

## Article

# Pain Detection in Biophysiological Signals: Knowledge Transfer from Short-Term to Long-Term Stimuli Based on Distance-Specific Segment Selection

Tobias Benjamin Ricken <sup>1,\*</sup>, Peter Bellmann <sup>1</sup>, Steffen Walter <sup>2</sup> and Friedhelm Schwenker <sup>1</sup>

<sup>1</sup> Institute of Neural Information Processing, Ulm University, James-Franck-Ring, 89081 Ulm, Germany

<sup>2</sup> Medical Psychology Group, University Clinic, 89075 Ulm, Germany

\* Correspondence: tobias-1.ricken@uni-ulm.de

**Abstract:** In this study, we analyze a signal segmentation-specific pain duration transfer task by applying knowledge transfer from short-term (phasic) pain stimuli to long-term (tonic) pain stimuli. To this end, we focus on the physiological signals of the X-ITE Pain Database. We evaluate different distance-based segment selection approaches with the aim of identifying individual segments of the corresponding tonic stimuli that lead to the best classification performance. The phasic domain is used to train the classification model. In the first main step, we compute class-specific prototypes for the phasic domain. In the second main step, we compute the distances between all segments of the tonic samples and each prototype. The segment with the lowest distance to the prototypes is then fed to the classifier. Our analysis includes the evaluation of a variety of distance metrics, namely the Euclidean, Bray–Curtis, Canberra, Chebyshev, City-Block and Wasserstein distances. Our results show that in combination with most of the metrics used, the distance-based selection of one individual segment outperforms the naive approach in which the tonic stimuli are fed to the phasic domain-based classification model without any adaptation. Moreover, most of the evaluated distance-based segment selection approaches lead to outcomes that are close to the classification performance, which is obtained by focusing on the respective best segments. For instance, for the trapezius (TRA) signal, in combination with the electric pain domain, we obtained an averaged accuracy of 68.0%, while the naive approach led to 66.0%. For the thermal pain domain, in combination with the electrodermal activity (EDA) signal, we obtained an averaged accuracy of 59.6%, outperforming the naive approach, which led to 53.2%.



**Citation:** Ricken, T.B.; Bellmann, P.; Walter, S.; Schwenker, F. Pain Detection in Biophysiological Signals: Knowledge Transfer from Short-Term to Long-Term Stimuli Based on Distance-Specific Segment Selection. *Computers* **2023**, *12*, 71. <https://doi.org/10.3390/computers12040071>

Academic Editors: Paolo Bellavista

Received: 15 December 2022

Revised: 21 March 2023

Accepted: 28 March 2023

Published: 31 March 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Physical pain is a sophisticated phenomenon [1]. Each individual learns its meaning by experience [2], mostly early in life. The experience of painful arousal is always subjective [2]. In general, self-reports are preferred for a precise assessment of pain, including its intensity [3]. However, self-reports are not always possible due to several circumstances [4], e.g., for comatose patients [3] and patients suffering from communication disorders or mental disorders such as dementia. Besides self-reports, there exist different techniques that are used by physicians for the assessment of pain [5]. However, the physician's perspective could be biased. More precisely, the human expert might be affected, for instance, by the patient's attractiveness [6]. Moreover, continuous pain monitoring is simply not possible for human experts [1,5], which might lead to erroneous pain treatments causing health risks [1].

In the field of automated pain recognition (APR), the aim is to address the issues discussed above [1,7], by using computer-aided detection systems in combination with machine learning techniques [5]. Note that most of the databases, including induced pain, focus on recordings specific to short-term (phasic) pain stimuli. However, patients are more

likely to experience long-term (tonic) pain episodes [8]. Unfortunately, it is very challenging to collect a sufficient amount of data specific to tonic pain sequences [9]. Therefore, the study objective is to use the knowledge of the phasic domain to detect pain-specific sequences in the tonic domain by partitioning the tonic samples into equally sized segments of the same length as the available phasic domain samples. To this end, we first compute class-specific prototypes for the phasic domain, which are used to measure the distances to the individual segments of the tonic samples. The segments leading to the smallest distances are then used to classify the corresponding tonic samples. Note that this is a follow-up study based on our previous work [10]. Our main contributions of the current work can be summarized as follows:

- We address the pain duration transfer learning task by applying knowledge transfer from phasic to tonic pain stimuli, based on the estimation of individual tonic segments, which constitute the shortest distance to the phasic domain.
- The segment estimation is based on class-specific prototypes computed for the phasic domain.
- We evaluate a set of various distance metrics, namely the Euclidean, Bray–Curtis, Canberra, Chebyshev, City-Block, Wasserstein and Hausdorff distances, whereby the Hausdorff distance uses the whole phasic domain to determine the distance to each of the tonic segments.

The rest of this study is structured as follows. In Section 2, we provide a brief overview of previous studies in the field of APR. The analyzed data set, including the preprocessing and feature extraction steps, is presented in Section 3. In Section 4, we define the evaluated distance metrics. The experimental settings as well as the evaluation approach are provided in Section 5. The obtained results are presented in Section 6 and discussed in detail in Section 7. Finally, this study is closed with a conclusion of our findings in Section 8.

## 2. Related Work

In this section, we summarize some of the previous works from the field of APR, including studies that cover the general context of automated pain recognition as well as the initial attempts at solving the task of signal segmentation-based pain duration transfer.

Over the last years, a lot of research has been performed in the field of APR, e.g., [11–14]. In [15], Thiam et al. conducted a time window analysis for the feature extraction process followed by the evaluation of three different fusion architectures to analyze different pain assessment tasks, including pain detection and pain intensity recognition. This study was evaluated in combination with the SenseEmotion Database [16], which includes recordings of physiological, audio and video data specific to subjects that underwent induced heat pain arousal. The database is partitioned into two parts, data separately collected by stimulating the participant's left forearm and by stimulating the participant's right forearm. A combined grid search over different shift steps of the stimulus starting point and the window length was performed to extract the most informative time window with respect to the pain assessment task. The evaluated window length ranged from 4 s to 6.5 s, whereas the shift of the starting point varied from 0 s (no shift) to 5 s with a step size of 1 s. From each extracted time window, various hand-crafted features were extracted. Each time window parameter combination was then evaluated in a subject-specific 10-fold cross-validation approach in combination with the task of no pain vs. the highest pain intensity level (pain tolerance level). The results show that a temporal shift of 4 s with a time window length of 6.5 s increased the accuracy values in combination with the physiological and audio signals. For the video data, a shift of 2 s in combination with a window length of 6.5 s led to the best results. Based on the optimal time window, from which multiple descriptors were extracted, three different fusion architectures were evaluated. In the first fusion approach, the descriptors specific to each signal were combined into one feature vector, which was then presented to the classification model. The second fusion architecture was a late fusion approach in which one model was designed per data channel (physiological signals, audio and video data). The last fusion approach was a late fusion approach in which one

model was designed for each single modality (audio, head pose, local binary patterns from three orthogonal planes, geometric data describing the subject's face, the electromyogram—measuring the muscle contraction, electrodermal activity—measuring the skin conductance level, electrocardiogram—measuring the heart activity, and respiration data). During the classification phase, in both late fusion approaches, the obtained scores from each model were combined to a final decision. Each fusion approach was additionally evaluated in combination with the leave-one-subject-out cross-validation (LOSO-CV) approach, in which the data of the left out subject are used to test the trained model in each iteration. The best classification performance for the task of no pain vs. the pain tolerance level was obtained in combination with the latter fusion approach (one model per modality), leading to the LOSO-CV mean accuracy values of 81.76% (left forearm) and 83.95% (right forearm).

Furthermore, in combination with the BioVid Heat Pain Database [17], Thiam et al. [18] reported state-of-the-art results obtained by end-to-end deep learning architectures. In the study, uni-modal architectures specific to each single modality from the BioVid Heat Pain Database as well as two fusion architectures were evaluated. Each architecture is based on convolutional neural networks, which are used to learn feature representations of each filtered signal, followed by fully connected layers. No hand-crafted features were used. The evaluated fusion approaches were defined as (a) late fusion of the uni-modality-specific learned feature representations which are concatenated and fed to a fully connected layer to receive a final classification decision; and (b) late fusion of the decisions of each uni-modal architecture, whereby each decision is weighted for the final output. As the evaluation protocol, they applied the LOSO-CV approach. For the binary classification task of no pain vs. the highest pain intensity level, the best performing single modality was the electrodermal activity signal for which they obtained a mean accuracy value of 84.57%. The best performing fusion approach was the late fusion (b) for which they obtained a mean accuracy value of 84.40%.

In one of our previous works [8], we also used the BioVid Heat Pain Database to discuss the possibilities of a segment-specific assessment of pain intensity levels. Note that the BioVid Heat Pain Database consists of biopotentials and video signals recorded in combination with induced short-term heat pain stimuli. We randomly concatenated different samples corresponding to one pain level and one participant, and defined the resulting feature vectors as long-term stimuli. The aim of the study was to analyze the course of the recognition rates with respect to the different lengths of the artificially constructed long-term stimuli. Our conclusions can be summarized as follows. Under the assumption that a person's physiological signals do not change significantly over time during a long-term experience of pain, the classification performance significantly increases with the number of available samples. However, we also discussed that this assumption is less likely to be observed in practice. Therefore, the results reported in [8] can be regarded as upper bound accuracy estimations specific to the corresponding classification scenarios, binary (no pain vs. pain tolerance) and multi-class (including no pain and four pain levels). Furthermore, in [19], we analyzed the complexity of the 5-class pain intensity recognition task constituted by the BioVid Heat Pain Database.

In [20], Badura et al. proposed a system for the rating of patient-specific pain responses during physiotherapy sessions, to prevent the patients from tissue damage, without frequently asking about the currently experienced pain level. For the study, the authors collected various signals during the physiotherapy sessions, whereby the assigned pain level are based on the patients' self-reported intensity levels. The intensity ranges from 0 (no pain) to 10 (maximum level of pain). For the feature extraction process, they applied the wavelet scattering transform. For the classification of the pain-specific sequences, the authors designed support vector machines [21] (SVMs) and decision tree-based adaptive boosting (AdaBoost) ensembles. The pain spectrum was divided into three categories: no pain, moderate pain and severe pain. For the binary classification task of no pain vs. severe pain, in combination with the LOSO-CV, the authors reported mean accuracy values of 0.82 and 0.56 for the AdaBoost and SVM approaches, respectively.

In [22], Othman et al. proposed a system for continuous pain monitoring based on the X-ITE Pain Database [23] in combination with classification- and regression-based random forests [24] and neural networks, analyzing a variety of modality fusions.

In [10], we presented initial knowledge transfer results based on the X-ITE Pain Database in combination with a signal segmentation-specific transfer from short-term to long-term stimuli. More precisely, applying a LOSO-CV evaluation, we analyzed the classification performance of random forests that were trained in combination with short-term stimuli and evaluated the models in the long-term-specific domain. To this end, we divided the long-term stimuli into equally sized segments of the same length as their short-term counterparts and analyzed different evaluation approaches. In the decision average approach, each of the extracted segments was classified separately by the classification model. The final label was then determined by the maximum average over the individual class scores. In the feature average approach, we computed the mean over the feature vectors, which were extracted from the individual segments. The resulting feature average was then fed to the classification model to obtain the final label. In addition, we evaluated the performance based on the individual segments of the long-term stimuli. Our results indicate that focusing on one of the segments outperforms the decision average and feature average approaches. Therefore, in the current study, we want to analyze to which extent it is possible to detect the segments which lead to the best classification performance.

For a larger overview of the field of APR, we kindly refer the reader to the survey articles of Werner et al. [1] and Al-Eidan et al. [25].

### 3. X-ITE Pain Database

In this section, we first briefly describe the X-ITE Pain Database [23], including the calibration and the main pain induction phases. Subsequently, we provide the preprocessing and feature extraction steps.

#### 3.1. Data Set Description

The X-ITE (Experimentally Induced Thermal and Electrical) Pain Database was collected at the Ulm University. A total of 134 healthy subjects, equally distributed among women and men, participated in the pain induction experiments. Pain was induced in form of heat (a thermode attached to one of the participants forearms) and electricity (a disposable Ag/AgCl electrode attached to the index finger and middle finger).

Each of the participants had to undergo an individual calibration phase to determine the pain threshold and the pain tolerance levels. The thermal pain threshold was defined by the temperature value at which the participant felt a change from heat to low pain. The electric pain threshold was defined by the milliamperc value at which the participant felt a change from tingling to low pain. For both pain sources, thermal and electric, the pain tolerance was defined as the point (corresponding value) at which the pain became unbearable. Both pain types were applied in combination with short-term (phasic) and long-term (tonic) stimuli. Therefore, for each participant, the calibration process was conducted four times, each one used to determine the corresponding pain threshold and pain tolerance levels. An intermediate level, defined as the mean between the pain threshold and pain tolerance, was introduced for each of the four combinations which consist of one pain duration type (phasic or tonic) and one pain source (thermal or electrical).

It was not allowed to exceed 50 °C and 49.5 °C for the phasic and tonic heat domains, respectively. For both domains, phasic and tonic electro, it was not allowed to exceed 25 mA to keep the participants constantly safe during the experimental phase (the ethics committee of the Ulm University approved the experiments (372/16)).

During the experimental phase, each of the participants was stimulated 30 times with each of the phasic pain levels, leading to 180 phasic stimuli per participant (90 thermal and 90 electrical). Moreover, each of the participants was stimulated once with each of the tonic pain levels, leading to 6 tonic stimuli per participant (3 thermal and 3 electrical). Each of the phasic thermal stimuli was held for four seconds, and each of the phasic electric

stimuli was held for five seconds. All phasic pain elicitations were always followed by a heat stimulation of 32 °C, with a random duration of eight to twelve seconds. Each of the tonic stimuli, thermal and electric, were held for sixty seconds, always followed by a heat stimulation of sixty seconds with 32 °C. All of the pain stimuli were applied during one session in randomized order.

During the pain elicitation experiments, the experimenters recorded audio signals, facial videos from three different angles, video signals from a body view camera and a thermal camera, as well as five physiological signals. The physiological signals include the electrodermal activity (EDA), which measured the skin conductance level, the electrocardiogram (ECG), which measured the participant's heart activity, as well as three electromyogram (EMG) signals. The EMG signals were measuring the participant's muscle activities of the musculus trapezius (TRA), corrugator supercilii (COR) and zygomaticus major (ZYG).

In this work, we focus on the physiological signals, i.e., ECG, EDA and the three EMG signals, COR, TRA, and ZYG. For additional data set information, we kindly refer the reader to [23].

### 3.2. Preprocessing and Feature Extraction

For the preprocessing and feature extraction process, we followed one of our previous works [26]. We were able to use the data of 125 out of the 134 available subjects. The data specific to the nine remaining subjects contained errors. Based on our time window analysis presented in [26], we applied the setup, which led to the highest mean accuracy values. More precisely, we shift each time window of a phasic thermal stimulus by 3 s. The time window of a phasic electric stimulus is shifted by 1 s. The length of both stimuli types is set to 4 s. For the tonic samples, we apply the same shifts. Each tonic sequence is additionally split into 14 equally sized segments with a length of 4 s. The last segment is ignored due to its length, which is less than 4 s. For each extracted time window, we applied a 3rd-order Butterworth bandpass filter for all three EMG channels and the ECG channel. For the non-periodic EDA signal, we did not conduct any filtering. For each EMG signal, we filtered out the frequencies below 20 Hz and above 250 Hz. The frequencies beneath 0.1 Hz and above 25 Hz were filtered out for the ECG signal. For the extraction of features, we followed our previous work [10]. In summary, each stimulus is represented by a vector of 412 features. For each pain domain, we extracted individual no-pain sequences (baselines). Each baseline followed a stimulus of the corresponding lowest pain intensity level. Note that in the pain elicitation experiments, the baseline stimuli were defined as thermal stimuli with the temperature set to 32 °C. We applied the same shifts to each extracted baseline stimulus, with respect to the electric and thermal pain domains.

Note that in this study, we focus on the data specific to the baseline level and the pain tolerance level. The number of features per single- and multi-modality is presented in Table 1. The number of available samples for each stimulus type is presented in Table 2.

**Table 1.** Number of features per single- and multi-modality. Corrugator (COR), Trapezius (TRA) and Zygomaticus (ZYG) are electromyogram (EMG) signals. ECG denotes the electrocardiogram, EDA denotes the electrodermal activity and EMG denotes the early fusion of COR, TRA, and ZYG. EFU denotes the early fusion over all modalities.

ECG	EDA	COR	TRA	ZYG	EMG	EFU
87	79	82	82	82	246	412

**Table 2.** Label distribution per pain domain. PB: phasic baseline stimuli (no pain), PP: phasic pain stimuli, TB: tonic baseline stimuli (no pain), TP: tonic pain stimuli.

Thermal Domain				Electric Domain			
PB	PP	TB	TP	PB	PP	TB	TP
3727	3716	121	124	3720	3719	123	123

#### 4. Methods

In this section, we describe the applied distance metrics that are used to find the segments with the smallest distance to the phasic pain domain and which will be used to classify the complete tonic samples. The distance between two vectors can be measured in multiple ways. At first, let  $x \in \mathbb{R}^n$  and  $y \in \mathbb{R}^n$  be  $n$ -dimensional feature vectors.

##### 4.1. Euclidean Distance

The Euclidean distance [27] is defined as the square root of the sum of squared differences between the elements of both vectors, i.e.,

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}. \quad (1)$$

##### 4.2. Manhattan (City-Block) Distance

The Manhattan or City-Block distance [27] is defined as the sum of absolute differences between the elements of the two vectors  $x, y$ , i.e.,

$$d(x, y) = \sum_{i=1}^n |x_i - y_i|. \quad (2)$$

##### 4.3. Chebyshev Distance

The Chebyshev distance [27] is defined as the maximum of the element-wise differences between the two vectors  $x, y$ , i.e.,

$$d(x, y) = \max_i |x_i - y_i|. \quad (3)$$

##### 4.4. Canberra Distance

Another distance measure, which can be seen as a weighted Manhattan distance, is the Canberra distance [27] defined as

$$d(x, y) = \sum_{i=1}^n \frac{|x_i - y_i|}{|x_i| + |y_i|}. \quad (4)$$

In the case of  $x_i = y_i = 0$ , the addend  $i$  is ignored in the summation.

##### 4.5. Bray–Curtis Dissimilarity

The Bray–Curtis dissimilarity [27] is defined as the sum over the absolute differences between the  $i$ -th elements of the two vectors  $x, y$  divided by the sum of the  $i$ -th elements of  $x, y$ , i.e.,

$$d(x, y) = \frac{\sum_{i=1}^n |x_i - y_i|}{\sum_{i=1}^n (x_i + y_i)}. \quad (5)$$

The denominator can be equal to zero in general settings, meaning that the corresponding Bray–Curtis distance is undefined. In special cases, e.g., for probability vectors, the Bray–Curtis is always well defined.

#### 4.6. Jensen–Shannon Divergence

For probability vectors  $x, y$ , the Jensen–Shannon divergence [27] is defined as

$$JSD(x, y) = \beta KL(x, z) + (1 - \beta)KL(y, z), \quad (6)$$

with  $\beta \in [0, 1]$  and  $z = \beta(x + y)$ . Here,  $KL$  denotes the Kullback–Leibler distance:

$$KL(x||y) = \sum_{i=1}^n x_i \ln \frac{x_i}{y_i}. \quad (7)$$

Setting  $\beta = 1/2$  and taking the square root of the  $JSD$  leads to a distance measure [27] defined as

$$d(x, y) = \sqrt{\frac{KL(x||z) + KL(y||z)}{2}}. \quad (8)$$

#### 4.7. Wasserstein Distance

The  $L_1$ -Wasserstein distance [27] is defined as

$$d(x, y) = \left( \int_0^1 |F_x^{-1}(u) - F_y^{-1}(u)| du \right), \quad (9)$$

whereby  $F_x, F_y$  denote the inverse cumulative distribution functions over the distributions  $x, y$ , respectively.

Bellemare et al. [28] showed that an alternative to the  $L_p$ -Wasserstein distance is the  $L_p$ -Cramér distance [29], for all  $p \geq 1$ , whereby the properties of the Wasserstein hold also for the Cramér distance. The Cramér distance is defined as

$$d_p(x, y) = \left( \int_{-\infty}^{\infty} |F_x(u) - F_y(u)|^p du \right)^{1/p}, \quad (10)$$

whereby  $F_x, F_y$  denote the cumulative distribution functions over the vectors  $x$  and  $y$ , respectively. As Bellemare et al. pointed out, the  $L_1$ -Wasserstein and  $L_1$ -Cramér measures lead to identical distances. For empirical distributions, the Cramér distance with  $p = 1$  is given by

$$d(x, y) = \sum_{i=1}^n |F_x(x_i) - F_y(y_i)|. \quad (11)$$

Therefore, the Cramér distance (the implementation of the Wasserstein distance in [30]) is used for the computation of the Wasserstein distance.

#### 4.8. Hausdorff Distance

The Hausdorff distance [27] measures the distance between two sets of samples  $X, Y$ . More precisely, the Hausdorff distance is defined in Equation (12), whereby the distance between the samples  $x, y$  is computed by the Euclidean norm as defined in Equation (13), i.e.,

$$d(X, Y) = \max(d_m(X, Y), d_m(Y, X)), \quad (12)$$

$$d_m(X, Y) = \max_{x \in X} \min_{y \in Y} \|x - y\|. \quad (13)$$

Note that in Equation (13), in addition to the Euclidean norm, one can also use other measures, such as the Chebyshev or Manhattan distances.

#### 4.9. Evaluation Protocol

As the evaluation protocol, we apply the leave-one-subject-out cross-validation (LOSO-CV). In each iteration of the LOSO-CV, the data of the left out subject are used to evaluate the classification model, which is trained in combination with the data specific to the remaining subjects.

### 5. Experimental Setup

In this section, we provide our experimental setup, including the parameter settings used to find and evaluate the tonic segments with the smallest distance to the phasic pain domain.

In this study, we focus on the biophysiological signals in combination with the binary classification task of no pain vs. the pain tolerance level, which we simply refer to as *no pain vs. pain*.

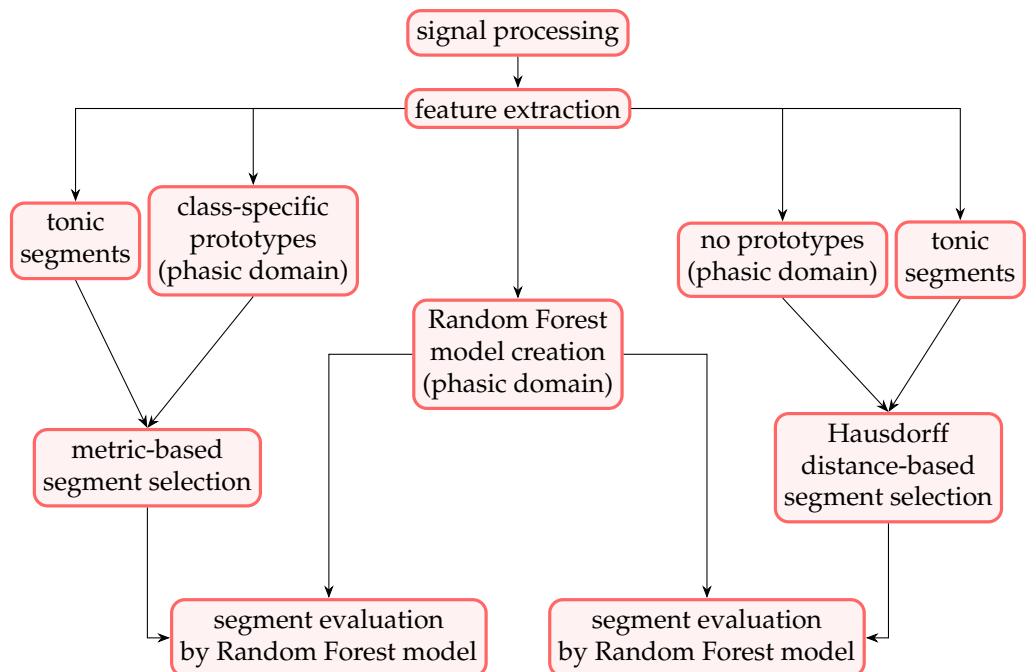
In the experiments, we evaluate the task of knowledge transfer from phasic to tonic pain stimuli in combination with the signal segmentation approach. We apply the random forest algorithm [24] with 100 decision trees [31]. Each tree is limited to ten nodes. The impurity is measured by the Gini-Index.

We evaluate each tonic segment individually in combination with the single- and multi-modal signals for both pain domains (electric and thermal), respectively. More precisely, for the evaluation of the distance metrics, one random forest model is trained in combination with the phasic pain stimuli and evaluated on one individual segment of the tonic samples, in the electric and thermal domain, respectively.

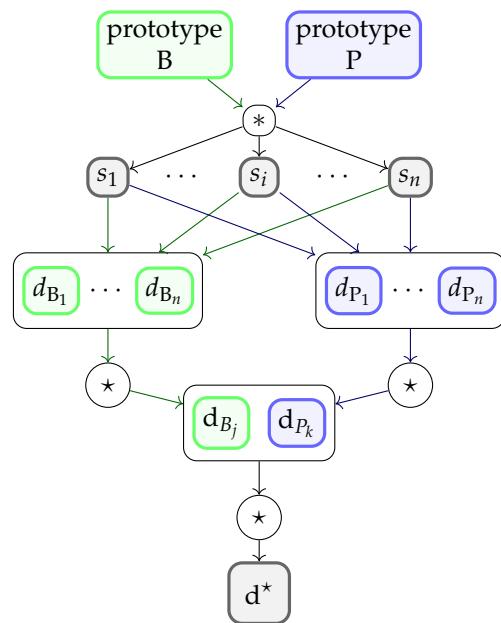
First, we compute two class-specific (no pain and pain) feature-average samples (prototypes). The distance between each segment of a tonic sample and each feature-average sample is computed by one of the distance metrics defined in Section 4. Based on the Jensen–Shannon distance, we first estimate the probability distribution over the given feature vectors by applying the kernel density estimation approach in combination with the radial basis function (RBF) kernel. The optimal bandwidth is estimated by performing a grid search over  $\gamma = 10^\delta, \delta \in \{-4, -3, -2, -1, 0, 1, 2\}$ . For each test sample, we obtain 28 distance values, one value for each segment (14 segments in total, 2 class-specific prototypes). Based on these distance values, the segment which has the shortest distance to the phasic prototypes is selected. In combination with the Hausdorff distance (implementation from <https://pypi.org/project/hausdorff/> (last accessed on 16 March 2023) which implements the algorithm presented in [32]), the complete phasic training set is used to determine the closest segment, instead of focusing on the prototype-specific representations.

The selected segment is then evaluated by the model and the predicted label is taken as the final output for the corresponding tonic sample. The complete evaluation process is depicted in Figure 1. Additional information for the prototype-based segment selection is depicted in Figure 2.

We conducted all experiments with the Python (<https://www.python.org/> (last accessed on 16 March 2023)) programming language and the corresponding packages of the data stack (NumPy [33], Scikit-learn [34], pandas [35], SciPy [30], Matplotlib [36] and seaborn [37]).



**Figure 1.** Complete process of the evaluated experiments. The random forest model is trained solely on the phasic domain samples. The evaluation follows one of the two available branches. The left branch represents the prototype-based evaluation, whereas the right branch represents the Hausdorff distance-based evaluation. Note that the Hausdorff distance is the only evaluated metric in which no prototypes are computed. The knowledge transfer from the phasic domain to the tonic domain is performed by the evaluation of the individual segments.



**Figure 2.** Segment selection process. The  $*$ -operator denotes the metric-based distance computation between the prototypes and each segment  $s_1, \dots, s_n$ . Parameters  $d_{B_1}, \dots, d_{B_n}$  and  $d_{P_1}, \dots, d_{P_n}$  denote the resulting distance values for the baseline (B) and pain (P) prototypes, respectively. The  $\star$ -operator denotes the selection operator (here: minimum) performed on the vectors of the obtained values.  $d_{B_j}$  and  $d_{P_k}$  denote the lowest distances to the prototypes. The overall lowest distance is obtained when the  $\star$ -operator is applied to  $d_{B_j}$  and  $d_{P_k}$ . The segment, to which the value  $d^*$  belongs to, is selected and evaluated by the model. (Note that this approach is applied in combination with all evaluated metrics defined above, except for the Hausdorff distance).

## 6. Results

In this section, we provide the results for the pain duration knowledge transfer task based on multiple distance measures, which are used to determine the individual segments of the tonic stimuli that are the closest to the phasic domain. The evaluated metrics are presented in Section 4.

First, for each signal, we provide the results, for which we always use the same segment for the classification of the tonic samples (evaluation of individual segments). In addition, for each signal, we provide the results, for which we evaluate the models in combination with the complete tonic samples (no segmentation). We will denote the evaluation without the segmentation as the naive approach. Moreover, as references for each signal, we additionally present the mean accuracy values, which are obtained by the models that are trained and evaluated in combination with the tonic stimuli without segmentation. The thermal and electric domains are always analyzed separately.

### 6.1. Electric Domain

In Table 3, we present the obtained mean accuracy values for the evaluation of the individual segments. The highest mean accuracy value of 74.0% is obtained with the ECG signal in combination with the first segment. Except for the EMG and TRA signals, the first segment leads to the highest mean accuracy values. For the EMG and TRA signals, the highest mean accuracy values are obtained with the 6th segment (68.0% and 68.4%, respectively).

**Table 3.** Electric domain: Evaluation of the individual segments of the tonic samples based on uni- and multi-modal signals. Parameter  $k$ -S specifies the evaluated segment. NA: Naive approach, in which the model is trained on phasic stimuli and evaluated on the unmodified tonic samples (no segmentation). Ref.: Reference values obtained in the tonic domain. The LOSO-CV accuracies are given in %. The segment leading to the best outcome is depicted in bold, respectively.

$k$ -S	1	2	3	4	5	6	7	8	9	10	11	12	13	14	NA	Ref.
ECG	<b>74.0</b>	64.8	59.2	54.0	55.2	56.0	49.6	48.4	40.8	41.6	42.4	46.4	45.6	40.0	50.0	83.2
EDA	<b>66.0</b>	60.8	54.0	54.4	48.4	44.4	50.4	50.4	45.2	50.0	49.2	50.4	51.2	47.2	55.6	84.8
EMG	53.6	46.0	50.4	55.6	56.0	<b>68.0</b>	60.8	61.2	61.6	60.4	55.6	58.8	52.0	50.4	66.0	90.8
COR	<b>61.2</b>	47.6	44.4	53.6	48.8	51.6	52.4	52.8	50.4	52.0	48.8	53.6	45.6	48.4	55.2	72.8
TRA	51.6	46.4	53.6	56.4	60.4	<b>68.4</b>	68.0	62.0	63.2	60.0	53.2	55.6	51.2	48.8	66.4	87.2
ZYG	<b>72.0</b>	66.8	52.4	55.6	54.0	56.4	53.2	50.0	47.2	46.8	43.6	44.0	42.4	42.0	43.2	74.8
EFU	<b>68.8</b>	50.8	46.8	54.0	55.2	53.6	56.8	53.2	52.8	51.2	49.2	56.4	48.8	48.0	56.8	93.6

In Table 4, we present the results for the distance metrics Euclidean, City-Block, Chebyshev, Canberra, Bray–Curtis, Wasserstein and Hausdorff. We evaluated the Hausdorff distance in combination with the Euclidean, Chebyshev and City-Block distances. Note that initial results indicated that on average, the Hausdorff metric performs best in combination with the Chebyshev distance. Therefore, for the Hausdorff distance, we focus on the outcomes which were obtained in combination with the Chebyshev metric. Across the distance-based approaches evaluated in the electric pain domain, the Chebyshev and the Canberra metrics led to the maximum of 68.0% mean accuracy in combination with the TRA signal. A slightly decreased mean accuracy value of 67.2% was obtained with the EMG signal in combination with the Wasserstein distance. The lowest mean accuracy value of 46.8%, which is below the chance level, was obtained with the COR signal in combination with the Hausdorff distance.

**Table 4.** Electro: Transfer from phasic to tonic stimuli. Evaluation of various distance metrics used to determine the closest segment. *k*-S: The highest accuracy value obtained in combination with the evaluation of the individual segments. NA: Naive approach in which the model is trained on phasic stimuli and evaluated on the unmodified tonic samples (no segmentation). Ref.: Reference values obtained in the tonic domain. The LOSO-CV accuracies are given in %. The metric leading to the best outcome is depicted in bold, respectively.

Metric	Euclidean	City-Block	Chebyshev	Canberra	Bray-Curtis	Wasserstein	Hausdorff	<i>k</i> -S	NA	Ref.
ECG	54.4	52.4	<b>56.8</b>	48.0	55.6	47.2	50.0	74.0	50.0	83.2
EDA	49.6	50.8	50.8	48.0	52.8	47.2	<b>54.0</b>	66.0	55.6	84.8
EMG	65.6	64.4	63.2	65.2	52.0	<b>67.2</b>	61.6	68.0	66.0	90.8
COR	<b>54.0</b>	50.4	52.0	53.2	50.8	52.0	46.8	61.2	55.2	72.8
TRA	65.6	66.4	<b>68.0</b>	<b>68.0</b>	67.2	66.4	57.2	68.4	66.4	87.2
ZYG	<b>54.8</b>	48.8	53.6	52.4	52.0	52.0	50.8	72.0	43.2	74.8
EFU	57.6	58.4	56.4	59.6	56.4	60.0	<b>60.8</b>	68.8	56.8	93.6

For the Jensen–Shannon distance, the obtained results are presented in Table 5. The maximum of 67.2% mean accuracy is obtained with  $\gamma \in \{10^1, 10^2\}$  in combination with the TRA signal. A slightly decreased mean accuracy value of 66.8% is obtained in combination with the TRA signal and  $\gamma = 10^0$ . The lowest mean accuracy value (42.4%) is obtained with  $\gamma = 10^{-1}$ , also in combination with the TRA signal.

**Table 5.** Electro: Transfer from phasic to tonic stimuli—Jensen–Shannon. Grid search over the bandwidth values for the kernel density estimation. *k*-S: The highest accuracy value obtained in combination with the evaluation of the individual segments. NA: Naive approach in which the model is trained on phasic stimuli and evaluated on the unmodified tonic samples (no segmentation). Ref.: Reference values obtained in the tonic domain. The LOSO-CV accuracies are given in %. The  $\gamma$ -value leading to the best outcome is depicted in bold, respectively.

$\gamma$	$10^{-4}$	$10^{-3}$	$10^{-2}$	$10^{-1}$	$10^0$	$10^1$	$10^2$	<i>k</i> -S	NA	Ref.
ECG	47.6	53.6	42.8	<b>54.4</b>	54.0	52.0	52.0	74.0	50.0	83.2
EDA	52.8	<b>51.2</b>	50.0	50.0	48.0	<b>53.6</b>	<b>53.6</b>	66.0	55.6	84.8
EMG	46.0	44.4	46.4	43.2	61.2	<b>65.2</b>	64.8	68.0	66.0	90.8
COR	47.6	<b>52.8</b>	51.2	52.0	52.4	52.4	52.4	61.2	55.2	72.8
TRA	60.0	47.6	44.4	42.4	66.8	<b>67.2</b>	<b>67.2</b>	68.4	66.4	87.2
ZYG	46.4	53.6	52.0	50.0	<b>54.4</b>	<b>54.4</b>	<b>54.4</b>	72.0	43.2	74.8
EFU	54.0	54.4	47.6	55.2	<b>58.8</b>	<b>58.8</b>	<b>58.8</b>	68.8	56.8	93.6

## 6.2. Thermal Domain

In Table 6, we present the obtained mean accuracy values for the evaluation of the individual segments with respect to the thermal domain. The highest mean accuracy value of 69.2% is obtained with the EDA signal in combination with the third segment of the tonic samples.

In Table 7, we present the results for the distance metrics Euclidean, City-Block, Chebyshev, Canberra, Bray–Curtis, Wasserstein and Hausdorff. For the heat pain domain, the highest mean accuracy value of 59.6% is obtained with the EDA signal in combination with the Wasserstein distance. The lowest mean accuracy value of 46.8%, which is below the chance level, is obtained with the ECG signal in combination with the Hausdorff distance.

**Table 6.** Thermal domain: Evaluation of the individual segments of the tonic samples in combination with the uni- and multi-modal signals. Parameter *k-S* specifies the evaluated segment. NA: Naive approach in which the model is trained on phasic stimuli and evaluated on the unmodified tonic samples (no segmentation). Ref.: Reference values obtained in the tonic domain. The LOSO-CV accuracies are given in %. The segment leading to the best outcome is depicted in bold, respectively.

<i>k-S</i>	1	2	3	4	5	6	7	8	9	10	11	12	13	14	NA	Ref.
ECG	42.4	50.0	54.8	53.2	53.6	48.4	50.4	48.0	48.8	<b>56.0</b>	50.8	48.4	50.8	50.8	48.8	81.6
EDA	34.8	58.8	<b>69.2</b>	57.6	52.0	50.0	46.0	49.6	50.0	55.6	55.6	52.0	54.8	50.4	53.2	84.4
EMG	42.8	48.4	51.2	51.2	<b>46.4</b>	49.6	53.2	54.8	52.8	57.2	56.8	<b>60.8</b>	56.8	54.8	62.8	70.8
COR	43.6	47.2	50.4	50.8	47.6	48.4	52.8	52.8	50.0	55.6	56.4	<b>61.6</b>	53.6	55.6	60.8	63.6
TRA	47.6	52.0	45.2	42.4	50.0	46.8	56.8	50.8	55.2	56.8	56.0	52.4	<b>60.8</b>	53.6	62.4	64.8
ZYG	<b>57.2</b>	<b>57.2</b>	54.0	50.0	45.2	46.4	51.6	54.0	53.6	52.8	54.0	55.6	52.8	53.6	57.6	58.4
EFU	33.6	51.6	<b>65.2</b>	56.0	48.8	43.6	51.2	52.8	51.6	58.4	56.4	58.0	53.6	52.0	59.6	86.0

**Table 7.** Thermal domain: Transfer from phasic to tonic stimuli. Evaluation of various distance metrics used to determine the closest segment. *k-S*: The highest accuracy value obtained in combination with the evaluation of the individual segments. NA: Naive approach in which the model is trained on phasic stimuli and evaluated on the unmodified tonic samples (no segmentation). Ref.: Reference values obtained in the tonic domain. The LOSO-CV accuracies are given in %. The metric leading to the best outcome is depicted in bold, respectively.

Metric	Euclidean	City-Block	Chebyshev	Canberra	Bray-Curtis	Wasserstein	Hausdorff	<i>k-S</i>	NA	Ref.
ECG	49.6	50.0	53.2	<b>55.2</b>	50.4	51.2	46.8	56.0	48.8	81.6
EDA	56.4	52.8	56.8	55.6	47.2	<b>59.6</b>	49.2	69.2	53.2	84.4
EMG	54.8	<b>57.6</b>	52.8	55.2	50.0	56.4	54.0	60.8	62.8	70.8
COR	50.8	50.0	52.4	53.6	51.6	54.0	<b>54.4</b>	61.6	60.8	63.6
TRA	55.2	<b>57.2</b>	51.6	54.4	54.4	54.0	56.0	60.8	62.4	64.8
ZYG	54.4	56.8	56.4	<b>58.8</b>	46.8	58.4	53.6	57.2	57.6	58.4
EFU	50.4	52.4	55.6	<b>57.6</b>	54.8	<b>57.6</b>	56.8	65.2	59.6	86.0

The results of the performed grid search in combination with the Jensen–Shannon distance are presented in Table 8. The maximum of 60.0% is obtained for the TRA signal in combination with  $\gamma = 10^{-4}$ . For  $\gamma = 10^0$ , the maximum of 57.6% mean accuracy is obtained in combination with the EDA signal. A slightly decreased value of 57.2% is obtained for the ZYG signal in combination with  $\gamma \in \{10^0, 10^1, 10^2\}$ . The lowest mean accuracy value of 45.2% (below chance level) is obtained when the COR signal is used in combination with  $\gamma$  set to  $10^{-3}$ .

**Table 8.** Thermal domain: Transfer from phasic to tonic stimuli—Jensen–Shannon. Grid search over the bandwidth values for the kernel density estimation. *k-S*: The highest obtained accuracy value in combination with the evaluation of individual segments. NA: Naive approach in which the model is trained on phasic stimuli and evaluated on the unmodified (no segmentation) tonic samples, Ref.: Reference values obtained in the tonic domain. The LOSO-CV results are given in %. The  $\gamma$ -value leading to the best outcome is depicted in bold, respectively.

$\gamma$	$10^{-4}$	$10^{-3}$	$10^{-2}$	$10^{-1}$	$10^0$	$10^1$	$10^2$	<i>k-S</i>	NA	Ref.
ECG	<b>53.2</b>	47.2	51.2	50.8	47.2	47.2	47.2	56.0	48.8	81.6
EDA	54.0	<b>48.4</b>	55.6	54.0	<b>57.6</b>	55.2	55.2	69.2	53.2	84.4
EMG	<b>56.8</b>	49.6	50.0	50.0	52.0	52.0	52.4	60.8	62.8	70.8
COR	<b>55.2</b>	45.2	52.0	51.6	51.6	51.6	51.6	61.6	60.8	63.6
TRA	<b>60.0</b>	50.4	52.0	55.6	54.0	54.0	54.0	60.8	62.4	64.8
ZYG	50.4	47.2	53.2	46.8	<b>57.2</b>	<b>57.2</b>	<b>57.2</b>	57.2	57.6	58.4
EFU	<b>57.2</b>	53.2	50.8	50.4	52.8	52.8	52.8	65.2	59.6	86.0

## 7. Discussion

In this study, we tried to determine the individual segments of the tonic samples which are expected to be the closest to the phasic domain. To this end, we evaluated different metrics used to compute the distances between all segments of the tonic samples and the class-specific prototypes of the phasic domain (except for the Hausdorff distance). Our findings indicate that using the knowledge obtained from the phasic domain to classify the metric-selected individual segments of the tonic domain samples performs well in the electric domain. However, in the thermal domain, the same task is more challenging. As also discussed in some previous studies [7,26,38], in combination with the electric stimuli, the classification tasks within the same pain duration domains, phasic and tonic, lead to higher accuracy values in comparison to the thermal domains, respectively. Moreover, for the electric domain, in combination with various signals, we are able to select the segments which are almost as good as the optimal segment that was found in the evaluation of all individual segments.

### 7.1. Electric Domain

In the individual segment evaluation, in combination with the electric domain, we were able to come close to or even to outperform the naive approach with various signals. The highest mean accuracy value of 68.0% is obtained with the 6th segment in combination with the EMG signal, which is 2.0% above the naive approach obtained in combination with the EMG signal.

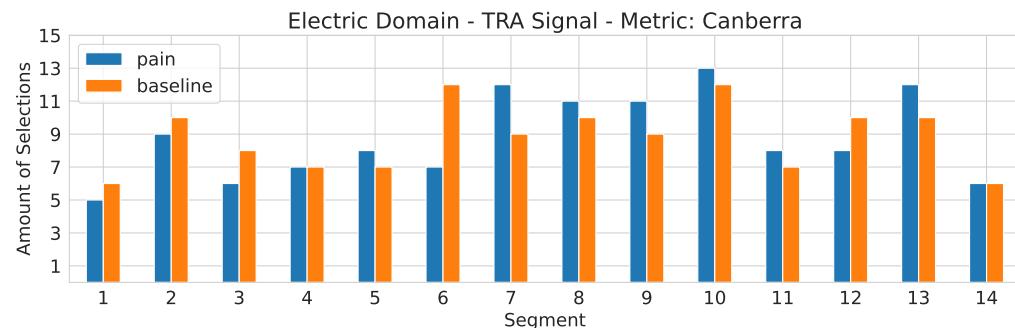
With the TRA signal, in combination with the 6th segment, we obtained a maximum of 68.4%, which is above the naive approach obtained in combination with the TRA signal (66.4%). The segment selection based on the Canberra and Chebychev metrics led to an almost equal result of 68.0% mean accuracy. Since the EMG modality is a composition of three signals (COR, TRA, and ZYG), the huge predictive power might be based in the TRA signal. The TRA signal led to a mean accuracy value of 68.4%, whereas the EMG signal is only 0.4% below the result obtained for the TRA signal, both in combination with the 6th segment. For the COR and ZYG signals, in combination with the 6th segment, a mean accuracy slightly above the chance level is obtained (51.6% and 56.4%, respectively).

In Figure 3, the selection frequency of each segment is depicted for the TRA signal, based on the Canberra metric. As we can observe, with simple metrics, it is not possible to determine the best-performing individual segment for each tonic sample. Since distance values are simple scalars which express a (linear) relationship between two vectors, the non-linear information which might be learned by the RF algorithm cannot be represented by these values. In general, the RF algorithm is able to use more information from the samples. Nevertheless, the segments selected based on the applied metrics include enough useful information, which leads to almost similar results.

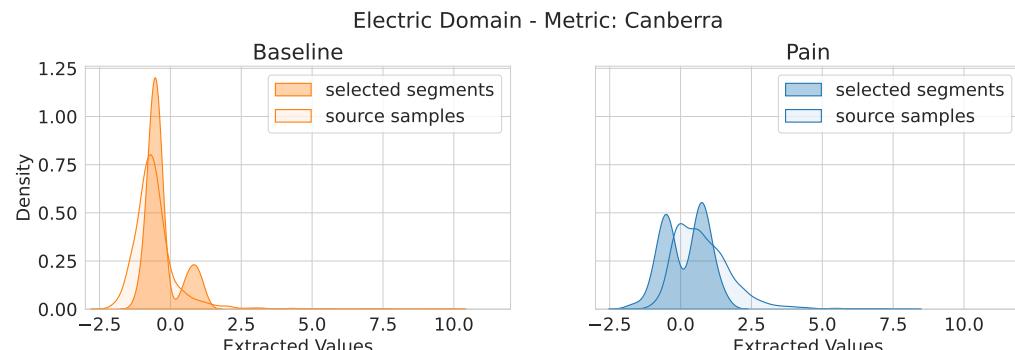
To further investigate the information contained in the selected segments, we collected the most important feature from the phasic electric domain by applying a 3-fold cross-validation in which each left-out fold contained the data of 1/3 of the available subjects. We analyzed the obtained RF model of each fold and gathered the most important feature from the three evaluations. For the obtained feature, we analyzed the extracted feature values, specific to each class for the phasic domain samples and the selected segments. The distributions of the extracted feature values are depicted in Figure 4. As it can be seen, for both baseline and pain, the distributions of the extracted feature values show similarities between the phasic samples and the selected segments. In Figure 5, we provide a similar observation for the TRA signal in combination with the Jensen–Shannon metric with the bandwidth set to ten.

The selection frequency of each segment is depicted in Figure 3, in combination with the TRA signal and the Canberra metric. As presented in Table 3, the 6th segment is the best-performing segment of the tonic samples. However, as it can be seen, not only the 6th segment is selected when the metric is applied. Therefore, under the assumption that the best results of the individual segment evaluation provide upper bounds for the

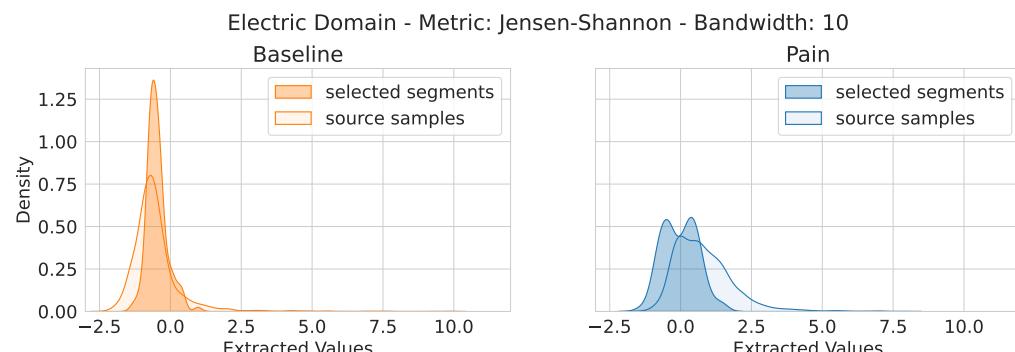
segment selection-based approaches, it is not possible to determine the optimal segments by computing the distances between the phasic electric prototypes and the corresponding segments. However, it seems that for the TRA signal, the segments which are determined in combination with the Chebychev and Canberra metrics, contain enough useful information for the analyzed classification task.



**Figure 3.** Electric domain. The segment selection frequency for the TRA signal in combination with the Canberra metric. On the *x*-axis, the segment number is depicted. The *y*-axis represents the segment counter.



**Figure 4.** Electric domain: TRA signal—Canberra metric. Distribution of extracted feature values for the feature trapezius\_second\_derivative\_features\_signal\_std of the TRA signal in combination with the selected segments with the Canberra metric and the distribution of the source domain samples (phasic domain). (**Left**) Value distributions for the baseline stimuli. (**Right**) Value distributions of the pain stimuli.



**Figure 5.** Electric domain: TRA signal—Jensen–Shannon metric with bandwidth set to ten. Distribution of extracted feature values for the feature trapezius\_second\_derivative\_features\_signal\_std of the TRA signal in combination with the selected segments with the Canberra metric and the distribution of the source domain samples (phasic domain). (**Left**) Value distributions for the baseline stimuli. (**Right**) Value distributions of the pain stimuli.

Note that for the electric pain domain, in combination with the signals ECG, EDA, COR, ZYG, and EFU, we observed a decline in the obtained accuracy values over time. As discussed in [8], the human body might adapt to the experienced pain over time, leading to a significant change in the physiological signals. Furthermore, as reported in [7,26] for the electric pain domain, the most informative part is found in the beginning of a pain-specific sequence due to the nature of electricity because electrically stimulated pain is felt instantly. This was also observed by the obtained results for the EFU signal in [10]. For the signals EMG and TRA, a similar behavior is observed but with a delay in time. Moreover, based on the results presented in Table 4, the TRA and EMG signals are the best-performing modalities with respect to the pain duration transfer task in the electric pain domain. These modalities also show good performance over time, as can be observed by the results presented in Table 3.

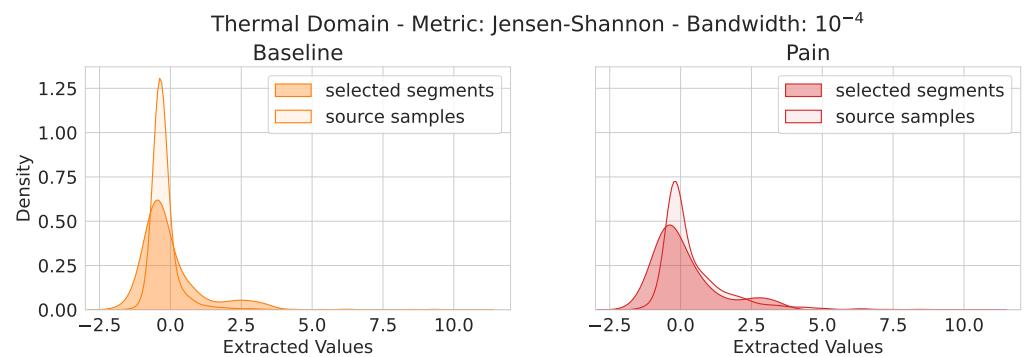
## 7.2. Thermal Domain

With the approach, in which each segment is evaluated individually, we were able to outperform the naive approach with the EDA and EFU signals. For the EDA signal, the highest mean accuracy value of 69.2% is obtained with the 3rd segment, whereas with the naive approach, the mean accuracy value is slightly above the chance level (53.2%). For the EFU signal, the maximum of 65.2% is also obtained with the 3rd segment (naive approach: 59.6%). As discussed in previous studies ([7,15,26]), the EDA signal is a good-performing signal for the thermal pain domain, in general. As already reported in [10,15,26], the human body shows a delayed reaction to the pain induction by thermal stimulation in the physiological signals. Based on the results presented in Table 6, the good performance of the EFU signal might originate from the effectiveness of the EDA signal since for each remaining signal (ECG, COR, TRA, and ZYG) the mean accuracy values are close to or even below the chance level.

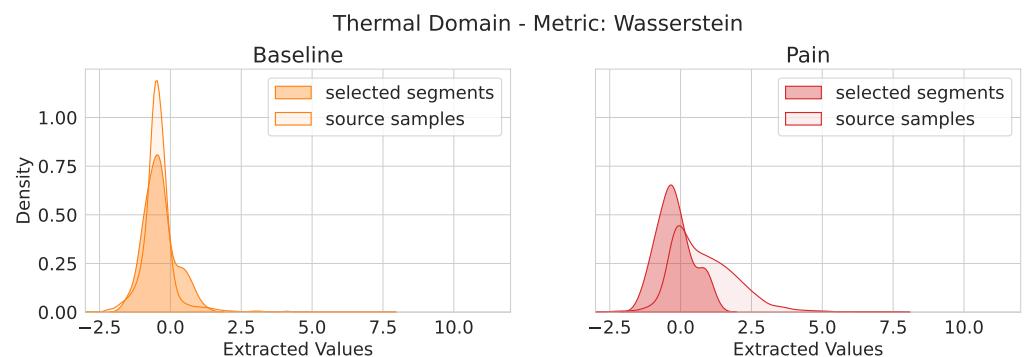
For the COR signal, in combination with the 12th segment, a mean accuracy value of 61.6% was obtained, which is 0.8% above the outcome obtained with the naive approach. For the TRA signal, in combination with the 13th segment, a maximum of 60.8% was obtained. With the TRA signal, in combination with the Jensen–Shannon metric and the bandwidth set to  $10^{-4}$ , we obtained a mean accuracy value of 60.0%, which is slightly below the result obtained for the 13th segment.

For the ZYG signal, we were able to reach the result which we obtained for the optimal segments (1st and 2nd: 57.2%). In combination with the Jensen–Shannon metric and the bandwidth set to 1, the ZYG signal led to a mean accuracy value of 57.2%. Moreover, with the obtained results for the optimal segments in combination with the Jensen–Shannon metric, we are only 1.2% below the reference value (58.4%).

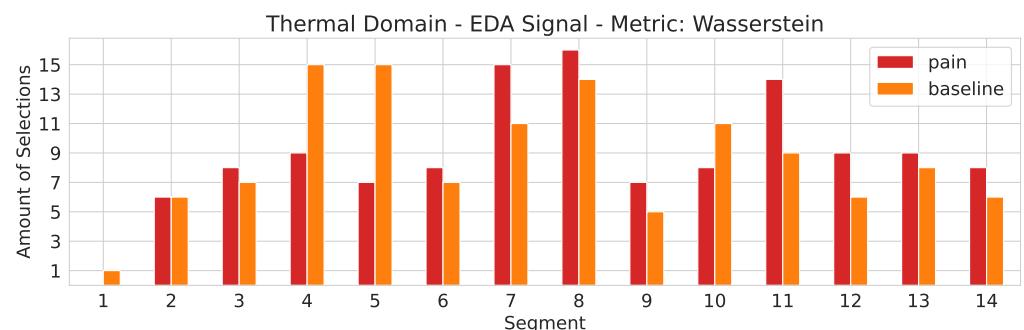
For the thermal pain domain, we observed a delayed body reaction to the induced pain stimuli. Similar to the phasic electric domain, we collected the most important feature, specific to the thermal pain domain, with the approach described in Section 7.1. The distributions of the extracted feature values for the most important feature in the thermal domain for the TRA signal in combination with the Jensen–Shannon metric and the bandwidth set to  $10^{-4}$  are depicted in Figure 6. As it can be seen, the distributions for the extracted feature values for both, baseline and pain, in combination with the phasic samples and the selected segments show similarities. In Figure 7, we provide a similar observation for EDA signal in combination with the Wasserstein metric. The selection frequency of each segment is depicted in Figure 8, in combination with the EDA signal and the Wasserstein metric.



**Figure 6.** Thermal Domain: TRA signal—Jensen–Shannon metric with a bandwidth of  $10^{-4}$ . Distribution of extracted feature values for the feature trapezius\_second\_derivative\_features\_signal\_split\_equal\_part\_std of the TRA signal in combination with the selected segments with the Canberra metric and the distribution of the source domain samples (phasic domain). (**Left**) Value distributions for the baseline stimuli. (**Right**) Value distributions of the pain stimuli.



**Figure 7.** Thermal Domain: EDA signal—Wasserstein metric. Distribution of extracted feature values for the feature trapezius\_mean\_value\_first\_diff of the EDA signal in combination with the selected segments with the Wasserstein metric and the distribution of the source domain samples (phasic domain). (**Left**) Value distributions for the baseline stimuli. (**Right**) Value distributions of the pain stimuli.



**Figure 8.** Thermal domain. The segment selection frequency for the EDA signal in combination with the Wasserstein metric. On the  $x$ -axis the segment number is depicted. The  $y$ -axis represents the segment counter.

## 8. Conclusions

In this study, we analyzed a pain duration transfer, in combination with electric and thermal pain domains, respectively. To this end, we evaluated various distance metrics, with the aim of finding optimal segments of the tonic samples, which are represented best by the phasic training stimuli. Our results indicate that the TRA signal is the most reliable signal for the knowledge transfer from phasic to tonic stimuli. In combination with the thermal domain, the pain duration transfer task is more challenging in comparison to the electric domain. Based on the electric pain domain, a maximum of 68.0% mean accuracy is obtained with the TRA signal in combination with the Canberra and Chebychev metrics. With the TRA signal, we were able to obtain outcomes close to the results, which were reported for the evaluation of the individual segments. Based on the thermal pain domain, the highest mean accuracy value of 60.0% was obtained with the Jensen–Shannon metric and the bandwidth of the kernel density estimator set to  $10^{-4}$ .

In combination with the applied metrics, we were able to select segments which show similar distributions for the most important feature in comparison to the phasic samples. However, the selected segments were not always equal to the ones which performed best in the evaluation of the individual segments. Nevertheless, the selected segments seem to contain enough important information that can be used by the model.

**Author Contributions:** Conceptualization, T.B.R. and P.B.; methodology, T.B.R.; software, T.B.R.; validation, T.B.R. and P.B.; formal analysis, T.B.R.; investigation, T.B.R.; writing—original draft preparation, T.B.R. and P.B.; writing—review and editing, T.B.R., P.B., S.W. and F.S.; visualization, T.B.R.; supervision, S.W. and F.S.; project administration, S.W. and F.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

APR	Automated Pain Recognition
SVM	Support Vector Machine
AdaBoost	Adaptive Boosting
ECG	Electrocardiogram
EDA	Electrodermalactivity
EMG	Electromyogram
TRA	Trapezius
COR	Corrugator
ZYG	Zygomaticus
EFU	Early Fusion
B	Baseline
P	Pain
LOSO-CV	Leave-One-Subject-Out Cross-Validation
RBF	Radial Basis Function
RF	Random Forest
PB	Phasic Baseline
PP	Phasic Pain
TB	Tonic Baseline
TP	Tonic Pain

## References

- Werner, P.; Lopez-Martinez, D.; Walter, S.; Al-Hamadi, A.; Gruss, S.; Picard, R. Automatic Recognition Methods Supporting Pain Assessment: A Survey. *IEEE Trans. Affect. Comput.* **2019**, *13*, 530–552. [[CrossRef](#)]
- Merskey, H.; Albe-Fessard, D.; Bonica, J.; Carmen, A.; Dubner, R.; Kerr, F.; Pagni, C. Editorial: The need of a taxonomy. *Pain* **1979**, *6*, 247–252.
- Herr, K.; Coyne, P.J.; McCaffery, M.; Manworren, R.; Merkel, S. Pain assessment in the patient unable to self-report: Position statement with clinical practice recommendations. *Pain Manag. Nurs.* **2011**, *12*, 230–250. [[CrossRef](#)] [[PubMed](#)]
- Craig, K.D. The facial expression of pain Better than a thousand words? *APS J.* **1992**, *1*, 153–162. [[CrossRef](#)]
- Walter, S.; Gruss, S.; Frisch, S.; Liter, J.; Jerg-Bretzke, L.; Zujalovic, B.; Barth, E. What about Automated Pain Recognition for Routine Clinical Use? A Survey of Physicians and Nursing Staff on Expectations, Requirements, and Acceptance. *Front. Med.* **2020**, *7*, 990. [[CrossRef](#)]
- Hadjistavropoulos, H.D.; Ross, M.A.; Von Baeyer, C.L. Are physicians' ratings of pain affected by patients' physical attractiveness? *Soc. Sci. Med.* **1990**, *31*, 69–72. [[CrossRef](#)]
- Werner, P.; Al-Hamadi, A.; Gruss, S.; Walter, S. Twofold-Multimodal Pain Recognition with the X-ITE Pain Database. In Proceedings of the 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos, ACII Workshops 2019, Cambridge, UK, 3–6 September 2019; pp. 290–296. [[CrossRef](#)]
- Bellmann, P.; Thiam, P.; Schwenker, F. Pain Intensity Recognition—An Analysis of Short-Time Sequences in a Real-World Scenario. In *Artificial Neural Networks in Pattern Recognition, Proceedings of the 9th IAPR TC3 Workshop, ANNPR 2020, Winterthur, Switzerland, 2–4 September 2020*; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2020; Volume 12294, pp. 149–161.
- Wally, Y.; Samaha, Y.; Yasser, Z.; Walter, S.; Schwenker, F. Personalized k-fold Cross-Validation Analysis with Transfer from Phasic to Tonic Pain Recognition on X-ITE Pain Database. In *Pattern Recognition, Proceedings of the ICPR International Workshops and Challenges, Virtual Event, 10–15 January 2021*; Bimbo, A.D., Cucchiara, R., Sclaroff, S., Farinella, G.M., Mei, T., Bertini, M., Escalante, H.J., Vezzani, R., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2020; Volume 12666, pp. 788–802. [[CrossRef](#)]
- Ricken, T.B.; Bellmann, P.; Walter, S.; Schwenker, F. Pain Detection in Biophysiological Signals: Transfer Learning from Short-Term to Long-Term Stimuli Based on Signal Segmentation. In Proceedings of the MPRSS—7th IAPR Workshop on Multimodal Pattern Recognition for Social Signal Processing in Human Computer Interaction, Montreal, QC, Canada, 21–25 August 2022; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2022.
- Kächle, M.; Thiam, P.; Amirian, M.; Schwenker, F.; Palm, G. Methods for person-centered continuous pain intensity assessment from bio-physiological channels. *IEEE J. Sel. Top. Signal Process.* **2016**, *10*, 854–864. [[CrossRef](#)]
- Kessler, V.; Thiam, P.; Amirian, M.; Schwenker, F. Pain recognition with camera photoplethysmography. In Proceedings of the IPTA 2017 Seventh International Conference on Image Processing Theory, Tools and Applications, Montreal, QC, Canada, 28 November–1 December 2017; pp. 1–5.
- Bellmann, P.; Lausser, L.; Kestler, H.A.; Schwenker, F. Introducing bidirectional ordinal classifier cascades based on a pain intensity recognition scenario. In Proceedings of the International Conference on Pattern Recognition, Milan, Italy, 10–15 January 2021; Springer: Berlin/Heidelberg, Germany, 2021; pp. 773–787.
- Thiam, P.; Hihn, H.; Braun, D.A.; Kestler, H.A.; Schwenker, F. Multi-Modal Pain Intensity Assessment Based on Physiological Signals: A Deep Learning Perspective. *Front. Physiol.* **2021**, *12*, 720464. [[CrossRef](#)]
- Thiam, P.; Kessler, V.; Amirian, M.; Bellmann, P.; Layher, G.; Zhang, Y.; Velana, M.; Gruss, S.; Walter, S.; Traue, H.C.; et al. Multi-Modal Pain Intensity Recognition Based on the SenseEmotion Database. *IEEE Trans. Affect. Comput.* **2021**, *12*, 743–760. [[CrossRef](#)]
- Velana, M.; Gruss, S.; Layher, G.; Thiam, P.; Zhang, Y.; Schork, D.; Kessler, V.; Meudt, S.; Neumann, H.; Kim, J.; et al. The SenseEmotion Database: A Multimodal Database for the Development and Systematic Validation of an Automatic Pain- and Emotion-Recognition System. In *Multimodal Pattern Recognition of Social Signals in Human-Computer-Interaction, Proceedings of the 4th IAPR TC 9 Workshop, MPRSS 2016, Cancun, Mexico, 4 December 2016*; Revised Selected Papers 4; Schwenker, F., Scherer, S., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2016; Volume 10183, pp. 127–139. [[CrossRef](#)]
- Walter, S.; Gruss, S.; Ehleiter, H.; Tan, J.; Traue, H.C.; Crawcour, S.C.; Werner, P.; Al-Hamadi, A.; Andrade, A.O. The biovid heat pain database data for the advancement and systematic validation of an automated pain recognition system. In Proceedings of the 2013 IEEE International Conference on Cybernetics, CYBCO 2013, Lausanne, Switzerland, 13–15 June 2013; pp. 128–131. [[CrossRef](#)]
- Thiam, P.; Bellmann, P.; Kestler, H.A.; Schwenker, F. Exploring Deep Physiological Models for Nociceptive Pain Recognition. *Sensors* **2019**, *19*, 4503. [[CrossRef](#)]
- Bellmann, P.; Thiam, P.; Kestler, H.A.; Schwenker, F. Machine Learning-Based Pain Intensity Estimation: Where Pattern Recognition Meets Chaos Theory—An Example Based on the BioVid Heat Pain Database. *IEEE Access* **2022**, *10*, 102770–102777. [[CrossRef](#)]
- Badura, A.; Maslowska, A.; Mysliwiec, A.; Pietka, E. Multimodal Signal Analysis for Pain Recognition in Physiotherapy Using Wavelet Scattering Transform. *Sensors* **2021**, *21*, 1311. [[CrossRef](#)] [[PubMed](#)]
- Vapnik, V. Universal learning technology: Support vector machines. *NEC J. Adv. Technol.* **2005**, *2*, 137–144.

22. Othman, E.; Werner, P.; Saxen, F.; Fiedler, M.; Al-Hamadi, A. An Automatic System for Continuous Pain Intensity Monitoring Based on Analyzing Data from Uni-, Bi-, and Multi-Modality. *Sensors* **2022**, *22*, 4992. [[CrossRef](#)]
23. Gruss, S.; Geiger, M.; Werner, P.; Wilhelm, O.; Traue, H.C.; Al-Hamadi, A.; Walter, S. Multi-modal signals for analyzing pain responses to thermal and electrical stimuli. *JoVE (J. Vis. Exp.)* **2019**, *146*, e59057.
24. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
25. Al-Eidan, R.M.; Al-Khalifa, H.; Al-Salman, A. Deep-learning-based models for pain recognition: A systematic review. *Appl. Sci.* **2020**, *10*, 5984. [[CrossRef](#)]
26. Ricken, T.; Steinert, A.; Bellmann, P.; Walter, S.; Schwenker, F. Feature Extraction: A Time Window Analysis Based on the X-ITE Pain Database. In *Artificial Neural Networks in Pattern Recognition, Proceedings of the 9th IAPR TC3 Workshop, ANNPR 2020, Winterthur, Switzerland, 2–4 September 2020*; Schilling, F.P., Stadelmann, T., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2020; Volume 12294, pp. 138–148. [[CrossRef](#)]
27. Deza, M.M.; Deza, E. *Encyclopedia of Distances*; Springer: Berlin/Heidelberg, Germany, 2016.
28. Bellemare, M.G.; Danihelka, I.; Dabney, W.; Mohamed, S.; Lakshminarayanan, B.; Hoyer, S.; Munos, R. The Cramer Distance as a Solution to Biased Wasserstein Gradients. *arXiv* **2017**, arXiv:1705.10743.
29. Cramér, H. On the composition of elementary errors. *Scand. Actuar. J.* **1928**, *1928*, 13–74. [[CrossRef](#)]
30. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17*, 261–272. [[CrossRef](#)]
31. Breiman, L.; Friedman, J.H.; Olshen, R.A.; Stone, C.J. *Classification and Regression Trees*; Routledge: London, UK, 2017.
32. Taha, A.A.; Hanbury, A. An Efficient Algorithm for Calculating the Exact Hausdorff Distance. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *37*, 2153–2163. [[CrossRef](#)] [[PubMed](#)]
33. Harris, C.R.; Millman, K.J.; van der Walt, S.J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N.J.; et al. Array programming with NumPy. *Nature* **2020**, *585*, 357–362. [[CrossRef](#)] [[PubMed](#)]
34. 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.* **2011**, *12*, 2825–2830.
35. McKinney, W. Data Structures for Statistical Computing in Python. In Proceedings of the 9th Python in Science Conference, Austin, TX, USA, 28 June–3 July 2010; van der Walt, S., Millman, J., Eds.; pp. 56–61. [[CrossRef](#)]
36. Hunter, J.D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95. [[CrossRef](#)]
37. Waskom, M.L. seaborn: Statistical data visualization. *J. Open Source Softw.* **2021**, *6*, 3021. [[CrossRef](#)]
38. Walter, S.; Al-Hamadi, A.; Gruss, S.; Frisch, S.; Traue, H.C.; Werner, P. Multimodale Erkennung von Schmerzintensität und- modalität mit maschinellen Lernverfahren. *Schmerz* **2020**, *34*, 400–409. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.