Ensemble Neural Networks for Multimodal Acute Pain Intensity Evaluation using Video and Physiological Signals

Manisha S. Patil[1], Hitendra D. Patil[2]

[1,2]Department of Computer Engineering, SSVPS's BSD College of Engineering, Dhule, 424002, KBCNMU, Jalgaon, India
Email: manishavpatil2007@gmail.com[1], hitendradpatil@gmail.com[2]

Received: 17.04.2024 Revised : 14.05.2024 Accepted: 27.05.2024

ABSTRACT
Automated pain recognition is essential in healthcare. Research done previously suggests that for conventional algorithms, automated pain recognition relies on features gathered from video and physiological inputs electrodermal activity (EDA). The article presents investigations into designing a collaborative neural network structure that combines a fine-tuned, 3-stream hybrid deep neural network (HEDLM) with Convolutional Neural Network (CNN) to extract face image and EDA signal characteristics and recognize and precisely identify pain level. The experiments we conducted demonstrate that multimodal data on context works much better than uni-level data on context. Our research results for pain task recognition in Part A of the BioVid Heat Pain database contain pain levels 0 and 1, pain levels 0 and 2, pain levels 0 and 3, and pain levels 0 and 4. During LOSO (leave one subject out cross validation techniques) research, the classification task among levels of pain 0 and 4 reached an average accuracy of 84.8% for 87 individuals. This suggested approach takes advantage of deep learning capacity utilization to perform better than standard techniques by integrating facial images and physiological signals.

Keywords: Signal, Feature, Pain, Classification, Image

1. INTRODUCTION

Pain assessment is one of the difficult topics to perform that healthcare professionals have to perform [1]. This becomes especially difficult when medical professionals have to evaluate patients suffering who are not able to communicate. Even though pain is seen as a subjective sensation that patients must self-report, this does not mean that a patient cannot be feeling pain and require therapy to manage it [2]. Factors such as physiological status, behaviour, mental state, pain from past events, and intellectual disability can all impact pain. Considering these features, an effective pain evaluation model must be developed [3]. Self-reported reporting by patients is the most trustworthy approach to evaluating pain in health care settings. This approach is based on the patient's capability to convey a self-analysis of pain, and physicians may utilize this information in imperative to clarify the patient's pain understanding. Patients who are asleep, intoxicated, recuperating from a stroke, on artificial breathing, or suffering from advanced dementia may struggle to self-report their discomfort.Analyzing acute nociceptive pain can shed light on internal health issues, including the breakdown of tissue [4]. Acute pain is small and tough to accurately measure due to its complex nature and subjective character. Nowadays, medical professionals and caregivers evaluate acute pain using pain scale assessments. These methods can be ineffective for evaluating pain in humans with mental health conditions and need extremely skilled medical professionals to analyses it. Applying these approaches subsequently can be tricky to select the pain-relieving medicine prescriptions for the patient in the critical care section. Despite that significance, acute pain is frequently neglected [5]. In most situations, the reason and response to pain are easily understood, yet it seldom fails. To detect acute pain, the pain evaluation must be effectively finished. Handle pain in hospitals; a practice of acute pain services (APS) has developed. APS is additionally accountable for addressing patients negative effects caused by medication for pain on a daily basis. The research will be piloted frequently to examine the success of the management. APS has been used in advanced nations for three decades with accurate pain guidelines and recommendations. In India, APS remains neglectedand skipped in regular medical practices [6]. Data from demonstrates that the APS method has not been efficiently implemented in India [7]. Only surgeons and anesthesia specialists regulate pain in the operating theatre and critical care units. Acute pain is frequently ignored [8].
Pain assessment methods that may be used to evaluate pain in nonverbal sufferers include the nonverbal pain scale (NVPS). These methods must have specially trained observer to evaluate pain [9]. Therefore, a

clear and significant disadvantage of these screening methods is their low reliability among observers due to assessment inaccuracy or misunderstanding. An additional drawback of these approaches is that they typically generate individual measurements. Conducting plenty of clinical observations is expensive, leading to increased pressure and burnout among nurses [10]. Consider such constraints since they may result in under- or over-treatment of pain, causing unneeded suffering. Pain from powerful medicines and/or painkiller misuse [11].Clinicians have a crucial role in evaluating and dealing with acute pain following surgery. According to reports, three-quarters of individuals who have surgical operations endure severe discomfort [12]. Acute pain is a kind that continues simply a few minutes and typically arises from an accident, disease, or damage to the tissue. In this situation, surgical treatment of pain is critical for speeding a patient's transition to ordinary utility, improving the comfort of patients, and avoiding subsequent difficulties [13]. Furthermore, inefficient acute pain treatment can induce psychological as well as physical suffering, such as sadness, nervousness, or chronic pain; it can also raise the likelihood of problems resulting in increased hospital admissions. Health issues can contribute to higher healthcare expenses for patients. Consequently, a tool that may be helpful medical professionals in determining acute pain needs to be developed, which ultimately leads to a more reliable, accurate, and trustworthy diagnosis of pain.

Consequently, humans need automated pain assessments. Pain recognition applications are utilized in medical applications to promote wellness and help recovery through physical treatment. Pain recognition methods utilize physiological and behavioural features to classify problems. Indicators involve physiological indications, expressions of the face and movements of the body, speech, or an amalgamation of them. Assessment of pain based on patient behaviour might not always be accurate. The patient can take control of their emotional expression. Moreover, patients show actions related to pain based on their individual traits. A few patients lose consciousness and fail to articulate uncomfortable emotions in a straightforward manner. It's challenging to determine pain using emotional actions. It is crucial to distinguish pain through bio-medical indicators and face images expressions. Based on concepts, suggested for this pain is associated with dysfunction of the autonomic nervous system [14]. The human brainstem subconsciously determines its action. It is divided into two distinct sections. These represent the Sympathetic Nervous Systems (SNS) and Parasympathetic Nervous System (PNS). The PNS becomes active under relaxation, whereas the SNS becomes active under stress or pain. However, these metrics are not effective [3]. The ANS stimulates and SNS, triggering substantial modifications in physiological indictors include electrodermal activity (EDA), photoplethysmography (PPG), Electrocardiography (ECG), Heart Rate Variability (HRV), Heart Rate (HR), etc.,[9] that may be utilize Consequently, humans need automated pain assessments. Pain recognition applications are utilized in medical applications to promote wellness and help recovery through physical treatment. Pain recognition methods utilize physiological and behavioural features to classify problems. Indicators involve physiological indications, expressions of the face and movements of the body, speech, or an amalgamation of them. Assessment of pain based on patient behaviour might not always be accurate.

This research work proposes to develop and assess a hybrid ensemble deep learning framework that recognizes the level of pain at several classification levels using patient video frame images and physiological signals as EDA. The proposed ensemble approach includes two stages: early fusion for feature extraction and late fusion for classification. In the initial phase of the fusion stage, feature extraction again has two stages; we use two modality, one is facial expression and the second is physiological signals as EDA. A feature extraction approach depending on the optimize VGGFace method, which combines Principal Component Analysis (PCA) systematically, has been applied to obtain features included on human face images. A single dimensional convolutional neural network (1D-CNN) is a neural network that works with single-dimensional data, such as sequence data. In this work, 1D CNNs have been utilized to retrieve relevant features from physiological signals using EDA. This can be done using one-dimensional convolution methods with filters. Such extracted relevant features are combined to produce a set of feature vectors. At the last stage of fusing, a three-stream CNN-RNN network is now created, along with the subsequent combined features (facial expression image + physiological signal EDA) applied as input to the ensemble classification model. Pain identification is a binary categorization that differentiates between painful and non-painful indicators. The present paper investigates the effectiveness of the presented model using Part A of the BioVid Heat Pain Database [32]. Part A of the BioVid Heat Pain database is split into 5 classes: 4 painful subclasses and a standard subclass that represents a non-painful class group. In this research, we present recommendations for all 87 subjects. We also compared it with earlier research that included all 87 subjects.

1.1. The following statements outline the research's motivations and contributions

A VGG-Face pre-trainer is utilized in a unique hybrid image classification technique through early fusion section to effectively and effectively extract features from facial images. The outputs are combined with PCA to minimize the dimensionality of the image dataset. 1D-CNNs have been used to extract representative features from physiological signals using EDA. This can be done using one-dimensional convolution methods with filters. This is a novel approach to classifying pain using physiological signals (EDA) and facial image features to develop a 3-stream ensemble CNN and RNN classifier structure. The outputs are fused through the late fusion stages, resulting in the final classification of the pain level into four classes. Our investigations indicate that contextual information with several levels of significance is more important than contextual information with only one level. The general structure is referred to as the hybrid ensemble deep learning approach (HEDLM), and it is trained and tested through Part A of the BioVid Heat Pain dataset, which are characterized by various pain modalities. The outcomes obtained can be utilized to benchmark HEDLM against the latest methods as the baseline model.

This paper covers a number of sections, which include: The second section discusses the fundamental concepts in deep neural networks implemented in the pain assessment automated methodology. The third section addresses the sub-net configurations used with our architecture, as well as the interpreting of multiple-level information regarding context. The fourth section provides the dataset, learning knowledge, illustrations, comparison outcomes of the recommended approach, and discussion. The final section include conclusion of the article.

2. RELATED WORK

The following sections detail specific investigations in pain recognition using facial expressions and physiological data such as EDA peak characteristics, along with a basic review of deep learning methods, earlier investigations, and ensemble neural networks.

2.1. Deep neural networks are utilized for expressions on the face and biopotantial signals

CNNs are capable of classifying images and recognizing faces and objects [33, 34]. CNNs, along with associated pre-trained algorithms, achieved significant results, especially for the extraction of features and classification of images [35]. Moreover, recently, CNN based algorithms have obtained superior results on the data set of ImageNet, such as AlexNet [36] and GoogLeNet [37]. Handcrafted features performed better with pre-trained CNNs in machine vision tasks involving object detection and emotion recognition. Although deep learning approaches are effective for task estimation, these are not adequate for analysing progressive data that includes voice or image data. Consequently, RNN evolved to analyse features by integrating data from all previous intervals and updating its depiction with recent data. [38].

Long short-term memory (LSTM) networks utilize the RNN framework and include feedback connections and similar feed forward neural network architectures. Conventional RNNs, similar to LSTMs, can gain knowledge from dependence over time. However, training systems can be challenging due to gradients that disappear or burst. LSTM contains three gates that govern the cell state: neglect, output, and input. Important data from previous phases is stored in the forget gate. The input gate updates the present stage's information. The resultant gate sets the next concealed stage condition [39, 40]. The Figure 1 represents the framework of an LSTM cell. Sequence labelling activities demand connection between current (left) and future elements. The hidden layer of the LSTM just takes in information from the prior frame and suppresses data from subsequent frames. Bidirectional LSTM (BiLSTM) [41] is an efficient approach that captures past and future information by presenting separately sequence both forward and backward as two independent hidden states.

2.2. Existing automatic pain recognition techniques based on facial expressions

To be able to recognize pain from expressions on the face, several deep learning (different hidden layers) and non-deep learning algorithms are being created, with notable advancements being achieved in this field under investigation at the present time. Support vector machines (SVM) based traditional non-deep learning approaches are being utilized to categorize facial expression features in a variety of classification challenges.

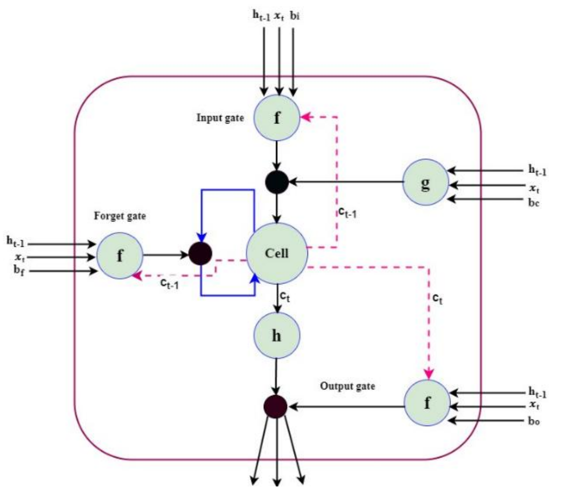

Figure 1. The structure of the LSTM component [39, 40]

Non-deep learning methods of feature extraction applied to the assignment included Active Appearance Models (AAM) and Active Shape Models (ASM), Gabor wavelets, and Local Binary Patterns (LBP). As an instance, [42, 43] integrated AAM-based characteristics with SVM classifiers for pain recognition [44, 45]. They used Gabor wavelets as the fundamental elements inside their filter banks and features of LBP with SVMs, separately. Deep learning algorithms, which have an excellent capacity to identify fundamentally hidden outlines in complicated images, are currently being used in feature extraction as well as classification problems, owing to significant advances in deep learning capacities and the increasing availability of large amounts of training data. Convolutional neural networks and deep belief models are instances of these deep models that have been shown to help improve the feature extraction process [36, 46]. In another study, [38] presented a way to acquire features from previously trained CNN and integrate them with RNN as a new model to create a continuous predictive framework that uses RNN to predict levels of pain from facial expressions. The same approach [47, 48] was used to make use of the time-dependent relationships between frames of video by extracting relent features of faces from previously trained VGGFaces and integrating them into an LSTM.

A 3-channel structure with three distinct feature extraction methods is shown in [49], in which Relevance Vector Machines (RVM) are utilized to compute the pain and the presence histogram of oriented gradients (HOG), CNN, and contour features are extracted using custom methods. The Automated Facial Expression Recognition (AFER) framework, which calculates frame-level scores of confidence for individual AUs, and the Multiple Instant Learning (MIL) framework, which uses contributions from a collection of AU pairings that are important to pain to conduct sequence-level pain prediction, are two machine learning mechanisms that are included in the a 3-channel structure with three distinct feature extraction methods is shown in [49], in which Relevance Vector Machines (RVM) are utilized to compute the pain and the presence histogram of oriented gradients (HOG), CNN, and contour features are extracted using custom methods. The Automated Facial Expression Recognition (AFER) framework, which calculates frame-level scores of confidence for individual AUs, and the Multiple Instant Learning (MIL) framework, which uses contributions from a collection of AU pairings that are important to pain to conduct sequence-level pain prediction, are two machine learning mechanisms that are included in the recommended automated pain recognition system [50]. A new overview publication provides further information regarding automated pain recognition techniques [51].

2.3. Existing automatic pain recognition techniques are based on physiological signals

This Several studies have also looked at single-modal techniques to recognize acute pain that rely on the electrocardiogram's heart signal. Recurrent neural networks (RNNs) were developed by Martinez and Picard [52] and inter-beat intervals from ECG data. Deep 1D CNN was utilized by Thiam et al. [53] for integrating galvanic skin response (GSR), electrocardiogram (ECG), and ECG signals. They studied both multimodal fusion methods and uni-modal approaches. Notably, [54] demonstrated excellent results in both binary and multiclass classification conditions by proposing a system to extract pseudo-heart rate data from images using a CNN. Features of heart rate instability were retrieved in [55], and significant pain recognition results were obtained utilizing a random forest classifier. Inter-beat intervals were used to determine a number of features, including heart rate, with some interesting results [56]. A significant enhancement in performance was seen in the subsequent research [57], following the creation of multi-task fully linked neural networks.

A feasible approach, emphasizing the multiple characteristics of pain, is to integrate modalities into a multimodal framework. Integrating information from a variety of sources may improve the pain assessment's sensitivity and specificity. The prediction capabilities for single modalities are positive, but when they are combined, the results are usually superior [58]. Furthermore, utilizing cues from many channels may be advantageous as well as necessary, especially in clinical situations when a modality may not be available for a variety of reasons. Utilizing several types of fusion approaches and utilizing a variety of information extracted from physiological and video sources, such as head movement, facial expression, GSR, EMG, and ECG, provided very encouraging results [59]. Integrated facial expressions and head position characteristics, along with several bio potential features obtained from ECG, EMG, and GSR [60], concentrated only on biomedical signals [61].Three of the mentioned physiological signals were used in [62], whereas different combinations of manually generated and acquired features were extracted from a BiLSTM model. To reduce the range of features, the minimal relevance method (MRMR) was first used, with encouraging results. The research presented in [63] generated a latent demonstration for every bio-potential signal using deep de-noising convolutional autoencoders (DDCAE). A weighting step followed after that, and the classification technique showed encouraging results. To achieve state-of-the-art results, [49] utilized a 3D CNN and combined computational facial features with pseudo-heart rate data.

3. METHODOLOGY

3.1. The proposed ensemble deep learning method

The present study presents an innovative hybrid ensemble deep learning approach (HEDLM) that can classify the intensity of pain over four levels using facial expression video frames and physiological signals as electrodermal activity (EDA).

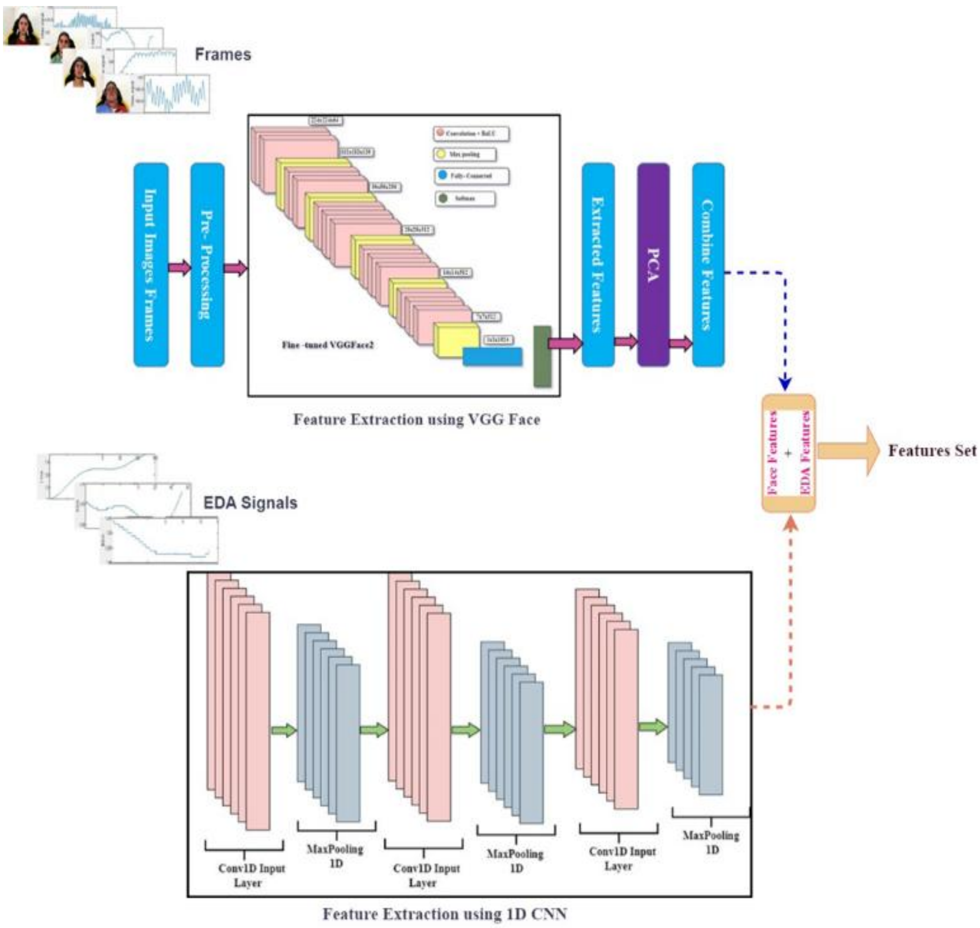

Figure 2. Proposed block diagram for feature extraction using VGG and 1D-CNN architecture

The proposed technique's schematic structure is shown in the Figure 2 the dataset involves initial processing and normalization before images are inserted within the presented deep neural network algorithm. The HEDLM contains two sections: feature extraction and classification. In the first phase, early fusion extracts and integrates features using a combination of linear PCA using previously trained CNN. The retrieved features are subsequently utilized in late fusion for classification. In the second phase, 1-D CNNs are utilized for obtaining characteristic features from physiological signals (EDA) in the time domains. Then, facial expression features and EDA signal features are combined and conveyed in the late fusion for pain level classification. Later fusion evaluates pain intensities into four categories using a combined three-stream CNN+RNN combination deep machine learning network.

3.2. Facial expression feature extraction

It is strongly suggested that comparison wih results from other published articles are provided to give more context and to strengthen the claim of novelty.To construct the proposed HEDLM approach, the initial step is to develop the initial fusion feature extraction segment. Additionally, pre-processed information is sent to the initial fusion process for feature extraction. The initial fusion part includes the adjusted VGG Face previously trained with facial images [64], and results are incorporated with PCA.Within visual computing, transfer learning frequently contains the practice of already pre-trained architecture, which includes an algorithm trained on a big standard database, to correct an issue. Numerous cutting-edge algorithms employ transfer learning to get image classification outputs [65].

The VGGFace model includes 5 blocks of convolution and 3 layers that are completely connected. In order to improve it, a deeply connected framework is created on top of the VGGFace framework, the convolution layers remain stationary, and input is frequently provided into the network [66]. Convolutional neuralnetwork approaches retrieve relevant features may from training data. However, the primary difficulty with the current research is that the elements of the derived image attributes may significantly improve with the integration of additional network layers [67]. To overcome a particular issue, after utilizing deep learning techniques to identify image features, the PCA approach is utilized in this investigation to reduce dimension. The research incorporates PCA, a way to decrease dimensionality that is beneficial for a range of applications [68]. This technique might assist us in selecting the most appropriate collection of information qualities that can help the framework act more effectively and enhance the accuracy of the algorithm's performance [69]. An aggregate of 125,280 features were extracted from the training data set and determined upon the provided input shape. The data to be trained is represented as (31,320, 4), with 34,800 referring to the number of training images which generates an output of 31320 x 4 = 125280. In addition, the 4 separated output features (for each image) acquired from the optimized VGG-Face are passed on into the PCA method with the objective to reduce the overall dimension of the retrieved features and improve classification performance.

3.3. Physiological signals (EDA) feature extraction

The single-dimensional convolutional neural network (1-D-CNN) is a neural network that works with single-dimensional data, which can be sequence data. This work employs 1-D CNNs to derive accurate characteristics from physiological signals (EDA) errors in time frame domains. This happens through 1D convolution processes with the filters. Figure 2 illustrates the CNN framework, which contains layers of convolutional filters, a MaxPooling1D layer, an FC layer, and a classification layer using a Rectified Linear Unit (RELU) function as the activation function. Dropouts and batch sizes have been used to prevent excessive overfitting. Equations (1-3) state the calculations: (RELU), which may result in nonlinear behaviour, is utilized as a CNN function of activation. Whereas deep neural networks initially became popular for image classification applications, they are now utilized for image frame evaluation and recognition. The applications of one-dimensional time series data for classification are very recent [70]. Physiological signals (EDA) may be viewed as a sequential modelling attempt; hence, 1D CNNs are a reasonable option. Compressed 1D-CNNs are appropriate for real-time applications because they have reduced processing needs.

$$x_{o,fl}^{l} = f\left(\sum_{im} x_i^{l-1} * k_{io,fl}^{l} + b^{l}\right) \quad (1)$$

$$x_o^{l} = f\left[\max\left(\sum_{im} x_i^{l-1}\right) + b^{l}\right] \quad (2)$$

$$x_0^{l} = \left(x_i^{l-1} * d_{io}^{l} + b^{l}\right) \quad (3)$$

Here x is a single dimensional inputs data matrix $(n \times 1)$, f (.), $k_{io,fl}^{l}$ is a kernel filter of size $(k \times 1)$, $x_{o,fl}^{l}$ is represented output of convolutional layer, vector bias is denoted by b and learning parameter is represented as a d.

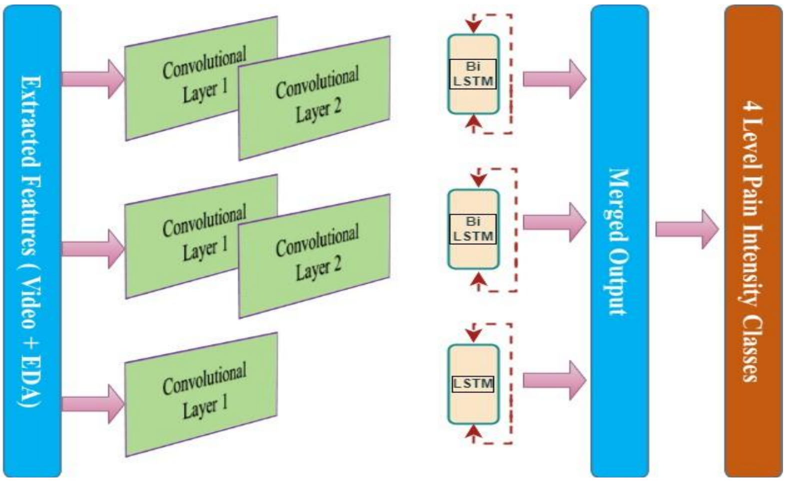

Figure 3. Propose framework of HEDLM method for pain recognition

3.4. Classification using ensemble deep neural network

Throughout the final fusion phase of the proposed HEDLM, a collective deep learning system with separated starting weights and network structure is developed for the purpose of classification. As indicated inside the related works section, collaborative learning is a useful technique that can enhance classification's capacity for generalization. Since the features from video image frames and accurate characteristics from physiological signals (EDA) feed to RNN. A technique for back propagation is utilized in RNN training [71].

A different iteration of the proposed approach was performed. The conclusions of the experiment indicated that systems that just contain RNN in the last fusion produce less precise results than those used in a hybrid CNN and RNN combination. Therefore, the results of three unique and amalgam CNN and RNN neural network techniques are combined. The combined result for evaluating the level of intensity of pain. Each of these independent, composite deep neural networks, DNN-1, DNN-2, and DNN-3, was created with various weights, parameters, and architectural configurations.Figure 3 illustrates the late fusion framework of the proposed HEDLM model.

3.5. Database and a setup for experiments

The purpose of the methodology (HEDLM) and all other comparative algorithms in this investigation have been created on a 32 GB RAM computer with an Intel Core i7 @ 3.5 GHz and a GTX 1060 NVIDIA graphics card. The framework building and prototype were done using the python framework, which contains an open-source deep learning library. Keras allows both continuous and convolutional networks, making prototyping quick and easy. Simulation data can be visualized and statistically analyzed using Matplotlib, a python 2D graphing package. The following provides a discussion of the evaluation measures and the selected database.

3.6. Part A of the biovid heat pain dataset

The approach presented has been evaluated using the (Part A) Bio-Vid Heat Pain database [76]. This multi-modal database includes 87 participants who have four differently measured levels of heat-induced pain (T1, T2, T3, and T4). Throughout the experiments, a number of modalities were recorded, including video streams, EMG, ECG, and EDA signals. The present work concentrates completely on recorded video frames (facial expression) and physiological signals (EDA). Every threshold of heat-induced pain was arbitrarily triggered 20 times. Every stimulation continued for 4 seconds, which included an arbitrary relaxation period of 8 to 12 seconds. Throughout the process of recovery, the baseline temperature (T0) was 320^0 C. In order to represent every one of the 87 individuals, an overall sample of 20 x 5 = 100 random samples was used. The raw dataset includes 8700 samples, all annotated with the intensity of heat-induced pain elicitation (P0, P1, P2, P3, and P4).

4. RESULTS AND DISCUSSIONS

We develop and evaluate our technique using the BioVid Pain dataset. Next, the analyzed outcomes are related to the benchmark methods and current research. We developed a HEDLM framework using a CNN and an RNN based on raw EDA signals and video frames. Firstly, we have used early fusion fine-tuned VGGFace and PCA to extract facial expression features from video signals and Second, the physiological signals (EDA) feature has been extracted using the 1D-CNN network. Last, facial expression and EDA

signal features are combined and transmitted in the last fusion for pain level classification. Late fusion classifies pain intensities into four distinct groups using a collective three-stream CNN+RNN composite deep neural network. We analyses the proposed approach using the BioVid Heat Pain Part A dataset as the foundation and employ a technique entitled LOSO cross validation in which one subject has been allocated as a test and the others can be used for learning. There are 87 lessons for 87 subjects. In order to compare the training phase average accuracy (in percentage terms) with earlier research, the LOSO cross-validation assessment and normative variation are also provided. We carry out four binary classification tasks to identify pain. There are additional experiments performed using various modalities. Table 1 displays the classification task performance for 87 individuals participating in the LOSO cross-validation. We reach the conclusion that the proposed approach for evaluating pain actually utilizes EDA. For classification tasks, the amalgamation of EDA and video data commonly performs better. The results demonstrate that multi-modal feature data performs much better than single-level data generated from physiological and video signals. Table 2 demonstrates an assessment of the proposed methodology based on the largest pain classification compared with earlier research. Table 2 presents the results of our EDA signal study, in which we were able to get an accuracy of 83.6% for 87 subjects. Figure 4 demonstrate graphical representation of accuracy of pain classifications in a LOSO cross-validation. Figure 5 illustrates a graphical comparison of results for classification accuracy evaluates with existing methods.
The results we obtained for the EDA signal classification of pain level 0 and pain level 4 are, in addition, the best. Further, our methodology, which alone combines EDA and video, outperformed methods that use both modalities together in terms of results. Fusion signal classification using the HEDLM framework approach gives an accuracy of 84.8% for 87 subjects. In this field, our technique performs effectively for the classification of pain level 0 and pain level 4.

Table 1. Pain classification accuracy in a loso cross-validation assessment using part A

Modality	Pain (0-1)	Pain (0-2)	Pain (0-3)	Pain (0-4)
EDA	60.6	67.5	76.2	83.6
ECG	51.2	50.3	53.6	63.2
Video + EDA	61.4	68.3	76.3	84.8

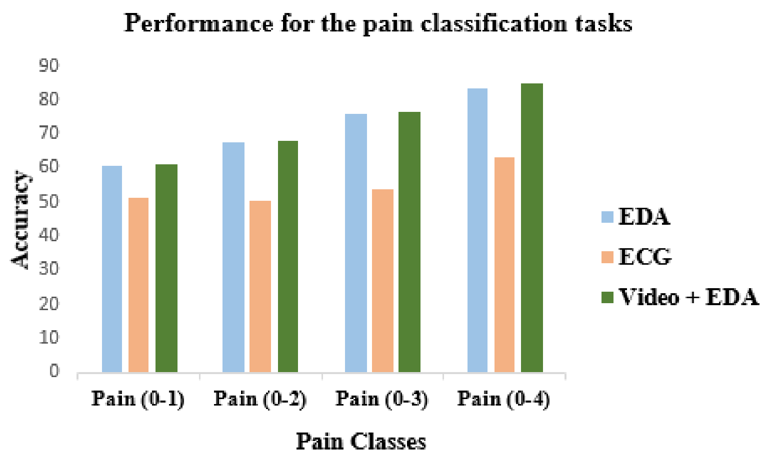

Figure 4. Accuracy of pain classifications in a LOSO cross-validation

The Table 1 demonstrates the way the proposed technique achieved classification tasks in comparison to several previous techniques. For further issues with classification, the proposed approach performs better than different methods in the majority of circumstances.

Table 2. Comparison of final results for taskclassification in a LOSO cross-validation analysis utilizing part A of the biovid heat pain

Method	Modality	CV Approach	Pain (0 vs 1)	Pain (0 vs 2)	Pain (0 vs 3)	Pain (0 vs 4)
Wamg et al. [72]	All Physiological	LOSO	58.5	64.2	75.1	83.3
Lopez et al. [73]	Physiological (One)	LOSO	56.4	59.4	66.00	74.2
Lopez et al.	Physiological	Ten	54.2	59.71	70.0	82.7

[74]	(Two)	FOLD				
Werner et al. [75]	Facial Image	LOSO	53.3	56.0	64.0	72.4
Werner et al. [76]	Physiological signals + Image	FIVE FOLD	49.6	60.5	72.0	80.6
Proposed Method	Video + EDA	LOSO	61.4	68.3	76.3	84.8

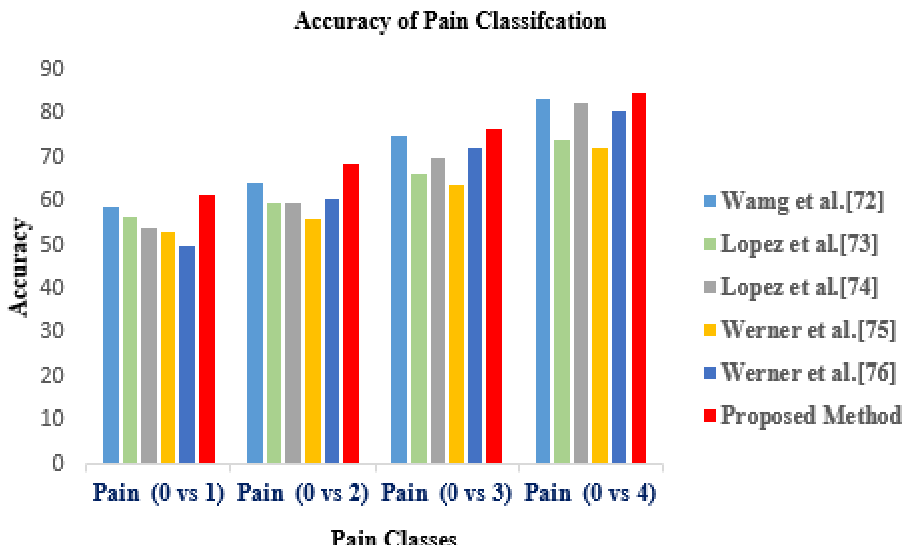

Figure 5. Comparison of results for the classification accuracy tasks existing methods

5. CONCLUSION

The present research provides a novel classification algorithm with a collaborative deep learning methodology to improve current initiatives in developing neural network technologies for pain identification utilizing physiological indicators (EDA) and facial expression images. Our approach executes the functions of extraction and classification of features, entirely substituting the difficult human process of feature extraction that requires professional skill. As a result, the three stream-separate CNN-RNN-based networks whose weights and structures have been seen to differ are integrated into the final HEDLM model, which represents features derived from EDA signals and facial images. The HEDLM methodology combines the customized VGGFace approach, PCA, and 1-D CNN techniques. The ensemble deep neural network approach, consisting of three separate CNN-RNNs, was subsequently developed and assessed for the precision of classification. The suggested HEDLM algorithm has been thoroughly evaluated for binary classification of initial and different pain levels using Part A of the BioVid Heat Pain database. The experiments we performed in Part A of the BioVid Heat Pain database demonstrate that multiple-level situation information is greater in significance than single-level context information. The results we obtain highlight the critical role of EDA and facial signals in pain classification. Integrating EDA signals and face images enhances the accuracy of the classification process utilizing Part A of the BioVid Heat Pain database. In conclusion, the approach used by deep learning takes greater promise to substitute traditional approaches to pain recognition challenges.

REFERENCES

[1] D. B. McGuire, K. S. Kaiser, M. E. Haisfield-Wolfe, and F. Iyamu, "Pain assessment in noncommunicative adult palliative care patients," Nursing Clinics, vol. 51, no. 3, pp. 397–431, 2016.

[2] H. Merskey, Ed., Classification of Chronic Pain: Descriptions of Chronic Pain Syndromes and Definitions of Pain Terms. Pain, 1986.

[3] P. Werner, D. Lopez-Martinez, S. Walter, A. Al-Hamadi, S. Gruss, and R. W. Picard, "Automatic recognition methods supporting pain assessment: A survey," IEEE Trans. Affective Comput., vol. 13, no. 1, pp. 530–552, 2019.

[4] Y. Chu, X. Zhao, J. Han, and Y. Su, "Physiological signal-based method for measurement of pain intensity," Front. Neurosci., vol. 11, no. MAY, pp. 1–13, 2017, doi: 10.3389/fnins.2017.00279.

[5] S. D. Subramaniam, B. Doss, L. D. Chanderasekar, A. Madhavan, and A. M. Rosary, "Scope of physiological and behavioural pain assessment techniques in children – A review," Healthcare Technol. Lett., vol. 5, no. 4, pp. 124–129, 2018, doi: 10.1049/htl.2017.0108.
[6] P. N. Jain, S. G. Bakshi, and R. S. Thota, "Acute pain services in India: A glimpse of the current scenario," J. Anaesthesiol. Clin. Pharmacol., vol. 31, no. 4, pp. 554–557, 2015, doi: 10.4103/0970-9185.169088.
[7] S. S. K. Khatib, S. S. Razvi, and S. Parab, "A multicentre survey of the current acute postoperative pain management practices in tertiary care teaching hospitals in Maharashtra," Indian J. Anaesth., vol. 49, no. 4, pp. 257–262, 2019, doi: 10.4103/ija.IJA.
[8] K. Deldar, R. Froutan, and A. Ebadi, "Challenges faced by nurses in using pain assessment scale in patients unable to communicate: A qualitative study," BMC Nurs., vol. 17, pp. 1–8, 2018, doi: 10.1186/s12912-018-0281-3.
[9] S. D. Subramaniam, B. Doss, L. D. Chanderasekar, A. Madhavan, and A. M. Rosary, "Scope of physiological, behavioural pain assessment techniques in children – a review," Healthcare Technol. Lett., vol. 5, pp. 124–129, 2018, doi: 10.1049/htl.2017.0108.
[10] J.-M. Roué, I. Morag, W. M. Haddad, B. Gholami, and K. J. Anand, "Using sensor-fusion, machine-learning algorithms to assess acute pain in non-verbal infants: A study protocol," BMJ Open, vol. 11, p. e039292, 2021, doi: 10.1136/bmjopen-2020-039292.
[11] K. Herr, P. J. Coyne, M. McCaffery, R. Manworren, and S. Merkel, "Pain assessment in the patient unable to self-report: Position statement with clinical practice recommendations," Pain Manag. Nurs., vol. 12, pp. 230–250, 2011, doi: 10.1016/j.pmn.2011.10.002.
[12] L. Köse Tamer and G. SucuDağ, "The assessment of pain and the quality of postoperative pain management in surgical patients," Sage Open, vol. 10, 2020.
[13] S. M. Ayasrah, T. M. O'Neill, M. S. Abdalrahim, M. M. Sutary, and M. S. Kharabsheh, "Pain assessment and management in critically ill intubated patients in Jordan: A prospective study," Int. J. Health Sci., vol. 8, pp. 287, 2014.
[14] S. N. Raja, "Role of the sympathetic nervous system in acute pain and inflammation," Ann. Med., vol. 27, pp. 241–246, 1995, doi: 10.1016/j.mpaic.2013.09.003.
[15] S. Koelstra, M. Pantic, and I. Patras, "A dynamic texture-based approach to recognition of facial actions and their temporal models," IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 11, pp. 1940–1954, 2010.
[16] Y. Tian, T. Kanade, and J. F. Cohn, "Facial expression recognition," in Handbook of Face Recognition, London: Springer, 2011, pp. 487–519.
[17] Z. Hammal, M. Kunz, M. Arguin, and F. Gosselin, "Spontaneous pain expression recognition in video sequences," in Proc. BCS Int. Acad. Conf., London, UK, 2008, pp. 191–210.
[18] Z. Ambadar, J. W. Schooler, and J. F. Cohn, "Deciphering the enigmatic face: The importance of facial dynamics in interpreting subtle facial expressions," Psychol. Sci., vol. 16, no. 5, pp. 403–410, 2005.
[19] P. Werner, A. Al-Hamadi, K. Limbrecht-Ecklundt, S. Walter, S. Gruss, and H. C. Traue, "Automatic pain assessment with facial activity descriptors," IEEE Trans. Affective Comput., vol. 8, no. 3, pp. 286–299, 2017.
[20] M. Webber and R. F. Rojas, "Human activity recognition with accelerometer, gyroscope: A data fusion approach," IEEE Sens. J., vol. 21, pp. 16979–16989, 2021, doi: 10.1109/JSEN.2021.3079883.
[21] N. Hirachan, A. Mathews, J. Romero, and R. F. Rojas, "Measuring cognitive workload using multimodal sensors," in Proc. 44th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), Glasgow, 2022, pp. 4921–4924, doi: 10.1109/EMBC48229.2022.9871308.
[22] M. Jiang et al., "Acute pain intensity monitoring with the classification of multiple physiological parameters," J. Clin. Monit. Comput., vol. 33, pp. 493–507, 2019, doi: 10.1007/s10877-018-0174-8.
[23] J. G. Chang, L. H. Ma, and S. L. Tsay, "Relationships between pain intensity and heart rate variability in patients after abdominal surgery: A pilot study," Chin. Med. J., vol. 125, no. 11, pp. 1964–1969, 2012.
[24] P. Dutta and M. N., "Facial pain expression recognition in real-time videos," J. Healthc. Eng., 2018, Art. no. 7961427, doi: 10.1155/2018/7961427.
[25] S. D. Subramaniam and B. Dass, "Automated nociceptive pain assessment using physiological signals and a hybrid deep learning network," IEEE Sens. J., vol. 21, no. 3, pp. 3335–3343, 2021, doi: 10.1109/JSEN.2020.3023656.
[26] F. Radenović, G. Tolias, and O. Chum, "Fine-tuning CNN image retrieval with no human annotation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 41, pp. 1655–1668, 2018.
[27] Y. Jiang et al., "A novel distributed multitask fuzzy clustering algorithm for automatic MR brain image segmentation," J. Med. Syst., vol. 43, p. 118, 2019.

[28] G. Tsoumakas and I. Vlahavas, "Random k-labelsets: An ensemble method for multilabel classification," in Proc. Eur. Conf. Mach. Learn., Warsaw, Poland, 2007, pp. 406–417
[29] B. Liu, Q. Cui, T. Jiang, and S. Ma, "A combinational feature selection and ensemble neural network method for classification of gene expression data," BMC Bioinformatics, vol. 5, no. 1, p. 136, 2004.
[30] T. G. Dietterich, "Ensemble methods in machine learning," in International Workshop on Multiple Classifier Systems, Günzburg, Germany, 2000, pp. 1-15. doi: 10.1007/3-540-45014-9_1.
[31] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in ICLR, San Diego, 2015, pp. 1-14.
[32] S. W. H. E. Walter and S. Gruss, "The BioVid Heat Pain Database: Data for the advancement and systematic validation of an automated pain recognition system," in IEEE International Conference on Cybernetics (CYBCONF), 2013, pp. 128-131.
[33] Y. Taigman, M. Yang, M. A. Ranzato, and L. Wolf, "DeepFace: Closing the gap to human-level performance in face verification," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 1701-1708.
[34] Y. Sun, X. Wang, and X. Tang, "Deeply learned face representations are sparse, selective, and robust," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 2892-2900.
[35] F. Hu, G.-S. Xia, J. Hu, and L. Zhang, "Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery," Remote Sensing, vol. 7, no. 11, pp. 14680-14707, 2015. doi: 10.3390/rs71114680.
[36] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in Advances in Neural Information Processing Systems, Nevada, USA, 2012, pp. 1097-1105.
[37] C. Szegedy et al., "Going deeper with convolutions," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1-9.
[38] J. Zhou, X. Hong, F. Su, and G. Zhao, "Recurrent convolutional neural network regression for continuous pain intensity estimation in video," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, Hawaii, USA, 2016, pp. 84-92.
[39] J. Schmidhuber, "Deep learning in neural networks: An overview," Neural Networks, vol. 61, pp. 85-117, 2015.
[40] F. A. Gers and J. Schmidhuber, "LSTM recurrent networks learn simple context-free and context-sensitive languages," IEEE Transactions on Neural Networks, vol. 12, no. 6, pp. 1333-1340, 2001. doi: 10.1109/72.963769.
[41] C. Dyer, M. Ballesteros, W. Ling, A. Matthews, and N. A. Smith, "Transition-based dependency parsing with stack long short-term memory," in ACL, Beijing, 2015.
[42] A. B. Ashraf et al., "The painful face–Pain expression recognition using active appearance models," Image and Vision Computing, vol. 27, no. 12, pp. 1788-1796, 2009. doi: 10.1016/j.imavis.2009.05.007.
[43] P. Lucey et al., "Automatically detecting pain in video through facial action units," IEEE Transactions on Systems, Man, and Cybernetics, Part B, vol. 41, no. 3, pp. 664-674, 2011.
[44] G. Littlewort-Ford, M. S. Bartlett, and J. R. Movellan, "Are your eyes smiling? Detecting genuine smiles with support vector machines and Gabor wavelets," in Proceedings of the 8th Joint Symposium on Neural Computation, 2001.
[45] C. Shan, "Learning local binary patterns for gender classification on real-world face images," Pattern Recognition Letters, vol. 33, no. 4, pp. 431-437, 2012.
[46] G. Hinton, N. Srivastava, and K. Swersky, "Neural networks for machine learning lecture 6a overview of mini-batch gradient descent," 2012.
[47] P. Rodriguez et al., "Deep pain: Exploiting long short-term memory networks for facial expression classification," IEEE Transactions on Cybernetics, 2017, pp. 1-11.
[48] G. Bargshady, J. Soar, X. Zhou, R. C. Deo, F. Whittaker, and H. Wang, "A joint deep neural network model for pain recognition from face," in Proceedings of the 4th IEEE International Conference on Computer and Communication Systems (ICCS 2019), Singapore, 2019, pp. 52-56.
[49] J. Egede, M. Valstar, M. T. Torres, and D. Sharkey, "Automatic neonatal pain estimation: An acute pain in neonates database," in 2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII), 2019, pp. 1-7.
[50] Z. Chen, R. Ansari, and D. Wilkie, "Learning pain from action unit combinations: A weakly supervised approach via multiple instance learning," IEEE Transactions on Affective Computing, 2019. doi: 10.1109/TAFFC.2019.2949314.

[51] 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, 2019.
[52] D. Lopez-Martinez and R. Picard, "Continuous pain intensity estimation from autonomic signals with recurrent neural networks," in 2018 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2018, pp. 5624-5627. doi: 10.1109/EMBC.2018.8513575.
[53] P. Thiam, P. Bellmann, H. A. Kestler, and F. Schwenker, "Exploring deep physiological models for nociceptive pain recognition," Sensors, vol. 19, no. 4503, 2019. doi: 10.3390/s19204503.
[54] D. Huang et al., "Spatio-temporal pain estimation network with measuring pseudo heart rate gain," IEEE Transactions on Multimedia, vol. 24, pp. 3300-3313, 2022. doi: 10.1109/TMM.2021.3096080.
[55] P. Werner, A. Al-Hamadi, R. Niese, S. Walter, S. Gruss, and H. C. Traue, "Automatic pain recognition from video and biomedical signals," in 2014 22nd International Conference on Pattern Recognition (ICPR), 2014, pp. 4582-4587.
[56] S. Gkikas et al., "Automatic pain intensity estimation based on electrocardiogram and demographic factors," in Proceedings of the 2022 15th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC), SciTePress, 2022, pp. 155-162.
[57] S. Gkikas, C. Chatzaki, and M. Tsiknakis, "Multi-task neural networks for pain intensity estimation using electrocardiogram and demographic factors," in Information, Communication Technologies for Ageing Well and e-Health, Springer Nature, 2023, pp. 324-337.
[58] 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, 2019.
[59] P. Werner, A. Al-Hamadi, R. Niese, S. Walter, S. Gruss, and H. C. Traue, "Automatic pain recognition from video and biomedical signals," in 2014 22nd International Conference on Pattern Recognition (ICPR), 2014, pp. 4582-4587.
[60] M. Kächele et al., "Bio-visual fusion for person-independent recognition of pain intensity," in Multiple Classifier Systems, Springer International Publishing, 2015, pp. 220-230.
[61] M. Kächele, P. Thiam, M. Amirian, F. Schwenker, and G. Palm, "Methods for person-centered continuous pain intensity assessment from bio-physiological channels," IEEE Journal of Selected Topics in Signal Processing, vol. 10, no. 5, pp. 854-864, 2016. doi: 10.1109/JSTSP.2016.2535962.
[62] R. Wang, K. Xu, H. Feng, and W. Chen, "Hybrid RNN-ANN based deep physiological network for pain recognition," in Proceedings of the 2020 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2020, pp. 5584-5587.
[63] P. Thiam, H. A. Kestler, and F. Schwenker, "Multimodal deep denoising convolutional autoencoders for pain intensity classification based on physiological signals," in ICPRAM 2020 – Proceedings of the 9th International Conference on Pattern Recognition Applications and Methods, SciTePress, 2020, pp. 289-296.
[64] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," in BMVC 2015 - Proceedings of the British Machine Vision Conference 2015, British Machine Vision Association, 2015.
[65] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in Neural Information Processing Systems 25, 2012.
[66] A. Gulli and S. Pal, Deep Learning with Keras, Packt Publishing Ltd., 2017.
[67] J. Ma and Y. Yuan, "Dimension reduction of image deep feature using PCA," Journal of Visual Communication and Image Representation, vol. 63, p. 102578, 2019.
[68] J. Li, B. Zhao, H. Zhang, and J. Jiao, "Face recognition system using SVM classifier and feature extraction by PCA and LDA combination," in 2009 International Conference on Computational Intelligence and Software Engineering, 2009, pp. 1-4.
[69] Y. Sun, Y. Chen, X. Wang, and X. Tang, "Deep learning face representation by joint identification-verification," in Advances in Neural Information Processing Systems 27, 2014.
[70] Y. A. M. Alsumaidaee et al., "Review of medium-voltage switchgear fault detection in a condition-based monitoring system by using deep learning," Energies, vol. 15, no. 18, p. 6762, 2022.
[71] C. Dyer, M. Ballesteros, W. Ling, A. Matthews, and N. A. Smith, "Transition-based dependency parsing with stack long short-term memory," arXiv preprint arXiv:1505.08075, 2015.
[72] R. Wang, K. Xu, H. Feng, and W. Chen, "Hybrid RNN-ANN based deep physiological network for pain recognition," in 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2020, pp. 5584-5587.
[73] D. Lopez-Martinez and R. Picard, "Continuous pain intensity estimation from autonomic signals with recurrent neural networks," in 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2018, pp. 5624-5627.

[74] 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.
[75] P. Werner et al., "Automatic pain assessment with facial activity descriptors," IEEE Transactions on Affective Computing, vol. 8, no. 3, pp. 286-299, 2016.
[76] P. Werner et al., "Automatic pain recognition from video and biomedical signals," in 2014 22nd International Conference on Pattern Recognition, 2014, pp. 4582-4587.