UbiComp Companion '26 · WellComp Workshop · Shanghai, China

CallSense: Sensing Call Engagement in Privacy-Constrained Call Center Environment with Commodity Smartwatches



Bonwoo Koo

Graduate School of Data Science



Junmo Lee

School of Computing



Uichin Lee

AI Computing Department

Korea Advanced Institute of Science and Technology (KAIST)

Motivation



Call Center Agents

Emotional labor in call centers

87%

rated workplace stress high or very high

Survey of 2,100 union-represented call center workers[1]

Repeated difficult interactions, emotional suppression, and limited autonomy create sustained strain.[2]

[1] Doellgast & O'Brady (2019), Making Call Center Jobs Better, executive summary. Cornell / CWA.
[2] Lee et al. (2026), Understanding Behind the Smile of Emotion Workers. CHI. doi:10.1145/3772318.3790803

Motivation

Stress Sensing



Estimate stress from
wearable physiological signals

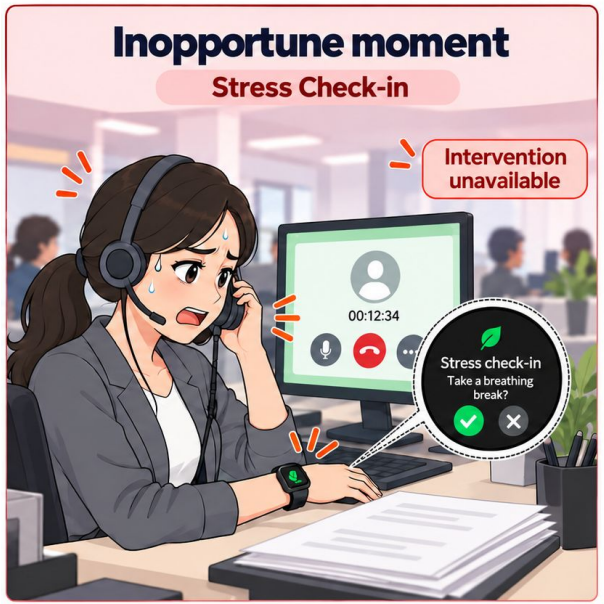
Just-in-Time Intervention



Offer brief well-being support,
such as a breathing exercise

Motivation





Just-in-time intervention must arrive at the right moment — not mid-call!

Motivation



Audio



Call logs

Privacy-constrained call center environment

We need a worker-owned wearable that finds the right moment on its own — no audio, no call logs

Existing Research

Work–break detection using passive physiological signals

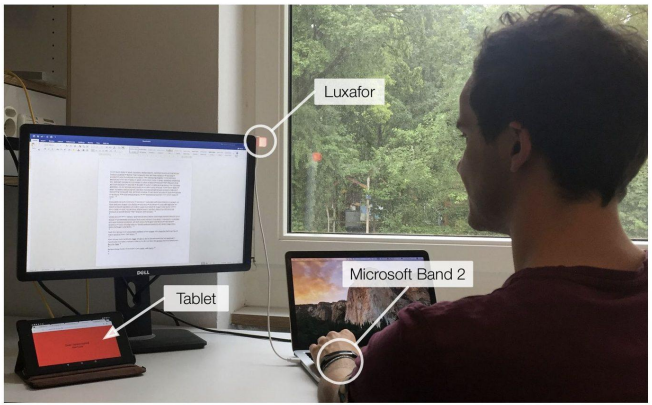


Fig. 6. The COLLINS system indicating that the office worker should not be interrupted.

Lab **Schaule et al., IMWUT 2018**

Consumer wearables → cognitive load → opportune moments for interruption in a controlled office

Table A.1. Summary of the 106 features extracted from the five sensory channels used in this paper. NN stands for the distance between BVP adjacent peaks. In addition to the features derived from the sensors we consider the hour of the day as in [62].

Sensory channel	Features extracted
EDA (mix-EDA + Phasic + Tonic) (68 features) From: [15, 20, 27, 57, 59, 65]	**Time-domain statistical features:** minimum, maximum, mean, median standard deviation, dynamic range, variance,slope, mean and standard deviation of the first and second derivative; **Wavelets features:** mean, standard deviation of wavelets coefficients at 1Hz, 2Hz and 4Hz; **SCR features:** number of peaks, peaks' amplitude, rise time, half-recovery time, width and area under the curve, maximum derivative of SCR;
BVP (13 features) From: [11, 15, 25, 44, 51]	**Time-domain statistical features:** mean, standard deviation, number of peaks, dynamic change; **Heart rate (HR);** **HRV statistical features:** mean of all NN, standard deviation of all NN (SDNN), standard deviation of differences between NN (SDSD), the square root of the mean of the sum of the squares of differences between NN (RMSSD); **BVP pulse features:** mean and standard deviation of pulses' length and amplitude;
ACC (7 features) From: [15, 55]	**Time-domain statistical features:** mean, standard deviation, dynamic range, slope, absolute value of the slope, mean and standard deviation of the first derivative;
Phone usage (2 features) From: [13, 53]	Number of unlock, number of screen on.
Laptop application usage (15 features)	**Duration of category of application:** Social Networking, Utilities, Uncategorized, Software Development, Shopping, News and Opinions, Entertainment, Learning and Reference, Design & Composition, Reference & Learning, Communication & Scheduling, Business, Non used, Most used category, Total duration of the laptop usage, Number of categories used.

In-the-wild **Di Lascio et al., IMWUT 2020**

Wrist EDA / BVP / ACC + phone & laptop usage → breaks vs. work, 5-min feature windows

Schaule, Florian, et al. "Employing consumer wearables to detect office workers' cognitive load for interruption management." Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 2.1 (2018): 1-20.
Di Lascio, Elena, et al. "A multi-sensor approach to automatically recognize breaks and work activities of knowledge workers in academia." Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 4.3 (2020): 1-20.

Research Goal · Temporal Granularity

Ordinary Workplace

Work	Break
~50 min	~10 min

Longer work–break cycles

Call Center

Work	Break	Work	Break
~2 min	~2 min	~2 min	~3 min

Frequent state changes within minutes

Call agents switch states every few minutes.

Research Goal · Behavioral Granularity

Ordinary Workplace

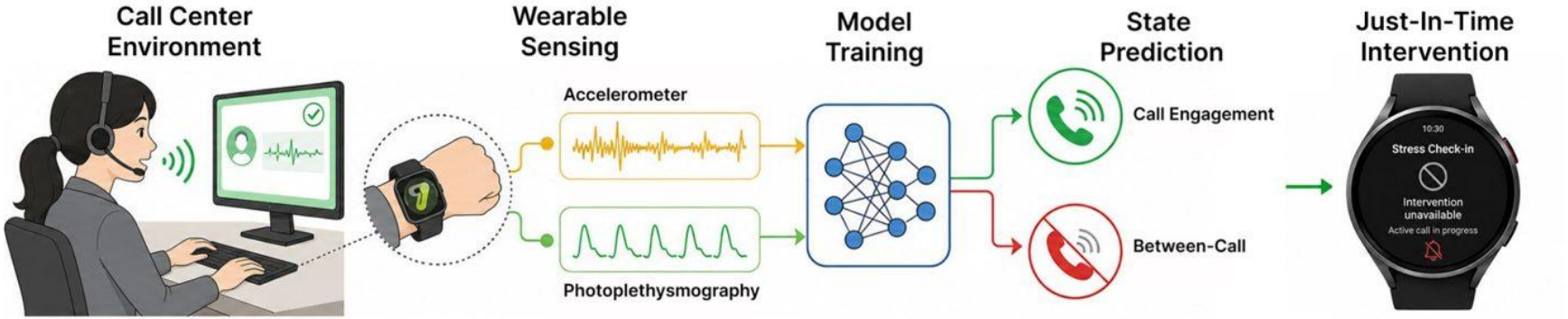
Work	Break
	
Typing	Rest · restroom · walk · stretch

Call Center

Work (call)	Break
	
At the desk	Still at the desk

CallSense: a smartwatch-based call engagement sensing system without audio or call logs for finer temporal and behavioral granularity

Overview



Methodology

- Galaxy Watch 7: ACC 25 Hz + PPG 25 Hz + HR 1 Hz
- Ground truth from call-log timestamps → CE / BC segments
- 60-s windows (50% overlap), 54 handcrafted features (28 ACC / 26 PPG)

Experiment

- Task: call engagement (CE) vs. between-call (BC) classification
- Models: XGB · CNN-BiLSTM · TabPFN (+ random baseline)
- Modalities: ACC · PPG · ACC + PPG
- User-independent (5-fold nested CV) & user-dependent (chronological 80/20)

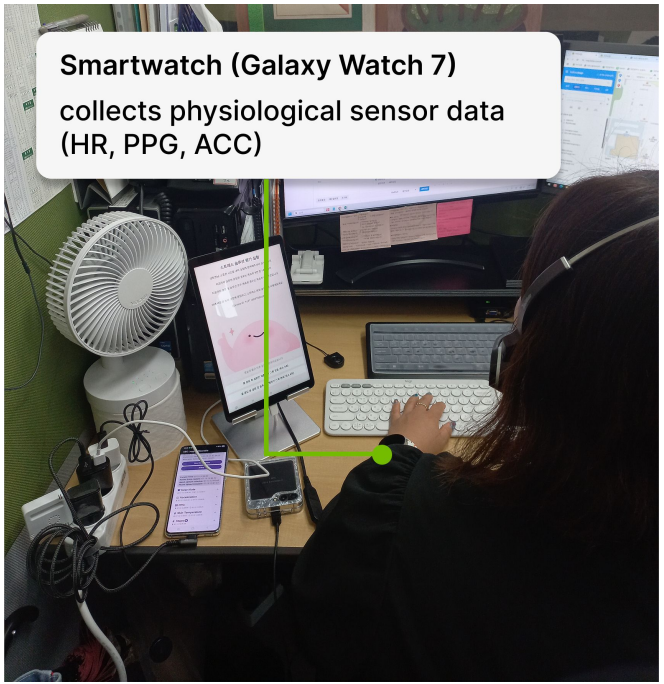
In-the-Wild Data Collection

Metropolitan call center, South Korea

19 **agents**
2 male / 17 female

38.9 **mean age**
SD 5.8 years

4 **weeks**
during regular shifts



In-the-wild data collection setup

Dataset

Preprocessing: 60-s windows · 50% overlap · ≥90% valid HR samples (Samsung SDK)

Call Engagement: 4,305 segments · ~13,000 windows

Between-call: 5,087 segments · ~52,000 windows



Call engagement/Between-call Segmentation and Labeling

Extracted Features

Feature Type	Extracted Features
ACC · 28	Mean, SD, range, slope, absolute slope, first-difference mean and SD (x, y, z, magnitude)
PPG · 2	Mean and SD of the PPG signal
Heart rate · 4	Median, minimum, maximum and SD of heart rate
HRV · 10	NN mean, SDNN, SDSD, RMSSD; NN count, median, minimum, maximum, SD and CoV
Pulse morphology · 10	Amplitude, inter-peak interval, peak / trough differences, pulse area (mean and SD of each)

Results

1 Every model beats chance

CallSense is feasible: smartwatch signals distinguish call engagement

2 TabPFN is the best model

AUROC 0.764 for unseen users and 0.854 for seen users

3 ACC carries the signal

PPG-only ≈ 0.6 AUROC;
ACC+PPG fusion ≈ ACC alone.

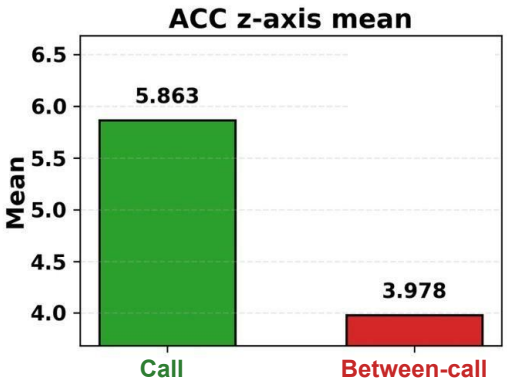
Table 2: Call Engagement Detection. Best AUROC per modality in bold. Threshold tuned on validation.

	ACC				PPG				ACC + PPG			
Model	Acc%	Sens%	Spec%	AUROC	Acc%	Sens%	Spec%	AUROC	Acc%	Sens%	Spec%	AUROC
User-independent validation												
Random Predictor	67.6	20.8	79.7	0.502	67.6	20.8	79.7	0.502	67.6	20.8	79.7	0.502
XGB	65.6	75.7	62.8	0.752	42.4	81.4	32.4	0.598	65.6	72.5	63.6	0.738
CNN-BiLSTM	57.7	79.8	51.9	0.697	38.6	82.2	27.5	0.566	62.6	72.0	60.2	0.712
TabPFN	66.9	73.0	65.1	**0.754**	43.8	81.5	34.1	**0.608**	68.7	70.7	68.0	**0.764**
User-dependent validation												
Random Predictor	67.8	19.9	78.3	0.491	67.8	19.9	78.3	0.491	67.8	19.9	78.3	0.491
XGB	77.1	67.7	79.0	0.828	47.8	78.5	40.4	0.652	77.1	66.8	79.1	0.812
CNN-BiLSTM	68.5	72.7	67.4	0.773	36.1	85.2	25.2	0.610	70.2	69.0	70.0	0.785
TabPFN	80.0	72.9	81.5	**0.854**	49.9	79.5	42.7	**0.673**	78.6	72.9	79.9	**0.848**

Discussion · Feature Importance

Wrist posture

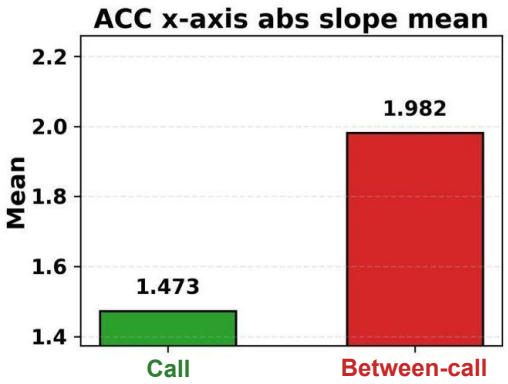
ACC z-axis mean



Higher during calls
Palm-down hands on the keyboard

Forearm movement

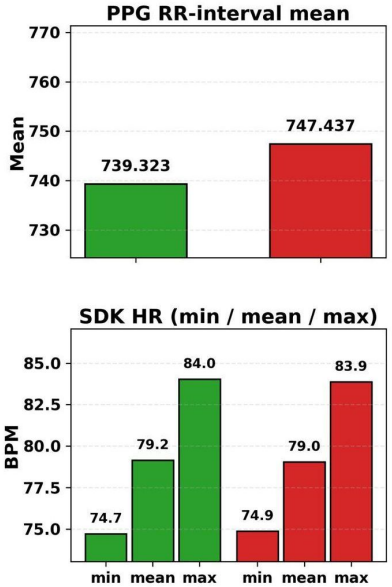
ACC x-axis abs. slope



Higher between calls
Freer, more dynamic movement

Cardiac signals

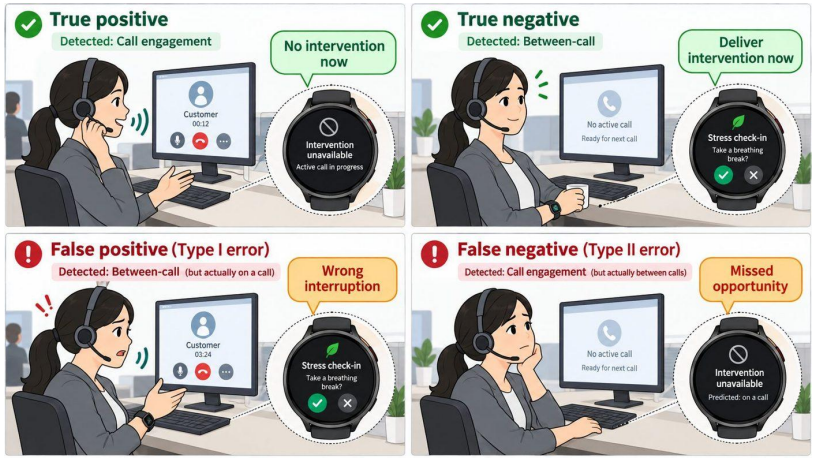
PPG interval and heart rate



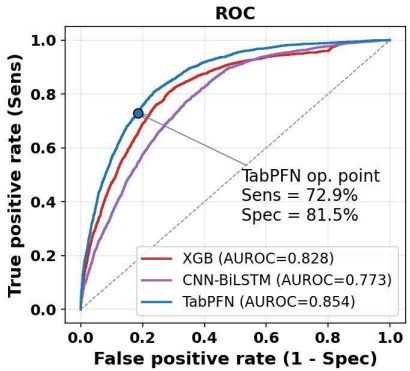
Small cardiac differences

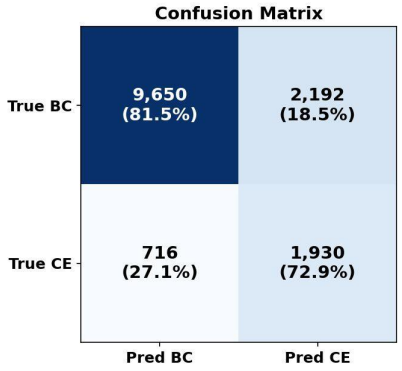
Discussion · Operating Point

Asymmetric cost of errors



Interrupting a call (FP) costs more than a missed window (FN)





TabPFN on ACC, user-dependent: AUROC 0.854 · Sens 72.9% · Spec 81.5%

- ≈ 27% of active-call windows would be interrupted; ≈ 19% of safe windows missed
- **A deployed system should shift toward sensitivity to cut mid-call prompts**
- Per-user labels via a brief enrollment (call logs or EMA), then label-free operation

Summary & Implication

CallSense is feasible



Smartwatch signals can distinguish call and between-call states using 60-second windows.

Posture and movement cues



Palm-down during calls
Freer movement between calls

Beyond call centers



Future Direction:
Interaction-bound service work