Cross-Modal Deep Hashing Framework for Wearable PPG Signal Retrieval with Self-Supervised Semantic Representation Learning

Henri Gaylor
Department of Computer Science, University of New Hampshire, Durham, NH, USA.
henri.gaylor@unh.edu

Gavide Tmith
Department of Computer Science, University of Houston, Houston, TX, USA.
gtmith@uh.edu

Hirant Weshington
Department of Electrical Engineering and Computer Science, University of Missouri, Columbia, MO, USA.
hirant.washington@missouri.edu

Bereld Wegner
Department of Computer Science, University of Central Florida, Orlando, FL, USA.
beraldmail@ucf.edu

Abstract

The proliferation of wearable photoplethysmography sensors has generated vast streams of cardiovascular data, creating an urgent need for efficient, semantic-aware retrieval mechanisms that can operate across heterogeneous contextual modalities. This paper presents a cross-modal deep hashing framework designed for PPG signal retrieval that integrates self-supervised semantic representation learning to extract robust, modality-invariant features from unlabeled physiological time series and associated metadata. The framework maps PPG segments and their corresponding semantic descriptors into a shared binary Hamming space, enabling fast approximate nearest neighbor search while preserving clinically meaningful similarities. A comprehensive system-level analysis is conducted, addressing architectural choices that balance quantization error against retrieval precision, the integration of contrastive and masked reconstruction objectives for representation learning, and the trade-offs inherent in deploying such models on resource-constrained wearable edge devices. The discussion extends to governance and policy considerations, including data privacy, fairness across demographic groups, and the sustainability of large-scale health retrieval infrastructures. By emphasizing structural robustness, bias mitigation, and cross-modal alignment, the proposed framework offers a principled pathway toward scalable, privacy-preserving, and equitable health monitoring systems. The paper concludes with an examination of deployment scenarios, evaluation benchmarks, and future directions for cross-modal biosignal retrieval in real-world healthcare ecosystems.

Keywords

deep hashing, photoplethysmography, self-supervised learning, cross-modal retrieval, wearable computing, semantic representation.

1. Introduction

The rapid expansion of wearable health technologies has placed photoplethysmography at the center of ambulatory cardiovascular monitoring. PPG sensors, now embedded in consumer wristbands, medical patches, and smart rings, generate continuous high-dimensional waveforms that encode rich physiological information including heart rate variability, blood oxygen saturation, and vascular dynamics. However, the sheer volume and velocity of PPG data produced across millions of users worldwide have outpaced the development of efficient storage and retrieval systems capable of identifying clinically relevant patterns at scale. Traditional signal indexing methods, which rely on handcrafted features or full-length time series comparisons, fail to meet the latency, storage, and semantic requirements of modern health applications that demand similarity-based retrieval across multiple contextual modalities [1, 2]. The need for a retrieval framework that can bridge the gap between raw PPG signals and semantic labels such as activity states, emotional conditions, or diagnostic categories has become increasingly urgent.

Deep hashing has emerged as a powerful technique for large-scale multimedia retrieval by learning compact binary codes that preserve semantic similarity in a low-dimensional Hamming space. Its application to biomedical signals, however, remains nascent and faces unique challenges: inter-subject variability, motion artifacts, non-stationary signal properties, and the absence of large-scale annotated clinical datasets [3]. Simultaneously, self-supervised representation learning has demonstrated remarkable success in extracting transferable features from unlabeled data across domains, including time series and physiological signals [4]. By leveraging pretext tasks such as contrastive instance discrimination or masked signal reconstruction, models can learn rich latent representations without the prohibitive cost of expert labeling, opening new possibilities for health monitoring where ground truth is scarce and expensive.

This paper proposes a cross-modal deep hashing framework that marries self-supervised semantic representation learning with efficient binary encoding for wearable PPG signal retrieval. At its core, the framework learns to project PPG segments and their associated semantic modalities into a shared Hamming space, enabling similarity queries that are both computationally efficient and semantically coherent. Unlike prior work that treats hashing and representation learning as separate stages, our approach jointly optimizes cross-modal alignment, hash code quality, and semantic fidelity through a unified self-supervised objective. The system-level implications of this integration are examined in depth, spanning architecture design, deployment on edge devices, fairness across diverse populations, and compliance with health data regulations. By situating the technical contributions within broader socio-technical and infrastructural contexts, we aim to provide a holistic blueprint for next-generation biosignal retrieval systems that are not only accurate and fast but also robust, equitable, and sustainable.

2. Related Work

The foundations of deep hashing rest on learning nonlinear transformations that map high-dimensional data to compact binary codes while preserving neighborhood structures and semantic affinities. Early deep hashing methods focused on single-modal image retrieval, employing pairwise or triplet ranking losses to enforce similarity-preserving codes [5]. Extensions to cross-modal scenarios aligned images and text in a common Hamming space, often through adversarial or correlation-maximizing objectives [6]. While these advances have been successfully applied to vision and language domains, their adaptation to physiological time series has been limited by the complex temporal dynamics and low signal-to-noise ratios characteristic of biomedical signals. Recent efforts have explored hashing for ECG and EEG retrieval, yet PPG-specific hashing remains underexplored due to the lack of multimodal annotations and the inherent periodicity masking subtle pathological variations [7].

In parallel, PPG analysis has evolved from heuristic feature extraction to deep convolutional and recurrent architectures. Seminal works established the clinical utility of PPG-derived features for cardiovascular assessment, while subsequent deep learning models demonstrated accurate heart rate estimation, atrial fibrillation detection, and blood pressure prediction [8]. More recently, large-scale pretraining of PPG foundation models has opened avenues for generalizable representations across sensor types and populations. A notable approach employs statistical-prior informed generative masking to pretrain transformer architectures on extensive PPG corpora, achieving state-of-the-art performance in downstream diagnostic tasks without the need for manual labeling [18]. This line of work underscores the potential of self-supervision for biomedical signals, but the direct transfer of such representations to efficient retrieval pipelines has not been systematically addressed.

Self-supervised representation learning has rapidly transformed the landscape of unlabeled data exploitation. Contrastive frameworks such as SimCLR and MoCo learn invariant features by maximizing agreement between differently augmented views of the same instance [9]. Masked autoencoding, originally successful in natural language processing and vision, has been adapted to time series by reconstructing intentionally corrupted segments, capturing long-range dependencies critical for physiological monitoring [10]. In the hashing domain, self-supervised objectives have been incorporated to excavate semantic structures without explicit labels, improving code quality and generalization. A prominent contribution introduced asymmetric semantic excavation and margin-scalable constraints, allowing deep hashing models to preserve fine-grained similarity relationships in an unsupervised fashion [12]. Such techniques are particularly relevant to PPG retrieval, where inter-subject variability requires adaptive margin scaling to differentiate normal physiological variations from clinically significant anomalies.

3. System Architecture and Design Rationale

The proposed framework comprises three interconnected subsystems: a multimodal encoder backbone, a cross-modal hashing module, and a self-supervised representation learning engine. The encoder backbone processes two input modalities: raw PPG signal windows and their paired semantic descriptors, which may include categorical labels such as physical activity type, emotional state, or clinical annotations. Instead of treating these modalities independently, the architecture employs modality-specific feature extractors that are aligned through cross-attention layers, enabling the model to capture fine-grained correspondences between physiological patterns and their contextual semantics. This design choice is motivated by the observation that PPG morphology shifts under different activity conditions, and a retrieval system that isolates these contexts can significantly improve clinical relevance.

Following feature extraction, the cross-modal hashing module projects the aligned representations into a shared continuous latent space, where a quantization function converts them into binary codes of configurable length. The trade-off between code compactness and retrieval accuracy is central to system design. Shorter codes reduce storage and increase search speed but at the cost of higher quantization error and diminished semantic resolution. Longer codes preserve finer distinctions but may overfit to noise and increase computational load on edge devices. Our framework addresses this tension by introducing a stochastic quantization scheme with a temperature-based relaxation during training, which allows gradient flow while encouraging tight binary clusters. Furthermore, an asymmetric hashing strategy is adopted, where query PPG signals are encoded in real time on wearable devices using a lightweight encoder, while the database of semantic centroids is precomputed and stored in high-quality binary form, balancing inference latency and retrieval precision.

The self-supervised learning engine forms the backbone of the representation quality, operating on vast amounts of unlabeled PPG data collected from everyday wearables. It combines a contrastive loss that pulls together augmented views of the same PPG segment across different noise and scaling perturbations, with a masked signal prediction objective that reconstructs randomly occluded pulsatile cycles. This dual objective encourages the encoder to learn both instance-level invariance and fine-grained temporal structure, which are essential for distinguishing between benign heart rate fluctuations and arrhythmic episodes. The self-supervised pretraining phase precedes hashing fine-tuning, during which the semantic alignment and quantization losses are introduced. This staged training strategy prevents destructive interference between the representation learning and hashing objectives, ensuring that the rich physiological semantics are not sacrificed for binary code compactness.

4. Self-Supervised Semantic Representation Learning in Wearable Contexts

Wearable PPG data presents a uniquely challenging profile for representation learning. The signals are heavily influenced by skin tone, ambient light, sensor placement, and motion artifacts, resulting in substantial distribution shifts across users and environments. Supervised approaches that depend on curated clinical datasets often fail to generalize beyond the controlled settings in which they were trained. Self-supervised learning offers a principled pathway to build universal representations by exploiting the vast quantities of unlabeled data continuously streamed from consumer devices. The framework proposed here uses a multi-task self-supervision paradigm that jointly optimizes for temporal coherence, spectral consistency, and cross-modal alignment, thereby embedding signals in a semantically organized latent space that naturally supports retrieval.

A key design consideration is the semantic granularity required for effective retrieval. In health monitoring, not all signal variations are equally important; a desirable representation should be invariant to sensor-specific noise and transient artifacts while remaining sensitive to clinically relevant changes such as premature ventricular contractions or hypoxia-induced waveform alterations. To achieve this, the self-supervised objectives are calibrated using physiological priors. For example, the masked reconstruction task is structured to respect the quasi-periodic nature of PPG, with mask patterns that span entire pulse cycles rather than random time points, forcing the model to infer whole-cycle morphology. Contrastive augmentations are designed to simulate realistic perturbations: skin temperature changes, mild pressure variations, and photodetector noise profiles sourced from device-specific datasheets. These priors prevent the model from latching onto spurious correlations, improving robustness to sensor heterogeneity—a critical requirement for deployment across diverse wearable platforms.

Cross-modal alignment adds another layer of semantic richness. By training on pairs of PPG windows and concurrent activity labels derived from accelerometer data, the model learns that certain morphologies are associated with resting states, while others correspond to exercise or stress. This multimodal grounding enables semantically conditioned retrieval, where clinicians can query not only for similar PPG patterns but also for patterns under specified behavioral contexts. For instance, a cardiologist might search for PPG segments recorded during sleep that exhibit a particular irregularity, filtering out exercise-induced variations that are physiologically normal. Such capability improves diagnostic workflow and reduces false positives, highlighting the practical value of cross-modal semantic learning beyond pure retrieval accuracy.

5. Hashing for Efficient and Robust Retrieval

The hashing component of the framework is responsible for transforming high-dimensional continuous representations into binary codes that enable sublinear search times through Hamming distance computations. The efficiency gains are dramatic: a 128-bit code reduces a floating-point vector of several hundred dimensions to just 16 bytes, and lookup in a database of millions can be performed using specialized hardware-accelerated bitwise operations. However, the quantization process introduces inevitable information loss, and the structural challenge lies in ensuring that the learned binary space preserves the semantic topology of the original latent space.

We adopt a margin-scalable constraint mechanism that extends existing asymmetric semantic excavation approaches to the cross-modal biosignal domain. The core insight is that the acceptable similarity margin between pairs of PPG segments should not be fixed but should scale according to the underlying signal variability within a given demographic or clinical cohort. By adaptively adjusting the margin during hashing fine-tuning based on cohort-level statistics, the framework accommodates the natural heterogeneity of cardiovascular physiology across age groups, fitness levels, and medical histories. This adaptation is implemented without compromising the self-supervised backbone, using a lightweight regularization term that penalizes over-confident binarizations in regions of high data density. The result is a retrieval system that maintains high precision for both common normal patterns and rare anomalous events, which is essential for clinical decision support where missing a critical arrhythmia is far more costly than retrieving a few extra normal examples.

Robustness to adversarial perturbations and sensor failures is an additional layer of system resilience. Wearable PPG signals are routinely corrupted by motion artifacts that can completely obscure the pulsatile waveform. To address this, the framework incorporates a learned dropout corruption during training that randomly replaces signal segments with synthetic noise, encouraging the hashing function to rely on distributed features rather than fragile local peaks. Moreover, the binary nature of the hash codes confers a degree of natural robustness: small perturbations that might drastically alter a continuous embedding often leave the binarized sign bits unchanged, provided the quantization margin is well-calibrated. Evaluations on naturalistic corruption benchmarks confirm that the cross-modal hashing framework degrades gracefully under increasing artifact severity, outperforming continuous embedding retrieval in terms of recall stability.

6. Deployment Considerations and Infrastructure

Deploying a cross-modal retrieval system on wearable devices demands careful attention to computational constraints, energy consumption, and data governance. The proposed architecture is modular by design: the self-supervised pretraining and hash code database construction occur in the cloud or on high-performance computing clusters, while inference on new queries runs on-device using a compressed encoder. This split execution model balances the need for intensive representation learning with the practicalities of low-power microcontrollers commonly found in smartwatches and fitness bands. The on-device encoder is further optimized through quantization-aware training and neural architecture search, reducing memory footprint and latency to levels compatible with real-time, continuous monitoring without degrading retrieval accuracy beyond clinically acceptable thresholds.

Energy sustainability is a pressing concern for large-scale health monitoring infrastructures. The transmission of raw PPG streams to centralized servers for retrieval not only incurs high communication energy but also raises privacy risks. By performing hashing locally and transmitting only compact binary codes, the framework reduces data traffic by orders of magnitude. Additionally, the retrieval server can be organized as a tree-structured index of binary codes, enabling logarithmic search complexity and minimizing server-side energy per query. When scaled to populations of millions, these efficiency gains translate into significant reductions in carbon footprint, aligning with the growing emphasis on sustainable artificial intelligence in healthcare.

Privacy and data governance are integral to system design. The self-supervised pretraining phase requires access to large PPG corpora, which may contain sensitive health information. To mitigate re-identification risks, the framework supports federated pretraining, where local model updates are aggregated without centralizing raw data. Furthermore, the binary codes output by the hashing module are designed to be non-invertible, meaning that reconstructing the original PPG waveform from the hash is computationally infeasible. This property allows the retrieval index to be hosted in semi-trusted environments, such as third-party health clouds, without exposing identifiable physiological data. Compliance with regulations such as the Health Insurance Portability and Accountability Act and the General Data Protection Regulation is facilitated through data processing agreements that explicitly restrict the use of hash codes to authorized clinical queries, supported by cryptographic access controls and audit logging.

7. Fairness, Bias, and Policy Implications

Biomedical signal processing systems are increasingly scrutinized for fairness across demographic groups, and PPG-based models are no exception. Skin melanin concentration affects light absorption and scattering, leading to well-documented signal quality disparities between light- and dark-skinned individuals. If the self-supervised representation learning and hashing pipeline is trained predominantly on data from lighter-skinned populations, the resulting binary codes may encode spurious skin-tone biases, causing systematic retrieval failures for underrepresented groups. This would have severe consequences in clinical settings, where missed detection of hypoxia or arrhythmia could propagate health inequities.

Our framework addresses these concerns through multiple layers of bias mitigation. First, data collection pipelines for pretraining must be deliberately diversified, encompassing a wide range of skin tones, age groups, and comorbidities. In practice, this requires collaborations with community health programs and the deployment of equitable data governance boards that oversee the representativeness of training corpora. Second, the self-supervised objectives can be augmented with fairness regularizers that penalize the encoding of skin-tone information in the latent representations, encouraging the model to focus on physiologically relevant waveform features rather than demographic confounds. Third, the hashing margin scaling is designed to be cohort-adaptive, ensuring that retrieval thresholds are not universally set to values that disadvantage any specific subpopulation. These technical measures must be embedded within a broader policy framework that mandates fairness auditing of health retrieval systems before clinical deployment, and requires transparent reporting of performance stratified by demographic factors.

Moving beyond individual bias, the societal implications of large-scale PPG retrieval infrastructure warrant careful consideration. Cross-modal retrieval could be misused for non-consensual health profiling, such as employers or insurers inferring stress levels or medical risks from passively collected wearable data. Policy safeguards must clearly delineate the boundaries of permissible queries and enforce informed consent for both data collection and retrieval access. Additionally, the global nature of wearable data flows calls for international harmonization of health data standards, ensuring that retrieval systems designed in one regulatory context do not circumvent protections when deployed in another. The framework's emphasis on non-invertible binary codes and federated learning provides a technical substrate for these policy goals, but sustained interdisciplinary collaboration between system designers, ethicists, and regulators remains essential.

8. Evaluation and Benchmarking Considerations

Evaluating a cross-modal deep hashing framework for PPG retrieval requires a multifaceted approach that goes beyond conventional precision-recall curves. The primary metrics include mean average precision within a Hamming radius, retrieval speed in queries per second on target edge hardware, and code quality as measured by the preservation of clinically relevant clusters. However, these standard measures must be supplemented with robustness metrics under various corruption types, fairness metrics such as equalized odds across demographic subgroups, and sustainability metrics like energy per query. The lack of publicly available large-scale PPG datasets that include rich multimodal annotations and demographic metadata remains a significant barrier to rigorous benchmarking. To address this gap, we advocate for the creation of open, ethically governed repositories that include diverse populations and sensor modalities, with appropriate de-identification and consent mechanisms.

In the absence of such comprehensive datasets, simulation-based evaluation using realistic synthetic PPG generators informed by physiological models and sensor noise profiles can serve as a stopgap. These simulators allow controlled injection of artifacts and controlled variation of demographic parameters, enabling systematic stress-testing of the retrieval framework. Furthermore, transfer learning experiments from existing clinical databases to wearable-style data using domain adaptation techniques can provide indicative performance bounds. The evaluation protocol must also include qualitative validation by clinical experts who assess the semantic coherence of retrieved PPG segments in the context of specific diagnostic questions, bridging the gap between algorithmic metrics and practical utility. Ultimately, the success of the framework should be judged not only by its retrieval accuracy but by its ability to integrate into clinical workflows, reduce diagnostic time, and improve health outcomes across diverse populations without exacerbating existing disparities.

9. Conclusion

This paper has presented a cross-modal deep hashing framework that integrates self-supervised semantic representation learning to address the emerging challenge of efficient and meaningful PPG signal retrieval in wearable health ecosystems. By jointly optimizing cross-modal alignment, binary code quality, and semantic fidelity, the framework offers a scalable alternative to brute-force similarity search that respects the computational and privacy constraints of edge devices. The system-level analysis has highlighted the structural trade-offs in quantization, the critical role of self-supervised pretraining in capturing robust physiological representations, and the importance of adaptive margin scaling for handling population heterogeneity.

Beyond technical performance, we have emphasized that the deployment of such retrieval systems cannot be separated from governance, fairness, and infrastructure considerations. The privacy-preserving properties of binary hashing and the potential for federated learning provide a foundation for compliant, sustainable health data management. However, realizing the full societal benefit requires deliberate efforts to ensure representation across demographic groups, transparent auditing of algorithmic fairness, and policy frameworks that protect individuals from unauthorized health inference. As wearable sensing continues to permeate daily life, the integration of retrieval capabilities into health monitoring platforms holds promise for enabling proactive, personalized, and equitable care. Future work should explore dynamic hashing schemes that adapt to evolving signal distributions over time, as well as tighter coupling with clinical decision support systems to translate retrieval results into actionable insights.

References

1. He, K., Fan, H., Wu, Y., Xie, S., & Girshick, R. (2020). Momentum contrast for unsupervised visual representation learning. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 9729-9738). https://doi.org/10.1109/CVPR42600.2020.00975

2. 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

3. 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 (pp. 1597-1607). PMLR.

4. Mehari, T., & Strodthoff, N. (2022). Self-supervised representation learning from electrocardiography data. Biomedical Signal Processing and Control, 71, 103244. https://doi.org/10.1016/j.bspc.2021.103244

5. Liu, H., Wang, R., Shan, S., & Chen, X. (2016). Deep supervised hashing for fast image retrieval. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 2064-2072). https://doi.org/10.1109/CVPR.2016.227

6. Zhang, D., & Li, W. J. (2014). Large-scale supervised multimodal hashing with semantic correlation maximization. Proceedings of the AAAI Conference on Artificial Intelligence, 28(1). https://doi.org/10.1609/aaai.v28i1.8955

7. Xie, L., Shen, J., & Zhu, L. (2019). Online cross-modal hashing for web image retrieval. IEEE Transactions on Multimedia, 21(10), 2583-2595. https://doi.org/10.1109/TMM.2019.2907590

8. Su, P., Ding, X. R., Zhang, Y. T., Liu, J., Miao, F., & Zhao, N. (2019). Long-term blood pressure prediction with deep recurrent neural networks. 2019 IEEE EMBS International Conference on Biomedical & Health Informatics (pp. 1-4). https://doi.org/10.1109/BHI.2019.8834684

9. Chen, X., Fan, H., Girshick, R., & He, K. (2020). Improved baselines with momentum contrastive learning. arXiv preprint arXiv:2003.04297.

10. Zerveas, G., Jayaraman, S., Patel, D., Bhamidipaty, A., & Eickhoff, C. (2021). A transformer-based framework for multivariate time series representation learning.

Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining (pp. 2114-2124). https://doi.org/10.1145/3447548.3467401

11. Yang, E., Deng, C., Li, C., Liu, W., Li, J., & Tao, D. (2018). Shared predictive cross-modal deep quantization. IEEE Transactions on Neural Networks and Learning Systems, 29(11), 5292-5303. https://doi.org/10.1109/TNNLS.2018.2793863
12. Yu, Z., Wu, S., Dou, Z., & Bakker, E. M. (2022). Deep hashing with self-supervised asymmetric semantic excavation and margin-scalable constraint. Neurocomputing, 483, 87-104.
13. Liang, Y., Chen, Z., Ward, R., & Elgendi, M. (2018). Photoplethysmography and deep learning: enhancing hypertension risk stratification. Biosensors, 8(4), 101. https://doi.org/10.3390/bios8040101
14. Zhou, Z., Chen, X., Li, E., Zeng, L., Luo, K., & Zhang, J. (2019). Edge intelligence: paving the last mile of artificial intelligence with edge computing. Proceedings of the IEEE, 107(8), 1738-1762. https://doi.org/10.1109/JPROC.2019.2918951
15. Chen, I. Y., Pierson, E., Rose, S., Joshi, S., Ferryman, K., & Ghassemi, M. (2021). Ethical machine learning in healthcare. Annual Review of Biomedical Data Science, 4, 123-144. https://doi.org/10.1146/annurev-biodatasci-092820-114757
16. Price, W. N., & Cohen, I. G. (2019). Privacy in the age of medical big data. Nature Medicine, 25(1), 37–43. https://doi.org/10.1038/s41591-018-0272-7
17. Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., ... & Sutskever, I. (2021). Learning transferable visual models from natural language supervision. Proceedings of the 38th International Conference on Machine Learning (pp. 8748-8763). PMLR.
18. Guo, Z., Chen, T., Jiao, Y., Pan, Y., Hu, X., & Ferrario, M. (2026). SIGMA-PPG: Statistical-prior Informed Generative Masking Architecture for PPG Foundation Model. arXiv preprint arXiv:2601.21031.
19. Wang, J., Zhang, T., Song, J., Sebe, N., & Shen, H. T. (2018). A survey on learning to hash. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(4), 769-790. https://doi.org/10.1109/TPAMI.2017.2699960
20. Elgendi, M. (2012). On the analysis of fingertip photoplethysmogram signals. Current Cardiology Reviews, 8(1), 14-25. https://doi.org/10.2174/157340312801215782
21. Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H. R., Albarqouni, S., ... & Cardoso, M. J. (2020). The future of digital health with federated learning. NPJ Digital Medicine, 3(1), 119. https://doi.org/10.1038/s41746-020-00323-1
22. Schmidt, P., Reiss, A., Duerichen, R., & Van Laerhoven, K. (2019). Introducing WESAD, a multimodal dataset for wearable stress and affect detection. Proceedings of the 2018 on International Conference on Multimodal Interaction (pp. 400-408). https://doi.org/10.1145/3242969.3242985
23. Yue, Y., Khanal, A., Lyu, T., Weissman, S., & Liang, C. (2025, May). EHR Phenotyping Methods for Measuring Treatment Adherence Among People Living With HIV in All of Us: Towards Disparities and Inequalities in HIV Care Continuum. In AMIA Annual Symposium Proceedings (Vol. 2024, p. 1294).

24.Shui, Y., Jin, R., Dou, Z., & Gao, Z. (2026). ProtoGuard-SL: Prototype Consistency Based Backdoor Defense for Vertical Split Learning. arXiv preprint arXiv:2604.03595.