VitaMon: Heart rate variability Measurement [1]

S. Huynh, R. Balan, J. Ko. et al.

Presented by
Zahid Hasan

[1] Sinh Huynh, Rajesh Krishna Balan, JeongGil Ko, and Youngki Lee. 2019. *VitaMon: measuring heart rate variability using smartphone front camera.* In Proceedings of the 17th Conference on Embedded Networked Sensor Systems (SenSys '19). Association for Computing Machinery, New York, NY, USA, 1–14. DOI:https://doi.org/10.1145/3356250.3360036

Problem statement

- Physiological information from Video
- HR
  - Number of Heart Beats per minute (Hz)
  - Less care about heart beat exact locations
- Heart Rate Variability
  - Variation in time between each heartbeats (ms)
  - Need exact peak locations (in ms)
    - Video frame restriction
  - Measured by time and frequency features
- Peak occurs in between video frame!
- Vital signal monitoring

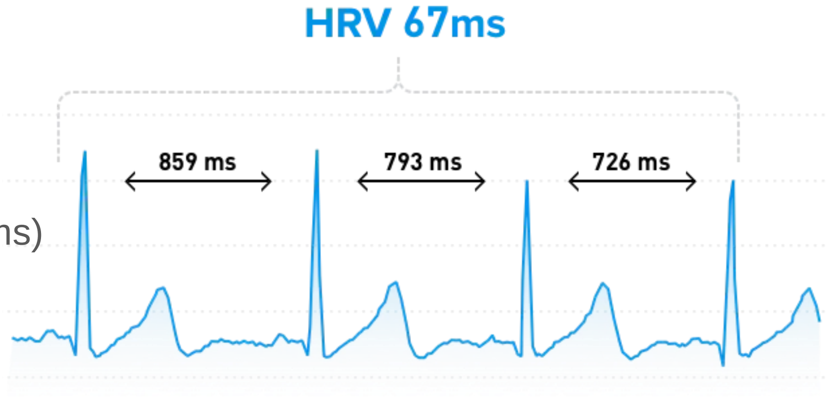


Figure: HRV [*]

*https://www.whoop.com/the-locker/heart-rate-variability-hrv/

Assumptions

- Human face contain multiple Cardiovascular pulse with phase difference
- Use both spatial-temporal features
- Correlation between PPG and ECG
  - Interpolation using phase information
- Encoded in Phase!
- Different tasks
  - Two networks
  - Dependent but not loss sharing

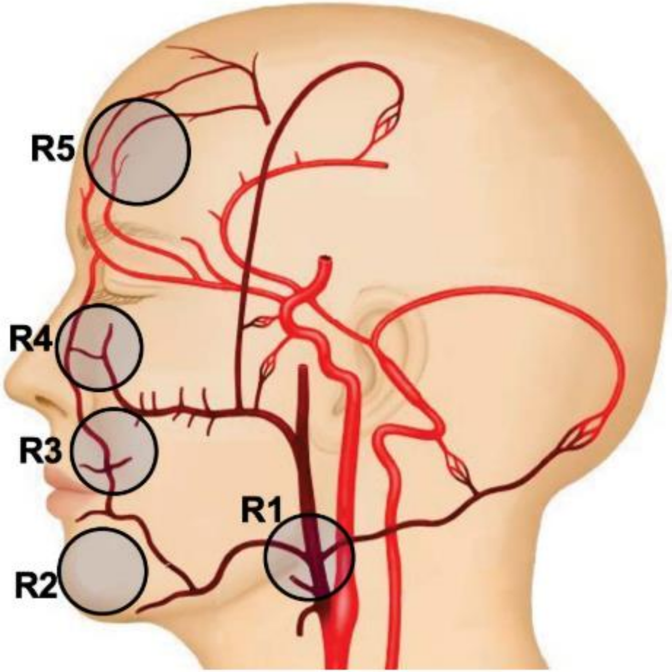


Figure: Facial Artery [1]

Contributions

- Use of Mobile Phone Camera
- Going higher granularities than Video frame
- CNN structure to extract HR
- HRV with high accuracy

Prior Arts

- PPG
  - Contact sensors
  - Low cost
  - Problem with ubiquitous
- HRV methods
  - ECG (Most accurate)
  - Expensive and contact
- Remote PPG
  - Video based PPG
  - Use for Heart rate
  - HRV limited by video frames

VitaMon Design

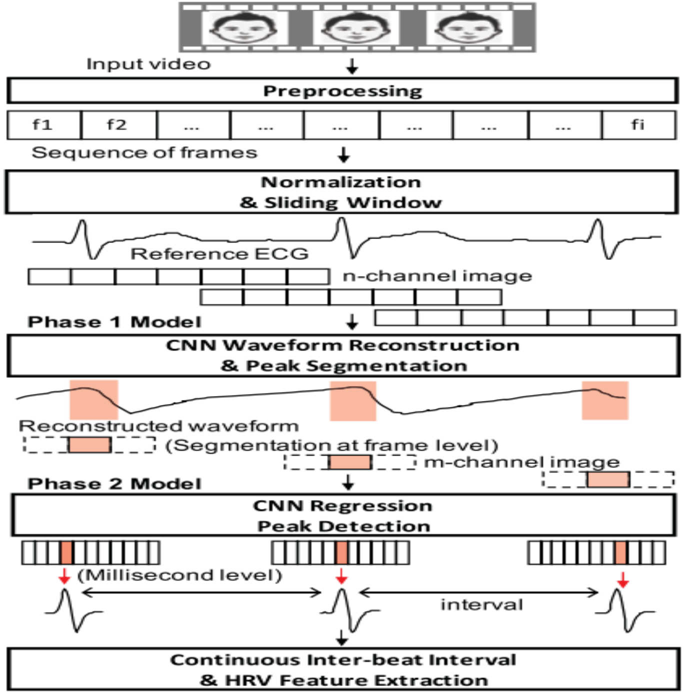
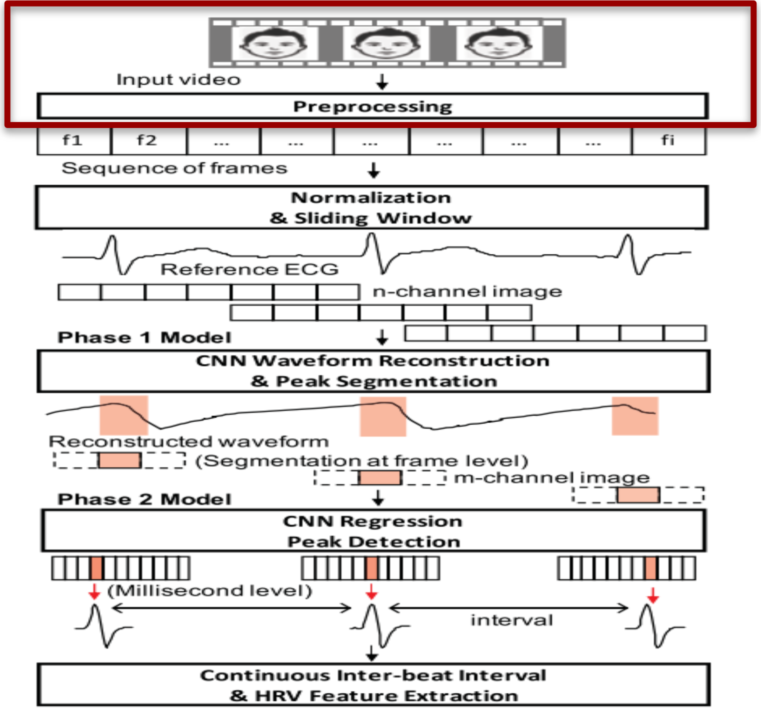


Figure: VitaMon methodology pipeline overview [1]

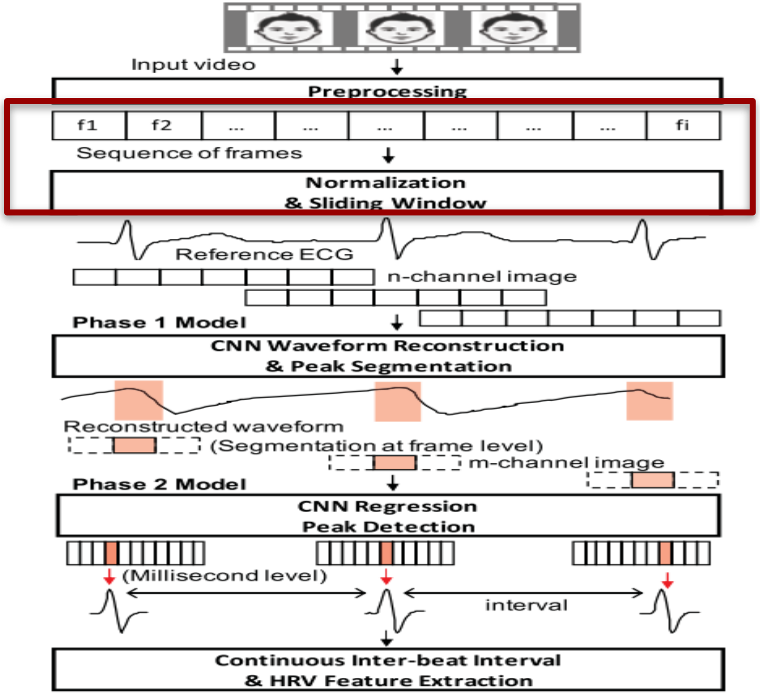
VitaMon Design



Preprocessing:

- Resize 224x224
- Select Green channel only
    - Highest Absorption

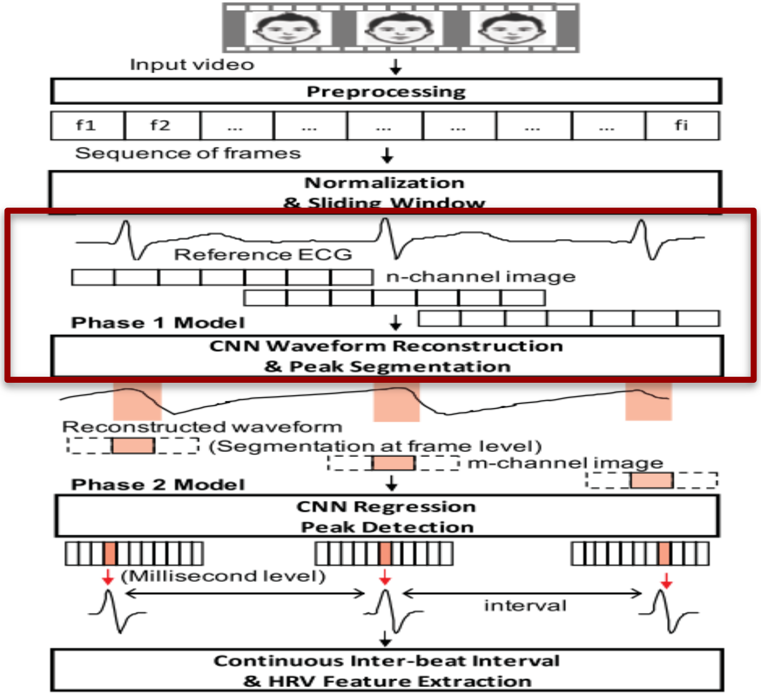
VitaMon Design



Normalizing and Sliding Window:

- Stack n = 25 frames (1.67s video) in one image of n-channel
  - Both spatial-temporal info in single image
  - Reduce complexity
- Normalization

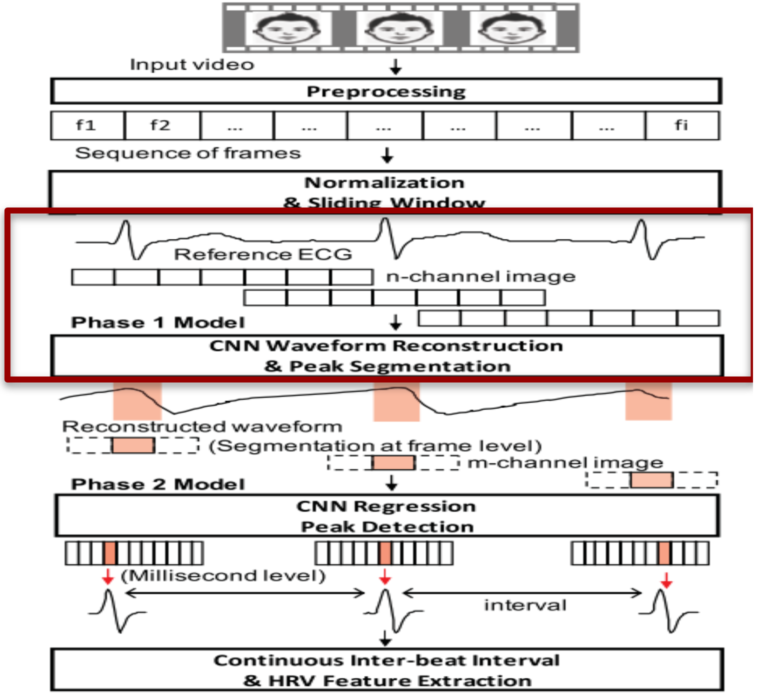
VitaMon Design



Phase 1: Reconstruction and Segmentation (HR)

- Need frame where Peak occurred
- Align 25 frames and ECG for labeling
  - For ECG peak in 13th frame 0
  - Or distance from 13th frame
- Construct Frame order waveforms
- Benefits
  - Normalized ECG representation
  - Distinguished between two peaks
- ~500 K parameters

VitaMon Design



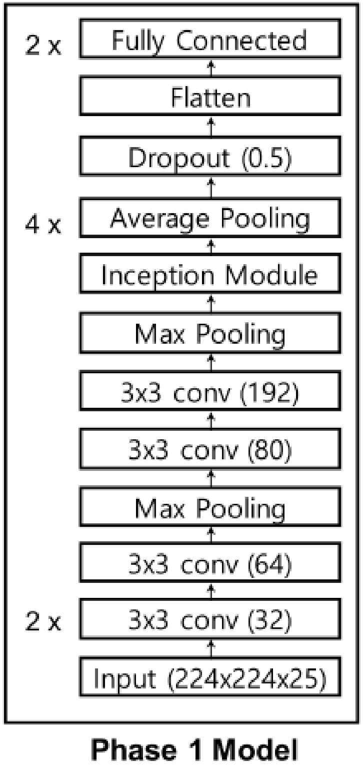


Figure: Architecture Phase 1 [1]

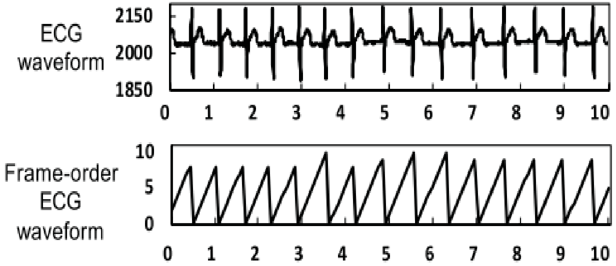
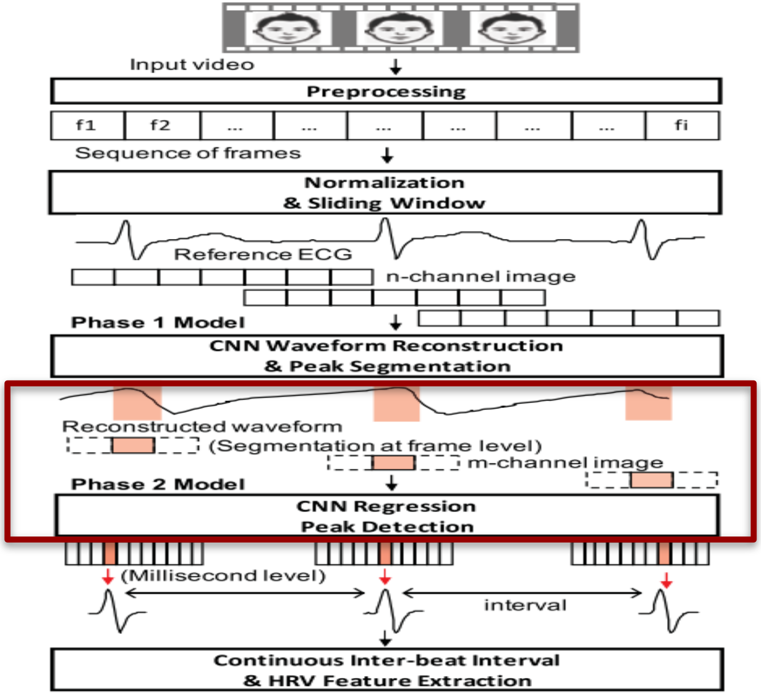


Figure: Baseline ECG and reconstructed ECE [1]

VitaMon Design



Phase 2: Peak Detection

- Choose 25 channels where CNN output 0
- Choose m = 7 channel from center.
- ~100K parameters
- Learn phase differences to learn actual R-wave Peak location
- Frame label to millisecond label
  - Interpolation

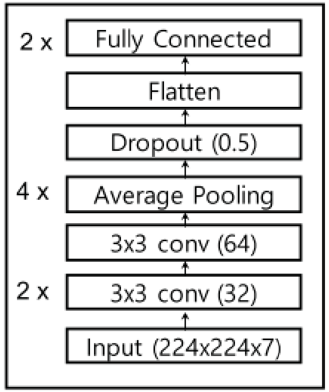
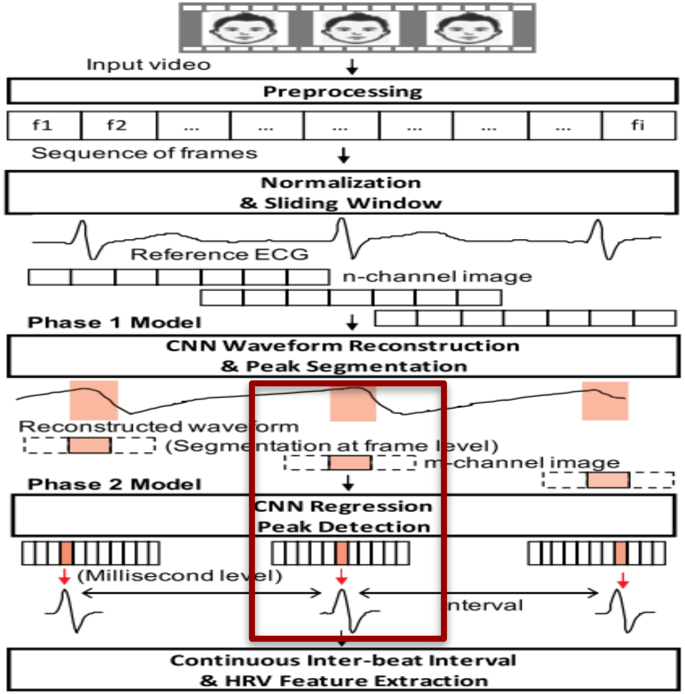


Figure: Phase 2 model Architecture [1]

VitaMon Design



Phase 2: Peak Detection

- HRV features Extraction
- Time domain and frequency domain analysis

Data Collection

- Sensors
  - Video-Lenovo Phab Pro2, Galaxy S8, Huawei P20 smartphone (Data)
    - Front Camera
    - 1920x1080, 8- MP, 15fps
    - Face video from 25-50 cm
  - ECG (Ground Truth)
    - Zephyr Bioharness 3 ECG strap
    - 250Hz
- Controlled Group
  - 30 Participants with different age, skin, race
  - Eight tasks; each for 5 minutes (No physical works)
    - Speaking, Counting, Head Motion, Manual phone holding
  - 5 light intensities (150 lux to 1000 lux)

Data Collection

- Real-world experiments
  - Passenger in driving Car
    - 5 Mins, handheld phone
  - Coffee shop
    - Dim light (40 lux)
- Stress data (Application - Test only)
  - 12 participants
  - Under Physiological stress

Results (HR)- Phase 1

		Light Condition				
Metric	Model	L1	L2	L3	L4	L5
HR MAE (bpm)	General	0.82	1.06	0.82	0.94	0.88
	Personalized	0.67	0.72	0.61	0.61	0.56
Peak Position MAE (frame)	General	0.78	0.98	0.76	0.80	0.84
	Personalized	0.63	0.72	0.65	0.72	0.62

		Motion Artifact Condition			
Metric	Model	M0	M1	M2	M3
HR MAE (bpm)	General	0.82	1.77	1.69	1.31
	Personalized	0.61	1.23	1.38	1.08
Peak Position MAE (frame)	General	0.76	1.33	1.45	1.32
	Personalized	0.65	1.02	1.19	1.18

Table 1: HR performance of Phase-1 network under light and motion artifact condition using General and Personalized Model [1]

UMBC

Result (HRV)- Phase 2

		HRV features								
Statistic	Source	RMSSD	SDNN	MRRI	NN50	PNN50	SD1	SD2	LFnu	HFnu
Mean	ECG	111.65	87.32	751.92	9.90	13.50	80.39	96.01	25.67	74.40
Mean	VitaMon	114.61	89.30	749.10	33.54	46.10	79.62	95.70	33.74	65.74
Standard Deviation	ECG	70.58	45.07	68.52	9.27	12.38	50.12	42.84	30.45	30.45
Standard Deviation	VitaMon	54.43	38.58	68.37	10.19	12.05	39.09	39.42	21.92	21.70
Correlation Coefficient		0.9817	0.9776	0.9943	0.4697	0.4317	0.9717	0.9710	0.72	0.72

Statistic	Source	RMSSD	SDNN	MRRI	NN50	PNN50	SD1	SD2	LFnu	HFnu
Mean	ECG	112.20	88.51	751.00	9.88	13.46	80.36	95.39	25.97	74.10
Mean	VitaMon	114.51	89.47	750.15	16.46	23.30	80.55	95.77	30.66	68.81
Standard Deviation	ECG	69.94	44.82	68.52	9.33	12.46	49.66	42.79	30.54	30.54
Standard Deviation	VitaMon	61.29	41.49	68.38	9.55	13.02	43.97	4.90	24.71	24.50
Correlation Coefficient		0.9879	0.9836	0.9855	0.7948	0.7394	0.9861	0.9830	0.8134	0.8139

Table 2: HRV monitoring performance of the General Model (Top) and Personal (Bottom) on different HRV features [1]

Metrics and Analysis

- Mean absolute error (MAE)
  - Any other metrics!
- Statistics of 9 HRV features and MAE
- Analysis of Light condition
- Analysis of Motion artifact
  - Small motion like speaking, head shaking

Paper Strength

- Considering frames to make a big image and apply CNN
- Peak Point frame detection
- Going Beyond Video Frame constraints by phase information
- 2 Stages Learning and Setup
    - Second stage used input refined from 1st stage

Potential Weakness

- Multiple peaks in 25 frames
- No Leave one out validation - Train and test
- No Physical Exercise Data
- Inception model and CNN - Different network, Hyperparameters
- Description of Loss function and training curve!
- Single person only- Whole face
- Experimental variables - MP, Person orientation, Faster movement
- Any face detections? Background effect
- Making sure that phase in contributing for interpolation!
- Benchmark with other methods!

Conclusion

- Idea of using phase difference in face PPG
- Data input construction
- Extraction of HR and HRV
- ECG reconstruction using Spatio-temporal features.

Thank You

Questions!