Few-Shot Personalization of Wearable Health Monitoring Systems

Anonymous Author
Anonymous Institution
anonymous@email.com

September 8, 2025

Abstract

[TODO: Complete after main results] Individual variability in physiological responses presents a significant challenge for deploying wearable health monitoring systems at scale. While companies collect extensive baseline data during initial deployments, the effectiveness of personalizing models using limited target participant data remains unclear. We present a systematic evaluation framework using Leave-One-Participant-Out (LOPO) cross-validation to assess few-shot personalization in smoking detection from accelerometer data. Our two-phase training methodology first learns general patterns from source participants, then adapts to target individuals using minimal data. [RESULTS PLACEHOLDER: Statistical significance, effect sizes, data efficiency findings]. These findings provide practical guidelines for companies implementing personalized health monitoring, demonstrating when and how personalization improves detection accuracy in real-world deployment scenarios.

Keywords: personalization, health monitoring, few-shot learning, wearables, smoking detection

1 Introduction

Wearable health monitoring systems have gained widespread adoption in both consumer and clinical settings, with applications ranging from fitness tracking to chronic disease management [4]. However, a fundamental challenge persists: individual physiological and behavioral differences lead to significant variability in sensor data patterns across users [2]. This variability directly impacts the accuracy of automated detection systems when models trained on population data are applied to new individuals.

1.1 Problem Motivation: Company Deployment Scenario

Consider a technology company deploying smoking cessation monitoring across a diverse user base. The company has extensive baseline data from initial participants but faces a critical decision: should they use a single population-level model, or invest in personalizing models for each new user? This decision involves trade-offs between deployment complexity, data collection requirements, and potential accuracy improvements.

1.2 Current Limitations

Existing personalization approaches in health monitoring suffer from several limitations:

- **Evaluation gaps:** Most studies use within-participant data splits rather than realistic cross-participant evaluation
- **Data requirements:** Unclear minimum target data needed for meaningful improvement
- **Individual variability:** Limited understanding of which participants benefit most from personalization

- **Deployment guidance:** Lack of practical guidelines for real-world implementation

1.3 Our Contributions

This work addresses these limitations through:

1. **Systematic LOPO Evaluation:** First comprehensive assessment of personalization effectiveness using realistic cross-participant evaluation
2. **Data Efficiency Analysis:** Quantification of minimum target data requirements for meaningful improvement
3. **Individual Characterization:** Identification of participant factors that predict personalization success
4. **Deployment Guidelines:** Practical recommendations for companies implementing personalized health monitoring

Our key hypotheses include: (H1) Customization will improve F1 scores by ¿10% on average, (H2) meaningful improvements are possible with ¡1 hour of target data, and (H3) base model performance predicts customization success.

2 Related Work

2.1 Personalization in Health Monitoring

Recent work has explored various approaches to personalizing health monitoring systems [3]. However, most studies focus on within-participant evaluation rather than realistic cross-participant scenarios.

2.2 Few-Shot Learning and Domain Adaptation

The challenge of adapting models to new participants with limited data shares similarities with few-shot learning and domain adaptation problems in computer vision and natural language processing [1].

3 Methods

3.1 Dataset and Participants

Our evaluation uses accelerometer data from [X] participants collected during smoking detection studies. Data consists of 50Hz tri-axial accelerometer measurements from wrist-worn devices, with expert-annotated smoking bouts providing ground truth labels.

[TODO: Add participant demographics and data statistics]

3.2 Two-Phase Training Methodology

Our approach consists of two distinct training phases designed to mirror real-world company deployment scenarios:

Algorithm 1 Two-Phase Personalization Training

Phase 1: Base Model Training
Train CNN model on source participants (all except target)
Optimize using early stopping on validation F1 score
Save base model parameters θ_{base}

Phase 2: Target Personalization
Initialize model with θ_{base}
Fine-tune on target participant data
Apply reduced learning rate and early stopping
Output personalized model $\theta_{personalized}$

3.3 Leave-One-Participant-Out (LOPO) Evaluation

We employ LOPO cross-validation to ensure realistic evaluation conditions that mirror company deployment scenarios. For each target participant:

1. Train base model on remaining [X-1] participants
2. Personalize model using target participant's data
3. Evaluate on held-out target participant test set

4. Compare personalized vs. base model performance

3.4 Model Architecture

Our smoking detection model uses a 1D convolutional neural network optimized for 60-second accelerometer windows (3000 samples at 50Hz). The architecture employs dilated convolutions to capture long-range temporal dependencies relevant to smoking gesture patterns.

3.5 Statistical Analysis

All comparisons use paired statistical tests accounting for participant-level dependencies. We report effect sizes (Cohen's d) with confidence intervals and apply multiple comparison corrections where appropriate.

4 Results

4.1 Main Findings: Personalization Effectiveness

[TODO: Add Figure 1 - Main Results Bar Chart]

[PLACEHOLDER: Key statistical findings] - Base model average F1: [X.XX] ± [X.XX] - Personalized model average F1: [X.XX] ± [X.XX] - Average improvement: [X.XX] ± [X.XX] (p ¡ [X.XXX]) - Effect size (Cohen's d): [X.XX] [95% CI: X.XX, X.XX]

4.2 Data Efficiency Analysis

[TODO: Add Figure 2 - Data Efficiency Curves]

[PLACEHOLDER: Data efficiency findings] Analysis of personalization effectiveness across different amounts of target data: - 25% target data: [X.XX] average improvement - 50% target data: [X.XX] average improvement - 75% target data: [X.XX] average improvement - 100% target data: [X.XX] average improvement

4.3 Individual Variability Patterns

[TODO: Add Figure 3 - Individual Analysis]

[PLACEHOLDER: Individual characterization results] Factors predicting personalization success: - Base model performance correlation: r = [X.XX] (p ¡ [X.XXX]) - Data quality metrics correlation: r = [X.XX] (p ¡ [X.XXX]) - Behavioral pattern consistency: [findings]

4.4 Failure Case Analysis

[PLACEHOLDER: Analysis of negative personalization cases] [X] out of [Y] participants showed decreased performance after personalization. Common patterns include: - Overfitting indicators: [analysis] - Data insufficiency markers: [analysis] - Base model ceiling effects: [analysis]

5 Discussion

5.1 Practical Deployment Guidelines

Based on our systematic evaluation, we provide the following recommendations for companies implementing personalized health monitoring:

- **When to personalize:** [Evidence-based criteria]
- **Data collection requirements:** [Minimum target data guidelines]
- **Success prediction:** [Participant screening approaches]
- **Implementation strategy:** [Phased deployment recommendations]

5.2 Theoretical Insights

[TODO: Develop theoretical framework] Our findings contribute to understanding of personalization in health monitoring by: - Quantifying individual vs. population-level model trade-offs - Establishing data efficiency baselines for few-shot health personalization - Identifying predictive factors for adaptation success

5.3 Limitations and Future Work

[TODO: Add limitations based on experimental findings] Key limitations include: - Single-task evaluation (smoking detection only) - Limited participant diversity [expand based on demographics] - [Other limitations discovered during experiments]

Future work should explore: - Multi-task personalization across different health behaviors - Advanced personalization methods beyond fine-tuning - Longitudinal adaptation as user patterns evolve

6 Conclusion

[PLACEHOLDER: Key takeaways and impact]

This work provides the first systematic evaluation of few-shot personalization in wearable health monitoring using realistic cross-participant validation. Our findings demonstrate [key results] and provide practical guidelines for companies deploying personalized health monitoring systems. [Impact statement and broader implications].

7 Acknowledgments

References

[1] TODO: Author. Few-shot learning: A survey. *TODO: Journal*, 2023. TODO: Add actual citation.

[2] TODO: Author. Individual variability in physiological sensor data: Challenges and opportunities. *TODO: Journal*, 2023. TODO: Add actual citation.

[3] TODO: Author. Personalization approaches in digital health monitoring. *TODO: Journal*, 2024. TODO: Add actual citation.

[4] TODO: Author. Wearable health monitoring systems: A comprehensive review. *TODO: Journal*, 2024. TODO: Add actual citation.

A Supplementary Methods

B Supplementary Results