Subject-independent Pain Recognition Using Physiological Signals and Para-linguistic Vocalizations

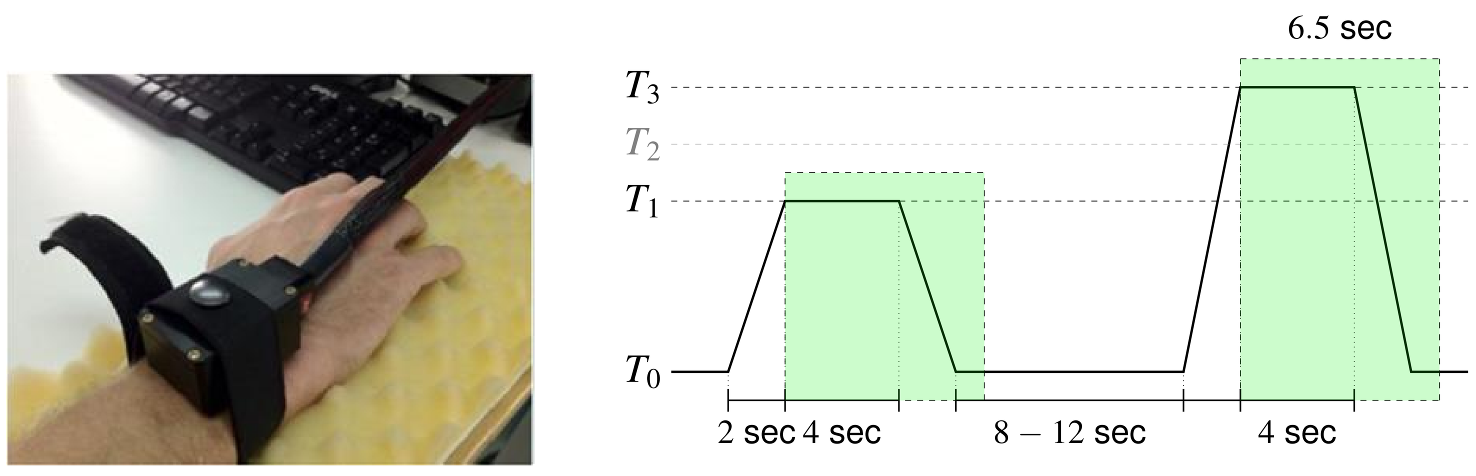
Nadeen Shoukry, Omar Elkilany, Patrick Thiam, Viktor Kessler, and Friedhelm Schwenker

Abstract

Pain is the result of a complex interaction among the various parts of the human nervous system. It plays an important role in the diagnosis and treatment of patients. The standard method for pain recognition is self-report; however, not all patients can communicate pain effectively. In this work, the task of automated pain recognition is addressed using para-linguistic and physiological data. Hand-crafted and automatically generated features are extracted and evaluated independently. Several state-of-the-art machine learning algorithms are applied to perform subject-independent binary classification. The *SenseEmotion* dataset is used for evaluation and comparison. Random forests trained on hand-crafted features from the physiological modalities achieved an accuracy of 82.61%, while support vector machines trained on hand-crafted features from the para-linguistic data achieved an accuracy of 63.86%. Hand-crafted features outperformed automatically generated features.

Dataset Description

The *SenseEmotion* dataset comprises multi-modal data from 40 healthy participants [1]. Pain was elicited in these participants using thermal stimuli (see Figure 1 for more details). A baseline temperature T_0 of 32 degrees Celsius was used in all of the experiments. At the start of an experimental session, calibration was conducted to determine the pain threshold T_1 and the tolerance threshold T_3 of each individual participant, where the pain threshold marks the point where the participant perceives pain and the tolerance threshold marks the point where the participant can no longer bear the pain. The mean of T_1 and T_3 was used to define another level labelled T_2. The resulting dataset consists of 120 events per participant, 30 for each heat level. In this work, the no-pain state corresponds to the baseline temperature T_0, while the pain state corresponds to the pain tolerance threshold T_3.



(a) (b)

Figure 1: Experimental settings. **Left:** Thermal simulator. **Right:** Pain stimulation. T_0: baseline temperature (32°C); T_1: pain threshold temperature; T_2: intermediate temperature; T_3: pain tolerance temperature.

The following modalities were used in this work:

- Audio recordings of para-linguistic vocalizations
- 4 Bio-physiological modalities (electromyographic activity of the trapezius muscle (EMG), galvanic skin response (GSR), respiration (RSP) and electrocardiogram (ECG))

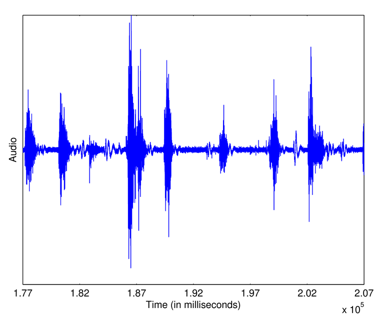
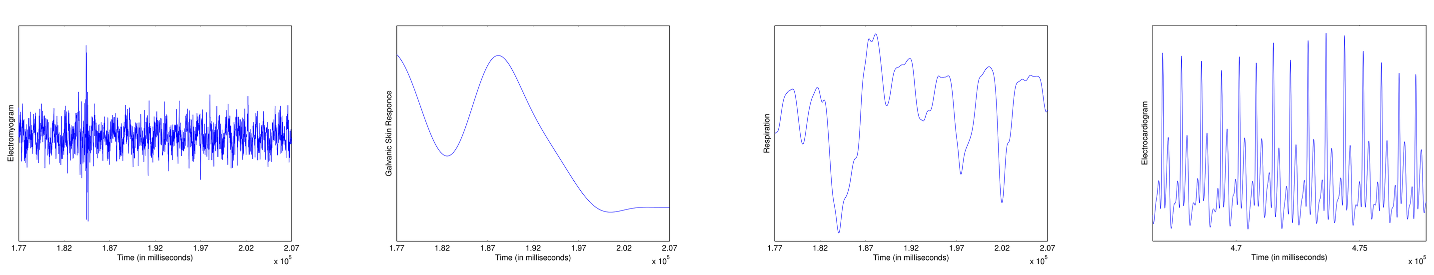


Figure 2: Audio signal.



(a) (b) (c) (d)

Figure 3: Bio-physiological modalities. (a) EMG, (b) GSR, (c) RSP, (d) ECG.

Feature Extraction

From the audio signal, several hand-crafted features were extracted. They were the Mel Frequency Cepstral Coefficients, Relative Spectral Perceptual Linear Predictive Coding, and the root mean square sound signal energy and logarithmic sound signal energy. Audio spectrograms (see Figure 4) were also created and fed to a pre-trained VGG16 model for automatic feature extraction.

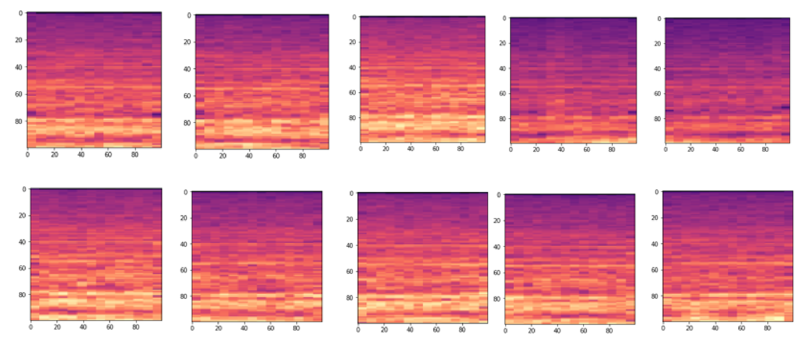
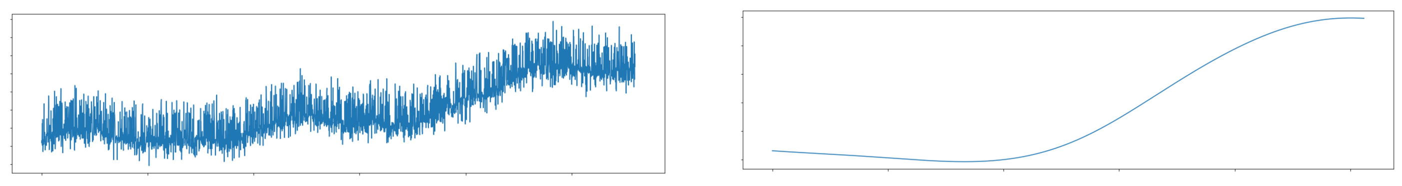


Figure 4: Data spectrograms for the para-linguistic audio signal during one of the pain events.

From the physiological signals, 331 hand-crafted features were extracted using the methodology presented in [2]. These features were then normalized before they were fed to classification algorithms. Furthermore, features were generated automatically using autoencoders, neural networks, and convolutional neural networks. To mitigate inter-subject variance, all the signals were normalized to fit in the range $[-1, 1]$ before automatic feature extraction (see Figure 5).



(a) (b)

Figure 5: The GSR signal recorded during one of the pain events. **Left:** before normalization. **Right:** after normalization.

Results

- Para-linguistic and bio-physiological data were used separately.
- Classification task: T_0 vs T_3.
- Evaluation setting: Leave-one-subject-out cross-validation.

Algorithm	Accuracy
SVM	63.86%
Voting Ensemble	63.28%
Naive Bayes	61.68%
Random Forests	61.20%
Bidirectional LSTM	50.12%

Table 1: A summary of the results obtained using the para-linguistic modalities.

Algorithm	Accuracy
Random Forests	$82.61 \pm 10.74\%$
Neural Network	$81.57 \pm 10.58\%$
RBF Kernel SVM	$81.34 \pm 10.62\%$
Polynomial Kernel SVM	$78.04 \pm 10.57\%$
1D CNN	$76.90 \pm 12.60\%$
Autoencoders	$70.00 \pm 12.00\%$

Table 2: A summary of the results obtained using the physiological modalities.

Conclusions

- MFCCs were the most important para-linguistic features.
- GSR was the most discriminative physiological modality.
- Hand-crafted features outperformed automatic feature extraction, but each produced different errors.
- Data collection should be coupled with the development of new methods to deal with inter-subject variance.
- A combination of both hand-crafted and automatically generated features may improve classification accuracy.

Contact

Patrick Thiam (patrick.thiam@uni-ulm.de)
Institute of Neural Information Processing
Ulm University, Germany
http://www.uni-ulm.de/en/in/institute-of-neural-information-processing.html

References

[1] Velana et al. *The SenseEmotion Database: A Multimodal Database for the Development and Systematic Validation of an Automatic Pain- and Emotion-Recognition System*, In: MPRSS 2016, 127–139

[2] Thiam et al. *Multi-modal Pain Intensity Recognition based on the SenseEmotion Database*, In: IEEE Transactions on Affective Computing, 2019