

Review

A Review of Voice-Based Pain Detection in Adults Using Artificial Intelligence

Sahar Borna ¹, Clifton R. Haider ², Karla C. Maita ¹, Ricardo A. Torres ¹, Francisco R. Avila ¹, John P. Garcia ¹, Gioacchino D. De Sario Velasquez ¹, Christopher J. McLeod ³, Charles J. Bruce ³, Rickey E. Carter ⁴, and Antonio J. Forte ^{1,*}

¹ Division of Plastic Surgery, Mayo Clinic, Jacksonville, FL 32224, USA

² Department of Physiology and Biomedical Engineering, Mayo Clinic, Rochester, MN 55902, USA

³ Department of Cardiovascular Medicine, Mayo Clinic, Jacksonville, FL 32224, USA

⁴ Department of Health Sciences Research, Mayo Clinic, Jacksonville, FL 32224, USA

* Correspondence: ajvforte@yahoo.com.br; Tel.: +1-904-953-2073

Abstract: Pain is a complex and subjective experience, and traditional methods of pain assessment can be limited by factors such as self-report bias and observer variability. Voice is frequently used to evaluate pain, occasionally in conjunction with other behaviors such as facial gestures. Compared to facial emotions, there is less available evidence linking pain with voice. This literature review synthesizes the current state of research on the use of voice recognition and voice analysis for pain detection in adults, with a specific focus on the role of artificial intelligence (AI) and machine learning (ML) techniques. We describe the previous works on pain recognition using voice and highlight the different approaches to voice as a tool for pain detection, such as a human effect or biosignal. Overall, studies have shown that AI-based voice analysis can be an effective tool for pain detection in adult patients with various types of pain, including chronic and acute pain. We highlight the high accuracy of the ML-based approaches used in studies and their limitations in terms of generalizability due to factors such as the nature of the pain and patient population characteristics. However, there are still potential challenges, such as the need for large datasets and the risk of bias in training models, which warrant further research.

Keywords: machine learning; pain assessment; voice analysis; vocal biomarkers



Citation: Borna, S.; Haider, C.R.; Maita, K.C.; Torres, R.A.; Avila, F.R.; Garcia, J.P.; De Sario Velasquez, G.D.; McLeod, C.J.; Bruce, C.J.; Carter, R.E.; et al. A Review of Voice-Based Pain Detection in Adults Using Artificial Intelligence. *Bioengineering* **2023**, *10*, 500. <https://doi.org/10.3390/bioengineering10040500>

Academic Editors: Chien-Hung Liao and Victor Hugo C. De Albuquerque

Received: 20 March 2023

Revised: 13 April 2023

Accepted: 19 April 2023

Published: 21 April 2023



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

1. Introduction

1.1. Background

It has been established that pain represents a significant issue in the realm of public health within the United States. However, its subjective nature is a challenge for effective pain management [1–3]. The International Society for the Study of Pain (IASP) defines pain as “an unpleasant sensory and emotional experience associated with, or resembling that associated with, actual or potential tissue damage” [4–6]. The previous definition was revised in light of the many variables that affect how intense one’s painful episodes are, including past painful experiences, cultural and social contexts, individual pain tolerance, gender, age, and mental or emotional state [2,7].

Despite a thorough understanding of the pathophysiological processes underlying the physical pain response and the technological advances to date, pain is often poorly managed. Misdiagnosis of pain levels through associated subjective biases, as it is currently performed, can increase unnecessary costs and risks. In addition, poor pain relief can result in emotional distress and is linked to several consequences, including chronic pain [8–10].

Self-reports and observations form the basis of well-defined instruments for measuring pain and pain-related factors. Visual analog scales (VAS), the McGill Pain Questionnaire, and the numeric rating scales (NRS) are examples of patient-reported outcome measures

(PROs, self-reports, or PROMs), which are frequently considered “the gold standard” for measuring acute and chronic pain. As these strategies depend on patients’ accounts, they are exclusively applicable to individuals with no verbal or cognitive impairments [1,6,11,12]. Because of this, all of these established techniques are inapplicable to newborns, delirious, sedated, and ventilated patients, or individuals with dementia or developmental and intellectual disorders [13]. Such patients are entirely dependent on others’ awareness of nonverbal pain cues. In light of this, observational pain scales are advised for use among adults in critical conditions when self-reported pain measurements are not feasible. Even so, the reliability and validity of these instruments are still limited, because even qualified raters cannot ensure an unbiased judgment [5,14–19].

Automatic pain recognition has transitioned from being a theory to becoming a highly important area of study in the last decade [10]. As a result, a few studies have concentrated on utilizing artificial intelligence (AI) to identify or classify pain levels using audio inputs. While face and body occlusions are frequently observed in newborns, these methods are especially relevant for examining their cries as the more accurate way to identify neonates’ discomfort [20]. However, there remains a significant disparity in voice-based pain detection among adult patients.

1.2. Research Question and Objectives

This study aimed to provide an updated and concise view of the use of AI with voice or audio to detect pain in adult patients. Specifically, this study aimed to (1) summarize the current state of new research in this area, (2) identify the limitations and challenges of existing approaches, and (3) propose potential solutions and future research directions to address these limitations and challenges.

1.3. Methodology

The literature search was performed in the PubMed, Medline, Scopus, Web of Science, CINAHL, and Google Scholar databases. We included all the studies on the use of AI with voice for pain detection in adults up to 2023. Studies that were published in the English language and peer-reviewed journals were included.

1.3.1. Inclusion and Exclusion Criteria

We included studies that focused on the use of AI algorithms to analyze vocal features of pain in adult human subjects. The studies used a variety of AI techniques, including deep learning, natural language processing, and machine learning. They also made use of various vocal samples, such as speech, moans, and groans. Studies were excluded if they only focused on animals, infants, pediatrics, and non-verbal or unconscious patients.

1.3.2. Search Strategy

The search was conducted using a combination of keywords and controlled vocabulary. The search terms were related to AI, pain, voice, and detection. The search was restricted to human research and articles in English. Two reviewers independently screened the search results, while a third reviewer resolved any conflicts.

1.3.3. Data Extraction and Analysis

Two independent reviewers extracted the data using a standard data extraction form. The extracted information covered the study’s methodology, sample size, type of ML employed, categories of speech and pain qualities examined, speech parameterization method, model validation method, and AI algorithm accuracy.

We begin by providing a general review of the use of voice to identify pain (Section 2) and the role of artificial intelligence in pain detection (Section 3). In Section 4, we review the relevant papers, and in Section 5, we discuss the limitations and difficulties of AI models for pain detection using voice. In Section 6, the paper is concluded.

2. Voice, Pain, and Artificial Intelligence

Using Voice to Detect Pain

To extract speech-specific features from the recorded voice, analysis techniques such as perceptual linear prediction (PLP), PLP-relative spectra (PLP-RASTA), and mel-frequency cepstral coefficients (MFCC), are required.

These techniques can be used for various tasks such as speech recognition, speaker identification, and emotion recognition. An unknown sample can be predicted by training a neural network with feature vectors, while classification and recognition can be performed using techniques such as Support Vector Machine (SVM), the Hidden Markov Model (HMM), Dynamic Time Warping (DTW), and Vector Quantization (VQ) [21–25].

According to the studies, verbal or acoustic cues can be used to identify and even categorize pain according to its intensity level. The mean and range of voice fundamental frequency (F0), loudness, and degree of periodicity are some of the acoustic characteristics of voice that have been shown to indicate pain levels. Pain-related vocalizations, such as screaming, groaning, moaning, and crying, are often accompanied by changes in these phonetic features and can be suggestive of pain [26–30].

Several studies on the use of voice for pain detection were reviewed by Helmer et al. in their 2014 paper [15]. Their research focused on investigating whether vocalizations alone could indicate the presence of pain.

Of the 35 studies that were examined, 29 of them provided evidence of a connection between vocalization and pain.

Increased pitch and louder vocalization were shown to be more closely related to pain than other phonetic traits.

They showed that in children who can only cry to express themselves, crying might be more related to distress than showing pain. However, overall, children's crying was proven to be a pain indicator.

Most studies found a link between moaning and pain, whereas screaming or verbalizing was only sometimes associated with pain. Additionally, groaning and sighing were frequently not correlated with pain.

The review also showed that the individual's personality and social context could impact communicative pain behavior. Therefore, vocalization cannot be solely attributed to pain. Still, considering the significance of distress and social interaction, it is reasonable to infer a correlation between vocalization and pain.

3. Artificial Intelligence in Pain Detection

3.1. AI Techniques Used in Pain Detection

Although breakthroughs in computing have previously been observed in regard to medical applications, it can be challenging to distinguish between various types of pain [14]. Chronic pain is more difficult to diagnose and necessitates the recognition of more subtle pain coping mechanisms, but the accompanying facial, behavioral, and verbal pain indications make it quite simple to recognize acute pain [31].

Artificial intelligence (AI) describes a system's ability to mimic human behavior and exhibit intelligence, which is now viewed as a branch of engineering. The adoption of low-cost, high-performance graphics processing units (GPUs) and tensor processing units (TPUs), quicker cloud computing platforms with a high digital data storage capacity, and model training with cost-effective infrastructure for AI applications have all contributed to AI's unprecedented processing power [32–35].

There are two main categories of AI applications in the field of medicine: physical and virtual. The virtual subfield involves machine learning (ML), natural language processing (NLP), and deep learning (DL). On the other hand, the physical subfield of AI in medicine involves medical equipment and care bots (intelligent robots), which assist in delivering medical care [36–39].

Because of their cutting-edge performance in tasks such as image classification, DL algorithms have become increasingly prominent in the past ten years. The ability of a

machine to comprehend text and voice is described as NLP. NLP has many practical applications, including speech recognition and sentiment analysis.

Machine learning algorithms may be classified into three categories: unsupervised (capability to recognize patterns), supervised (classification and prediction algorithms based on previous data), and reinforcement learning (the use of reward and punishment patterns to build a practical plan for a specific problem space). ML has been applied to a variety of medical fields, including pain management [37,39,40].

To establish an automated pain evaluation system, it is essential to document the pertinent input data channels. Modality is the term used to describe the behavioral or physiological sources of information. The principal behavioral modalities are auditory, body language, tactile, and facial expressions [20,41].

The availability of a few databases with precise and representative data linked to pain has allowed for recent developments in the field of automatic pain assessment [42], with most works focusing on the modeling of facial expressions [7].

Numerous databases have been established by pooling data from diverse modalities acquired from distinct cohorts of healthy individuals and patients. Most studies use the following publicly available databases: UNBC-McMaster, BioVid Heat Pain, MIntPAIN, iCOPE, iCOPEvid, NPAD-I, APN-db, EmoPain, Emobase 2010, SenseEmotion, and X-ITE.

The BioVid Heat Pain Database is the second most frequently used dataset for pain detection after the UNBC-McMaster Shoulder Pain Archive Database. In the first case, video recordings of 90 healthy people under the influence of heat pain applied to the forearm are collected. The second is made up of 200 videos of participants' faces as they experience pain through physical manipulation of the shoulder [20,43–45].

Voice, meanwhile, has thus far received little consideration. Out of all the publicly available databases, BioVid Heat Pain, SenseEmotion, X-ITE, and Emobase use audio as one of their modalities. To fully leverage the potential of these databases, it is imperative to employ deep learning algorithms. Systems built on deep learning operate in two stages: training and inference. During training, the system is presented with a large dataset so as to teach it to recognize patterns and make predictions. Then, the trained model is used for inference, creating predictions based on the new data [20,43,44,46].

3.2. AI Models Used in Pain Detection from Voice

Multiple artificial neurons are merged to create an artificial neural network (ANN). These artificial neurons imitate the behavior and structure of biological neurons. Furthermore, to enhance the efficiency and precision of the ANN, the neurons are organized into layers for ease of manipulation and precise mathematical representation.

The operation of artificial neurons is governed by three fundamental principles: multiplication, summation, and activation. Initially, each input value is multiplied by a distinctive weight. Subsequently, a summation function aggregates all the weighted inputs. Finally, at the output layer, a transfer function transmits the sum of the previous inputs and bias [47,48].

The topology of an ANN refers to the way in which different artificial neurons are coupled with one another [47]. Different ANN topographies are appropriate for addressing different issues.

The most critical topologies are as follows:

- (1) Recurrent Artificial Neural Networks (RNNs).
- (2) Feed-Forward Artificial Neural Networks (FNNs).
- (3) Convolutional Neural Networks (CNNs): These use multiple layers to automatically learn features from the input data.
- (4) Long Short-Term Memory (LSTMs): It can handle vanishing and exploding gradients, which are common problems in the training of RNNs.
- (5) Multitask Neural Network (MT-NN): This employs the sharing of representations across associated tasks to yield a more advanced generalization model.

Bi-directional Artificial Neural Networks (Bi-ANN) and Self-Organizing Maps (SOM) can be named as other types [47–49].

An artificial neural network can solve a problem once its topology has been selected and it has been tuned and learned the right behavior [47,48].

Lately, there has been significant focus on artificial intelligence algorithms in sound event classification and voice recognition, as shown in Figure 1.

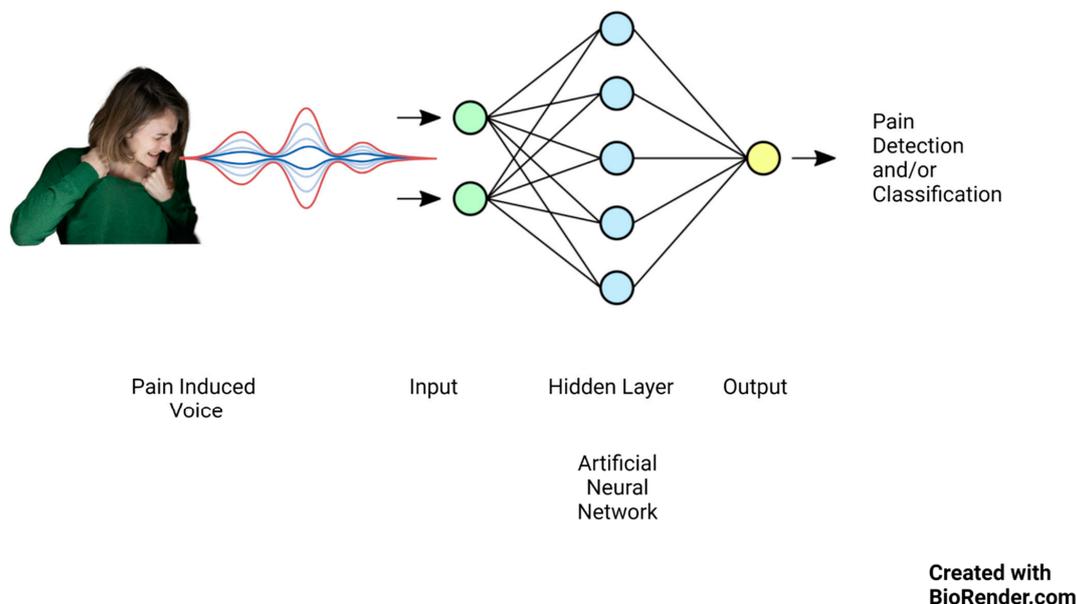


Figure 1. Artificial neural network mechanism of action on pain-induced vocalization.

Keyword spotting (KWS), wake-up word (WUW), and speech command recognition (SCR) are three essential techniques in speech processing that enable machines to recognize spoken words and respond accordingly.

Keyword identification technology is an automated approach to recognizing specific keywords within an uninterrupted spoken language and vocalization flow.

KWS systems are less reliant on high-quality audio inputs. They are created to be cheap and flexible and to run accurately and reliably on low-resource gadgets such as embedded edge devices [50–53].

Researchers have begun to create algorithms in order to automate pain level assessment using speech due to developments in signal processing and machine learning methods. As an example, Tsai et al. [14] employed bottleneck LSTM to detect pain based on a subset of the triage dataset, specifically prosodic signals. Later, Li et al. [7] introduced age and gender factors into a variational acoustic model.

4. Review of the Studies

4.1. AI and Pain Triage

Acute and chronic pain management differ significantly in the respective approaches. While acute pain management can be addressed using various medication options ranging from over-the-counter drugs to opioids, chronic pain management primarily relies on psychiatric interventions to control the underlying cause, which may sometimes even require surgical interventions. The authors of [54–56] conducted a study on chronic pain detection and followed patients with spinal/brain injuries for six months. In contrast, the authors of [50] focused on acute MI pain and achieved the highest overall accuracy using their CNN model. Other studies [14,57] followed patients for one hour to assess their pain levels and conduct an additional audio session, suggesting that the nature of their pain was acute (Table 1).

Table 1. Reviewed studies’ characteristics.

Author	Subjects	Modality Type	Speech Parameterization Method	AI/ML Type	Model Validation Method	Pain Characteristics	Metric Score
Oshrat et al., 2016 [54]	Patients with spinal cord and/or brain injuries 27 cases (20 male, 7 female) 400 sound files	Audio	OpenSMILE toolkit RASTA-PLP MFCC logMelFreqBand lspFreq	CFS SMO on SVM	Cross-validation and five folds	Not significant (pain levels ≤ 2) Significant (pain levels ≥ 2.5)	CCI ratio 73.75% OpenSMILE 77.25% OpenSMILE+ new features Kappa 39.81% First Group 46.97% Second Group
Tsai et al., 2016 [57]	On-boarding emergency patients 117 cases (205 sound files)	Audio–video Physiological (HR, SBP, DBP) vital sign data Other physiologically relevant results (analgesic prescription and patient disposition)	LLD	2-Class SVM 3-Class SVM Linear regression model (supervised)	Leave one patient out Cross-validation	Binary and tertiary pain severity classification	Accuracy 72.3% Binary Classification 51.6% Ternary Classification
Tsai et al., 2017 [14]	On-boarding emergency patients 63 cases (126 sound files)	Audio–video Physiological (HR, SBP, DBP) vital sign data Other physiologically relevant results	LLD	LSTMs with stacked bottlenecks	Leave one patient out Cross-validation	Binary and tertiary pain severity classification	Accuracy 72.3% Binary Classification 54.2% Tertiary Classification
Li et al., 2018 [7]	On-boarding emergency patients 141 cases 335 sound files (201 male, 134 female)	Audio–video	LLD (MFCC)	MMD-CVAE (unsupervised) linear-kernel SVM	Leave one speaker out Cross-validation	Binary and tertiary pain severity classification	UAR 70.7% Binary Classification 47.4% Tertiary Classification
Tsai et al., 2019 [2]	On-boarding emergency patients 184 cases (323 sound files)	Audio–video Physiological (HR, SBP, DBP) vital sign data other physiologically relevant results	OpenSMILE toolkit LLD	EEMN	Leave one patient out Cross-validation	Binary and tertiary pain severity classification	UAR 70.1% Multimodal Binary 52.1% Multimodal Tertiary
Mohan et al., 2022 [50]	Healthy adults 60 cases (360 sound files)	Audio	LLD (MFCC)	CNN	N/A	Pain Detection	TP 89% to 100% 97.91% Overall Accuracy

Abbreviations: OpenSMILE (Open Speech and Music Interpretation by Large Space Extraction), RASTA-PLP (RelAtive SpecTral PLP), MFCC (mel-frequency cepstral coefficients), logMelFreqBand (logarithmic power of mel-frequency band), LspFreq (Line Spectral Pair Frequencies), LLD (low-level descriptors), CFS (Correlation-Based Feature Selection), SMO (Sequential Minimal Optimization), SVM (Support Vector Machine), LSTM (Long Short-Term Memory), MMD-CVAE (maximum mean discrepancy variational autoencoder), EEMN (emotion-enriched multitask network), CNN (Convolutional Neural Network), CCI (Correctly Classified Instances), UAR (Unweighted Average Recall), TP (True Positive).

Accurate pain assessment is critical for effective pain management, and improvements in the precision of pain assessment can reduce the risk of incorrect management, especially in opioid prescription, which carries a risk of addiction and can pose a threat to patients' lives [8,58,59].

In 2016, Tsai et al. [57] conducted a study evaluating their framework against the NRS. They used the audio–video system in combination with the NRS and various measures (age, vital signs, pain levels, pain levels predicted using SVM models), as well as the two clinical outcome variables of painkiller prescription and disposition. The preprocessing steps undertaken in the study comprised manual segmentation of the recorded data followed by the extraction of LLDs. Subsequently, z-normalization was applied to each patient. Given that the duration of each session was 30 s, two encoding methods were employed to create a constant-length feature vector for each session: the computation of statistical functionals on the LLDs and adoption k-means bag of word (BoW) coding.

Linear regression models were trained using these measures to generate scores for each outcome. The study found that even when considering the best available medical instruments, such as physiological data and the NRS, the use of audio–video-based pain level assessment improves the accuracy in predicting whether a doctor will prescribe analgesics or hospitalize the patient.

The method of “embedding stacked bottleneck features”, a specific type of deep neural network architecture, provides a major advantage in audio processing. They are computationally more efficient and significantly enhance the performance of AI models when confronted with complex data containing numerous variables [14,60–62].

Tsai et al. conducted one of the first and most influential studies in the field of AI use to detect pain with voice in real adult patients in 2017 [14]. They employed a stacked bottleneck LSTM model, intending to classify the patients' pain levels.

Patients who initially presented with symptoms of chest, abdomen, limbic pain, and headaches had their NRS pain scores recorded for the follow-up period. During these sessions, nurses conversed with the patients regarding their pain location and characteristics.

Prosodic (pitch, intensity, etc.) and spectral (frequency content of the audio signal) elements were extracted from the vocal data to create two different subgroups. The authors created a compressed data representation with reduced dimensions and utilized stacked bottleneck acoustic attributes with the LSTMs to generate deep bottleneck features (DBFs).

The DaAi database was used to train an unsupervised LSTM sequence-to-sequence autoencoder. After that, the authors fine-tuned the stacked bottleneck layer by encoding the patients' sentences through different network layers. Lastly, they used the SVM to classify the pain levels. They showed that this novel approach could obtain a good accuracy in binary (severe, mild) and tertiary (severe, moderate, and mild) pain classification tasks.

The regular monitoring of individuals with heart conditions and provision of immediate medical attention can prevent emergencies from occurring.

For example, over 80% of individuals who experience a heart attack report a tight, squeezing, or crushing sensation in the chest, which can also radiate to the arms, shoulders, neck, jaw, or back [50,63].

The main objective of the study of Mohan et al. in 2021 was to design an audio-based intelligent system to recognize distress signals during MI in an interior environment with a low-power embedded GPU (graphics processing unit) edge device coupled with deep learning (Jetson Nano).

This device is small and portable. It includes a GPU for conducting accelerated computing operations at a network's edge rather than in a centralized data center or cloud environment.

The authors instructed 60 people to scream and use the Kannada language to communicate as if they were having a heart attack.

The study included three stages: In stage 1, audio data samples were gathered and prepared. A CNN model was trained during stage 2, and stage 3 involved the deployment

of the model trained in Jetson Nano. Finally, the researchers also tested the CNN model on a separate, private audio dataset to validate their approach.

Overall, the study results suggest that the proposed supervised audio classification scheme can effectively detect sound distress in an indoor smart space, even in challenging environments with background noise [50].

4.2. Pain as a Complex Affection

The experience of pain is intrinsically linked to an individual's emotional state, which can manifest in various forms of expressive behavior. Research has indicated that emotions are instrumental in shaping both the pain experience and the associated behavioral responses [64–67]. However, there has not been as much focus on understanding how a person's internal emotional state influences the pain they report.

In 2019, two studies were carried out by Tsai et al. to examine the link between self-reported pain levels and emotional states (Experiment 1) as well as the efficiency of an emotion-enriched multitask network (EEMN) in enhancing pain level detection by incorporating emotion-related data (Experiment 2). They also used 'activation' as the level of energy and intensity of the emotion and 'valence' as the negativity or positivity of the emotional expression.

Patients with headaches and abdomen, chest limbic, and lower back discomfort from the emergency department's triage were included. The information included vital signs, audio–visual recordings, and other relevant outcomes.

The findings indicated a substantial inverse relationship between valence mood and pain severity, while no such relationship existed between activation and the pain level [2].

For each recorded sample, the authors obtained acoustic characteristics utilizing the eGeMAPS feature set available in the openSMILE toolkit.

In the second trial, a multimodal fusion model's capacity for pain level recognition was enhanced using the EEMN. The model was trained with a multitasking network to focus on the primary objective of pain intensity and an auxiliary task of either valence or activation. By adding valence as an additional task to detect emotions, the study found that both speech and facial modalities had generally high recognition rates for pain levels.

4.3. Can the Speech Prosody Act as a Biosignal?

Various studies have established a correlation between pain and physiological indicators, such as systolic blood pressure and heart rate. Moreover, research has highlighted the influences of physiological factors on human vocalization [68–74].

In 2016, Oshrat et al. [54] conducted a study to determine whether the pain level can be measured based on voice biosignal parameters. The authors interviewed 27 adult participants with spinal cord or brain injuries and injury-related pain. Data were collected over six months through short voice-recorded interviews, and pain levels were evaluated using NRS. The voice samples were processed through manual cutting and feature extraction utilizing openSMILE software.

Machine learning analysis was performed using the WEKA suite, and the classifier SMO was used for classification. In addition, the authors developed two features: sequence indication and heat map thresholding.

Sequence indications refer to methods of analyzing sequences of values to identify patterns or trends in data, which can be used to measure changes in voice parameters as an indication of pain.

The authors created two different lists of sequence-length value indicators for audio samples. Then, they deployed these indicators in the PLP analysis of speech to investigate variations in the voice characteristics' pace of change. Next, they developed a feature extraction technique called heat map thresholding, which involved the display of coefficients extracted from a voice sample in a two-dimensional matrix and the production of heat map images. They compared pairs of such images of participants with low and high reported pain levels and observed differences in the color patterns. Finally, they ran

two classification experiments of “no significant pain” versus “significant pain”. In the first case, only the OpenSMILE features were retrieved; in the second, they added newly generated features.

The study indicated a correlation between the concurrent self-reported pain intensity and the quantifiable speech biosignal attributes, similar to the findings of the other two studies which demonstrated the effectiveness of prosodic low-level descriptors for spectral LLDs in identifying pain [14,57].

4.4. Age, Gender, and Pain

Recent research has indicated that ethnicity, race, gender, and age may all impact how someone experiences pain. For example, women tend to encounter more intense pain that occurs more frequently and lasts longer than men. While evidence suggests a decline in pain perception as individuals age, older adults may experience heightened sensitivity to pressure-induced pain [75–80].

Li et al. [7] proposed a method for automatic pain level recognition that involves the learning of latent acoustic representations embedded with attributes of gender and age. They achieved this by utilizing a conditional variational autoencoder optimized with the criterion of the maximum mean discrepancy (MMD-CVAE). For the preprocessing task, the authors utilized the acoustic LLD set called eGeMAPS. The features were subjected to z-normalization for every speaker, known as speaker normalization. Subsequently, context window expansion was implemented. The acoustic LLDs were also transformed into hidden features using a CVAE.

Typically, personal attribute dependency in pain recognition is addressed by training multiple independent models, such as gender-specific or age-specific models. However, the authors’ proposed MMD-CVAE approach directly embeds this information in the encoded acoustic space.

The authors demonstrated that, compared to females, male patients with severe pain express their pain with limited maximum jitter (differences in the pitch period of a speech waveform). However, female voices are more variable in pitch and have less harmonic content than male patients. In addition, elderly patients’ voices, in the context of pain, are less variable in pitch and amplitude than non-elderly voices.

The authors also showed that encoding the LLDs for latent representation using VAEs can improve the recognition UAR compared to other models that rely only on functionals.

They also demonstrated the effectiveness of their proposed approach using a large pain database of real patients and achieved significant improvements in the classification accuracy compared to existing methods. Their findings highlight the potential of using MMD-CVAE for personalized pain recognition and the importance of considering personal attributes in pain perception research.

5. Limitations and Challenges of AI Models in Pain Detection from Voice

The difficulty of automatic pain recognition arises from its intricate and subjective nature, which is susceptible to multifaceted influences, including but not limited to an individual’s personality, social environment, and past experiences. This matter remains the subject of ongoing scholarly discourse [81].

In comparison to healthy adults, people who are younger, older, in the intensive care unit, or have dementia have different pain thresholds and levels of pain tolerance. Previous studies have found that physiological indications can distinguish between “no pain” and “pain”. Still, researchers have had difficulty in differentiating between pain levels [43,54,69–71,75,77,78,80,82]. Moreover, nonverbal expressions of pain may not necessarily be consistent with verbal reports of pain [83].

Deep learning algorithms have a significant advantage in representing complex problems and subjective measurements, such as pain.

However, the majority of the research in this field has been on exploiting visual clues, including facial expressions, to identify pain.

Even though vocalizations are a well-recognized sign of pain and are frequently utilized in clinical contexts to measure pain intensity, the application of AI with voice or audio to identify pain has received much less attention. The difficulties in interpreting voice or audio data may be one factor affecting the limited number of studies in this area. For instance, the pitch, length, and other acoustic characteristics of pain vocalizations can vary greatly, making it challenging to create robust algorithms for pain detection based on these signals. Additionally, it can be challenging to separate pain vocalizations from other voices due to background noise and other confusing elements. For example, the authors of [50] showed that the limitation in automatic MI detection was the fact that their obtained model needs to distinguish between speech and various pain states in order to minimize false alarms.

In all of the studies regarding the detection of pain using AI, the accuracy of pain measurement could be improved by creating a combined multimodal framework that incorporates facial expressions, body language, biological signals (ECG, EMG, and GSR of the trapezius muscle), audio, and vocalization data. Furthermore, our detailed comprehension of the described medical emergency application could be improved by using a large number of validated datasets [81]. This is similar to the study in [50], in which the broadening of the database could lead to improved classification outcomes that could be applied to real-world situations.

Pain is influenced by different personal and social circumstances, which is another reason explaining why there is not more research in this field. This variability can make the development of accurate and generalizable algorithms for the detection of pain challenging.

Though the 13 findings of Tsai et al.'s study were encouraging, there were still constraints on the scope of the existing databases, specifically in regard to the collection of the actual triage sessions and the use of the DaAi background corpus to enhance the behavior patterns noted and represented in their framework [14].

In [50], the researchers had difficulties in creating and developing a dataset for automatic pain identification that fulfills benchmarking norms. This highlights the urgent need for a comprehensive reference dataset of various ethnicities and languages to reliably determine pain levels using audio signals. The inclusion of diverse voices with background noises and sounds is also essential to ensure that the model is reliable and can be generalized to various scenarios.

6. Conclusions

6.1. Summary of Findings

In 2016, Tsai et al. [57] found that combining the audio–video-based pain scale with other clinical measures results in more accurate analgesic prescription and patient disposition predictions than using the NRS scale alone. Additionally, in 2019, in emergency triage, where there is natural spoken contact between a medical practitioner and a patient, the authors of [14] concentrated on using machine learning algorithms to estimate the pain intensity. A similar process was carried out in [2], whose authors created an emotion-enriched multitask network (EEMN) by simultaneous optimization using effect states and multimodal behavior data to enhance self-reported NRS classification (acoustic and facial expressions). They showed that pain can be assumed to be a complex affect, and the use of AI to detect pain with this new definition can change traditional pain assessments. Additionally, the CNN method was employed for audio-based emergency detection during MI by the authors of [50]. They demonstrated that their lightweight CNN architecture was well-adapted to deliver an immediate response to heart vital attack signs in real-time processing. Oshrat et al. [54] discovered that in addition to the patient's report, speech in pain could be obtained as a biosignal for categorizing pain episodes. They successfully classified the pain levels into “no severe pain” and “severe pain” using auditory characteristics and heat map threshold image analysis for the first time.

The study conducted by Li et al. in 2018 proposed a novel approach for recognizing pain levels using vocal-based information. The approach involved embedding personal

attributes such as age and gender directly into the latent layer of a machine learning network called MMD-CVAE. This enabled the encoding of acoustic LLDs into the latent representation, improving pain level recognition accuracy [7].

Mohan et al. [50] achieved the highest accuracy of 97.91% using the CNN algorithm, while Tsai et al. [57] obtained the lowest accuracy for tertiary pain classification using the SVM and linear regression models (51.6%). Oshrat et al.'s CSF and SMO model showed a promising increase in the CCI ratio after the addition of new features to openSMILE [54]. The authors of [2,7] achieved a UAR of over 70% for binary pain classification using MMD-CVAE and EEMN, respectively. However, they showed weaker results in tertiary classification due to its more ambiguous nature, with an accuracy ranging from 47.4% to 52.1%.

6.2. Clinical Applications

The results of this review have essential and wide-ranging potential implications. According to the studies we evaluated, it is possible to efficiently and accurately identify various forms of pain by combining vocalizations and audio recordings with cutting-edge machine learning algorithms.

These discoveries may lead to new instruments and technologies that use voice analysis to identify and manage pain. In healthcare settings, where precise and prompt pain identification is essential for optimal patient treatment, such technologies could be very helpful. In addition, voice-based pain-monitoring devices could be integrated into surveillance systems for patients who cannot self-report pain, such as nonverbal or comatose patients.

Another implication is the employment of voice-based pain detection systems in research settings to better comprehend the complicated nature of pain.

Overall, the results of the research that has been reviewed have significant implications for the creation of new technologies and therapies that make better use of voice and machine learning to detect and control pain. While AI-based tools can reduce healthcare costs and save time, these effects may significantly influence pain treatment and the lives of patients with acute and chronic pain.

6.3. Potential Areas for Future Research

Upcoming research utilizing voice recognition and AI to identify pain offers various exciting directions for investigation.

First, to construct AI models, it is crucial to consider pain as an emotive experience. Since pain is a subjective experience, future research may benefit from a better comprehension of the subtleties of pain perception and the significance of accurately capturing the complexity of this experience.

Secondly, the precision and efficiency of prediction models could be greatly improved by combining several modalities and thorough data collection, including personal attributes.

Moreover, expanded objective metrics for the assessment of AI models are required. For instance, to improve the accuracy of pain identification by voice analysis, future studies might investigate the combination of additional metrics, such as voice, with physiological parameters such as the skin conductance level or heart rate. Additionally, the inclusion of routine practice scenarios could offer a more realistic environment for evaluating the efficacy of AI models in pain detection.

Finally, there is still a shortage of studies on the use of AI for pain recognition in adult populations, despite the rising corpus of literature on the topic in pediatric and neonatal populations. It is therefore essential to address this knowledge gap and broaden the research focus to include adult populations. Future research could then contribute to the creation of focused strategies to enhance pain management in clinical settings and offer a more thorough understanding of the effectiveness of AI-assisted pain detection in various patient populations.

6.4. Limitations

The main drawback of our review is the lack of information about the use of speech and audio modalities in adult patients in real-world settings. Since engineers and scientists have only recently become interested in this field of inquiry, there are still plenty of prospects for additional study and improvement. However, the small amount of research that has been performed on this subject makes it difficult to comprehend all its subtleties and complexities in full. In addition, although ethical considerations represent a crucial area of focus in AI research, this review did not cover them due to its broad scope.

6.5. Final Remarks

Despite the challenges, there is growing interest in the utilization of speech or audio with AI to detect pain, as it could have a wide range of clinical and research applications. For example, it can be used to track patients with chronic pain or to evaluate pain in patients who cannot express it verbally. However, further research is needed in order to develop and assess precise algorithms that can recognize pain in adults through voice or audio signals. Additionally, more investigation is necessary in order to understand the factors that affect the correlation between vocalizations and pain.

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

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found in the reference section.

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

References

1. Keskinarkaus, A.; Yang, R.; Fylakis, A.; Mostafa, S.E.; Hautala, A.; Hu, Y.; Peng, J.; Zhao, G.; Seppänen, T.; Karppinen, J. Pain fingerprinting using multimodal sensing: Pilot study. *Multimed. Tools Appl.* **2021**, *81*, 5717–5742. [[CrossRef](#)]
2. Tsai, F.-S.; Weng, Y.-M.; Ng, C.-J.; Lee, C.-C. Pain versus Affect? An Investigation in the Relationship between Observed Emotional States and Self-Reported Pain. In Proceedings of the 2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), Lanzhou, China, 18–21 November 2019.
3. Duca, L.M.; Helmick, C.G.; Barbour, K.E.; Nahin, R.L.; Von Korff, M.; Murphy, L.B.; Theis, K.; Guglielmo, D.; Dahlhamer, J.; Porter, L.; et al. A Review of Potential National Chronic Pain Surveillance Systems in the United States. *J. Pain* **2022**, *23*, 1492–1509. [[CrossRef](#)] [[PubMed](#)]
4. Shu, L.; Xie, J.; Yang, M.; Li, Z.; Li, Z.; Liao, D.; Xu, X.; Yang, X. A Review of Emotion Recognition Using Physiological Signals. *Sensors* **2018**, *18*, 2074. [[CrossRef](#)] [[PubMed](#)]
5. Nazari, R.; Sharif, S.P.; Allen, K.; Nia, H.S.; Yee, B.-L.; Yaghoobzadeh, A. Behavioral Pain Indicators in Patients with Traumatic Brain Injury Admitted to an Intensive Care Unit. *J. Caring Sci.* **2018**, *7*, 197. [[CrossRef](#)]
6. Berger, S.E.; Baria, A.T. Assessing Pain Research: A Narrative Review of Emerging Pain Methods, Their Technosocial Implications, and Opportunities for Multidisciplinary Approaches. *Front. Pain Res.* **2022**, *3*, 896276. [[CrossRef](#)]
7. Li, J.-L.; Weng, Y.-M.; Ng, C.-J.; Lee, C.-C. Learning Conditional Acoustic Latent Representation with Gender and Age Attributes for Automatic Pain Level Recognition. In Proceedings of the Interspeech 2018, Hyderabad, India, 2–6 September 2018. [[CrossRef](#)]
8. Han, J.S.; Bird, G.C.; Li, W.; Jones, J.; Neugebauer, V. Computerized analysis of audible and ultrasonic vocalizations of rats as a standardized measure of pain-related behavior. *J. Neurosci. Methods* **2005**, *141*, 261–269. [[CrossRef](#)]
9. Ren, Z.; Cummins, N.; Han, J.; Schnieder, S.; Krajewski, J.; Schuller, B. Evaluation of the pain level from speech: Introducing a novel pain database and benchmarks. In Proceedings of the 13th ITG-Symposium, Oldenburg, Germany, 10–12 October 2018; Speech Communication.
10. Werner, P.; Lopez-Martinez, D.; Walter, S.; Al-Hamadi, A.; Gruss, S.; Picard, R.W. Automatic Recognition Methods Supporting Pain Assessment: A Survey. *IEEE Trans. Affect. Comput.* **2019**, *13*, 530–552. [[CrossRef](#)]

11. Breau, L.M.; McGrath, P.; Camfield, C.; Rosmus, C.; Finley, G.A. Preliminary validation of an observational checklist for persons with cognitive impairments and inability to communicate verbally. *Dev. Med. Child Neurol.* **2000**, *42*, 609–616. [[CrossRef](#)]
12. Gruss, S.; Geiger, M.; Werner, P.; Wilhelm, O.; Traue, H.C.; Al-Hamadi, A.; Walter, S. Multi-modal signals for analyzing pain responses to thermal and electrical stimuli. *JoVE (J. Vis. Exp.)* **2019**, *146*, e59057.
13. Takai, Y.; Yamamoto-Mitani, N.; Ko, A.; Heilemann, M.V. Differences in Pain Measures by Mini-Mental State Examination Scores of Residents in Aged Care Facilities: Examining the Usability of the Abbey Pain Scale–Japanese Version. *Pain Manag. Nurs.* **2014**, *15*, 236–245. [[CrossRef](#)]
14. Tsai, F.-S.; Weng, Y.-M.; Ng, C.-J.; Lee, C.-C. Embedding stacked bottleneck vocal features in a LSTM architecture for automatic pain level classification during emergency triage. In Proceedings of the 2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII), San Antonio, TX, USA, 23–26 October 2017. [[CrossRef](#)]
15. Helmer, L.M.; Weijenberg, R.A.; de Vries, R.; Achterberg, W.P.; Lautenbacher, S.; Sampson, E.L.; Lobbezoo, F. Crying out in pain—A systematic review into the validity of vocalization as an indicator for pain. *Eur. J. Pain* **2020**, *24*, 1703–1715. [[CrossRef](#)] [[PubMed](#)]
16. Roulin, M.-J.; Ramelet, A.-S. Generating and Selecting Pain Indicators for Brain-Injured Critical Care Patients. *Pain Manag. Nurs.* **2015**, *16*, 221–232. [[CrossRef](#)]
17. Manfredi, P.L.; Breuer, B.; Meier, D.E.; Libow, L. Pain Assessment in Elderly Patients with Severe Dementia. *J. Pain Symptom Manag.* **2003**, *25*, 48–52. [[CrossRef](#)] [[PubMed](#)]
18. Hong, H.-T.; Li, J.-L.; Weng, Y.-M.; Ng, C.-J.; Lee, C.-C. Investigating the Variability of Voice Quality and Pain Levels as a Function of Multiple Clinical Parameters. *Age* **2019**, *762*, 6.
19. Lotan, M.; Icht, M. Diagnosing Pain in Individuals with Intellectual and Developmental Disabilities: Current State and Novel Technological Solutions. *Diagnostics* **2023**, *13*, 401. [[CrossRef](#)]
20. Gkikas, S.; Tsiknakis, M. Automatic assessment of pain based on deep learning methods: A systematic review. *Comput. Methods Programs Biomed.* **2023**, *231*, 107365. [[CrossRef](#)]
21. Dave, N. Feature extraction methods LPC, PLP and MFCC in speech recognition. *Int. J. Adv. Res. Eng. Technol.* **2013**, *1*, 1–4.
22. Htet, Y.; Zin, T.T.; Tin, P.; Tamura, H.; Kondo, K.; Chosa, E. HMM-Based Action Recognition System for Elderly Healthcare by Colorizing Depth Map. *Int. J. Environ. Res. Public Health* **2022**, *19*, 12055. [[CrossRef](#)]
23. Mahmut, E.-E.; Nicola, S.; Stoicu-Tivadar, V. Support-Vector Machine-Based Classifier of Cross-Correlated Phoneme Segments for Speech Sound Disorder Screening. *Stud. Health Technol. Inform.* **2022**, *294*, 455–459. [[CrossRef](#)]
24. Liang, Y.-C.; Wijaya, I.; Yang, M.-T.; Juarez, J.R.C.; Chang, H.-T. Deep Learning for Infant Cry Recognition. *Int. J. Environ. Res. Public Health* **2022**, *19*, 6311. [[CrossRef](#)]
25. Hitchcock, E.R.; Ochs, L.C.; Swartz, M.T.; Leece, M.C.; Preston, J.L.; McAllister, T. Tutorial: Using Visual–Acoustic Biofeedback for Speech Sound Training. *Am. J. Speech-Lang. Pathol.* **2023**, *32*, 18–36. [[CrossRef](#)] [[PubMed](#)]
26. Icht, M.; Rassis-Tal, H.W.; Lotan, M. Can the Vocal Expression of Intellectually Disabled Individuals Be Used as a Pain Indicator? Initial Findings Supporting a Possible Novice Assessment Method. *Front. Psychol.* **2021**, *12*, 655202. [[CrossRef](#)]
27. Lautenbacher, S.; Salinas-Ranneberg, M.; Niebuhr, O.; Kunz, M. Phonetic characteristics of vocalizations during pain. *Pain Rep.* **2017**, *2*, e597. [[CrossRef](#)] [[PubMed](#)]
28. Raine, J.; Pisanski, K.; Simner, J.; Reby, D. Vocal communication of simulated pain. *Bioacoustics* **2018**, *28*, 404–426. [[CrossRef](#)]
29. Carter, B.; McArthur, E.; Cunliffe, M. Dealing with uncertainty: Parental assessment of pain in their children with profound special needs. *J. Adv. Nurs.* **2002**, *38*, 449–457. [[CrossRef](#)] [[PubMed](#)]
30. Salinas-Ranneberg, M.; Niebuhr, O.; Kunz, M.; Lautenbacher, S.; Barbosa, P. Effects of pain on vowel production—Towards a new way of pain-level estimation based on acoustic speech-signal analyses. In Proceedings of the 43 Deutsche Jahrestagung für Akustik, Kiel, Germany, 6–9 March 2017.
31. Corbett, A.; Achterberg, W.P.; Husebo, B.S.; Lobbezoo, F.; de Vet, H.; Kunz, M.; Strand, L.I.; Constantinou, M.; Tudose, C.; Kappesser, J.; et al. An international road map to improve pain assessment in people with impaired cognition: The development of the Pain Assessment in Impaired Cognition (PAIC) meta-tool. *BMC Neurol.* **2014**, *14*, 229. [[CrossRef](#)]
32. Nagireddi, J.N.; Vyas, A.K.; Sanapati, M.R.; Sooin, A.; Manchikanti, L. The Analysis of Pain Research through the Lens of Artificial Intelligence and Machine Learning. *Pain Physician* **2022**, *25*, e211.
33. Mundial, F.E.; Schwab, K. The Fourth Industrial Revolution: What it means, how to respond. In Proceedings of the 2016 World Economic Forum, Dubai, United Arab Emirates, 11–12 November 2016.
34. House of Lords Select Committee A. *AI in the UK: Ready, Willing and Able*; House of Lords: London, UK, 2018; p. 36.
35. De Neufville, R.; Baum, S.D. Collective action on artificial intelligence: A primer and review. *Technol. Soc.* **2021**, *66*, 101649. [[CrossRef](#)]
36. Cornet, G. Chapter 4. Robot companions and ethics: A pragmatic approach of ethical design. *J. Int. Bioéth.* **2013**, *24*, 49–58. [[CrossRef](#)]
37. Hamet, P.; Tremblay, J. Artificial intelligence in medicine. *Metabolism* **2017**, *69*, S36–S40. [[CrossRef](#)]
38. Jung, W.; Lee, K.E.; Suh, B.J.; Seok, H.; Lee, D.W. Deep learning for osteoarthritis classification in temporomandibular joint. *Oral Dis.* **2023**, *29*, 1050–1059. [[CrossRef](#)] [[PubMed](#)]
39. Liu, Z.; He, M.; Jiang, Z.; Wu, Z.; Dai, H.; Zhang, L.; Luo, S.; Han, T.; Li, X.; Jiang, X.; et al. Survey on natural language processing in medical image analysis. *Zhong Nan Da Xue Xue Bao Yi Xue Ban* **2022**, *47*, 981–993. [[PubMed](#)]

40. Theofilatos, K.; Pavlopoulou, N.; Papasavvas, C.; Likothanassis, S.; Dimitrakopoulos, C.; Georgopoulos, E.; Moschopoulos, C.; Mavroudi, S. Predicting protein complexes from weighted protein–protein interaction graphs with a novel unsupervised methodology: Evolutionary enhanced Markov clustering. *Artif. Intell. Med.* **2015**, *63*, 181–189. [[CrossRef](#)]
41. Qin, K.; Chen, W.; Cui, J.; Zeng, X.; Li, Y.; Li, Y.; You, X. The influence of time structure on prediction motion in visual and auditory modalities. *Atten. Percept. Psychophys.* **2022**, *84*, 1994–2001. [[CrossRef](#)] [[PubMed](#)]
42. Thiam, P.; Kessler, V.; Amirian, M.; Bellmann, P.; Layher, G.; Zhang, Y.; Velana, M.; Gruss, S.; Walter, S.; Traue, H.C.; et al. Multi-Modal Pain Intensity Recognition Based on the SenseEmotion Database. *IEEE Trans. Affect. Comput.* **2019**, *12*, 743–760. [[CrossRef](#)]
43. Walter, S.; Gruss, S.; Ehleiter, H.; Tan, J.; Traue, H.C.; Crawcour, S.; Werner, P.; Al-Hamadi, A.; Andrade, A.O. The biovid heat pain database data for the advancement and systematic validation of an automated pain recognition system. In Proceedings of the 2013 IEEE International Conference on Cybernetics (CYBCO), Lausanne, Switzerland, 13–15 June 2013. [[CrossRef](#)]
44. Lucey, P.; Cohn, J.F.; Prkachin, K.M.; Solomon, P.E.; Matthews, I. Painful data: The UNBC-McMaster shoulder pain expression archive database. In Proceedings of the 2011 IEEE International Conference on Automatic Face & Gesture Recognition (FG), Santa Barbara, CA, USA, 21–23 March 2011.
45. Rodriguez, P.; Cucurull, G.; Gonzalez, J.; Gonfaus, J.M.; Nasrollahi, K.; Moeslund, T.B.; Roca, F.X. Deep Pain: Exploiting Long Short-Term Memory Networks for Facial Expression Classification. *IEEE Trans. Cybern.* **2022**, *52*, 3314–3324. [[CrossRef](#)]
46. Gouverneur, P.; Li, F.; Shirahama, K.; Luebke, L.; Adamczyk, W.M.; Szikszay, T.M.; Luedtke, K.; Grzegorzec, M. Explainable Artificial Intelligence (XAI) in Pain Research: Understanding the Role of Electrodermal Activity for Automated Pain Recognition. *Sensors* **2023**, *23*, 1959. [[CrossRef](#)]
47. Krenker, A.; Bešter, J.; Kos, A. Introduction to the artificial neural networks. In *Artificial Neural Networks: Methodological Advances and Biomedical Applications*; InTech: Rijeka, Croatia, 2011; pp. 1–18.
48. O’Shea, K.; Nash, R. An introduction to convolutional neural networks. *arXiv* **2015**, arXiv:151108458.
49. Tsai, S.-T.; Fields, E.; Xu, Y.; Kuo, E.-J.; Tiwary, P. Path sampling of recurrent neural networks by incorporating known physics. *Nat. Commun.* **2022**, *13*, 7231. [[CrossRef](#)]
50. Mohan, H.M.; Anitha, S. Real Time Audio-Based Distress Signal Detection as Vital Signs of Myocardial Infarction Using Convolutional Neural Networks. *J. Adv. Inf. Technol.* **2022**, *13*, 106–116. [[CrossRef](#)]
51. Zehetner, A.; Hagmüller, M.; Pernkopf, F. Wake-up-word spotting for mobile systems. In Proceedings of the 2014 22nd European Signal Processing Conference (EUSIPCO), Lisbon, Portugal, 1–5 September 2014.
52. Hou, J.; Xie, L.; Zhang, S. Two-stage streaming keyword detection and localization with multi-scale depthwise temporal convolution. *Neural Netw.* **2022**, *150*, 28–42. [[CrossRef](#)] [[PubMed](#)]
53. Qin, H.; Ma, X.; Ding, Y.; Li, X.; Zhang, Y.; Ma, Z.; Wang, J.; Luo, J.; Liu, X. BiFSMNv2: Pushing Binary Neural Networks for Keyword Spotting to Real-Network Performance. *IEEE Trans. Neural Netw. Learn. Syst.* **2023**, 1–13. [[CrossRef](#)] [[PubMed](#)]
54. Oshrat, Y.; Bloch, A.; Lerner, A.; Cohen, A.; Avigal, M.; Zeilig, G. Speech prosody as a biosignal for physical pain detection. In Proceedings of the 8th Conference on Speech Prosody, Boston, MA, USA, 31 May–3 June 2016.
55. Kinoshita, M.; Nascimento, I.J.B.D.; Styrnisdóttir, L.; Bruschetini, M. Systemic opioid regimens for postoperative pain in neonates. *Cochrane Database Syst. Rev.* **2023**, *4*, Cd015016. [[CrossRef](#)] [[PubMed](#)]
56. Carroll, C.P.; Brandow, A.M. Chronic Pain: Prevalence and Management. *Hematol. Oncol. Clin. N. Am.* **2022**, *36*, 1151–1165. [[CrossRef](#)]
57. Tsai, F.-S.; Hsu, Y.-L.; Chen, W.-C.; Weng, Y.-M.; Ng, C.-J.; Lee, C.-C. Toward Development and Evaluation of Pain Level-Rating Scale for Emergency Triage based on Vocal Characteristics and Facial Expressions. In Proceedings of the Interspeech 2016, San Francisco, CA, USA, 8–12 September 2016. [[CrossRef](#)]
58. Lavand’homme, P.M.; Kehlet, H.; Rawal, N.; Joshi, G.P. Pain Management after Total Knee Arthroplasty: Procedure Specific Postoperative Pain Management Recommendations. *Eur. J. Anaesthesiol.* **2022**, *39*, 743–757. [[CrossRef](#)]
59. Ganguly, A.; Michael, M.; Goschin, S.; Harris, K.; McFarland, D.C. Cancer Pain and Opioid Use Disorder. *Oncology* **2022**, *36*, 535–541.
60. Grezl, F.; Karafiat, M.; Vesely, K. Adaptation of multilingual stacked bottle-neck neural network structure for new language. In Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, 4–9 May 2014.
61. Wu, Z.; Valentini-Botinhao, C.; Watts, O.; King, S. Deep neural networks employing multi-task learning and stacked bottleneck features for speech synthesis. In Proceedings of the 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), South Brisbane, Australia, 19–24 April 2015. [[CrossRef](#)]
62. Zhang, Y.; Chuangsuwanich, E.; Glass, J.R. Language ID-based training of multilingual stacked bottleneck features. In Proceedings of the Interspeech 2014, Singapore, 14–18 September 2014. [[CrossRef](#)]
63. Smith, K.L.; Cameron, P.; Meyer, A.; McNeil, J.J. Knowledge of heart attack symptoms in a community survey of Victoria. *Emerg. Med. Australas.* **2002**, *14*, 255–260. [[CrossRef](#)]
64. Carter, L.; McNeil, D.W.; Vowles, K.; Sorrell, J.T.; Turk, C.L.; Ries, B.J.; Hopko, D.R. Effects of Emotion on Pain Reports, Tolerance and Physiology. *Pain Res. Manag.* **2002**, *7*, 21–30. [[CrossRef](#)]
65. Keefe, F.J.; Lumley, M.; Anderson, T.; Lynch, T.; Carson, K.L. Pain and emotion: New research directions. *J. Clin. Psychol.* **2001**, *57*, 587–607. [[CrossRef](#)]

66. Melzack, R. *Wall PD the Challenge of Pain*; Basic Books: New York, NY, USA, 1982.
67. Robinson, M.E.; Riley, J.L., III. The role of emotion in pain. In *Psychosocial Factors in Pain: Critical Perspectives*; The Guilford Press: New York, NY, USA, 1999.
68. Loggia, M.L.; Juneau, M.; Bushnell, C.M. Autonomic responses to heat pain: Heart rate, skin conductance, and their relation to verbal ratings and stimulus intensity. *Pain* **2011**, *152*, 592–598. [[CrossRef](#)]
69. Möltner, A.; Hölzl, R.; Strian, F. Heart rate changes as an autonomic component of the pain response. *Pain* **1990**, *43*, 81–89. [[CrossRef](#)] [[PubMed](#)]
70. Robinson, B.F. Relation of Heart Rate and Systolic Blood Pressure to the Onset of Pain in Angina Pectoris. *Circulation* **1967**, *35*, 1073–1083. [[CrossRef](#)] [[PubMed](#)]
71. Storm, H. Changes in skin conductance as a tool to monitor nociceptive stimulation and pain. *Curr. Opin. Anaesthesiol.* **2008**, *21*, 796–804. [[CrossRef](#)] [[PubMed](#)]
72. Ledowski, T.; Bromilow, J.; Wu, J.; Paech, M.J.; Storm, H.; Schug, S.A. The assessment of postoperative pain by monitoring skin conductance: Results of a prospective study. *Anaesthesia* **2007**, *62*, 989–993. [[CrossRef](#)]
73. Rivasi, G.; Menale, S.; Turrin, G.; Coscarelli, A.; Giordano, A.; Ungar, A. The Effects of Pain and Analgesic Medications on Blood Pressure. *Curr. Hypertens. Rep.* **2022**, *24*, 385–394. [[CrossRef](#)]
74. Saloni, R.; Sharma, K.; Gupta, A.K. Classification of high blood pressure persons vs. normal blood pressure persons using voice analysis. *IJ Image Graph. Signal Process.* **2014**, *1*, 47–52.
75. Wandner, L.D.; Scipio, C.D.; Hirsh, A.T.; Torres, C.A.; Robinson, M.E. The Perception of Pain in Others: How Gender, Race, and Age Influence Pain Expectations. *J. Pain* **2012**, *13*, 220–227. [[CrossRef](#)]
76. Unruh, A.M. Gender variations in clinical pain experience. *Pain* **1996**, *65*, 123–167. [[CrossRef](#)]
77. Horgas, A.L.; Elliott, A.F. Pain assessment and management in persons with dementia. *Nurs. Clin. N. Am.* **2004**, *39*, 593–606. [[CrossRef](#)]
78. Lautenbacher, S.; Kunz, M.; Strate, P.; Nielsen, J.; Arendt-Nielsen, L. Age effects on pain thresholds, temporal summation and spatial summation of heat and pressure pain. *Pain* **2005**, *115*, 410–418. [[CrossRef](#)]
79. Rittger, H.; Rieber, J.; Breithardt, O.; Dücker, M.; Schmidt, M.; Abbara, S.; Sinha, A.; Jakob, A.; Nölker, G.; Brachmann, J. Influence of age on pain perception in acute myocardial ischemia: A possible cause for delayed treatment in elderly patients. *Int. J. Cardiol.* **2011**, *149*, 63–67. [[CrossRef](#)] [[PubMed](#)]
80. Robinson, M.E.; Dannecker, E.A.; George, S.; Otis, J.; Atchison, J.W.; Fillingim, R.B. Sex Differences in the Associations among Psychological Factors and Pain Report: A Novel Psychophysical Study of Patients with Chronic Low Back Pain. *J. Pain* **2005**, *6*, 463–470. [[CrossRef](#)] [[PubMed](#)]
81. Werner, P.; Al-Hamadi, A.; Niese, R.; Walter, S.; Gruss, S.; Traue, H.C. Automatic pain recognition from video and biomedical signals. In Proceedings of the 2014 22nd International Conference on Pattern Recognition, Stockholm, Sweden, 24–28 August 2014.
82. Van Iersel, T.; Timmerman, D.; Mullie, A. Introduction of a pain scale for palliative care patients with cognitive impairment. *Int. J. Palliat. Nurs.* **2006**, *12*, 54–59. [[CrossRef](#)] [[PubMed](#)]
83. Waters, S.J.; Riordan, P.A.; Keefe, F.J.; Lefebvre, J.C. Pain Behavior in Rheumatoid Arthritis Patients: Identification of Pain Behavior Subgroups. *J. Pain Symptom Manag.* **2008**, *36*, 69–78. [[CrossRef](#)] [[PubMed](#)]

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