

Domain Adaptation for Cross-Device Robustness in Wearable Photoplethysmographic Health Monitoring

Mukesh G. Mathur

Department of Computer Science, Colorado State University, Fort Collins, CO, USA.
mukesh.work@colostate.edu

Abstract

Wearable photoplethysmographic sensors have become central to continuous health monitoring, yet their clinical and consumer utility remains constrained by substantial variability across device types, skin tones, sensor placements, and motion conditions. Domain adaptation offers a systems-level strategy for learning physiological representations that remain stable across heterogeneous data sources. This paper examines cross-device robustness from an interdisciplinary perspective, integrating signal processing, machine learning, deployment architecture, and governance. It argues that cross-device generalization is not only an algorithmic problem but also a property of the wider socio-technical infrastructure in which wearable monitors are embedded. The analysis reviews characteristic sources of photoplethysmographic distributional shift, including optical design differences, tissue interface variability, and environmental motion artifacts. It then connects these sources to formal domain adaptation perspectives and discusses adversarial, discrepancy-based, and self-supervised methods as components of a resilient monitoring architecture. The paper further considers how data curation, fairness, regulatory oversight, and lifecycle maintenance shape the robustness of deployed systems. Robustness mechanisms must be evaluated not only under controlled benchmarks but also through continuous postmarket evidence, equity audits, and transparent uncertainty reporting. The discussion concludes that durable cross-device robustness emerges from the alignment of technical adaptation capacity with organizational governance, device certification frameworks, and participatory design practices. Future systems should treat domain adaptation as a continuous infrastructure function rather than a one-time model correction.

Keywords

domain adaptation, wearable photoplethysmography, cross-device generalization, physiological monitoring, health equity, regulatory governance.

1. Introduction

Wearable photoplethysmography has become a key modality for unobtrusive cardiovascular and respiratory monitoring. The optical measurement of blood volume pulsations supports heart rate tracking, oxygen saturation estimation, and increasingly complex inference tasks associated with arrhythmia and hemodynamic state. Early work established the physiological basis and clinical utility of photoplethysmography across diverse measurement sites [1]. Subsequent analyses expanded understanding of signal morphology and analysis choices [2]. Despite these advances, the translation of photoplethysmographic algorithms from controlled laboratory settings to heterogeneous wearable ecosystems remains difficult. A central obstacle is domain shift: a model trained on one sensor configuration, population, or acquisition protocol often degrades when deployed under new conditions. This paper addresses that obstacle through the lens of domain adaptation, treating cross-device robustness as a systems

property shaped by architecture, infrastructure, and governance as much as by statistical learning.

The proliferation of wrist-worn devices, earbuds, rings, and chest patches has generated an installed base with widely varying optical components, sampling rates, filter characteristics, and firmware preprocessing. These differences produce data distributions that cannot be fully harmonized by simple calibration. Transfer learning provides a general framework for reusing knowledge across related tasks or domains [3]. Theoretical treatments of learning from different domains clarify why empirical risk minimization under source distributions does not guarantee target performance unless explicit assumptions about domain divergence are managed [4]. The present paper extends these insights into the specific context of wearable photoplethysmography and argues that robust health monitoring requires coordinated design across sensing hardware, representation learning, deployment feedback loops, and regulatory oversight.

2. Wearable PPG Signal Acquisition and Cross-Device Variability

Photoplethysmographic acquisition is governed by optical physics at the tissue interface. Light-emitting diodes inject radiation into capillary beds, and photodetectors measure reflected or transmitted intensity. The resulting waveform encodes volumetric changes influenced by cardiac pulsation, respiration, vasomotor activity, and local tissue properties [1]. However, the measurement is highly sensitive to contact pressure, skin pigmentation, ambient light, sensor geometry, and emitter wavelength. Algorithms designed for fingertip sensors do not automatically transfer to wrist-worn devices because the signal-to-noise ratio, morphology, and motion artifact profile differ substantially [2]. These differences are structural rather than incidental because they arise from the geometry of the optical path and the dynamics of the peripheral vascular bed.

Cross-device variability becomes especially pronounced when commercial platforms update hardware or when data are aggregated across device generations. A source model may learn device-specific preprocessing artifacts or sensor noise signatures that are absent in target devices. This phenomenon parallels the broader challenge of covariate shift in computational perception, where superficial characteristics of the input space distort model generalization [5]. In parallel, adversarial learning methods have shown that representations can be trained to confuse a domain discriminator, thereby reducing the influence of acquisition-specific information [6][7]. Other approaches align second-order statistics or use deep adaptation layers to transfer features across domains [8]. Within photoplethysmographic monitoring, these methods must operate under tight power, latency, and memory constraints, a tension that shapes architectural choices more than in unconstrained vision tasks.

Remote and contactless photoplethysmography has contributed additional insight into algorithmic principles for robust optical sensing [9]. Although remote measurement differs from wearable contact sensors, many of the same issues of illumination, skin reflectance, and movement apply. At the same time, quality assessment methods are essential for unsupervised telehealth deployments because corrupted photoplethysmographic segments can silently degrade downstream decision-making [10]. Large-scale challenges in physiological signal classification have demonstrated that algorithmic performance can be highly sensitive to data source and recording conditions [11]. A review of breathing rate estimation from electrocardiogram and photoplethysmogram data further documents variability across sensor types and patient groups [12]. Together, these findings indicate that cross-device robustness

cannot be achieved by single-site data collection alone. It demands explicit mechanisms for domain alignment and continuous quality monitoring.

3. Domain Adaptation as a Structural Response to Distributional Shift

Domain adaptation provides formal and practical tools for aligning source and target distributions without requiring extensive labeled data in the target domain. In wearable health monitoring, labels are expensive and often require clinical adjudication. Transfer learning research has shown that reusable representations can reduce target sample complexity when source and target tasks share underlying structure [3]. The theoretical analysis of domain divergence emphasizes that successful adaptation depends on the distance between source and target distributions and the capacity of the hypothesis class [4]. For wearable photoplethysmography, domain divergence may arise from optical hardware, sensor placement, user behavior, and demographic physiology. The system designer must decide whether to address such divergence through invariant representation learning, data augmentation, or model personalization.

Adversarial domain adaptation is one prominent family of methods. It trains a feature extractor jointly with a domain classifier that attempts to distinguish source from target examples. The feature extractor is optimized to prevent the classifier from succeeding, thereby producing domain-invariant features [6]. A related approach separates the stages of feature encoding and adversarial alignment, with empirical success in visual and physiological applications [7]. Discrepancy-based methods, by contrast, minimize explicit distance measures between source and target feature distributions, often using kernel embeddings or normalized statistics [8]. A foundational early method showed that simple batch normalization and linear adaptation can achieve surprising gains across domain shifts [5]. These approaches form a design space in which monitoring systems can trade off complexity, stability, and interpretability.

For wearable photoplethysmography, pre-training on large-scale physiological datasets is becoming a viable route to robust representations. Transfer learning in medical imaging has shown that pre-training on broad datasets can outperform task-specific training when target labels are scarce, but only when domain differences are carefully managed [13]. A study on adaptive notch filtration during physical activity illustrates a complementary signal-level intervention that recovers photoplethysmographic signals corrupted by motion, thereby reducing the domain gap before learning takes place [14]. At the representation level, contrastive objectives originally developed for visual data have demonstrated that self-supervised encoders can generalize across downstream tasks, and they can be adapted to physiological signals through augmentations that mimic motion artifacts, sensor noise, and skin tone variation [15]. These approaches form a layered strategy in which signal conditioning and representation learning jointly address distributional shift.

4. System Architecture and Deployment Considerations

Deploying domain-adapted photoplethysmographic models in real-world health monitoring requires careful architectural design. Modern wearable systems frequently separate on-device inference from cloud-based model updates. On-device components must operate within strict energy and memory budgets, while cloud components can support computationally intensive retraining and population-level monitoring. The success of deep learning in health applications has expanded expectations for wearable analytics but also increased the risk that models will rely on brittle correlations [16]. A resilient architecture therefore embeds domain

adaptation not as a standalone module but as part of a continuous learning pipeline that can respond to new devices, firmware updates, and changing user populations.

Continuous adaptation raises operational questions about model versioning, data provenance, and fallback behavior. When a target domain differs substantially from all available training sources, the system should detect the shift and reduce reliance on high-confidence predictions. Quality assessment modules can gate the flow of data into adaptation algorithms, ensuring that corrupted segments do not contaminate updated representations [10]. The architecture should also distinguish between short-term personalization and long-term population-level adaptation. Personalization can adjust a model to an individual user but may overfit to transient physiological states. Population-level adaptation can improve equity across demographic groups but requires centralized aggregation and careful privacy controls. Both layers must coexist in a governance framework that sets acceptable risk thresholds and mandates human review for safety-critical outputs.

Cloud-edge hybrid designs provide one route through these trade-offs. Lightweight feature extractors on the wearable can compress raw photoplethysmographic signals into privacy-preserving embeddings before transmission. The cloud can then maintain a domain adaptation engine that aligns embeddings from many device types and demographics. This structure reduces raw data egress and supports federated updates. However, it also introduces dependency on connectivity and raises concerns about vendor lock-in. An alternative is to embed adaptation entirely on-device using small adaptation heads updated during charging cycles. Such an approach preserves autonomy but limits the diversity of data available for adaptation. The choice between cloud-centric, edge-centric, and hybrid architectures is therefore not purely technical; it shapes data ownership, regulatory obligations, and the sustainability of the monitoring ecosystem.

5. Data Governance, Fairness, and Regulatory Implications

Cross-device robustness is closely connected to fairness. If a model performs well on light skin tones but degrades on darker skin tones because of melanin absorption differences, domain adaptation may reduce but not eliminate the disparity. Responsible machine learning frameworks emphasize that health models must be evaluated across subgroups and deployed with explicit monitoring for harm [17]. Explainability research similarly highlights that clinicians and users need to understand when adaptation has changed model behavior in ways that affect care decisions [18]. For wearable photoplethysmography, fairness requires representative data collection across skin tones, age groups, body mass distributions, and common comorbidities. Without such representation, domain adaptation can learn to ignore protected attributes while still failing to serve minority populations.

The regulatory environment for adaptive physiological monitors is evolving. Traditional medical device regulation assumes a relatively static model with premarket validation, but continuous domain adaptation introduces postmarket model change. Regulatory scholars have argued that adaptive algorithms challenge the distinction between device and software update [19]. In the United States, the Food and Drug Administration has begun to consider frameworks for predetermined change control plans and continual learning devices [20]. A National Academy of Medicine report on artificial intelligence in health care underscores the need for lifecycle governance that integrates clinical, technical, and institutional perspectives [21]. These developments matter for wearable photoplethysmography because the same sensor may serve wellness and medical use cases, each with different evidentiary standards.

Privacy is another governance dimension. Domain adaptation often requires aggregating physiological data across users and devices, which raises concerns about reidentification and secondary use. Legal scholarship on medical big data warns that anonymization is difficult for high-dimensional biosignals and that consent frameworks may not anticipate federated learning or embedding transfer [22]. Organizations must therefore design data infrastructure with technical controls such as differential privacy, secure aggregation, and audit logging, but also with institutional mechanisms for user notification and withdrawal. The governance challenge is to maintain the data diversity required for robust adaptation while preserving individual autonomy. This tension cannot be resolved by privacy-enhancing technology alone; it requires clear policies about data use, retention, and benefit sharing.

6. Evaluation, Sustainability, and Lifecycle Robustness

Evaluation of domain adaptation in wearable photoplethysmography must go beyond accuracy on a single held-out test set. The health machine learning literature has repeatedly observed that benchmark performance can overstate real-world utility when data distributions are narrow and evaluation protocols fail to capture deployment shifts [23]. For cross-device robustness, evaluation should include multiple source and target devices, realistic motion protocols, and demographic stratification. It should also report calibration, uncertainty, and failure modes under low-perfusion conditions. Metrics such as heart rate error or arrhythmia detection sensitivity must be conditioned on signal quality and clinical context. A model that performs well on clean segments may fail precisely when monitoring is most needed, namely during physical activity, sleep disturbance, or acute physiological stress.

Fairness audits should be integrated into lifecycle evaluation. The field of algorithmic fairness has produced taxonomies of bias that apply to health data, including representation bias, measurement bias, and evaluation bias [24]. Cross-device photoplethysmographic systems are vulnerable to measurement bias because optical sensing interacts with skin pigmentation and tissue composition. Evaluation protocols must measure performance across protected and clinically relevant subgroups using intersectional categories rather than single-axis demographic variables. In addition, the sustainability of monitoring systems depends on their ability to maintain performance as device fleets age. Optical components drift, batteries degrade, and firmware updates alter preprocessing. Domain adaptation can help manage this lifecycle drift, but only if the system collects longitudinal evidence of performance and triggers recalibration when drift exceeds predefined bounds.

Sustainability also includes organizational and environmental dimensions. Training large adaptation models has energy and carbon costs, and frequent model updates can strain data centers and device battery life. A sustainable cross-device monitoring infrastructure should favor smaller task-adapted models, reuse pre-trained representations, and schedule retraining based on observed drift rather than fixed intervals. The same principle applies to data governance: adaptation should not justify unlimited collection of sensitive physiological data. Instead, systems should collect the minimum data needed to maintain safety and equity, and should periodically delete or aggregate older data. These practices align with broader calls for responsible machine learning in health care [17]. They also reinforce the central claim of this paper: cross-device robustness is an emergent property of technical, organizational, and regulatory design choices rather than a standalone algorithmic achievement.

7. Conclusion

This paper has examined domain adaptation for cross-device robustness in wearable photoplethysmographic health monitoring as a systems problem. It traced the sources of photoplethysmographic distributional shift, including optical hardware variation, tissue interface differences, and motion artifacts. It then positioned domain adaptation methods within a broader architecture that includes signal conditioning, representation learning, deployment feedback, and quality control. The discussion emphasized that adversarial, discrepancy-based, and self-supervised methods offer meaningful but partial responses. Their effectiveness depends on how they are embedded in wearable devices, cloud services, and regulatory processes.

The governance analysis highlighted that fairness and privacy are not external constraints but first-order design requirements. Cross-device performance gaps along demographic lines can undermine trust and clinical utility even when aggregate accuracy appears acceptable. Adaptive monitoring systems must include equity audits, explainable model behavior, and clear pathways for regulatory oversight. Continuous adaptation suggests that models are no longer static products; they are evolving services that require lifecycle governance, postmarket surveillance, and user-centered consent mechanisms. These responsibilities are amplified by the physiological nature of photoplethysmographic data and the increasing use of wearables for clinical decision support.

Future work should develop open cross-device photoplethysmographic benchmarks that include diverse populations, realistic motion, and longitudinal drift. It should also investigate lightweight adaptation techniques suitable for on-device execution, methods for detecting domain shift in real time, and governance models for federated adaptation across institutions. The long-term goal is not a single universal model but a resilient socio-technical infrastructure in which monitoring systems can adapt safely, fairly, and sustainably to the heterogeneity of human physiology and device ecosystems. The required integration of adaptive signal processing [14] into this infrastructure illustrates that progress will emerge from the convergence of sensing science, machine learning, and health system governance.

References

1. Allen, J. (2007). Photoplethysmography and its application in clinical physiological measurement. *Physiological Measurement*, 28(3), R1–R39. <https://doi.org/10.1088/0967-3334/28/3/R01>
2. Elgendi, M. (2012). On the analysis of fingertip photoplethysmogram signals. *Current Cardiology Reviews*, 8(1), 14–25.
3. Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359.
4. Ben-David, S., Blitzer, J., Crammer, K., Kulesza, A., Pereira, F., & Vaughan, J. W. (2010). A theory of learning from different domains. *Machine Learning*, 79(1–2), 151–175.
5. Sun, B., Feng, J., & Saenko, K. (2016). Return of frustratingly easy domain adaptation. *Proceedings of the AAAI Conference on Artificial Intelligence*, 30(1), 2058–2065.
6. Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., Marchand, M., & Lempitsky, V. (2016). Domain-adversarial training of neural networks. *Journal of Machine Learning Research*, 17(59), 1–35.

7. Tzeng, E., Hoffman, J., Saenko, K., & Darrell, T. (2017). Adversarial discriminative domain adaptation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 7167–7176.
8. Long, M., Cao, Y., Wang, J., & Jordan, M. I. (2015). Learning transferable features with deep adaptation networks. *Proceedings of the 32nd International Conference on Machine Learning*, 97–105.
9. Wang, W., den Brinker, A. C., Stuijk, S., & de Haan, G. (2017). Algorithmic principles of remote PPG. *IEEE Transactions on Biomedical Engineering*, 64(7), 1479–1491.
10. Lovell, N. H., Redmond, S. J., Basilakis, J., & Celler, B. G. (2010). Biosignal quality detection: An essential feature for unsupervised telehealth applications. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2010, 2131–2134.
11. Clifford, G. D., Liu, C., Moody, B., Lehman, L. H., Silva, I., Li, Q., Johnson, A. E. W., & Mark, R. G. (2017). AF classification from a short single lead ECG recording: The PhysioNet/Computing in Cardiology Challenge 2017. *Computing in Cardiology*, 44, 1–4.
12. Charlton, P. H., Birrenkott, D. A., Bonnici, T., Pimentel, M. A. F., Johnson, A. E. W., Alastruey, J., Tarassenko, L., Watkinson, P. J., Beale, R., & Clifton, D. A. (2018). Breathing rate estimation from the electrocardiogram and photoplethysmogram: A review. *IEEE Reviews in Biomedical Engineering*, 11, 2–20. <https://doi.org/10.1109/RBME.2017.2763681>
13. Raghu, M., Zhang, C., Kleinberg, J., & Bengio, S. (2019). Transfusion: Understanding transfer learning for medical imaging. *Advances in Neural Information Processing Systems*, 32, 3347–3357.
14. Zheng, X., Dwyer, V. M., Barrett, L. A., Derakhshani, M., & Hu, S. (2022). Adaptive notch-filtration to effectively recover photoplethysmographic signals during physical activity. *Biomedical Signal Processing and Control*, 72, 103303.
15. Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020). A simple framework for contrastive learning of visual representations. *Proceedings of the 37th International Conference on Machine Learning*, 119, 1597–1607.
16. Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., Cui, C., Corrado, G., Thrun, S., & Dean, J. (2019). A guide to deep learning in healthcare. *Nature Medicine*, 25(1), 24–29. <https://doi.org/10.1038/s41591-018-0316-z>
17. Wiens, J., Saria, S., Sendak, M., Ghassemi, M., Liu, V. X., Doshi-Velez, F., Jung, K., Heller, K., Kale, D., Saeed, M., Ossorio, P. N., Thadaney-Israni, S., & Goldenberg, A. (2019). Do no harm: A roadmap for responsible machine learning for health care. *Nature Medicine*, 25(9), 1337–1340. <https://doi.org/10.1038/s41591-019-0548-6>
18. Tonekaboni, S., Joshi, S., McCradden, M. D., & Goldenberg, A. (2019). What clinicians want: Contextualizing explainable machine learning for clinical end use. *Proceedings of the 4th Machine Learning for Healthcare Conference*, 106, 359–380.
19. Gerke, S., Babic, B., Evgeniou, T., & Cohen, I. G. (2020). The need for a system view to regulate artificial intelligence/machine learning-based software as medical device. *npj Digital Medicine*, 3, 53. <https://doi.org/10.1038/s41746-020-0262-2>

20. U.S. Food and Drug Administration. (2021). Proposed regulatory framework for modifications to artificial intelligence/machine learning (AI/ML)-based software as a medical device (SaMD): Discussion paper. U.S. Department of Health and Human Services.
21. Matheny, M. E., Whicher, D., & Israni, S. T. (2019). Artificial intelligence in health care: The hope, the hype, the promise, the peril. National Academy of Medicine.
22. Cohen, I. G., & Mello, M. M. (2018). Big data, big tech, and protecting patient privacy. JAMA, 319(13), 1331–1332. <https://doi.org/10.1001/jama.2018.2026>
23. Zech, J. R., Badgeley, M. A., Liu, M., Costa, A. B., Titano, J. J., & Oermann, E. K. (2018). Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study. PLoS Medicine, 15(11), e1002683. <https://doi.org/10.1371/journal.pmed.1002683>
24. Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. ACM Computing Surveys, 54(6), Article 115. <https://doi.org/10.1145/3457607>