Wrocław University of Science and Technology
Doctoral School

InterDocSchool Project

POLISH NATIONAL AGENCY FOR ACADEMIC EXCHANGE

Research Trip Summary Report

Task 2. Foreign mobility of WUST doctoral students

I. Data of the doctoral student

1. Full name: **Dominika Kunc**

2. Year of studies: **I year**

3. Educational discipline: **Information and Communication Technology**

II. Foreign research trip (research visit)

1.Research institute in which the foreign research was implemented: **University of Southern California, Los Angeles, the United States of America**

2. Name and surname of the host person (mentor): **prof. Shrikanth (Shri) Narayanan**

3. Dates of the research trip: **26.03.2023-23.04.2023**

4. Title and date of a seminar delivered during the research trip:
"Physiological Affective Map", 17.04.2023

5. Description of work carried out during the research trip:

During my research visit to the Signal Analysis and Interpretation Laboratory (SAIL) at the University of Southern California, Los Angeles, together with the Signal Analysis and Interpretation Lab (SAIL) team, we worked on the representation learning of physiological signals collected in everyday life using wearable devices. As I was visiting SAIL in December 2022, we discussed the preliminary setup for our research. After discussing the approaches from the literature, we decided to start with 15 seconds long ECG signals collected in real life using wearable devices from the TILES 2018 dataset. As a one-month research visit is not long enough to perform the whole range of research experiments, we decided to split the work. The part I was mainly focused on was validating the representations delivered by the SAIL team, which were trained using a contrastive learning approach.

To validate these representations, I used visualization and clustering methods, along with distance metrics. I tested several methods for representation visualization (TSNE, PacMAP, UMAP, and PCA) and clustering (KMeans, DBSCAN, and Agglomerative Clustering). I also measured the distances between representations to determine whether samples from the same participant were close to each other in the representation space. Additionally, I tested the representations on a different dataset, CASE, to evaluate their performance on new data and emotion-related labels.

I was also involved in another, more brainstorming part of the research, focused on special events and their influence on human behavior - specifically how hospital residents react and behave after

the patient's death. Together with my supervisor, dr. inż. Stanisław Saganowski and the SAIL Team, we discussed possible research directions and approaches to this problem.

6. Description of the main results obtained:

The main findings from my research trip were that the physiology representations tended to group with other samples from the same participant. This suggests that each individual has unique patterns and characteristics in their physiological signals, which is a reasonable basis for personalization methods. It was confirmed by the visualization methods and the clustering metrics - the best clustering results were obtained for the number of clusters close to the number of participants (elbow method - the trade-off between complexity/granularity and performance).

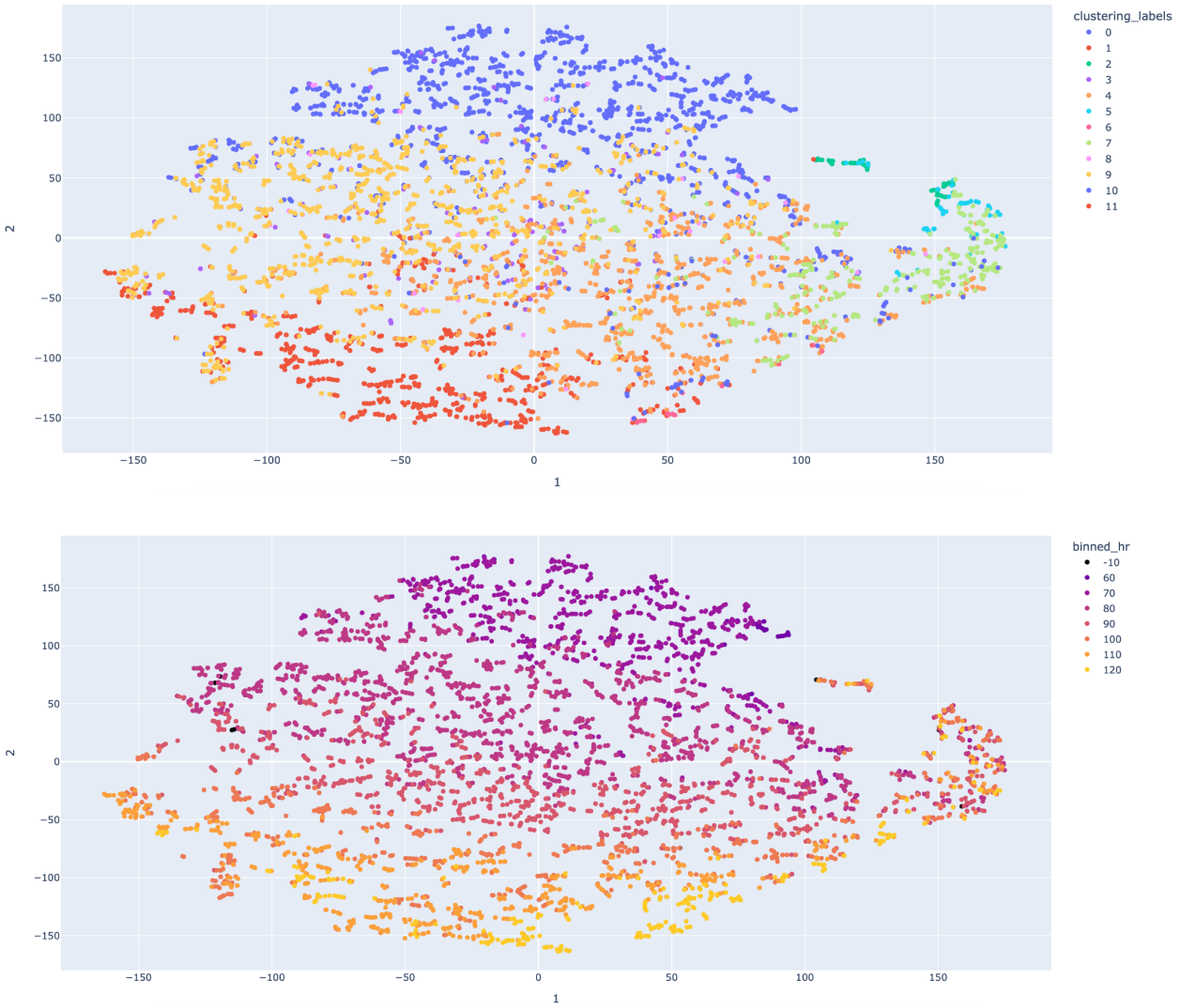

Fig. 1. Visualization of TILES participant's representations colored by: upper - clustering label from best resulting Agglomerative Clustering method, lower - value of heart rate (bins of 10)

Another important finding was that samples with similar heart rate values were placed near each other in the representation space, indicating that the trained representation preserved information about heartbeats from the original ECG signals. It showed that the self-supervised representation learning method maintained relevant information. Moreover, the unsupervised clustering algorithms (without specifying the number of clusters) tend to create clusters similar to the groups of heart rates [Fig. 1]. The number of clusters is greater than the number of heart rate bins, which should not be surprising, as the clustering methods might catch the relations between samples from the same bin.

Tests conducted on the CASE dataset revealed another important findings: 1) even though the representation learning model was trained on a different dataset (and so different people) the participants' samples are mostly grouped near each other, 2) samples collected during the presentation of stimuli designed to arouse negative valence tended to occur in one part of the embedding space, even though they were from different people. This indicates that although there is individuality in human physiology, some affective states can be reflected in physiology in similar ways.

However, the study had some limitations. I only tested one representation learning approach on two datasets, so testing these representations on more datasets and downstream tasks is crucial. Additionally, there is a need to establish more formal metrics for validation, as it is challenging to evaluate the quality of derived representations when there are no consistent labels in different datasets.

Future research will focus on developing new representation learning methods that incorporate personalization and other physiological wearable-based modalities, such as blood volume pulse, heart rate variability, or accelerometer data. Additionally, further tests on more datasets will be conducted to validate the representations on a broader range of cases and downstream tasks.

7. Future collaborations (if applicable):

The joint research on self-supervised representation learning will be continued remotely. We plan to validate another representation learning model and compare the two to see how the training objectives influence the representations. Hopefully, we will have more opportunities to collaborate during other research visits.

8. Title and date of a seminar presenting the results of the trip delivered at Wroclaw University of Science and Technology after returning from the research trip:

"Representation learning for physiological signals collected in everyday life using wearables - a summary of research trip to University of Southern California", 22.05.2023

III. Doctoral student's signature

……..

…….

(Date) (doctoral student's signature)

IV. Confirmation and information from the host

1. Confirmation of compliance of the information contained in the report: I CONFIRM / DO NOT CONFIRM. (*In justified cases, the confirmation of the host may be sent by e-mail to the Dean's Office of the Doctoral School email: interdocschool@pwr.edu.pl*)

2. Additional information and comments

……
……
……..…..

……

……..……

(Date) (signature(s) of Host)