E-PREVENTION: THE ICASSP-2023 CHALLENGE ON PERSON IDENTIFICATION AND RELAPSE DETECTION FROM CONTINUOUS RECORDINGS OF BIOSIGNALS

A. Zlatintsi[1], *P. P. Filntisis*[1], *N. Efthymiou*[1], *C. Garoufis*[1], *G. Retsinas*[1], *T. Sounapoglou*[2], *I. Maglogiannis*[3], *P. Tsanakas*[1], *N. Smyrnis*[4] *and P. Maragos*[1]

[1]School of ECE, National Technical University of Athens, 15773 Athens, Greece,

[2]Blockachain PC, 55535 Thessaloniki, Greece, [3]Department of Digital Systems, University of Piraeus, 18534 Pireas, Greece,

[4] Laboratory of Cognitive Neuroscience and Sensorimotor Control, UMHRI "COSTAS STEFANIS, 11527 Athens, Greece,

{nzlat, maragos}@cs.ntua.gr, {nefthymiou, filby, gretsinas}@central.ntua.gr, cgaroufis@mail.ntua.gr

ABSTRACT

The e-Prevention challenge concerns the analysis and processing of long-term continuous recordings of biosignals recorded from wearable sensors, i.e., accelerometers, gyroscopes and heart rate monitors embedded in smartwatches, as well as sleep information and daily step count, in order to extract high-level representations of the wearer's activity and behavior, termed as digital phenotypes. The ability of these digital phenotypes to quantify behavioral patterns and traits will be evaluated in two different tasks: 1) Person Identification, and 2) Relapse Detection in patients in the psychotic spectrum. The long-term data that will be used in this challenge have been acquired during the course of the e-Prevention project, an innovative integrated system for medical support that facilitates effective monitoring and relapse prevention in patients with mental disorders (i.e, schizophrenia and bipolar disorder). Specifically, the data were continuously collected from patients for a monitoring period of up to 2.5 years, while from the control subgroup for a period of 3 months, constituting one of the largest of its kind ever recorded.

***Index Terms*—** Person Identification, Relapse Detection, Psychotic Disorders, Smartwatch Wearables, Data Challenge

1. INTRODUCTION

Nowadays, wearable technologies foster unique opportunities for designing novel personalized intelligent electronic services that can address various well-being issues, and improve the relapse course in psychotic patients, thus having the potential to revolutionize psychiatry and its clinical practice [1, 2, 3]. Indeed, the broad adoption of wearable products, such as smartwatches and fitness trackers, has led to the emergence of the interdisciplinary field of digital phenotyping [4], which encompasses the *in situ* quantification of human behavior and traits (the "phenotype") by utilizing the sensors included in these low-cost devices [5] for the collection of physiological signals. Such wearables collect multimodal data, usually using accelerometers, gyroscopes and heart rate monitors among others, to measure the user's physical and kinetic activity. Other works in the field utilize sensorial data from smartphones [6], offering promising evidence for quantifying digital human phenotypes, as well. This abundance of sensory data has actually kickstarted the development of several applications focused on health monitoring, as well as other tasks, such as emotional well-being, sleep tracking, and physical activity detection [7, 8].

This research has been financed by the European Regional Development Fund of the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH–CREATE–INNOVATE (project acronym: e-Prevention, code: T1EDK-02890/MIS: 5032797)

Concerning person identification from wearable signals, recent approaches utilize supervised learning or representation learning using CNNs [9], an approach that effectively generalizes in user-diverse datasets, while being able to disentangle the recorded signals and sensor noise with proper data augmentation. On the other hand, a suitable approach for relapse detection is non-supervised sensor-based anomaly detection, the importance of which has been highlighted during recent years and the pandemic, through the clinical mass adoption of telehealth [10]. This approach is especially suitable for mental health monitoring, where the availability of data corresponding to relapsing states is scarce, and has been applied on data collected from various passive sensors [11, 12].

However, the public access to large user-diverse datasets of physiological signals is limited, especially in conjunction with mental health indicators. As a result, the e-Prevention challenge[1] is designed to encourage and foster research on a large-scale collection of raw biosignals from both a group of patients in the psychotic spectrum and a group of healthy controls, in two different, yet complementary tasks of major importance: i) Studying the correlation of the raw signals to user-specific behavioral patterns via **person identification**, and ii) using those signals as biomarkers of psychotic symptomatology through **relapse detection** in psychotic patients. The ability to detect relapsing states from unobtrusive sensors is a first step towards the emergent task of determining biomarkers that correlate with the mental state of psychotic patients, and could eventually lead to relapse prediction and finally their prevention.

2. DATASET

During the course of the e-Prevention project[2], a total of 62 people (39 patients in the psychotic spectrum and 23 healthy controls) were recruited at the University Mental Health, Neurosciences and Precision Medicine Research Institute "Costas Stefanis" (UMHRI) in Greece, and the protocol of the project was approved by the Ethics Committee of the Institution. All participants were provided with a Samsung Gear S3 smartwatch that monitored the user's linear acceleration and angular velocity (m/s^2 and deg/s, sampled at 20Hz), heart rate variability and RR intervals (sampled at 5Hz), sleeping schedule, and steps. This information was continuously collected from the patients for a monitoring period of up to 2.5 years, while the same data were collected from the control subgroup for a provisional period of 3 months. The resulting dataset [1] is one of the largest of its kind ever recorded, with a total of approximately 20000 human-days of collected data spread among all participants. The data were anonymized, and each participant in the study was assigned a unique

[1]Challenge website: https://robotics.ntua.gr/eprevention-sp-challenge/

[2]e-Prevention Project website: http://eprevention.gr

ID as an identifier. The clinicians annotated patients' relapse periods according to their monthly assessments and communication with the attending physician or the family. For the purposes of this challenge, we provide two subsets of the dataset, one for each challenge track.

3. CHALLENGE TRACKS AND BASELINES

3.1. First Track - Person Identification

The goal in this task is to identify the watch wearer by forming and classifying their digital phenotypes from the recorded biosignals.

Data Format and Evaluation Metrics: We provide a stratified split of the dataset (both patients and controls), consisting of 46 users, and including data of approximately two and a half month per user. The data are divided into training, validation and testing splits. The training and validation splits contain the raw sensor recordings, sleeping and walking information, as well as the unique ID corresponding to the identity of the respective watch wearer as the ground truth. The testing data consist solely of the raw recordings and the sleeping/walking information. As the evaluation metric, we use the unweighted accuracy (UA, %).

Data Preprocessing: We normalized the accelerometer and gyroscope values into the 0-1 interval, after removing incorrect sensor measurements. Similarly, we kept heart rate values smaller than 255, and RR interval values smaller than 2000.

Baseline Implementation: For the baseline of the first track we trained an 1D CNN with 5 convolutional layers, including Batch Normalization (BN) and ReLU activations. After the last BN layer we use Adaptive Average Pooling and a final fully connected layer to predict the user logits. During training, we sample a random contiguous 3H segment of the day, provided that it included at least 2.5H of valid data. Then, we impute any missing timestamps using nearest neighbour interpolation, resulting in a sequence of 721×8 features (we did not use sleeping, temporal or step information). Using this imputed sequence, the ID of the respective user is predicted. During evaluation, we select all contiguous 3H segments with at least 1H of valid data, impute them again with nearest neighbors, and use voting over all segments to select the final predicted user ID of the day, resulting in a final validation score of ca. 62%.

3.2. Second Track - Relapse Detection

In this task, we want to detect the appearance of patients' relapses, based on the smartwatch measurements.

Data Format and Evaluation Metrics: The provided dataset constitutes a subset of the full dataset, with data derived from 10 patients. Similar to the first track, the data are split into training, validation and testing data. The training/validation data contain the raw sensor recordings and sleeping/walking data matched with the unique ID of the patient, as well as timestamps and the starting and ending dates for every relapse, while the testing data include sensor recordings and patient IDs, but no information related to relapses. In this case, the training split contains only data acquired while the patient condition was stable, while the validation and testing splits span both stable and relapsing periods. As evaluation metrics, we utilize the PR-AUC and ROC-AUC scores, applied on the daily anomaly scores; days corresponding to relapsing periods are expected to record higher anomaly scores.

Data Preprocessing: The mean norm of the linear acceleration and the angular velocity, the mean heart rate and RR intervals, the width of the Poincare ellipse, and the normalized low and high frequency powers of the Lomb-Scargle periodogram were extracted from 5-min slices of the original data. Moreover, daily sinusoidal encoding was applied on the timestamps, and the percentage of valid 5-sec measurements was also calculated for each 5-min slice, resulting in 10 features per slice. Missing features were handled by median interpolation for intervals up to 3H; larger intervals of missing values were discarded. These features were stacked into 2D tensors of size 48×10, thus covering 4H each, with an 1H hop length. Finally, they were standardized per-patient and flattened into a 480-D vector.

Baseline Implementation: An 1-layer linear autoencoder was trained to reconstruct the input vectors, with a bottleneck dimension equal to $N = 60$ and a LeakyReLU activation. During inference, input tensors corresponding to 4H intervals were first standardized according to their respective per-patient transform, and then fed into the autoencoder. Afterwards, for each input vector, we calculate the anomaly score as the distance between the per-feature mean of the autoencoder output and the estimated multivatiate normal distribution of the feature vectors in the training set. Finally, the daily anomaly score is computed as the median of the anomaly scores corresponding to the 4H feature vectors of each day. Application of the above methodology in the provided validation set yields a PR-AUC score of 0.635 and a ROC-AUC score of 0.578.

4. CONCLUSIONS

In this paper, we introduce e-Prevention, the ICASSP-2023 Grand Challenge on Person Identification and Relapse Detection from continuous recordings of biosignals acquired from smartwatches, describing the datasets, data preprocessing, baselines and evaluation protocols employed for both its tracks. Through the release of a large-scale open dataset of biosignal recordings, we intend to advance research in the field of digital phenotyping towards personalized mental health monitoring.

5. REFERENCES

[1] A. Zlatintsi et al., "E-Prevention: Advanced support system for monitoring and relapse prevention in patients with psychotic disorders analyzing long-term multimodal data from wearables and video captures," *Sensors*, vol. 22(19), 7544, Oct. 2022, DOI:https://doi.org/10.3390/s22197544.

[2] E. Kalisperakis et al., "Smartwatch digital phenotypes predict positive and negative symptom variation in a longitudinal monitoring study of patients with psychotic disorders," *Front. Psychiatry*, vol. 14, Mar. 2023, DOI:https://doi.org/10.3389/fpsyt.2023.1024965.

[3] P. McGorry et al., "Biomarkers and clinical staging in psychiatry," *World Psychiatry*, vol. 13, pp. 211–223, 2014.

[4] J. Torous et al., "New tools for new research in psychiatry: a scalable and customizable platform to empower data driven smartphone research," *JMIR Mental Health*, vol. 3(2), pp. e16, 2016.

[5] D. Robotham et al., "Do we still have a digital divide in mental health? A five-year survey follow-up," *J. Medical Internet Research*, vol. 18, pp. e309, 2016.

[6] I. Barnett, J. Torous, P. Staples, et al., "Relapse prediction in schizophrenia through digital phenotyping: a pilot study," *Neuropsychopharmacology*, vol. 43, pp. 1660–1666, 2018.

[7] K. Burghardt et al., "Having a bad day? detecting the impact of atypical events using wearable sensors," in *Proc. SBP-BRiMS*, 2021.

[8] A. Saeed et al., "Multi-task self-supervised learning for human activity detection," *ACM on IMWUT*, vol. 3, pp. 1–30, 2019.

[9] G. Retsinas et al., "Person identification using deep convolutional neural networks on short-term signals from wearable sensors," in *Proc. ICASSP*, 2020.

[10] V.A. Canady, "Psychiatrists, patients favor telemedicine; challenges exist," *Mental Health Weekly*, vol. 30(20), pp. 4–5, 2020.

[11] M. Panagiotou et al., "A comparative study of autoencoder architectures for mental health analysis using wearable sensors data," in *Proc. EUSIPCO*, Belgrade, Serbia, 2022.

[12] D. A. Adler et al., "Predicting early warning signs of psychotic relapse from passive sensing data: an approach using encoder-decoder neural networks," *JMIR mHealth and uHealth*, vol. 8(8), pp. e19962, 2020.