Pain fingerprinting using multimodal sensing: pilot study

Anja Keskinarkaus[1] · Ruijing Yang[1,2] · Angelos Fylakis[1] · Md. Surat-E-Mostafa[1] ·
Arto Hautala[1,3] · Yong Hu[4] · Jinye Peng[2] · Guoying Zhao[1] · Tapio Seppänen[1] ·
Jaro Karppinen[5]

Abstract

Pain is a complex phenomenon, the experience of which varies widely across individuals. At worst, chronic pain can lead to anxiety and depression. Cost-effective strategies are urgently needed to improve the treatment of pain, and thus we propose a novel home-based pain measurement system for the longitudinal monitoring of pain experience and variation in different patients with chronic low back pain. The autonomous nervous system and audio-visual features are analyzed from heart rate signals, voice characteristics and facial expressions using a unique measurement protocol. Self-reporting is utilized for the follow-up of changes in pain intensity, induced by well-designed physical maneuvers, and for studying the consecutive trends in pain. We describe the study protocol, including hospital measurements and questionnaires and the implementation of the home measurement devices. We also present different methods for analyzing the multimodal data: electroencephalography, audio, video and heart rate. Our intention is to provide new insights using technical methodologies that will be beneficial in the future not only for patients with low back pain but also patients suffering from any chronic pain.

Keywords Low back pain · Machine learning · Facial expression · Audio analysis · Heart rate · Electroencephalography

1 Introduction

In recent years, the understanding of the role of the brain in pain processing has increased due to non-invasive brain imaging methodologies. Thus, tremendous efforts have been made to understand the cellular and molecular basis of chronic pain [4, 9, 14]. Current evidence shows no clear relationship between the amount of tissue damage and the degree of discomfort or functional disability [4]. Furthermore, for any individual, the pain experience varies across different experiments, and even within a single experiment depending upon the environment, experimenter, instructions, stimulus, and procedural design [1], which makes the problem even more challenging. However, pain intensity, pain-related disability,

✉ Anja Keskinarkaus
    anja.keskinarkaus@oulu.fi

Extended author information available on the last page of the article

pain duration and pain effects are the aspects that define pain and its influence on quality of life [12]. Meanwhile, pain monitoring of patients in intensive care (ICU), or home care is mainly a manual check by a clinician to make any required adjustments to medication or treatment [10]. This is a huge workload for clinicians and the consistency and reliability are cannot be guaranteed. Therefore, determining the mechanism of pain and designing a system that could automatically monitor pain to reduce the heavy workload of clinicians is paramount, to provide them with a point of reference for accurate treatment, and to further improve people's quality of life. Although it is possible to identify neural activity that ordinarily causes pain, there is no direct or objective way to measure pain. Awareness of pain is a perception and, therefore, subjective [9, 22]. Well-defined instruments for assessing pain and pain-related variables are ultimately based on self-reports, observations, or both [4, 12, 22]. One of the golden standards for assessing or monitoring pain is through self-reporting, when a patient is asked the intensity of pain using a 10-cm visual analog scale (VAS).

According to a review on emotion recognition from physiological signals, emotion arises spontaneously rather than through conscious effort and is often accompanied by physical and physiological changes that are relevant to the human organs and tissues such as the brain, heart, skin, blood flow, muscles, facial expressions and voice [24]. Thus, pain is a complex sensory and emotional experience and has a potential relationship with these physiological changes.

In [31], multimodal data, with bio-potentials and video recordings, were collected from pain patients. Heat stimulation was utilized to induce pain using a thermode attached to the arm. The stimulation temperatures were adjusted on the basis of the participant-specific pain threshold and pain tolerance was divided into four levels. Data were collected from 90 healthy adults. The database is available for non-commercial research use. Lucey et al. [18] describe methods for collecting data and extracting facial features using two cameras. The participants with shoulder pain were self-identified and recruited from physiotherapy clinics and through advertisements. The participants underwent a series of range-of-motion tests. In all, 200 videos with facial expressions of 129 (63 males, 66 females) participants were collected. A subset of 25 participants has been made publicly available from the original data. The dataset features annotation with self-reported and observed measures of pain intensity at video level and facial action coding at frame level.

In [34] a data acquisition recording system setup is described for collecting color video of the face, 3D and thermal video, and physiological signals. The designed protocol includes social interviews, film watching, physical experience, and controlled activities. Pain was stimulated once per person by a cold pressor task. The study included 140 healthy volunteers. In addition to the data acquisition system and protocol, some initial analyses to validate the data are also presented.

Velana et al. [30] describe a method for collecting biopotentials, camera images of the facial region, and audio signals. The data were recorded among healthy adults, when heat stimulation was applied to elicit pain. Emotional states were elicited using image and sound stimuli. For each participant, 30 min of multimodal sensory data were recorded. In addition to data collection, the aim of the study was to detect patterns of heat pain intensities under the influence of emotional stimuli. However, the work concentrated on the design of the experiment and procedure of the experiments. Aung et al. [2] explain in detail the patient recruitment process and the actual trial for collecting an emotion-related database. Recording devices are explained in detail, with a configuration of eight high resolution cameras, and electromyographic sensors for motion tracking attachments. The recorded data was initially analyzed, but the main focus was on the possibility of automatically recognizing the facial expression of pain.

In [13] the procedure of data collection, experimental protocol and the database structure for collecting electromyographic, RGB, depth and thermal (RGBDT) imagery is described. Twenty healthy participants were used in the experiments, with pain elicited by electrical stimulation to the participants' muscles. Visual modalities, RGB, depth and thermal (RGBDT) imagery, and their fusion were considered in the evaluation. Furthermore, Gruss et al. [11] describe a protocol to elicit pain and simultaneously record physiological responses (electrocardiogram, electromyography, skin conductance level) as well as video and audio data. A total of 134 healthy adults underwent experimental pain stimulation with heat and electrical stimuli. Participant recruitment and selection, preparations for the pain elicitation experiment, and details of the actual data acquisition system are described in detail. In [7], data were collected for a 30-min period each day from six healthy participants who came for measurements every day for one week. Electrical stimulation was used to induce pain and three-channel biosensors were used to obtain blood volume pulse, electrocardiogram and skin conductance from the study participants.

Considering the different modalities, relations between facial expressions and pain have attracted considerable interest in the research community. In [15], a review evaluated the most pain-related action units across clinical and experimental settings and demonstrated a consistent subset of action units (AU) that emerged during pain, which consisted of lowering brows (AU4), cheek raising/lid tightening (AU6_7), nose wrinkling/raising the upper lip (AUs9_10), and opening the mouth (AUs25_26_27). In [16], individualized maneuvers were used to exacerbate clinical pain in patients with chronic low back pain, thereby experimentally producing different levels of pain. Machine-learning models were built from central and autonomic parameters (heart rate variability) collected before and after pain exacerbation. As a result, within-patient (participant-specific) relatively lower and higher clinical pain intensity states were classified. The UNBC-McMaster database [18], with facial images with rich annotation, has provided the premises for many of the related research [32].

Since the release of the BioVid database [31], contact-based sensor approaches have gained more attention in pain-related research [32] Voice has also been deployed as a cue for pain analysis, but has so far received less attention. In [21], voice recordings of patients, made during an interview in a medical center, were used to separate significant pain from non-significant pain using machine learning methods. Tsai et al. [29] and Li et al. [17] analyzed audio signals recorded during clinical interviews in an emergency triage situation. A Triage Pain-Level Multimodal database was collected and used in the study. In [27], the authors suggest a combination of three distinctive modalities (audio, video, physiology) for the recognition of artificially induced pain intensities.

A literature review revealed that experimental data utilized in pain-related research are mostly collected in a laboratory setting and generally long-term data are not available or utilized. Recording data in real environment is challenging, as is the lack of long-term monitoring data. Several recent studies have raised this issue [19, 27, 32].

In this study, we developed new technology by combining existing technology in a pilot study in Finland. The developed technology enables pain monitoring (smartphone/laptop/tablet applications) in both clinical and home settings, which is novel in the sense that most previous studies have been conducted in controlled laboratory environments. We describe the experimental setup and protocol, including hospital/home measurements, in which changes in pain intensity are induced by well-designed physical maneuvers for monitoring pain experience and variation among patients with chronic low back pain. The proposed technology will help collect genuine, spontaneous pain-related emotion and further be beneficial for affective data collection and related research. The developed software will make

it possible to label the emotions of the participants at the time of recording, also enabling long-term data collection.

There is a need to better understand the problem of pain recognition. Furthermore, to build a recognition model, it is important to extract the most suitable features. Tools and methods are developed and utilized for the collected pilot data, to preprocess, discriminate features and analyze the data. We have taken advantage of an extensive audio software package, developed in our team to analyze prosodic features from audio data, and also implemented tools for electroencephalography (EEG) and heart rate analyses. For videos, a machine learning system for facial expression analysis has been implemented to estimate pain. Furthermore, correlation analysis, for EEG data, audio, video and Autonomic Nervous System (ANS) parameters against reported pain intensities have revealed important relations, which will be discussed.

A cost-effective home-based pain-monitoring device would be effective in daily clinical practice not only among low back pain patients but also among patients suffering from any chronic pain. We present the results of our preliminary studies and discuss the potentials of the developed methods and techniques, as well as points for future development.

2 Pilot study

2.1 Materials

This study is part of the large "Pain Fingerprint using Multimodal Sensing" (PASE) study, conducted in the Center for Life Course Health Research and the Center for Machine Vision and Signal Analysis at the University of Oulu, Finland. The Ethical Committee of the Northern Ostrobothnia Hospital District, Oulu, Finland approved the PASE study protocol, and all voluntary patients and healthy participants gave their informed consent. Confidentiality, data management and storage is governed by the terms of the informed consent. The study population consisted of 14 participants (aged 32 to 53 years, 7 men and 7 women). Measurements consisted of measurements carried out in the Oulu University Hospital and at participants' homes.

2.2 Study design overview

The overall procedure of 1–2-month follow-up of participants' measurements, illustrated in Fig. 1, was constructed of three phases. In Phase 1, the participants were instructed to fill in a questionnaire before their first visit to the hospital. The questionnaires covered health-related questions on background, exercise and sleeping habits, intensities of low back pain and possible radiating leg pain, pain drawing, and the short version of the Örebro Musculoskeletal Pain Screening Questionnaire, questionnaire for physical activity, STarT Back Tool and the Oswestry Disability Index (ODI) [8]. At the hospital, EEG signals were measured and each participant was instructed in how to use the developed home measurement device. For each participant, the first measurement using the software was made with the nurse's guidance at the hospital.

In home measurements, we applied new technologies to report pain experience with the use of a tablet application. During home measurements (Phase 2), the participants used the application to report weekly subjective pain intensity and audio-visual and heart rate data. The intensity of pain was measured using 10-cm VAS [3], where '0' represents no pain and
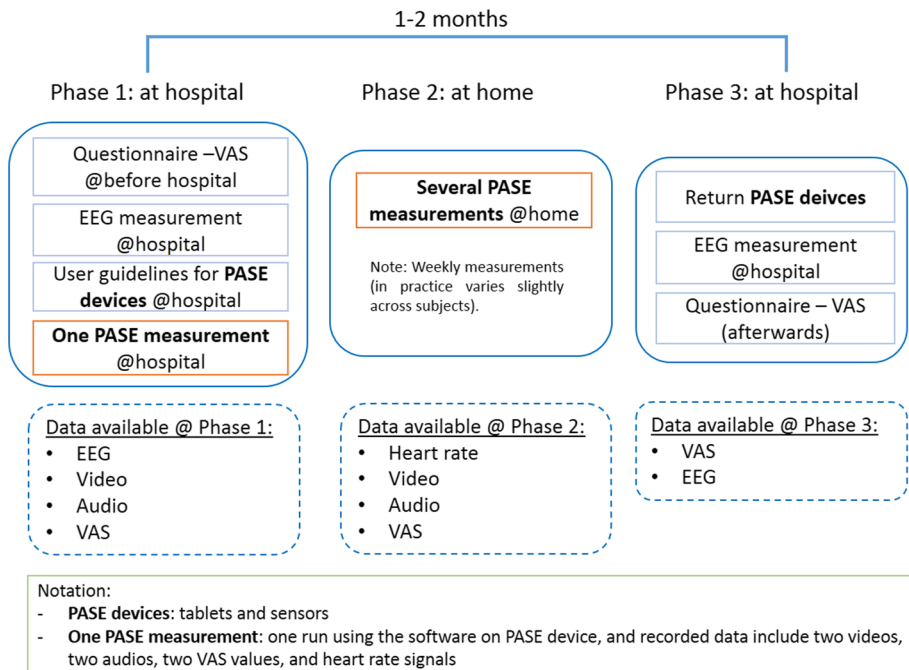
Fig. 1 Overview of participant study and measurements. VAS = Visual Analog Scale for pain intensity assessment

'10' stands for the worst imaginable pain. In Phase 3, after the 1–2-month follow-up, EEG was recorded again along with pain intensity and disability (ODI) assessments. The home measurement protocol and devices were defined as illustrated in Fig. 2.
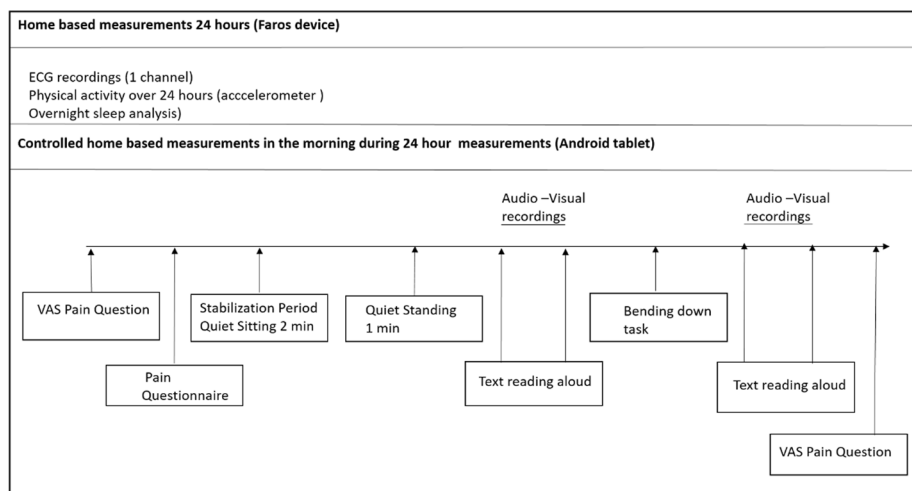

Fig. 2 Home measurement protocol

2.3 Home measurements

Pain experience was monitored using an Android application. The application runs on 7'
Lenovo Tab3 7 tablets (Android 6.0, Mediatek MT8161 chipset, 1 GB RAM, 7'' display, 5
MP / 2 MP, frontal / rear cameras). The criteria for choosing the devices were the follow-
ing: a relatively compact device that can be easily distributed to the users; and a frontal
camera of 2 MP for fair resolution video-recording for the home measurement protocol.
Android was selected as the operating system since it is highly compatible with different
devices. Bluetooth connectivity also enables future development and the possibility of
testing other external measuring devices. The application collected data from pain-related
questions requiring user input combined with audio-visual recording. The PASE home
measurement Graphical User Interface (GUI) is illustrated in Fig. 3b).

The answers to the questions were recorded in the database files together with the times-
tamps indicating when each answer was given or when the videos were recorded. The vid-
eos were recorded in MPEG-4 format.

Patients were given general instructions for the use of devices. These included general
instructions for using the devices and instructions for charging. The home measurement
devices are illustrated in Fig. 3a). Weekly measurement started by attaching the Bittium
Faros device (Bittium Corporation, Oulu, Finland) to the breast area and pressing the start
button before starting the home measurement software on the tablet. Heart rate was meas-
ured using the Faros device for 24 h once a week. The tablet was placed in an upright
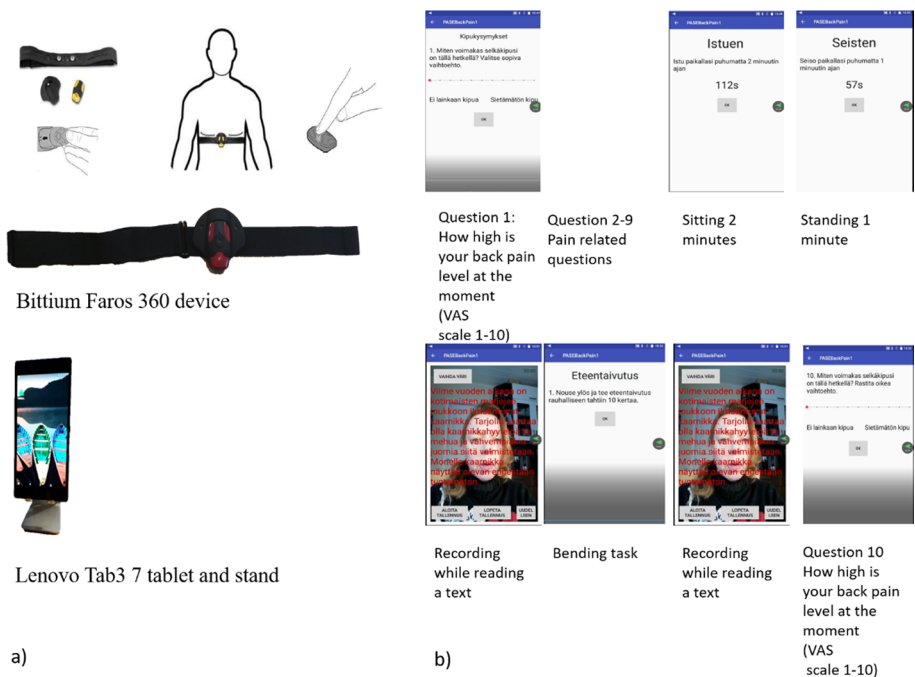

Fig. 3 a) PASE home measurement devices b) PASE measurement GUI. Question 1: Rate your current
back pain level (VAS scale 0–10). Questions 2–9: pain-related questions. Sitting two minutes. Sitting one
minute. Question 10: Rate your current back pain level (VAS scale 0–10)

position to keep it stable during measurements, about 50–100 cm from the face, and with back lightning. Each home measurement followed a procedure and was guided by the software:

1   The participant is asked to indicate their current pain intensity level (Question 1), by selecting the discrete value using the touch screen (VAS 0–10)
2   To gather a more comprehensive picture of the participant's recent condition, the following questions are asked:

- Have you had low back pain during the last week?
- What was the intensity of your low back pain during last week?
- What was your general level of discomfort at work?
- What was your general level of discomfort in your social life?
- What was your general level of discomfort when doing housework?
- Do you have pain radiating from your back to your leg?
- What is the intensity of your leg pain?

3   The participant is instructed to sit for one minute and stand for two minutes in order to stabilize the measurement situation.
4   Video 1 is recorded. The participant performs a reading task – the text is shown on the tablet screen. This promotes the collection of a facial expressions database, with audio data recorded simultaneously. The text selected was emotionally neutral (see, 3.1).
5   In order to induce pain, the participant is asked to do a series of bending exercises.
6   After the body movements, video 2 is recorded, and the participant reads the same text again from the screen.
7   To assess the participant's pain intensity, Question 10 (identical to Question 1) is asked.

The above seven-step process was repeated with each participant weekly, for 1–2 months in total.

The questionnaire responses were stored in an SQL database. Figure 4 illustrates a sample table from the database, which shows the responses under the "value" field. The timestamps were recorded upon entering each one of the measurement sections. The format of timestamps that we used was Unix Time, in milliseconds. Timestamps make it easier to

Fig. 4 Database structure

index	timestamp	section	value
1	1485500520110	start	
2	1485500529811	Question1	7
3	1485500535843	Question2	Kyllä
4	1485500548957	Question3	9
5	1485500561768	Question4	9
6	1485500568531	Question5	9
7	1485500573771	Question6	9
8	1485500579529	Question7	9
9	1485500584554	Question8	Kyllä
10	1485500595796	Question9	7
11	1485500858261	Video1	video_1485500812068.mp4
12	1485500980420	Video2	video_1485500928089.mp4
13	1485500992501	Question10	8

associate the heart rate recording with the respective section of the session during which they were recorded.

3 Analysis methods

Audio-visual reporting protocols with advanced affective computing methods were used in the analyses. The various affective parameters were computed from the expressions of the face and from speech by using a large software package developed in earlier projects by the team. Data-processing methods were also developed for analyzing facial expressions, voice, heart rate and EEG. Each modality was characterized by specific properties, which are assumed to provide valuable and distinctive insights into the level of pain. The input data were processed in order to find and extract multiple types of features. Prior to the extraction of descriptors from each of the recorded modalities, an individual pre-processing step was undertaken. In the audio analysis, we extracted features based on ideas that have proved successful in emotion domains but are not specifically adapted for pain recognition. Camera-based facial pain expression recognition includes localizing facial landmarks (points along mouth, eyes, and eyebrows, etc.) and registering landmarks and/or facial texture to gain invariance to translation, scale, and rotation. We extracted Local Binary Pattern (LBP) and facial distance features. In the heart rate analysis, we studied heart rate variability as an indicator of pain. Moreover, in the EEG analysis, the feature in the brain signal alpha channel was discriminated from the data, and the eyes open/closed ratio was extracted, based on an evolving idea from emotion studies. The labels (pain intensity level) were extracted from the answers to the in Phase 1 and Phase 3 VAS questionnaire (Fig. 1) and from the collected home measurement data, Question 1 and Question 10 (Fig. 3b). Video key landmarks' distance, VAS value and texture based spatio-temporal methods (LBP-TOP) were used to estimate low back pain. In order to find discriminative features in the audio analysis, prosodic features and their correlation to pain intensity level were considered. Heart rate analysis concentrated on long-term pain, which is has not been studied so often by analyzing differences between the high pain level group and low pain level group. The EEG analysis examined the correlation of pain intensity level with the Eyes Open/Eyes Closed ratio of alpha power in the defined region of interest. The details of the pre-processing, feature extraction and methods are explained in the following subchapters.

3.1 Audio analysis

The audio analysis was performed using prosodic data and mainly by observing the F0 fundamental frequency. F0 frequency is defined as the lowest frequency of a periodic waveform. Being the lowest frequency, it is perceived as the pitch of the spoken voice. Studies have shown that the pitch is the most important acoustic parameter in terms of the identification of emotion or attitude. Other relevant parameters are the duration of changes in pitch, energy and the ratios between speech and silence [20].

The audio recordings of the text that the users read during the measurements were extracted from the videos. The selected text was identical for all the users and was emotionally neutral, representing matter-of-fact newspaper prose. It has been originally used for discriminating emotion in spoken Finnish [28].
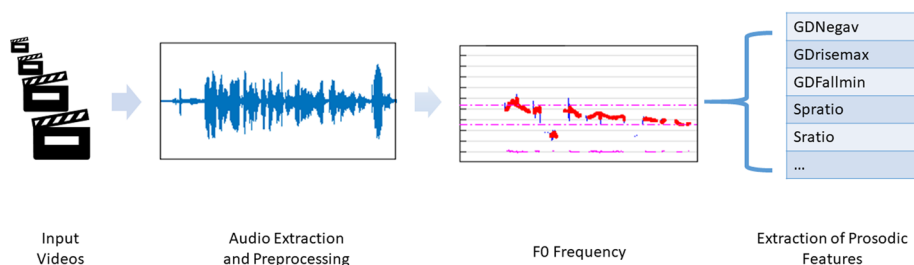
Fig. 5 Prosodic Feature Extraction Process

Pre-processing After the audio tracks were extracted from each video, they were converted into wav format. After this, in the preprocessing procedure, all the tracks shared the same characteristics in the analyses. Specifically, the sampling frequency of the audio was converted to 11,025 Hz, then the audio channels were merged into one and the mean for each recording was multiplied by 10.

Methods Specifically, the prosodic features included modeling the logarithmic fundamental F0 frequency, energy and duration of the voiced and unvoiced segments, which are the most important parameters in speech. Several types of correlates were calculated, constructing a set of 41 prosodic features, including features describing F0 frequency (the short-term energy maximum, mean of short-term energy, F0 maximum, F0 minimum, F0 mean, and F0 standard deviation), spectral, temporal and intensity properties of the speech samples. The feature extraction process is illustrated in Fig. 5. The following features were found to be the most relevant: the average decrease in F0 frequency during the continuous voiced segments (GDnegav), the maximum increase in F0 during the continuous voiced segments (GDrisemax), the maximum decrease in F0 during the continuous voiced segment (GDfallmin), the ratio of speech (voiced + unvoiced < 300 ms) to pauses (Spratio), and the ratio of silence to speech (Sratio).

3.2 Video analysis

Video analysis was performed by first evaluating the correlation between the geometric changes in the face and pain, together with the spatio-temporal facial descriptors to distinguish between no pain and pain. Geometric changes in faces are mainly measured by the changes in the distance between key landmarks (e.g. the distance between the center of the eye center and the tip of the nose). Spatio-temporal features are extracted from video sequences to encode the dynamic changes throughout the spatial and temporal dimensions by using LBP-TOP [35].

Pre-processing During this step, the video data and corresponding labels were processed respectively, as shown in Fig. 6. For video data, the raw videos were first registered, and then the key facial landmarks were determined according to the template shown in Fig. 7. According to the labeled key landmarks, the face regions were cropped and warped. At the same time, the VAS values collected from the tablets were extracted and matched with the corresponding video data. In addition, an average image of all labeled frames was
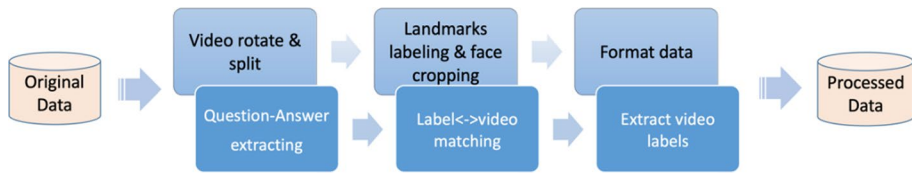
Fig. 6 Data pre-processing procedure

computed, to check the quality of the labeled landmarks. The average image of the pilot data is depicted in Fig. 7.

Methods In order to reveal the relationship between low back pain and facial expression, geometric distance changes in the key landmarks on the face were measured first. Specifically, the distances between the center of the eye and tip of the nose, as shown in Fig. 8 were measured during pain. Changes in such distances directly reflect the appearance of the expression induced by pain.

To capture the facial dynamics that describe the evolution of pain expressions, Local Binary Pattern on Three Orthogonal Planes (LBP-TOP) [35] was adopted to extract the spatio-temporal features. For a video sequence, the spatio-temporal information can be viewed as a set of volumes in (X, Y, T) space, where X and Y represent spatial coordinates and T is the temporal axis. LBP-TOP decomposes the three-dimensional volume into three orthogonal planes: XY, XT, and YT. Then local binary patterns are computed for every pixel on each orthogonal plane by comparing a center pixel with its neighborhoods, denoted as XY-LBP, XT-LBP, YT-LBP. Encoding and concatenating the co-occurrence statistics on the three planes generated the representation for a video sequence.

For the classification, Support Vector Machine [5] was utilized, which separated the samples from two classes (no pain vs. pain) by an optimal hyperplane/decision boundary that had the largest margin to the nearest samples of any class.
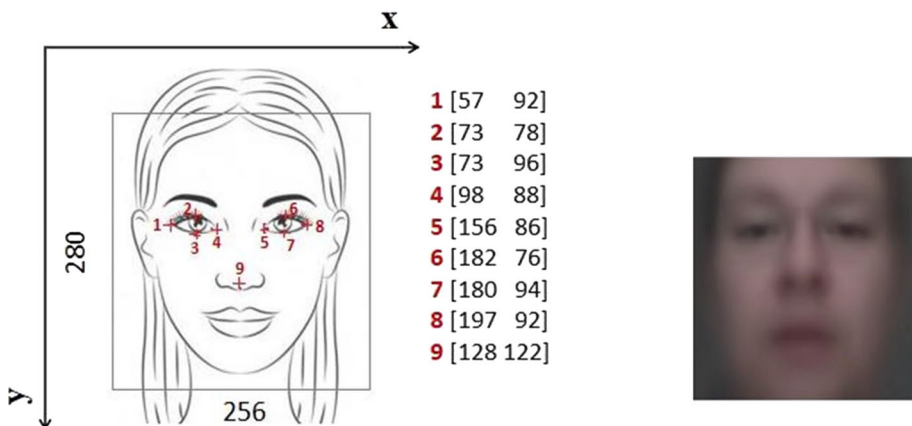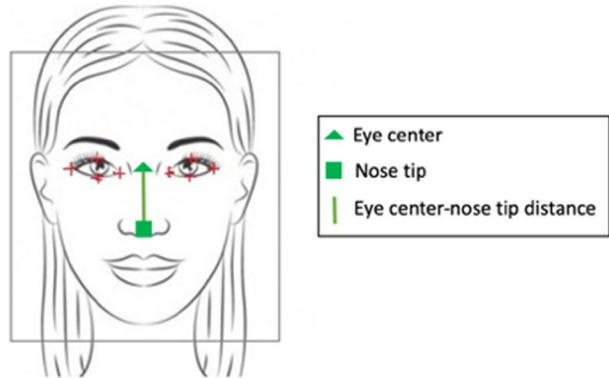

Fig. 7 Key landmarks template and average image of pilot data

Fig. 8 Geometric changes in face

3.3 Heart rate analysis

Previous cross-sectional studies have associated chronic pain with altered ANS regulation. However, less data are available for long-term follow-up of pain and ANS. Therefore, we evaluated whether the autonomic regulation, analyzed using the heart rate variability (HRV) method, of the low back pain patients with severe pain differed from that of those with mild pain during the 1–2-month follow-up, according to the home-based HRV measurements.

Pre-processing Kubios HRV software (University of Kuopio, Kuopio, Finland) [26] was utilized for pre-processing and HRV analysis. An example of the heart rate data recording over 24 h is shown in Fig. 9. Kubios HRV includes two methods for correcting any artifact and ectopic beats in the RR interval data. In our analysis, we used the first method, the threshold-based correction method, in which the artifacts and ectopic beats are simply corrected by comparing every RR interval value against a local average interval. We specified the extreme threshold (extreme: 0.05 s). This indicates that the "extreme" correction level identifies all RR intervals that are 0.05 s longer/shorter than the local average. The correction was made by replacing the identified artifacts with interpolated values using a cubic spline interpolation. In addition, we used smoothness priors based on the detrending approach for removing baseline wandering, using a cut-off frequency of 0.035 Hz and a regularization parameter of 500.

Methods The heart rate data of 10 volunteers (aged 45 ± 10 years, 4 men and 6 women) who had suffered from low back pain at average of 37 ± 21 months were available. Measurements were carried out in the home of every patient. The Bittium Faros 360 device (Bittium Corporation, Oulu, Finland) was used to extract the RR interval and accelerometer data for each low back pain patient for 24 h once a week. After waking in the morning, the patients' pain experience on the VAS scale was assessed using question 10: How severe is your back pain at the moment? The pain recordings were saved on the tablet and RR interval recordings on a computer for further analysis of HRV. From the collected RR interval data we analyzed one-hour data during sleeping hours confirmed by the accelerometer data (the average of four measurement points). Finally, the pre-processed time-series data were analyzed using the Kubios HRV software.
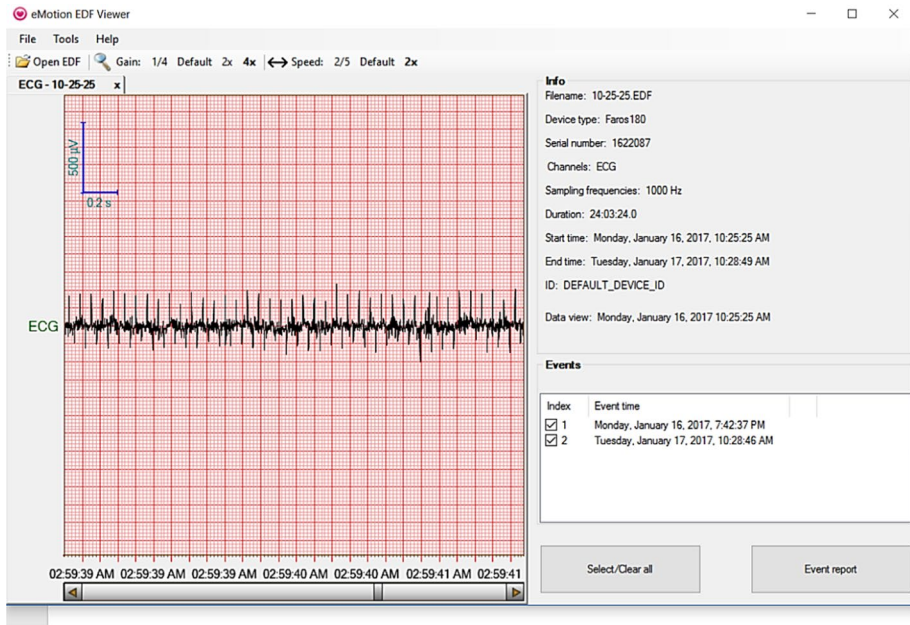
**Fig. 9** Representative example of raw heart rate data in Kubios HRV software

3.4 EEG analysis

Pre-processing The EEG signals were pre-processed using standard methods by segmenting the EEG recordings (the first and the last minute were excluded from the 5-min EEG recording). From the middle period of 3 min, 10 subsets of EEG signals, each lasting one second, were selected, and down-sampled to 1000 Hz. Baseline correction, filtering (a low-pass filter with a 50-Hz cut-off frequency was used to remove powerline noise) and artifact rejection was realized by Independent Component Analysis (to remove the eye movement component). A spectral analysis by Fast Fourier Transform (FFT) was applied to signals, focusing on the alpha band (8–13 Hz).

Methods Previous studies [6, 23] have applied a feature of alpha symmetry to study the effect of music on emotion. The alpha asymmetries in the electrode pairs of F3/F4, F5/F6, C3/C4, C5/C6, and T8/T7 were studied. The alpha asymmetry index (AI) is calculated by the Eq. (1):

$$AI = \left(P_{right} - P_{left}\right)/P_{baseline}, \tag{1}$$

where baseline is the average of alpha waves from all electrode sites in both the left and right hemispheres [23]. In both studies, the emotions were music induced. In [6], the effect of eye states on the modulation of the association between music and EEG markers was also studied.

Fig. 10 ROI Selection
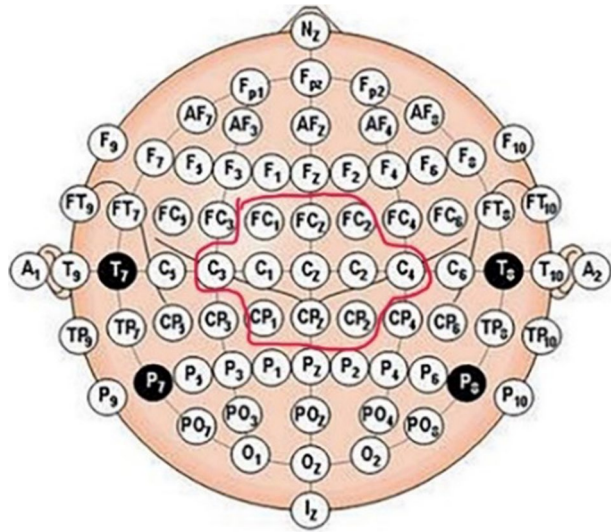
In our analysis, in which the pain intensity was static throughout the EEG trial, and not externally induced, we evaluated the relationship between pain intensity and the examined properties in the alpha channel. We calculated the alpha modulation in Eyes-Open (EO) to Eyes-Closed (EC) with a power ratio of the alpha band (8-13 Hz) between EO to EC in a selection ROI (See Fig. 10), as defined by Eq. (2):

$$Modulation\ ratio = \frac{Power_{\alpha}(EO)}{Power_{\alpha}(EC)}, \tag{2}$$

4 Experimental results

We evaluated the subjective ratings of the pain level during the study period, in which around 10–24 measurements were taken by every patient at home. In Phase 1, the intensity of low back pain was reported by the patient (Baseline VAS, Table 1). The baseline VAS values ranged from 0 to 10 with an average of 4.9 (average VAS values that are higher than 4 are highlighted in yellow). In Table 1 below we report the minimum and maximum values along with the averages of the patient data for both before (Q1) and after the bending task (Q10). For comparison, the 1–2-month follow-up VAS values evaluated in Phase 3 are also reported in the table.

In total, all 14 participants' data (273 audio/video recordings) are used in this study. The histogram of pain level of all the data is indicated in Fig. 11. In the EEG analysis we used altogether 28 recordings from 14 patients. The portion of the heart rate data was excluded due to artifacts, and data from 10 patients are used.

Table 1 Statistics of reported pain intensities during experiments

Patient	Baseline VAS	min Q1	max Q1	min Q10	max Q10	AVG Q1	AVG Q10	Follow-up VAS
Data_1	4	1	5	2	4	2.4	3	6
Data_2	7	6	8	6	8	7.2	8	4
Data_3	0	0	3	0	4	1.5	2	0
Data_4	3	0	3	0	2	1.5	1	2
Data_5	5	0	2	0	3	1	1	2
Data_6	4	3	6	2	6	4	4	4
Data_7	5	0	3	0	4	1.3	2	2
Data_8	10	1	4	1	3	2.7	2	6
Data_9	8	4	7	3	8	5.4	6	4
Data_10	8	2	8	4	8	4	5	7
Data_11	6	0	5	0	5	2.7	3	5
Data_12	5	3	8	3	9	5.8	7	6
Data_13	2	0	3	0	4	1.5	2	2
Data_14	2	0	1	0	2	0.6	1	1

4.1 Audio analysis results

The pre-processed audio files underwent a final processing round using the F0 frequency tools, which extracted the prosodic features of the audio data. Figures 12–14 demonstrate how the difference between the selected prosodic features correlated with the difference in the VAS values. Specifically, this was $dVAS = VAS_2 - VAS_1$ where VAS_1 is the value before and VAS_2 the value after the bending exercise. Since the patient with ID 12 reported the largest increase in pain between the responses, he was selected as the following example. The patient repeated the test 10 times and the results are shown in Figs. 12–14 in the form of scatter graphs with 10 points and a line that depicts the linear fit.
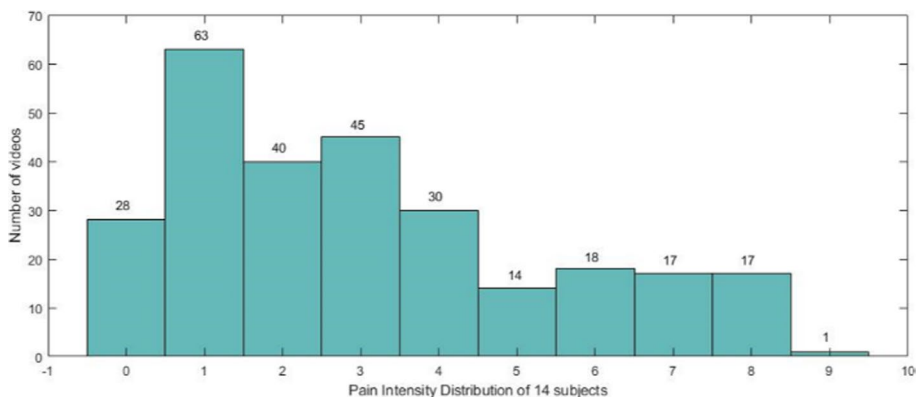

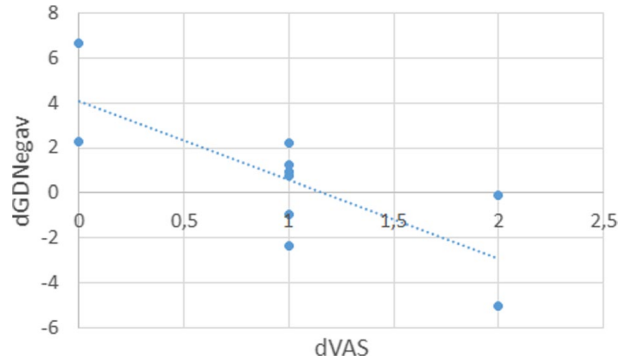**Fig. 11** Histogram of pain intensities

Fig. 12 Difference in average F0 decrease from patient ID12

Figure 12 depicts the difference in GDNegav which is equal to $dGDNEgav = GDNegav_2 - GDNegav_1$, where $GDNegav_1$ is the value before the bending exercise and $GDNegav_2$ after the bending exercise. GDNegav is the average decrease in the F0 frequency during continuous segments and is measured in Hz. The figure demonstrates how this difference decreases and eventually becomes negative when the pain intensities are higher after the bending exercise.

Similarly, Fig. 13 shows the difference in GDfallmin, which is the maximum decrease in the F0 frequency.

The reason for the observation is explained as the pitch of the voice tending to increase when a person experiences pain. As a result, the fundamental or F0 frequency tends to decline less when the pain level is high.

Figure 14 depicts the difference in Spratio, which is the ratio of speech to pauses. The ratio decreases as the pain increases. Unfortunately, the specific example did not contain many samples with pain increasing by two units, so this is not perfectly perceived from this example, but Fig. 15 makes this clearer by including the average results of all 14 patients.

Specifically, Fig. 15 demonstrates the ratio of speech using the average speech to pauses ratios and average VAS responses before and after the bending exercises for all 14 patients.

Here it becomes more obvious that higher VAS values correspond to smaller ratios of speech to pauses. This is because people tend to stutter and generally make pauses in their speech more often when they suffer from pain.
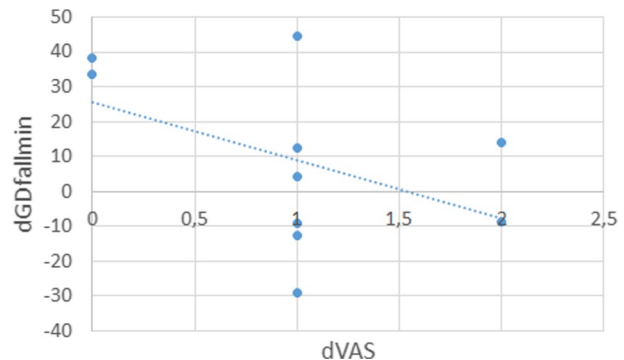
Fig. 13 Difference in maximum F0 decrease from patient ID12

Fig. 14 Difference in ratio of speech to pauses from patient ID12
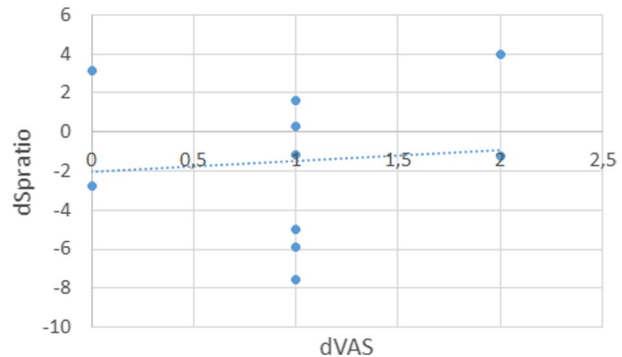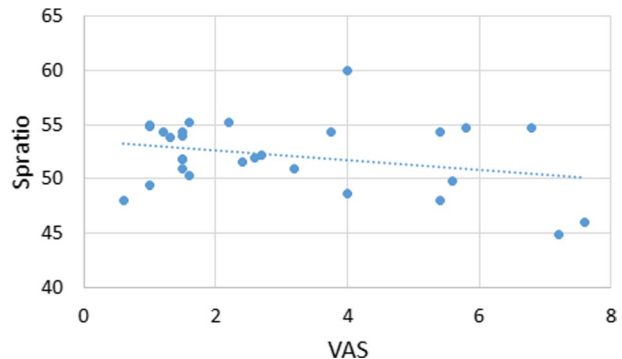

Fig. 15 Ratio of speech to pauses, average values from all 14 patients


4.2 Video analysis results

In order to demonstrate how low back pain reflects on facial expression, we first analyzed the correlation between key landmarks' distance and VAS value. The center of the eye was computed from the landmarks of the left eye and right eye and mostly located in the nasal area. Then the distances from the center of the eye to the nose, the center of the left eye to the nose, the center of the right eye to the nose were calculated and normalized. Finally, the correlations between the distance from the center of the eye to the nose and the VAS values were analyzed, as changes in the nasal area are vital signs of expressing pain. The stronger the pain, the smaller the distance between the center of the eye and the nose. Figure 16 shows that the distances from the center of the eye to the nose decreased when the patient had a higher level of pain, which validates the assumption.

Texture based spatio-temporal methods (LBP-TOP) were also used to estimate low back pain. Since the distribution of pain intensities (Fig. 11) was highly imbalanced, we divided the pain and no pain classes according to whether the VAS value was greater than 2 or less than/equal to 2. Hence it became a binary classification problem to distinguish pain and no pain. The evaluation metrics, training strategies and corresponding results are demonstrated as follows.

Fig. 16 Correlation between distance from the center of the eye to the nose and VAS values for patient ID14
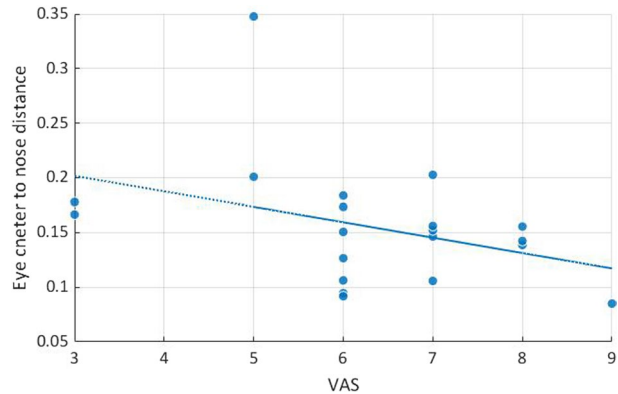
Evaluation metrics Accuracy and F1 score were used as the evaluation methods in video analysis. Based on the confusion matrix listed in Table 2, accuracy and F1 score were calculated according to Eq. (3) and Eq. (6).

$$ACC = \frac{TP + TN}{P + N} \tag{3}$$

$$precision = \frac{TP}{TP + FP} \tag{4}$$

$$recall = \frac{TP}{TP + FN} \tag{5}$$

$$F_1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} \tag{6}$$

Training details Two training strategies were used in the video analysis: leave-one-participant-out cross validation and tenfold cross validation. During the training time, data from one participant (fold) were left as testing data, whereas the data on the rest of the participants (folds) were used as training data.

Results analysis Tables 3 and 4 show the results of different LBP-TOP parameters, where p is the number of neighboring points sampled on XY plane, XT plane, and YT plane, respectively, r represents the radius on local neighboring pixels that form a circularly symmetry neighbor set, and bloc is the volume size of each spatio-temporal block. For instance, parameters with bloc = 6*6*5 represented the number of divisions in X, Y and T directions were 6, 6, 5 respectively.

Table 2 Confusion matrix

	Ground truth positive	Ground truth negative
Predicted condition positive	True positive (TP)	False positive (FP) I Error
Predicted condition negative	False negative (FN) II Error	True negative (TN)

Table 3 Performance in video data using Leave-one-participant-out strategy

Method	Parameters	ACC	F1 score
LBP-TOP	p=8,r=1,bloc=6*6*5	0.5055	0.5055
	p=8,r=1,bloc=8*6*6	0.5018	0.5108
	p=8,r=1,bloc=10*7*5	0.5018	0.4769
	p=8,r=2,bloc=6*6*5	0.4945	0.4773
	p=8,r=2,bloc=8*6*6	0.4359	0.4420
	p=8,r=2,bloc=10*7*5	0.4615	0.4411

Table 4 Performance in video data using tenfold cross validation strategy

Method	Parameters	ACC	F1 score
LBP-TOP	p=8,r=1,bloc=6*6*5	0.8422	0.8512
	p=8,r=1,bloc=8*6*6	0.8235	0.8298
	p=8,r=1,bloc=10*7*5	0.8280	0.8385
	p=8,r=2,bloc=6*6*5	0.8459	0.8542
	p=8,r=2,bloc=8*6*6	0.8317	0.8425
	p=8,r=2,bloc=10*7*5	0.8493	0.8581

For the leave-one-participant-out strategy, parameters with p=8, r=1, the block division of 6*6*5 achieved the best accuracy of 0.5055, whereas for the tenfold cross validation, parameters with p=8, r=2, the block division of 10*7*5 received the best performance with an accuracy of 0.8493 and an F1 score of 0.8581. The huge difference between the two protocols reveals that it is rather challenging to distinguish no pain and pain for an unseen participant with small scale data size, as in the case of the leave-one-participant-out protocol.

4.3 Heart rate analysis

The patients with low back pain were divided into two groups according to their average level of pain during Phase 2: high pain level group (HPG n=4, VAS>4) and low pain level group (LPG n=6, VAS<4).

The average pain level during follow-up was 6.5 ± 1.6 for the HPG and 1.4 ± 0.6 for the LPG ($p < 0.0001$). Heart rate during night hours showed a tendency be lower ($p=0.062$) for the LPG (70 ± 8 bpm) than for the HPG (80 ± 6 bpm). The average of vagally mediated high frequency power was significantly lower in the HPG than in the LPG (1.8 ± 0.3 ms2 vs. 2.2 ± 0.2 ms2, $p=0.028$, respectively) and indicator of sympathovagal balance (low to high frequency ratio) was higher in the HPG than in the LPG (4.7 ± 2.0 vs. 2.2 ± 0.8, $p=0.039$, respectively). There was also a clear trend toward lower values of high frequency power for the HPG compared to the LPG when analyzed week by week (Fig. 17) The groups did not differ from each other in terms of gender, smoking or pain duration ($p=ns$ for all). Interestingly, three of the four patients with low back pain in the HPG had some diagnosis of cardiometabolic disease as a comorbidity, whereas none were found in the LPG ($p=0.033$).

The results of this study show that a higher level of long-term pain is associated with decreased vagal activity and increased sympathetic activity towards the heart. As altered ANS regulation has been associated with several diseases and prognosis, the results of this preliminary follow-up study of low back pain patients highlight the importance of effective treatment of pain.

Fig. 17 A trend towards lower values of vagally mediated high frequency power (HFln2) for the HPG (VAS > 4) compared to the LPG (VAS < 4) was observed (general linear model for main effect p = 0.110) during Phase 2 (1–2 months follow-up)
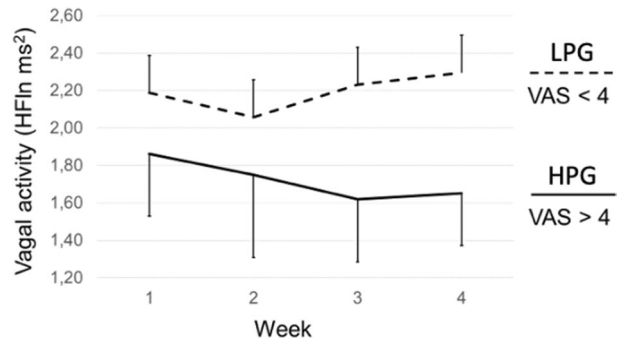
4.4 EEG analysis results

Evaluation metrics The correlation coefficient between the alpha modulation ratio and the VAS value of low back pain intensity was calculated at Phase 1 (Baseline VAS) and Phase 3 (Follow-up VAS).

Results analysis The correlation coefficient showed -0.575 in the baseline experiment and -0.622 in the follow-up experiment (Fig. 18). The alpha oscillatory in resting state intrinsic EEG showed high power in the eyes-closed with increasing pain intensity, whereas the eyes-open alpha rhythm had low power in pain. That is, the higher the intensity of pain the lower the EO/EC ratio of the alpha power.

Fig.18 Correlation analysis; baseline (trial 1) and follow-up (trial 2)
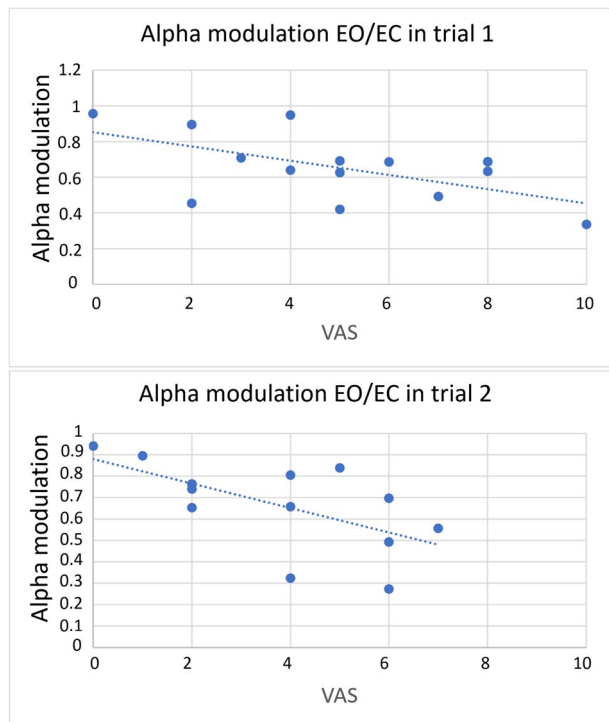
5 Discussion

Automatic pain recognition and assessment has been a research area of growing interest in the last decade, containing activity on developing methods for the data acquisition and computation. Data acquisition includes many issues that are related to the devices and sensors used, setup of the experiments and the processing of the signals. Moreover, defining the protocol for data collection is crucial, including participant recruitment and selection and preparation of the pain provocation experiments and context, requiring expertise and collaboration of various branches of science. The main research in the area has been on exploiting the available databases [2, 11, 13, 18, 30, 31, 34]. Appendix Table 5 contains the basic properties and the scale of the experiments described in the state-of-the-art literature, illustrating the current status of existing databases and studies. It also shows the challenges in pain-related research: the selection of appropriate data, devices, pain stimuli, protocol design, and data labelling. Additional challenges are the experimental environment, the guidance required, patient recruitment, and ethical issues. All the studies (Appendix Table 5) were conducted in a laboratory environment. In [11, 13, 30, 31, 34] and [7] heat or electrical stimulation was used to induce pain. In these cases, it was possible to calibrate the stimulus for annotation per person. [18] and [2] use motion (controlled movement) as a stimulus. In these self-reports, the VAS scale and both expert and lay observers were used to annotate the data. Overall, a typical system design in the related literature had several devices, and the position of the cameras and the placement of sensors were configured very precisely. The measurements were also guided by experts in a highly controlled environment.

Despite the increasing interest in automatic pain recognition, a major challenge for advancing this area of research is still the lack of research in real environments [19, 27, 32]. In home measurements, pain provocation must be designed so that an application can guide it and it must be applicable without continuous supervision. The selection of appropriate data, devices, pain stimuli, protocol design, data labelling, and the implementation of the questionnaires require additional attention. In this sense, the home environment or any other real environment poses even more challenges. In this study, a pilot data set of 14 patients was collected with defined protocol, measurement setup and application in order to evaluate the validity of the proposed approach especially in the home environment, which differs significantly from the laboratory settings in [2, 11, 13, 18, 30, 31, 34]. Each participant was measured weekly for 1–2 months.

In addition to the data acquisition, we present analyses to validate the data. Discriminative features were extracted from the collected data, based on ideas that have proven to be successful in other domains but not specifically utilized in pain research. Our preliminary experiments showed that some features extracted from the data can indicate the existence of low back pain or trends in low back pain. For example, a decrease in the distances between the center of the eye and the nose indicates the appearance of pain in facial expressions, which is consistent with the studies in [33]. Earlier studies [21] have also demonstrated that one of the symptoms of pain is less rapid human responses, which also affects some voice parameters. Similarly, our results indicate that F0 frequency tends to decline less when the pain level is high and the ratio of speech to pauses decreases as the pain level increases. Interesting findings were also found regarding the heart rate and EEG data. For the heart rate analysis, we hypothesize that relief of individual pain experience may facilitate more normalized and healthier ANS regulation [25]. Moreover, previous studies have indicated that frontal, parietal, and temporal alpha symmetries could reflect emotional valence [6, 23]. Our results indicated a correlation between the level of pain intensity and EO/EC ratio of alpha power in EEG signal.

In the experiments we were able to discover features that correlated with self-reported VAS values, also providing preliminary classification results regarding pain recognition. However,

these findings should be confirmed in studies with larger databases. The performance of the whole system could be improved by appropriately combining the information provided by several modalities using classifier fusion strategies [27]. Fusion strategies are used to combine different modalities, features, decision scores, or other information sources to obtain a single final prediction. Although feature fusion is common, performance should be compared with the results obtained from the individual inputs or other fusion methods. Decision fusion methods combine the outputs of multiple models, either by a fixed rule or another trained model. The results of the feature fusion or decision fusion strategies depend on the selected classifier and data [32]. Thus, to obtain optimal results, several techniques should be compared, especially in the case of multimodal data, when various features can be extracted from each modality. For this, more data need to be gathered and evaluated and different options tested. We feel that the performance of a multimodal system benefits from better input, and accordingly improving recognition within each single modality is valuable. In general, one of the biggest challenges is overcoming the fact that different individuals react differently to pain. Methods should be developed to confront this issue by more efficiently performing the analyses, for example, by normalizing the data, grouping the data into different categories, or developing person-specific models.

Studies require obtaining informed consent from research participants and the anonymization of identifiable information. The GDPR contains regulations concerning individuals' data protection (https://gdpr-info.eu/). This requires paying extra attention to the management of information, such as videos or images, in which individuals can be identified. Unfortunately, this may affect the availability of images and video information for research in public databases in the future.

6 Conclusions

This paper presented a novel protocol for low back pain estimation based on the data collected from both hospital and home measurements, casting light on the potential of pain monitoring in natural scenarios. All the data, including questionnaires completed by patients before the measurements, videos, audio recordings, heart rate information, and feedback were collected through a user-friendly application on a tablet, which eased the process of home monitoring of pain. The data, consisting of multiple modalities, were analyzed to determine the correlations between features or trends in the data and the intensity or amount of increase/decrease in pain.

The analysis of the pilot data was a useful source of information and feedback to dictate how the current setup should evolve to be able to collect larger volumes of data. Combining physiological responses and audio-visual responses with self-reported pain intensity, questionnaires, and developing a regression model can be considered a future step for quantitatively assessing pain experience. The characterization of bio-signals from individual pain patients is crucial for identifying relevant domains of pain experience to achieve more targeted treatments and thereby prevent the development of chronic pain. The use of modern technology, such as home-based monitoring applications may enable us to assess pain experience more granularly at the individual level.

Enabling the monitoring of pain experience (smartphone/laptop/tablet applications) during home measurements provides huge potential for improving the treatment of pain patients. Our research extends what has been done so far by providing a more efficient model for measurements. Specifically, it shows that recordings can also take place at home with the use of a mobile device and a user-friendly application designed to collect multimodal data without a person supervising the process. This allows long-term follow-up of patients for more individualized treatment decisions.

Appendix

Table 5 Existing databases and studies in state-of-the-art literature

	Selection of data	Devices	Pain stimulus and protocol design	Data labelling	Environment/ guidance	Patient recruitment
[31]	-video of face -SCL -ECG -sEMG -EEG	-Kinect camera -three video cameras -Nexus-32 amplifier (SCL, ECG, sEMG EEG)	-heat pain -4 intensi-ties × 20 repetitions -emotion elicitation	-stimulus calibrated per person -four pain levels	-laboratory -controlled	-90 healthy subjects -study approved by eth-ics committee
[18]	-video of face	-two Sony dig-ital cameras	-eight motion tests	-self-report (VAS, sensory and affective verbal scales) -observer assessed pain inten-sity (OPI) -FACS coding	-laboratory room -controlled -physi-otherapist involved for passive tests	-129 adult shoulder pain patients -recruited from 3 physi-otherapy clinics and by advertisements -informed consent
[34]	-video of face -EDA -EGC -RSP -BVP	-3D camera -infrared camera -BioPac (EDA, EGC, RSP, BVP)	-cold pressor test	-stimulus -FACS coding	-laboratory -controlled	-141 healthy subjects -study approved by IRB -informed consent
[30]	-video of face -audio -EDA -ECG -sEMG -RSP	-3 video cameras -MOBIlab (EMG,ECG) -GSRsensor (SCL) -RSP sensor -digital wire-less headset microphone -directional microphone	-heat pain, -image stimu-lus -heat pain (3 intensi-ties × 30 rep-etitions × 2 stimulus sites) -emotion elicitation	-pain and emotion stimulus -pain cali-brated per person -self-report (VAS)	-laboratory -controlled	-45 healthy subjects -recruited by advertise-ments -informed consent -study approved by eth-ics committee
[2]	-video -audio, -sEMG	-8 video cameras -customized motion cap-ture suit -wireless sEMG probes	-physical exercises designed by physiothera-pists with expertise in treating chronic low back pain	-self report (1–10 scale) -pain intensity assessed by naive observers -presence of pain behaviors assessed by experts	-laboratory -controlled -followed by physiothera-pist or psy-chologist	-22 chronic lower back pain patients -28 healthy controls identified by health care staff -informed consent -study approved by eth-ics committee

Table 5 (continued)

	Selection of data	Devices	Pain stimulus and protocol design	Data labelling	Environment/ guidance	Patient recruitment
[13]	-RGB, depth and thermal video of face -EMG	-Kinect camera -thermal camera -surface electrodes (EMG)	-electrical pain stimulation -two trial for no pain and for four pain levels	-stimulus calibrated per person -self-report (VAS)		-20 healthy subjects
[11]	-video of face -audio -ECG -SCL -sEMG	-4 video cameras -BioPac (ECG,SCL, sEMG)	-heat and electrical pain -3 intensities	-pain stimulus calibrated per person	-laboratory -camera-monitored, temperature controlled, low-noise calibration and monitoring room -controlled	-134 healthy subjects -study approved by ethics committee
[7]	-BVP -ECG -SCL	-three-channel biosensors (BVP,ECG, SCL)	-electrical stimulation -repetition to obtain 7 day data set	-stimulus intensity -five pain states	-laboratory -controlled	-six healthy subjects

SCL: skin conductance level, ECG: electrocardiogram, sEMG: surface electromyography, EEG: electroencephalography, EDA: electrodermal activity, IRB: Institutional Review Board, FACS: Facial Action Coding System, RSP: Respiration, BVP: blood volume pulse.

Availability of data and material Not available.

Code availability Not available.

Declarations

Ethics approval The Ethical Committee of the Northern Ostrobothnia. Hospital District, Oulu, Finland, approved the PASE study protocol.

Consent to participate All voluntary patients and healthy participants gave their informed consent.

Conflicts of interest/Competing interests The authors declare that they have no known competing financial interests or personal relationships that could appear to influence the work reported in this paper.

References

1. Apkarian AV, Bushnell MC, Treede RD, Zubieta JK (2005) Human brain mechanisms of pain perception and regulation in health and disease. Eur J Pain 9:463–484
2. Aung MS, Kaltwang S, Romera-Paredes B, Martinez B, Singh A, Cella M, Valstar M, Meng H, Kemp A, Shafizadeh M (2015) The automatic detection of chronic pain-related expression: requirements, challenges and the multimodal EmoPain dataset. IEEE Trans Affect Comput 7:435–451
3. Bijur PE, Silver W, Gallagher EJ (2001) Reliability of the visual analog scale for measurement of acute pain. Acad Emerg Med 8:1153–1157
4. Caraceni A, Cherny N, Fainsinger R, Kaasa S, Poulain P, Radbruch L, De Conno F (2002) Pain measurement tools and methods in clinical research in palliative care: recommendations of an Expert Working Group of the European Association of Palliative Care. J Pain Symptom Manage 23:239–255
5. Chang CC, Lin CJ (2011) LIBSVM: A library for support vector machines. ACM Trans Intell Syst Technol (TIST) 2:27
6. Chang Y, Lee Y, Liang K, Chen I (2015) Tsai C and Hsieh S (2015) Experiencing affective music in eyes-closed and eyes-open states: an electroencephalography study. Front Psychol 6:1–9
7. Chu Y, Zhao X, Yao J, Zhao Y, Wu Z (2014) Physiological signals based quantitative evaluation method of the pain. IFAC Proceedings Volumes 47:2981–2986
8. Fairbank J, Pynsent P (2000) The Oswestry disability index. Spine 25:2940–2953
9. Garland EL (2012) Pain processing in the human nervous system: a selective review of nociceptive and biobehavioral pathways. Prim Care 39:561–571
10. Gawande A (2010) Checklist manifesto, the (HB). Penguin Books IndiaA. Gawande, Checklist manifesto, the (HB)., Penguin Books India., 2010
11. Gruss S, Geiger M, Werner P, Wilhelm O, Traue HC, Al-Hamadi A, Walter S (2016) Multi-modal signals for analyzing pain responses to thermal and electrical stimuli. J Vis ExpS 146:e59057
12. Haefeli M, Elfering A (2006) Pain assessment. Eur Spine J 15:17–24
13. Haque MA, Bautista RB, Noroozi F, Kulkarni K, Laursen CB, Irani R, Bellantonio M, Escalera S, Anbarjafari G, Nasrollahi K (2018) Deep multimodal pain recognition: a database and comparison of spatio-temporal visual modalities. 13th IEEE Int Conf Autom Face Gesture Recognit: 250–257
14. Kuner R (2010) Central mechanisms of pathological pain. Nat Med 16:1258–1266
15. Kunz M, Meixner D, Lautenbacher S (2019) Facial muscle movements encoding pain-a systematic review. Pain 160:535–549
16. Lee J, Mawla I, Kim J, Loggia ML, Ortiz A, Jung C, Chan ST, Gerber J, Schmithorst VJ, Edwards RR et al (2019) Machine learning-based prediction of clinical pain using multimodal neuroimaging and autonomic metrics. PAIN 160:550–560
17. Li J, Weng Y, Ng C, Lee C (2018) Learning conditional acoustic latent representation with gender and age attributes for automatic pain level recognition. Proc Interspeech 3438–3442. https://doi.org/10.21437/Interspeech.2018-1298
18. Lucey P, Cohn J, Prkachin K, Solomon P, Iain M (2011) Painful data: the UNBC-McMaster shoulder pain expression archive database. IEEE International Conference on Automatic Face and Gesture Recognition and Workshops 57–64
19 M Al-Eidan R, Al-Khalifa H, Al-Salman A (2020) Deep-Learning-Based Models for Pain Recognition: A Systematic Review. Appl Sci 10:5984
20. Mozzicanacci SJ, Hermes DJ (1997) A study of intonation patterns in speech expressing emotion or attitude: production and perception. IPO Annual Progress Report 32:154–160S. J
21. Oshrat Y, Bloch A, Lerner A, Cohen A, Avigal M, Zeilig G (2016) Speech prosody as a biosignal for physical pain detection. Conferece Proceedings of 8th Speech Prosody 420–424
22. Osterweis M, Kleinman A, Mechanic D (1987) Pain and disability: clinical, behavioral, and public policy perspectives. National Academy Press, Washington, DC

23. Schmidt B, Hanslmayr S (2009) Resting frontal EEG alpha-asymmetry predicts the evaluation of affective musical stimuli. Neurosci Lett 460:237–240
24. Shu L, Xie J, Yang M (2018) A review of emotion recognition using physiological signals. Sensors (Basel) 18:2074
25. Södervall J, Karppinen J, Puolitaival J, Kyllönen E, Siltala J, Kiviniemi AM, Tulppo MP (2013) Hautala AJ (2013) Heart rate variability in sciatica patients referred to spine surgery. BMC Musculoskelet Disord 26(14):149. https://doi.org/10.1186/1471-2474-14-149
26. Tarvainen MP, Niskanen JP (2008) User's Guide of Kubios HRV version 2. Biosignal Analysis and Medical Imaging Group, BSAMIG
27. Thiam P, Kessler V, Amirian M, Bellmann P, Layher G, Zhang Y, Velana M, Gruss S, Walter S, Traue HC (2021) Multi-modal pain intensity recognition based on the senseemotion database. IEEE Trans Affect Comput 12(3):743–760
28. Toivanen J, Väyrynen E, Seppänen T (2004) Automatic Discrimination of Emotion from Spoken Finnish. Lang Speech 47:383–412
29. Tsai F, Weng Y, Ng C, Lee C (2017) Embedding stacked bottleneck vocal features in a LSTM architecture for automatic pain level classification during emergency triage. International Conference on Affective Computing and Intelligent Interaction: 313–318
30. Velana M, Gruss S, Layher G, Thiam P, Zhang Y, Schork D, Kessler V, Meudt S, Neumann H, Kim J (2016) The senseemotion database: A multimodal database for the development and systematic validation of an automatic pain-and emotion-recognition system. IAPR Workshop on Multimodal Pattern Recognition of Social Signals in Human-Computer Interaction:127–139
31. Walter S, Gruss S, Ehleiter H, Tan J, Traue HC, Werner P, Al-Hamadi A, Crawcour S, Andrade AO, Silva GMd (2013) The biovid heat pain database data for the advancement and systematic validation of an automated pain recognition system. IEEE International Conference on Cybernetics: 128–131
32. Werner P, Lopez-Martinez D, Walter S, Al-Hamadi A, Gruss S, Picard R (2019) Automatic recognition methods supporting pain assessment: a survey. IEEE Trans Affect Comput P. https://doi.org/10.1109/TAFFC.2019.2946774
33. Werner P, Al-Hamadi A, Limbrecht-Ecklundt K, Walter S, Gruss S, Traue HC (2016) Automatic pain assessment with facial activity descriptors. IEEE Trans Affect Comput 8:286–299
34. Zhang Z, Girard JM, Wu Y, Zhang X, Liu P, Ciftci U, Canavan S, Reale M, Horowitz A, Yang H (2016) Multimodal spontaneous emotion corpus for human behavior analysis. Conference on Computer Vision and Pattern Recognition (CVPR): 3438–3446
35. Zhao G, Pietikäinen M (2007) Dynamic texture recognition using local binary patterns with an application to facial expressions. IEEE Trans Pattern Anal Mach Intell 6:915–928

Authors and Affiliations

Anja Keskinarkaus[1] [ORCID] **· Ruijing Yang[1,2] · Angelos Fylakis[1] · Md. Surat-E-Mostafa[1] · Arto Hautala[1,3] · Yong Hu[4] · Jinye Peng[2] · Guoying Zhao[1] · Tapio Seppänen[1] · Jaro Karppinen[5]**

Ruijing Yang
ruijing.yang@oulu.fi

Angelos Fylakis
angelos.fylakis@oulu.fi

Md. Surat-E-Mostafa
mmostafa@student.oulu.fi

Arto Hautala
arto.j.hautala@jyu.fi

Yong Hu
yhud@hku.hk

Jinye Peng
pjy@nwu.edu.cn

Guoying Zhao
guoying.zhao@oulu.fi

Tapio Seppänen
tapio.seppanen@oulu.fi

Jaro Karppinen
jaro.karppinen@oulu.fi

1   Center for Machine Vision and Signal Analysis (CMVS), Faculty of Information Technology
    and Electrical Engineering, University of Oulu, Oulu, Finland

2   School of Information Science and Technology, Northwest University, Xi'an, China

3   Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, Finland

4   Neural Eng. & Clin Electrophysiology Laboratory, Department of Orthopaedics & Traumatology,
    Li Ka Shing Faculty of Medicine, University of Hong Kong, Pok Fu Lam, Hong Kong

5   Medical Research Center Oulu, Oulu University Hospital and University of Oulu, Oulu, Finland