses decision-making to analyze the new data. Essentially, a decision tree is a process of classifying data through a series of rules. Because of essential advantage of interpretability tree-based algorithms hold, they make machine learning process jumps out of the “black box” [166]. On the other hand, due to the simple structure of tree-based models, over-fitting easily happened on tree-based models [167]. Besides, they lack the ability to deal with missing data due to the continuity of tree structure.

Random forest

Random forest is an algorithm that integrates multiple trees through the idea of ensemble learning. Its basic unit is a decision tree, and essentially, it belongs to a large branch of machine learning “Ensemble Learning” method. Intuitively, each decision tree is a classifier, then for an input sample, N decision trees will have N classification results. Random forest integrates all classification voting results and designates the category with the most votes as the final output, which is a “bagging” idea. With the tree base and bagging theory random forest holds, it has advantages such as preventing overfitting, easy to parallelize, friendly with high-dimensional data, etc. [168]. In contrast, random forests have longer trained and predicting time compared to decision trees. Vijayakumar et al. [121] applied random forest on 25 subjects’ EEG data and obtained 89.45% accuracy. Naeini et al. [169] utilized random forest on BioVid dataset and got accuracy 79%. Werner et al. [170] used RF on their new dataset “X- ITE” dataset and achieved 94.3% accuracy for phasic electrical pain classification.

Neural networks

Neural networks have also been used by scholars for automatic pain detection [159], [160]. NN abstracts the human brain neuron network from the perspective of information processing, establishes a certain simple model, and composes different networks according to different connection structures. Benefits from the development of digital society, the amount of data available for machine learning has become huge. NN, which can go deep in its layers structure, can reveal implicit information from data. Therefore, as the amount of data grows, the performance of NN keeps increasing while traditional algorithms such as SVM, RF are limited. Nevertheless, NN has the defect of “black box” characteristic. Such un-interpretability keeps NN from blooming in certain fields such as text and code analysis [171], judicial decision and AI medicine because such fields require clear, understandable and interpretable decision-making process. Martinez et al. [172] utilized NN on BioVid dataset and obtained 82.75% accuracy on multi-task classification. Jiang et al. [48] applied artificial neural networks (ANN) on 30 subjects and gained an average accuracy of 83.3%. The deviation of neural networks is widely used in automated pain assessment such as convolutional neural network (CNN) [157], recurrent neural network (RNN) [173], long short-term memory neural network (LSTM) [174], etc.

Audio Analysis

Infant cry is a common sign of discomfort, hunger, or pain. It conveys information that helps caregivers to assess the infant's emotional state and react appropriately. Crying analysis can be divided into two main stages: (1) signal processing stage, which includes preprocessing the signal and extracting representative features; and (2) the classification stage. We classified the existing methods of signal processing stage into: (1) Time-domain methods; (2) Frequency-domain methods; and (3) Cepstral-domain methods.

Time domain analysis is the analysis of a signal with respect to time (i.e., the variation of a signal's amplitude over time). Linear Prediction Coding (LPC) is one of the most common time-domain methods for analyzing sounds. The main concept behind LPC is the use of a linear combination of the past time-domain samples to predict the current time-domain sample. Other time-domain features that are commonly used for infants' sound analysis are energy, amplitude, and pause duration. Vempada et al. [89] presented a time-domain method to detect discomfort-relevant cries. The proposed method was evaluated on a dataset consists of 120 cry corpuses collected during pain (30 corpuses), hunger (60 corpuses), and wetdiaper (30 corpuses). We want to note that the paper does not provide information about the stimulus that triggered the pain state nor the data collection procedure. The infants' age ranges from 12 - 40 weeks old. All corpses were recorded using a Sony digital recorder with sampling rate of 44.1 kHz. In the feature extraction stage, two features were calculated: 1) Short-time energy (STE), which is the average of the square of the sample values in a suitable window; and 2) Pause duration within the crying segment. Part of these features were used to build SVM, and the remaining were used to evaluate its performance. The recognition performance of pain cry, hunger cry, and wet-diaper cry were 83.33%, 27.78%, and 61.11% respectively. The average recognition rate was 57.41%.

Pupil Size

The measurement of changes in pupil size has been shown to be a promising physiological indicator of pain intensity. Pupil size can be used to monitor the effects of painful stimuli in the brain. The pupil dilates in response to pain, due to the activation of the sympathetic branch, which releases norepinephrine, and the inhibition of the parasympathetic branch, which is responsible for constriction of the pupil. This section discusses the mechanism of using pupil dilation as pain indicator and literature reviews of using pupil dilation on automated pain assessment.

The pupil dilation is a complex physiological response regulated automatically by two muscles in the eye, the sphincter pupillae and the dilator pupillae. The sphincter pupillae is controlled by the parasympathetic system to contract the pupil while the dilator pupillae is dominated by the sympathetic system to dilate the pupil [90].

Höfle et al. [91] investigated the influence of different luminance conditions on pupillometry for pain detection and found that the baseline pupil size values significantly differed under different luminance conditions while the peak dilation remained the same. Bertrand et al. [175] explored the influence of gender and anxiety on pupil dilation for pain detection and concluded that pupil dilation changes similarly in both men and women and are exacerbated in the presence of anxiety. Connelly et al. [92] conducted an experiment on 30 children undergoing elective surgical correction of pectus excavatum and found that maximum pupil size, percent change in pupil size, and maximum constriction velocity were the most related features to pain intensity. Chapman et al. [176] reported a delay of 1.25 seconds on 20 adult volunteers under noxious stimulation, while Eisenacha et al. [177] reported a peak in pupil size with a lag of 4.25 seconds after the onset of heat pain on 28 adult

volunteers. Wang et al. [178] found that the pupillary response together with machine learning algorithms could be a promising method of objective pain level assessment by measuring pupillary response during induced cold pain on 32 subjects.

Multi-modal Pain Detection

Including more modalities can possibly increase information density which lead to increased accuracy. Thus, researchers have been increasingly turning to multi-modal approaches to enhance the accuracy and reliability of automated pain assessment systems. These approaches combine information from multiple modalities, such as biomedical signals and facial expressions, to provide a more comprehensive understanding of the patient's pain experience. Furthermore, a multi-modal approach can capture a more nuanced and diverse range of pain responses, which is particularly important given the wide variation in pain perception among individuals with different characteristics and cultural backgrounds. Fig. 5 presents a typical flow of multi-modal pain assessment.

Fusion strategies commonly used in multi-modal pain assessment can be categorized into early fusion and late fusion. Early fusion involves the combination of features from different modalities before the training of a classifier, while late or decision fusion combines the predictions of individual classifiers after training. Common methods of combining predictions include fixed methods like taking the mean or product, and trainable methods like using a pseudo-inverse. Figure 6 illustrates the early and late fusion strategies. Some research has explored combining early and decision fusion by merging specific features at the feature level and then fusing those with other features at the decision level [45].

The first study to combine video and physiological signals for automated pain detection is presented in [146], where an early fusion strategy is used to concatenate features from both modalities. The optimal fusion set is found to be the combination of all video and physiological signals, achieving accuracies of 80.6% and 77.8% for person-specific and generic classifiers, respectively, in detecting baseline and highest tolerable pain using a random forest ensemble-based classifier. Kachele et al. [179] applies both early and late fusion strategies using SVM with linear kernel and random forest for recognizing baseline and highest tolerable pain, achieving accuracies of 68.2% and 76.6% for early and late fusion, respectively.

Continuing the BioVid dataset, Kachele et al. [180] applies early and late fusion techniques with new features included, achieving slightly better results with late fusion (83.1%) than early fusion (82.7%). Thiam et al. [181] proposes a hierarchical fusion architecture that splits multimodal data into three subsets and uses them for the first layer of random forest training, pseudo-inverse mapping, MLP mapping, and a final layer of both pseudo-inverse and MLP fusion mapping, respectively. Kessler et al. [182] takes advantage of the fusion strategy proposed in [181] and applies it to rPPG.

Other studies focus on incorporating additional modalities, such as audio. Velana et al. [37] publishes the SenseEmotion database, which captures video, physiological signals, and audio for the first time. Thiam et al. [183] merges feature from video, physiological signal, and audio data on the SenseEmotion dataset, exploring different data fusion strategies including early fusion, group late fusion, and individual late fusion. Results show that individual late fusion outperforms other strategies slightly on leave-subject-out experiment while group late fusion slightly outperforms on user-specific task. There is also dataset on neonatal pain assessment with video, audio, and physiological signals introduced [45], [173].

Recent studies have explored new fusion approaches. In [184], Bellmann et al. propose a dominant channel fusion approach that identifies the most relevant input channel and combines it with the remaining channels to create an ensemble of classifiers. Bellman et al. [185] proposes a novel late fusion approach that combines a mixture of experts and stacked generalization approaches and is assessed on different datasets involving the bio-physiological modalities EMG, ECG, and EDA. Thiam et al. [160] proposes an information theoretic approach that uses a deep denoising convolutional auto-encoder to learn and aggregate latent representations based on each input channel.

However, it is evident that late fusion, utilizing multiple models as part of an ensemble learning approach, requires significantly more computational power and storage space compared to early fusion methods. As pain assessment is an emerging field, the current focus is predominantly on enhancing predictive accuracy rather than on resource utilization, and discussions on model complexity are relatively scarce. Yet, with the advent of Tiny ML and the rise of edge computing [186], running large models on microprocessors becomes challenging. Consequently, early fusion might gain popularity on edge devices, where the ability to run simpler, more compact models efficiently is crucial. This shift could make early and lightweight fusion approaches more viable and preferred in scenarios where computational resources are limited. Also, with the increasing inclusion of multimodal data, we can envisage future fusion methods potentially incorporating recently developed self-attention algorithms [187].

Discussion

The pain assessment field is faced with several challenges and opportunities for future development. This section will focus on three areas of concern: data, machine learning techniques, and ethical considerations and then propose future research directions.

Data

Automatic pain assessment is challenged by the limited availability of clinical pain data, as most studies have focused on experimental or induced pain. Widely used datasets like BioVid, BP4D+, and X-ITE are collected from healthy volunteers and use external thermal or electrical pain. These studies are conducted under consistent experimental conditions that differ from real-world scenarios. Furthermore, induced pain has different mechanisms than disease pain, which encompasses different types of pain, such as nociceptive and central pain. It is therefore important to test models trained on experimental data on clinical disease pain data. Additionally, more clinical pain data should be collected to facilitate the development of automatic pain assessment models and enable their use in clinical trials.

Pupil dilation has been identified as a promising indicator of brain activity and pain levels. However, in previous studies, pain was often used as the stimuli for measuring brain activity, rather than the focus of the study. Consequently, only a few studies have directly correlated pupil dilation with pain levels. A potential research direction is to include pupil dilation in the automatic pain assessment modality family. Pupil dilation has been shown to be effective in affective computing, with datasets such as the MAHNOB-HCI and SEED datasets containing eye-tracking data that demonstrate the contribution of pupil data to arousal detection. As pain can also be regarded as physiological arousal, transferring pupil dilation to automatic pain assessment studies is a worthwhile area of research.

Personalization of Pain Responses

In the following subsection, we explore personalized pain detection, focusing on the considerable differences in pain experiences among individuals. Pain perception varies widely due to a mix of biological factors and social-psychological influences. These differences are shaped by demographics like gender, age, and ethnicity, which are linked to varying rates of chronic pain. Additionally, factors such as genetic predispositions and psychological processes also significantly impact pain responses, whether in clinical settings or experimental scenarios. Importantly, these elements interact in complex ways, crafting the unique pain experiences of everyone. Research has highlighted that genetic markers associated with pain can differ across genders and ethnicities and interact with psychological aspects like stress, affecting pain perception. These myriads of interacting factors culminates in a distinctive set of influences for each person's experience of pain [188].

In [189], Jiang et al. (2024) introduce a method that enhances pain assessment by incorporating personalized features. They utilize machine learning to analyze individual pain data, enabling the model to tailor its predictions to each patient's unique physiological and psychological characteristics. This approach improves the accuracy of pain management by adapting to personal pain profiles. In [190], Casti et al. (2020) developed a platform to improve pain diagnosis by leveraging personalized data. Using a combination of visual, speech, and physiological indicators, they employ machine learning techniques to tailor assessments to individual patient profiles, enhancing the precision and effectiveness of pain management strategies. Martinez et al. (2017) propose a method to refine pain estimation by integrating personalized features. They employ machine learning to analyze individual facial expressions, allowing the model to adjust its predictions based on each person's unique facial expressiveness score. This approach enhances the accuracy of Visual Analog Scale (VAS) estimations by adapting to individual pain profiles [191].

Most papers on personalized pain assessment claim personalization at the model level, focusing on enhancing machine learning (ML) models to suit individualized approaches or using ML techniques to delve deeper into databases for extracting personalized information to improve predictions. The predominant reliance on public databases for research is evident, as most researchers use these readily available datasets. This reliance restricts personalization efforts to the data provided by these databases, making highly tailored training challenging. Additionally, most pain-related datasets globally are derived from experiments involving artificially induced pain, which must pass rigorous ethical or clinical trial reviews, further limiting the quantity of available data. Looking to the future, personalization will undoubtedly be a crucial focus. It is foreseeable that researchers will collect more personalized data during experiments, including variables like personality traits and ethnicity. This will likely lead to the generation of more nuanced datasets that include varied physiological responses to different pain stimuli, enhancing the granularity and effectiveness of personalized pain management solutions.

Real-time Pain Detection

Building on our earlier discussion about the personalization of pain responses, it's essential to delve into another critically relevant clinical application: real-time monitoring [192]. The goal of such monitoring is not just to detect pain but to enable timely and effective interventions that can significantly enhance patient outcomes. Real-time monitoring of pain becomes particularly crucial in postoperative care, where accurately gauging a patient's pain levels is vital for adjusting analgesic dosages. This not only helps in managing the pain effectively but also minimizes the risk of both undermedication and over-medication, which can lead to complications such as opioid dependency

or inadequate pain relief. In intensive care units (ICUs), the stakes are even higher. Many patients in ICUs are unable to communicate due to their conditions or sedation, making verbal reports of pain unreliable. Here, real-time monitoring systems can play a transformative role by continuously tracking pain indicators through physiological signals such as heart rate, blood pressure, and facial expressions. These data can then be analyzed to provide a dynamic, real-time assessment of pain, informing caregivers when an intervention is necessary. Moreover, real- time monitoring integrates seamlessly with the concept of personalized pain management. By continuously collecting and analyzing data specific to each patient, healthcare providers can tailor their interventions more precisely to the individual's pain profile and response to treatment. This approach not only improves the quality of care but also enhances patient comfort and satisfaction. As technology advances, the potential for real-time pain monitoring grows. Innovations in wearable technology, machine learning algorithms, and data integration are paving the way for even more accurate and responsive pain management systems. These systems promise to transform how pain is managed in healthcare settings, making care more proactive, patient-centered, and effective.

In the academic sphere, the development of real-time pain monitoring is primarily concentrated on two aspects: improving model efficiency to enable fast judgments suitable for real-time applications and developing practical tools like wearable devices and mobile apps to facilitate widespread implementation. Enhancing the processing speed of models involves not only maintaining accuracy but also integrating advanced machine learning technologies such as deep learning. Meanwhile, the development of tools like wearables and mobile applications allows for the non-invasive collection of physiological data and real-time analysis, helping patients and healthcare providers to promptly assess pain levels and treatment effectiveness. This combination of improved models and practical tools is driving pain management towards more precise, personalized, and proactive solutions. In [193], Kong et al. (2021) introduce a smartphone application that enhances real-time pain detection using electrodermal activity (EDA) signals collected from a wrist-worn device. They tested the app with thermal grill and electrical pulse data, demonstrating high accuracy in pain detection with a random forest model. This approach offers a practical solution for objective, near real-time pain assessment in everyday settings. In [79], Dai et al. (2019) address automatic pain detection using a mix of pain and emotion datasets to enhance model robustness, achieving 88.4% accuracy. They critique CNNs for overfitting on biased data and validate their method through experiments on a humanoid robot in physiotherapy, emphasizing the importance of real-time, real-world testing and assessing the system's practical utility and accuracy.

In summary, the advancement of real-time pain monitoring represents a significant enhancement in healthcare, enabling precise and timely interventions that are tailored to the unique needs of each patient. This technology not only improves the accuracy of pain assessments but also enriches the quality of care by integrating cutting-edge machine learning models and wearable technologies. As this field continues to evolve, it holds the promise of transforming pain management into a more responsive, personalized, and patient-centered practice.

ML Techniques

Although deep learning has revolutionized computer vision and physiological signal analysis, traditional machine learning algorithms still dominate the field of physiological signal- based automatic pain assessment. One possible reason for this is that deep learning requires extensive data, which is time-consuming and resource-intensive to collect. As a result, studies often involve only tens of participants, making it difficult to collect deep datasets.

In this context, transfer learning, a prominent topic in artificial intelligence, offers a promising

alternative solution. Transfer learning involves reusing knowledge from a source domain to a new target domain, which can be particularly useful in scenarios where data collection is challenging. Differing data distributions between the source and target domains can lead to performance degradation if models are applied directly. Transfer learning helps bridge this gap, ensuring better model performance across different settings [194].

Kächele et al. [195] proposed an adaptive confidence learning method for personalizing pain intensity estimation systems, demonstrating the efficacy of transfer learning in this field. Feature extraction involved specific preprocessing steps for each signal type, such as bandpass filtering and artifact correction for EMG. A multi-stage ensemble classifier was applied to learn the confidence of a regression system. This method involved selecting confident samples from unlabeled data of the test participants to iteratively adapt the model. Their experiments showed that the adaptive learning approach significantly improved the performance of pain intensity estimation.

Chen et al. [196] implemented "TrAdaboost," a transfer learning algorithm, to improve facial expression recognition, including pain expressions. They used the PAINFUL database, which contains video sequences of 25 patients with shoulder injuries, encompassing 48,398 frames of spontaneous pain expressions. The primary challenge addressed was the variability in pain expressions across different individuals. They proposed an inductive transfer learning algorithm to develop person-specific models. This algorithm first trains a set of weak classifiers on source data from multiple subjects and then selects the most relevant classifiers for the target subject. Experimental results showed that inductive transfer learning significantly improved pain detection accuracy. For example, the Area Under the Curve (AUC) for pain detection increased from 0.769 to 0.782 with just 10 target samples and reached 0.891 with 100 samples. Furthermore, this approach drastically reduced training time compared to traditional methods, making it feasible for rapid retraining in clinical settings.

While traditional machine learning remains prevalent in automatic pain assessment due to data constraints, transfer learning presents a viable alternative. It addresses the challenges associated with varying data distributions and limited dataset sizes, enhancing model robustness and performance. Future research should explore the potential of transfer learning algorithms further, integrating them into clinical practice to improve pain management outcomes.

Ethical Considerations

Automatic pain assessment raises several ethical concerns that need to be addressed. One primary concern is the privacy and security of patients' health data. The use of physiological signals, such as facial expressions, speech patterns, and pupil dilation, to assess pain levels can lead to the collection of sensitive health data. Therefore, it is essential to ensure that the data collected is secure and protected from unauthorized access.

Another ethical consideration is the potential for bias in automatic pain assessment models. Machine learning models are only as good as the data they are trained on, and if the training data is biased, the model will be biased too. Bias can result in inaccurate pain assessment, leading to inadequate pain management and, in some cases, even harm to patients. Therefore, it is crucial to ensure that the data used to train the models is representative and unbiased.

Future Directions

Automated pain assessment has made significant strides in recent years, leveraging technological advancements and data-driven approaches to enhance the accuracy and efficiency of pain detection. However, several promising directions for future research remain unexplored. Addressing these areas could lead to the development of more sophisticated and reliable automated pain assessment systems.

Firstly, integrating data from various sources, such as pupil dilation, voice analysis, and body movement, could offer a more comprehensive understanding of pain. This requires a more comprehensive, clinical and clean database to be released. Secondly, exploring novel deep learning architectures, including transformer-based models and generative adversarial networks (GANs), may yield improved performance in pain assessment tasks. These architectures could capture intricate patterns and dependencies within pain-related data, leading to enhanced predictive capabilities. Thirdly, collaboration with healthcare professionals is crucial to validate the effectiveness and reliability of automated pain assessment systems in real-world clinical settings. Integrating these systems into clinical workflows could provide valuable insights and assist healthcare providers in making informed decisions. Finally, using transfer learning can provide new insights. In scenarios where large, annotated datasets are scarce, exploring transfer learning techniques and methods to adapt models to smaller datasets could prove beneficial. These approaches could enable the development of accurate pain assessment models even with limited training data.

Conclusion

This survey reviewed the current advancements in automated pain assessment using machine learning techniques. Traditional pain assessment methods, reliant on self-reports and observational scales, face significant limitations, particularly for non-communicative patients. We explored various modalities for automated pain detection, including facial expressions, physiological signals, audio, and pupil dilation. While each modality has its strengths, combining multiple modalities can enhance accuracy but also introduces challenges in data fusion and model complexity. Despite progress, challenges remain, such as the scarcity of diverse clinical pain datasets and ethical concerns regarding patient privacy. Personalized pain assessment models are also necessary due to variability in pain perception across populations. Future research should focus on developing more robust algorithms and leveraging deep learning and transfer learning. Collaborative efforts to create comprehensive pain datasets are crucial, as is integrating real-time pain monitoring into clinical practice. In summary, automated pain assessment has the potential to transform pain management. Continued interdisciplinary research and collaboration are key to overcoming current challenges and fully realizing these technologies' benefits.

Acknowledgements

RF is responsible for writing abstract, introduction physiological signals, pupil size, multi-modal study, discussion and conclusion, organizing and formatting the paper. EH is responsible for writing the introduction, facial expression part. RZ is responsible for writing pain mechanism, EDA parts. SR is responsible for collecting information, reviewing and final editing. HH is responsible for reviewing and funding acquisition.

Conflicts of Interest

None declared.

Abbreviations

IASP: International Association for the Study of Pain
VRS: Verbal Rating Scale
VAS: Visual Rating Scale
NRS: Numerical Rating Scale
WBS: Wong-Baker FACES Scale
FPS-R: Faces Pain Scales-Revised
BPS: Behavioral Pain Scale
PAINAD: Pain Assessment in Advanced Dementia
NIPS: Neonatal Infant Pain Scale
PNS: Peripheral Nervous System
CNS: Central Nervous System
TRPV1: Transient receptor potential cation channel subfamily V member 1
TRPM8: Transient receptor potential cation channel subfamily M member 8
ICD-11: International Classification of Diseases-11
TENS: Transcutaneous Electrical Nerve Stimulation
SNS: Sympathetic Nervous System
ML: Machine Learning
FLACC: Face, Legs, Activity, Cry, Consolability Scale
FACS: Facial Action Coding System
AU: Action Units
LBP: Local Binary Pastterns
HOG: Histogram of Oriented Gradients
LSTM: Long Short-Term Momory
AAM: Active Appearance Model
SVM: Support Vector Machine
RVM: Relevance Vector Machine
DCT: Discrete Cosine Transform
RVR: Relevance Vector Regression
PHOG: Pyramid Histogram of Orientation Gradients
PFDL: Personalized Federated Deep Learning
CNN: Convolutional Neural Network
ECCNet: Ensemble of Compact Convolutional Neural Network
CDBN: Convolutional Deep Belief Network
FEAPAS: Facial Expressions-based Automatic Pain Assessment System
N-CNN: Neonatal Convolutional Neural Network
EDLM: Ensemble Deep Learning Model
TCN: Temporal Convolutional Network

C3D: Convolutional 3D
SVR: Support Vector Regression
DMSN: Decomposed Multiscale Spatiotemporal Network
EEG: Electroencephalogram
fMRI: Functional magnetic resonance imaging
fNIRS: Functional Near-Infrared Spectroscopy
ECG: Electrocardiogram
PPG: Photoplethysmograph
SCL: Skin Conductance Level
sEMG: Surface electromyogram
RSP: Respiration
ANN: Artificial Neural Network
RF: Random Forest
FDA: The United States Food and Drug Administration
HR: Heart Rate
BVP: Blood Volumn Pulse
WHOS: Wavelet Higher-Order Spectral
BOLD: Blood Oxygen Level-Dependent
PCA: Principal Component Analysis
LPC: Linear Prediction Coding
STE: Short-Time Energy
MLP: Multilayer Perceptron
GAN: Generative Adversarial Network

Multimedia Appendix 1

Summary of studies table (total 3 pages).

References

1. H. Merskey, “The definition of pain,” European psychiatry, vol. 6, no. 4, pp. 153–159, 1991.
2. A. C. d. C. Williams and K. D. Craig, “Updating the definition of pain,” Pain, vol. 157, no. 11, pp. 2420–2423, 2016.
3. R. J. Yong, P. M. Mullins, and N. Bhattacharyya, “Prevalence of chronic pain among adults in the united states,” Pain, vol. 163, no. 2, pp. e328–e332, 2022.
4. D. J. Gaskin and P. Richard, “The economic costs of pain in the united states,” The journal of pain, vol. 13, no. 8, pp. 715–724, 2012.
5. L. Manchikanti, S. H. B. Fellows, J. W. Janata, V. Pampati, J. S. Grider, and M. V. Boswell, “Opioid epidemic in the united states,” Pain physician, vol. 15, no. 3S, p. ES9, 2012.
6. R. Fink, “Pain assessment: the cornerstone to optimal pain management,” in Baylor university medical center proceedings, vol. 13, no. 3. Taylor & Francis, 2000, pp. 236–239.
7. R. H. Gracely, P. McGrath, and R. Dubner, “Ratio scales of sensory and affective verbal pain descriptors,” Pain, vol. 5, no. 1, pp. 5–18, 1978.
8. H. M. McCormack, J. d. L. David, and S. Sheather, “Clinical applications of visual analogue scales: a critical review,” Psychological medicine, vol. 18, no. 4, pp. 1007–1019, 1988.
9. W. Downie, P. Leatham, V. Rhind, V. Wright, J. Branco, and J. Anderson, “Studies with pain rating scales.” Annals of the rheumatic diseases, vol. 37, no. 4, pp. 378–381, 1978.
10. D. L. Wong and C. M. Baker, “Smiling face as anchor for pain intensity scales,” Pain, vol. 89, no. 2, pp. 295–297, 2001.
11. H. Dehghani, H. Tavangar, and A. Ghandehari, “Validity and reliability of behavioral pain scale in patients with low level of consciousness due to head trauma hospitalized in intensive

care unit," Archives of trauma research, vol. 3, no. 1, 2014.
12. V. Warden, A. C. Hurley, and L. Volicer, "Development and psychometric evaluation of the pain assessment in advanced dementia (painad) scale," Journal of the American Medical Directors Association, vol. 4, no. 1, pp. 9–15, 2003.
13. J. Lawrence, D. Alcock, P. McGrath, J. Kay, S. B. MacMurray, and C. Dulberg, "The development of a tool to assess neonatal pain." Neonatal network: NN, vol. 12, no. 6, pp. 59–66, 1993.
14. J. Kappesser and A. C. d. C. Williams, "Pain estimation: asking the right questions," Pain, vol. 148, no. 2, pp. 184–187, 2010.
15. H. Merskey, D. Albe-Fessard, J. Bonica, A. Carmen, R. Dubner, F. Kerr, and C. Pagni, "Editorial: The need of a taxonomy," Pain, vol. 6, no. 3, pp. 247–252, 1979.
16. R. Gorczyca, R. Filip, and E. Walczak, "Psychological aspects of pain," Ann Agric Environ Med, vol. 1, no. Special issue, pp. 23–27, 2013.
17. E. L. Garland, "Pain processing in the human nervous system: a selective review of nociceptive and biobehavioral pathways," Primary Care: Clinics in Office Practice, vol. 39, no. 3, pp. 561–571, 2012.
18. N. R. Council et al., "Recognition and alleviation of pain in laboratory animals," 2010.
19. E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. Siegelbaum, A. J. Hudspeth, S. Mack et al., Principles of neural science. McGrawhill New York, 2000, vol. 4.
20. D. Julius and A. I. Basbaum, "Molecular mechanisms of nociception," Nature, vol. 413, no. 6852, pp. 203–210, 2001.
21. R.-D. Treede, W. Rief, A. Barke, Q. Aziz, M. I. Bennett, R. Benoliel, M. Cohen, S. Evers, N. B. Finnerup, M. B. First et al., "A classification of chronic pain for icd-11," Pain, vol. 156, no. 6, p. 1003, 2015.
22. J. A. Markenson, "Mechanisms of chronic pain," The American journal of medicine, vol. 101, pp. S6–S18, 1996.
23. D. Borsook, "A future without chronic pain: neuroscience and clinical research," in Cerebrum: the Dana forum on brain science, vol. 2012. Dana Foundation, 2012.
24. S. Mee, B. G. Bunney, C. Reist, S. G. Potkin, and W. E. Bunney, "Psychological pain: a review of evidence," Journal of Psychiatric Research, vol. 40, no. 8, pp. 680–690, 2006.
25. M. J. Bair, R. L. Robinson, W. Katon, and K. Kroenke, "Depression and pain comorbidity: a literature review," Archives of internal medicine, vol. 163, no. 20, pp. 2433–2445, 2003.
26. M. Von Korff and G. Simon, "The relationship between pain and depression," The British journal of psychiatry, vol. 168, no. S30, pp. 101–108, 1996.
27. G. L. Engel, ""psychogenic" pain and the pain-prone patient," The American journal of medicine, vol. 26, no. 6, pp. 899–918, 1959.
28. M. Bassler, H. Krauthauser, and S. Hoffmann, "Inpatient psychotherapy with chronic psychogenic pain patients," Psychotherapie, Psychosomatik, medizinische Psychologie, vol. 44, no. 9-10, pp. 299–307, 1994.
29. S. L. Paxton, "Clinical uses of tens: A survey of physical therapists," Physical therapy, vol. 60, no. 1, pp. 38–44, 1980.
30. T. Ziemssen and S. Kern, "Psychoneuroimmunology–cross-talk between the immune and nervous systems," Journal of neurology, vol. 254, no. 2, pp. II8–II11, 2007.
31. K. L. Teff, "Visceral nerves: vagal and sympathetic innervation," Journal of Parenteral and Enteral Nutrition, vol. 32, no. 5, pp. 569– 571, 2008.
32. S. Singaram, K. Ramakrishnan, J. Selvam, M. Senthil, and V. Narayanamurthy, "Sweat gland morphology and physiology in diabetes, neuropathy, and nephropathy: a review," Archives of physiology and biochemistry, pp. 1–15, 2022.
33. P. Lucey, J. F. Cohn, K. M. Prkachin, P. E. Solomon, and I. Matthews, "Painful data: The unbc-mcmaster shoulder pain expression archive database," in 2011 IEEE International

Conference on Automatic Face & Gesture Recognition (FG). IEEE, 2011, pp. 57–64.
34. S. Walter, S. Gruss, H. Ehleiter, J. Tan, H. C. Traue, P. Werner, A. AlHamadi, S. Crawcour, A. O. Andrade, and G. M. da Silva, “The biovid heat pain database data for the advancement and systematic validation of an automated pain recognition system,” in 2013 IEEE international conference on cybernetics (CYBCO). IEEE, 2013, pp. 128–131.
35. M. A. Haque, R. B. Bautista, F. Noroozi, K. Kulkarni, C. B. Laursen, R. Irani, M. Bellantonio, S. Escalera, G. Anbarjafari, K. Nasrollahi et al., “Deep multimodal pain recognition: a database and comparison of spatio-temporal visual modalities,” in 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018). IEEE, 2018, pp. 250–257.
36. M. S. Aung, S. Kaltwang, B. Romera-Paredes, B. Martinez, A. Singh, M. Cella, M. Valstar, H. Meng, A. Kemp, M. Shafizadeh et al., “The automatic detection of chronic pain-related expression: requirements, challenges and the multimodal emopain dataset,” IEEE transactions on affective computing, vol. 7, no. 4, pp. 435–451, 2015.
37. M. Velana, S. Gruss, G. Layher, P. Thiam, Y. Zhang, D. Schork, V. Kessler, S. Meudt, H. Neumann, J. Kim et al., “The senseemotion database: A multimodal database for the development and systematic validation of an automatic pain-and emotion-recognition system,” in IAPR Workshop on Multimodal Pattern Recognition of Social Signals in Human-Computer Interaction. Springer, 2017, pp. 127–139.
38. S. Gruss, M. Geiger, P. Werner, O. Wilhelm, H. C. Traue, A. AlHamadi, and S. Walter, “Multi-modal signals for analyzing pain responses to thermal and electrical stimuli,” JoVE (Journal of Visualized Experiments), no. 146, p. e59057, 2019.
39. X. Zhang, L. Yin, J. F. Cohn, S. Canavan, M. Reale, A. Horowitz, P. Liu, and J. M. Girard, “Bp4d-spontaneous: a high-resolution spontaneous 3d dynamic facial expression database,” Image and Vision Computing, vol. 32, no. 10, pp. 692–706, 2014.
40. Z. Zhang, J. M. Girard, Y. Wu, X. Zhang, P. Liu, U. Ciftci, S. Canavan, M. Reale, A. Horowitz, H. Yang et al., “Multimodal spontaneous emotion corpus for human behavior analysis,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 3438–3446.
41. S. Brahnam, C.-F. Chuang, F. Y. Shih, and M. R. Slack, “Svm classification of neonatal facial images of pain,” in Fuzzy Logic and Applications: 6th International Workshop, WILF 2005, Crema, Italy, September 15-17, 2005, Revised Selected Papers 6. Springer, 2006, pp. 121–128.
42. D. Harrison, M. Sampson, J. Reszel, K. Abdulla, N. Barrowman, J. Cumber, A. Fuller, C. Li, S. Nicholls, and C. M. Pound, “Too many crying babies: a systematic review of pain management practices during immunizations on youtube,” BMC pediatrics, vol. 14, no. 1, pp. 1–8, 2014.
43. J. Egede, M. Valstar, M. T. Torres, and D. Sharkey, “Automatic neonatal pain estimation: An acute pain in neonates database,” in 2019 8th international conference on affective computing and intelligent interaction (ACII). IEEE, 2019, pp. 1–7.
44. S. Brahnam, L. Nanni, S. McMurtrey, A. Lumini, R. Brattin, M. Slack, and T. Barrier, “Neonatal pain detection in videos using the icopevid dataset and an ensemble of descriptors extracted from gaussian of local descriptors,” Applied Computing and Informatics, vol. 19, no. 1/2, pp. 122–143, 2023.
45. M. S. Salekin, G. Zamzmi, J. Hausmann, D. Goldgof, R. Kasturi, M. Kneusel, T. Ashmeade, T. Ho, and Y. Sun, “Multimodal neonatal procedural and postoperative pain assessment dataset,” Data in Brief, vol. 35, p. 106796, 2021.
46. G. Zamzmi, C.-Y. Pai, D. Goldgof, R. Kasturi, T. Ashmeade, and Y. Sun, “A comprehensive and context-sensitive neonatal pain assessment using computer vision,” IEEE Transactions on Affective Computing, vol. 13, no. 1, pp. 28–45, 2019.

47. M. S. Salekin, G. Zamzmi, D. Goldgof, R. Kasturi, T. Ho, and Y. Sun, “First investigation into the use of deep learning for continuous assessment of neonatal postoperative pain,” in 2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020). IEEE, 2020, pp. 415–419.
48. M. Jiang, R. Mieronkoski, E. Syrjälä, A. Anzanpour, V. Terävä, A. M. Rahmani, S. Salanterä, R. Aantaa, N. Hagelberg, and P. Liljeberg, “Acute pain intensity monitoring with the classification of multiple physiological parameters,” Journal of clinical monitoring and computing, vol. 33, no. 3, pp. 493–507, 2019.
49. M. Jiang, R. Mieronkoski, A. M. Rahmani, N. Hagelberg, S. Salanterä, and P. Liljeberg, “Ultra-short-term analysis of heart rate variability for real-time acute pain monitoring with wearable electronics,” in 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2017, pp. 1025–1032.
50. B. T. Susam, M. Akcakaya, H. Nezamfar, D. Diaz, X. Xu, V. R. de Sa, K. D. Craig, J. S. Huang, and M. S. Goodwin, “Automated pain assessment using electrodermal activity data and machine learning,” in 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2018, pp. 372–375.
51. M. L. Loggia, M. Juneau, and M. C. Bushnell, “Autonomic responses to heat pain: Heart rate, skin conductance, and their relation to verbal ratings and stimulus intensity,” PAIN®, vol. 152, no. 3, pp. 592–598, 2011.
52. L. J. Hadjileontiadis, “Eeg-based tonic cold pain characterization using wavelet higher order spectral features,” IEEE Transactions on Biomedical Engineering, vol. 62, no. 8, pp. 1981–1991, 2015.
53. A. Badura, A. Masłowska, A. My ́sliwiec, and E. Pi ̨etka, “Multimodal signal analysis for pain recognition in physiotherapy using wavelet scattering transform,” Sensors, vol. 21, no. 4, p. 1311, 2021.
54. Y. Chu, X. Zhao, J. Han, and Y. Su, “Physiological signal-based method for measurement of pain intensity,” Frontiers in neuroscience, vol. 11, p. 279, 2017.
55. Y. Chu, X. Zhao, J. Yao, Y. Zhao, and Z. Wu, “Physiological signals based quantitative evaluation method of the pain,” IFAC Proceedings Volumes, vol. 47, no. 3, pp. 2981–2986, 2014.
56. A. J. Hautala, J. Karppinen, and T. Seppänen, “Short-term assessment of autonomic nervous system as a potential tool to quantify pain experience,” in 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2016, pp. 2684–2687.
57. P. Ekman and W. V. Friesen, Facial action coding system: Investigator’s guide. Consulting Psychologists Press, 1978.
58. K. M. Prkachin and P. E. Solomon, “The structure, reliability and validity of pain expression: Evidence from patients with shoulder pain,” Pain, vol. 139, no. 2, pp. 267–274, 2008.
59. A. d. Williams, “Facial expression of pain: an evolutionary account.” Behav Brain Sci, vol. 25, no. 4, pp. 439–455, 2002.
60. A. B. Ashraf, S. Lucey, J. F. Cohn, T. Chen, Z. Ambadar, K. M. Prkachin, and P. E. Solomon, “The painful face–pain expression recognition using active appearance models,” Image and vision computing, vol. 27, no. 12, pp. 1788–1796, 2009.
61. P. Lucey, J. F. Cohn, I. Matthews, S. Lucey, S. Sridharan, J. Howlett, and K. M. Prkachin, “Automatically detecting pain in video through facial action units,” IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 41, no. 3, pp. 664–674, 2010.
62. B. Gholami, W. M. Haddad, and A. R. Tannenbaum, “Relevance vector machine learning for neonate pain intensity assessment using digital imaging,” IEEE Transactions on biomedical engineering, vol. 57, no. 6, pp. 1457–1466, 2010.
63. Z. Hammal and J. F. Cohn, “Automatic detection of pain intensity,” in Proceedings of the

14th ACM international conference on Multimodal interaction, 2012, pp. 47–52.
64. S. Kaltwang, O. Rudovic, and M. Pantic, “Continuous pain intensity estimation from facial expressions,” in International Symposium on Visual Computing. Springer, 2012, pp. 368–377.
65. R. A. Khan, A. Meyer, H. Konik, and S. Bouakaz, “Pain detection through shape and appearance features,” in 2013 IEEE International Conference on Multimedia and Expo (ICME). IEEE, 2013, pp. 1–6.
66. H. Pedersen, “Learning appearance features for pain detection using the unbc-mcmaster shoulder pain expression archive database,” in Computer Vision Systems: 10th International Conference, ICVS 2015, Copenhagen, Denmark, July 6-9, 2015, Proceedings 10. Springer, 2015, pp. 128–136.
67. J. O. Egede, S. Song, T. A. Olugbade, C. Wang, C. D. C. Amanda, H. Meng, M. Aung, N. D. Lane, M. Valstar, and N. Bianchi-Berthouze, “Emopain challenge 2020: Multimodal pain evaluation from facial and bodily expressions,” in 2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020). IEEE, 2020, pp. 849–856.
68. K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” arXiv preprint arXiv:1409.1556, 2014.
69. K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.
70. O. Rudovic, N. Tobis, S. Kaltwang, B. Schuller, D. Rueckert, J. F. Cohn, and R. W. Picard, “Personalized federated deep learning for pain estimation from face images,” arXiv preprint arXiv:2101.04800, 2021.
71. E. Hosseini, R. Fang, R. Zhang, C.-N. Chuah, M. Orooji, S. Rafatirad, S. Rafatirad, and H. Homayoun, “Convolution neural network for pain intensity assessment from facial expression,” in 2022 44th annual international conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, 2022, pp. 2697–2702.
72. E. Barsoum, C. Zhang, C. C. Ferrer, and Z. Zhang, “Training deep networks for facial expression recognition with crowd-sourced label distribution,” in Proceedings of the 18th ACM International Conference on Multimodal Interaction, 2016, pp. 279–283.
73. D. Huang, Z. Xia, L. Li, K. Wang, and X. Feng, “Pain-awareness multistream convolutional neural network for pain estimation,” Journal of Electronic Imaging, vol. 28, no. 4, pp. 043 008–043 008, 2019.
74. A. Semwal and N. D. Londhe, “Eccnet: an ensemble of compact convolution neural network for pain severity assessment from face images,” in 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence). IEEE, 2021, pp. 761–766.
75. R. Kharghanian, A. Peiravi, and F. Moradi, “Pain detection from facial images using unsupervised feature learning approach,” in 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2016, pp. 419–422.
76. R. Kharghanian, A. Peiravi, F. Moradi, and A. Iosifidis, “Pain detection using batch normalized discriminant restricted boltzmann machine layers,” Journal of Visual Communication and Image Representation, vol. 76, p. 103062, 2021.
77. A. Semwal and N. D. Londhe, “Mvfnet: A multi-view fusion network for pain intensity assessment in unconstrained environment,” Biomedical Signal Processing and Control, vol. 67, p. 102537, 2021.
78. T. Alghamdi and G. Alaghband, “Facial expressions based automatic pain assessment system,” Applied Sciences, vol. 12, no. 13, p. 6423, 2022.
79. L. Dai, J. Broekens, and K. P. Truong, “Real-time pain detection in facial expressions for health robotics,” in 2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW). IEEE, 2019, pp. 277–283.

80. I. Karamitsos, I. Seladji, and S. Modak, “A modified cnn network for automatic pain identification using facial expressions,” Journal of Software Engineering and Applications, vol. 14, no. 8, pp. 400–417, 2021.
81. P. D. Barua, N. Baygin, S. Dogan, M. Baygin, N. Arunkumar, H. Fujita, T. Tuncer, R.-S. Tan, E. Palmer, M. M. B. Azizan et al., “Automated detection of pain levels using deep feature extraction from shutter blinds-based dynamic-sized horizontal patches with facial images,” Scientific reports, vol. 12, no. 1, p. 17297, 2022.
82. C. Li, A. Pourtaherian, L. van Onzenoort, W. T. a Ten, and P. De With, “Infant facial expression analysis: towards a real-time video monitoring system using r-cnn and hmm,” IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 5, pp. 1429–1440, 2020.
83. G. Zamzmi, R. Paul, D. Goldgof, R. Kasturi, and Y. Sun, “Pain assessment from facial expression: Neonatal convolutional neural network (n-cnn),” in 2019 International Joint Conference on Neural Networks (IJCNN). IEEE, 2019, pp. 1–7.
84. M. A. Witherow, M. D. Samad, N. Diawara, H. Y. Bar, and K. M. Iftekharuddin, “Deep adaptation of adult-child facial expressions by fusing landmark features,” IEEE Transactions on Affective Computing, 2023.
85. G. Bargshady, X. Zhou, R. C. Deo, J. Soar, F. Whittaker, and H. Wang, “Ensemble neural network approach detecting pain intensity from facial expressions,” Artificial Intelligence in Medicine, vol. 109, p. 101954, 2020.
86. T. Alghamdi and G. Alaghband, “Safepa: An expandable multi-pose facial expressions pain assessment method,” Applied Sciences, vol. 13, no. 12, p. 7206, 2023.
87. G. Bargshady, X. Zhou, R. C. Deo, J. Soar, F. Whittaker, and H. Wang, “The modeling of human facial pain intensity based on temporal convolutional networks trained with video frames in hsv color space,” Applied Soft Computing, vol. 97, p. 106805, 2020.
88. ——, “Enhanced deep learning algorithm development to detect pain intensity from facial expression images,” Expert Systems with Applications, vol. 149, p. 113305, 2020.
89. K. S. Rao, S. G. Koolagudi, and R. R. Vempada, “Emotion recognition from speech using global and local prosodic features,” International journal of speech technology, vol. 16, pp. 143–160, 2013.
90. S. A. Zambach, C. Cai, H. C. C. Helms, B. O. Hald, Y. Dong, J. C. Fordsmann, R. M. Nielsen, J. Hu, M. Lønstrup, B. Brodin et al., “Precapillary sphincters and pericytes at first-order capillaries as key regulators for brain capillary perfusion,” Proceedings of the National Academy of Sciences, vol. 118, no. 26, p. e2023749118, 2021.
91. M. Höfle, R. Kenntner-Mabiala, P. Pauli, and G. W. Alpers, “You can see pain in the eye: pupillometry as an index of pain intensity under different luminance conditions,” International journal of psychophysiology, vol. 70, no. 3, pp. 171–175, 2008.
92. M. A. Connelly, J. T. Brown, G. L. Kearns, R. A. Anderson, S. D. St Peter, and K. A. Neville, “Pupillometry: a non-invasive technique for pain assessment in paediatric patients,” Archives of disease in childhood, vol. 99, no. 12, pp. 1125–1131, 2014.
93. J. H. Medicine, “Electroencephalogram (eeg).”
94. D. Rissacher, R. Dowman, and S. Schuckers, “Identifying frequencydomain features for an eeg-based pain measurement system,” in 2007 IEEE 33rd Annual Northeast Bioengineering Conference. IEEE, 2007, pp. 114–115.
95. T. Adjei, W. Von Rosenberg, V. Goverdovsky, K. Powezka, U. Jaffer, and D. P. Mandic, “Pain prediction from ecg in vascular surgery,” IEEE journal of Translational Engineering in Health and Medicine, vol. 5, pp. 1–10, 2017.
96. D. J. Heeger and D. Ress, “What does fmri tell us about neuronal activity?” Nature reviews neuroscience, vol. 3, no. 2, pp. 142–151, 2002.
97. M. E. Robinson, A. M. O’Shea, J. G. Craggs, D. D. Price, J. E. Letzen, and R. Staud, “Comparison of machine classification algorithms for fibromyalgia: neuroimages versus self-

report," The Journal of Pain, vol. 16, no. 5, pp. 472–477, 2015.
98. Y. Tu, Z. Fu, A. Tan, G. Huang, L. Hu, Y. S. Hung, and Z. Zhang, "A novel and effective fmri decoding approach based on sliced inverse regression and its application to pain prediction," Neurocomputing, vol. 273, pp. 373–384, 2018.
99. W. Shen, Y. Tu, R. L. Gollub, A. Ortiz, V. Napadow, S. Yu, G. Wilson, J. Park, C. Lang, M. Jung et al., "Visual network alterations in brain functional connectivity in chronic low back pain: A resting state functional connectivity and machine learning study," NeuroImage: Clinical, vol. 22, p. 101775, 2019.
100. M. Ferrari and V. Quaresima, "A brief review on the history of human functional near-infrared spectroscopy (fnirs) development and fields of application," Neuroimage, vol. 63, no. 2, pp. 921–935, 2012.
101. K. D. Karunakaran, K. Peng, D. Berry, S. Green, R. Labadie, B. Kussman, and D. Borsook, "Nirs measures in pain and analgesia: fundamentals, features, and function," Neuroscience & Biobehavioral Reviews, vol. 120, pp. 335–353, 2021.
102. R. Fernandez Rojas, X. Huang, and K.-L. Ou, "A machine learning approach for the identification of a biomarker of human pain using fnirs," Scientific reports, vol. 9, no. 1, pp. 1–12, 2019.
103. A. H. Association, "Electrocardiogram (ecg or ekg)."
104. P. Werner, A. Al-Hamadi, R. Niese, S. Walter, S. Gruss, and H. C. Traue, "Towards pain monitoring: Facial expression, head pose, a new database, an automatic system and remaining challenges," in Proceedings of the British Machine Vision Conference, 2013, pp. 1–13.
105. J. Allen, "Photoplethysmography and its application in clinical physiological measurement," Physiological measurement, vol. 28, no. 3, p. R1, 2007.
106. D. T. Lykken and P. H. Venables, "Direct measurement of skin conductance: A proposal for standardization," Psychophysiology, vol. 8, no. 5, pp. 656–672, 1971.
107. S. L. Pullman, D. S. Goodin, A. I. Marquinez, S. Tabbal, and M. Rubin, "Clinical utility of surface emg: report of the therapeutics and technology assessment subcommittee of the american academy of neurology," Neurology, vol. 55, no. 2, pp. 171–177, 2000.
108. J. N. Mak, Y. Hu, and K. Luk, "Ica-based ecg removal from surface electromyography and its effect on low back pain assessment," in 2007 3rd International IEEE/EMBS Conference on Neural Engineering. IEEE, 2007, pp. 646–649.
109. N. Aguila-Roca, "Essential physiology."
110. M. Tavakolian and A. Hadid, "Deep spatiotemporal representation of the face for automatic pain intensity estimation," in 2018 24th International Conference on Pattern Recognition (ICPR). IEEE, 2018, pp. 350–354.
111. ——, "A spatiotemporal convolutional neural network for automatic pain intensity estimation from facial dynamics," International Journal of Computer Vision, vol. 127, pp. 1413–1425, 2019.
112. Y. Huang, L. Qing, S. Xu, L. Wang, and Y. Peng, "Hybnet: a hybrid network structure for pain intensity estimation," The Visual Computer, pp. 1–12, 2022.
113. J. Wang and H. Sun, "Pain intensity estimation using deep spatiotemporal and handcrafted features," IEICE TRANSACTIONS on Information and Systems, vol. 101, no. 6, pp. 1572–1580, 2018.
114. W. C. de Melo, E. Granger, and M. B. Lopez, "Facial expression analysis using decomposed multiscale spatiotemporal networks," arXiv preprint arXiv:2203.11111, 2022.
115. E. Granger, P. Cardinal et al., "Deep domain adaptation for ordinal regression of pain intensity estimation using weakly-labelled videos," arXiv preprint arXiv:2008.06392, 2020.
116. R. G. Praveen, E. Granger, and P. Cardinal, "Deep weakly supervised domain adaptation for pain localization in videos," in 2020 15th IEEE International Conference on

Automatic Face and Gesture Recognition (FG 2020). IEEE, 2020, pp. 473–480.
117. J. Carreira and A. Zisserman, “Quo vadis, action recognition? a new model and the kinetics dataset,” in proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 6299–6308.
118. L. Shu, J. Xie, M. Yang, Z. Li, Z. Li, D. Liao, X. Xu, and X. Yang, “A review of emotion recognition using physiological signals,” Sensors, vol. 18, no. 7, p. 2074, 2018.
119. W. Li, Z. Zhang, and A. Song, “Physiological-signal-based emotion recognition: An odyssey from methodology to philosophy,” Measurement, vol. 172, p. 108747, 2021.
120. P. Panavaranan and Y. Wongsawat, “Eeg-based pain estimation via fuzzy logic and polynomial kernel support vector machine,” in The 6th 2013 Biomedical Engineering International Conference. IEEE, 2013, pp. 1–4.
121. V. Vijayakumar, M. Case, S. Shirinpour, and B. He, “Quantifying and characterizing tonic thermal pain across subjects from eeg data using random forest models,” IEEE Transactions on Biomedical Engineering, vol. 64, no. 12, pp. 2988–2996, 2017.
122. T. D. Wager, L. Y. Atlas, M. A. Lindquist, M. Roy, C.-W. Woo, and E. Kross, “An fmri-based neurologic signature of physical pain,” New England Journal of Medicine, vol. 368, no. 15, pp. 1388–1397, 2013.
123. J. J. Meeuse, M. S. Löwik, S. A. Löwik, E. Aarden, A. M. van Roon, R. O. Gans, M. van Wijhe, J. D. Lefrandt, and A. K. Reyners, “Heart rate variability parameters do not correlate with pain intensity in healthy volunteers,” Pain Medicine, vol. 14, no. 8, pp. 1192–1201, 2013.
124. E. Hosseini, R. Fang, R. Zhang, S. Rafatirad, and H. Homayoun, “Emotion and stress recognition utilizing galvanic skin response and wearable technology: A real-time approach for mental health care,” in 2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2023, pp. 1125–1131.
125. E. Hosseini, R. Fang, R. Zhang, A. Parenteau, S. Hang, S. Rafatirad, C. Hostinar, M. Orooji, and H. Homayoun, “A low cost eda-based stress detection using machine learning,” in 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2022, pp. 2619–2623.
126. R. Merletti and D. Farina, Surface electromyography: physiology, engineering, and applications. John Wiley & Sons, 2016.
127. J. Srinivasan and V. Balasubramanian, “Low back pain and muscle fatigue due to road cycling—an semg study,” Journal of Bodywork and Movement Therapies, vol. 11, no. 3, pp. 260–266, 2007.
128. M. Jiang, A.-M. Rahmani, T. Westerlund, P. Liljeberg, and H. Tenhunen, “Facial expression recognition with semg method,” in 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing. IEEE, 2015, pp. 981–988.
129. Z. Zhang, R. Zhang, C.-W. Chang, Y. Guo, Y.-W. Chi, and T. Pan, “iwrap: A theranostic wearable device with real-time vital monitoring and auto-adjustable compression level for venous thromboembolism,” IEEE Transactions on Biomedical Engineering, vol. 68, no. 9, pp. 2776–2786, 2021.
130. R. Zhang, R. Fang, C. Fang, H. Homayoun, and G. Goncu Berk, “Privee: A wearable for real-time bladder monitoring system,” in Adjunct Proceedings of the 2023 ACM International Joint Conference on Pervasive and Ubiquitous Computing & the 2023 ACM International Symposium on Wearable Computing, 2023, pp. 291–295.
131. EMPATICA, “Empatica e4 wristband,” https://www.empatica.com/en-int/research/e4/.
132. EMOTIV, “Emotiv epoc ecg measurement headset,” https://www. emotiv.com/epoc/.

133. BIOHARNESS, “Bioharness 3,” https://www.zephyranywhere.com/ media/download/ bioharness3-user-manual.pdf.
134. T. A. Ajayi, L. Salongo, Y. Zang, N. Wineinger, and S. Steinhubl, “Mobile health-collected biophysical markers in children with serious illness-related pain,” Journal of Palliative Medicine, vol. 24, no. 4, pp. 580–588, 2021.
135. G. Nazari, J. C. MacDermid, K. E. Sinden, J. Richardson, and A. Tang, “Reliability of zephyr bioharness and fitbit charge measures of heart rate and activity at rest, during the modified canadian aerobic fitness test, and recovery,” The Journal of Strength & Conditioning Research, vol. 33, no. 2, pp. 559–571, 2019.
136. J. C. Rawstorn, N. Gant, I. Warren, R. N. Doughty, N. Lever, K. K. Poppe, and R. Maddison, “Measurement and data transmission validity of a multi-biosensor system for real-time remote exercise monitoring among cardiac patients,” JMIR rehabilitation and assistive technologies, vol. 2, no. 1, p. e3633, 2015.
137. F. Løberg, V. Goebel, and T. Plagemann, “Quantifying the signal quality of low-cost respiratory effort sensors for sleep apnea monitoring,” in Proceedings of the 3rd International Workshop on Multimedia for Personal Health and Health Care, 2018, pp. 3–11.
138. R. Fang, R. Zhang, E. Hosseini, C. Fang, S. Rafatirad, and H. Homayoun, “Introducing an open-source python toolkit for machine learning research in physiological signal based affective computing,” in 2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2023, pp. 1890–1894.
139. D. Makowski, T. Pham, Z. J. Lau, J. C. Brammer, F. Lespinasse, H. Pham, C. Schölzel, and S. A. Chen, “Neurokit2: A python toolbox for neurophysiological signal processing,” Behavior research methods, pp. 1–8, 2021.
140. L. Cabañero-Gomez, R. Hervas, I. Gonzalez, and L. Rodriguez-Benitez, “eeglib: A python module for eeg feature extraction,” SoftwareX, vol. 15, p. 100745, 2021.
141. V. Iashin and other contributors, “Video features,” https://github.com/ v-iashin/video_features, 2020.
142. R. Lenain, J. Weston, A. Shivkumar et al., “Surfboard: Audio feature extraction for modern machine learning [internet], 2020. arxiv. org.”
143. N. Behzadfar, “A brief overview on analysis and feature extraction of electroencephalogram signals,” Signal Processing and Renewable Energy, vol. 6, no. 1, pp. 39–64, 2022.
144. M. M. Van Der Miesen, M. A. Lindquist, and T. D. Wager, “Neuroimaging-based biomarkers for pain: state of the field and current directions,” Pain reports, vol. 4, no. 4, 2019.
145. F. Shaffer and J. P. Ginsberg, “An overview of heart rate variability metrics and norms,” Frontiers in public health, p. 258, 2017.
146. P. Werner, A. Al-Hamadi, R. Niese, S. Walter, S. Gruss, and H. C. Traue, “Automatic pain recognition from video and biomedical signals,” in 2014 22nd International Conference on Pattern Recognition. IEEE, 2014, pp. 4582–4587.
147. S. Gruss, R. Treister, P. Werner, H. C. Traue, S. Crawcour, A. Andrade, and S. Walter, “Pain intensity recognition rates via biopotential feature patterns with support vector machines,” PloS one, vol. 10, no. 10, p. e0140330, 2015.
148. E. Campbell, A. Phinyomark, and E. Scheme, “Feature extraction and selection for pain recognition using peripheral physiological signals,” Frontiers in neuroscience, vol. 13, p. 437, 2019.
149. M. Kächele, P. Thiam, M. Amirian, F. Schwenker, and G. Palm, “Methods for person-centered continuous pain intensity assessment from bio-physiological channels,” IEEE Journal of Selected Topics in Signal Processing, vol. 10, no. 5, pp. 854–864, 2016.
150. R. Fang, R. Zhang, S. M. Hosseini, M. Faghih, S. Rafatirad, S. Rafatirad, and H.

Homayoun, “Pain level modeling of intensive care unit patients with machine learning methods: An effective congeneric clustering-based approach,” in 2022 4th International Conference on Intelligent Medicine and Image Processing, 2022, pp. 89–95.
151. S. Walter, S. Gruss, K. Limbrecht-Ecklundt, H. C. Traue, P. Werner, A. Al-Hamadi, N. Diniz, G. M. d. Silva, and A. O. Andrade, “Automatic pain quantification using autonomic parameters,” Psychology & Neuroscience, vol. 7, no. 3, pp. 363–380, 2014.
152. K. Nakano, Y. Ota, H. Ukai, K. Nakamura, and H. Fujita, “Frequency detection method based on recursive dft algorithm,” in Proceedings of the 14th Power Systems Computation. Conference (PSCC’02), Paper, vol. 5, 2002, pp. 1–7.
153. C. Cao and S. Slobounov, “Application of a novel measure of eeg non-stationarity as ‘shannon-entropy of the peak frequency shifting’for detecting residual abnormalities in concussed individuals,” Clinical Neurophysiology, vol. 122, no. 7, pp. 1314–1321, 2011.
154. W. Chen, J. Zhuang, W. Yu, and Z. Wang, “Measuring complexity using fuzzyen, apen, and sampen,” Medical engineering & physics, vol. 31, no. 1, pp. 61–68, 2009.
155. H.-M. Chen, P. K. Varshney, and M. K. Arora, “Performance of mutual information similarity measure for registration of multitemporal remote sensing images,” IEEE transactions on Geoscience and Remote Sensing, vol. 41, no. 11, pp. 2445–2454, 2003.
156. P. Bellmann, P. Thiam, H. A. Kestler, and F. Schwenker, “Machine learning-based pain intensity estimation: Where pattern recognition meets chaos theory—an example based on the biovid heat pain database,” IEEE Access, vol. 10, pp. 102 770–102 777, 2022.
157. P. Gouverneur, F. Li, W. M. Adamczyk, T. M. Szikszay, K. Luedtke, and M. Grzegorzek, “Comparison of feature extraction methods for physiological signals for heat-based pain recognition,” Sensors, vol. 21, no. 14, p. 4838, 2021.
158. E. Othman, P. Werner, F. Saxen, M.-A. Fiedler, and A. Al-Hamadi, “An automatic system for continuous pain intensity monitoring based on analyzing data from uni-, bi-, and multi-modality,” Sensors, vol. 22, no. 13, p. 4992, 2022.
159. F. Pouromran, Y. Lin, and S. Kamarthi, “Personalized deep bi-lstm rnn based model for pain intensity classification using eda signal,” Sensors, vol. 22, no. 21, p. 8087, 2022.
160. P. Thiam, H. Hihn, D. A. Braun, H. A. Kestler, and F. Schwenker, “Multi-modal pain intensity assessment based on physiological signals: A deep learning perspective,” Frontiers in Physiology, vol. 12, p. 720464, 2021.
161. D. H. Wolpert and W. G. Macready, “No free lunch theorems for optimization,” IEEE transactions on evolutionary computation, vol. 1, no. 1, pp. 67–82, 1997.
162. C. Cortes and V. Vapnik, “Support vector machine,” Machine learning, vol. 20, no. 3, pp. 273–297, 1995.
163. A. Verikas, A. Gelzinis, and M. Bacauskiene, “Mining data with random forests: A survey and results of new tests,” Pattern recognition, vol. 44, no. 2, pp. 330–349, 2011.
164. L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, Classification and regression trees. Routledge, 2017.
165. L. Breiman, “Random forests,” Machine learning, vol. 45, no. 1, pp. 5–32, 2001.
166. L. H. Gilpin, D. Bau, B. Z. Yuan, A. Bajwa, M. Specter, and L. Kagal, “Explaining explanations: An overview of interpretability of machine learning,” in 2018 IEEE 5th International Conference on data science and advanced analytics (DSAA). IEEE, 2018, pp. 80–89.
167. M. Pal and P. M. Mather, “An assessment of the effectiveness of decision tree methods for land cover classification,” Remote sensing of environment, vol. 86, no. 4, pp. 554–565, 2003.
168. R. Fang, R. Zhang, E. Hosseini, A. M. Parenteau, S. Hang, S. Rafatirad, C. E. Hostinar, M. Orooji, and H. Homayoun, “Prevent over-fitting and redundancy in physiological signal analyses for stress detection,” in 2022 IEEE International Conference on

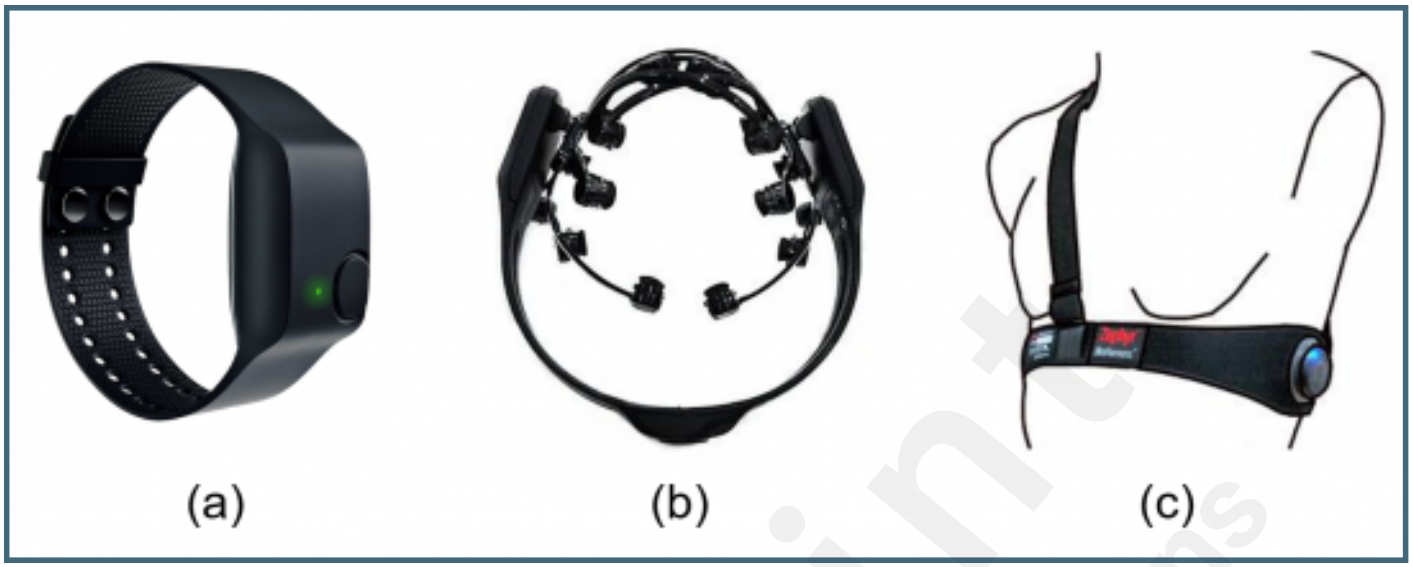
Bioinformatics and Biomedicine (BIBM). IEEE, 2022, pp. 2585–2588.
169. E. K. Naeini, S. Shahhosseini, A. Subramanian, T. Yin, A. M. Rahmani, and N. Dutt, “An edge-assisted and smart system for real-time pain monitoring,” in 2019 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE). IEEE, 2019, pp. 47–52.
170. P. Werner, A. Al-Hamadi, S. Gruss, and S. Walter, “Twofoldmultimodal pain recognition with the x-ite pain database,” in 2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW). IEEE, 2019, pp. 290– 296.
171. C. Fang, N. Miao, S. Srivastav, J. Liu, R. Zhang, R. Fang, A. Asmita, R. Tsang, N. Nazari, H. Wang et al., “Large language models for code analysis: Do llms really do their job?” arXiv preprint arXiv:2310.12357, 2023.
172. D. Lopez-Martinez and R. Picard, “Multi-task neural networks for personalized pain recognition from physiological signals,” in 2017 Seventh International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW). IEEE, 2017, pp. 181–184.
173. M. S. Salekin, G. Zamzmi, D. Goldgof, R. Kasturi, T. Ho, and Y. Sun, “Multimodal spatio-temporal deep learning approach for neonatal postoperative pain assessment,” Computers in Biology and Medicine, vol. 129, p. 104150, 2021.
174. J. O. Pinzon-Arenas, Y. Kong, K. H. Chon, and H. F. Posada-Quintero, “Design and evaluation of deep learning models for continuous acute pain detection based on phasic electrodermal activity,” IEEE Journal of Biomedical and Health Informatics, 2023.
175. A. L. Bertrand, J. B. S. Garcia, E. B. Viera, A. M. Santos, and R. H. Bertrand, “Pupillometry: the influence of gender and anxiety on the pain response,” Pain Physician, vol. 16, no. 3, p. E257, 2013.
176. C. R. Chapman, S. Oka, D. H. Bradshaw, R. C. Jacobson, and G. W. Donaldson, “Phasic pupil dilation response to noxious stimulation in normal volunteers: relationship to brain evoked potentials and pain report,” Psychophysiology, vol. 36, no. 1, pp. 44–52, 1999.
177. J. C. Eisenach, R. Curry, C. A. Aschenbrenner, R. C. Coghill, and T. T. Houle, “Pupil responses and pain ratings to heat stimuli: Reliability and effects of expectations and a conditioning pain stimulus,” Journal of neuroscience methods, vol. 279, pp. 52–59, 2017.
178. L. Wang, Y. Guo, B. Dalip, Y. Xiao, R. D. Urman, and Y. Lin, “An experimental study of objective pain measurement using pupillary response based on genetic algorithm and artificial neural network,” Applied Intelligence, vol. 52, no. 2, pp. 1145–1156, 2022.
179. M. Kächele, P. Werner, A. Al-Hamadi, G. Palm, S. Walter, and F. Schwenker, “Bio-visual fusion for person-independent recognition of pain intensity,” in International Workshop on Multiple Classifier Systems. Springer, 2015, pp. 220–230.
180. M. Kächele, P. Thiam, M. Amirian, P. Werner, S. Walter, F. Schwenker, and G. Palm, “Multimodal data fusion for person-independent, continuous estimation of pain intensity,” in International Conference on Engineering Applications of Neural Networks. Springer, 2015, pp. 275–285.
181. P. Thiam, V. Kessler, and F. Schwenker, “Hierarchical combination of video features for personalised pain level recognition.” in ESANN, 2017, pp. 465–470.
182. V. Kessler, P. Thiam, M. Amirian, and F. Schwenker, “Multimodal fusion including camera photoplethysmography for pain recognition,” in 2017 International Conference on Companion Technology (ICCT). IEEE, 2017, pp. 1–4.
183. P. Thiam and F. Schwenker, “Multi-modal data fusion for pain intensity assessment and classification,” in 2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA). IEEE, 2017, pp. 1–6.
184. P. Bellmann, P. Thiam, and F. Schwenker, “Dominant channel fusion architectures-an

intelligent late fusion approach," in 2020 International Joint Conference on Neural Networks (IJCNN). IEEE, 2020, pp. 1–8.
185. ——, "Using meta labels for the training of weighting models in a sample-specific late fusion classification architecture," in 2020 25th International Conference on Pattern Recognition (ICPR). IEEE, 2021, pp. 2604–2611.
186. F. Oliveira, D. G. Costa, F. Assis, and I. Silva, "Internet of intelligent things: A convergence of embedded systems, edge computing and machine learning," Internet of Things, p. 101153, 2024.
187. Y. Xiong, Z. Zeng, R. Chakraborty, M. Tan, G. Fung, Y. Li, and V. Singh, "Nyströmformer: A nyström-based algorithm for approximating self-attention," in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35, no. 16, 2021, pp. 14 138–14 148.
188. C. S. Nielsen, R. Staud, and D. D. Price, "Individual differences in pain sensitivity: measurement, causation, and consequences," The journal of pain, vol. 10, no. 3, pp. 231–237, 2009.
189. M. Jiang, R. Rosio, S. Salanterä, A. M. Rahmani, P. Liljeberg, D. S. da Silva, V. H. C. de Albuquerque, and W. Wu, "Personalized and adaptive neural networks for pain detection from multi-modal physiological features," Expert Systems with Applications, vol. 235, p. 121082, 2024.
190. P. Casti, A. Mencattini, J. Filippi, M. D'Orazio, M. C. Comes, D. Di Giuseppe, and E. Martinelli, "A personalized assessment platform for non-invasive monitoring of pain," in 2020 IEEE International Symposium on Medical Measurements and Applications (MeMeA). IEEE, 2020, pp. 1–5.
191. D. Lopez Martinez, R. Picard et al., "Personalized automatic estimation of self-reported pain intensity from facial expressions," in Proceedings of the IEEE conference on computer vision and pattern recognition workshops, 2017, pp. 70–79.
192. R. Zhang, R. Fang, Z. Zhang, E. Hosseini, M. Orooji, H. Homayoun, and G. Goncu-Berk, "Short: Real-time bladder monitoring by bioimpedance analysis to aid urinary incontinence," in Proceedings of the 8th ACM/IEEE International Conference on Connected Health: Applications, Systems and Engineering Technologies, 2023, pp. 138– 142.
193. Y. Kong, H. F. Posada-Quintero, and K. H. Chon, "Real-time high-level acute pain detection using a smartphone and a wrist-worn electrodermal activity sensor," Sensors, vol. 21, no. 12, p. 3956, 2021.
194. R. Fang, R. Zhang, E. Hosseini, A. M. Parenteau, S. Hang, S. Rafatirad, C. E. Hostinar, M. Orooji, and H. Homayoun, "Towards generalized ml model in automated physiological arousal computing: A transfer learning-based domain generalization approach," in 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2022, pp. 2577–2584.
195. M. Kächele, M. Amirian, P. Thiam, P. Werner, S. Walter, G. Palm, and F. Schwenker, "Adaptive confidence learning for the personalization of pain intensity estimation systems," Evolving Systems, vol. 8, no. 1, pp. 71–83, 2017.
196. J. Chen, X. Liu, P. Tu, and A. Aragones, "Learning person-specific models for facial expression and action unit recognition," Pattern Recognition Letters, vol. 34, no. 15, pp. 1964–1970, 2013.
197. A. Phinyomark, P. Phukpattaranont, and C. Limsakul, "Feature reduction and selection for emg signal classification," Expert systems with applications, vol. 39, no. 8, pp. 7420–7431, 2012.
198. A. Phinyomark and E. Scheme, "An investigation of temporally inspired time domain features for electromyographic pattern recognition," in 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2018,

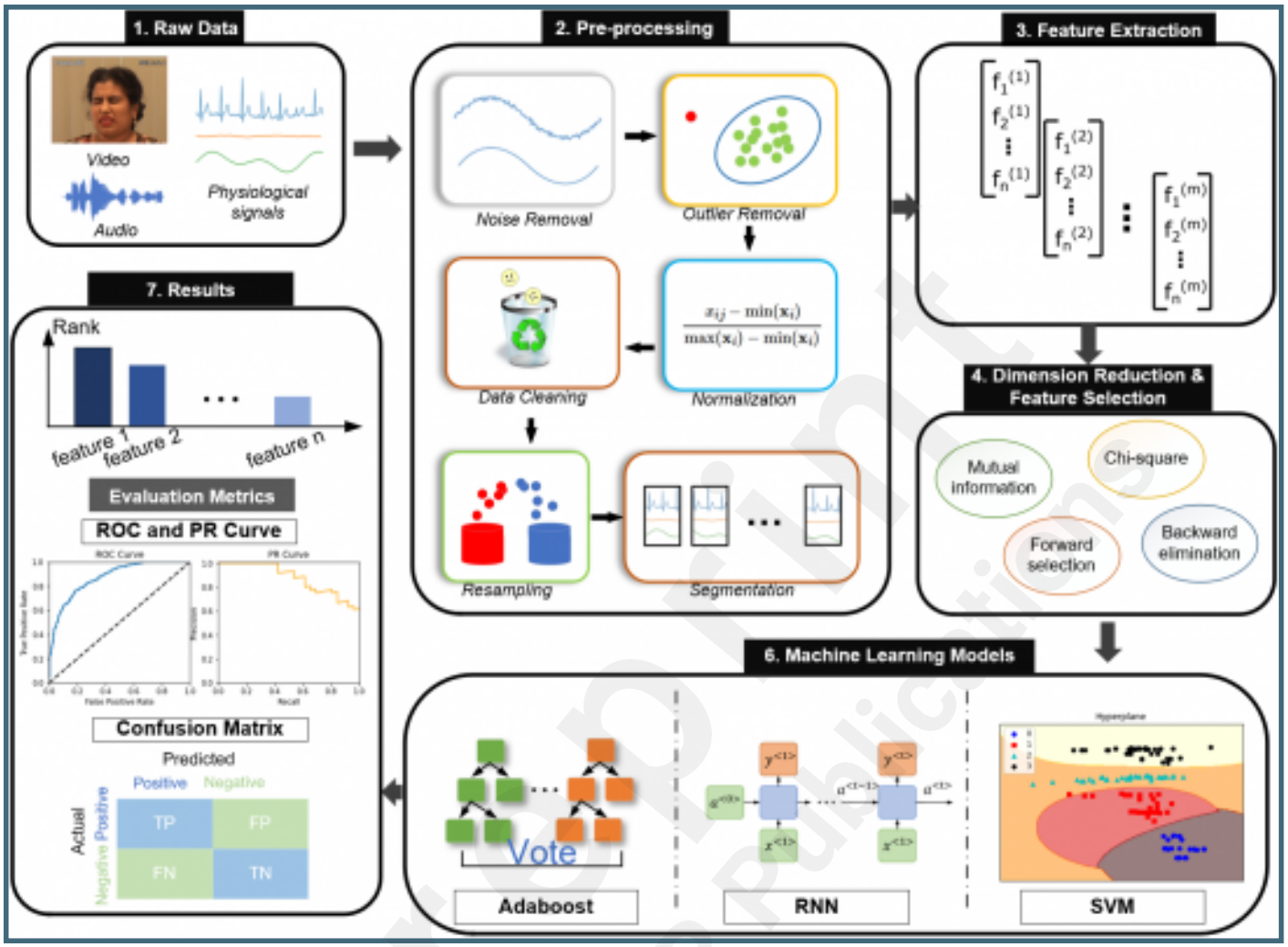
pp. 5236– 5240.
199. S. M. Pincus, “Approximate entropy as a measure of system complexity.” Proceedings of the National Academy of Sciences, vol. 88, no. 6, pp. 2297–2301, 1991.
200. B. Kosko, “Fuzzy entropy and conditioning,” Information sciences, vol. 40, no. 2, pp. 165–174, 1986.
201. J. S. Richman and J. R. Moorman, “Physiological time-series analysis using approximate entropy and sample entropy,” American Journal of Physiology-Heart and Circulatory Physiology, 2000.
202. J. Lin, “Divergence measures based on the shannon entropy,” IEEE Transactions on Information theory, vol. 37, no. 1, pp. 145–151, 1991.
203. A. Zhang, B. Yang, and L. Huang, “Feature extraction of eeg signals using power spectral entropy,” in 2008 international conference on BioMedical engineering and informatics, vol. 2. IEEE, 2008, pp. 435–439.
204. H. L. Kennedy, A new statistical measure of signal similarity. IEEE, 2007.
205. S. Dukic, P. M. Iyer, K. Mohr, O. Hardiman, E. C. Lalor, and B. Nasseroleslami, “Estimation of coherence using the median is robust against eeg artefacts,” in 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2017, pp. 3949–3952.

Supplementary Files

Untitled.

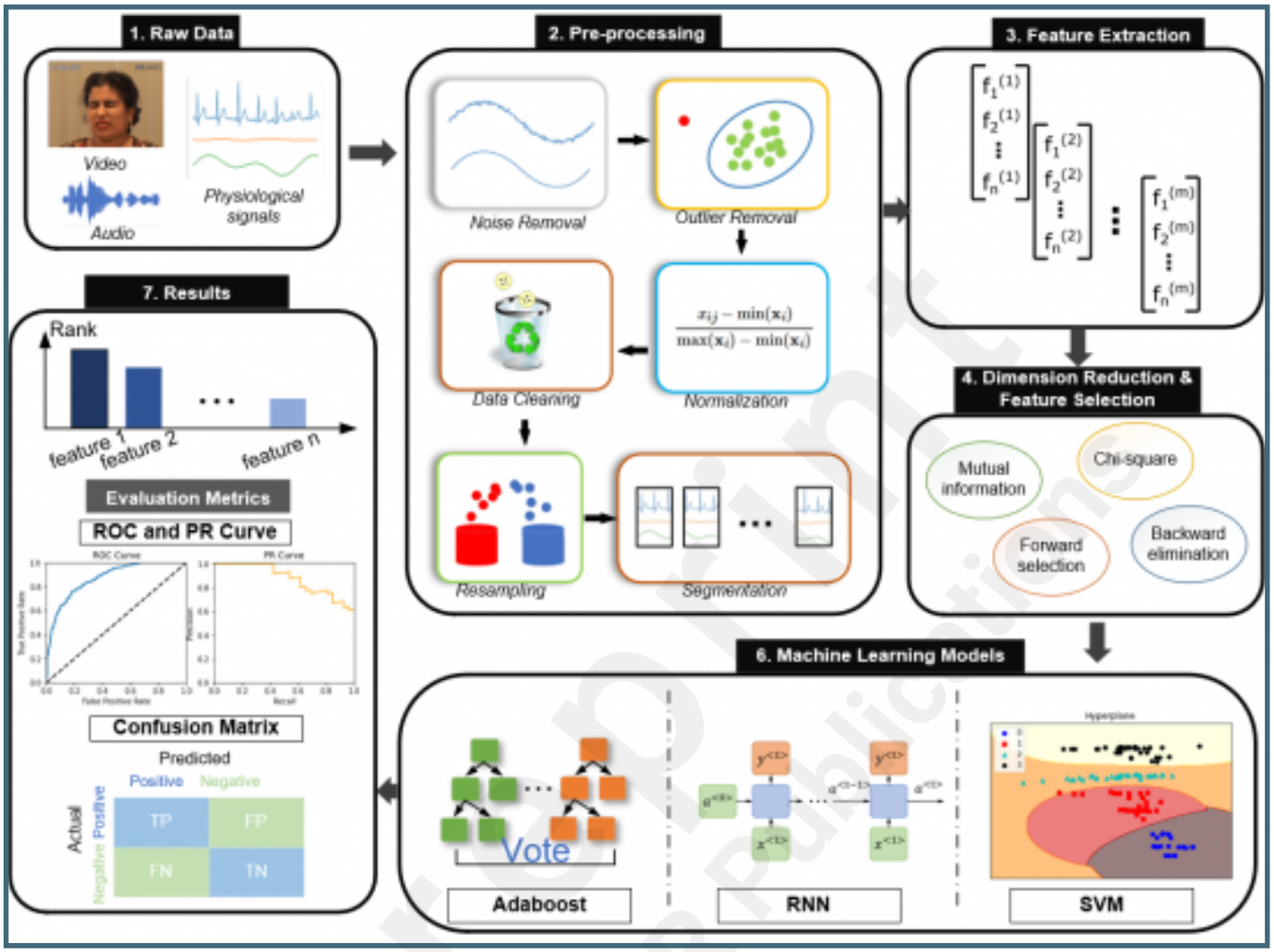


Untitled.


1. Raw Data
Video
Audio
Physiological signals
2. Pre-processing
Noise Removal
Outlier Removal
Data Cleaning
Normalization
Resampling
Segmentation
3. Feature Extraction
4. Dimension Reduction & Feature Selection
Mutual information
Chi-square
Forward selection
Backward elimination
6. Machine Learning Models
Vote
Adaboost
RNN
SVM
7. Results
Rank
feature 1
feature 2
feature n
Evaluation Metrics
ROC and PR Curve
Confusion Matrix
Predicted
Positive
Negative
Actual
TP
FP
FN
TN

Figures

Automated pain assessment typical pipeline.



Pain mechanism.

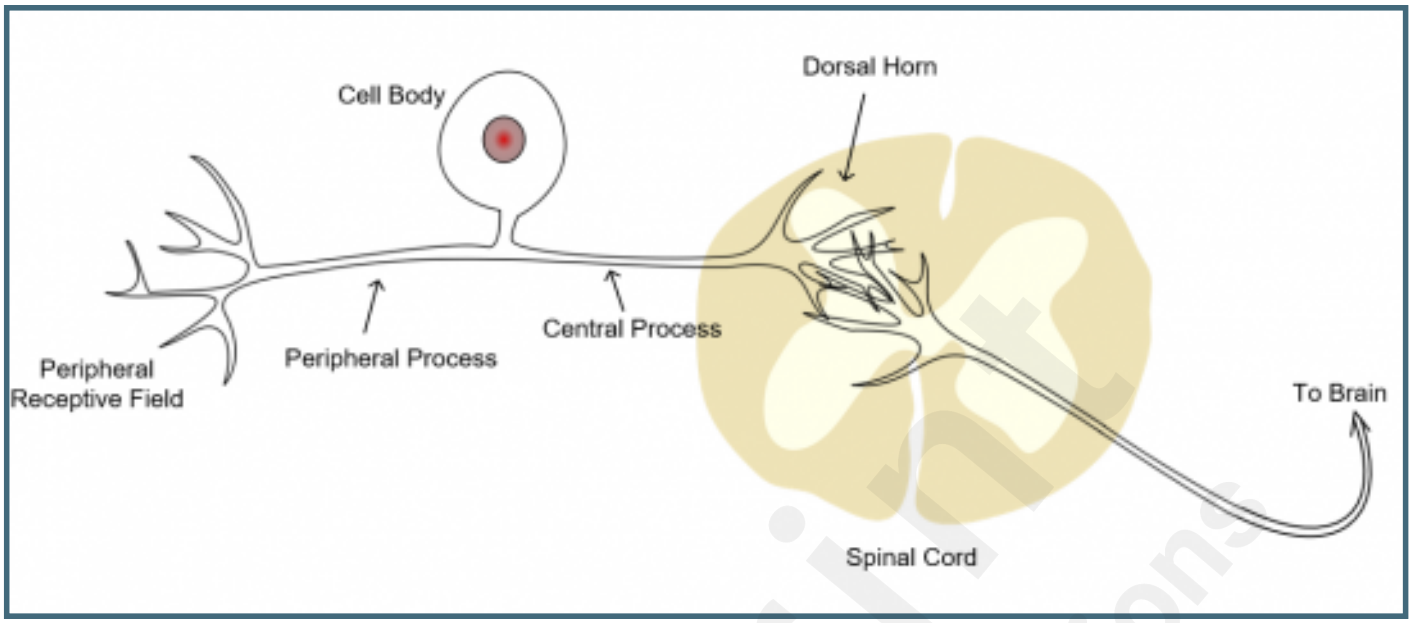
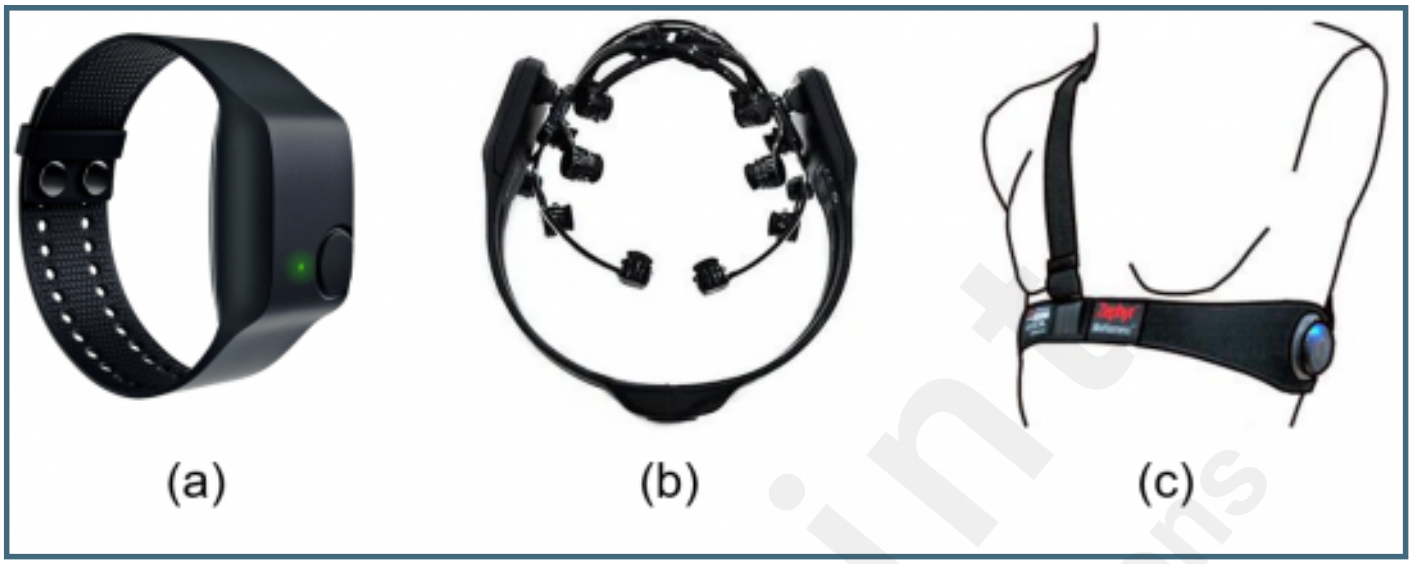
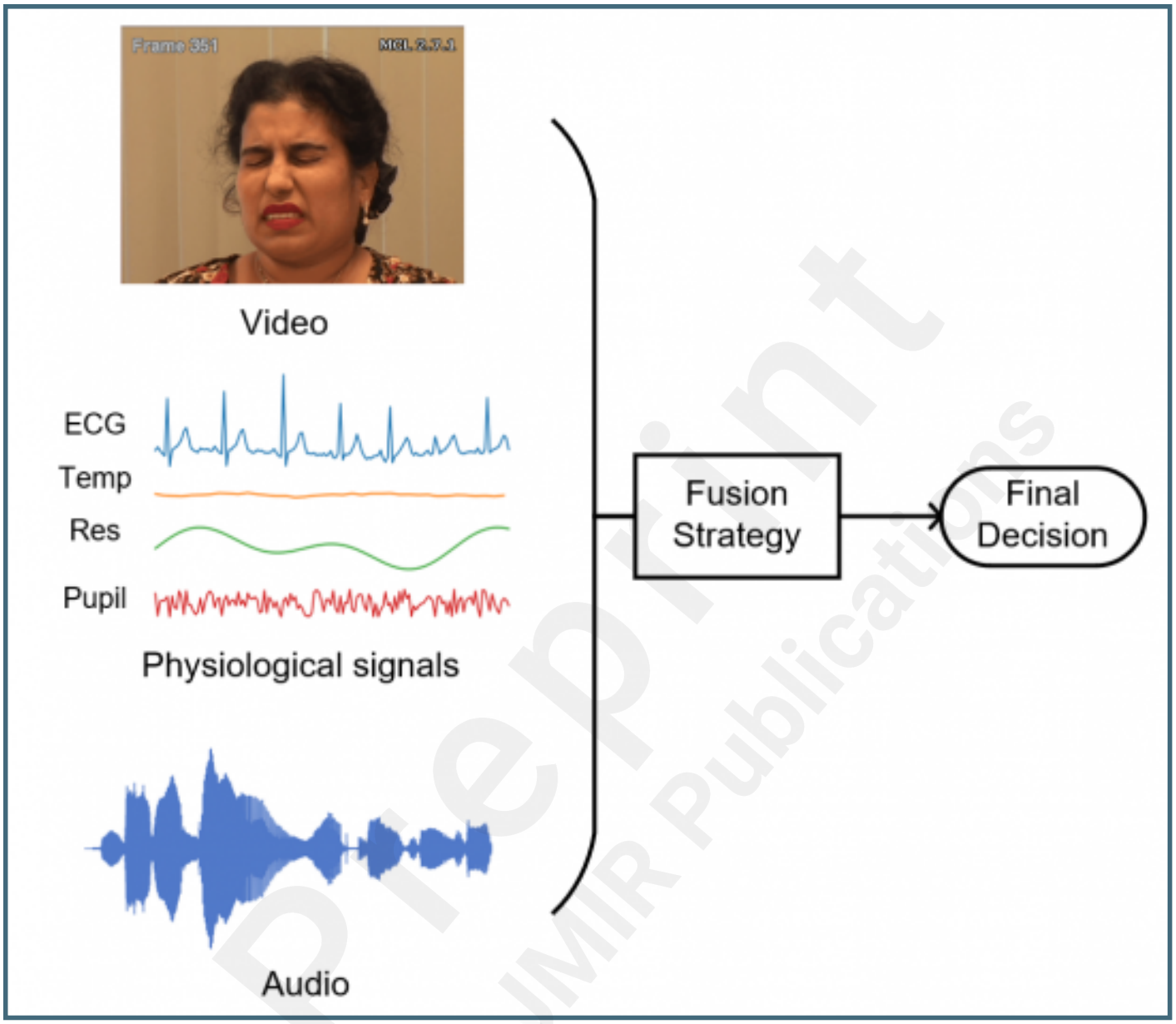


Image frame samples of the UNBC-McMaster shoulder pain database.

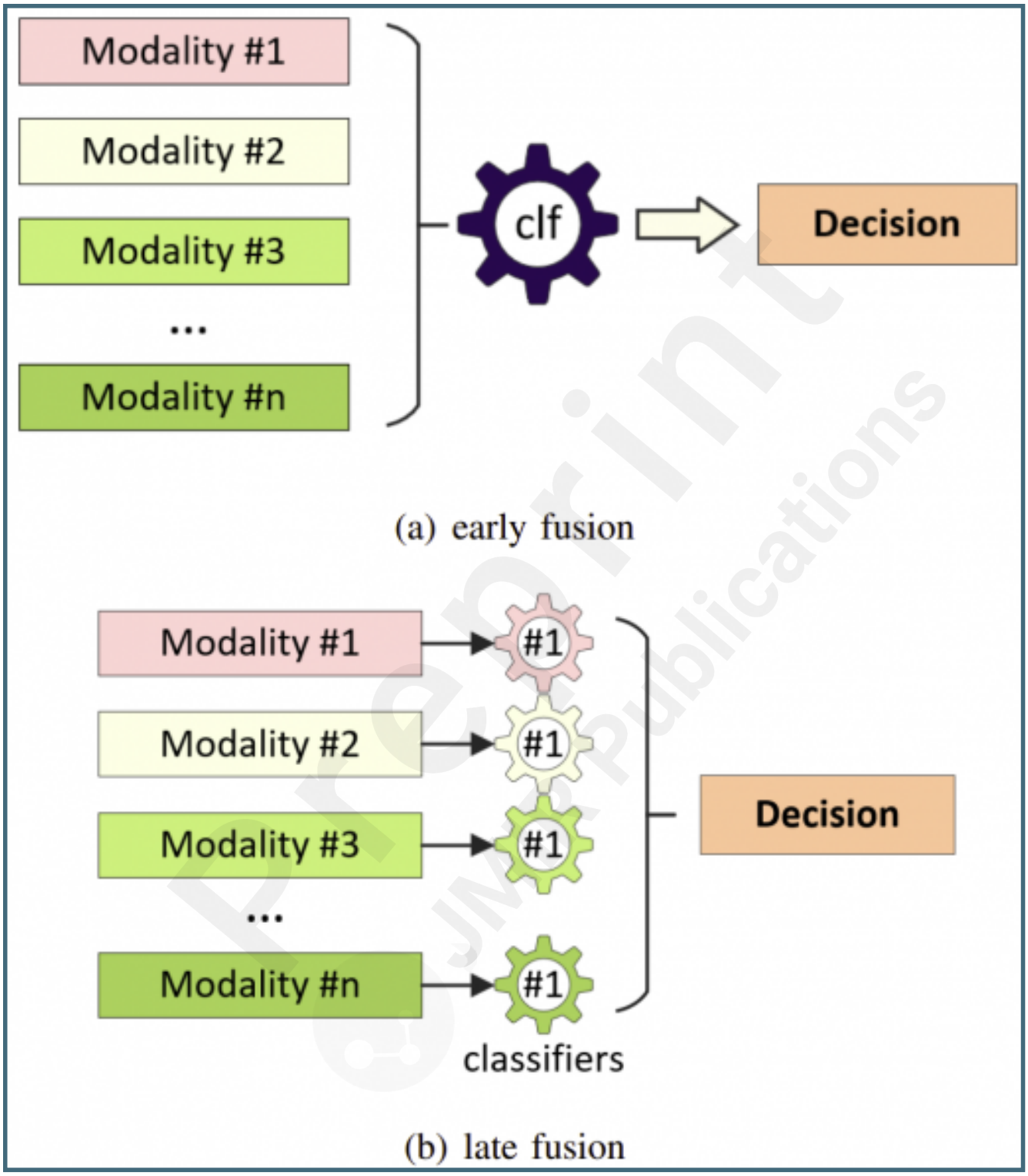
Devices used in physiological signal based pain assessment: (a) wristband - Empatica E4 [125], (b) headset - Emotiv EPOC [126], (c) chestband - Bioharness 3 [127].



Multimodal pain assessment.



Fusion strategies.



Multimedia Appendixes

Summary of studies.
URL: http://asset.jmir.pub/assets/df2756a9eb337573ad24f5a6f75b5846.pdf