applied sciences

IMPACT FACTOR 2.5

CITESCORE 5.5

Article

The Face of Low Back Pain: A Preliminary Method for Quantifying Pain-Related Facial Expressions

Franciele Parolini, Ricardo Pires, Sara Dereste dos Santos, Márcio F. Goethel, Klaus Becker, João Paulo Vilas-Boas, Rubim Santos and Ulysses F. Ervilha

MDPI

https://doi.org/10.3390/app16062830

applied sciences

MDPI

Article

The Face of Low Back Pain: A Preliminary Method for Quantifying Pain-Related Facial Expressions

Franciele Parolini [1,2,*], Ricardo Pires [3], Sara Dereste dos Santos [3], Márcio F. Goethel [2,4], Klaus Becker [2,4], João Paulo Vilas-Boas [2,4], Rubim Santos [1,*] and Ulysses F. Ervilha [2,5]

1 CIR, ESS, Polytechnic of Porto, Rua Dr. António Bernardino de Almeida, n° 400, 4200-072 Porto, Portugal
2 Porto Biomechanics Laboratory, University of Porto, 4200-450 Porto, Portugal; gbiomech@fade.up.pt (M.F.G.); klausmagnobecker@gmail.com (K.B.); jpvb@fade.up.pt (J.P.V.-B.); ulyervil@usp.br (U.F.E.)
3 Electrical Department, São Paulo Campus, Federal Institute of Education, Science and Technology of São Paulo, São Paulo 01109-010, Brazil; ricardo_pires@ifsp.edu.br (R.P.); sarad@ifsp.edu.br (S.D.d.S.)
4 Center of Research, Education, Innovation and Intervention in Sport, Faculty of Sport, University of Porto, 4200-450 Porto, Portugal
5 Laboratory of Physical Activity Sciences, School of Arts, Sciences and Humanities, University of São Paulo, São Paulo 05508-220, Brazil
* Correspondence: fcsp@ess.ipp.pt (F.P.); rss@ess.ipp.pt (R.S.)

Abstract

Background: Facial expressions of pain are essential for pain assessment, yet subjective pain reports often vary between sexes. Traditional self-report measures are prone to bias, and objective methods are needed for more reliable pain evaluation. Objective: To develop and validate a subjectivity-free automated tool to assess acute low back pain using facial expressions recorded during a functional spinal extension task. Participants: Thirty healthy adults, aged 18–40 years. Methods: Participants received intramuscular injections of hypertonic (pain) and isotonic (placebo) saline in the lumbar region during separate sessions. Facial expressions were video-recorded during a submaximal lumbar extension task and analyzed using a custom software based on Haar Cascade and Local Binary Pattern Histogram algorithms, which are techniques that do not require neither training data nor subjective labeling, contrary to what happens in deep learning solutions. Results: The tool successfully detected significant differences in facial expressions between pain, placebo, and pain-free conditions ($p < 0.001$). Test–retest reliability was good (ICC = 0.85). While both sexes showed similar facial expression patterns during pain, males reported higher pain scores on the numeric rating scale ($p < 0.01$). Pain significantly reduced steadiness of force in both sexes. Conclusion: The automated tool objectively quantified facial expressions associated with acute low back pain and revealed sex-related differences in subjective pain perception. This multimodal approach integrating expression analysis, physical performance, and self-report may enhance the accuracy of pain assessment in physiotherapy settings.

Keywords: facial expression of pain; acute low back pain; force; artificial intelligence

Academic Editor: Thomas Lindner

Received: 2 February 2026
Revised: 1 March 2026
Accepted: 11 March 2026
Published: 16 March 2026

1. Introduction

Accurate pain assessment remains a significant challenge in both clinical and research settings. Traditional methods, such as the visual analog scale and numeric rating scale, despite being widely used and validated, are inherently subjective and susceptible to biases, including underreporting or overreporting influenced by emotional, cultural, or contextual

factors [1–3]. To address these limitations, considerable efforts have been directed towards developing objective tools for pain assessment, particularly through the analysis of facial expressions using technologies such as electromyography, infrared thermography, and computer vision [4,5]. These methodologies are designed to overcome the limitations of subjective reporting, providing a more consistent and reliable means of evaluating pain [2,6–9].

Recent advancements have focused on automated facial expression analysis employing machine learning and deep learning techniques. Studies have utilized Local Binary Patterns (LBP) [10], histograms of topographic features [11], and recurrent convolutional neural networks [12] to estimate pain intensity [13] and have demonstrated the potential of combining handcrafted features with deep learning to enhance accuracy in detecting subtle facial cues. However, despite the technical sophistication of these methods, many require substantial computational power and complex training procedures, limiting their applicability in real-time or resource-constrained clinical settings.

Contemporary theories of pain communication emphasize that pain is expressed through multiple behavioral channels, including verbal reports, postural changes, and facial expressions [14]. From an evolutionary perspective, these expressions serve adaptive functions to signal distress and promote survival [15,16]. While verbal pain reports remain the gold standard in clinical practice [17], they are subject to significant interindividual variability and psychological modulation [18,19]. Facial expressions, often undervalued in clinical practice, provide complementary data that can help triangulate pain assessment, offering a more comprehensive understanding of the patient's condition [20,21].

Despite the significant advancements in objective facial pain expression analysis, most studies have focused on specific populations, such as infants, children, or individuals with cognitive impairments [22–24], with limited exploration in healthy adults or individuals with musculoskeletal pain conditions such as low back pain. To date, few studies have examined the connection between facial expressions and experimentally induced acute low back pain during physical tasks, such as submaximal force generation, which is a critical gap in the literature. Musculoskeletal pain, especially in the context of low back pain, often manifests during physical tasks and has a significant impact on motor behavior [25,26]. Therefore, understanding facial expressions in this context is essential for improving pain assessment methodologies.

This study aims to develop and examine the performance of an automated tool for detecting facial expression variations associated with experimentally induced acute low back pain during a functional lumbar extension task. The tool integrates Haar Cascade and Local Binary Pattern Histogram subjectivity-free algorithms with self-reported pain intensity and objective force measures, including maximal voluntary contraction and force steadiness. Unlike deep learning methods, which require large amounts of data and subjective labeling for training, the techniques chosen here require neither training data nor manual annotation. This approach provides a computationally efficient method for multimodal pain-related assessment under controlled experimental conditions. The proposed method may contribute to the development of complementary objective tools for investigating pain-related motor adaptations in future clinical research.

2. Materials and Methods

2.1. Study Design

This is an experimental study that involved participants attending two visits, with a seven-day interval between them. The research protocol received ethical approval from the ethics committee of the Faculty of Sport of the University of Porto (CEFADE 28-2023). All objectives and methodologies were meticulously developed and executed in accordance

with the guidelines established by the Declaration of Helsinki. Prior to participation, all volunteers provided informed consent by signing the Informed Consent Form, ensuring their understanding of the objectives and procedures of the study.

2.2. Sample

The inclusion criteria were defined as healthy individuals aged between 18 and 40, with no history of musculoskeletal diseases or recurrent pain in the past six weeks. Additionally, participants should not have used anti-inflammatory or analgesic medication in the 24 h prior to the experiment. Exclusion criteria included individuals with a history of lumbar injury or surgery, existing chronic pathologies such as disk degeneration, osteoarthritis of the lumbar spine, facial paralysis, disk herniation, ankylosing spondylitis, and other conditions that could affect muscle function and facial expression, as well as pregnant women. These criteria were established to ensure that the participants represented a homogeneous population, thereby minimizing the impact of confounding variables on the study results.

The sample consisted of 30 young adults (average 29.96 $\pm$ 5.96 years). Among the 30 volunteers, 8 were women, with an average height of 164.66 $\pm$ 6.3 cm and an average age of 29.77 $\pm$ 6.2 years. The other 22 participants were men, with an average height of 175.20 $\pm$ 5.5 cm and an average age of 28.79 $\pm$ 5.9 years. The study included participants from Brazil and Portugal, and one participant from Cape Verde, representing both white and black ethnicities.

2.3. Digital Image Representation and Algorithm Elements

In a digital system, such as a computer, an image is represented as an array of pixels (picture elements). A digital video can be treated as a series of frames, each having this representation. When an image is in color, each pixel usually contains three numbers to represent the levels of its color components. When an image is represented in grayscale, each pixel can be represented by a single number proportional to its lightness. In grayscale, a commonly used system is one in which a completely dark pixel is represented by the number 0 and a pixel with maximum lightness by the number 255. Intermediate levels of lightness are represented between 0 and 255 [27].

2.3.1. Face Detection with Haar Cascade

Viola and Jones proposed a method for detecting objects in images, which they claimed was robust and very fast [28]. The motivation for its development was face detection, but it has more general application. This method uses a type of calculation called by the authors integral image. It is performed on grayscale images, using x and y to represent the coordinates of a pixel, with x increasing from left to right, y increasing from top to bottom in the image, and i (x, y) representing the intensity of the pixel at (x, y); the value of the integral image ii (x, y) at coordinates (x, y) is the value of the sum of all the pixel values that are in the rectangle whose diagonal runs from the upper leftmost pixel to the pixel at (x, y).

$$ii(x,y) = \sum i(x',y') \quad (1)$$

$$(x' \leq x, y' \leq y)$$

The values of ii (x, y) can be calculated progressively for all pixels in an image, moving from left to right and top to bottom. In this calculation, on each advancement to the next pixel, previous results calculated for nearby pixels can be used, avoiding having to restart the sum from the top left corner.

Once all the values of ii (x, y) for an image have been calculated, the calculation of the sum of the pixel values for any rectangle in the same image is quick, using the known

values of ii (x, y), even if this new rectangle does not have one of its ends in the upper left corner. Figure 1 provides an example of this. Considering the red rectangle, the sum of the values of its pixels is equal to the sum of those in the blue rectangle minus those in the green rectangle minus those in the white rectangle and replacing once those where the green overlaps with the white, since they were subtracted twice. In this example, using the previously calculated values of ii (x, y), this calculation would be:

$$\text{total in the red rectangle} = \text{ii}\,(7, 4) - \text{ii}\,(7, 2) - \text{ii}\,(3, 4) + \text{ii}\,(3, 2)$$

where

ii (7, 4) is the sum in the blue rectangle;
ii (7, 2) is the sum in the green rectangle;
ii (3, 4) is the sum in the white rectangle;
ii (3, 2) is the sum in the overlap of green and white.

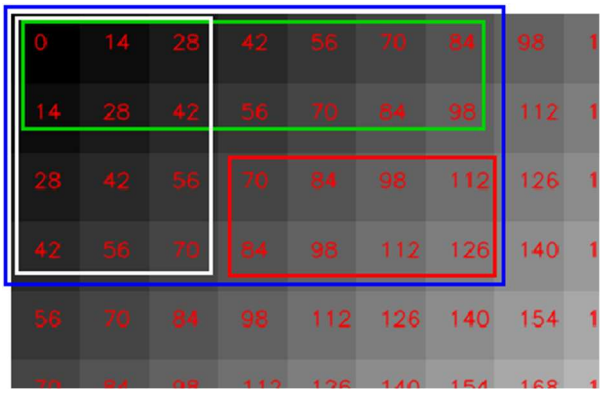

Figure 1. Example of any rectangle in the image, in red. The sum of the values of its pixels is equal to the sum of those in the blue rectangle minus those in the green rectangle minus those in the white rectangle and replacing once those in the overlap between the green and the white, since they were subtracted twice.

It can thus be seen that the sum of the pixel values in any rectangle can be quickly calculated from the previously calculated results for ii (x, y). This fact is exploited in the Viola and Jones [28] method, using ideas such as in the following example.

Each Figure 2a or Figure 2b has a grayscale face, with a rectangle positioned over it. This is a detector rectangle for a particular feature of any face: the transition region between the forehead and the eyes. It is common in images of faces, due to lighting, for the forehead to appear lighter than the eyes region. The detector rectangle is divided in half by a horizontal line. It must slide across the image, traversing it step by step. At each step, the sum *ssup* of the pixel values under the upper half of the rectangle (marked with +) is calculated and, separately, the sum *sinf* of its lower half (marked with −) is calculated. Finally, the difference $d = ssup - sinf$ is calculated. With the rectangle positioned over the forehead, the value of d will tend to be relatively small, because, since both halves are light, both *ssup* and *sinf* will have large, similar values. However, in Figure 2b, the rectangle is in a position where *ssup* is much larger than *sinf*, since the upper part is much lighter than the lower part. This makes d relatively large, indicating that this feature of a face may have been detected there.

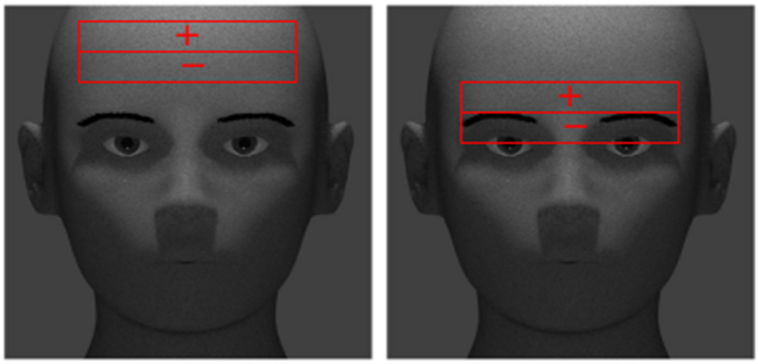

Figure 2. Example of a rectangle detecting a facial feature (transition from forehead to eyes). Legend: (**a**) Rectangle in a position without detecting the feature. (**b**) Rectangle in a position detecting the feature.

However, this rectangle, in this process, is not capable of determining on its own whether what is underneath it is, in fact, a feature of a face. This simple pattern could be found in other objects in the image. Therefore, Viola and Jones' [28] method uses many rectangles, each of which searches for a type of pattern. To detect a nose, a rectangle divided into three parts by vertical lines can be used, as in Figure 3. The calculation associated with it is the sum of the values of the pixels under its central part (marked with +) minus the sums of the values of the pixels under its lateral parts (marked with −). This takes advantage of the fact that, normally, due to lighting, the central part of the image of a nose appears brighter than its lateral parts. The higher the result of this calculation, the greater the probability that there is a nose under the rectangle.

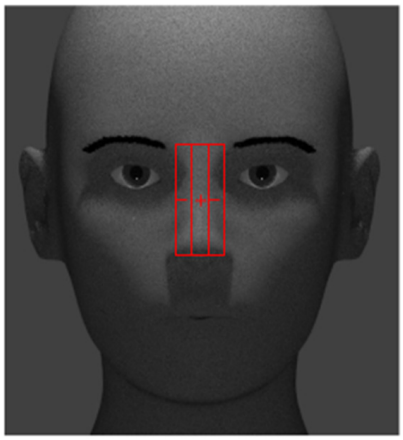

Figure 3. Example of using the nose detector rectangle.

The Viola and Jones [28] method applies several of these appropriate detector rectangles to each region of the image. The sequence of their application is planned so that first, in each region, rectangles are applied that are more reliable in declaring whether that region may contain the desired object. If high-confidence rectangles do not detect the pattern that they can detect in a region, it is decided that region does not contain the desired pattern, saving the time of applying the other rectangles. Otherwise, less reliable detector rectangles are applied successively. Finally, it is decided whether that region contains the desired pattern, considering the results obtained by all the detector rectangles applied.

The succession of rectangle applications described here, with evaluations of whether the process should continue in a region after each application, uses a strategy called "cascade". The type of calculation performed under the rectangles is similar to that of the so-called Haar functions. Therefore, this object detection process is called Haar Cascade. The Haar Cascade method is not specific to face detection; it is more generally applicable. It can be adjusted to detect other types of objects. For each type of object to be detected, appropriate detector rectangles must be used in an arrangement adjusted to the specific case. In other words, the method is general purpose, but the rectangles to be used and their arrangement in the cascade are specific. For common use cases, such as faces detection, specific arrangements of these rectangles and information about their application cascade are available as public computer files and ready for use.

2.3.2. Feature Extraction with LBPH (LBP)

This subsection briefly describes the form of representation of face images presented in [10]. Its basic element is the Local Binary Pattern. Figure 4 shows a small region containing nine pixels of a matrix that represents a grayscale image. In this set, the central pixel, with a value of 140, is taken as a reference. Then, each of its eight neighboring pixels receives a binary value (bit) associated with it, equal to 1 if this pixel has a value greater than the reference (140) and 0 otherwise. Finally, a number is formed, represented in the binary numbering system, by chaining those eight bits, as indicated by the arrows: 00011100. This value is 28 in the decimal numbering system. It is the LBP associated with the central pixel of that set. In the binary system, numbers with eight bits can be from 0 (00000000) to 255 (11111111) in the decimal system.

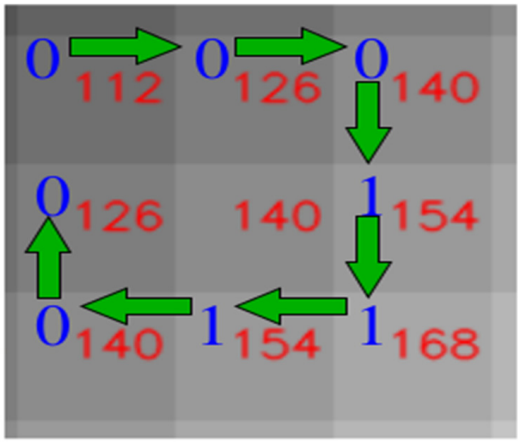

Figure 4. Example of LBP calculation for a pixel (the central one of the nine considered).

An important feature of this value is that it depends only on the results of comparing the brightness of the central pixel with its neighbors, not on these absolute values. If the image were equally brightened or darkened, this value would not change. For example, if the entire image were darkened so that each pixel value in it was reduced by 10 units, the LBP value of the central pixel would still be 28, because it would still be darker than its three neighbors below and to the right and lighter than (or as bright as) its other neighbors. The LBP is robust to variations in brightness. Another important feature is that this value is highly dependent on the relationships of the central pixel to its neighbors in terms of the pattern they form together. The brightness does not matter, but the pattern does.

When applying LBP to a grayscale image, an LBP matrix with the same geometry as the image matrix is created. The LBP value is calculated for each pixel in the image. This value is placed in the LBP matrix in the position corresponding to the position of the pixel for which the LBP was calculated in the image. This creates a new matrix, containing values between 0 and 255, each element of which contains information about the image pattern in the small region around the corresponding pixel.

As an example, suppose that the hypothetical 10 × 10 LBP matrix in Table 1 was obtained from a grayscale image of a face containing 10 × 10 pixels (this is a very small number of pixels compared to real examples, but it is easier to present here). Once such a matrix has been obtained, the next step in the representation method of [10] is to divide the matrix into regions. In this example, it has been divided into four regions (submatrices) of size 5 × 5 each, with the boundaries between them marked by the lines in the matrix in Table 1. The division could be done into submatrices of different sizes, and they would not need to be square.

Table 1. Hypothetical example of LBP matrix.

78	40	133	135	167	176	217	50	166	38
241	227	246	190	161	96	118	145	237	5
133	31	86	174	21	52	2	216	74	252
219	171	227	119	127	80	173	72	61	122
157	113	214	1	180	0	5	112	17	182
105	2	128	190	30	173	54	112	87	132
204	113	202	133	99	92	151	36	54	37
179	1	71	236	145	126	42	9	98	60
145	118	84	45	198	176	133	43	241	105
219	234	51	92	45	89	68	168	19	185

After dividing the LBP matrix, a histogram of the values that appeared in it is assembled for each submatrix, as shown in Figure 5. The number of histogram classes and the limit values of each class are freely chosen. In this example, eight classes were used for each histogram. Each class of each histogram contains a band with an amplitude of 32. The first band, centered on the value 16, has the count of occurrences of the values from 0 to 31 in the corresponding submatrix. The second band is centered on 48 and contains the count of occurrences of the values from 32 to 63, and so on. In Figure 5, the histograms are in positions corresponding to the positions of the four submatrices in Table 1 from which they originated. Finally, the histograms are concatenated (placed side by side), as shown in Figure 6. Since each of the four histograms had eight classes, the result of the concatenation is a sequence containing 4 × 8 = 32 numbers:

$$(3, 1, 2, 3, \ldots, 4, 0, 1)$$

This is the result of the LBPH (Local Binary Pattern Histogram). This sequence of numbers (32, in this example) describes patterns of relationships between pixels in an image. It can then be used by machine learning algorithms, or other, as a representation of the image, bringing important information about it.

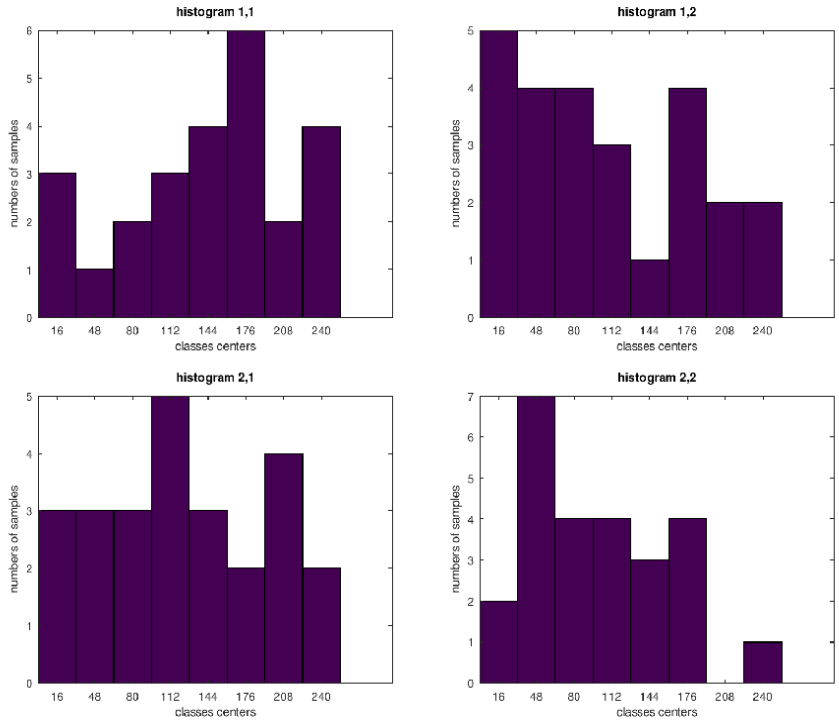

Figure 5. Histograms of the occurrences of values in the four submatrices of Table 1.

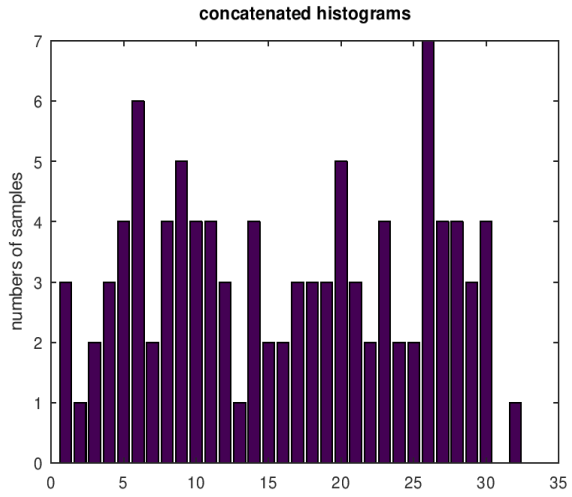

Figure 6. Result of concatenating the histograms in Figure 5.

2.3.3. Development of Software for Analysis of Facial Expressions with Experimentally Induced Acute Low Back Pain

A computer program was developed in Python language version 3.9, using the OpenCV library version 4.5.1 [29]. This has many algorithms implemented for image processing. The program must access two video files, in which the face of the person whose pain is to be assessed appears. In the first video, which is relatively short, the face should appear with a preparation phase of expression. This is called the learning video by the program. It is used so that the program can learn what the preparation phase of expression of that person is. In the second video, called the application video, the program measures, frame by frame, the difference between the facial expression in the frame and the average expression learned in the frames of the learning video. The program interface window is shown in Figure 7. It was developed using the Python package tkinter [30].

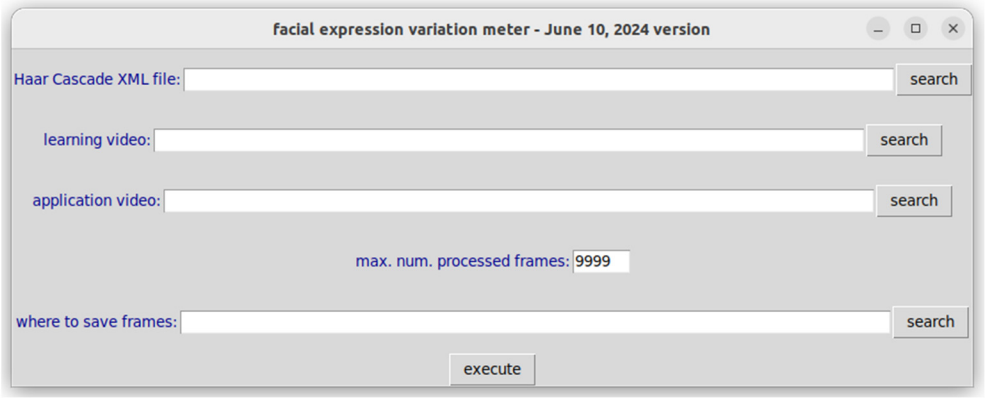

Figure 7. Program interface window.

The first information requested from the program user in this window is the name and location on the computer of a file in XML format that contains the details of the face detection by Haar Cascade, as discussed at the end of Section 2.3.2. Next, the files with the videos for learning and for application are requested. Then, the user is asked for the limit number of frames of the video for application to be processed. This number is useful when the user wants to do a quick test of the processing by entering a small value there. Finally, the user is asked for the location on the computer where the files with the results will be saved. One of these files is an annotated video, frame by frame, with a number proportional to the difference between the expression detected in it and the average expression learned in the video for learning. The other is a file with a record, in plain text format, of the number of each frame next to the result of the measurement of the difference in expression between that of the face in it and the average learned. These numbers can then be entered into spreadsheets or provided to other programs for analysis and visualization of the evolution of the difference over time.

Algorithm 1 is a representation of the algorithm in pseudocode.

Face detection in video frames is performed by an object of the Cascade Classifier class of the OpenCV library [29], named face_cascade, configured to use the XML file indicated by the user in the program interface window. During the learning phase (in Algorithm 1, lines 5 to 25), the program processes each frame of the first video by converting it to grayscale and attempting to detect faces using the face_cascade detector. If no face is detected in each frame, the program proceeds to the next one. If multiple faces are detected, processing continues only for the face with the largest area, as this is likely the primary subject, while smaller faces are assumed to belong to people passing through the background, images on the walls, or other irrelevant elements. Once the largest face is isolated by the face_cascade detector, a bounding rectangle is drawn around it. This rectangle is then divided into 9 rows and 8 columns, as illustrated in green in Figure 8. The choice of this division (9 rows and 8 columns) was determined through empirical testing conducted by the authors using the JAFFE (Japanese Female Facial Expression) facial expression database [8,31]. Since the leftmost column and the rightmost column have several rectangles that contain parts of the image that are not on the face, they are not used in the further processing, as they may suffer spurious effects from the scene background.

Algorithm 1. Facial Expression Difference Meter

1: Receive the names of the video files from the user.
2: numCols ← 8
3: numLins ← 9
4: Configure Haar Cascade using XML file.
Learning phase:
5: Read video file to learning.
6: numHistLearn ← 0
7: histConcatLearnSum ← zeros
8: while there are frames in the video file do
9: frame ← video frame
10: imgGray ← Convert frame to grayscale.
11: faces ← coordinates of rectangles having faces
12: if number of faces > 0 then
13: numHistLearn ← numHistLearn + 1
14: imgAFace ← rectangle delimiting the biggest face
15: smallRects ← cut imgAFace in numCols × numLins smaller rectangles
16: for rect in smallRects from the second to the penultimate column do
17: LBPrect ← LBP of rect
18: histRect ← histogram of LBPrect
19: histNorm ← normalize histRect
20: end for
21: histConcat ← concatenation of histNorm built in the for loop above
22: histConcatLearnSum ← histConcatLearnSum + histConcat
23: end if
24: end while
25: histConcatLearnAverage ← histConcatLearnSum/numHistLearn
Application phase:
26: Read video file to application phase.
27: while there are frames in the video file do
28: frame ← a video frame
29: imgGray ← Convert frame to grayscale.
30: faces ← coordinates of rectangles having faces
31: if number of faces > 0 then
32: imgAFace ← rectangle delimiting the biggest face
33: smallRects ← cut imgAFace in numCols × numLins smaller rectangles
34: for rect in smallRects from the second to the penultimate column do
35: LBPrect ← LBP of rect
36: histRect ← histogram of LBPrect
37: histNorm ← normalize histRect
38: end for
39: histConcat ← concatenation of histNorm built in the for loop above
40: modDif ← average of the elements ($|\text{histConcat} - \text{histConcatLearnAverage}|$)
41: Show 1000 × modDif
42: end if
43: end while

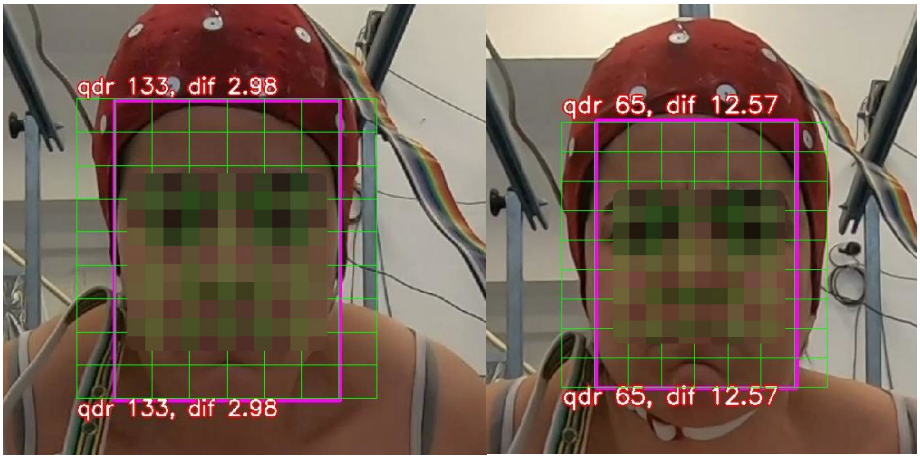

Figure 8. Rectangles containing faces detected by Haar Cascade divided into rows and columns (in green). Only the part delimited in violet is used in the following processing. Legend: (**left picture**) pain-free, (**right picture**) with low back pain.

Following the process outlined in Section 2.3.3, the Local Binary Pattern (LBP) is computed for each of the remaining 9 × 6 = 54 small rectangles. The LBP implementation used is from the skimage package in Python [1]. The LBP histogram for each small rectangle is created using the NumPy package [30]. Each histogram consists of 30 classes and is normalized such that the sum of its values equals 1, irrespective of the area covered by the rectangle. Once the LBP histograms for all small rectangles are assembled and normalized, the 54 histograms are concatenated, resulting in a sequence containing 54 × 30 = 1620 numbers. This sequence serves as the Local Binary Pattern Histogram (LBPH) representation of the largest face detected in the video frame.

Once the learning video is finished, the average LBPH representation of its frames is calculated, according to the expression:

$$reprLBPH_{neutral} = \frac{\left(\sum frames\ with\ faces\right) * \left(reprLBPH_{frame}\right)}{number\ of\ frames\ with\ faces} \quad (2)$$

Thus, reprLBPHpreparation phase is a sequence containing 1620 numbers and represents the average face from the learning video, specifically corresponding to the preparation phase expression of the individual. During the application phase (in Algorithm 1, lines 26 to 43), similar steps are followed for each frame of the video: the image is converted to grayscale, the largest face is detected, the face is divided into small rectangles, and histograms are assembled and concatenated. This results in a sequence of 1620 numbers, denoted as reprLBPHframe, associated with the face in that particular frame. The difference between this sequence and reprLBPHpreparation phase is then calculated to quantify how much the face in the current frame deviates from the preparation phase of expression.

$$\text{diffLBPHframe} = \text{reprLBPHframe} - \text{reprLBPHpreparation phase} \quad (3)$$

This difference is calculated using the mean_absolute_error function from the Python sklearn package [32]. To understand how this calculation works, here is a generic numerical example. Given the following sequences of numbers:

$$a = (10, 20, 30, 40, 50)$$

and

$$b = (1, 1, 2, 2, 3)$$

where a is in the role of the sequence reprLBPHframe and b is in the role of the sequence reprLBPHpreparation phase, the difference between a and b will be:

$$d = (9, 19, 28, 38, 47),$$

The difference is calculated position by position: the first element in d is the absolute difference between the first element in a and the first in b, the second element in d is the difference between the second in a and the second in b, and so on. After calculating the absolute values of these differences, the average of the numbers in d is computed, which in this case is 28.2. This type of calculation follows expression (3), producing a number that indicates how much the face in a frame deviates from the average preparation phase of the face. Experimentally, it was observed that the result is typically on the order of thousandths or hundredths. To make the output more readable, the number shown in the result video and the corresponding text file is the result of multiplying difLBPHframe by 1000. This scaling ensures that the value presented to the user is in the range of units of tens, rather than thousandths or hundredths. The number does not have a physical unit of measurement, such as meters or seconds, since it represents an average of counting results of pure numerical differences. It is important to emphasize that the software does not implement a threshold-based classification of facial expressions, for example, as "pain" or "no pain." Instead, it generates a continuous quantitative index representing the magnitude of expressive deviation relative to the individual's baseline (preparation phase).

Part of the expected time for the algorithm to execute on each video frame is due to the execution of the Haar Cascade. Due to its characteristics, this is difficult to analyze in terms of conventional algorithm complexity analysis. However, in their original paper, Viola and Jones (2001) [28] reported that operating on 384 by 288 pixel images, faces were detected at 15 frames per second on a conventional (of that time) 700 MHz Intel Pentium 111. Still on the complexity of the algorithm, it has "while" loops, but each one of them treats one frame at a time in the entire video, without affecting the analysis time per frame. The analysis time per frame in the application phase is dominated by the "for" loop on line 34 in Algorithm 1. The dependence of the execution time on the pixel resolution of the frame for a fixed number of rectangles is quadratic, because, if, for example, the resolution doubles both horizontally and vertically, the number of pixels in a rectangle will be four times greater, which will make the time to assemble the histogram for each rectangle four times greater.

2.4. Experimental Protocol

The experimental protocol was designed to assess the effects of acute low back pain induction through intramuscular injections of hypertonic and isotonic saline solutions. The randomized, crossover design allowed participants to be exposed to both pain (hypertonic) and control (isotonic) conditions. The evaluation included measures of facial expression, muscle strength, and self-reported pain, providing a comprehensive analysis of both physiological and subjective responses to acute low back pain.

2.4.1. Experimentally Induced Acute Low Back Pain

Experimental acute low back pain was induced by intramuscular injections of hypertonic and isotonic saline solution, administered 7 days apart. This interval was based on previous studies indicating that a 7-day washout is adequate to prevent carry-over effects in crossover pain models [33,34]. During the injections, the participants were positioned in the prone position with a 20° inclination, which helped to promote muscle relaxation. The study used a randomized crossover design, in which each participant received a single dose of 2.5 mL [35], to ensure that the test was performed during the period of pain, of hypertonic saline solution (6.0%), designed to induce pain, or isotonic saline solution (0.9%) [36], used

as a placebo. The injections were administered in the intervertebral space between L3 and L4, at the level of the spinous process, located approximately 2 to 3 cm from the midline. After cleaning the area with 70% alcohol, the solution was injected into the dominant side of the body using a 25 × 28 mm needle, at a depth of 30 mm from the skin surface, by a qualified professional. The participants were assessed in three different conditions: baseline, experimentally induced pain (hypertonic solution), and control condition (isotonic solution) while performing a spinal extension task in the prone position.

2.4.2. Force Test

Force steadiness testing was included in the experimental protocol to investigate the association between facial expressions and motor control during physical tasks, focusing on the force capacity and stability of the trunk extensor muscles during spinal extension. Participants first completed a familiarization phase, consisting of five repetitions of the maximal voluntary contraction (MVC) test and three repetitions of a submaximal isometric test using a trapezoidal force signal to assess steadiness [37–39]. After a 2 min rest, one MVC trial was performed, followed by a 1 min rest and one repetition of the submaximal trapezoidal force signal test. The protocol consisted of two assessments conducted by an experienced physiotherapist:

- Test 1 (MVC) Assessing maximal force capacity.
- Test 2 (Submaximal isometric contraction at 20% of MVC; assessing steadiness).

These tests were administered during two visits, separated by a 7-day interval, with the same procedures repeated on both occasions. Familiarization ensured consistency and minimized learning effects on performance [39].

2.4.3. Evaluation of Facial Expressions

Facial expressions were captured by a GoPro Hero 8 Black camera (San Mateo, CA, USA) with 1080p resolution at 60 frames per second. Ambient lighting was standardized across all sessions, using fixed ceiling fluorescent lights with consistent intensity and color temperature. The distance and angle of the camera, as well as participant positioning, were maintained constant to ensure uniform conditions during data acquisition. The videos were analyzed using the OpenCV library in the Python programming language [40].

The participant's faces were recorded in video frontal view while they performed a spinal extension in prone position, sustained for 20 s with a submaximal load corresponding to 20% of the maximum force. To ensure that the face remained aligned and visible during all experimental conditions, participants were instructed to fix their gaze on a computer screen positioned directly in front of them and to concentrate on maintaining the trapezoidal force signal projected on the screen for the entire 20 s period. In addition, Participants were instructed to maintain silence throughout the procedure to prevent potential interference in the analysis of facial expressions and to ensure unbiased detection of spontaneous pain expressions [6,41,42]. All participants were informed in advance about the recording and provided written informed consent authorizing the capture and use of the images.

To determine specific expressive variations associated with pain, the software calculates, frame by frame, the mean absolute difference between each facial representation (LBPHframe) and the average pattern extracted from the preparation phase video (LBPH-preparation phase). This value is then multiplied by a factor of 1000, generating a numerical index proportional to expressive intensity. A consistent increase in this index, especially during the induced pain condition, was interpreted as indicative of pain-related facial activation. These values were jointly analyzed with muscle force and self-report data, and the method's sensitivity was empirically validated by the distinct differences observed across experimental conditions (pain, placebo, and pain-free), as illustrated in Figure 8.

2.4.4. Evaluation of Subjective Responses to Pain

On a desktop computer with Intel(R) Core (TM) i5 CPU M 520 @ 2.40GHz, 2 cores, 4 Gbytes RAM, with Debian GNU/Linux 11 operating system, in the algorithm application phase, the average analysis time per frame was 0.97 s. Under these conditions, in order for all frames of a video to be analyzed, a file with this video must be recorded and the analysis then performed on the finished video, not in real-time. In general, for real-time analysis, only 1 at each N frames could be processed, with N being chosen by the user, reducing the temporal resolution.

2.4.5. Evaluation of Subjective Responses to Pain

Pain intensity ratings were obtained immediately after each stimulus using the Numeric Rating Scale (NRS) [3], a widely validated and sensitive method for assessing subjective pain perception. The scale was standardized with extreme points labeled 0 = "no pain" and 10 = "extremely severe pain." Participants received standardized verbal instructions from the same researcher and self-reported their pain intensity, ensuring consistent comprehension and execution. Assessments were conducted at 1 min intervals over a 7 min period, ensuring reliable and consistent data collection. The NRS is recognized for its simplicity and effectiveness in detecting subtle changes in pain perception, contributing to the robustness of the study results.

2.5. Statistical Analysis

Descriptive statistics were calculated for the NRS and facial expression measurements recorded during each experimental condition and session. Normality of the data was assessed using the Shapiro–Wilk test, which indicated non-normal distributions; therefore, non-parametric tests were applied. To compare the three experimental conditions within the same participants (baseline/no pain, placebo injection, and hypertonic saline injection/pain), the Friedman test was used. Pairwise comparisons were performed using the Wilcoxon signed-rank test with Bonferroni correction to control for multiple comparisons. In addition, Wilcoxon tests were applied to examine sex-related differences in facial expression changes within each condition.

The intraclass correlation coefficient (ICC 3,1) was calculated using a two-way mixed-effects model with absolute agreement to assess the consistency of software-based facial expression measurements across the two experimental days (no pain condition), as recommended for test–retest reliability. According to Koo and Li (2016), ICC values below 0.50 indicate poor reliability, 0.50–0.75 moderate, 0.75–0.90 good, and above 0.90 excellent [43]. Cohen's d was calculated as a measure of effect size for all comparisons and interpreted as small ($d \geq 0.2$), moderate ($d \geq 0.5$), or large ($d \geq 0.8$). These effect magnitudes suggest adequate statistical sensitivity to detect the reported differences. All analyses were performed using SPSS Statistics (IBM, Version 27) with a significance level of $\alpha = 0.05$.

3. Results

The test–retest reliability of facial expression measurements was assessed based on data obtained on two consecutive days in the pain-free condition. The means (±SD) were 4.98 ± 2.96 on Day 1 and 4.92 ± 2.89 on Day 2, with no statistically significant difference between the days ($p = 0.28$). The intraclass correlation coefficient (ICC = 0.85; 95% CI = 0.68–0.97) indicated good reliability between sessions. According to the 95% confidence interval, reliability ranged from moderate to excellent, reflecting some variability in the precision of the estimate.

3.1. Measuring Facial Expressions

The videos were analyzed using the OpenCV library in the Python programming language [40]. The results (Figure 9) indicated statistically significant differences ($p < 0.001$) between all conditions, confirming the effectiveness of the software in detecting variations in facial expressions. The analysis revealed that facial expression in men showed a greater mean difference during pain compared to the placebo and pain-free conditions, with significant differences observed between all conditions ($p < 0.01$). Women also showed an increase in facial expression during pain, although the effect was smaller than in men, with significant differences observed between the pain, pain-free and placebo conditions ($p < 0.05$). However, when directly comparing the male and female groups, no statistically significant differences were observed between the sexes in any condition. Furthermore, Cohen's d value of 0.588 suggests a moderate effect size, further emphasizing the robustness of the results given the sample size [44].

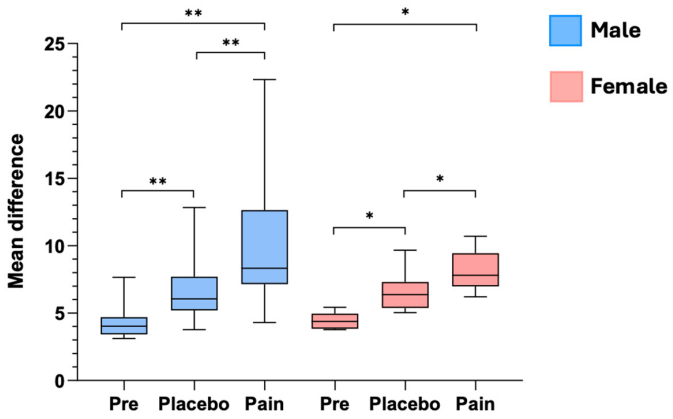

Figure 9. Graph of the mean difference in facial expressions at pre (pain-free), placebo, and during experimentally induced acute low back pain. Legend: ** $p < 0.001$; * $p < 0.05$.

Force Test

When maximal voluntary contraction was analyzed, male and female participants showed significant intra-group differences between the experimental conditions (Figure 10a). For women, the MVC was highest during preparation, followed by placebo and the pain period, with the Friedman test revealing a significant effect ($\chi^2 = 16.00$, $p < 0.001$). Among men (N = 22), MVC was slightly higher during the placebo period compared to stimulation, being the lowest during the control period, also showing significant differences ($\chi^2 = 32.818$, $p < 0.001$). In addition, more detailed analyses indicate that despite these intra-group variations, no significant differences were found between genders, suggesting that both sexes showed similar responses to experimental conditions in terms of changes in MVC.

Analysis of the steadiness figure (Figure 10b) revealed distinct patterns of response to the experimental conditions. Friedman's test indicated statistically significant differences between the conditions ($\chi^2 = 44.000$, $p < 0.001$), suggesting robust variations in steadiness under different experimental influences. Comparisons between conditions revealed that steadiness was significantly lower during the pain condition compared to pre and placebo for both sexes, indicating a pronounced impact of the pain condition on motor stability. In men, steadiness was significantly higher during the placebo condition compared to both baseline (pre) and pain conditions. Additionally, baseline steadiness was significantly higher than during the pain condition, which presented the lowest values. These results suggest that the experimental condition modulated force steadiness, reflecting a progressive impairment in neuromuscular control from placebo to baseline, and most markedly under pain.

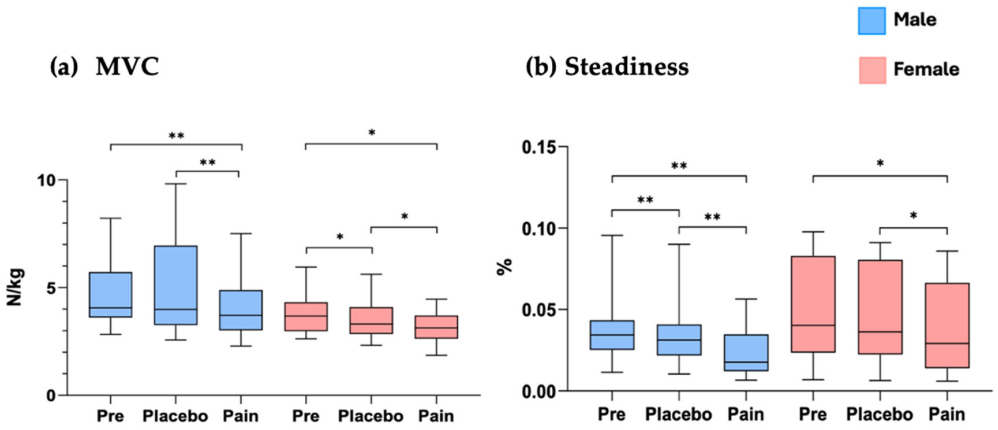

Figure 10. Maximum strength and steadiness under pre, placebo and pain conditions in both men and women. Legend: Pre (pain-free), placebo and during pain, ** $p < 0.001$; * $p < 0.05$.

3.2. Gender Differences in NRS Responses During Acute Low Back Pain

On the NRS, men reported pain as more intense compared to women, on a scale of 0 to 10. The standard deviation is low for both groups, but is lower for women, suggesting that there is less variation in pain ratings among women than among men, as seen in Figure 11. Analysis of the NRS revealed significant differences in pain perception between the experimental conditions and between sexes. For the total group of 30 participants, both men and women reported significantly greater pain during the pain condition compared to the placebo condition. However, a more detailed analysis showed significant differences between the sexes. For men, the average NRS with pain was 5.12, significantly higher than the placebo NRS average of 1.68 ($Z = -4.108$, $p < 0.001$). In contrast, for women, the mean NRS with pain was 3.75, while the mean NRS for placebo was 1.25 ($Z = -2.58$, $p = 0.01$). These results indicate not only that the intensity of perceived pain is higher during the pain condition compared to placebo, but also that there are significant differences in the way men and women report pain. Men's reported verbal pain perception was consistently higher than women's, suggesting that gender influences the intensity of reported pain. These results emphasize the importance of considering both experimental variables and gender differences when assessing and interpreting pain in clinical studies.

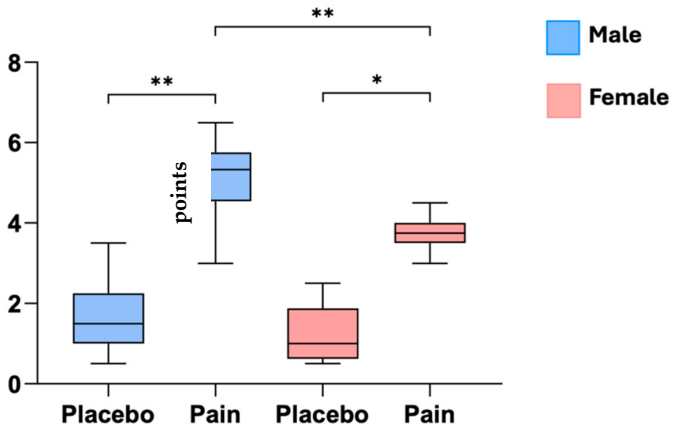

Figure 11. NRS results obtained under pre, placebo and pain conditions in both men and women. Legend: Numeric rating scale (NRS) 0–10, ** $p < 0.001$; * $p < 0.05$.

4. Discussion

The innovation of this study lies in the combination of digital imaging techniques with image processing algorithms for analyzing facial expressions, offering a more objective approach to the understanding of pain under controlled experimental conditions. The soft-

ware developed captures changes in facial expressions and integrates them with physical (MVC and steadiness) and subjective responses to pain (NRS) in both male and female subjects. The software was able to detect variations in facial expressions during episodes of experimentally induced acute low back pain compared to pre and placebo conditions in males and females, providing a multidimensional view of the pain experience.

The recent literature has focused on the application of deep learning models for the automatic estimation of facial pain intensity, exploiting both spatial and temporal feature extractions to capture static and dynamic aspects of pain-related facial expressions [12,13]. Egede et al. [13] proposed a hybrid model that integrates representations extracted by deep convolutional neural networks with manual descriptors of action units, obtaining a root mean square error of less than one point on a 16-level scale and a Pearson correlation of 0.673 with clinical pain scores [12]. A real-time regression framework was developed based on a recurrent convolutional neural network to estimate pain intensity in facial videos, considering image sequences aligned by AAM and achieving promising results in terms of accuracy and speed on the UNBC-McMaster shoulder pain expression archive dataset.

Other studies have provided a solid basis for validating these algorithms [9] made available the UNBC-McMaster shoulder pain expression archive database, which has established itself as an international reference for benchmarking facial pain estimation. In addition, the psychometric properties of facial pain expressions in patients with shoulder pain were examined, demonstrating the concurrent validity and test–retest reliability of an action units-based index (ICC = 0.87; 95% CI: 0.79–0.92) [45]. In contrast, the present study adopts classic computer vision approaches, using Haar Cascade for face detection and LBPH for texture characterization, integrated with an optimized image pre-processing and normalization flow. This results in a significant reduction in average processing time, estimated at less than 50 ms per frame on conventional processors, while maintaining interpretability and robustness. The test–retest reliability results (ICC = 0.85; 95% CI: 0.68–0.97) indicate good reliability; however, the relatively wide confidence interval suggests some degree of uncertainty, likely associated with the modest sample size, and therefore these findings should be interpreted with appropriate caution.

The results also explored potential differences in pain reporting between sexes. Although men reported higher subjective pain levels on the NRS compared to women, this perception was not consistently reflected in changes in facial expressions. Both men and women exhibited similar facial expressiveness during experimentally induced acute low back pain, as well as in pre-pain and placebo conditions. These findings align with previous studies indicating no significant sex differences in facial expressiveness of pain [6,46]. However, given the relatively small female subsample, any interpretation of sex-related patterns should be considered exploratory. Although men reported significantly higher pain levels on the NRS compared to women, contrasting with findings from previous studies [14,20], this result suggests that sex influences the relationship between self-reported pain and facial expression of emoticons [47,48]. Thus, gender may influence the relationship between self-reported pain and facial expressions, highlighting the need for tailored approaches in pain management for men and women [49].

Another interesting point not previously addressed in the literature [14,20,50,51] is the integration of changes in facial expressions during pain with reductions in muscle strength and steadiness. The expressive changes in facial dynamics captured by the software were frequently associated with a decrease in both muscle strength and sustained contraction stability in both sexes. This association was particularly pronounced during the acute pain condition, where greater variability in steadiness was observed, alongside statistically significant differences in MVC between the pain and placebo conditions ($\chi^2 = 32.8$, $p < 0.001$ for males; $\chi^2 = 16.0$, $p < 0.001$ for females). Taken together, these findings support the

hypothesis that changes in facial expressions during episodes of experimentally induced acute low back pain may reflect physiological components underlying the pain experience, especially when considered in conjunction with reductions in motor performance, rather than constituting an established clinical biomarker at this stage [16].

The motor areas at the frontoparietal level are intrinsically linked to pain processing, integrating both sensory and affective components, thus emphasizing the complex interaction between pain perception and motor response [7]. These regions may suggest common mechanisms in prediction and emotional regulation of pain [7,16]. Recruitment of the frontoparietal network may also play a role in the conscious representation of the pain context, making nociceptive information accessible for integration into decision-making processes [16,52–54]. This link is reflected in both facial expressions and subjective self-reports of individuals, suggesting that motor areas not only contribute to the perception and physical response to pain but also play a role in emotional modulation [54,55].

Systematic analyses have consistently demonstrated that high levels of pain-related fear, catastrophizing, and depression are linked to a reduction in spinal range of motion and an increase in trunk muscle activity in individuals with low back pain, with these associations occurring independently of pain intensity [53,54,56]. However, the small effect sizes question the relevance of these psychological factors as major determinants of a stiffer spinal motor behavior. Furthermore, the literature [56,57] has shown very small associations between pain intensity and decreased spinal amplitudes, as well as increased trunk muscle activity. These findings suggest that future research should incorporate specific and individualized measures for psychological factors, pain intensity, and spinal motor behavior [56]. The integrated analysis of facial expressions captured through our software, subjective responses, and strength data conducted in our study contributes to capturing subtle aspects of the pain experience that motor measures alone may not fully represent.

Therefore, it is important for health professionals to recognize that self-reporting in response to pain can vary substantially between individuals due to a combination of biopsychosocial factors, with gender being only one of the multiple determinants [58,59]. Although men and women displayed similar facial expression patterns during experimentally induced acute low back pain in this sample, differences were observed in subjective pain reports. To develop innovative therapeutic approaches for individuals suffering from low back pain, it is crucial to assess pain across multiple dimensions [59]. Future studies should further investigate the relationship between facial expression changes, psychological factors, and peripheral and central neural mechanisms in motor control, particularly through steadiness during spinal movement in individuals with low back pain [38,57]. These elements may contribute to a better understanding of pain-related motor adaptations and inform future research directions.

4.1. Limitations

An important limitation of this study is the relatively small sample size of 30 participants, which may limit the generalizability of the results to wider and more diverse populations. An additional limitation is the gender imbalance in the sample, with only eight women (26.7%). This distribution reflects a common challenge in human experimental pain studies, as the previous literature indicates that male participants are often more likely to volunteer for such studies, leading to sex/gender sampling bias [60–62]. Consequently, findings related to sex differences should be interpreted with caution and may not fully generalize to populations with a more balanced gender representation.

Furthermore, the sample was restricted to young, healthy individuals with no history of chronic conditions or musculoskeletal pathologies, which reduces the applicability of the findings to clinical populations with chronic low back pain or other comorbidities. Another

limitation refers to the experimental nature of the induced pain, which may not accurately replicate the complexity and variability of pain experienced in real clinical scenarios. Finally, despite the use of advanced software to detect facial expressions, automated analysis still depends on external factors, such as the positioning and quality of the camera (number of pixels), which can influence the accuracy of the data captured. The computer program presented here measures differences between facial expressions, regardless of their nature. Therefore, its results regarding pain measurement only make sense if the variations in facial expression are due solely to the pain itself, not to laughter or other causes. It was also not designed to detect pain feigning. Therefore, for the purpose presented here, its use should only occur in controlled situations.

4.2. Future Research in Clinical Practice

Future research utilizing the software developed in this study could explore innovative avenues to enhance pain assessment and management in clinical and rehabilitation contexts. One promising direction is applying the software to clinical populations experiencing chronic pain, such as persistent low back pain, to compare facial expressions observed in experimental acute pain scenarios with chronic pain patterns. Additionally, using larger samples, diverse clinical populations, and more widely adopted experimental pain models will be crucial to validate and extend our findings. In clinical settings, the software may be tested as a complementary tool during functional assessments and therapeutic exercises, potentially supporting monitoring of pain exacerbation, adjustment of exercise intensity, and longitudinal tracking of pain-related motor adaptations. For practical implementation in routine care, computational efficiency and real-time processing capability will be essential.

In this context, technical improvements may include implementation on higher-performance hardware, graphics processing unit acceleration, and algorithmic optimizations such as parallel processing and resolution adjustment, which could substantially reduce processing time and facilitate integration into clinical environments. The integration of additional biomarkers, such as electromyography data, may further broaden the scope of the software toward a multimodal pain assessment model. Another promising direction involves the exploration of hybrid or deep learning-based architectures to complement the current subjectivity-free framework. Advanced feature extraction and pooling strategies, such as singular pooling, could enhance the preservation of discriminative facial features and improve sensitivity to subtle pain-related facial dynamics [63]. Although the present study intentionally employed classical computer vision techniques to avoid training dependency and subjective labeling, subsequent implementations incorporating such strategies may further optimize performance in larger-scale datasets and real-time clinical applications.

5. Conclusions

The developed software demonstrated sensitivity to subtle variations in facial expressions during experimentally induced acute low back pain in a fully automated and subjectivity-free manner. Unlike deep learning approaches that require large, labeled datasets for training, the proposed method operates without prior training data or manual annotation. The system detected significant differences across pain, placebo, and baseline conditions and showed consistent associations with subjective pain reports and force steadiness measures. This methodological approach provides a promising complementary metric for objectively quantifying pain-related expressive changes under controlled conditions. Further validation in larger and clinical populations is warranted to determine its translational applicability.

Supplementary Materials: Software_ Method for Quantifying Pain-Related Facial Expressions, https://doi.org/10.6084/m9.figshare.29330033.

Author Contributions: Conceptualization, F.P. and R.S.; methodology, F.P., M.F.G., U.F.E. and J.P.V.-B.; software, F.P.; validation, F.P., R.P., S.D.d.S. and U.F.E.; formal analysis, F.P.; investigation, F.P., R.P. and S.D.d.S.; resources, R.S.; data curation, F.P.; writing—original draft preparation, F.P.; writing—review and editing, F.P., R.S., M.F.G., U.F.E., J.P.V.-B. and K.B.; visualization, F.P.; supervision, R.S.; co-supervision, M.F.G., U.F.E. and J.P.V.-B.; project administration, R.S.; funding acquisition, R.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Rehabilitation Research Center-Foundation for Science and Technology (FCT) through R&D Units funding UI/BD/151415/2021.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and the research protocol received ethical approval from the Ethics Committee of the Faculty of Sport, University of Porto (CEFADE 28-2023, 12 July 2023).

Informed Consent Statement: Informed consent was obtained from all individual participants included in the study. Informed consent for publication was obtained from all identifiable human participants.

Data Availability Statement: The data presented in this study are available on request from the corresponding authors. The data are not publicly available due to privacy. The complete source code of the software will be provided as Supplementary Materials.

Acknowledgments: The authors thank the professors and colleagues for their support and contributions throughout the development of this study. This manuscript is partially based on the Ph.D. thesis of Franciele Carvalho Santos Parolini, entitled "Exploring Acute Low Back Pain: Connections Between Neuroscience, Biomechanics, and Psychology", submitted to the Universidade do Porto, Porto, Portugal, in 2024. The authors also acknowledge the scientific collaboration of the Pain Research Group (School of Arts, Sciences and Humanities, University of São Paulo) and the support provided by LABIOMEP (Porto Biomechanics Laboratory) and CIR (Rehabilitation Research Centre), particularly regarding technical and infrastructure resources.

Conflicts of Interest: The authors have no conflicts of interest to declare.

Abbreviations

The following abbreviations are used in this manuscript:

AAM	Active Appearance Model
AI	Artificial Intelligence
CC BY	Creative Commons Attribution License
CEFADE	Ethics Committee of the Faculty of Sport, University of Porto
CIR	Center for Rehabilitation Research
CNN	Convolutional Neural Network
ICC	Intraclass Correlation Coefficient
JAFFE	Japanese Female Facial Expression Database
LBP	Local Binary Pattern
LBPH	Local Binary Pattern Histogram
MVC	Maximal Voluntary Contraction
NRS	Numeric Rating Scale
OpenCV	Open Source Computer Vision Library
SD	Standard Deviation
SPSS	Statistical Package for the Social Sciences

References

1. Price, D.D.; McGrath, P.A.; Rafii, A.; Buckingham, B. The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. *Pain* **1983**, *17*, 45–56. [CrossRef]
2. Hadjistavropoulos, T.; Craig, K.D.; Duck, S.; Cano, A.; Goubert, L.; Jackson, P.L.; Mogil, J.S.; Rainville, P.; Sullivan, M.J.L.; Williams, A.C.C.; et al. A biopsychosocial formulation of pain communication. *Psychol. Bull.* **2011**, *137*, 910–939. [CrossRef]
3. Ferreira-Valente, M.A.; Pais-Ribeiro, J.L.; Jensen, M.P. Validity of four pain intensity rating scales. *Pain* **2011**, *152*, 2399–2404. [CrossRef]
4. Werner, P.; Al-Hamadi, A.; Limbrecht-Ecklundt, K.; Walter, S.; Gruss, S.; Traue, H. Automatic Pain Assessment with Facial Activity Descriptors. *IEEE Trans. Affect. Comput.* **2016**, *8*, 286–299. [CrossRef]
5. Hammal, Z.; Cohn, J.F. Automatic detection of pain intensity. In Proceedings of the 14th ACM International Conference on Multimodal Interaction, New York, NY, USA, 26 October 2012; Volume 2012, pp. 47–52. [CrossRef]
6. Prkachin, K.M. The consistency of facial expressions of pain: A comparison across modalities. *Pain* **1992**, *51*, 297–306. [CrossRef] [PubMed]
7. Picard, M.-E.; Kunz, M.; Chen, J.-I.; Coll, M.-P.; Vachon-Presseau, É.; Wager, T.D.; Rainville, P. A distributed brain response predicting the facial expression of acute nociceptive pain. *eLife* **2024**, *12*, RP87962. [CrossRef] [PubMed]
8. Lyons, M.; Kamachi, M.; Gyoba, J. Coding Facial Expressions with Gabor Wavelets (IVC Special Issue). *arXiv* **2020**, arXiv:2009.05938v1. [CrossRef]
9. Lucey, P.; Cohn, J.F.; Matthews, I.; Lucey, S.; Sridharan, S.; Howlett, J.; Prkachin, K.M. Automatically detecting pain in video through facial action units. *IEEE Trans. Syst. Man Cybern. Part B* **2010**, *41*, 664–674. [CrossRef]
10. Ahonen, T.; Hadid, A.; Pietikäinen, M. Face Recognition with Local Binary Patterns. In Proceedings of the 8th European Conference on Computer Vision (ECCV 2004), Prague, Czech Republic, 11–14 May 2004; Volume 3021, pp. 469–481.
11. Florea, C.; Florea, L.; Vertan, C. Learning pain from emotion: Transferred hot data representation for pain intensity estimation. In *Proceedings of the Computer Vision-ECCV 2014 Workshops, Zurich, Switzerland, 6, 7, 12 September 2014, Proceedings, Part III 13*; Springer: Cham, Switzerland, 2015; pp. 778–790.
12. Zhou, J.; Hong, X.; Su, F.; Zhao, G. Recurrent convolutional neural network regression for continuous pain intensity estimation in video. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 84–92.
13. Egede, J.; Valstar, M.; Martinez, B. Fusing deep learned and hand-crafted features of appearance, shape, and dynamics for automatic pain estimation. In Proceedings of the 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017), Washington, DC, USA, 30 May–3 June 2017; pp. 689–696.
14. Schneider, P.; Lautenbacher, S.; Kunz, M. Sex differences in facial expressions of pain: Results from a combined sample. *Pain* **2024**, *165*, 1784–1792. [CrossRef]
15. Sneddon, L.U. Evolution of nociception and pain: Evidence from fish models. *Philos. Trans. R. Soc. B Biol. Sci.* **2019**, *374*, 20190290. [CrossRef]
16. Zheng, W.; Woo, C.-W.; Yao, Z.; Goldstein, P.; Atlas, L.Y.; Roy, M.; Schmidt, L.; Krishnan, A.; Jepma, M.; Hu, B.; et al. Pain-Evoked Reorganization in Functional Brain Networks. *Cereb. Cortex* **2020**, *30*, 2804–2822. [CrossRef] [PubMed]
17. Lynch, M. Pain as the fifth vital sign. *J. Intraven. Nurs. Off. Publ. Intraven. Nurses Soc.* **2001**, *24*, 85–94.
18. Tavakolian, M.; Hadid, A. Deep Binary Representation of Facial Expressions: A Novel Framework for Automatic Pain Intensity Recognition. In Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 7–10 October 2018; pp. 1952–1956.
19. Tavakolian, M.; Hadid, A. A Spatiotemporal Convolutional Neural Network for Automatic Pain Intensity Estimation from Facial Dynamics. *Int. J. Comput. Vis.* **2019**, *127*, 1413–1425. [CrossRef]
20. Kunz, M.; Gruber, A.; Lautenbacher, S. Sex differences in facial encoding of pain. *J. Pain* **2006**, *7*, 915–928. [CrossRef]
21. Kunz, M.; Chen, J.-I.; Rainville, P. Keeping an eye on pain expression in primary somatosensory cortex. *NeuroImage* **2020**, *217*, 116885. [CrossRef]
22. Atee, M.; Hoti, K.; Chivers, P.; Hughes, J.D. Faces of pain in dementia: Learnings from a real-world study using a technology-enabled pain assessment tool. *Front. Pain Res.* **2022**, *3*, 827551. [CrossRef]
23. Symons, F.J.; Harper, V.; Shinde, S.K.; Clary, J.; Bodfish, J.W. Evaluating a sham-controlled sensory-testing protocol for nonverbal adults with neurodevelopmental disorders: Self-injury and gender effects. *J. Pain* **2010**, *11*, 773–781. [CrossRef]
24. Stevens, B.J.; Johnston, C.C.; Horton, L. Factors that influence the behavioral pain responses of premature infants. *Pain* **1994**, *59*, 101–109. [CrossRef]
25. Sterling, M.; Jull, G.; Wright, A. The effect of musculoskeletal pain on motor activity and control. *J. Pain Off. J. Am. Pain Soc.* **2001**, *2*, 135–145. [CrossRef]

26. Sanderson, A.; Wang, S.; Elgueta-Cancino, E.; Martinez-Valdes, E.; Sanchis-Sánchez, E.; Liew, B.; Falla, D. The effect of experimental and clinical musculoskeletal pain on spinal and supraspinal projections to motoneurons and motor unit properties in humans: A systematic review. *Eur. J. Pain* **2021**, *25*, 1668–1701. [CrossRef]
27. Gonzalez, R.; Faisal, Z. *Digital Image Processing Second Edition*; Oxford University Press: Oxford, UK, 2019.
28. Viola, P.; Jones, M. Rapid object detection using a boosted cascade of simple features. In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, Kauai, HI, USA, 8–14 December 2001; p. I.
29. Team, OpenCV. OpenCV. 2024. Available online: https://github.com/opencv/opencv/wiki/2024 (accessed on 1 February 2026).
30. NumPy. NumPy Team. 2024. Available online: https://numpy.org/teams/ (accessed on 1 February 2026).
31. Lyons, M. "Excavating AI" Re-Excavated: Debunking a Fallacious Account of the Jaffe Dataset. *arXiv* **2021**, arXiv:2107.13998. [CrossRef]
32. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2012**, *12*, 2825–2830.
33. Kellgren, J.H. Referred Pains from Muscle. *Br. Med. J.* **1938**, *1*, 325–327. [CrossRef] [PubMed]
34. Okkerse, P.; van Amerongen, G.; de Kam, M.L.; Stevens, J.; Butt, R.P.; Gurrell, R.; Dahan, A.; van Gerven, J.M.; Hay, J.L.; Groeneveld, G.J. The use of a battery of pain models to detect analgesic properties of compounds: A two-part four-way crossover study. *Br. J. Clin. Pharmacol.* **2017**, *83*, 976–990. [CrossRef]
35. Canestri, R.; Franco-Alvarenga, P.E.; Brietzke, C.; Vinícius, Í.; Smith, S.A.; Mauger, A.R.; Goethel, M.F.; Pires, F.O. Effects of experimentally induced muscle pain on endurance performance: A proof-of-concept study assessing neurophysiological and perceptual responses. *Psychophysiology* **2021**, *58*, e13810. [CrossRef]
36. Graven-Nielsen, T. Fundamentals of muscle pain, referred pain, and deep tissue hyperalgesia. *Scand. J. Rheumatol.* **2006**, *35*, 1–43. [CrossRef]
37. Valentin, G.; Maribo, T. Hand-held dynamometry fixated with a tripod is reliable for assessment of back extensor strength in women with osteoporosis. *Osteoporos. Int.* **2014**, *25*, 2143–2149. [CrossRef]
38. Arvanitidis, M.; Falla, D.; Sanderson, A.; Martinez-Valdes, E. Does pain influence control of muscle force? A systematic review and meta-analysis. *Eur. J. Pain* **2024**, *29*, e4716. [CrossRef]
39. Parolini, F.; Goethel, M.; Robalino, J.; Becker, K.; Sousa, M.; Pulcineli, B.C.; Ervilha, U.F.; Vilas-Boas, J.P.; Santos, R. Precision and Reliability of a Dynamometer for Trunk Extension Strength and Steadiness Assessment. *Appl. Sci.* **2025**, *15*, 4081. [CrossRef]
40. van der Walt, S.; Schönberger, J.L.; Nunez-Iglesias, J.; Boulogne, F.; Warner, J.D.; Yager, N.; Gouillart, E.; Yu, T. scikit-image: Image processing in Python. *PeerJ* **2014**, *2*, e453. [CrossRef]
41. Gross, J.J.; Levenson, R.W. Emotional suppression: Physiology, self-report, and expressive behavior. *J. Personal. Soc. Psychol.* **1993**, *64*, 970–986. [CrossRef]
42. Karmann, A.J.; Lautenbacher, S.; Bauer, F.; Kunz, M. The influence of communicative relations on facial responses to pain: Does it matter who is watching? *Pain Res. Manag.* **2014**, *19*, 15–22. [CrossRef] [PubMed]
43. Koo, T.K.; Li, M.Y. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* **2016**, *15*, 155–163. [CrossRef] [PubMed]
44. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*; Routledge: New York, NY, USA, 2013.
45. Prkachin, K.M.; Solomon, P.E. The structure, reliability and validity of pain expression: Evidence from patients with shoulder pain. *Pain* **2008**, *139*, 267–274. [CrossRef] [PubMed]
46. Craig, K.D.; Hyde, S.A.; Patrick, C.J. Genuine, suppressed and faked facial behavior during exacerbation of chronic low back pain. *Pain* **1991**, *46*, 161–171. [CrossRef]
47. Kring, A.M.; Gordon, A.H. Sex differences in emotion: Expression, experience, and physiology. *J. Personal. Soc. Psychol.* **1998**, *74*, 686. [CrossRef]
48. Brody, L.R. Gender differences in emotional development: A review of theories and research. *J. Personal.* **1985**, *53*, 102–149. [CrossRef]
49. Merksey, H.; Bogduk, N. Classification of chronic pain. In *Descriptions of Chronic Pain Syndromes and Definitions of Pain Term*; International Association for the Study of Pain: Washington, DC, USA, 1994; pp. 59–71.
50. Karmann, A.J.; Lautenbacher, S.; Kunz, M. The role of inhibitory mechanisms in the regulation of facial expressiveness during pain. *Biol. Psychol.* **2015**, *104*, 82–89. [CrossRef]
51. Keogh, E. Gender differences in the nonverbal communication of pain: A new direction for sex, gender, and pain research? *Pain* **2014**, *155*, 1927–1931. [CrossRef]
52. Coll, M.-P.; Slimani, H.; Woo, C.-W.; Wager, T.D.; Rainville, P.; Vachon-Presseau, É.; Roy, M. The neural signature of the decision value of future pain. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2119931119. [CrossRef]
53. Del Cul, A.; Baillet, S.; Dehaene, S. Brain Dynamics Underlying the Nonlinear Threshold for Access to Consciousness. *PLoS Biol.* **2007**, *5*, e260. [CrossRef] [PubMed]

54. Burt, J.B.; Helmer, M.; Shinn, M.; Anticevic, A.; Murray, J.D. Generative modeling of brain maps with spatial autocorrelation. *NeuroImage* **2020**, *220*, 117038. [CrossRef]
55. Pan, J.; Zhan, L.; Hu, C.; Yang, J.; Wang, C.; Gu, L.; Zhong, S.; Huang, Y.; Wu, Q.; Xie, X.; et al. Emotion Regulation and Complex Brain Networks: Association Between Expressive Suppression and Efficiency in the Fronto-Parietal Network and Default-Mode Network. *Front. Hum. Neurosci.* **2018**, *12*, 70. [CrossRef] [PubMed]
56. Christe, G.; Crombez, G.; Edd, S.; Opsommer, E.; Jolles, B.M.; Favre, J. Relationship between psychological factors and spinal motor behaviour in low back pain: A systematic review and meta-analysis. *Pain* **2021**, *162*, 672–686. [CrossRef] [PubMed]
57. Parolini, F.; Becker, K.; Ervilha, U.F.; Santos, R.; Vilas-Boas, J.P.; Goethel, M.F. Disrupted Corticomuscular Coherence and Force Steadiness During Acute Low Back Pain. *Bioengineering* **2025**, *12*, 1269. [CrossRef]
58. Sullivan, M.J.; Tripp, D.A.; Santor, D. Gender differences in pain and pain behavior: The role of catastrophizing. *Cogn. Ther. Res.* **2000**, *24*, 121–134. [CrossRef]
59. Parolini, F.; Goethel, M.; Becker, K.; Fernandes, C.; Fernandes, R.J.; Ervilha, U.F.; Santos, R.; Vilas-Boas, J.P. Breaking Barriers: Artificial Intelligence Interpreting the Interplay between Mental Illness and Pain as Defined by the International Association for the Study of Pain. *Biomedicines* **2023**, *11*, 2042. [CrossRef]
60. Risberg, G.; Johansson, E.E.; Hamberg, K. A theoretical model for analysing gender bias in medicine. *Int. J. Equity Health* **2009**, *8*, 28. [CrossRef]
61. Mogil, J.S. Qualitative sex differences in pain processing: Emerging evidence of a biased literature. *Nat. Rev. Neurosci.* **2020**, *21*, 353–365. [CrossRef]
62. Samulowitz, A.; Gremyr, I.; Eriksson, E.; Hensing, G. "Brave Men" and "Emotional Women": A Theory-Guided Literature Review on Gender Bias in Health Care and Gendered Norms towards Patients with Chronic Pain. *Pain Res. Manag.* **2018**, *2018*, 6358624. [CrossRef]
63. Zhu, S.; Cai, J.; Xiong, R.; Zheng, L.; Ma, D. Singular Pooling: A Spectral Pooling Paradigm for Second-Trimester Prenatal Level II Ultrasound Standard Fetal Plane Identification. *IEEE Trans. Circuits Syst. Video Technol.* **2025**, *35*, 12508–12523. [CrossRef]