Mental State Estimation of People with PIMD using Physiological Signals

Gašper Slapničar
gasper.slapnicar@ijs.si
Jožef Stefan Institute, Jožef Stefan IPS
Jamova cesta 39
Ljubljana, Slovenia

Erik Dovgan
erik.dovgan@ijs.si
Jožef Stefan Institute
Jamova cesta 39
Ljubljana, Slovenia

Jakob Valič
jakob.valic@ijs.si
Jožef Stefan Institute
Jamova cesta 39
Ljubljana, Slovenia

Mitja Luštrek
mitja.lustrek@ijs.si
Jožef Stefan Institute
Jamova cesta 39
Ljubljana, Slovenia

ABSTRACT

People with profound intellectual and multiple disabilities are a very diverse and vulnerable group of people. Their disabilities are cognitive, motor and sensory, and they are also incapable of symbolic communication, making them heavily reliant on caregivers. We investigated the connection between physiological signals and inner states as well as communication attempts of people with PIMD, using signal processing and machine learning techniques. The inner states were annotated by expert caregivers, and several heart rate variability features were computed from photoplethysmogram. We then fed the features into hyper-parameter-tuned classification models. We achieved the highest accuracy of 62% and F1-score of 0.59 for inner state (pleasure, displeasure, neutral) classification using Extreme Gradient Boosting, which notably surpassed the baseline.

KEYWORDS

PIMD, mental state, physiological signals, classification

1 INTRODUCTION

People with profound intellectual and multiple disabilities (PIMD) often face extreme difficulties in their day-to-day life due to severe cognitive, motor and sensory disabilities. They require a nearly everpresent caregiver to help them with most tasks. Additionally, they are unable to communicate their feelings or express their current mental state in a traditional symbolic way. This causes a gap between a caregiver and the care recipient, as it can take an extended period of time for the caregiver to recognize any potential patterns and their relationship with the mental state of the care recipient.

The aforementioned reasons call for a technological solution that might help bridge the gap between the caregiver and the care recipient and help the former better understand the latter. The INSENSION project [8] aims to develop such assistive technology, which takes into account many aspects of the care recipient. The aim is to both bridge the previously mentioned gap as well as empower the people with PIMD to be able to interact with their surroundings through technology. One part of the system considers the patterns in a person's gestures and facial expressions, which might have some significance and correlation to their behavioural and mental state, or their communication attempt. The initial solution dealing with this part was already described by Cigale et al. [1, 2]. In this paper, we instead focus on exploring the relationship between the physiological response of the body and the mental state of the people with PIMD by using features computed from photoplethysmogram (PPG). PPG is a periodic signal, where each cycle corresponds to a single heart beat. We obtained the PPG in two different ways: 1.) by using a high-quality wearable Empatica E4 with an optical sensor measuring the reflection of light from the skin and 2.) by using a contact-free RGB camera mounted on a wall, which records the color changes of the skin pixels. The features were then used to train classification models, which predicted the person's inner state or communication attempt.

The rest of this paper is structured as follows: we first investigate the related work in Section 2, then we describe the data collected and used in the experiments in Section 3. We continue with the methodology and experimental setup description in Section 4, and conclude with results and discussion in Section 5.

Information society '20, October 5–9, 2020, Ljubljana, Slovenia

2 RELATED WORK

The connection between physiological parameters and mental states is a mature and highly-researched field when it comes to average healthy people.

Schachter et al. [6] investigated the emotional state of people as a function of cognitive, social and physiological state. Several propositions were made and experimentally confirmed, supporting the overall connection between emotional and physiological state.

Cigale et al. [1, 2] explored the communication signals of people with PIMD, which are atypical and idiosyncratic. They highlighted the challenging interpretation of these signals and their meaning and suggested how technology could help overcome the gap between caregivers and care recipients. Some models were proposed that take the person's non-verbal signals (NVS) as input and classify their inner state or communication attempt.

Kramer et al. [3] highlighted the challenges of analysing the NVS in people with PIMD, as they are difficult to discern, instead focusing on physiological body responses. They conducted a research in which the expressions of three emotional states of one person with PIMD were recorded during nine emotion-triggering situations. They collected heart rate (HR) and skin conductance level (SCL), and investigated the connection between these two physiological signals and the emotional state. They found higher SCL activity during anger or happiness and lower SCL activity during relaxation or neutral state.

Vos et al. confirmed that HR and skin temperature allow the same conclusions in people with PIMD and people without disabilities, regarding positive and negative emotion. This finding gives additional motivation to our work, showing that the connection between physiological and mental state also holds for people with PIMD [9].

3 DATA

We created a recording setup in the INSENSION project, which uses two Logitech C920 cameras capable of recording full HD (1920x1080) resolution video at 30 frames per second (fps). The cameras were setup perpendicular to one another to record from two distinct angles, allowing for decent facial exposure even when the face changes direction. The caregivers were instructed to attempt to conduct their activity in front of one of the cameras whenever possible. Additionally, the subjects were given an Empatica E4 wristband, which served both as the ground truth for PPG, as well as a fall-back mechanism for obtaining physiological signals in cases when camera is unreliable or unavailable. The wristband records PPG at 64 Hz, allowing for capture of reasonable morphological details. The temporal synchronization between the video and ground truth was ensured to the best of our abilities using suitable protocols and checks.

With the described setup, we obtained 48 recording sessions, each lasting between 10 and 30 minutes. Five sessions were eliminated immediately, as there was a large mismatch between the duration of the video and the duration of the ground truth, which may happen due to several reasons, such as a caregiver forgetting to turn on the wristband during a session or the wristband losing connection.

It is important to note that the recordings were made in a natural way, as the caregivers were not given any additional restrictions other than to be in front of the camera when possible. In practice this means that large parts of some recordings might be useless due to the person with PIMD being turned away or the caregiver blocking them. Examples of good and bad sessions are shown in Figure 1.

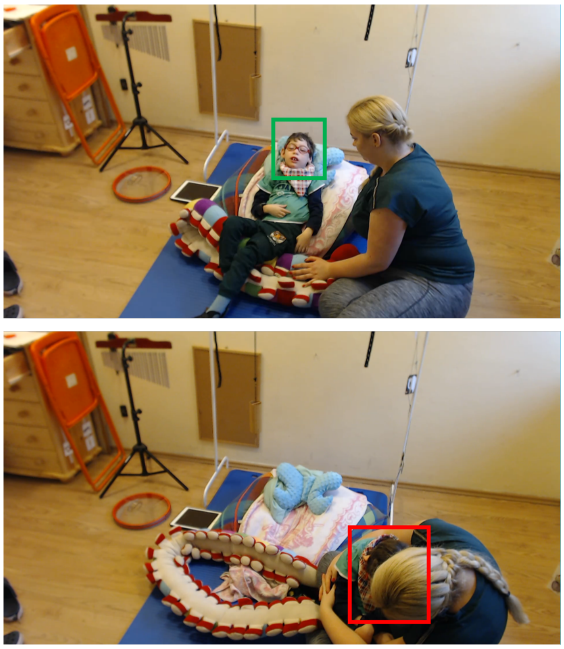

Figure 1: Example of good (green) and bad (red) video recordings.

3.1 Annotating the ground truth

In order to classify mental states of people with PIMD, we first required the ground truth annotations. As it is generally difficult to obtain such ground truth, we relied on the expert knowledge of partners in the project who specialize in education of people with special needs, alongside the caregivers, who know their care recipients the best. Together they devised an annotation schema, in which they annotated inner states and communication attempts of people with PIMD and can take the values given in Equations 1 and 2.

$$InnerState = \begin{cases} displeasure & \text{if 1, 2 or 3} \\ neutral & \text{if 4, 5 or 6} \\ pleasure & \text{if 7, 8 or 9} \end{cases} \quad (1)$$

The three numbers within each mental state indicate the intensity, where a lower number for displeasure means more intense displeasure, and a higher number for pleasure indicates more intense pleasure.

$$CommAttempt = \begin{cases} protest \\ comment \\ demand \end{cases} \quad (2)$$

The caregivers were tasked with annotation of videos, looking at camera recordings and marking inner states and communication attempts in time, always marking the start and end of each recognized state, regardless of duration (can be a few seconds or a few minutes). Naturally, large periods remained where nothing was annotated, as the experts were either not sure or did not recognize any of the pre-defined states. This does not mean that nothing is happening in those periods, but simply that the inner experience of the person with PIMD is unknown. Thus, we added an additional class value for the areas where nothing was annotated – unknown.

4 METHODOLOGY OF MENTAL STATE ESTIMATION

Having both the ground truth annotations and physiological data and videos, we then investigated two approaches: 1.) we attempted to reconstruct PPG from the camera recordings in a contact-free manner and use the reconstructed rPPG (remote PPG) to calculate features and to classify inner state and communication attempt and 2.) we directly used the Empatica ground truth PPG to calculate features to be used in the same classification task.

4.1 Using rPPG Reconstruction

In order to obtain the remote PPG, we used a rather standard pipeline, which was updated with a convolutional neural network in order to further enhance the rPPG. At a high level, the pipeline consists of detection or region of interest (ROI), extraction of red, green and blue signal components (RGB), detrending and band-pass filtering of RGB, rPPG reconstruction using the Plane

Orthogonal to Skin (POS) algorithm, band-pass rPPG filtering (0.5 to 4.0 Hz), and rPPG enhancement via deep learning. Details were already described in our previous work [7] and are not subject of this paper.

We ran the pipeline described above on 30-second segments of video using a sliding window without overlap. We decided to use 30 seconds due to the nature of some frequency features that we chose, as frequency analysis makes sense once a reasonable number of periods are available - in our case this means that a sufficient number of heart cycles must be available. Additionally, this length makes sense as we are primarily attempting to predict inner states, which do not change extremely in such a short time span. An example output of the pipeline is shown in Figure 2.

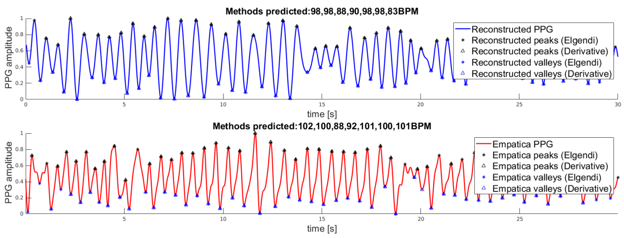

Figure 2: Example of a good rPPG segment obtained with our pipeline.

We then used the rPPG to compute several heart rate variability (HRV) features. These are known to be well-correlated with stress, cognitive load, conflict experience and other inner states [5, 4]. A detailed list of computed features is given in Table 1.

4.2 Using Empatica PPG

The Empatica records PPG directly on the skin, thus making the raw PPG readily available, without the need for additional reconstruction. Still, due to subject arm and wrist movements, we opted to use similar preprocessing steps used previously, namely detrending and band-pass filtering, as the signal can sometimes be quite noisy.

We computed the same set of features and window length as before (see Table 1), and used them in the same classification task, attempting to recognize inner states and communication attempts.

5 EXPERIMENTS AND RESULTS

Once both the input (HRV features) and output (annotations) were known, we investigated six classification algorithms (k Nearest Neighbours, Decision Trees, Random Forest, Support Vector Machines, AdaBoost and Extreme Gradient Boosting) for this task, always training separate models for inner state and communication attempt. We always compared each algorithm against a baseline majority vote classifier using two metrics, accuracy and F1-score.

5.1 Using Empatica PPG

We started our evaluation using the Empatica data, as it is more reliable, since the PPG reconstruction is not needed. At the time of evaluation, we had annotations for 15 recording sessions in which 2 different people with PIMD are present. Using the chosen 30-second window, we initially had 417 segments of Empatica PPG available. The unknown class label heavily skewed the data for both classes, and there is no way to know which (other) class label it actually belongs to, so we decided to exclude it from evaluation. This left us with 272 instances for class inner state and 80 instances for class communication attempt, which was annotated more sparsely. The final distributions for each class are shown in Figure 3.

Table 1: List of computed HRV features.

Feature	Description
HRmean	$60/mean(NN)$
HRmedian	$60/median(NN)$
IBImedian	$median(NN)$
SDNN	$std(NN)$
SDSD	$std(abs(NN'))$
RMSSD	$sqrt(mean((NN')^2))$
NN20 and NN50	The number of pairs of successive NNs that differ by more than 20ms and 50ms
pNN20 and pNN50	The proportion of NN20 and NN50 divided by total number of NNs
SDbonus1	$sqrt(0.5) * SDNN$
SDbonus2	$sqrt(abs(2 * SDSD^2 - 0.5 * SDSD^2))$
VLF	Area under periodogram in the very low frequencies
LF	Area under periodogram in the low frequencies
HF	Area under periodogram in the high frequencies
LFnorm and HFnorm	Area under periodogram in the low and high frequencies, normalized by the whole area under periodogram
LFdHF	LF/HF
where *std*	is standard deviation,
abs	is absolute value,
X'	is the first order derivative,
sqrt	is the square root and
NN	are the beat-to-beat intervals.

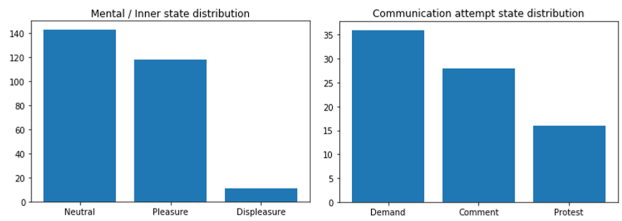

Figure 3: Distributions of both classes.

Initially we conducted a 5-fold cross validation (CV) to investigate the best hyper-parameters using a grid search. Once the hyper-parameters were determined, we ran a separate experiment, using the best overall hyper-parameters for each model. Again, we ran a 5-fold CV with the best hyper-parameter settings obtained on the full data to validate the performance. All the investigated algorithms (from the Scikit-learn and XGBoost packages) and their corresponding sets of optimized parameters with the best values are available from the authors, but are not listed here due to space restrictions. Results of our evaluation in terms of accuracy and F1-score for both classes are given in Table 2.

Table 2: Accuracy and F1 score for both classes.

Algorithm	$ACC_{mentalstate}$	$F1_{mentalstate}$
Baseline (majority)	0.52	0.36
kNN	0.55	0.55
Tree	0.54	0.56
RF	0.57	0.56
SVM	0.55	0.52
AdaBoost	0.59	0.56
XGB	**0.62**	**0.59**
Algorithm	$ACC_{commattempt}$	$F1_{commattempt}$
Baseline (majority)	0.45	0.27
kNN	0.42	0.42
Tree	0.41	0.39
RF	0.46	0.43
SVM	0.43	0.34
AdaBoost	0.43	0.41
XGB	**0.48**	**0.45**

5.2 Using rPPG reconstruction

Using the rPPG for evaluation proved to be more difficult, as we only had limited amount of good subsequent facial crops from the videos, while also having a limited amount of annotations. This meant that the overlap between the two was very small – we had only 12 such 30-second segments for inner state and only 6 for communication attempt. Such a low amount of data is infeasible to be used in a realistic evaluation scheme (not even all three different class labels were present), so we instead decided to use the models previously trained on the Empatica data, to classify these instances obtained via the rPPG. We achieved reasonably high accuracy of 75% and F1-score of 0.84 for inner state and low accuracy of 33% and F1-score of 0.33 for communication attempt. Confusion matrices are shown in Figure 4.

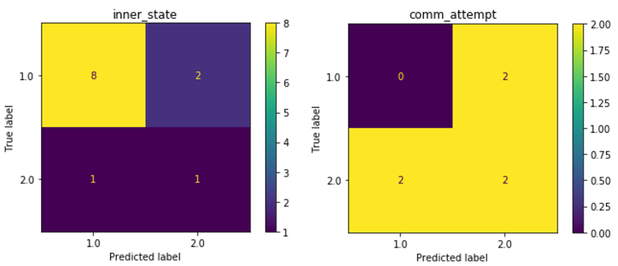

Figure 4: Confusion matrices for classifying rPPG instances using models trained on Empatica data. For inner state, the class values are 1.0="neutral" and 2.0="pleasure". For communication attempt 1.0="comment" and 2.0="demand".

6 CONCLUSION

We conducted an initial investigation of the connection between physiological signals and mental states of people with PIMD, attempting to classify their inner states and communication attempts. We used HRV features computed from the PPG obtained with an Empatica E4 wristband and investigated the performance of such models on instances obtained via rPPG. XGB has shown the best performance, achieving accuracy of 62% and F1 score of 0.59 for inner state, and accuracy of 48% and F1 score of 0.45 for communication attempt, notably surpassing the baseline majority classifier.

Limitations of our work lie in low number of instances for communication attempt and little variety in subjects, having just two for which annotations were available. Additionally, the evaluation using the rPPG is limited, as we had very few instances for which both high-quality segments of video and annotations were available. Thus, the focus of future work should be on gathering more data and conducting a more extensive evaluation of the methods, which is planned in the trial stage of the INSENSION project.

ACKNOWLEDGMENTS

This work is part of the INSENSION project that has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No. 780819. The authors also acknowledge the financial support from the Slovenian Research Agency (ARRS).

REFERENCES

[1] Matej Cigale and Mitja Luštrek. 2019. Multiple knowledge categorising behavioural states and communication attempts in people with profound intellectual and multiple disabilities. In *AmI (Workshops/Posters)*, 46–54.

[2] Matej Cigale, Mitja Luštrek, Matjaž Gams, Torsten Krämer, Meike Engelhardt, and Peter Zentel. 2018. The quest for understanding: helping people with PIMD to communicate with their caregivers. *INFORMATION SOCIETY-IS 2018*.

[3] Torsten Krämer and Peter Zentel. 2020. Expression of emotions of people with profound intellectual and multiple disabilities. A single-case design including physiological data. *Psychoeducational Assessment, Intervention and Rehabilitation*, 2, 1, 15–29.

[4] Richard D Lane, Kateri McRae, Eric M Reiman, Kewei Chen, Geoffrey L Ahern, and Julian F Thayer. 2009. Neural correlates of heart rate variability during emotion. *Neuroimage*, 44, 1, 213–222.

[5] Junoš Lukan, Martin Gjoreski, Heidi Mauersberger, Annekatrin Hoppe, Ursula Hess, and Mitja Luštrek. 2018. Analysing physiology of interpersonal conflicts using a wrist device. In *European Conference on Ambient Intelligence*. Springer, 162–167.

[6] Stanley Schachter. 1964. The interaction of cognitive and physiological determinants of emotional state. In *Advances in experimental social psychology*. Volume 1. Elsevier, 49–80.

[7] Gašper Slapničar, Erik Dovgan, Pia Čuk, and Mitja Luštrek. 2019. Contact-free monitoring of physiological parameters in people with profound intellectual and multiple disabilities. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*.

[8] Poznań Supercomputing and Networking Center. 2017. The INSENSION project. https://www.insension.eu/.

[9] Pieter Vos, Paul De Cock, Vera Munde, Katja Petry, Wim Van Den Noortgate, and Bea Maes. 2012. The tell-tale: what do heart rate; skin temperature and skin conductance reveal about emotions of people with severe and profound intellectual disabilities? *Research in developmental disabilities*, 33, 4, 1117–1127.