Wang Yi (王熠)

Research Portfolio for PhD / RA Outreach · Wearable Health Sensing, Physiological Signals, Mobile Health, and Reliable Multimodal Systems

PPG / BVP / rPPG / BCG / ECG	Wearable & Mobile Health	Reliability-Aware ML	CS / ECE / Biomedical AI / HCI-Health

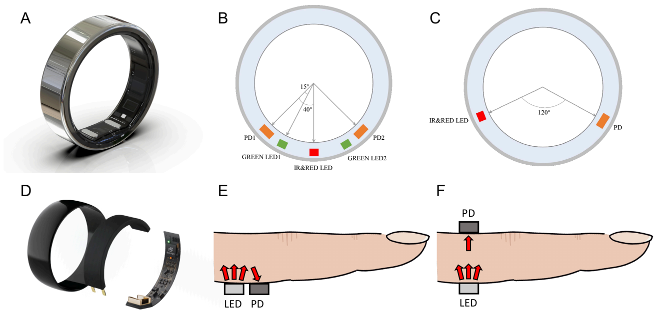
Research Identity

I build wearable and physiological sensing systems that connect real-world data collection, signal processing, machine learning, reliability-aware evaluation, and deployment-constrained prototypes. My current work focuses on PPG / rPPG / BCG / ECG, sleep and cardiorespiratory signals, motion-robust modeling, and mobile health systems.

Core Project: CITRUS

CITRUS studies motion-robust heart-rate estimation from ring PPG/BVP signals. It treats motion-corrupted monitoring as candidate disambiguation plus reliability-aware reporting, rather than only direct denoising or scalar regression.

- tau-Ring benchmark with 54 participants and 49.76 hours.
- IR/red PPG and 3-axis accelerometry at 25 Hz.
- Clinical references for HR, RR, SpO2, and blood pressure.
- Causal decoding and accept / hold / reject reporting for streaming deployment.



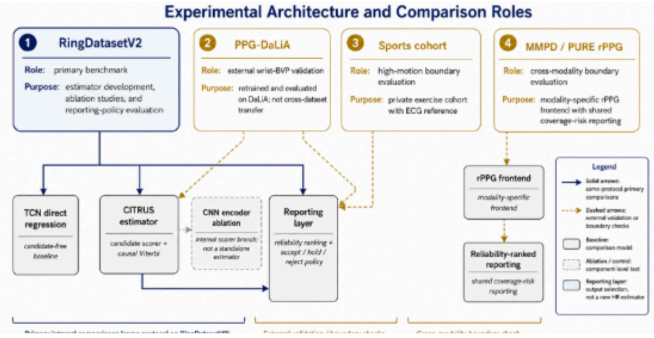
Real paper figure: ring hardware and optical sensor layout for physiology sensing.

54	49.76 h	10.91 BPM	0.898
participants in the tau-Ring benchmark	wearable physiological recordings	motion MAE at full coverage	AUC for detecting error > 10 BPM

> Portfolio goal: show a coherent research trajectory from wearable sensing hardware and benchmark design to robust physiological modeling, reproducible evaluation, and deployment-aware health AI.

CITRUS: Method and Evaluation Design

Motion-robust wearable PPG/BVP monitoring with reliability-aware reporting.



Real thesis figure: primary ring benchmark, external wrist-BVP validation, private sports cohort, and rPPG boundary checks.

Problem Framing

Wearable PPG fails under motion because true pulse peaks, motion peaks, harmonics, and contact artifacts can overlap in the same cardiac frequency band. A practical wearable stream needs both an internal estimate and a decision about whether that estimate should be reported.

Method

- Generate multiple plausible HR candidates per analysis window.
- Rank candidates with source, spectral, consensus, harmonic, and temporal features.
- Decode a causal physiologically consistent HR path.
- Estimate reliability and apply accept / hold / reject reporting.

Research Contribution

The contribution is not only a lower error number. The thesis decomposes a deployable wearable monitoring problem into benchmark construction, candidate-based inference, temporal consistency, reliability modeling, and coverage-error reporting.

Why It Transfers

The same research pattern is relevant to CS, ECE, Biomedical AI, Digital Health, and HCI-health: noisy multimodal signals, temporal models, reliability under real-world conditions, and deployment-aware evaluation.

CITRUS: Results, Ablations, and Deployment

Evidence that the system is evaluated as a stream, not only an offline model.

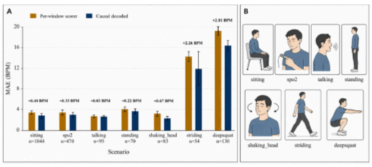


Figure 4.1: Scenario-wise RingDatasetV2 MAE for the per-window scorer and the causal

Real thesis figure: scenario-wise MAE for per-window scoring and causal decoded CITRUS.

Table 5.2: Ablation summary for the motion-robust HR system.

Component	Control	Conclusion	Key result
Candidate source	DSP candidates vs. uniform grid	signal-derived structure matters	10.91 vs. 12.06 motion
Estimator form	TCN regression vs. candidate path	candidate path improves motion	11.62 vs. 10.91 motion
Temporal decoding	row-wise scorer vs. Viterbi	sequence path selection helps	12.94 → 10.91 motion
Decoder latency	L=0 vs. offline	no future context required	10.66 vs. 10.88 motion
Scorer encoder	disable raw-waveform CNN	encoder redundant	Δoverall −0.01 BPM
ACC in candidates	remove ACC slice/features	useful upstream	10.73 vs. 11.76 motion
Reliability features	PPG vs. PPG+ACC vs. ACC	reliability is PPG-driven	AUC 0.898/0.888/0.710
Reporting policy	estimator stream vs. policy stream	hold/reject improves emitted HR	10.91 → 9.99 motion

Ablation evidence: candidate sources, candidate path, temporal decoding, ACC features, reliability features, and reporting policy.

Candidate generation	current PPG/ACC window	DSP operations	no trainable parameters
Candidate scorer	candidate features and optional waveform slice	85k params; 28k without CNN	small learned scorer
Causal Viterbi	previous DP state + current candidates	per-window path update	past-only, streamable
Reliability model	decoded HR and quality features	HGB model dominates memory	compressible or replaceable
Reporting policy	reliability bin + hold count	negligible	accept/hold/reject output stream

Deployment evidence: streaming inputs, small learned scorer, causal Viterbi state, and negligible reporting policy cost.

14.39 -> 10.91
motion MAE reduction versus LSTM baseline

3.73
overall MAE in BPM

2.90
motion MAE at 31% motion coverage

17.67 ms
p95 CPU replay latency under 2 s hop

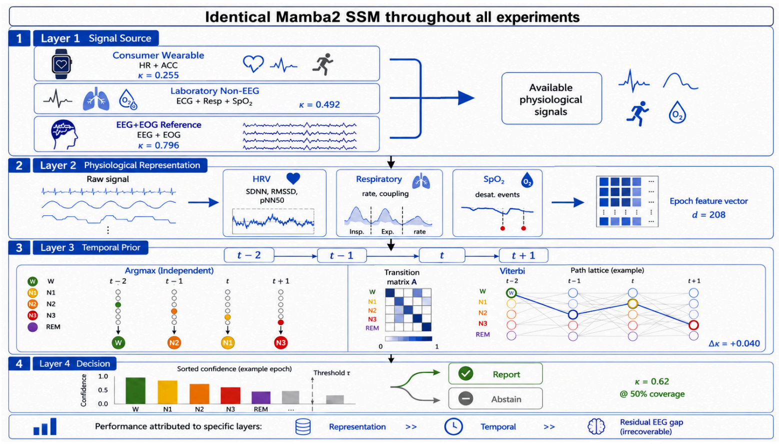
This evidence supports a research profile centered on reliable wearable streams: data collection, model design, temporal decoding, reliability-aware reporting, and deployment constraints are treated together instead of as disconnected tasks.

Wearable-Compatible Sleep Staging

Non-EEG cardiorespiratory sleep staging with controlled decomposition and confidence-aware reporting.

BIBM 2026 Draft: How Far Can Wearable-Compatible Signals Go?

This paper studies how far consumer- and laboratory-grade non-EEG physiological signals can go for sleep staging. The contribution is not an overclaimed SOTA model; it is a controlled decomposition of where performance comes from: signal source, physiological representation, temporal prior, and decision / abstention.



Four-layer decomposition: signal source, physiological representation, temporal prior, and decision / abstention.

0.253
naive non-EEG baseline kappa

0.485
physiology-aware SHHS rpoint200 kappa

0.616
kappa at 50% coverage in abstention analysis

195
matched SHHS subjects with cardiorespiratory signals

Why It Matters

Consumer wearables often report sleep stages from HR, accelerometry, PPG-derived features, and limited physiological proxies rather than EEG. This work asks what those signals can honestly recover, what remains ambiguous, and when a device should abstain instead of reporting a hard label. The research logic matches CITRUS: decompose a noisy physiological inference problem, quantify each component's contribution, and use reliability / confidence to separate trustworthy outputs from uncertain windows or epochs.

PhD fit: temporal modeling, signal processing, abstention, physiological state estimation, reproducible biomedical AI, digital health, and HCI-health.