Deep Learning-Based Models for Pain Recognition: A Systematic Review

Rasha M. Al-Eidan,[1] Hend Al-Khalifa,[1] and AbdulMalik Al-Salman[1]

[1] College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia

Correspondence should be address to Rasha M. Al-Eidan; raleidan@ksu.edu.sa

ABSTRACT

The traditional standards employed for pain assessment have many limitations. One such limitation is reliability because of inter-observer variability. Therefore, there have been many approaches to automate the task of pain recognition. Recently, deep-learning methods have appeared to solve many challenges, such as feature selection and cases with a small number of data sets. This study provides a systematic review of pain-recognition systems that are based on deep-learning models for the last two years only. Furthermore, it presents the major deep-learning methods that were used in review papers. Finally, it provides a discussion of the challenges and open issues.

INDEX TERMS   pain assessment, pain recognition, deep learning, neural network, review, dataset

I.   INTRODUCTION

Deep learning is an important branch of machine learning that is used to solve many applications, and it is a powerful way of achieving supervised learning. The history of deep learning has undergone different naming conventions and developments [1]. It was first referred to as *cybernetics* in the 1940s–1960s. Then, from the 1980s–1990s, it became known as *connectionism.* Finally, the phrase *deep learning* began to be utilized in 2006. There are also some names that are based on human knowledge. For example, an artificial neural network (ANN) refers to the biological aspect of deep learning. However, some problems cannot be solved by neural inspiration, and many deep-learning perspectives therefore depend on recently developed statistics and applied mathematical principles.

Automated pain recognition is a complex problem that needs to be solved by a powerful method based on deep learning. Recently, some studies of pain recognition have employed this approach to machine learning. Therefore, this study aims to review deep-

learning applications for pain recognition only. It provides a systematic review that incorporates a search strategy and inclusion criteria, and we describe it in the next section. To best of our knowledge, there are only one review paper on pain recognition models in general [2]. This study focuses on the deep models for pain recognition.

The main contributions of this study are summarized as follows:

- Review of the pain-recognition studies that are based on deep learning.
- Presentation and discussion of the main deep-learning methods employed in the reviewed papers.
- Review of the available data sets for pain recognition.
- Discussion of some challenges and future works.

II. Methodology

A. *Search Strategy*

First, we identified important search terms for pain-recognition systems. These included 'pain recognition', 'pain assessment', and 'pain diagnosis'. Secondly, the period spanned publications in 2017 to 2019 only. These two constraints with respect to terminology and time were used to perform searching in the popular databases IEEE, ACM, and Web of Science (WOS).

B. *Inclusion and Exclusion Criteria*

These criteria include related fields of research, and utilized the deep-learning method only. After the first search iteration, many papers that required filtering were listed.

For example, the WOS search engine has first 634 studies. We performed filtering by selecting the ones that are related to computer science fields only. As a result, the second iteration had 58 papers. Moreover, we selected only the studies that are based on deep-learning models. The final number of suitable studies from the WOS search engine was 13.

C. *Classification Method*

The resulting papers were scanned rapidly in order to obtain a suitable classification method. Therefore, we first classified the paper into a single model or multi-model. After that, the single model has four categories, which are described in the figure 1 below.

- Pain recognition and deep-learning models
  - Single model
    - Physiological signals
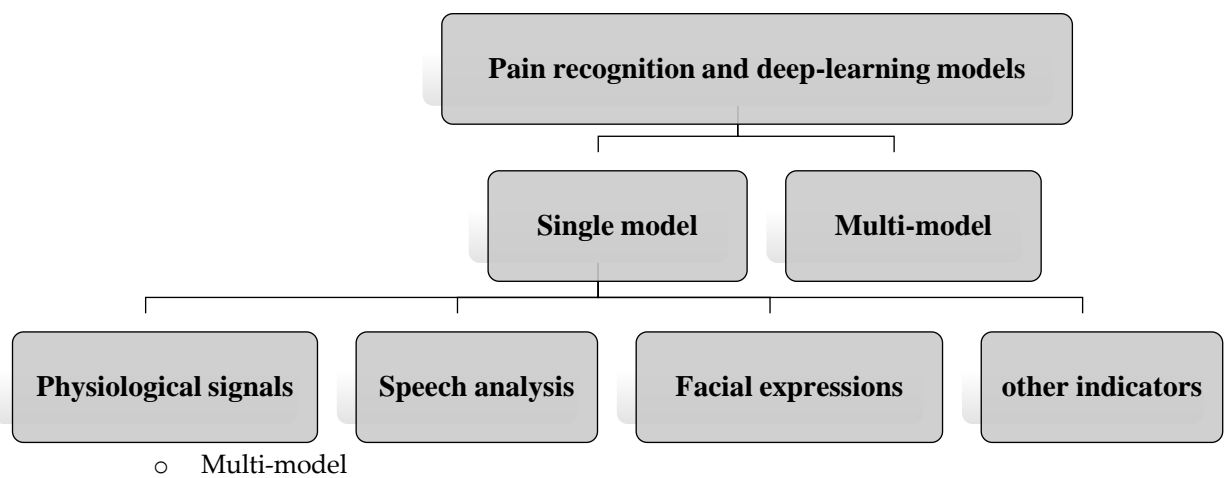   - Speech analysis
    - Facial expressions

  - Multi-model

Figure 1. Classification schema

3

III.   Review Papers

A.   *Single-Model-Based Pain Recognition*

Single-model-based pain-recognition systems can be defined as systems that use a single kind of measure to classify the pain level. These measures are physiological, speech, body movements, and facial expressions. Based on our literature review, we grouped the measures into three categories that are based on deep-learning methods. Next, we present details of the previous studies for each single model.

1)   PHYSIOLOGICAL SIGNALS

Physiological signals are among the most important measures used to describe the pain level from the physiological response of the body. These signals include vital signs (such as blood pressure, respiration rate, heart rate, and muscle activity) or brain dynamics.

In addition to the lack of using physiological signals for pain detection, Lopez-Martinez and Picard in 2017 [3] proposed a model using neural network (NN) techniques. Their proposed approach implements a multi-task learning method to tackle the problem of personal differences through shared layers in NNs. Their system involves multi-model data that depend on physiological signals from skin conductance (SC) and electrocardiograms (ECGs) only. They used the available data set, called a BioVid dataset, to build a model and conduct their experiments. Then, they compared their model with other machine-learning methods (logistic regression (LR), and support-vector machine with both linear kernel (SVM-L) and radial basis function kernel (SVM-RBF)). They reported that this NN approach outperforms others by around 82.75%. However, they reported no experimental testing of the model on new data sets. In the future, this method can be easily adapted in real clinical settings because of the simple use of two features that can be acquired from wrist-wearable devices.

2)   SPEECH ANALYSIS

Another study [4] recognized the pain level based on speech analysis and the use of long short-term memory (LSTM) NNs. First, they employed an unsupervised learning NN to extract the vocal features using the Chinese corpus. Then, they fine-tuned it based on NNs and an emergency triage database to output the specific acoustic representation at the sentence-level for each patient. Finally, they performed a pain-level classification to two or three classes using the SVM. As a result, their methods achieved weighted average recall (WAR) values of 72.3% in binary-class and 54.2% in three-class pain-intensity classification tasks. To the best of our knowledge, this work is the first study that used the speech only to detect the pain level.

3)   FACIAL EXPRESSIONS

Facial expressions are signals that have received attention from researchers for many applications, such as face recognition in the field of biometrics.

Deep learning is used directly to estimate the pain from face expressions. One of the distinct approaches is to estimate the pain from the self-reported visual analogue scale (VAS) pain levels to understand the individual differences [5]. Their method includes two learning stages. The first one, which is performed by learning the recurrent NNs (RNNs), estimates the Prkachin and Solomon pain intensity (PSPI) levels from face images. Then, personalized hidden conditional random fields (HCRFs) used the previous output to estimate the VAS for each person. By making comparisons with non-personalized approaches, this approach achieved high performance, and the score for a single-sequence test is the best.

4

Deep learning has been mainly used to extract the important features, as was recently done by [6] for pain detection based on facial expressions. Their approach is based on three steps: First, convolutional NNs (CNNs) are used to extract the features from VGG_Faces. After that, the result of the feature map is used to train the LSTM. This is a type of RNN used to find the binary pain estimation (pain, no pain). They provided a summary of previous works done on the popular data set of pain detection using faces. This data set is called a UNBC-McMaster database, and has 200 video sequences from 25 patients who suffered from shoulder pain. Their experiments on this data set showed that their approach outperforms all previous works with an area under the curve (AUC) performance of 93.3%. Their model can also be generalized for application to other facial emotion recognition. This ability was realized when applying their model to the Cohn Kanade+ facial expression database, and resulted in a competitive score (AUC = 97.2%).

In the same manner, in 2017, Egede, Valstar, and Martinez [7] proposed a pain-estimation model that combined learned features obtained from deep learning and other handcrafted features. Their idea comes from the hardness to obtain a data set for pain estimation in a large area in order to work well with deep learning. Therefore, they extracted handcrafted features directly from the face image, and used a CNN to learn some features. Their features include appearance, shape, and dynamics information. Finally, they classified the pain level using the linear regression model on the combined features and individual. Their results outperformed the state-of-the-art methods in terms of the root mean-square error (RMSE) of 0.99 and Pearson correlation (CORR) of 0.67. A limitation of this approach and all face-based approaches is that they consider only the front of the face without capturing several combinations of indicators, such as audio, body movements, and physiological signals.

Another solution that aims to deal with small data sets and deep learning was proposed in 2017 by Wang et al. [8]. They fine-tuned a small pain data set using a face verification network that is trained by the WebFace dataset, which has 500,000 face images. Then, they fitted a problem as a regression problem by applying a regression loss regularized with the center loss. Their performance was evaluated based on new proposed metrics to avoid the use of imbalance data. Based on the results, this method achieved a high performance compared with the state-of-art methods using both weighted metrics (mean absolute error (MAE): 0.389, (mean squared error) MSE: 0.804, (Pearson correlation coefficient) PCC: 0.651) and new proposed metrics (weighted MAE 0.991, weighted MSE 1.720). However, pain is temporal and involves subjective information, and no such information and stimulus knowledge are used in this method, which requires further investigation.

In contrast to previous solutions, the cumulative attribute (CA) method was used in 2018 [9] as a good solution to overcome the imbalance of data in pain estimation datasets. The cumulative attribute is defined as 'an intermediate representation C_i obtained by transforming the original labels y_i into a vector'. In this study, the deep CNN is used with the cumulative attributes in two steps. In the first step, the cumulative attribute vector is outputted from a trained CNN. In the second step, the regression model is trained to produce the final real output. Their approach performed tests on a pain estimation dataset and estimated the age. Their pain estimation results obtained higher values for CA-CNN experiments compared with non-CA-CNN experiments. In addition to the use of a CA layer trained with a Log-loss function, it significantly outperforms a CA layer trained with the Euclidean loss. Their approach has the advantage of using the CNN framework without any additional annotations. However, the requirement to build the annotated dataset for pain estimation is important to overcome most of the problems in classification tasks.

Therefore, Haque et al. [10] built a new database with RGB, depth, and thermal (RGBDT) images of the face for pain-level recognition in sequences. Their approach of elicitation is different from previous datasets obtained by stimulated healthy people with electrical pulses. Twenty subjects participated in the data collection to determine the pain recognition based on five levels of pain (0 for no pain and four for severe pain). After collecting the data sets, they constructed a baseline model for pain recognition based on spatio–temporal features and deep learning. First, they preprocessed the video frames by cropping only the face region based on their previous proposed method [11] for RGB images. Then, they used homography matrix codes to crop other depth and thermal images. After that, they applied deep learning based on two approaches for individual modalities or the fusion between them. The main idea of the proposed method is based on two steps. First, they used 2D-CNN for frame-features extraction and pain recognition. Secondly, they used LSTM to find the temporal relation between frames and sequence level pain

recognition. From the results, the fusion approach exhibited the greatest performance compared with the individual modalities. In addition, early fusion, which is achieved by integrating the input from all models to be fed to classifiers is better than late fusion, which integrates the output of each model to input the second classifier.

4) OTHER INDICATORS

In 2017, there was an attempt [12] to use deep learning method for LBP recognition. Their methods were based on segmentation and classification of LBP X-Ray data. They got a Tomography lumbar spine pictures from Meta Picture (MHD) arrange and gave them five vertebral levels. Then, deep learning framework extracted the features and classified the LBP based on five severity levels (normal, mild, crush, wedge, severe and biconcavity). They reached around 65% of accuracy that needs further improvements. However, using X-ray data of patients is not enough to recognize LBP.

Recently, [13] Hu et al. proposed a deep learning method to recognize the low back pain in the context of static standing. Their system depends on kinematic data that was acquired using three attached motion sensors on different places of human skin. After preprocessing, they used the data as an input to LSTM network. Using 22 healthy people and 22 LBP patients, they got 1073 time series for training and 107 time series for testing. Their results showed a high accuracy of 97.2% as they mentioned. The disadvantage of this study is the using of only kinematic data and ignores the EMG data, which is important to differentiate LBP patients.

B. *Multi-Model-Based Pain Recognition*

A recent attempt [14] to recognize pain based on a multi-model approach involves combining face and physiological signals (ECC and SC). The personalization of pain estimation is their main goal, and is achieved based on clustering subjects into different profiles, and not for each individual, as in previous works. Then, they used the multitask NN (MT-NN) approach, where each task corresponds to one profile. For their experiment, they used the available data set BioVid Heat Pain database. The results exhibited a better performance for high clusters (C = 4), which shows the need for further investigation in future using a larger number of clusters.

In addition, [15] performed fusion between physiological signals (EMG, ECG, and SC) and face videos. Their approach involves implementing the idea of adaptively using the system to test unknown individuals based on unlabelled data. Therefore, they used multi-stage ensemble-based classification, and were based on the BioVid Heat Pain database. NNs were used at the confidence-estimation stage, which was trained using three different inputs, namely the prediction of one-vs-one classifiers, the continuous pain-level estimation of the regressor, and the variance of a bagged ensemble of random forests. Then, the use of NNs will enable us to determine the confidence level of samples by employing a random regressor as an input. After the evaluation of different combinations of inputs to the NN, the highest values of the correlation coefficient and RMSE were 0.183 and 0.347, respectively. They found that the adaptation process is not an easy task because of individual differences in response to pain stimuli, and they therefore require more investigation.

More recently, Thiam et al. [16] explored several CNNs architectures based on the available BioVid Heat Pain Database (Part A). They used the three modalities of signals: EDA, ECG and EMG. They tried two kinds of input to network which are 1D and 2D. In addition, different deep fusion architectures were presented and tested. Their results reached 84.57% and 84.40% for binary classification. Next, we present in Table 1 a summery of all pain-recognition studies that based on deep learning.

(Abbreviations: Weighted average recall (WAR), Area under the curve (AUC), Pearson correlation (CORR), Mean-square error (MSE), Root mean-square error (RMSE), Intra-class correlation (ICC), Correlation coefficient (CC), Mean absolute error (MAE), Pearson correlation coefficient (PC)

Study	Deep-learning approaches	Task	Features - devices	Dataset	Metric-score
[3] 2017	Multi-task neural network (MT-NN)	Classification	Skin conductance (SC) and heart-rate features (ECG) only	Available: BioVid Heat Pain Database	Accuracy 82.75%
[4] 2017	Long-short term memory neural networks (LSTMs)	Feature extraction	Vocal from audio Face from video Device: Sony HDR handy cam	Collected: Triage Pain-Level Multimodal database Available: Speech data: Chinese corpus: The DaAi database Three-class (severe, moderate, and mild)	WAR 72.3%: binary classes 54.2%: three-class classes
[5] 2017	LSTMs	Classification	Face	Available: UNBC-MacMaster Shoulder Pain Expression Archive Database	MAE 2.47 (0.18) ICC 0.36 (0.08) Confusion matrices
[6] 2018	- Convolutional neural networks (CNNs) - LSTMs	-Feature extraction -Classification	Face	Available: UNBC-MacMaster Shoulder Pain Expression Archive Database Cohn Kanade + facial expression database	AUC: 93.3%
[7] 2017	CNN	Feature extraction	Face	Available: UNBC-MacMaster Shoulder Pain Expression Archive Database	CORR 0.67 RMSE 0.99
[8] 2017	CNN- Fine-tuning- regularizing	Classification	Face	Available: UNBC-MacMaster Shoulder Pain Expression Archive Database The face verification network [12] is trained on CASIA- WebFace dataset [16], which contains 494,414 training images from 10,575 identities	Unweighted Metrics MAE 0.389 MSE 0.804 PCC 0.651 Weighted Metrics: Weighted MAE 0.991 Weighted MSE 1.720
[9] 2018	Cumulative attributes (CA)-CNN	Classification	Face	Available: UNBC-MacMaster Shoulder Pain Expression Archive Database	Regression: PCC (0.47, 0.53) RMSE (1.20, 1.23) Multi-class: PCC (0.36, 0.41) RMSE (1.17, 1.19)
[10] 2018	CNN and LSTM	Feature extraction And Classification	Face - Microsoft Kinect Version2. - Axis Q1922 thermal camera	Collected: Multimodal Intensity Pain (MIntPAIN)' database Healthy subjects Classes: 5	5-fold cross-validation Accuracy The confusion matrix
[14] 2017	MT-NN	Classification	ECG, SC Face	Available: BioVid Heat Pain Database	MAE, RMSE , ICC
[15] 2017	NN	Confidence estimation	ECG, SC, EMG Face	Available: BioVid Heat Pain Database	Cross validation RMSE: 0.347 CC: 0.183
[12] 2017	LSTM (Tensor Flow)	Classification	Tomography lumbar spine pictures –from Meta Picture (MHD) arrange	Classes: 6	65%
[13] 2018	LSTM	Classification	- Kinematic data - Motion sensors	22 healthy people and 22 LBP patients	Accuracy : 97.2%
[16] 2019	CNNs	Feature extraction And Classification	EDA, ECG, EMG	Available: BioVid Heat Pain Database (part1)	Accuracy : 84.40%

IV. Primary Deep-Learning Methods Employed for Pain Recognition

In the previous section, we presented the literature review of deep-learning methods employed for pain recognition over the last two years. In this section, we present and discuss the main deep-learning models used in pain-recognition studies in addition to the metrics used to evaluate such models.

A. Convolutional Neural Networks (CNNs)

CNNs were derived from the first model on an NN that was invented in 1998 by Yann LeCun et al. [17]. In general, conventional networks perform logistic regression by applying a filter to the input. In addition to the filter size (f), the important parameters required to build a deep CNN are the stride (s) and padding (p). CNNs have many types of layers, namely conventional, pooling, and fully connected layers [18]. We can illustrate the CNN architecture as in Figure 2
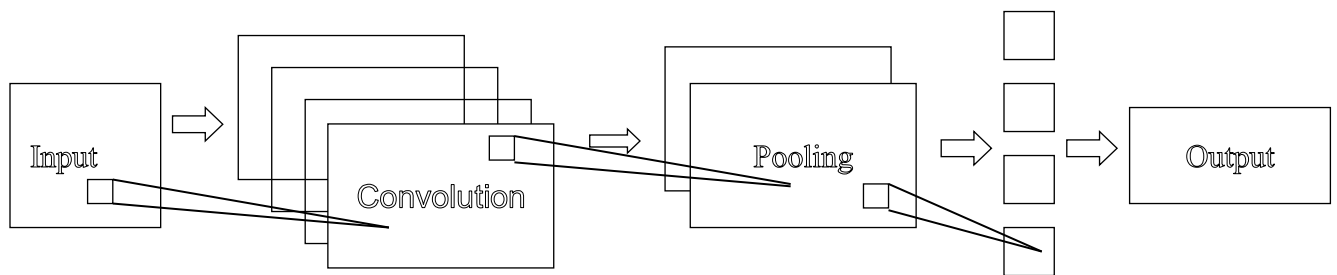

Figure 2. CNN architecture [18]

The CNN has an advantage of generalization compared with MLP [19]. In addition, it has fewer parameters than the fully connected layers in MLP.

As we can see from the literature, CNNs are mostly used for feature extraction. Only two studies have employed them for classification tasks. However, they were used only for pain recognition based on facial expressions without considering physiological signals and speech analysis. For the evaluation tasks, the commonly used metrics are AUC2, PCC, MAE, RMSE, and MSE, and the accuracy measure is rarely used.

B. Recurrent Neural Networks (RNNs)

Recurrent neural networks (RNNs) are another form of NN that is suitable for processing sequences of unequal length [17]. The use of unidirectional RNNs solves the problem of different lengths between the input and output. However, they are predicted using only earlier data. Therefore, the bidirectional model is used for prediction from both directions. In general, RNNs are capable of using three different settings [17].

- The first setting is *standard*, which learns from labelled data and predicts the output.
- The second setting is called *sequence setting,* and is able to learn data from multiple labels. It has sequences with combinations of different kinds of data, and cannot break it. Therefore, it takes a full sequence to predict the next state and more.
- The third setting is called *predict next setting,* and can take unlabelled data or implicitly labelling, such as words in a sentence. In this example application, the RNN breaks the words down into subsequences, and considers the next subsequence as a target.

The first successful model of an RNN was called a Hopfield network. In 1997, the most usable and important RNNs were LSTM NNs, which were proposed by Hochreiter and Schmidhuber. Next, we explain this kind of NN in more detail.

C. Long-Short Term Memory Neural Networks (LSTM-NNs)

LSTM is one of the most important architectures of NNs, and were proposed in 1998 by Hochreiter and Schmidhuber [17]. It is

usually used for sequential tasks such as time analysis and natural language processing [17]. In addition to the main link from one unit to the next in RNN, the LSTM has another link, i.e., the cell sate, that keeps or removes data through different gates [17]. The architecture of LSTM is shown in Figure 3.
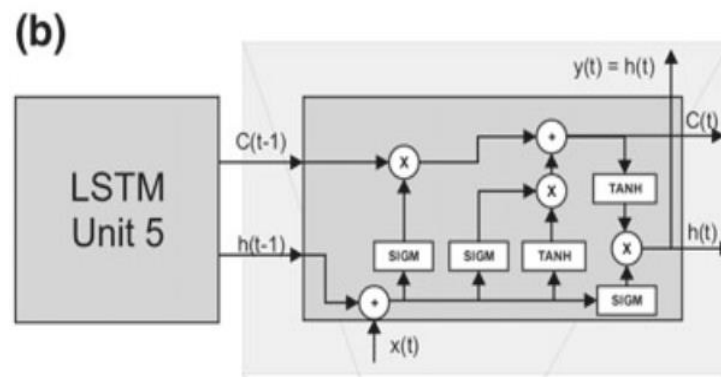

Figure 3. LSTM architecture [17].

In pain-recognition studies, researchers used this model for both feature extraction and classification tasks. The measures used for pain recognition are obtained only from facial expressions and speech analysis. They achieved a good performance when used for feature extraction [4] and also classification (Martinez, Rudovic, and Picard 2017) . However, they had a better result when used in classification, and combined with CNN for feature extraction, as in [6] . This comparison of the evaluation depends on metrics that are used in LSTM pain-recognition studies, namely, AUC, AUC2, MAE, ICC, and confusion matrices. We can compare the results obtained for CNN and LTSM in terms of MAE only, which had a better result (0.18) in LSTM than CNN (0.0389). In contrast, the weighted MAE had a better result for CNN (0.991).

D. Multi-Task Neural Network (MT-NN)

The multi-task learning (MTL) approach functions by sharing representation between related tasks in order to have a better generalization model [17]. For NNs, it is based on having a single NN that is able to perform many tasks better than performing individual tasks. Two techniques are mostly used for MTL-NN. The first method is the hard parameter sharing, which is based on sharing the hidden layers for all tasks, and then having different outputs. The second method is soft parameter sharing, which is based on modelling each task alone, and trying to achieve regularization in order to have similar parameters.

From the previous studies of pain recognition, we observe that this approach was rarely used. The reason is based on the difficulty of this approach, which has to perform many tasks using a single neuron. Therefore, the most applicable and widely used approach is transfer learning, which helps to have a small data set problem from the big data.

9

IV.   Datasets for Pain Recognition

In this section, we present the available data set for pain recognition. After reviewing the literature, we found eleven[20]–[30] data sets created for pain-recognition systems. The data sets were different in terms of their selected features, number of participants, stimuli method, pain levels, and the devices used for sensing the data. The main differences are summarized in Table 2 for comparison and discussion.

We can see that only three datasets were applied to patients with shoulder, back, and neck pain. The number of participants for all datasets ranged from 20–129 persons. The main stimuli methods employed for healthy people came from heat and cold sensing. As a new method, electrical stimuli have been used recently [10] for healthy people. The facial expression data set had more than two levels of pain. In contrast, when the biosignals and body movements were included in the datasets, they had only 2–4 levels. This was because of recent studies of pain recognition from body movements and biosignals, which require more studies and experiments.

V.   Challenges and Future Directions

Deep-learning methods are widely used in many applications. In recent times, pain recognition has been associated with sensory data rather than only facial expressions. Deep learning has shown promising results in many e-health systems with sensory data [2]. This achievement has resulted in greater interest in this approach by the pain-recognition community based on sensing data.

Based on the review of state-of-the-art with respect to the use of deep learning in pain recognition, we have found few recent studies. Most pain-recognition studies used CNNs and LSTMS owing to their ability to resolve many challenges related to deep-learning methods. However, many challenges and future directions need to be solved and fixed.

There is a need to better understand the problem of pain recognition. This will help to define the tools and additional methods such as the use of similarity measures and meta information [1].

Furthermore, it is important to extract the most suitable features, and deep learning will help in this case. Moreover, the use of a single model depending on facial expression is limited, and does not provide an accurate pain level. Therefore, the multimodality is considered as a solution and another challenge. This requires a combination of different kinds of signals, such as physiological signals with behaviour expression [4].

Moreover, it is important to find a new objective measure of pain based on sensory data and experiments, compared with the use of only instruments in the existing measures [4].

The available datasets ranged from 22–142 participants, which is not enough to generalize the model of deep learning. This is considered an open issue that needs to be addressed, and a dataset should be developed with several hundred participants.

In addition, the challenge that faces deep-learning methods in pain recognition is strongly related to data acquisition. Data acquisition includes many issues that are related to: (1) the devices and sensors used, (2) processing of signals, (3) dimensionality of data [18], and (4) reliability of data [18]. With respect to the sensors, it is more convenient to deal with wearables and mobile devices, but such tools also have many challenges in terms of short battery life, and more. More information about this issue can be found on our previous studies [31] for such kind of wearable devices.

Finally, most of the works were performed in a laboratory setting. However, the main challenge involves investigating the models in real clinical settings.

VI.   Conclusions

This study presented a systematic review of pain-recognition systems that are based on deep-learning methods only. Based on the papers reviewed, the new taxonomy of classification was presented based on the kind of data used. These pain-recognition data were obtained from facial expression, speech, or physiological signals.

Furthermore, this study described the primary deep-learning method that was used in review papers. Finally, the main challenges and future direction were discussed.

VII.   Data Availability

No data were used to support this study.

Dataset name	Title	Features	Devices	Stimuli	Participants	Classes
UNBC 2011 [20]	PAINFUL DATA: The UNBC-McMaster Shoulder Pain Expression Archive Database	Facial expression RGB	Two Sony digital cameras	Natural shoulder pain	129 Shoulder pain patients (63 males, 66 females)	0-16 (PSPI) and 0-10 (VAS)
BioVid 2013 [21]	Data for the Advancement and Systematic Validation of an Automated Pain Recognition System	- Video: Facial expression RGB - Biopotential signals (SCL, ECG, sEMG, EEG)	- Kinect camera - Nexus-32 amplifier	Heat pain at right forearm thermode (PATHWAY, http://www.medoc-web.com)	90 Healthy	4 levels of pains
BP4D- **Spontaneous** Database (BP4D) 2014 [22]	BP4D-Spontaneous: a high-resolution spontaneous 3D dynamic facial expression database	-Facial expression	Two stereo cameras and one texture video camera	Cold pressor test with left arm.	41 healthy	8 classes of pain as one the emotions (Happiness/amusement Sadness, Startle, Embarrassment Fear, Physical pain, Anger, Disgust)
BP4D+ 2016 [23]	Multimodal Spontaneous Emotion Corpus for Human Behavior Analysis,	-Facial expression -EDA, heart rate, respiration rate, blood pressure	-3D camera Di3D -infrared camera FLIR -BioPac	Same as before	141 healthy	Same as before
SenseEmotion 2016 [24]	The SenseEmotion Database: A Multimodal Database for the Development and Systematic Validation of an Automatic Pain- and Emotion-Recognition System	- Facial expressions: - Biosignals - ECG, EMG - GSR - RSP - Audio	- 3 cameras (iDS UI-3060CP-C-HQ) - g.MOBIlab - g.GSRsensor - Piezoelectric crystal sensor (chest respiration waveforms) www.gtec.at/Products/Electrodes-and-Sensors/g.Sensors-Specs-Features - Digital wireless headset microphone (Line6 XD-V75HS) + directional microphone (Rode M3)	Heat pain Medoc Pathway thermal simulator	40 heathy (20 male , 20 female)	5 (no pain, 4 levels of pain)
EmpPain 2016 [25]	The Automatic Detection of Chronic Pain-Related Expression: Requirements, Challenges and the Multimodal EmoPain Dataset	- Audio, Facial expressions - Body movements - sEMG	- 8 cameras - Animzaoo IGS-190 - BTS FREEEMG 300	Natural while doing physical exercises.	22 chronic low back pain (CLBP) (7 male, 15 female)	2 for face 6 for body behaviours Combined: binary
MIntPAIN 2018 [26]	Deep Multimodal Pain Recognition: A Database and Comparison of Spatio-Temporal Visual Modalities	Facial expression - RGB, Depth - Thermal	- Microsoft Kinect Version2 - Axis Q1922 thermal camera	Electrical pain	20 healthy	5 classes (0-4)
X-ITE pain 2019 [27]	Multi-Modal Signals for Analyzing Pain Responses to Thermal and Electrical Stimuli	- Audio, Facial expressions - ECG, SCL, sEMG (trapezius, corrugator, zygomaticus)	- 4 cameras - BioPac	Heat and electrical.	134 healthy adults	3
COPE 2005 [28]	SVM Classification of Neonatal Facial Images of Pain	Facial expression	heel lancing	heel lancing for blood collection	26 neonates (age 18-36 hours)	5 (pain, rest, cry, air puff, or friction)
YouTube 2014 [29]	Too many crying babies: a systematic review of pain management practices during immunizations on YouTube.	-Video -Audio	injection	immunizations (injection)	142 infants	FLACC observer pain assessment
IIIT-S ICSD 2016 [30]	Discriminating the Infant Cry Sounds Due to Pain vs. Discomfort Towards Assisted Clinical Diagnosis	-Audio	injection	immunizations (injection)	33 infants	6 (pain, discomfort, hunger/thirst, and three others)

Table 2. Summary of pain recognition data sets.

References

[1]   I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[2]   P. Werner, D. Lopez-Martinez, S. Walter, A. Al-Hamadi, S. Gruss, and R. Picard, 'Automatic Recognition Methods Supporting Pain Assessment: A Survey', *IEEE Transactions on Affective Computing*, pp. 1–1, 2019.

[3]   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)*, 2017, pp. 181–184.

[4]   F. S. Tsai, Y. M. Weng, C. J. Ng, and C. C. Lee, 'Embedding stacked bottleneck vocal features in a LSTM architecture for automatic pain level classification during emergency triage', in *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*, 2017, pp. 313–318.

[5]   D. L. Martinez, O. Rudovic, and R. Picard, 'Personalized Automatic Estimation of Self-reported Pain Intensity from Facial Expressions', *arXiv:1706.07154 [cs]*, Jun. 2017.

[6]   P. Rodriguez *et al.*, 'Deep Pain: Exploiting Long Short-Term Memory Networks for Facial Expression Classification', *IEEE Transactions on Cybernetics*, pp. 1–11, 2018.

[7]   J. Egede, M. Valstar, and B. Martinez, 'Fusing Deep Learned and Hand-Crafted Features of Appearance, Shape, and Dynamics for Automatic Pain Estimation', in *2017 12th IEEE International Conference on Automatic Face Gesture Recognition (FG 2017)*, 2017, pp. 689–696.

[8]   F. Wang *et al.*, 'Regularizing face verification nets for pain intensity regression', in *2017 IEEE International Conference on Image Processing (ICIP)*, 2017, pp. 1087–1091.

[9]   S. Jaiswal, J. Egede, and M. Valstar, 'Deep Learned Cumulative Attribute Regression', in *2018 13th IEEE International Conference on Automatic Face Gesture Recognition (FG 2018)*, 2018, pp. 715–722.

[10]   M. A. Haque *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)*, 2018, pp. 250–257.

[11]   M. Bellantonio *et al.*, 'Spatio-temporal Pain Recognition in CNN-Based Super-Resolved Facial Images', in *Video Analytics. Face and Facial Expression Recognition and Audience Measurement*, Springer, Cham, 2016, pp. 151–162.

[12]   K. R. Kulkarni, A. Gaonkar, V. Vijayarajan, and K. Manikandan, 'Analysis of lower back pain disorder using deep learning', *IOP Conf. Ser.: Mater. Sci. Eng.*, vol. 263, no. 4, p. 042086, 2017.

[13]   B. Hu, C. Kim, X. Ning, and X. Xu, 'Using a deep learning network to recognise low back pain in static standing', *Ergonomics*, pp. 1–9, Jul. 2018.

[14]   D. Lopez-Martinez, O. Rudovic, and R. Picard, 'Physiological and behavioral profiling for nociceptive pain estimation using personalized multitask learning', *arXiv:1711.04036 [cs]*, Nov. 2017.

[15]   M. Kächele *et al.*, 'Adaptive confidence learning for the personalization of pain intensity estimation systems', *Evolving Systems*, vol. 8, no. 1, pp. 71–83, Mar. 2017.

[16]   P. Thiam, P. Bellmann, H. A. Kestler, and F. Schwenker, 'Exploring Deep Physiological Models for Nociceptive Pain Recognition', *bioRxiv*, p. 622431, Apr. 2019.

[17]   S. Skansi, *Introduction to Deep Learning: From Logical Calculus to Artificial Intelligence*. Springer, 2018.

[18]   A. Obinikpo, B. Kantarci, A. A. Obinikpo, and B. Kantarci, 'Big Sensed Data Meets Deep Learning for Smarter Health Care in Smart Cities', *Journal of Sensor and Actuator Networks*, vol. 6, no. 4, p. 26, Nov. 2017.

[19]   A. Sathyanarayana *et al.*, 'Sleep Quality Prediction From Wearable Data Using Deep Learning', *JMIR mHealth and uHealth*, vol. 4, no. 4, p. e125, Nov. 2016.

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

[21]   S. Walter *et al.*, '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)*, 2013, pp. 128–131.

[22]   X. Zhang *et al.*, 'BP4D-Spontaneous: a high-resolution spontaneous 3D dynamic facial expression database', *Image and Vision Computing*, vol. 32, no. 10, pp. 692–706, Oct. 2014.

[23]   Z. Zhang *et al.*, 'Multimodal Spontaneous Emotion Corpus for Human Behavior Analysis', presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 3438–3446.

[24]   M. Velana *et al.*, 'The SenseEmotion Database: A Multimodal Database for the Development and Systematic Validation of an Automatic Pain- and Emotion-Recognition System', in *Multimodal Pattern Recognition of Social Signals in Human-Computer-Interaction*, 2016, pp. 127–139.

[25]   M. S. H. Aung *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, Oct. 2016.

[26]   M. A. Haque *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)*, 2018, pp. 250–257.

[27]   S. Gruss *et al.*, 'Multi-Modal Signals for Analyzing Pain Responses to Thermal and Electrical Stimuli', *J Vis Exp*, no. 146, Apr. 2019.

[28]   S. Brahnam, C.-F. Chuang, F. Shih, and M. Slack, 'SVM Classification of Neonatal Facial Images of Pain', 2005, pp. 121–128.

[29]   D. Harrison *et al.*, 'Too many crying babies: a systematic review of pain management practices during immunizations on YouTube', *BMC Pediatr*, vol. 14, p. 134, May 2014.

[30]   V. Mittal, 'Discriminating the Infant Cry Sounds Due to Pain vs. Discomfort Towards Assisted Clinical Diagnosis', 2016, pp. 37–42.

[31]   R. M. Al-Eidan, H. Al-Khalifa, and A. M. Al-Salman, 'A Review of Wrist-Worn Wearable: Sensors, Models, and Challenges', *Journal of Sensors*, 2018. [Online]. Available: https://www.hindawi.com/journals/js/2018/5853917/abs/. [Accessed: 07-Feb-2019].