Title :

An AI-Enhanced Framework for Robust Remote Photoplethysmography (rPPG) Using Multimodal Signal Processing and Fine-Tuned Gemma-Medical

Contactless HR & HRV

Standard RGB camera ~30 FPS

Multimodal + medical AI

Full name of Speaker :

Dr. Phonepadith PHOUMMAVONG

Institution :

Department of Computer Engineering, Faculty of Engineering, National University of Laos (Lao PDR)

Project Title: Background

Background :

- **Remote photoplethysmography (rPPG)** estimates heart rate (HR) and heart rate variability (HRV) from facial video captured by a conventional RGB camera.
- Low-cost and contactless, it has strong potential for remote health monitoring, telemedicine and digital healthcare.
- In Laos and across the ASEAN region, this can bring preliminary health assessment to remote and underserved communities using hardware that is already available.

Why is rPPG still unreliable in real conditions?

- **Weak pulse signal** Small colour changes are easily buried in noise
- **Head motion** Movement creates strong motion artifacts
- **Illumination variation** Lighting changes distort the skin colour signal
- **Facial occlusion** Hair, hands or masks hide skin regions
- **Skin-region changes** Informative facial areas shift over time

Gap: conventional rPPG relies on fixed signal-processing steps and predefined filter parameters, which cannot adapt to complex, non-linear signal distortions. → **An adaptive, AI-enhanced approach is needed.**

Project Title: Targets

Targets:

Primary objective: develop and evaluate an AI-assisted framework for robust rPPG-based physiological monitoring from standard RGB video at approximately 30 frames per second.

01

Facial acquisition & alignment

Robust face region acquisition and alignment for continuous rPPG analysis

02

Multimodal inputs

Combine RGB temporal signals, POS pulse signal and 3-D facial motion

03

AI signal cleaning

Skin-region weighting, motion-artifact suppression and signal denoising

04

Adaptive frequency bands

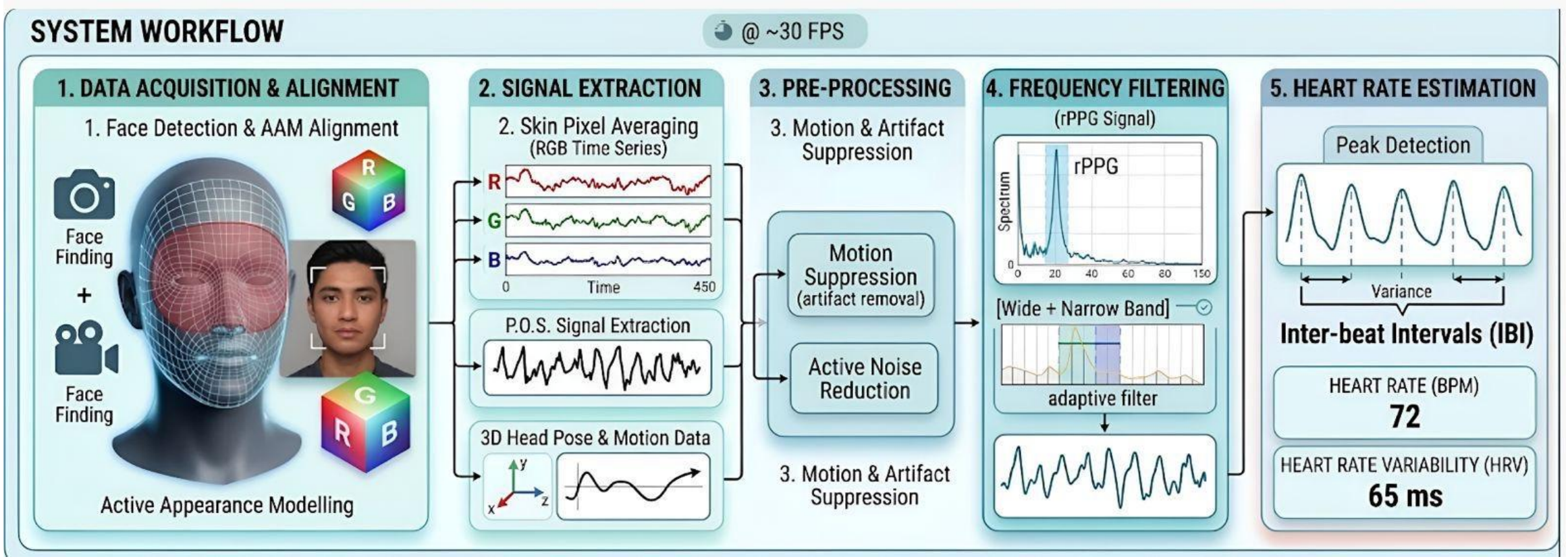
AI-assisted selection of the most informative frequency band

05

HR / HRV + confidence

Estimate IBI, HR, HRV and signal quality, with feedback to earlier stages

Proposed Method: (1) Scientific & technological — five-stage workflow



Input
Standard RGB camera, ~30 FPS

Multimodal signals
RGB · POS pulse · 3-D head pose

AI in every stage
Classical processing + learned modules

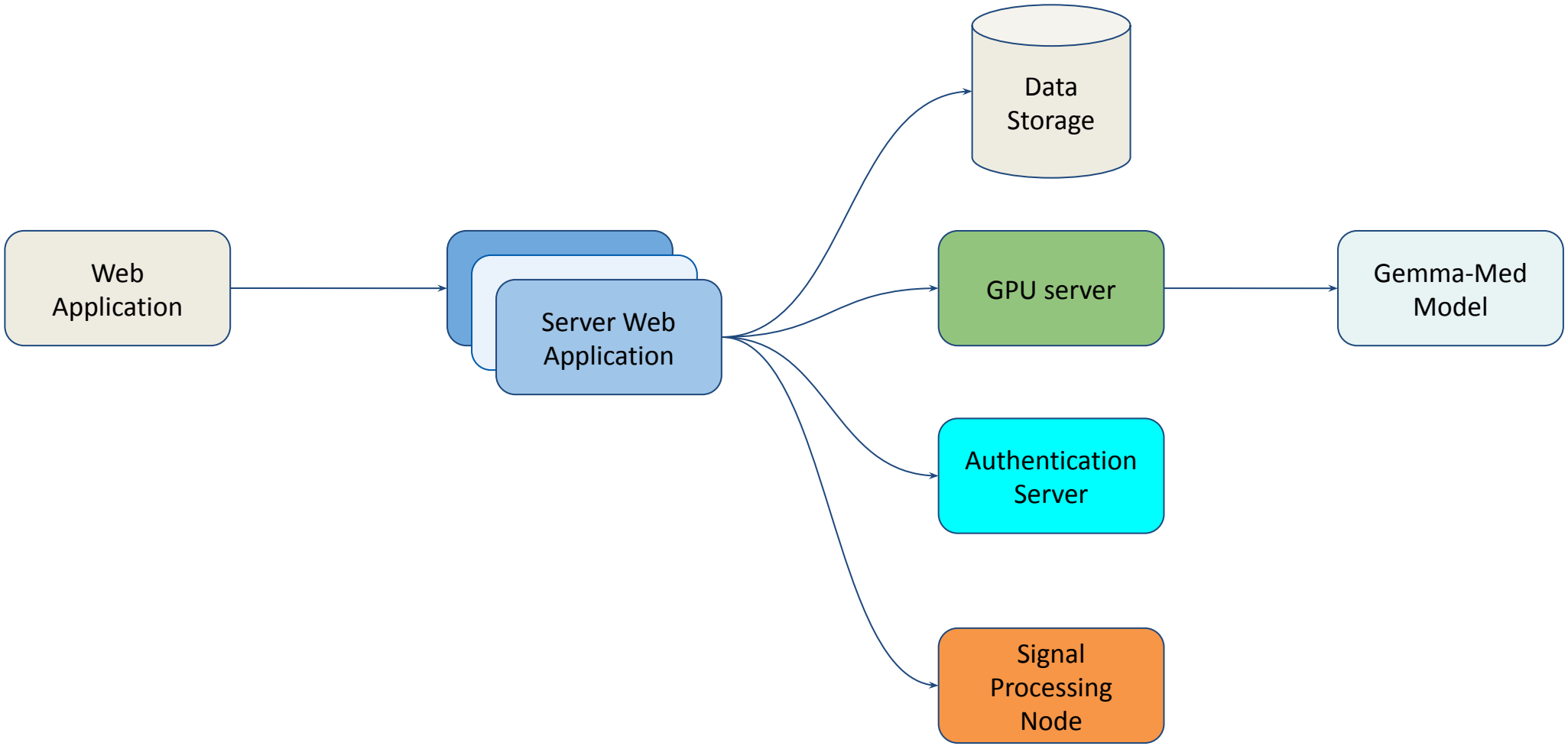
Outputs
IBI · HR · HRV · confidence

Proposed Method: (2) Prototype Device for testing rPPG





Proposed Method: (2) Software Infrastructure



Proposed Method: (2) Where AI enhances each stage

Stage	Classical processing	AI enhancement
1 Acquisition & alignment	Face detection, landmarks, AAM alignment	Deep face detector + landmark-tracking CNN for robust position / orientation
2 Signal extraction	Skin-pixel RGB averaging, POS pulse, 3-D head pose	Skin segmentation with attention-based ROI weighting
3 Pre-processing	Motion suppression, active noise reduction	Denoising autoencoder learns and removes complex artifacts
4 Frequency filtering	Spectrum analysis, wide + narrow band filter	Spectral attention selects and adjusts the best frequency band
5 HR / HRV estimation	Peak detection → inter-beat intervals (IBI)	Temporal model (LSTM / Transformer): HR, HRV + signal-quality score

Closed loop: the estimated signal quality feeds back to re-select facial regions (Stage 2) and re-tune frequency bands (Stage 4).

Proposed Method: (3) Implementation & experiments

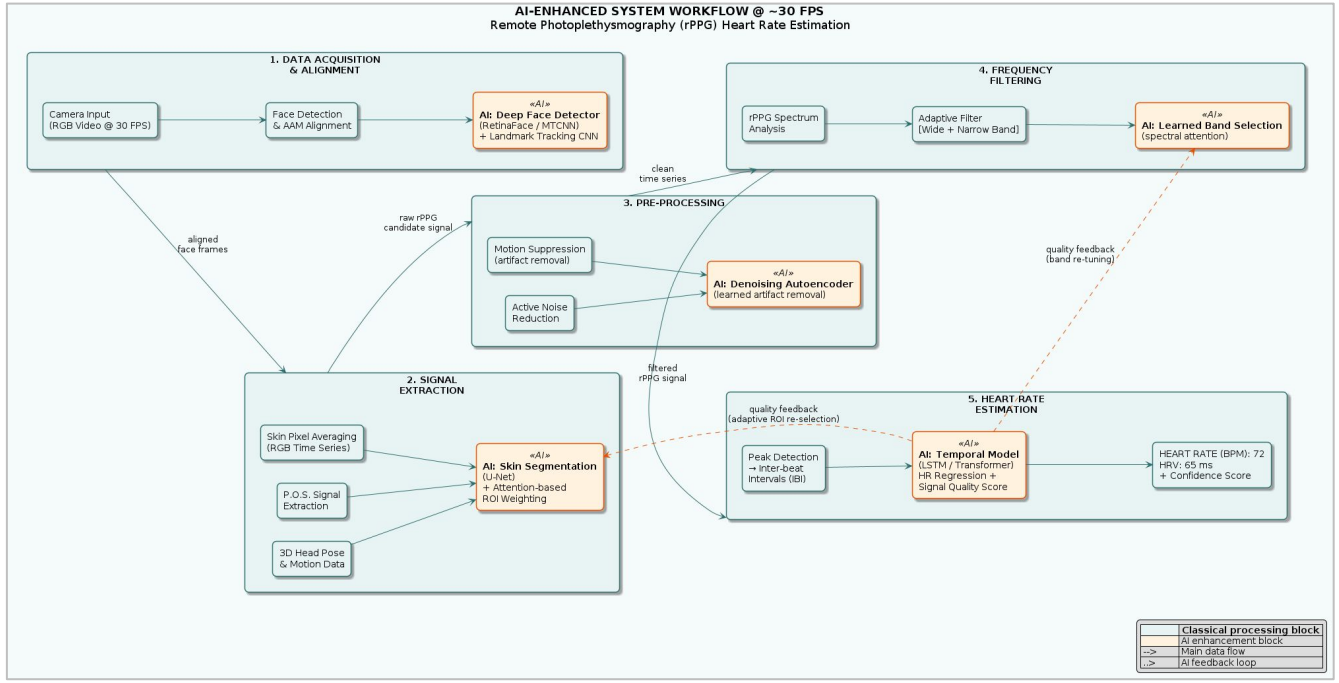


Figure 2. AI-enhanced workflow: classical blocks (teal) and AI blocks (orange) with feedback loops.

AI blocks: deep face detector · skin segmentation (U-Net) · denoising autoencoder · spectral attention · temporal model (LSTM / Transformer)

Implementation

- Standard RGB camera at ~30 FPS — no contact sensors
- Fine-tuned Gemma-Medical as the medical-AI layer for adaptive processing and physiological parameter estimation
- Signal-quality feedback adapts ROI selection and frequency-band parameters
- Model optimisation for lightweight edge-computing devices

Experiments & field testing

- Compare HR / HRV with a contact reference (ECG or finger PPG)
- Test under head motion, illumination change and partial occlusion
- Metrics: HR error (MAE, RMSE), IBI error, HRV agreement, confidence reliability
- Pilot on low-cost screening kiosks in remote communities

Impact: (1) Scientific & technological

Conventional rPPG	Proposed AI-enhanced framework
Parameters: Fixed, predefined filter settings	**Parameters:** Adaptive — AI selects frequency bands
Noise & motion: Hand-designed filters	**Noise & motion:** Learned denoising of non-linear artifacts
Skin region: Static region of interest	**Skin region:** AI-weighted, informative skin regions
Information: Mainly colour signal	**Information:** Facial, physiological, motion, spectral & temporal
Output: Heart rate	**Output:** IBI, HR, HRV + signal-quality confidence
Architecture: Open-loop pipeline	**Architecture:** Closed loop with quality feedback

Impact: (2) Societal & collaborative



Telehealth & remote monitoring

- Contactless vital-sign checks with an ordinary webcam or phone camera
- Supports telemedicine and digital-healthcare services



Access for underserved communities

- Accessible preliminary health assessment in remote areas
- Low-cost screening kiosks — no specialised medical sensors



ASEAN collaboration

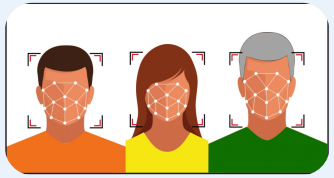
- Joint development of localized AI for digital healthcare
- Shared research platform among ASEAN institutions

Output/Outcome: (1) Scientific — new technologies & applications

New technology

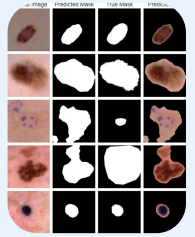
A unified, closed-loop AI-enhanced rPPG framework that combines classical signal processing with fine-tuned medical AI.

New application: robust contactless HR / HRV monitoring from commodity RGB cameras.



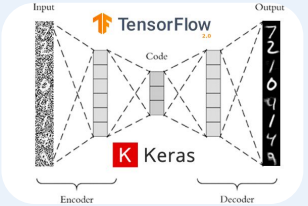
Robust face alignment

Deep detection + landmark tracking with AAM



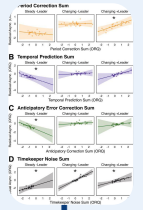
Skin ROI weighting

Segmentation with attention-based region weights



Learned denoising

Autoencoder for motion & noise artifacts



Adaptive band + temporal model

Spectral attention; HR, HRV & confidence

Output/Outcome: (2) Societal & collaborative — roadmap



Framework

AI-enhanced multimodal rPPG pipeline (this work)

Edge deployment

Model optimisation for lightweight edge-computing platforms

Screening kiosks

Low-cost medical screening kiosks for remote, underserved communities

ASEAN network

Collaboration on localized AI for digital healthcare & remote monitoring

Societal: technology that can be transferred to telehealth providers and kiosk / device makers.

Collaborative: new partners among ASEAN research institutions working on accessible digital healthcare.

Conclusion:

Targets

Reliable, contactless HR and HRV monitoring from a standard RGB camera (~30 FPS), robust to motion, lighting changes and occlusion.

Method (idea)

Five-stage rPPG pipeline — alignment, multimodal extraction, AI pre-processing, adaptive filtering, HR/HRV estimation — with AI in every stage and a quality-feedback loop.

Scientific impact

Moves rPPG from fixed parameters to adaptive processing by unifying facial, physiological, motion, spectral and temporal information.

Societal impact

Low-cost telehealth and screening kiosks for remote communities, and a basis for ASEAN collaboration on localized healthcare AI.

Thank you

Questions & discussion

Dr. Phonepadith PHOUMMAVONG

Department of Computer Engineering, Faculty of Engineering, National University of Laos