

Clinical Pain Research

Marco Cascella*, Vincenzo Norman Vitale, Fabio Mariani, Manuel Iuorio and Francesco Cutugno

Development of a binary classifier model from extended facial codes toward video-based pain recognition in cancer patients

<https://doi.org/10.1515/sjpain-2023-0011>

Received January 22, 2023; accepted August 12, 2023;

published online September 5, 2023

Abstract

Objectives: The Automatic Pain Assessment (APA) relies on the exploitation of objective methods to evaluate the severity of pain and other pain-related characteristics. Facial expressions are the most investigated pain behavior features for APA. We constructed a binary classifier model for discriminating between the absence and presence of pain through video analysis.

Methods: A brief interview lasting approximately two-minute was conducted with cancer patients, and video recordings were taken during the session. The Delaware Pain Database and UNBC-McMaster Shoulder Pain dataset were used for training. A set of 17 Action Units (AUs) was adopted. For each image, the OpenFace toolkit was used to extract the considered AUs. The collected data were grouped and split into train and test sets: 80 % of the data was used as a training set and the remaining 20 % as the validation set. For continuous estimation, the entire patient video with frame prediction values of 0 (no pain) or 1 (pain), was imported into an annotator (ELAN 6.4). The developed Neural Network classifier consists of two dense layers. The first layer contains 17 nodes associated with the facial AUs extracted by OpenFace for each image. The output layer is a classification label of “pain” (1) or “no pain” (0).

Results: The classifier obtained an accuracy of ~94 % after about 400 training epochs. The Area Under the ROC curve (AUROC) value was approximately 0.98.

Conclusions: This study demonstrated that the use of a binary classifier model developed from selected AUs can be an effective tool for evaluating cancer pain. The implementation of an

APA classifier can be useful for detecting potential pain fluctuations. In the context of APA research, further investigations are necessary to refine the process and particularly to combine this data with multi-parameter analyses such as speech analysis, text analysis, and data obtained from physiological parameters.

Keywords: automatic pain assessment; pain; cancer pain; artificial intelligence; binary classifier model

Introduction

According to the International Association for the Study of Pain (IASP), chronic pain refers to persistent or recurring pain that lasts for an extended period, typically lasting for three months or more. It is a complex multidimensional experience that can have a significant impact on an individual's physical, emotional, and social well-being. This condition can severely compromise the patient's quality of life, often limiting the ability to work, sleep, and affecting social interactions with friends and family [1]. Cancer pain is a common problem among cancer patients, and its prevalence can vary depending on the stage and type of cancer, as well as the individual patient's pain tolerance and pain management strategies. The World Health Organization (WHO) reported that approximately 30–50 % of patients with cancer experience moderate to severe pain, and 70–90 % of individuals with advanced cancer experience pain [2].

A comprehensive approach to pain treatment usually involves a combination of pharmacological and non-pharmacological interventions and should be tailored to the individual needs of each patient. Careful pain assessment, therefore, is of paramount importance to guide the decision-making process pertaining to therapy [3]. Nevertheless, clinicians define treatment based on subjective pain scales and continuously refine drug dosages based on the patient's responses [4]. On the other hand, pain is a complex interplay of sensory and emotional components that is often challenging to verbally articulate [5, 6]. Furthermore, although

*Corresponding author: Marco Cascella, Department of Anesthesia and Pain Medicine, Istituto Nazionale Tumori, IRCCS – Fondazione G Pascale, Napoli, Naples, Italy, E-mail: m.cascella@istitutotumori.na.it

Vincenzo Norman Vitale, Fabio Mariani, Manuel Iuorio and Francesco Cutugno, DIETI, University of Naples “Federico II”, Naples, Italy

self-report quantitative methods such as the Numeric Rating Scale (NRS) and Visual Analog Scale (VAS) are widely regarded as the standard for pain assessment [7], they are prone to reporting bias. This bias stems from the inherent nature of self-reporting and can be influenced by various psychosocial factors, including patients' tendencies to catastrophize or underreport their pain [8].

Notably, in cancer patients, pain perception and expression are influenced by a wide range of variables, including psychosocial issues and distress, as well as the type and progression of the primary tumor [9], and pain pathophysiology [10]. Moreover, the individual's pain tolerance, communication abilities [11], coping strategies, and emotional state [12] further complicate the clinical evaluation. In this complex scenario, the limits of subjective pain assessment emerge and consequently, it is important to develop models for the study of pain focused on automatic and objective responses.

Automatic pain assessment (APA) relies on the use of objective measures to assess pain intensity and other pain-related factors [13]. These methods aim to provide a more objective measure of pain, as opposed to relying on subjective pain scales. Strategies for APA may be particularly useful in situations in which obtaining reliable self-report data is difficult, such as in children with cognitive disabilities [9], as well as patients of all ages who face communication challenges [14], including individuals with dementia or those who are non-verbal or intubated [15]. These techniques are also particularly useful for research purposes, enabling the study of pain experiences across diverse populations and settings [16].

There are different approaches to APA, including the use of physiological indicators, such as heart rate or electrodermal activity, and the application of brain imaging techniques, such as functional magnetic resonance imaging (fMRI) or electroencephalography (EEG), to measure pain-related brain activity [17]. Behavioral pain assessment is another group of APA methods. It involves observing and recording an individual's behavior due to pain experience and analyzing the data collected to identify patterns and trends. Behavioral-based approaches encompass facial expressions, qualitative (e.g., content analysis) and quantitative (e.g., computational linguistics) methods for linguistic analyses [18], and non-verbal physical indicators of pain such as body movements and gestures [19]. Current methodologies tend to incorporate multimodal techniques that integrate observable behaviors with neurophysiological elements [20].

Facial expressions are a crucial aspect of non-verbal communication, especially in conveying pain to others. Therefore, extensive research has focused on studying facial expressions as a pain behavior to enable automatic and

objective pain assessment [21, 22]. Moreover, the facial expression of pain shows consistency across ages, genders, cognitive states (e.g., noncommunicative patients), and different types of pain and may correlate with self-report of pain [23]. Facial expressions that indicate pain can be classified into voluntary and involuntary physical manifestations. The former is intentionally triggered by the cortex and can be consciously controlled in real time. However, these expressions tend to be less smooth and exhibit more variations in movement. Involuntary facial expressions, on the other hand, are unconsciously triggered by subcortical processes and are recognized by their symmetrical, synchronized, reflex-like movements of facial muscles, which give them a seamless appearance.

Action Units (AUs) are a set of specific facial muscle movements or activations that are used to describe and analyze facial expressions. The concept of AUs was developed, in 1978, by psychologists Paul Ekman and Wallace V. Friesen as part of their Facial Action Coding System (FACS), which is a comprehensive tool for describing and analyzing human facial expressions [24]. There are 46 individual AUs that are defined and categorized by the FACS. Each AU represents a specific movement or combination of movements of the facial muscles, such as raising the eyebrows, wrinkling the nose, or pursing the lips. The manual analysis of the process is often exceedingly intricate. Nevertheless, with the advent of artificial intelligence (AI) techniques, there has been a shift toward computer-mediated automatic detection of pain-related behaviors and various AI-based image-processing strategies have been implemented [25]. Consequently, the utilization of AI methodologies can offer a significant and valuable application in the field of APA, particularly when conducting assessments based on facial behaviors.

Although the growing interest in studying physiological signals to investigate APA in cancer patients [13, 20], there is insufficient research on pain behavior associated with APA in this population. On these premises, we investigated facial pain behaviors (i.e., AUs) related to APA and developed a binary classifier to distinguish between "pain" and "no-pain" using videos collected from cancer patients.

Methods

The study was performed in accordance with the Declaration of Helsinki and the patients provided written informed consent. The Medical Ethics Committee of the Istituto Nazionale Tumori, Fondazione Pascale (Napoli, Italy) approved the research (Pascale study, Pain Assessment in Cancer Patients by Machine Learning). Protocol code 41/20 Oss (26 November 2020). The study is part of a multimodal APA

project encompassing computer vision and natural language processing approaches, as well as physiological signals and clinical investigations.

Interview and video recording

A brief interview lasting approximately two-minute was conducted with oncology patients, and video recordings were taken during the session. The interview addressed demographic data, clinical information, pain features, and personal data (Table 1).

Datasets implemented

The publicly available Delaware Pain Database [26], and UNBC-McMaster Shoulder Pain dataset [27] were used for training. The Delaware Pain Database is a repository comprising photos of painful expressions. For its development, 240 individual targets (127 females and 113 males) posed with neutral facial expressions, followed by facial expressions representing their responses to five potentially painful scenarios. For each scenario, different pain levels were established. Subsequently, the set of collected images was processed by rating emotional features and implementing the FACS. Finally, the investigators selected an amount of 229 unique painful expressions.

The UNBC-McMaster Shoulder Pain dataset is a comprehensive collection of video recordings featuring 129 participants (66 females and 63 males), suffering from shoulder pain. The recordings document a series of active and passive range-of-motion tests performed by the

participants on their affected and unaffected limbs, during two distinct sessions. Throughout the recordings, the participants' facial expressions were captured and analyzed by certified FACS coders at the AU intensity for each frame. To supplement the visual data, both self-report and observer measures were taken at the sequence level. This involved capturing the participants' subjective experiences of pain and discomfort, as well as the observations made by the researchers during the testing process. Although the dataset was originally collected in the form of a video, for our purposes, we converted the dataset into images, taking one out of every 30 and tagging them with the pain label.

Based on indications from the considered datasets, the images were divided into two groups namely "pain" (marked as p in the Delaware dataset) and "no-pain" (neutral) images.

Action Units selection and analysis

For AUs analysis, OpenFace 2.0 was adopted. It is a free and open-source toolkit for facial recognition and facial landmark detection. The tool uses Deep Neural Networks (DNN) to perform facial recognition and facial landmark detection, and it can be used to build applications for tasks such as facial recognition, facial tracking, and facial expression analysis [28, 29].

The Neural Network classifier used the intensities (r) of 17 AUs that were specifically chosen. As stated in the official OpenFace documentation [28], these intensities are measured on a five-point scale with an additional point, where a score of 1 indicates minimal presence, 5 indicates maximum presence, and 0 indicates no presence at all.

According to Prkachin et al. [30], we considered the AUs related to brow lowering (AU4), orbital tightening (AU6 and AU7), nose wrinkler (AU9), upper lip raiser (AU10), and eye closure (AU43) and representing the "core" actions of pain expression. Such a core set of six AUs has been extended with 11 AUs considered in Littlewort's study on spontaneous pain expressions [31]. They included inner brow raiser (AU01), outer brow raiser (AU02), upper lid raiser (AU05), lip corner puller (AU12), dimpler (AU14), lip corner depressor (AU15), chin raiser (AU17), lip stretcher (AU20), lip tightener (AU23), lips part (AU25) and jaw drop (AU26) (Figure 1).

For each image, OpenFace extracts the considered AUs. It takes the input video and extracts AUs frame by frame. Then, the collected data were grouped, according to the associated class, and split into train and test sets, respectively 80 and 20 % of extracted data. Subsequently, 20 % of the train set was reserved for the validation set.

Neural Network classifier

This developed Neural Network classifier is based on the Shallow Network architecture. This structure usually consists of an understandable combination of two layers and has been demonstrated to be as effective as DNN on small-scale datasets [32]. The dataset used for model training and validation, which includes parameters and AUs extracted using the OpenFace toolkit, is available at reference [33].

The first layer of the classifier takes as input the 17 intensity values related to the considered AUs extracted by OpenFace, for each image, with a ReLU/sigmoid as an activation function. The last layer (output) utilizes the sigmoid activation function to generate an output of a classification label, specifically indicating pain (1) or no-pain (0) (Figure 2). The proposed network uses a binary cross-entropy loss function, the standard Adam optimizer (from the Keras library) was employed to minimize the validation loss.

Table 1: Interview "Expression of Pain" used for video recording.

Demographics
How old are you?
What is your occupation?
Are you married?
How many people are in your family?/Do you have children?/How old are your children?
Clinical information
When were you diagnosed?
What type of therapy are you currently undergoing or have undergone?
Pain features
Can you try to describe your pain?
Where in your body do you feel pain? Where is it localized?
Can you try to describe the episode in which you experienced the worst pain (when it happened, what type of pain it was, how long it lasted)?
What does it resemble? Could you describe it using a word, an adjective, a symbol, or an object?
If we could describe your pain using a color, what color would it be?
Final questions
Do you have any hobbies?
What do you enjoy doing in your free time?
What would you like to do today?
Thank you for the information provided.

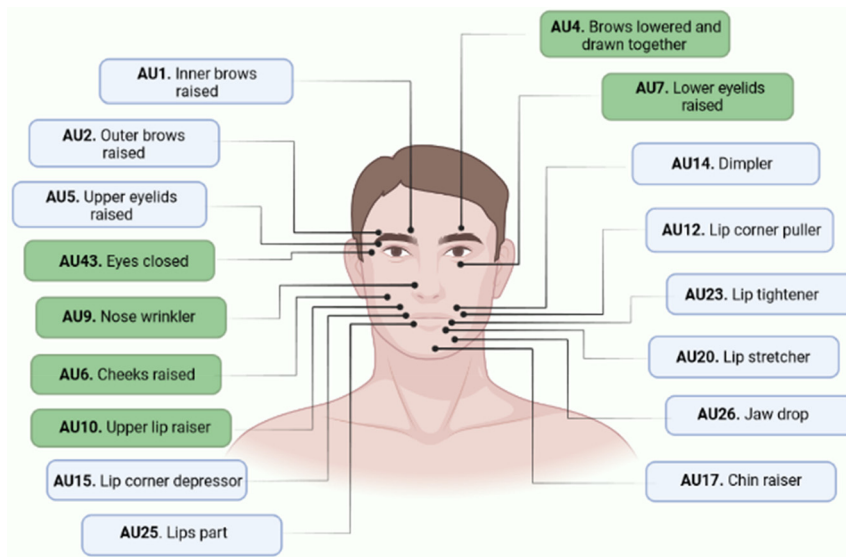


Figure 1: The action units (AUs) employed. The “core” set of six pain expressions (AU4, AU6, AU7, AU9, AU10, and AU43) [28] (green boxes) was extended with 11 selected AUs (light blue boxes), including inner brow raiser (AU1), outer brow raiser (AU2), upper lid raiser (AU5), lip corner puller (AU12), dimpler (AU14), lip corner depressor (AU15), chin raiser (AU17), lip stretcher (AU20), lip tightener (AU23), lips part (AU25), and jaw drop (AU26). Created with BioRender.com (accessed on 25 April 2023).

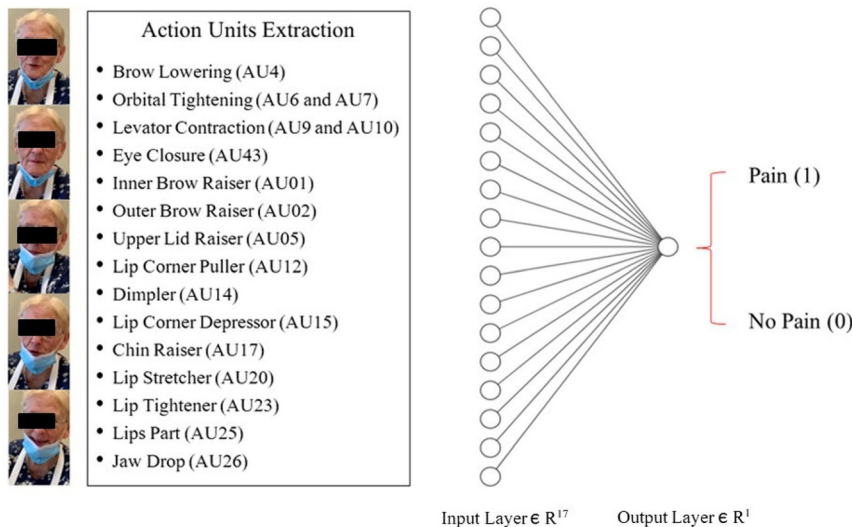


Figure 2: The binary classifier model. The classifier is made up of two dense layers. The first layer consists of 17 nodes associated with the facial Action Units (AUs) extracted by OpenFace for each image. The final layer (i.e., output layer), uses a sigmoid activation function to output a classification label of either “pain” (1) or “no-pain” (0), based on the outputs from the neurons in the previous layer. Patient consent was acquired for the study (ClinicalTrials.gov Identifier: NCT04726228) and scientific divulgation of personal identifiable information.

Video annotation

For video annotation, we used Eudico Linguistic Annotator (ELAN), version 6.4 [34]. It is a software tool for creating, editing, and analyzing multimedia data like video, audio, and text. Time-aligned annotations enable the assessment of the temporal relationship between events.

In order to perform continuous estimation, the entire patient video, along with its corresponding frame prediction values of 0 (no-pain) or 1 (pain), was imported into ELAN (Figure 3).

Dataset preparation

In relation to the technique employed for extracting frames from the videos and for training and validating the neural network, as well as the approach used to implement the dataset of images from cancer patients for testing the neural network, we have utilized the following strategy:

- Concerning the McMaster dataset, we utilized the pre-processed version, which was already divided into frames, comprising a total

of 48,398 images. Each subject appeared in only one of the three phrases, and we made sure to maintain a balanced ratio of males and females in each phase.

- As for the Delaware dataset, we followed the same approach while also preserving the proportion of different ethnicities.
- Lastly, for the data collected by our team, we recorded high-definition videos at 30 fps and extracted the set of AUs for each frame.

We partitioned the data while carefully preserving an equal proportion of males/females, Americans/Latinos/African Americans, etc., ensuring that each subject appeared in only one of the three phases and that every category was represented in each phase.

Model evaluation

All samples were subjected to OpenFace processing and converted into AU intensities. Subsequently, the processed datasets consisting of AUs were combined, and the sets were split using the standard shuffling

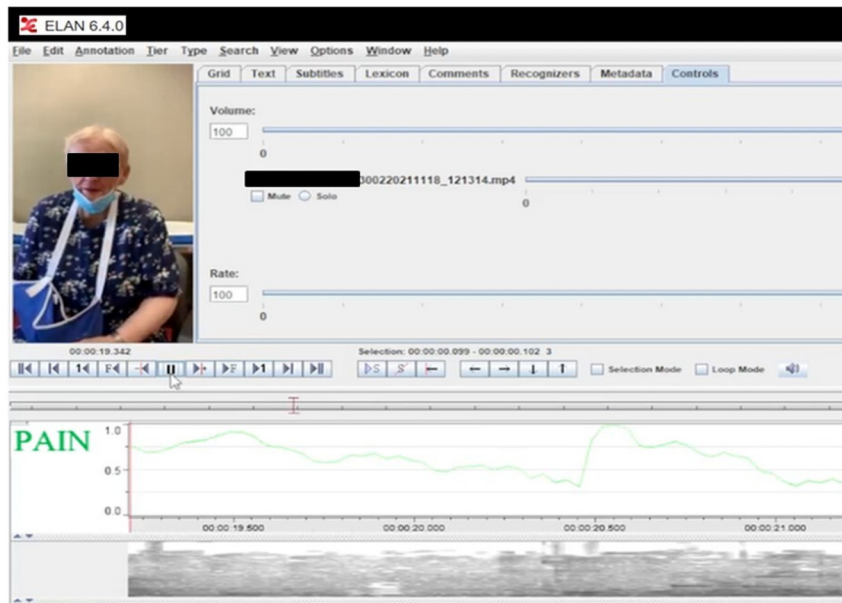


Figure 3: Frame prediction values of “pain” ranging from 0 to 1 (red line). The annotation tool for audio and video recordings ELAN 6.4 was implemented. Patient consent was acquired for the study (ClinicalTrials.gov Identifier: NCT04726228) and scientific divulgation of personal identifiable information.

feature provided by the Keras framework. The classifier was trained with a maximum of 10,000 epochs, implementing an early stopping policy based on validation loss with patience of 200 epochs. This means that if the validation loss does not decrease within 200 epochs, the training process will stop, and the best model achieved will be retained.

The Receiver Operating Characteristic (ROC) curve is a graphical representation used to assess the performance of a binary classification model. It illustrates the relationship between the true positive rate (sensitivity) and the false positive rate (1-specificity) at various threshold settings for the classification model. The Area Under the ROC Curve (AUROC) is a metric derived from the ROC curve that quantifies the overall performance of the classification model [35]. It represents the probability that the model will rank a randomly selected positive instance higher than a randomly selected negative instance. The AUROC value ranges between 0 and 1, with higher values indicating better performance. An AUROC value of 0.5 indicates a model that performs no better than random guessing, while a value of 1 represents a perfect classifier. In our model, it measures how good the classifier is in distinguishing between pain and no-pain.

Accuracy is another metric used to measure the overall correctness of a classification model. It calculates the proportion of correctly classified instances out of the total number of instances. While it provides a simple assessment of performance, it may not be sufficient in cases of imbalanced datasets. Therefore, other metrics like precision (quantifies the model’s accuracy in identifying true positives while minimizing false positives), and recall (sensitivity or true positive rate), may be necessary for a more comprehensive evaluation. The choice of evaluation metrics depends on the specific characteristics of the classification problem [36].

Binary cross-entropy is a measure of error commonly used in binary classification tasks. It quantifies the difference between predicted probabilities and true binary labels. The goal is to minimize this error during model training to improve classification accuracy.

Results

The binary classifier obtained an accuracy of 0.9448 after about 400 training epochs. In addition to the achieved accuracy, the AUROC value of ~0.98 (Figure 4).

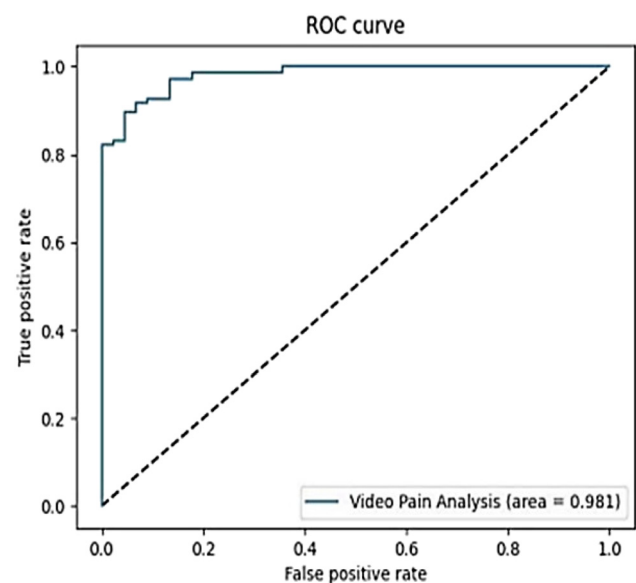


Figure 4: The area under the receiver operating characteristic (ROC) curve (AUC) of the considered model.

The measure of the error (binary cross-entropy) on the validation set is 0.1581. Precision and recall measures are 0.9528 and 0.968, respectively. The performance of the classifier is shown in Table 2.

Table 2: Performance of the binary classifier model.

	Validation loss	Accuracy	Precision	Recall
Performance	0.1581	0.9448	0.9528	0.968

Discussion

In the field of APA, the study of facial expressions is of paramount importance. We used two datasets to develop a binary classifier capable of distinguishing between the categories “pain=1” and “non-pain=0”. The complete patient video, together with its associated frame prediction values (0 and 1), was loaded into ELAN. This process allowed for the analysis of the video in greater detail, as the frame prediction values provided a quantitative measure of the pain levels present in each individual frame of the video. This enabled a more accurate and precise estimation of the overall pain levels experienced by the patient, as well as the specific moments and duration of pain throughout the video. To our knowledge, this is the first study conducted for the APA using AUs in a setting of cancer patients. The systems, originally trained using publicly available datasets with an expanded set of AUs to better capture pain expressions, have ultimately been tested on actual cancer pain patients in real-world scenarios.

Compared to other classifiers based on the Shallow network architecture [34], the proposed model employed meaningful features rather than extracting them through a feature processing layer, which is supposed to be trained along with the classification layer. The model’s validation revealed a high level of accuracy (~94 %). This is not surprising, as error rates tend to decrease on smaller datasets when using shallow networks [37].

Other experiments were conducted with the aim of extracting pain characteristics from videos through the analysis of AUs. For example, Lucey et al. [38] used the UNBC-McMaster Shoulder Pain database and implemented an active appearance model for automatically detecting the frames in which a patient is in pain. They were able to extract visual features and derived 3D parameters. The features were utilized to classify individual AUs through a linear support vector machine (SVM). To merge the SVMs’ outputs for the AUs, linear logistic regression was employed, allowing for a supervised calibration of the scores into log-likelihood scores. The same authors also worked on AUs and performed a frame-by-frame analysis. A combination of a convolutional neural network (CNN) and recurrent neural network (RNN) was utilized to analyze the spatial characteristics and dynamic changes in video frames. The CNN was responsible for learning spatial features from each frame, while the RNN captured the temporal dynamics between consecutive frames. Following this analysis, a regression model was employed to estimate pain intensity scores for each video [39].

The FACS is a comprehensive system for describing all observable facial movements resulting from the activation

of individual muscles. Despite FACS, by its nature, does not provide descriptors specific to emotions, the system is widely used in the field of psychology as a means of studying and exploring different aspects related to emotions as well as other human states [40]. Nevertheless, the specifics of the codes (e.g., number and types of AUs, namely “pain relevant AUs”) and methods (e.g., automated vs. semi-automated systems) used for pain research have yet to be determined [41]. Interestingly, Skiendziel et al. [42] utilized FaceReader 7, an automated emotion and AUs coding system, to analyze a dataset of standardized facial expressions of six basic emotions. In addition to the clinical setting, another peculiarity of our study is the use of an extended combination of AUs. The original “core” of six pain expressions [28] was expanded to obtain a higher accuracy of the data (i.e., a higher number of nodes of the input layer of the classifier). Additionally, the capability to analyze a video frame-by-frame can enable the detection of pain fluctuations. This aspect is crucial in cancer pain assessment and management, for example for addressing specific cancer pain phenomena such as breakthrough cancer pain [43].

Research in the field of APA for the management of cancer pain must address several issues. Probably, a key aspect is multimodal data collection. In this regard, the combination of data obtained from the analysis of language and text, as well as physiological data, will allow for the development of more accurate predictive models for the diagnosis of various components of pain in cancer patients [44]. This is because pain is a multidimensional experience, and it is influenced by not only physiological factors but also psychological and emotional elements [45]. By integrating these different types of data, more targeted and effective treatment plans could be designed [46].

Study limitations

The study in question has several limitations that should be acknowledged. The major limitation pertains to the datasets utilized in the research, which were not explicitly intended for studying cancer pain. This aspect could possibly have an impact on the precision and validity of the outcomes. For example, the datasets implemented in this study may not have included all the necessary variables to thoroughly analyze and understand cancer pain phenomena. Moreover, the inclusion criteria and sampling techniques for these datasets may not have been optimized for cancer pain research, which could have led to some bias or inaccuracies in the data. Therefore, caution should be exercised when drawing conclusions from the results, as the limitations of the datasets utilized could potentially affect the generalizability

and applicability of the findings. Future research endeavors may need to focus on collecting more targeted and comprehensive datasets to overcome this limitation and improve the accuracy and reliability of the results.

Additionally, the model used in the study is limited in its ability to conduct temporal analysis as it focuses on a single time instant not considering contextual parameters. These limitations suggest that a more complex classifier would be needed in order to improve the accuracy of the study's findings and address these gaps in the analysis. In order to further validate the model and increase the generalizability of the results, it is important to test the model on a larger sample of patients, including a control group. This will allow for a more robust evaluation of the model's ability to accurately predict and quantify cancer pain. Moreover, it is also important to compare the results obtained from the model with clinical measurements, specifically pain scales, and psychometric measures. This will enable a more comprehensive evaluation of the model's precision. Therefore, the next stage of the research would be to conduct a clinical study to validate these findings.

Conclusions

Despite APA methods are not a substitute for subjective pain assessments and should be used in conjunction with other methods to provide a more comprehensive understanding of an individual's pain experience, the study of facial behavior using automatic systems offers important perspectives. The proposed Shallow Network classifier combined with the augmented set of AUs demonstrated to be a fast and accurate method to identify pain based on facial expressions. More research is needed to improve the process and especially to integrate this data with multi-parameter analyses such as speech analysis, text analysis, and those obtained from physiological parameters.

Research ethics: The Medical Ethics Committee of the Istituto Nazionale Tumori, Fondazione Pascale (Napoli, Italy) approved the research (Protocol code 41/20 Oss, 26 November 2020). The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Informed consent: Informed consent was obtained from all individuals included in this study, or their legal guardians or wards.

Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Competing interests: The authors state no conflict of interest.

Research funding: None declared.

Data availability: The raw data can be obtained on request from the corresponding author.

References

1. Treede RD, Rief W, Barke A, Aziz Q, Bennett MI, Benoliel R, et al. Chronic pain as a symptom or a disease: the IASP Classification of Chronic Pain for the International Classification of Diseases (ICD-11). *Pain* 2019;160:19–27.
2. World Health Organization (WHO). Cancer pain management; 2018. <https://www.who.int/cancer/palliative/painladder/en/> [Accessed 1 Jan 2023].
3. Shkodra M, Brunelli C, Zecca E, Infante G, Miceli R, Caputo M, et al. Cancer pain: results of a prospective study on prognostic indicators of pain intensity including pain syndromes assessment. *Palliat Med* 2022;36:1396–407.
4. Brunelli C, Borreani C, Caraceni A, Roli A, Bellazzi M, Lombi L, et al. PATIENT VOICES study group. PATIENT VOICES, a project for the integration of the systematic assessment of patient reported outcomes and experiences within a comprehensive cancer center: a protocol for a mixed method feasibility study. *Health Qual Life Outcome* 2020;18:252.
5. Haroun R, Wood JN, Sikandar S. Mechanisms of cancer pain. *Front Pain Res (Lausanne)* 2023;3:1030899.
6. Boselie JJLM, Peters ML. Shifting the perspective: how positive thinking can help diminish the negative effects of pain. *Scand J Pain* 2023;23:452–63.
7. Chien CW, Bagraith KS, Khan A, Deen M, Strong J. Comparative responsiveness of verbal and numerical rating scales to measure pain intensity in patients with chronic pain. *J Pain* 2013;14:1653–62.
8. Kristiansen FL, Olesen AE, Brock C, Gazerani P, Petrini L, Mogil JS, et al. The role of pain catastrophizing in experimental pain perception. *Pain Pract* 2014;14:E136–45.
9. Caraceni A, Shkodra M. Cancer pain assessment and classification. *Cancers (Basel)* 2019;11:510.
10. Cascella M, Muzio MR, Monaco F, Nocerino D, Ottaiano A, Perri F, et al. Pathophysiology of nociception and rare genetic disorders with increased pain threshold or pain insensitivity. *Pathophysiology* 2022;29:435–52.
11. Cascella M, Bimonte S, Saettini F, Muzio MR. The challenge of pain assessment in children with cognitive disabilities: features and clinical applicability of different observational tools. *J Paediatr Child Health* 2019;55:129–35.
12. Yessick LR, Tanguay J, Gandhi W, Harrison R, Dinu R, Chakrabarti B, et al. Investigating the relationship between pain indicators and observers' judgements of pain. *Eur J Pain* 2023;27:223–33.
13. Moscato S, Orlandi S, Giannelli A, Ostan R, Chiari L. Automatic pain assessment on cancer patients using physiological signals recorded in real-world contexts. *Annu Int Conf IEEE Eng Med Biol Soc* 2022;2022:1931–4.
14. Wang J, Cheng Z, Kim Y, Yu F, Heffner KL, Quiñones-Cordero MM, et al. Pain and the Alzheimer's disease and related dementia spectrum in community-dwelling older Americans: a nationally representative study. *J Pain Symptom Manage* 2022;63:654–64.
15. Deldar K, Froutan R, Ebadi A. Challenges faced by nurses in using pain assessment scale in patients unable to communicate: a qualitative study. *BMC Nurs* 2018;17:11.
16. Pouroumran F, Radhakrishnan S, Kamarthi S. Exploration of physiological sensors, features, and machine learning models for pain intensity estimation. *PLoS One* 2021;16:e0254108.

17. Misra G, Wang WE, Archer DB, Roy A, Coombes SA. Automated classification of pain perception using high-density electroencephalography data. *J Neurophysiol* 2017;117:786–95.
18. Voytovich L, Greenberg C. Natural language processing: practical applications in medicine and investigation of contextual autocomplete. *Acta Neurochir Suppl* 2022;134:207–14.
19. Walsh J, Eccleston C, Keogh E. Pain communication through body posture: the development and validation of a stimulus set. *Pain* 2014;155:2282–90.
20. Gkikas S, Tsiknakis M. Automatic assessment of pain based on deep learning methods: a systematic review. *Comput Methods Progr Biomed* 2023;231:107365.
21. Hassan T, Seus D, Wollenberg J, Weitz K, Kunz M, Lautenbacher S, et al. Automatic detection of pain from facial expressions: a survey. *IEEE Trans Pattern Anal Mach Intell* 2021;43:1815–31.
22. Werner P, Lopez-Martinez D, Walter A-HA, Gruss S, Picard R. Automatic recognition methods supporting pain assessment: a survey. *IEEE Trans Affect Comput* 2019;13:530–52.
23. Chambers CT, Mogil JS. Ontogeny and phylogeny of facial expression of pain. *Pain* 2015;156:798–9.
24. Ekman P, Friesen WV. Facial action coding system: a technique for the measurement of facial movement. Edina, MN, USA: Consulting Psychologists Press; 1978.
25. Tian Y. Artificial intelligence image recognition method based on convolutional neural network algorithm. *IEEE Access* 2020;8:125731–44.
26. Mende-Siedlecki P, Qu-Lee J, Lin J, Drain A, Goharзад A. The Delaware Pain Database: a set of painful expressions and corresponding norming data. *Pain Rep* 2020;5:e853.
27. Lucey P, Cohn JF, Prkachin KM, Solomon PE, Matthews I. Painful data: the UNBC-McMaster shoulder pain expression archive database. In: 2011 IEEE International Conference on Automatic Face & Gesture Recognition (FG); 2011:57–64 pp.
28. OpenFace. <https://cmusatyalab.github.io/openface/> [Accessed 6 Jan 2023].
29. Baltrušaitis T, Robinson P, Morency LP. OpenFace: an open source facial behavior analysis toolkit. In: 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Placid, NY, USA; 2016:1–10 pp.
30. Prkachin KM, Solomon PE. The structure, reliability and validity of pain expression: evidence from patients with shoulder pain. *Pain* 2008;139:267–74.
31. Littlewort GC, Bartlett MC, Lee K. Faces of pain: automated measurement of spontaneous all facial expressions of genuine and posed pain. In: Proceedings of the 9th International Conference on Multimodal Interfaces, ICMI 2007, Nagoya, Aichi, Japan, November 12–15; 2007.
32. Schindler A, Lidy T, Rauber A. Comparing shallow versus deep neural network architectures for automatic music genre classification. In: Aigner W, Schmiedl G, Blumenstein K, Zeppelzauer M, editors. Proceedings of the 9th forum media technology 2016. Austria: St. Polten University of Applied Sciences, Institute of Creative\Media\ Technologies; 2016. Available from: <http://ceur-ws.org>.
33. Cascella M. Dataset for binary classifier. Pain 2023. <https://doi.org/10.5281/zenodo.7557362>.
34. ELAN. <https://archive.mpi.nl/tla/elan/download> [Accessed 21 Jan 2023].
35. Bellini V, Cascella M, Cutugno F, Russo M, Lanza R, Compagnone C, et al. Understanding basic principles of Artificial Intelligence: a practical guide for intensivists. *Acta Biomed* 2022;93:e2022297.
36. Greener JG, Kandathil SM, Moffat L, Jones DT. A guide to machine learning for biologists. *Nat Rev Mol Cell Biol* 2022;23:40–55.
37. McDonnell MD, Vladusich T. Enhanced image classification with a fast-learning shallow convolutional neural network. In: 2015 International Joint Conference on Neural Networks (IJCNN). IEEE; 2015.
38. Lucey P, Cohn JF, Matthews I, Lucey S, Sridharan S, Howlett J, et al. Automatically detecting pain in video through facial action units. *IEEE Trans Syst Man Cybern B Cybern* 2011;41:664–74.
39. Lucey P, Cohn JF, Prkachin KM, Solomon PE, Chew S, Matthews I. Painful monitoring: automatic pain monitoring using the UNBC-McMaster shoulder pain expression archive database. *Image Vis Comput* 2012;30:197–205.
40. Clark EA, Kessinger J, Duncan SE, Bell MA, Lahne J, Gallagher DL, et al. The facial action coding system for characterization of human affective response to consumer product-based stimuli: a systematic review. *Front Psychol* 2020;11:920.
41. Lautenbacher S, Hassan T, Seuss D, Loy FW, JU G, Schmid U, et al. Automatic coding of facial expressions of pain: are we there yet? *Pain Res Manag* 2022;2022:6635496.
42. Skiendziel T, Rösch AG, Schultheiss OC. Assessing the convergent validity between the automated emotion recognition software Noldus FaceReader 7 and facial action coding system scoring. *PLoS One* 2019;14:e0223905.
43. Cuomo A, Cascella M, Forte CA, Bimonte S, Esposito G, De Santis S, et al. Careful breakthrough cancer pain treatment through rapid-onset transmucosal fentanyl improves the quality of life in cancer patients: results from the BEST multicenter study. *J Clin Med* 2020;9:1003.
44. Gehr NL, Bennedsgaard K, Ventzel L, Finnerup NB. Assessing pain after cancer treatment. *Scand J Pain* 2022;22:676–8.
45. Nummenmaa L. Mapping emotions on the body. *Scand J Pain* 2022;22:667–9.
46. Bäckryd E. Pain assessment 3 × 3: a clinical reasoning framework for healthcare professionals. *Scand J Pain* 2023;23:268–72.