CorNET: Deep Learning framework for PPG based Heart Rate Estimation and Biometric Identification in Ambulant Environment

Dwaipayan Biswas, Luke Everson, Muqing Liu, Madhuri Panwar, Bram-Ernst Verhoef, Shrishail Patki, Chris H. Kim, Amit Acharyya, Chris Van Hoof, Mario Konijnenburg, Nick Van Helleputte

Abstract— **Advancements in wireless sensor network technologies have enabled the proliferation of miniaturized body-worn sensors, capable of long-term pervasive biomedical signal monitoring. Remote cardiovascular monitoring has been one of the beneficiaries of this development, resulting in non-invasive, photoplethysmography (PPG) sensors being used in ambulatory settings. Wrist-worn PPG, although a popular alternative to electrocardiogram (ECG), suffers from motion artifacts inherent in daily life. Hence, in this paper, we present a novel deep learning framework (CorNET) to efficiently estimate heart rate (HR) information and perform biometric identification (BId) using only wrist-worn, single-channel PPG signal collected in ambulant environment. We have formulated a completely personalized data-driven approach, using a four-layer deep neural network. Two convolution neural network layers are used in conjunction with two long short-term memory layers, followed by a dense output layer for modelling the temporal sequence inherent within the pulsatile signal representative of cardiac activity. The final dense layer is customized with respect to the application, functioning as: a) regression layer - having a single neuron to predict HR; b) classification layer - two neurons which identifies a subject among a group. The proposed network was evaluated on the TROIKA dataset having 22 PPG records collected during various physical activities. We achieve a mean absolute error of 1.47±3.37 BPM for HR estimation and an average accuracy of 96% for BId on 20 subjects. CorNET was further evaluated successfully in an ambulant use-case scenario with custom sensors for two subjects.**

Index Terms— **Average heart rate, biometric, PPG, deep learning, convolutional neural network, long short-term memory.**

I. INTRODUCTION

CARDIOVASCULAR monitoring using wearable sensors has primarily focused on processing electrocardiography (ECG) for key applications – heart rate (*HR*) monitoring, disease detection/prognosis, sports, biometric identification (*BId*), etc. [1-5]. ECG is limited in terms of its placement (requires ground and reference sensors proximal to chest) for signal fidelity, making it inefficient in terms of wearability in ambulant daily living conditions. PPG sensors provide the distinct advantage of having a small form factor and incur low-cost, making them a popular alternative for embedding onto wearable devices (e.g. smart watch) [6-8]. PPG signals are obtained from pulse oximeters, emitting light (using light emitting diode) on the skin and measuring (using photodiode) the miniature variations in reflected or transmitted light intensity. The periodicity of light corresponds to the cardiac rhythm, often used for HR estimation. PPG sensors provide a convenient solution as they can be acquired from peripheral positions such as earlobes, fingertips or wrist, with the latter considered favorable for unobtrusive daily usage. However, data collected from a wrist-worn device is vulnerable to motion artifacts (MA), which correspondingly distorts the signal fidelity, inhibiting the robust estimation of vital parameters [9].

MA are caused by several factors – physical activity, ambient light leaking through the widening gap between sensor and the skin surface during motion and change in blood flow due to movements. This causes the spectral component of the MA to overpower the heart-beat related PPG component [10]. A host of signal processing techniques have been proposed to remove/attenuate MA using adaptive filtering [11], Kalman filtering [12], wiener filtering [13], independent component analysis [14], empirical mode decomposition (EMD) [15], spectral subtraction [10] and feature-engineering based learning algorithms [16-17]. Such methods have often used a motion reference signal from an external sensor (e.g. accelerometer), for detecting and removing MA resulting from motion. The majority of this research was propelled by the IEEE Signal Processing Cup (SPC) 2015. It focused on HR estimation from wrist-worn, two green light illuminated PPG channels, collected during vigorous physical activities [10]. Although successful, the reported improvements in performance are usually accompanied by heuristic thresholds or a large number of expertly-tuned free parameters which could prevent generalization of the developed methodologies.

In this paper, we propose a learning-based framework,

CorNET, for reliable *HR* estimation, using only a single channel wrist-worn PPG signal, collected in ambulant conditions. The proposed framework is based on the fundamentals of Deep Neural Network (DNN). DNN obviates the requirement for feature engineering (extraction/selection of hand-crafted features necessitating domain knowledge), and has been successful in a wide range of applications. Our exploration has been further motivated by the success of a DNN based approach [18], reported recently for *BId* using one-dimensional PPG signals, collected in ambulant environment. A data-driven methodology, *BiometricNET* [18], has been successfully used to model the underlying temporal sequence within the pulsatile cardiac signal based on a convolution neural network (CNN) with long and short term memory (LSTM).

The contributions in this paper are built upon [18], reporting a generic framework, *CorNET*, comprising of two-layer CNN, two-layer LSTM and a final dense layer (DL). The work in [18] has been extended on two aspects:

• new target application, i.e. HR monitoring and
• validation on 22 SPC PPG records and 2 custom records.

This necessitated customizing the DL for two key applications, functioning as – a) regression layer, having a single neuron to estimate *HR* and b) classification layer, having two neurons for *BId*. The framework operating on a single green channel PPG signal, was evaluated on the SPC database. It achieves a mean absolute error of 1.47±3.37 beats per minute (BPM) for *HR* estimation on 22 PPG records collected during various physical activities and 96% accuracy for *BId* on 20 subjects (larger cohort compared to [18]). We further successfully evaluate *CorNET* in an ambulant daily living scenario, using bespoke sensor platforms, on two subjects having different skin types, over a longer monitoring duration. The paper is structured as follows. Section II describes the motivation and problem formulation using a learning-based approach. The proposed methodology highlighting the DNN fundamentals and the developed network architecture, *CorNET*, have been detailed in section III. The results, comparison with state-of-the-art approaches and complexity analysis for *CorNET* have been presented in section IV. The paper has been concluded and future research avenues have been discussed in section V.

II. MOTIVATION AND PROBLEM FORMULATION

State-of-the-art research using wrist-PPG has primarily focused on using green light [10], having a shorter wavelength (in comparison to red/infra-red), as the illuminating source. It provides a distinct advantage of producing large intensity variations to cardiac modulation and yields a better signal-to-noise ratio (SNR) [19]. Moreover, reflective system, with LED and photodetector (PD) on the same side is the preferred mode (in comparison to transmitive) due to user comfort [20]. An illustration of the effect of MA on *HR* estimation has been shown in Fig. 1, with an example PPG segment collected during walking and transition from walking to running. The spectral information, reflected in the peak due to MA is higher compared to the *HR* peak (encircled in bottom plot). It is evident from Fig. 1, that the widely used periodogram, computed

through FFT, has high variance in spectrum estimation and suffers from leakage effects. A TROIKA framework [10], based on signal decomposition, sparsity-based spectral estimation and peak tracking, was successful in estimating HR every 2 seconds (s) from 8s of MA-affected PPG windows, reporting 2.34 BPM error on 12 subjects (SPC). It was followed by an improvement, formulating a multiple measurement vector model, computing the spectra of PPG and acceleration jointly, reporting 1.28 BPM on the same dataset [21]. Recently, an approach based on Wiener filtering and phase vocoder (WFPV) has produced comparative results with 1.02 and 1.97 BPM on 12 and 23 SPC PPG recordings respectively [13].
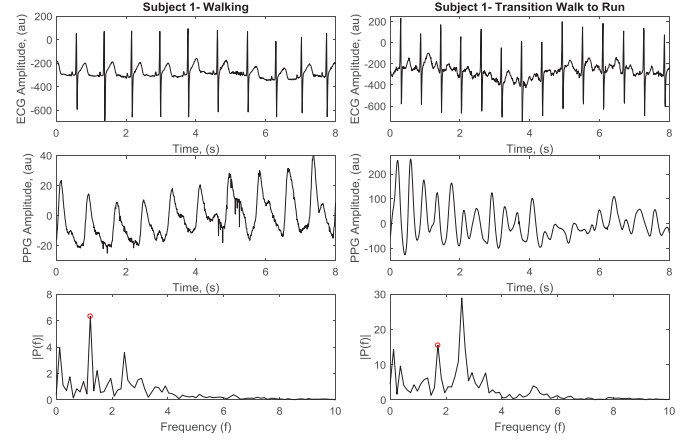

Fig. 1. Raw ECG, PPG signal and spectrum while walking (left) and during transition from walking to running (right) respectively. The highest PPG spectral peak does not coincide with true HR (encircled) during intense motion.

Further developments based on short-time Fourier transform [22], adaptive normalized least mean square (NLMS) filters [23], singular value decomposition [24] and wavelet decomposition [25], using both PPG channels and/or accelerometer data (as motion reference) have been successfully used. A learning algorithm (Random Forest) in conjunction with features, were used to detect beat vs inter-beat samples, allowing HR estimation with 2.86% classification error [16]. Another recent approach has used probabilistic methods, feature extraction and a 3-layer multi-layer perceptron with 22 neurons, reporting 2.81 BPM on 23 PPG recordings of SPC [17]. The algorithms based on spectral processing, involve heuristic thresholds and custom post-processing steps, whereas the learning algorithms have relied on feature engineering.

We propose to use an ECG-assisted supervised framework for *HR* estimation from PPG data. During a dedicated training phase, the relationship between each PPG window (frame) and the *HR* computed from corresponding ECG frame (considered as ground truth) are learnt. Once trained, the model predicts *HR* for a test PPG dataset. The predictions are verified against ground truth *HR* and the error magnitude is averaged over the number of observations and reported in BPM. We formulate a subject-specific (personalized) use-case scenario as opposed to a subject-independent (generalized) validation for our proposed framework. The basic premise of this use-case rests on the fact that biological signals are influenced by various physiological factors – age, sex, height, weight/body-mass-index, etc. and

most importantly the cardiac condition of a given subject. The unified *CorNet* framework (network layers, filter sizes) is trained for each subject, ensuring proportional representation of training examples, which is envisaged for a real-life use-case scenario as illustrated in Fig. 2. We aim to collect baseline data (PPG, ECG) for a given subject, incorporating all possible variations for better generalization, use it to *train* a cross-validated model and employ it for *inference* during daily living.
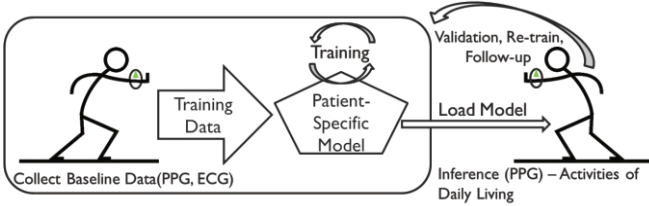

Fig. 2. Personalized use-case evaluation scenario for *CorNET* framework.

For this exploration, we evaluate our algorithm on two databases: a) IEEE SPC 2015 and b) IMEC database (*IMEC-Db*). We use the SPC database in view of its popularity for state-of-the-art research. It comprises PPG signals of 5-minute duration, from 20 healthy subjects, age ranging 18 to 58 [10]. Each subject's data contains simultaneous recordings of - two channels of PPG from the wrist (dorsal) using a pulse oximeter with green LED (wavelength: 515 nm); tri-axial accelerometer signals from the wrist, and a channel of ECG from the chest using wet ECG electrodes. The ECG signal acts as the ground-truth for PPG-based *HR* estimation. All signals were sampled at 125 Hz and transmitted to a computer through Bluetooth. PPG window (frame) lengths considered for this exploration was 8s (sliding by 2s), like ECG-HR computation. The subjects performed three types of activities. First, Type1 (T1), performed by subjects 1-12, involving walking or running on a treadmill with the following speeds in order: 1–2 km/h for 0.5 min, 6–8 km/h for 1 min, 12–15 km/h for 1 min, 6–8 km/h for 1 min, 12–15 km/h for 1 min, and 1–2 km/h for 0.5 min. The subjects used their hand (with wristband) to pull clothes, wipe sweat on forehead, and push buttons on the treadmill. Second, Type2 (T2), performed by subjects 14, 15, 18 and 20, involved in forearm/upper arm exercises (e.g. shake hands, stretch, push, running, jump, and push-ups). Last, Type3 (T3), performed by subjects 15, 16, 17, 18 and 19, involving intense arm movements (e.g. boxing). Hence, we have 20 subjects and 22 records in total, since subjects 15 and 18 were involved in both T2 and T3.

The *IMEC-Db* serves as an experimental evaluation platform, collected in complete ambulatory settings with bespoke sensor platforms. Experiments were performed at IMEC, Belgium, with a customized wristband (since to the best of knowledge, most commercial versions do not provide access to raw PPG data, quintessential for research) and a custom 3-channel ECG patch (IMEC AFE) [26]. The PPG-based wrist module uses an Osram SFH 7060 LED [27] ('green' and 'infra-red') and PD module with TIAFE 4403 [28] as the front end. Both sensing platforms, shown in Fig. 3, are IMEC research prototypes and suited our requirement for a *proof-of-concept* evaluation of *CorNET*. The ECG and PPG signals were

sampled at 256 Hz and saved on a SD card and an eMMC respectively, on these devices. The *IMEC-Db* comprises data collected from two subjects (different ethnic backgrounds), male/female, age ranging 25-35, with their informed consent, in natural office environment, where they were encouraged to carry on with their daily activities. Both subjects started with an initial rest period of five minutes which helps to synchronize the PPG and ECG data streams. The synchronization was based on the accelerometers (sensitivity: ±2g, sampling at 32 Hz), present in both sensor modules. Activities performed generally varied from sitting and working with a computer (keyboard typing, mouse handling, etc.), having coffee, washroom use, taking stairs up/down three floors (several times), talking and performing hand gestures. All activities were performed in a completely voluntary manner, ensuring variability within the data. A naturalistic-ambulant experiment protocol was selected, since it best captures situations where PPG signal quality is affected by micro/fine-grain motion performed within a working environment. Such movements (e.g. finger tapping, pronation/supination, with wrist in relatively stable position) are difficult to be detected by an accelerometer integrated in a wristband. The wristband module performs continuous automatic calibration, ensuring optimal PPG signal for a given skin tone. Data was collected from Subject 1 and 2 for approximately two and seven hours respectively. The differences in duration for the two subjects were aimed at observing the effect of training duration on *HR* estimation.

Fig. 3. Customized chest patch supporting two-lead ECG acquisition (left) and wristband with green LED PPG component.

Besides extracting clinical information, biological signals have long been used for automated recognition of individuals, commonly termed as Biometrics by the International Organization for Standardization (ISO) [29]. Common characteristics include fingerprints, iris, DNA, etc. However, signals such as PPG [30], ECG [31] and electroencephalograms (EEG) [32], provide distinct biological characteristics and an insight into the clinical condition, which paves the way for personalization and identification of individuals. Furthermore, such signals can be captured for long durations without manual intervention, enabling continuous authentication systems. Given the form factor and ease of unobtrusive use in daily life, PPG based *BId* has emerged as an alternative to ECG, however majority of PPG-based work have used signals collected in controlled clinical settings which are less prone to MA, hence, rendering these models unsuitable for ambulatory usage.

PPG-based identification using data from finger has used

frequency domain (Fourier) analysis, correlation analysis, learning based methodology employing hand crafted features with a k-nn classifier [33], fuzzy-logic [30][34], and Linear Discriminant Analysis [35], yielding high accuracies between 90-95%. A recent work focused on wrist-worn, 'green' PPG under motion, using a two-stage procedure, involving clustering and learning techniques (Restricted Boltzmann Machines and Deep Belief Networks) [36]. Although promising, the method had various shortcomings such as potential non-unique clustering of subjects, use of hand-crafted features, significant overhead in cluster formation, which have been discussed in [18] and correspondingly a fully data-driven, personalized, deep learning approach has been proposed in [18].

A binary classification (one vs rest) approach is adopted in [18], to identify a subject against a group based on their PPG signal. The errors of the target class are weighed according to the unbalanced class distribution (e.g. the imbalance of 19:1 for a cohort of 20 subjects in SPC) during training, illustrated in Fig. 4, allowing the network to learn the underlying data distribution in contrary to predicting the dominant zero-class (i.e. 'rest'-class). The network learns subject specific features from respective PPG and hence requires training for each new user. Although a disadvantage, since it necessitates subject-specific data collection over an initial period, network training and weight downloading back to a commercial device, a multi-class classification paradigm would not be feasible to implement. This is because subject-specific training examples would not be available during initial training and having a class for each subject, would cause an unbounded number of classes. Motivated by this problem formulation, we adapt the network topology [18] to extend it for both *HR* estimation and *BId* (using only PPG) as outlined in Section III.
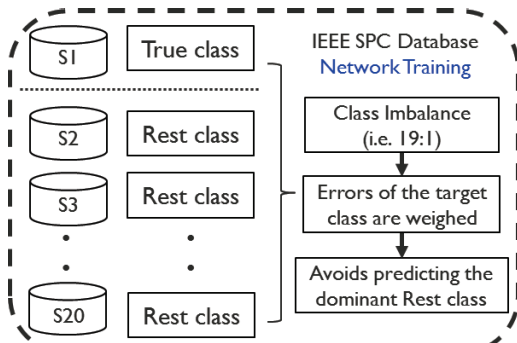

Fig. 4. Binary (one vs rest) classification adopted for *BId* on SPC.

III. PROPOSED FRAMEWORK

An overview of our framework for *HR* estimation and *BId* is shown in Fig. 5, illustrating the ECG-HR assisted training and validation for the former application. The PPG data samples are pre-processed with a band-pass 4th order Butterworth filter having cut-off frequencies 0.1 – 18 Hz. It restricts the high frequency noise component and DC drifts from the signal of interest. The filtered signal is further normalized to zero mean and unit variance.

A. *CNN* and *LSTM*

The taxonomy of DNN primarily includes multi-layer perceptron's (MLP), CNN, Recurrent Neural Networks (RNN). They enable learning of task-adapted feature representations from the data [37]. CNNs are characterized by an initial layer of convolutional filters (set of weights which slides over the input), followed by non-linearity (activation function – rectified linear units), sub-sampling (pooling), and fully connected layer which realizes the classification [38]. The stacking of multiple convolutional layers helps achieving automatic feature extraction, where downstream layers capture more complex or differentiating features. This aids to integrate information from different filters and various levels of abstraction.

RNNs are an effective choice for analyzing time series data for inferring sequential/time-variant information, since they incorporate contextual information from past inputs and are robust to localized distortions in the input sequence along time. A bottleneck for deep CNN structures (with increased network capability) is the vanishing gradient (long chain problem) wherein information from previous computations is rapidly attenuated with progression through the data flow [39]. RNNs applied to long sequential data suffer similarly, as all time steps have the same weight and consequently, the contribution of an input in the hidden state is subjected to exponential decay. LSTM, a variant of RNN, uses a memory block [39] inspired by a computer memory cell, where context-dependent input, output, and forget gates control what is written, read and kept in the cell in each time-step. Hence, it becomes convenient for the network to store a given input over many time steps, in effect helping LSTM layers to capture temporal properties.

B. *CorNET architecture*

The data driven approach is realized by allowing the network to learn discriminatory features, accomplished by using a two layer 1-D CNN, which can be thought of as a feature extractor. The input is convolved with the filters to generate points in the temporal-feature domain. Corresponding layers use these features to convolve with additional filters to generate the final features from the time-series PPG input. One potential drawback of CNN is that the generated features are not completely phase invariant. Depending on the time of occurrence of the heart beat relative to the beginning of the sample, the relevant features will be slightly shifted. Hence, LSTM being instrumental in capturing the temporal dependency amidst the sequence of historical local trends of the underlying cardiac activity in PPG signals, further helps to recover from the phase offset. The output from CNN is fed into two LSTM layers and finally, its output is passed through a DL which is customized for the application. DL consist of a set of weights for each neuron that is multiplied by the input and summed to give the neuron's activation. Since, it is critical to apply the correct activation function to DL, a single neuron with linear transfer function (activation) is used on the LSTM output to estimate a continuous HR value. For *BId*, two neurons are used with SoftMax transfer function to generate a probability estimation between the target and rest classes.
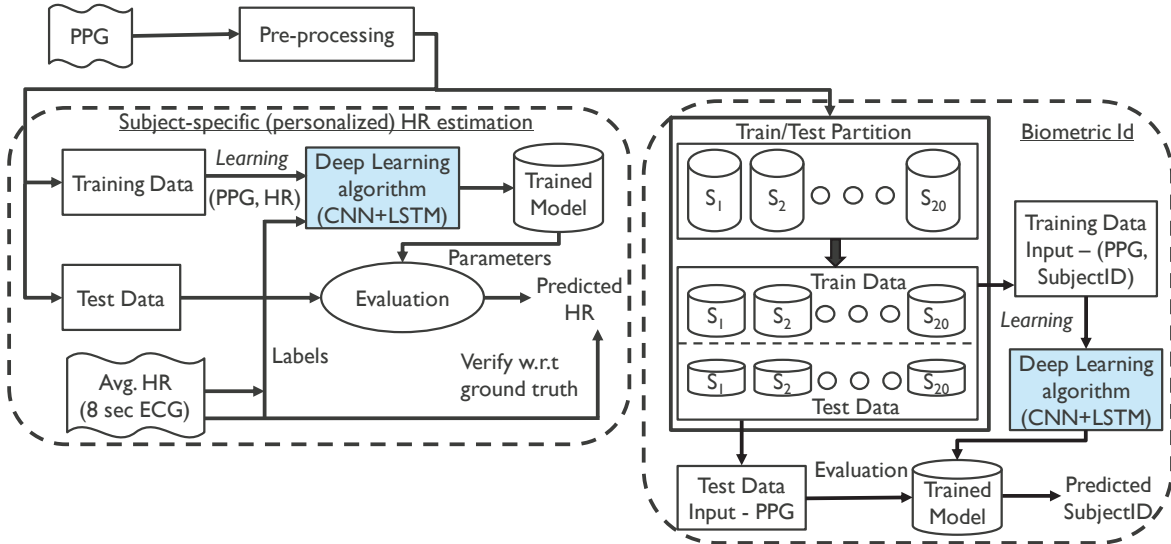
Fig. 5. Overview of the proposed methodology for *HR* estimation and *BId* using a single 'green' channel wrist-worn PPG data collected in ambulant environment.

C. Implementation

CorNET for *HR* estimation and *BId* was trained on an Nvidia Tesla K80 GPU and is modeled in Keras version 2.0.4 [40], configured to use theano [41] version 0.0.9 as the backend. Each CNN layer consists of a 1-D CNN operation with Rectified Linear Unit (ReLU) activation [42], whereas each LSTM layer used hyperbolic tangent (tanh) function. There is a batch normalization layer following the RELU activation. The max-pooling layers used a pool size of 4 and dropout layer with rate 0.1. Root Mean Square Propagation (RMSProp) is used as the optimizer with the default hyperparameters which is recommended for training recurrent networks [43]. For *BId*, as described earlier (section II), the class loss is weighted to offset the class imbalance. Training batch size is set to 25 to balance the training time and sensitivity to individual inputs.
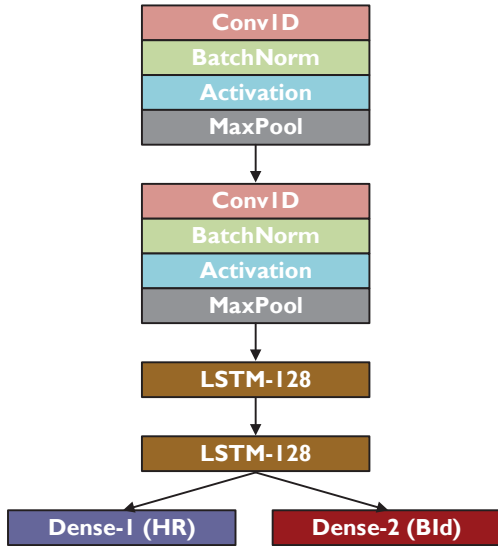

Fig. 6. *CorNET* topology using 2 CNN and 2 LSTM layers with a dense layer having a single neuron for *HR* estimation and two neurons for *BId*.

The proposed *CorNET* topology, illustrated in Fig. 6, uses a filter size of 40 (*sF*) in both CNN layers, with number of filters as 32 (*nF*). A size of 128 units (*nU*) was used for both LSTM

layers working in conjunction with CNN. These hyperparameters were optimized using a heuristic grid search method and selected parameters yielded the best performance. Increasing *sF*, *nF*, *nU* or number of layers (*nL*) beyond the ones selected here did not result in an improved overall performance.

IV. RESULTS

The SPC database comprises temporally interleaved data representing sequence of activities (e.g. T1) and hence we perform, *leave-one-window-out* validation. Here, each window is tested upon and the rest of the windows are considered for training, with a gap of three windows maintained between train-test data, accounting for the 6s overlap. This selection strategy ensures that there is no data overlap between train and test sets and furthermore, the use of max-pooling and dropout layer in conjunction with CNN, helps to avoid overfitting. The performance of *CorNET* has been evaluated with respect to standard metrics [10], which have been briefly described here. We compute absolute error (*AE*), as the absolute difference between true HR_E (from ECG) and estimated HR_P (from PPG), as shown in (1), *i* being the respective window.

$$AE_i = abs(HR_{E_i} - HR_{P_i}) \qquad (1)$$

Correspondingly, the mean absolute error (*MAE*) and the standard deviation of the absolute error (*SDAE*) over all processed windows, are computed and compared with state-of-the-art work (for SPC), as summarized in Table I. Here, the metrics have been computed for the first 12 subjects, performing T1, records 14-23, involved in T2 and T3 and finally the results for all 22 records, enabling a comparative evaluation with [10], [13], [15] and [21]. We achieve a *MAE±SDAE* of 1.99±4.64, 0.86±1.86 and 1.47±3.37 for records 1-12, 14-23 and 1-23 respectively, reflecting the improvement compared to state-of-the-art methodologies. The correlation between HR_E and HR_P over all records is shown in Fig. 7 (a), having a correlation coefficient of 0.998. Furthermore, a comparison of HR_E and HR_P for records 5 and 1 have been

shown in Fig 7 (b) and (c), demonstrating the best and the worst performance of *CorNET* over the first 12 subjects of the experimental cohort. It is interesting to note that *MAE* for records 14-23 is less compared to 1-12, although the former involves random and intense arm movements. On closer observation, it is evident that higher *MAE* (records 1-12) is particularly dominated by record 1 (6.23 BPM), with the initial phases of 1-2, 6-8 and 12-15 km/h, incurring maximal error (cf. Fig. 7c). An activity-wise (T1) analysis of *MAE*, shown in Table I, highlights the relatively high *MAE* for the first three phases, which recovers on the later three phases.
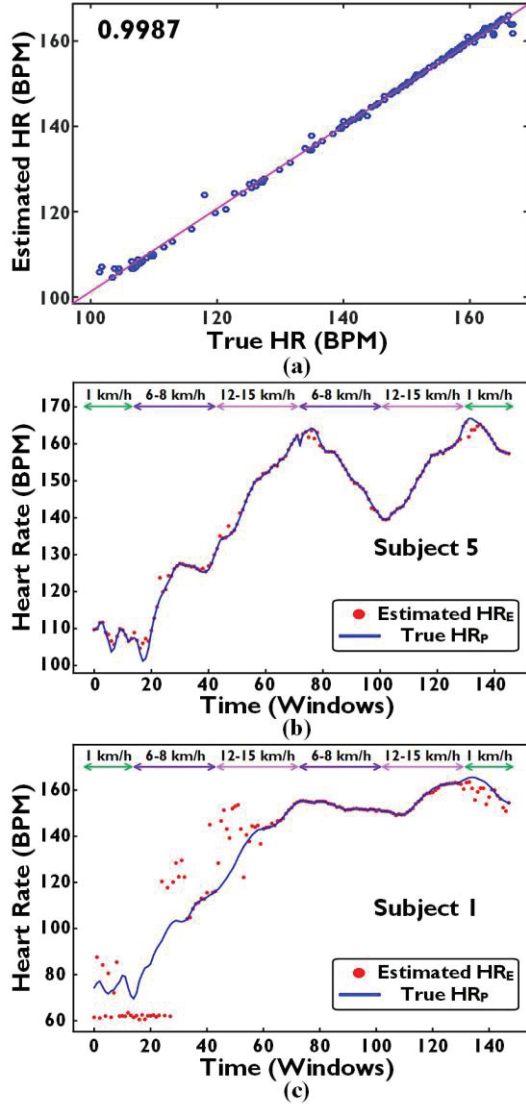

Fig. 7. (a) Pearson correlation coefficient (≈0.998) between the estimated and true HR, estimated vs true HR for Subject 5 (b) and Subject 1 (c) respectively.

The motion artifacts for this subject (record 1) presents difficulties for the model to capture the high variance in the underlying data morphology, especially during the initial phase when the model lacks proportional representative examples to learn from. High variance during initial phase of the data could pose problems for the model to learn and adapt accordingly. Hence it is important to formulate a cross-validated model incorporating maximal variance in the training data, which we have performed on IMEC-Db, described in following section.

The results for *BId*, (evaluated only on SPC) also shown in Table I, represent the average outcomes of 5-fold CV calculated for 20 subjects. The metrics used for evaluation are: a) *Accuracy* - ratio of correctly classified observation (subject to be identified) to the total observations; b) *Precision* - ratio of correctly predicted positive observations to the total predicted positive observations; c) *Recall* - ratio of correctly predicted positive observations to all observations in actual class; d) *F1 Score* - weighted average of precision and recall. We achieve an average accuracy, f1 score, recall, and precision of 0.96, 0.72, 0.86 and 0.67 respectively over 20 subjects. The low f1 score could be attributed towards the class imbalance (1:19) between true and rest classes. It can be observed that subjects 10 and 20 have the best performance in terms of average accuracy and precision. We envisage the PPG based identification to augment traditional identification systems (fingerprint, iris, DNA, etc.) and act as a pre-cursor to these traditional methods in real-life settings.

A. Experimental Evaluation – IMEC Database (IMEC-Db)

The ECG data (collected with chest patch) was processed using a CWT-based beat detection algorithm [44], deemed robust in ambulatory environment, to compute ground-truth HR_E values for validation. The range (minimum and maximum) HR_E values for both subjects varied within $55 - 118$ BPM. The PPG data was windowed in the same manner as SPC, 8s (6s overlap), thereby ensuring a new HR computation every 2s. The PPG windows (each having an associated HR_E), were split into completely independent *training* and *validation datasets* (ratio of 80%-20%). The best model chosen using five-fold cross-validation on the *training dataset* was prospectively evaluated on the *validation dataset*.

The results, in terms of *MAE±SDAE* and *Min-Max* error (BPM) in comparison with WFPV [13] are shown in Table II and a subject-specific plot highlighting HR_E, HR_P and WFPV is shown in Fig. 9. It can be observed that the errors are higher in comparison with SPC results (cf. Table I), which could be attributed to either factors – a) the *IMEC-Db* comprises naturalistic/spontaneous activities performed in an ambulant environment and b) reference ECG and beat detection yielding HR_E, is prone to anomalies in an ambulant environment. However, *CorNET* framework performs better than state-of-the-art WFPV [13] (which performs better in comparison to [21]). Moreover, the trend of HR_P follows that of HR_E and the range of HR_P (Subject1: $58 - 103$ BPM; Subject2: $77 - 109$ BPM) is inbound to that of HR_E (Subject1: $55 - 118$ BPM; Subject2: $72 - 110$ BPM). Lastly the reduced error margin (*MAE±SDAE, Min-Max*) for Subject2, reflects the advantage of a longer training duration which helps the model to incorporate possible variabilities in data distribution and generalize better. The proposed model could be retrained periodically depending on the changing physiology of the subject on a longitudinal scale, and new model parameters hence estimated can be re-used for the desired functionality. Initial results are motivating to drive future research towards longitudinal evaluation and hence we have analyzed the complexity of the *inference* mode aimed at real-time processing in the next sub-section.

TABLE I
PERFORMANCE EVALUATION OF CORNET FOR HR ESTIMATION, BID AND ERROR ANALYSIS FOR SUBJECT1 (SPC)

Metric	Record	Troika [10]	Joss [21]	EEMD [15]	WFPV [13]	CorNET (proposed)	SubjectID	Average (5-CV) Accuracy	F1 Score	Recall	Precision
	1	2.29±2.18	1.33±1.19	1.70	1.25±1.15	6.23±9.44	S1	0.97	0.70	0.71	0.71
	2	2.19±2.37	1.75±1.66	0.84	1.41±1.30	1.83±5.18	S2	0.97	0.74	0.79	0.74
	3	2.00±1.50	1.47±1.27	0.56	0.71±0.59	0.89±3.49	S3	0.98	0.81	0.81	0.83
	4	2.15±2.00	1.48±1.41	1.15	0.97±0.88	0.49±2.29	S4	0.98	0.85	0.92	0.79
	5	2.01±1.22	0.69±0.51	0.77	0.75±0.57	0.40±1.01	S5	0.98	0.81	0.88	0.76
	6	2.76±2.51	1.32±1.09	1.06	0.92±0.75	3.08±6.47	S6	0.96	0.69	0.84	0.63
	7	1.67±1.27	0.71±0.54	0.63	0.65±0.50	1.34±4.42	S7	0.97	0.73	0.94	0.60
	8	1.93±1.47	0.56±0.47	0.53	0.97±0.83	3.64±10.19	S8	0.98	0.78	0.83	0.77
	9	1.86±1.28	0.49±0.41	0.52	0.55±0.48	3.30±6.81	S9	0.98	0.86	0.92	0.83
	10	4.70±2.49	3.81±2.43	2.56	2.06±1.29	1.77±3.96	S10	1.00	0.97	0.99	0.95
	11	1.72±1.29	0.78±0.51	1.05	1.03±0.68	0.41±1.37	S11	0.98	0.83	0.89	0.81
	12	2.84±2.30	1.04±0.81	0.91	0.99±0.70	0.50±1.05	S12	0.98	0.82	0.95	0.72
	13	-	-	-	3.54±4.08	-	S13	0.96	0.47	0.65	0.57
	14	6.63±8.76	8.07±10.9	-	9.59±12.2	1.60±2.29	S14	0.98	0.80	0.96	0.70
	15	1.94±2.56	1.61±2.01	-	2.57±3.16	0.24±0.56	S15	0.82	0.50	0.95	0.34
	16	1.35±1.04	3.10±2.69	-	2.25±1.87	1.60±3.87	S16	0.95	0.61	0.80	0.51
	17	7.82±4.88	7.01±4.49	-	3.01±1.99	2.04±5.02	S17	0.92	0.51	0.79	0.39
	18	2.46±2.00	2.99±2.52	-	2.73±2.29	0.95±3.02	S18	0.90	0.64	0.92	0.51
	19	1.73±1.28	1.67±1.23	-	1.57±1.15	0.28±0.60	S19	0.93	0.46	0.68	0.36
	20	3.33±3.90	2.80±3.46	-	2.10±2.41	0.28±0.65	S20	0.99	0.90	0.92	0.91
	21	3.41±2.43	1.88±1.32	-	3.44±2.45	0.67±1.09					
	22	2.69±2.12	0.92±0.74	-	1.61±1.26	0.42±0.73	**Mean**	0.96	0.72	0.86	0.67
	23	0.51±0.59	0.49±0.57	-	0.75±0.88	0.57 ±0.8					

Record 1-12 (T1)

Metric	Troika [10]	Joss [21]	EEMD [15]	WFPV [13]	CorNET (proposed)
MAE± SDAE	2.34±2.47	1.28±2.61	1.02±1.79	1.02±1.25	1.99±4.64

Record 14-23 (T2 and T3)

Metric	Troika [10]	Joss [21]	EEMD [15]	WFPV [13]	CorNET (proposed)
MAE± SDAE	3.19±3.61	3.05±3.35	-	2.95±3.71	0.86±1.86

Record 1-23 (T1, T2, T3)

Metric	Troika [10]	Joss [21]	EEMD [15]	WFPV [13]	CorNET (proposed)
MAE± SDAE	-	-	-	1.97±2.48	1.47±3.37

Subject1: Activity (T1) - Error Breakup (HR Estimation Analysis)

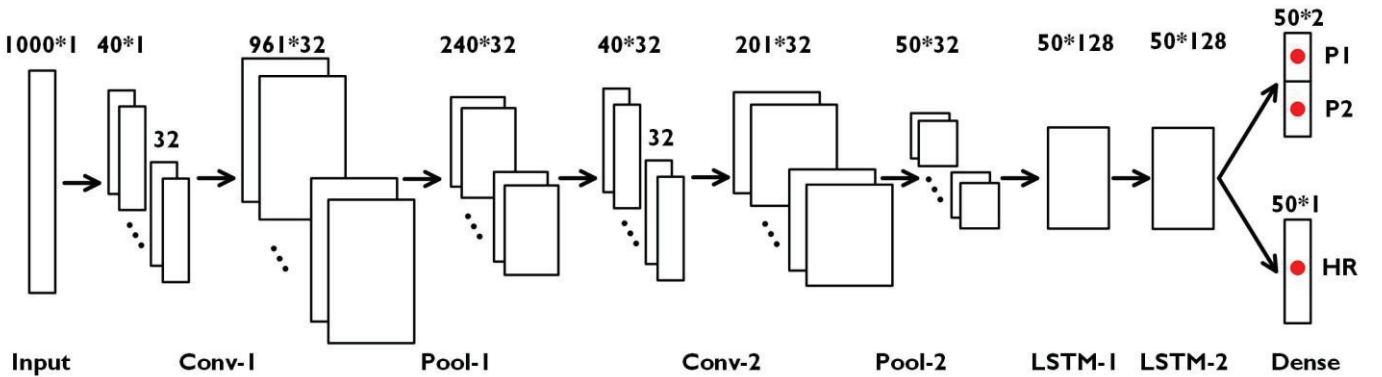| Activity (km/h)	MAE±SDAE
1 - 2	11.27±3.98
6 - 8	4.14±3.35
12 - 15	7.20±10.12
6 - 8	0.07±0.13
12 - 15	1.08±0.55
1 - 2	3.50±3.12
Mean	**6.23±9.44**

Fig. 8. *CorNET* operation, illustrating an input dimensionality of 1000 in conjunction with a filter size (sF) of 40 in CNN layer1 and 2 respectively, each having number of filters (nF) as 32 and 128 LSTM units. P1 and P2 represent the probability of classified subject ID's. The remaining numbers are calculated based on a convolution operation with the input data.

TABLE II
IMEC DATABASE EVALUATION

IMEC-Db	WFPV [13]		CorNET (proposed)	
	MAE±SDAE	Min-Max Error	MAE±SDAE	Min-Max Error
Subject1	10.22±11.33	0.001-86	8.54±7.63	0.02-47
Subject2	14.79±14.09	0.002-67	5.93±4.63	0-28

B. Complexity Analysis

Recent trends in remote cardiovascular disease monitoring, highlights the importance of 'on-node' processing of sensor data. This aids continuous/real-time monitoring, negating continuous data transmission and elongating the life of battery-operated sensor nodes [45]. In such systems, it is quintessential to adopt low-complexity data processing strategies for energy

efficient operation of the sensor nodes. Hence, a complexity analysis of the proposed architecture (cf. Table III), illustrates the number of multiply-and-accumulate (MAC) operations and trainable parameters for each component layer. Efforts to reduce the trainable parameters in conjunction with a reduced sample size of PPG windows, by downsampling (similar to [13]), resulted in a degraded performance, a further testimony to our hyperparameter selection. Current trends in architecture development for DNN hold promise and make the numbers achievable [46-49]. Further leverage could be obtained by adopting an offline-online processing strategy [50], whereby the time and compute intensive training is carried out offline on a GPU and a processor/accelerator performs an online inference on real-time sensor data.
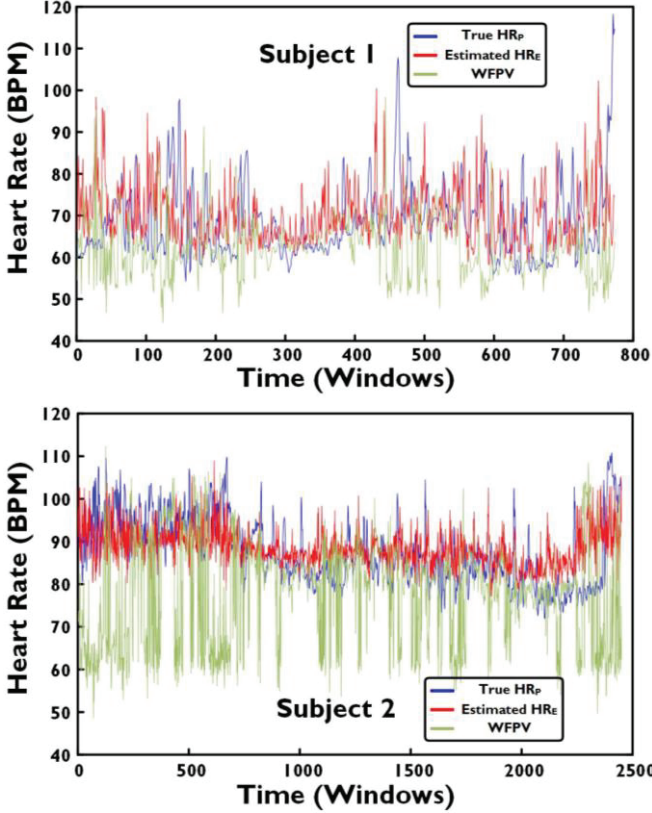

Fig. 9. Illustration of estimated (proposed *CorNET* and WFPV) HR_P and True HR_E for the two subjects of *IMEC-Db*.

TABLE III
CORNET: COMPLEXITY ANALYSIS

Layer	Metadata	MACs	Trainable Parameters
CNN-1	sF=40, nF=32	1.2 M	1312
CNN-2	sF=40, nF=32	8.2 M	40992
LSTM-1	nU=128	4.1 M	82432
LSTM-2	nU=128	6.6 M	131584
Dense-1	n=2	258	258

V.   CONCLUSION

In this paper, we have presented a first of its kind exploration, for performing *HR* estimation and *BId*, reporting a personalized data-driven approach using DNN on wearable PPG signals collected in ambulatory situation. This negates the use of heuristics/thresholds, post-processing steps, auxiliary sensor data and extraction of hand-crafted features. The proposed

CorNET topology, using two layers of CNN and LSTM is customized in conjunction with a dense layer for either of the investigated applications. We achieve an average error of 1.47±3.37 BPM for *HR* on all 22 PPG recordings and an average accuracy of 96% for *BId* on all 20 subjects. The results could be considered favorable since to the best of knowledge, these represent the best accuracy in comparison to state-of-the-art methods on the given application area. We also successfully evaluate *CorNET* on two subjects in an ambulant environment using custom sensors for a longer duration. The present exploration focusses on estimating average HR, predicting a new value every 2s using 8s PPG frame, thereby missing out on instantaneous information. Hence, we would like to extend our investigation towards heart rate variability measures, which could provide insights into functioning of the sympathetic nervous system and help in disease prognosis (e.g. myocardial infarction). Lastly, future research would focus on energy-efficient execution of the algorithm on wearable devices in real-time on a processor or hardware solutions (ASIC) using schemes proposed in [46-47].

REFERENCES

[1] E. B. Mazomenos *et al*., "A Low-Complexity ECG Feature Extraction Algorithm for Mobile Healthcare Applications," in *IEEE Journal of Biomed. and Health Inf.*, vol. 17, no. 2, pp. 459-469, March 2013.

[2] H. Kim *et al*., "A configurable and low-power mixed signal SoC for portable ECG monitoring applications", *IEEE Trans. Biomed. Circuits Syst.*, vol. 8, no. 2, pp. 257-267, Apr. 2014.

[3] H. So and K. Chan, "Development of QRS detection method for real-time ambulatory cardiac monitor," in Engineering in Medicine and Biology Society, 1997. Proceedings of the 19th Annual International Conference of the IEEE, 1997, vol. 1, pp. 289–292.

[4] Y. Zou, J. Han, S. Xuan, S. Huang, X. Weng, D. Fang, and X. Zeng, "An energy-efficient design for ECG recording and R-peak detection based on wavelet transform," IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 62, no. 2, pp. 119–123, 2015.

[5] E. J. da S. Luz, W. R. Schwartz, G. Cámara-Chávez, and D. Menotti, "ECG-based heartbeat classification for arrhythmia detection: A survey," Computer methods and programs in biomedicine, vol. 127, pp. 144–164, 2016.

[6] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," Phys Meas, v. 28, pp. 1-39, 2007.

[7] L. Wang, B. P. Lo, and G.-Z. Yang, "Multichannel reflective PPG earpiece sensor with passive motion cancellation," IEEE transactions on biomedical circuits and systems, vol. 1, no. 4, pp. 235–241, 2007.

[8] J. Lee, D.-H. Jang, S. Park, and S. Cho, "A Low-power Photoplethysmogram-based Heart Rate Sensor using Heart Beat Locked Loop," IEEE transactions on biomedical circuits and systems, 2018.

[9] T. Aoyagi and K. Miyasaka, "Pulse oximetry: its invention, contribution to medicine, and future tasks.," Anesthesia and analgesia, vol. 94, no. 1 Suppl, p. S1, 2002.

[10] Zhang, Zhilin *et al*., "TROIKA: A general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise." *IEEE Trans. Biomed. Eng.*, 62.2 (2015): 522-531.

[11] M. Mashhadi *et al*., "Heart Rate Tracking using Wrist-Type Photoplethysmographic (PPG) Signals during Physical Exercise with Simultaneous Accelerometry", *IEEE Signal Processing Letters*, v. 23, 2016.

[12] B. Lee *et al*., "Improved elimination of motion artifacts from a photoplethysmographic signal using a kalman smoother with simultaneous accelerometry," *Physiol Meas*., vol. 31, no. 12, pp. 1585–1603, 2010.

[13] A. Temko, "Accurate Heart Rate Monitoring During Physical Exercises Using PPG.," *IEEE Trans. Biomed Eng*, vol. 64, no. 9, pp. 2016–2024, 2017.

[14] B. Kim, S. Yoo, "Motion artifact reduction in photoplethysmography using independent component analysis," *IEEE Trans. Biomed. Eng.*, v. 53, pp. 566–568, 2006.

[15] E. Khan *et al.*, "A Robust Heart Rate Monitoring Scheme Using Photoplethysmographic Signals Corrupted by Intense Motion Artifacts," *IEEE Trans. Biomed. Eng.*, v.63, pp. 550-562, 2016.

[16] E. Grisan *et al.*, "A supervised learning approach for the robust detection of heart beat in plethysmographic data," in *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, pp. 5825–5828, 2015.

[17] M. Essalat *et al.*, "Supervised heart rate tracking using wrist-type photoplethysmographic (PPG) signals during physical exercise without simultaneous acceleration signals," in *Signal and Information Processing (GlobalSIP), 2016 IEEE Global Conference on*, pp. 1166–1170, 2016.

[18] L. Everson et al., "BiometricNet: Deep Learning based Biometrie Identification using Wrist-Worn PPG," in *IEEE Int. Symp. Circuits Syst. (ISCAS)*, Florence, Italy, pp. 1-5, 2018.

[19] Y. Meada *et al.*, "The advantage of green reflected photoplethysmograph," *J. Med. Syst.*, vol. 35, pp. 829–834, 2011.

[20] Y. Sun and N. Thakor, "Photoplethysmography revisited: from contact to noncontact, from point to imaging," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 3, pp. 463–477, 2016.

[21] Z. Zhang, "Photoplethysmography-Based Heart Rate Monitoring in Physical Activities via Joint Sparse Spectrum Reconstruction," *IEEE Trans. Biomed. Eng.*, v.62, pp. 1902-1910, 2015.

[22] D. Zhao *et al.*, "Sfst: A robust framework for heart rate monitoring from photoplethysmography signals during physical activities," *Biomedical Signal Processing and Control*, vol. 33, pp. 316–324, 2017.

[23] T. Schäck *et al.*, "A new method for heart rate monitoring during physical exercise using photoplethysmographic signals," *in Signal Processing Conference (EUSIPCO), 2015 23rd European*, pp. 2666–2670, 2015.

[24] M. B. Mashhadi *et al.*, "Heart rate tracking using wrist-type photoplethysmographic (PPG) signals during physical exercise with simultaneous accelerometry," *IEEE Signal Processing Letters*, vol. 23, no. 2, pp. 227–231, 2016.

[25] M. Raghuram *et al.*, "Evaluation of wavelets for reduction of motion artifacts in photoplethysmographic signals," *in Proc. 10th Int. Conf. Inform. Sci. Signal Process. Appl.*, pp. 460–463, 2010.

[26] https://www.imec-int.com/en/Biomedical-AFE [online]

[27] https://www.osram.com/os/ecat/SFH%207060/com/en/class_pim_web_catalog_103489/global/prd_pim_device_2220014/ [online]

[28] http://www.ti.com/product/AFE4403# [online]

[29] A. Bonissi *et al.*, "A preliminary study on continuous authentication methods for photoplethysmographic biometrics", *2013 IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications*, Napoli, pp. 28-33, 2013.

[30] Y. Gu *et al.*, "A novel biometric approach in human verification by photoplethysmographic signals", *4th International IEEE EMBS Special Topic Conference on Information Technology Applications in Biomedicine,* pp. 13-14, 2003.

[31] A. Lourenço *et al.*, "Unveiling the biometric potential of finger-based ecg signals", *Computational intelligence and neuroscience*, p5, 2011.

[32] RB. Paranjape *et al.*, "The electroencephalogram as a biometric", *in Electrical and Computer Engineering, 2001. Canadian Conference on*, IEEE, Vol. 2, pp. 1363-1366, 2001.

[33] A.R. Kavsaoğlu *et al.*, "A novel feature ranking algorithm for biometric recognition with PPG signals." *Computers in biology and medicine, vol. 49*, pp.1-14., 2014.

[34] Y. Gu, Y. Zhang, "Photoplethysmographic authentication through fuzzy logic", *4th International IEEE EMBS Special Topic Conference on Information Technology Applications in Biomedicine,* pp. 136-137, 2003.

[35] P. Spachos *et al.*, "Feasibility study of photo-plethysmographic signals for biometric identification", In *Digital Signal Processing (DSP), 2011 17th International Conference on, IEEE,* pp. 1-5, 2011.

[36] V. Jindal *et al.*, "An adaptive deep learning approach for PPG-based identification." In *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the IEEE.*, pp. 6401-6404, 2016.

[37] Y. Bengio, "Learning deep architectures for AI." Foundations and trends® in Machine Learning 2.1 (2009): 1-127.

[38] Y. LeCun *et al.*, "Gradient-based learning applied to document recognition." *Proceedings of the IEEE*, vol. 86(11), pp. 2278-2324, 1998.

[39] S. Hochreiter and J. Schmidhuber, "Long short-term memory", *Neural computation,* vol. 9(8), pp.1735-1780, 1997.

[40] https://keras.io/ [online]

[41] http://deeplearning.net/software/theano/ [online]

[42] Klambauer, Günter, et al. "Self-Normalizing Neural Networks." *arXiv preprint:1706.02515* (2017).

[43] http://climin.readthedocs.io/en/latest/rmsprop.html [online]

[44] I. Romero *et al.*, "Low-power robust beat detection in ambulatory cardiac monitoring," *in Biomedical Circuits and Systems Conference, BioCAS 2009. IEEE,* pp. 249-252, 2009.

[45] E. De Giovanni *et al.*, "Ultra-low power estimation of heart rate under physical activity using a wearable photoplethysmographic system," *in Proceedings of the 19th IEEE/Euromicro Conference On Digital System Design (DSD 2016)*, vol. 1, pp. 1-10, 2016.

[46] R. Andri *et al.*, "YodaNN: An Architecture for Ultralow Power Binary-Weight CNN Acceleration," *IEEE Trans. on Computer-Aided Design of Int. Circuits and Systems*, 37(1), pp.48-60, 2018.

[47] M. Verhelst and B. Moons, "Embedded Deep Neural Network Processing: Algorithmic and Processor Techniques Bring Deep Learning to IoT and Edge Devices," *IEEE Solid-State Circuits Magazine*, 9(4), pp.55-65, 2017.

[48] M. Panwar *et al.*, "Modified Distributed Arithmetic Based Low Complexity CNN Architecture Design Methodology", *in Circuit Theory and Design (ECCTD), 2017 European Conference on*. IEEE, pp. 1-4, 2017.

[49] L. R. Everson *et al*, "A 104.8TOPS/W One-Shot Time-Based Neuromorphic Chip Employing Dynamic Threshold Error Correction in 65nm," *2018 IEEE Asian Solid-State Circuits Conference (A-SSCC)*, Tainan, Taiwan, 2018, pp. 273-276.

[50] D. Biswas *et al.*, "Low-Complexity Framework for Movement Classification Using Body-Worn Sensors," *IEEE Trans. on Very Large Scale Int. (VLSI) Sys.*, vol. 25(4), pp.1537-1548, 2017.

Dwaipayan Biswas received his MSc in System on Chip, 2011 and his Ph.D in Electrical Engineering, 2015, from University of Southampton (UoS), UK. He went on to work as a post-doctoral research fellow at UoS, 2015 – 2016. His area of research includes low-power VLSI design, biomedical signal processing, machine learning, brain computer interface and computer architecture. On November, 2016, he joined IMEC, Belgium, as a researcher on digital IC design for biomedical applications. He has authored 9 journals, over 20 conference publications and three book chapters.

Luke Everson received the BSEE and MSEE from the University of Minnesota, Minneapolis, MN, USA in 2015 and 2016 respectively. He is currently pursuing the Ph.D. at UMN in the VLSI Research Lab. He is currently working on developing time-based architectures for machine learning and biosignal recording applications. During the summer of 2017 he was a Visiting International Scholar at imec, Belgium, where he applied machine learning to biomedical signals.
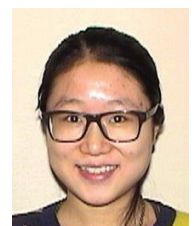
Muqing Liu received the B.S. degree in applied physics from Tongji University, Shanghai, China in 2013 and the M.S. degree in electrical engineering from Columbia University, New York, NY, USA in 2015. She is currently pursuing the Ph.D. degree in electrical engineering at

University of Minnesota, Minneapolis, MN, USA. She joined the VLSI Research Laboratory at University of Minnesota in 2015, focusing on hardware implementation of neural networks and hardware security, such as neuromorphic circuits design, physical unclonable functions (PUFs) and counterfeit electronic sensors design.

Madhuri Panwar received integrated degree in VLSI from Jayoti Vidhyapeeth Women's University, Jaipur, India in 2014. Currently, she is pursuing her Ph.D. from Indian Institute of Technology, Hyderabad. Her research interest includes low-complexity deep learning algorithm and circuit design for biomedical applications.

Bram-Ernst Verhoef received a master in theoretical psychology and a master of science in statistics from KU Leuven, Belgium. He obtained his Ph.D. in biomedical sciences from the same institute in 2010. Following his Ph.D., he performed post-doctoral research on both experimental and computational neuroscience; first at Harvard Medical School (Boston, MA, USA; till 2014), and subsequently at the University of Chicago (Chicago, IL, USA; till 2017). In 2017 he joined IMEC, Belgium, as an R&D engineer, developing deep learning algorithms for neuromorphic chips.

Shrishail Patki received his bachelor degree in Instrumentation Engineering from University of Mumbai and MS degree from University of Miami in Biomedical Engineering. He spent one and half years working at Bio-Signal Group Corp where he worked on EEG system development . In 2010, he joined imec-nl/Holst Centre, where he is currently working on developing different generations of wearables for acquisition of biopotential and physiological signals focusing on miniaturized and wireless system design that is optimized for low noise and low power.

Chris H. Kim received the B.S. and M.S. degrees from Seoul National University, Seoul, South Korea, and the Ph.D. degree from Purdue University, Lafayette, IN, USA. He was with Intel Corporation, Hillsboro, OR, USA, where he performed research on variation-tolerant circuits, on-die leakage sensor design, and crosstalk noise analysis. He joined the Electrical and Computer Engineering Faculty, University of Minnesota, Minneapolis, MN, USA, in 2004, where he is currently a Professor. He has authored or co-authored over 200 journal and conference papers. His current research interests include digital mixed-signal, and memory circuit design in silicon and non-silicon (organic TFT and spin) technologies.

Dr. Kim was a recipient of the SRC Technical Excellence Award, the Council of Graduate Students Outstanding Faculty Award, the NSF CAREER Award, a Mcknight Foundation Land-Grant Professorship, the 3M Non-tenured Faculty Award, DAC/ ISSCC Student Design Contest Awards, the IBM Faculty Partnership Awards, the IEEE Circuits and Systems Society Outstanding Young Author Award, and the ISLPED Low Power Design Contest Awards.

Amit Acharyya (M11) received the Ph.D. degree in 2011 from the School of Electronics and Computer Science in the University of Southampton, UK. Currently he is working as an Associate Professor in the Indian Institute of Technology (IIT), Hyderabad, India. His research interests include signal processing algorithms, VLSI architectures, low power design techniques, computer arithmetic, numerical analysis, linear algebra, bio-informatics and electronic aspects of pervasive computing.

Chris Van Hoof received the Ph.D. degree in electrical engineering from the University of Leuven, Belgium, in 1992. He is Director of Wearable Healthcare at imec and is also imec Fellow. He has a track record of over 25 years of initiating, executing, and leading national and international contract R&D at imec. His work resulted in three startups (two in the healthcare domain) and he has delivered sensor flight hardware to two cornerstone European Space Agency missions. After receiving his Ph.D. from the University of Leuven in 1992 in collaboration with imec, he has held positions at imec as manager and director in diverse technical fields (sensors and imagers, MEMS and autonomous microsystems, wireless sensors, and body-area networks, wearable health). He has published over 600 papers in journals and conference proceedings and given over 60 invited talks. He is also full professor at the University of Leuven (KULeuven).

Mario Konijnenburg (M'08) received the M.S. degree in electrical engineering and the Ph.D. degree from the Delft University of Technology, Delft, The Netherlands, in 1993 and 1999, respectively. While working on the Ph.D. degree, he wrote the paper, Automatic Test Pattern Generation for Sequential Circuits. He joined Philips Research/NXP Semiconductors and worked on methodology and tool development to improve testability of designs and ease Design for Test (DFT). Currently, he is a System Architect and R&D Engineer at Holst Centre/imec, Eindhoven, The Netherlands, where he works on development of ultra-low-power (wireless sensor node) designs for mainly biomedical applications. His research interests are design for low-power and DFT in DSPs.

Nick Van Helleputte received the MS degree in electrical engineering in 2004 from the Katholieke Universiteit Leuven, Belgium. He received his Ph.D. degree from the same institute in 2009 (MICAS research group). His PhD research focused on low-power ultra-wide-band analog front-end receivers for ranging applications. He joined imec in 2009 as an Analog R&D Design Engineer. He is currently team leader of the biomedical circuits and systems team. His research focus is on ultra-low-power circuits for biomedical applications. He has been involved in analog and mixed-signal ASIC design for wearable and implantable healthcare applications. Nick is an IEEE member and served on the technical program committee of VLSI circuits symposium and ISSCC.