PAL: Privacy-preserving Audio, Visual, and Physiological Contexts for Wearable Context-aware Behavior Change Support

Mina Khan[a], Glenn Fernandes[a] and Pattie Maes[a]

[a]*MIT Media Lab, 75 Amherst St, Cambridge, MA, USA. 02139*

Abstract

Context tracking and context-aware interventions are key for behavior change. However, there is a lack of context-aware wearable systems for just-in-time interventions using privacy-preserving on-device deep learning. We created a wearable device with multimodal context sensing (using an egocentric camera, microphone, and optical pulse sensor), an on-device deep learning accelerator (for privacy-preserving context detection), and open-ear audio output (for seamless just-in-time interventions). We also added a mobile phone app for sensor control, intervention set-up, data visualization, and geolocation and physical activity sensing. Our system supports custom context tracking and custom context-based interventions. Our open-source system can be combined with different deep learning models for context tracking and context-based interventions in different behavioral, environmental, and physiological contexts using audio, visual, and heart-rate sensing. Our system combines multimodal context sensing, privacy-preserving on-device deep learning, and just-in-time interventions for a more convergent, holistic, and translational approach towards real-world context-aware behavior change support.

Keywords

microphone, heart-rate, egocentric camera, privacy-preserving, just-in-time, audio interventions, context-based, on-device deep learning,

1. Introduction

Context tracking and context-aware interventions are key for behavior change, and research has used emotional, behavioral, and environmental contexts for user tracking, user modeling, and context-aware interventions [1, 2]. Recent advances in deep learning allow us to recognize different user contexts, especially using computer vision and audio processing. However, sending audio or visual data to the cloud for deep learning raises privacy concerns.

We present a wearable system, called PAL, with multimodal context sensing (audio, egocentric visual, and heart-rate), on-device deep learning (for privacy-preserving context detection), and open-ear audio output (for seamless just-in-time interventions). Even though there are several multimodal context-sensing systems, none combine audio, visual, and physiological context sensing with privacy-preserving on-device deep learning and just-in-time interventions. PAL also includes a mobile app for sensor control, intervention setup, and geolocation and physical activity sensing.

We have made our wearable system open-source, modular, and extensible so that developers, researchers, and users may use it for

College Station '21: Joint Proceedings of the ACM IUI 2021 Workshops, April 13-17, 2021, College Station, USA
minakhan01@gmail.com (M. Khan); glennfer@mit.edu (G. Fernandes); pattie@media.mit.edu (P. Maes)

custom context tracking and behavior change interventions in the real world. Our modular and open-source system facilitates a convergent, holistic, and translational approach towards real-world behavior change to meet the multiple, diverse, and customized behavior change support needs of users and researchers [3].

2. Related Work

There are several self-tracking applications, e.g., for memory support and activity tracking [2] and mental health tracking [4]. Behavior change interventions are also common [5], including context-aware interventions [1]. Deep learning is also used, e.g., for activity and context recognition [6]. However, there are no deep learning systems for just-in-time behavior change interventions using privacy-preserving on-device deep learning, multimodal audio, visual, and physiological context sensing, and open-ear interventions. We created PAL, an open-source wearable system with multimodal context sensing, privacy-preserving on-device deep learning, and open-ear audio interventions, to enable convergent and customizable deep learning-based context-aware behavior change support.

3. Design

We aimed to leverage deep learning for personalized and privacy-preserving context-aware behavior change support. We chose multimodal context sensors, on-device deep learning, and open-ear audio interventions for context-aware support, and enable custom goal-setting, intervention contexts, and intervention messages using a customizable if-CONTEXT-then-INTERVENTION framework.

We decided on five types of contexts for sensing – egocentric camera view, audio input, heart rate, geolocation, and physical activity. Egocentric visual context and audio input can give key information about the user's behavioral and environmental contexts, especially using deep learning. We also integrated an optical pulse sensor as heart rate variability is a helpful sensor for detecting user's physiological states [4]. Finally, we added geolocation and physical activity sensing as they are commonly used by existing applications [4] and could be easily accessed via a mobile phone.

We added an on-device deep learning processor so that sensitive user data, especially audio-visual data, could be processed on device, not on cloud, in a privacy-preserving, offline, and real-time manner.

We added open-ear audio output to provide real-time, minimally-disruptive, and private interventions in real-world mobile contexts.

4. Implementation

We created a wearable system, called PAL (Figure 1). PAL's wearable device supports context-tracking and interventions, and PAL's mobile/ web app enables data visualizations, sensor control, and intervention-setting. The open-source implementation is available here: https://github.com/minakhan01/PAL_Wearable_Opensource/.

We used Google Coral for on-device deep learning. We placed a camera on the ear to capture similar contexts as the eyes, without being too prominent on the face. We added an optical heart rate sensor on the earlobe as earlobe Pulse Photoplethysmographic (PPG) Heart Rate Variability (HRV) is comparable to Electrocardiographic (ECG) HRV [7]. We used a mini-speaker for open-ear audio output, a microphone for audio input, and Google's Places and Activity Recognition APIs for geolocation and physical activity sensing.

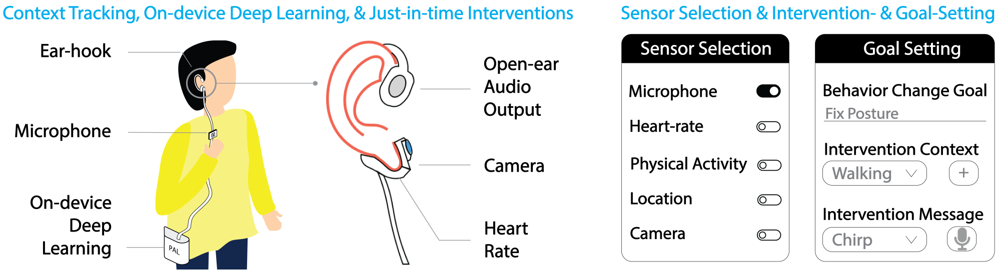

Figure 1: Wearable device with on-device deep learning, camera, open-ear audio output, microphone, and button.

	Correlation
Apple Watch & PAL	91%
Zephyr Biopatch ECG & PAL	75%
Empatica E4 & Zephyr	69%

Table 1
PAL's Heart Rate Variability (HRV) comparison

The on-ear camera captures ~70% of a visual context ~1m away. We performed initial evaluations of PAL's pulse rate sensor with 5 people (3 male, 2 female, μ=23.4 years, σ=2.56 years). We compared PAL's PPG-HRV to HRV from Series 4 Apple Watch ECG and Zephyr Biopatch ECG. Apple Watch does 30s recordings and we did 10 30s recordings per participant, with each participant simultaneously wearing PAL and Apple Watch. For Zephyr Biopatch, we had 30 minute recordings per participant, with each participant simultaneously wearing PAL, Empatica E4, and Zephyr. The results are in Table 1. PAL consumes ~0.3A and a 2500 mAh battery lasts ~5 hours.

5. Applications

We used PAL with low-shot, custom-trainable, and on-device deep learning models for visual context detection and our findings show that interventions in personalized visual contexts improve real-world habit-support interventions [8]. PAL can be further used for more context tracking, e.g., food tracking using food detection models or mood tracking using speech processing, and for context-based behavior change support, e.g., using just-in-time adaptive interventions tailored to the user's contexts [1] . More sensors can also be added to PAL's wearable system or paired via Bluetooth, e.g. for continuous glucose monitoring along with visual food tracking.

6. Conclusion

We combined previous research in context-aware behavior change support and recent advances in on-device deep learning to create a wearable system, called PAL, with multimodal context sensing, privacy-preserving on-device deep learning, and open-ear audio interventions. PAL supports customizable context tracking and context-aware interventions to enable diverse and personalized context-aware behavior change support. We used PAL to improve real-world habit-support interventions using personalized visual contexts [8], and have open-sourced PAL so that other users, researchers, and developers can use PAL's multimodal context sensing, privacy-preserving on-device deep learning, and open-ear audio interventions for diverse, convergent, and customizable context-aware behavior change support in wearable contexts.

References

[1] I. Nahum-Shani, S. N. Smith, B. J. Spring, L. M. Collins, K. Witkiewitz, A. Tewari, S. A. Murphy, Just-in-Time adaptive interventions (JITAIs) in mobile health: Key components and design principles for ongoing health behavior support, Ann. Behav. Med. 52 (2018) 446–462.

[2] C. Xia, P. Maes, The design of artifacts for augmenting intellect, in: Proceedings of the 4th Augmented Human International Conference, AH '13, Association for Computing Machinery, New York, NY, USA, 2013, pp. 154–161. URL: https://doi.org/10.1145/2459236.2459263. doi:10.1145/2459236.2459263.

[3] M. Khan, G. Fernandes, P. Maes, Users want diverse, multiple, and personalized behavior change support: Need-finding survey, in: International Conference on Persuasive Technology, Springer, 2021.

[4] E. Reinertsen, G. D. Clifford, A review of physiological and behavioral monitoring with digital sensors for neuropsychiatric illnesses, Physiol. Meas. 39 (2018) 05TR01.

[5] C. Pinder, J. Vermeulen, B. R. Cowan, R. Beale, Digital behaviour change interventions to break and form habits, ACM Trans. Comput. -Hum. Interact. 25 (2018) 15:1–15:66.

[6] V. Radu, C. Tong, S. Bhattacharya, N. D. Lane, C. Mascolo, M. K. Marina, F. Kawsar, Multimodal Deep Learning for Activity and Context Recognition, Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 1 (2018) 157:1–157:27. URL: https://doi.org/10.1145/3161174. doi:10.1145/3161174.

[7] B. Vescio, M. Salsone, A. Gambardella, A. Quattrone, Comparison between electrocardiographic and earlobe pulse photoplethysmographic detection for evaluating heart rate variability in healthy subjects in short- and Long-Term recordings, Sensors 18 (2018).

[8] M. Khan, G. Fernandes, A. Vaish, M. Manuja, A. Stibe, P. Maes, Improving context-aware habit-support interventions using egocentric visual contexts, in: International Conference on Persuasive Technology, Springer, 2021.