Comprehensive Analysis of Cardiogenic Vibrations for Automated Detection of Atrial Fibrillation Using Smartphone Mechanocardiograms

Mojtaba Jafari Tadi, Saeed Mehrang, Matti Kaisti, Olli Lahdenoja, Tero Hurnanen, Jussi Jaakkola, Samuli Jaakkola, Tuija Vasankari, Tuomas Kiviniemi, Juhani Airaksinen, Timo Knuutila, Eero Lehtonen, Tero Koivisto, and Mikko Pänkäälä

***Abstract*— Atrial fibrillation (AFib) is the most common sustained heart arrhythmia and is characterized by irregular and excessively frequent ventricular contractions. Early diagnosis of AFib is a key step in the prevention of stroke and heart failure. In this paper, we present a comprehensive time–frequency pattern analysis approach for automated detection of AFib from smartphone-derived seismocardiography (SCG) and gyrocardiography (GCG) signals. We sought to assess the diagnostic performance of a smartphone mechanocardiogram (MCG) by considering joint SCG–GCG recordings from 435 subjects including 190 AFib and 245 sinus rhythm cases. A fully automated AFib detection algorithm consisting of various signal processing and multidisciplinary feature engineering techniques was developed and evaluated through a large set of cross-validation (CV) data including 300 (AFib = 150) cardiac patients. The trained model was further tested on an unseen set of recordings including 135 (AFib = 40) subjects considered as cross-database (CD). The experimental results showed accuracy, sensitivity, and specificity of approximately 97%, 99%, and 95% for the CV study and up to 95%, 93%, and 97% for the CD test, respectively. The F_1 scores were 97% and 96% for the CV and CD, respectively. A positive predictive value of approximately 95% and 92% was obtained, respectively, for the validation and test sets suggesting high reproducibility and repeatability for mobile AFib detection. Moreover, the kappa coefficient of the method was 0.94 indicating a near-perfect agreement in rhythm classification between the smartphone algorithm and visual interpretation of telemetry recordings. The results support the feasibility of self-monitoring via easy-to-use and accessible MCGs.**

***Index Terms*— Atrial fibrillation, seismocardiography, gyrocardiography, smartphone mechanocardiography, machine learning.**

Manuscript received September 30, 2018; revised November 15, 2018; accepted November 16, 2018. Date of publication November 26, 2018; date of current version February 15, 2019. This work was supported in part by the Academy of Finland under Grant 290930 and in part by Business Finland (former Finnish Funding Agency for Innovation) under Grant 2496/31/2014. The work of M. Jafari Tadi was supported in part by the Academy of Finland through NoStrokes Project under Grant 290930, in part by the Nokia Foundation Awards under Grants 201710299 and 201810316, and in part by Tekniikan Edistämissäätiön (Finnish Foundation for Technology Promotion). The associate editor coordinating the review of this paper and approving it for publication was Prof. Aime Lay-Ekuakille. *(Corresponding author: Mojtaba Jafari Tadi.)*

M. Jafari Tadi is with the Department of Future Technologies, University of Turku, 20014 Turku, Finland, and also with the Department of Biomedicine, University of Turku, 20014 Turku, Finland (e-mail: mojjaf@utu.fi).

S. Mehrang, M. Kaisti, O. Lahdenoja, T. Hurnanen, T. Knuutila, E. Lehtonen, T. Koivisto, and M. Pänkäälä are with the Department of Future Technologies, University of Turku, 20014 Turku, Finland.

J. Jaakkola, S. Jaakkola, T. Vasankari, T. Kiviniemi, and J. Airaksinen are with the Heart Center, Turku University Central Hospital, 20521 Turku, Finland.

Digital Object Identifier 10.1109/JSEN.2018.2882874

I. Introduction

ATRIAL fibrillation (AFib) is a prevalent heart arrhythmia which is characterized by random uncoordinated atrial activations (supraventricular tachyarrhythmia) resulting failure of atrial mechanical function [1]. Prevalence of AFib increases with aging and approximately occures in 3% of the adults aged 20 years or older [2]. Despite the clinical importance of this problem, AFib detection remains challenging, because it can appear episodic and asymptomatic ("silent atrial fibrillation") in many cases before it manifests itself with hemodynamic impairment and thromboembolic events. Silent AFib can potentially lead to formation of blood clot(s), heart failure, acute strokes, and other heart-related complications, if it is not diagnosed and treated in time. Undiagnosed AFib is linked to considerable morbidity, mortality, and financial costs [1].

Standard diagnostic tools such as electrocardiography (ECG) [1], [3] and echocardiography [4] have been widely used for AFib detection. In addition to the standard methods, phonocardiography (PCG) [5], and photoplethysmography (PPG) [6] have also been investigated recently. Handheld ECG devices, mobile phones [7], smart watches [8], weighing scales [9], smart wristbands [10], earlobe sensors [11], and web-camera sensing [12] have been introduced as new screening techniques which aim to streamline and enhance personal health monitoring by enabling individuals to monitor themselves. AFib is, however, known to remain sporadic and absent in some of medical examinations requiring long-term continuous monitoring for an in time and accurate diagnosis [3]. In order to detect symptoms appearing at periodic or random intervals, a capability for longer-term monitoring is needed, e.g. for several days or weeks at a time.

Readily available and clinically validated mobile/wearable devices such as AliveCore Kardia [13], Zenicor EKG [14], Mydiagnostick [15], Zio Patch [16], Medtronic implantable loop recorders and wearable SEEQ mobile cardiac telemetry [17], iPhone optical sensor [7], Kardia Band [18], and Samsung Simband smartwatch [19] have been recently presented offering possibilities for long term monitoring. However, these methods, most of which are ECG-based, always require additional hardware which might be inaccessible for some users. PPG is an alternative technique to detect AFib with a smartphone (by placing a fingertip on the camera) or a smartwatch optical sensor [19], while no additional hardware is needed.

However, PPG recordings measure peripheral capillary oxygen saturation (SpO2) in the blood from which heart rate may be derived, but is susceptible to variations in ambient light and skin contact. PPG recorders, although easy to use, do not give an immediate feedback due to the heart's electromechanical coupling (atrial and ventricular excitation-contraction) as bio-impedance/bio-mechanical generated signals provide [20].

Sensor-rich smartphones are today available for most people worldwide. Almost all modern smartphones contain accurate accelerometers and gyroscopes to detect axis- and angular orientation, and subsequently, they can be used to measure precordial vibrations on the surface of the chest wall skin with high resolution. These precordial vibrations originate from complex myocardial movements traveling throughout the surface of the chest wall. Seismocardiography (SCG) [21] and gyrocardiography (GCG) [22] techniques jointly constitute the concept of mechanocardiography (MCG) by which cardiogenic signals can be collected by placing the smartphone on the chest of the individuals while they are in supine/seating position without any additional hardware requirement. SCG is an accelerometer-based method for measuring translational precordial vibrations, while GCG is a method based on angular rate sensors for measuring rotational precordial activities. AFib detection based on the analysis of smartphone-derived SCG and GCG signals has been introduced recently and it has yielded encouraging results that support the feasibilty of self-monitoring [23]. Prior to this, conventional ballisto-cardiography (BCG) and SCG methods have been used for AFib detection [24], [25]. Much of the SCG/BCG research has focused on evaluating the methods based on relatively small number of subjects outside clinical environments. Recent findings in SCG/BCG/GCG further support the hypothesis, that mechanical signals have potential in cardiac performance assessment, e.g. related to heart arrhythmia, myocardial ischemia, cardiac resynchronization therapy, and heart failure [22], [26].

In this study, we consider a large dataset (n = 435 subjects) and utilize various feature engineering techniques to comprehensively analyze cardiogenic vibration patterns in detecting and quantifying AFib. To the best of our knowledge, this is the first time that this number of subjects is studied using only smartphone MCG for arrythmia detection. Moreover, we consider a wide spectrum of time-frequency features that adequately describe mechanical characteristics of the heart during arrythmia and normal rhythm. In [23], we demonstrated a proof-of-concept machine learning pipeline for AFib detection via smartphone accelerometer and gyroscopes, and recently, in a blinded offline analysis, a previously developed (knowlegde-based) algorithm similar to [25] was employed to the SCG-GCG recordings to classify the rhythm as either AFib or SR among 300 patients [27]. During our previous works, we have either assumed that 1) the control group consists of healthy young male adults and the AFib group consists of hospitalized (older) patients (which can be vulnerable to classification bias) [23] or 2) both the control group and the AFib group consist of hospitalized (and thus older) patients with other heart diseases as well [27]. In this paper, our goal is to extend our smartphone AFib detection framework towards efficient operation among the full adult population range, and enhance classification performance of machine learning models by employing multidisciplinary signal processing and machine learning techniques that effectively contribute in AFib screening.

II. Materials and Methods

A. Human Study Subjects

In this study a sample of size 435 subjects was recruited in two sets (cross-validation and cross-database) as follows:

1) Cross-Validation Dataset: Of the total number of subjects, 300 patients (150 of which were labeled as AFib) were considered as the derivation population from MODE-AF study (ClinicalTrials.gov Identifier: NCT03274583) [27]. These patients were hospitalized for cardiac indications. A continuous 5-lead telemetry ECG (Philips IntelliVue MX40) recording was acquired simultaneously with the MCG recordings and was used as the comparison method to assess cardiac rhythm and the number of possible supraventricular (SVES) or ventricular extrasystoles (VES). The rhythm of each telemetry ECG was labeled as sinus rhythm (SR), AFib, or "other" by two independent cardiologists. In cases of inconsistency in labeling of the two cardiologists, a third independent cardiologist also blinded to the original rhythm classification made the final decision. Medical history of the subjects was investigated by inspecting the electronic patient records. A full description of the patients' clinical characteristics is given in Table I.

2) Cross-Database Dataset: Another subset with a sample size of 135 including 95 healthy adults with SR and 40 elders with positive AFib arrhythmia (labeled by cardiologists) was considered as validation population. Informed consent was collected from all subjects. Demographic information of the patient group (n = 40) can be found in our previous contribution [23]. With the SR group, smartphone measurements were obtained from young adults in an opportunistic data collection fashion – without collecting demographic information – at the University of Turku campus. Pre-study investigations on this group of young healthy subjects revealed substantial heart rate variabilities, e.g. due to respiratory sinus arrhythmia, which motivated us to consider this as a challenge for differentiating AFib from SR. Helsinki declaration was strictly followed in all phases and data collections were made with the permission of Ethical Committee of the Hospital District of South-Western Finland.

B. Data Acquisition and Measurement Protocol

All the clinical data recordings including subjects' enrolment, data acquisition, and data documentation were made by a (single) study investigator. Following consent collection, background data was gathered (e.g. blood pressure, pulse, chest circumference, height, and weight) as reported in Table I. Afterwards, a recording of three minutes was obtained using a Sony Xperia Z1 or Z5 smartphone and custom designed Android application. A sampling frequency of 200 Hz was configured for smartphone accelerometer and gyroscope data acquisition and only one device was used for each recording.

TABLE I

CLINICAL CHARACTERISTICS OF THE SUBJECTS WITH SR OR AFIB DURING THE MCG RECORDINGS

Clinical variable	SR (n=150)	AFib (n=150)	P-Value
Age, y	74.5 (73.0-76.1)	75.0 (73.5-76.6)	0.660
Female sex	66 (44.0)	66 (44.0)	1.000
Chest circumference, cm	103 (101-104)	105 (103-107)	0.043
BMI, kg/m2	27.5 (26.7-28.2)	29.0 (28.1-29.9)	0.013
CHA_2DS_2VASc	3.7 (3.5-4.0)	4.1 (3.8-4.4)	0.049
History of heart failure	24 (16.0)	73 (48.7)	<0.001
Hypertension	101 (67.3)	103 (68.7)	0.902
Diabetes	27 (18.0)	40 (26.7)	0.096
History of ischemic stroke	12 (8.0)	10 (12.7)	0.255
Coronary artery disease	90 (60.0)	54 (36.0)	<0.001
History of AMI	68 (45.3)	26 (17.3)	<0.001
Peripheral artery disease	7 (4.7)	13 (8.7)	0.247
Prior CABG	16 (10.7)	23 (15.3)	0.303
Pulmonary disease	24 (16.0)	23 (15.3)	1.000
History of AFib	30 (20.0)	150 (100)	<0.001
LVEF			0.021
<40%	13 (40.6)	19 (59.4)	
40-49%	16 (44.4)	20 (55.6)	
50%	75 (63.6)	43 (36.4)	
Heart rate, beats/min	70.8 (68.7-73.0)	88.0 (84.6-91.4)	<0.001
Respiratory rate	16.9 (16.2-17.5)	19.2 (18.2-20.2)	<0.001
Systolic blood pressure, mmHg	145 (141-149)	137 (134-141)	0.003
Diastolic blood pressure, mmHg	70.0 (67.7-72.2)	80.9 (77.7-84.1)	<0.001
LBBB configuration in ECG	15 (10.0)	12 (8.0)	0.687
RBBB configuration in ECG	10 (6.7)	10 (6.7)	0.597
Edema in Chest X-ray	38 (37.6)	67 (54.5)	0.015
ProBNP (pg/mL)	1095 [3596]	2965 [5603]	<0.001
Patients with SVES during MCG	47 (31.3)	0 (0.0)	<0.001
Patients with VES during MCG	52 (34.7)	73 (49.0)	0.014

*AMI indicates acute myocardial infarction; BMI, body mass index; CABG, coronary artery bypass grafting; CHA_2DS_2VASc, congestive heart failure, hypertension, 75 years of age and older, diabetes mellitus, previous stroke or transient ischemic attack, vascular disease, 65 to 74 years of age, female; ECG, electrocardiogram; LBBB, left bundle branch block; LVEF, left ventricular ejection fraction; proBNP, pro-brain natriuretic peptide; RBBB, right bundle branch block; SVES, supraventricular extrasystolia; VES, ventricular extrasystolia Values are presented as number (%), median [interquartile range] or mean (95% confidence interval)

The smartphone was placed (without any chest belt/strap) on the subject's chest longitudinally while the screen was facing upwards and the bottom edge of the phone at the level of the lower edge of the body of sternum. Chest micro-motions in x-axis corresponds to the right-to-left lateral movements, y-axis to the head-to-foot aligned movements, and z-axis to the dorso-ventral movements caused by precordial vibrations. During the recording, the subjects were advised to stay calm, silent, and motionless. When three minutes had passed from the initiation of the recording the developed application automatically stopped the recording.

III. MACHINE LEARNING PIPELINE

Fig.1 shows the general schematic of our machine learning pipeline including pre-processing, feature engineering, feature matrix, and classification methods. We perform a preliminary signal pre-processing by first segmenting the signals from each of the six axes *(AccX, AccY, . . ., GyroZ)* into certain lengths, then extract various features and build three feature matrices, and finally classify AFib from SR subjects. Pre-processing, feature extraction, and classification were performed with Matlab Version 2017.

A. Data Pre-Processing

1) Singular Spectrum Analysis (SSA): The frequency spectrum of noise may not be known a priori. Deploying conventional frequency domain filters — that require setting the pass-band frequency limits a priori — is likely to cause loss of vital information. A noise removal technique that decomposes the signal into its favorable and unwanted components is therefore needed. Singular spectrum analysis (SSA) is an algorithm of time series analysis which consists of two complementary stages: decomposition and reconstruction [28]. Decomposition stage includes two steps, i) Embedding the original time series of length N into a sequence of lagged vectors of size L by forming $K = N - L + 1$ principal components, where L is the window length ($1 < L < N$), and creating a trajectory matrix A (Hankel matrix), and ii) Singular value decomposition (SVD) of the trajectory matrix A by setting $S = AA^T$ and delivering it as a sum of rank-one bi-orthogonal elementary matrices. Reconstruction stage consists of two steps as well: i) Eigentriple grouping by splitting the elementary matrices A_i into m disjoint subsets $I_1, \ldots, I_m$ and summing the matrices within each group, and ii) Diagonal averaging by transforming each matrix A_I of the grouped decomposition into a new series of length N.

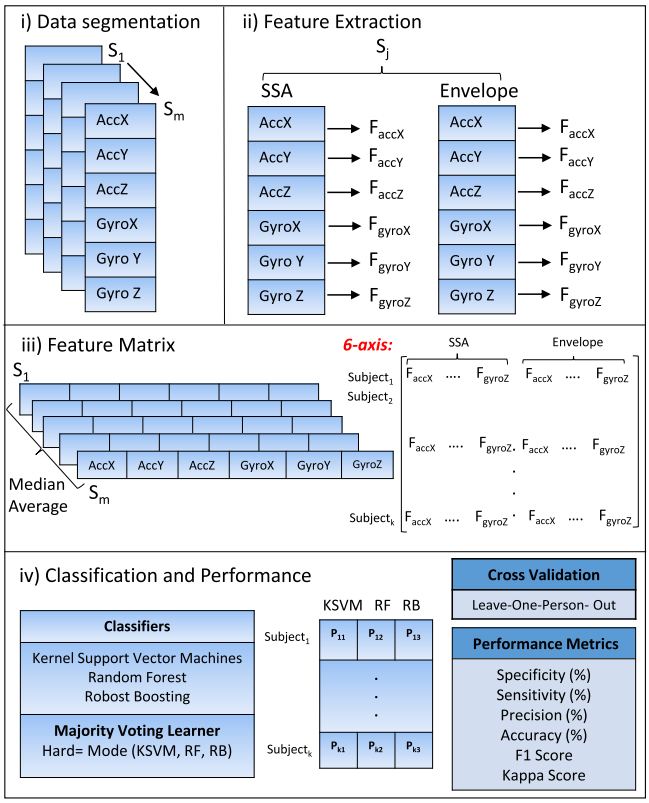

Fig. 1. General schematic of machine learning pipeline.

2) Envelope Detection: In addition to the SSA filter, we consider the pulse amplitude by computing the envelope of the signals. This envelope detection algorithm operates based on a moving-average filtering as described in the following. To apply this filter to each data point, we construct our coefficients so that each point is equally weighted and contributes 1/4 to the total average. Subsequently, we apply

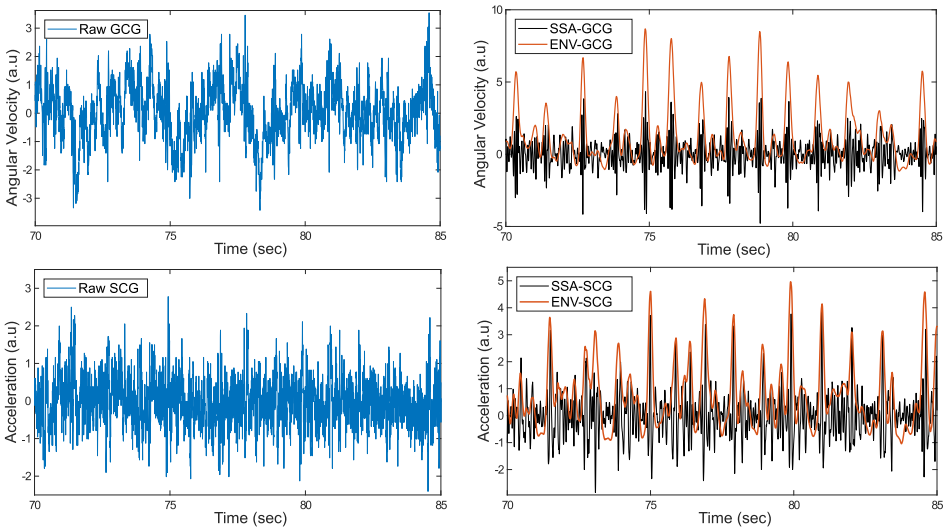

Fig. 2. An example of AFib signal where SSA (black line) and envelope detection filters (red line) are applied on poor quality raw SCG (Z-axis) and GCG (Y-axis) signals (blue line). Dominant heartbeats are visible after SSA filtering.

a moving-average filter with a triangle shaped window and length of 8 samples to the accelerometer and gyroscope signals. The resulting signal is filtered with another moving average filter with triangle shaped window of the length of 51 samples, equal to 250 ms. The final signal envelope forms reproducible cardiac pulsatile waveforms that contains dominant physiological characteristics stable enough to assess heart rhythm. Fig. 2 shows an example of a noisy AFib signal where SSA and envelope detection are applied on GCG and SCG signals. It is apparent from this figure that SSA filter has significantly improved the signal quality by removing noise and baseline wander in SCG and GCG signals. Moreover, the signal envelope provides a pulse-shape waveform that shows the cardiac rhythm and its variations vigorously.

3) Signal Segmentation: The feature extraction starts by dividing individual axes/channels of each sensor into a sequence of discrete segments in order to reveal the underlying properties of cardiogenic signals. Segments were obtained with overlapping windows (75%) of lengths 10, 20, 30, 40, 50, and 60 seconds. The aim of this experiment was to find out the best window length for AFib detection for which the best detection accuracy is obtained. This allows us to realize which window length is most suitable to stationarize the mechanical cardiac signals. With stationarization, we aim at preserving the statistical properties (e.g. mean, signal deviation, etc.) of these segments that are almost constant over time. Moreover, the overlapping segmentation, although demands more processing power, can result in a smoother change of the parameter values between the frames, e.g. processing in frequency domain analysis will be more robust to noise when frames in each motion axis is processed in parallel. A Hamming window with the same size as the segmented frame was applied to each frame to facilitate the spectral analysis via fast Fourier transform (FFT), e.g. in terms of spectral leakage, side lobe amplitude, and width of the central peak [29].

B. Feature Engineering

We consider a variety of time- and frequency-domain features for AFib detection with smartphone SCG-GCG. Our method benefits from various feature extraction techniques which have been previously used for different analysis applications such as speech processing and pattern recognition. We also deploy features such as heart rate variability (HRV) and associated parameters to measure the complexity and randomness of the SCG-GCG signals. In the following, we explain segment by segment analysis of smartphone-derived SCG and GCG signals and describe selected features that contribute in the classification of AFib.

1) Zero Crossing Rate: Zero-crossing rate (ZCR) for a signal segment s of length N is defined as

$$ZCR_t = \frac{1}{2N}\sum_{n=1}^{N} |sign[s(n)] - sign[s(n-1)]|, \qquad (1)$$

where $sgn[s(n)] = 1$ if the signal has a positive amplitude and -1 for negative amplitudes in frame t. ZCR provides a measure of the noisiness/randomness of the signal.

2) Energy Entropy of the Segmented Signal: The energy of the segmented discrete-time signal is defined as

$$E_s = \sum_{n=1}^{N} s(n)^2. \qquad (2)$$

Entropy of the signal energy, Ent_e is a measure of irregularity of energy a single segment carries. In our considerations, the Ent_e parameter is calculated by sub-segmenting a frame into 10 shorter intervals to compute normalized sub-frame energies, and entropy of the energies as

$$P_i = \frac{subsegments_i^2}{(E_s)}, \qquad (3)$$

where P_i is the probability of energy segment in the i^{th} sub-segments (i = 1 to 10) and E_s is the total energy of

the segment. The entropy of the energy Ent_e is then defined as

$$Ent_e = -\sum P_i \times \log(P_i). \tag{4}$$

3) Spectral Centroid and Spread: Spectral centroid calculates the *center of mass* of the signal spectrum. The individual centroid of a spectral frame is calculated as the mean frequency weighted by frequency amplitudes and divided by total sum of the amplitudes, that is:

$$SP_{centroid} = \frac{\sum_{k=1}^{K}(M(k) \times F[k])}{\sum_{k=1}^{K} F[k]}, \tag{5}$$

where $M(k)$ is the sample range up to the center frequency ($F_c = 100$ Hz) and $F[k]$ is the frequency amplitude (Fourier transform of a frame s_n or $|FFT(s_n)|$ with a bin size of $K = 1000$) of the kth spectrum bin (with a frequency resolution $(F_c/K) = 0.1$ Hz). Spectral spread gives the mean variation of the rate-map (computed for individual frequency bin) around its centroid, and is typically dependent on the bandwidth of the signal. It is defined as

$$SP_{spread} = \frac{\sqrt{\sum_{k=1}^{K}(M(k) - SP_{centroid})^2 \times F[k]}}{\sum_{k=1}^{K} F[k]}. \tag{6}$$

4) Spectral Flux and Roll-Off: Spectral flux measures how fast the power spectrum changes in a signal (temporal variation) and is calculated by comparing the power spectrum of one frame against the power spectrum of the previous frame as

$$SP_{flux} = \sum_{n=2}^{Nseg}(FFT(s_n) - FFT(s_n - 1))^2, \tag{7}$$

where $Nseg$ is the total number of segments and $FFT(s_n)$ is the Fourier transform of n^{th} segment.

Spectral roll-off is another measure indicating the amount of the right-skewness of the power spectrum, and is obtained by the fraction of the bins in the power spectrum at which certain (C_p) percentage of the total power is at lower frequencies [30]. In this study the single-sided power spectral density (PSD) was estimated using the Welch's spectral density computation method.

$$argmin\ f_R \in \{1, \ldots, K\} \sum_{k=1}^{f_R} PS_n \geq C_p \sum_{k=1}^{K} PS_n. \tag{8}$$

Eq. 8 computes the spectral roll-off where k is the frequency bin index, PS_k, is the corresponding spectral magnitude, f_R is the spectral roll-off frequency and K is the total number of frequency bins. The spectral roll-off is defined as the frequency, f_R, that a specified proportion, C_p, of the total spectral energy is accumulated [30].

5) Fundamental Frequency and Heart Rate: The spectrum of a signal can be harmonic or non-harmonic. In a harmonic signal the envelope of the frequency spectrum may feature a comb-structure where identifying first peak gives the fundamental frequency. This feature is commonly obtained by autocorrelation of the signal spectrum and finding the first side peak. The fundamental frequency (F_0) of the segment can be estimated by finding the first side peak of autocorrelation signal as explained in [31].

The F_0 is the lowest frequency of a periodic waveform and in the case of SCG-GCG it literally reflects the heart rate (HR) in time domain, i.e. $HR = 60 \times F_0$. In order to calculate instantaneous heart rates in a segment, we consider short-term autocorrelation (S-AC) algorithm which is able to analyze the periodicity of a signal [25]. To achieve this, we compute the autocorrelation $R(s(j), i)$ of the segment as

$$R(s(j), i) = \sum s(j)s(j + i), \tag{9}$$

where j is a discrete variable denoting the frequency steps, and indices $j + i$ up to the length of segment(N) are taken into account. The fundamental frequency (F_0) of the segment can be estimated by finding the first side peak of $R(s(j), i)$. This is performed by localizing the index of the first side peak

$$i_{\text{first peak}} = \arg_i \max(R(s(j), i). \tag{10}$$

We calculate the S-AC for each episode/segment by dividing each into sub-segments with the duration of 2.5 second. The algorithm can then return cardiac cycles — the total number of cycles depends on the length of the segment — from one single segment.

6) Harmonic-to-Noise Ratio: Harmonic-to-Noise ratio (HNR) is a measure of the overall periodicity of the segmented signal obtained by quantifying the ratio between the harmonic part (considered as periodic) and non-harmonic part (considered as aperiodic or noise components). Assuming a signal as a mixture of a harmonic signal and a noise signal, we can decompose it as:

$$s[t] = h[t] + n[t], \tag{11}$$

where $s[t]$ is the signal, $H[t]$ is a harmonic component, and $N[t]$ is a noise component. In real situation, especially in the case of biological signals, the amplitudes and the fundamental frequency are slowly changing over time, unless some phenomena like AFib happens. To take this into account, we segmented the signals into small enough time frames that we may assume to be quasi-stationary. Subsequently, we calculate the HNR as the ratio of the component energies (in frequency domain):

$$HNR = \frac{\sum_{t=0}^{N-1} n[t]^2}{\sum_{t=0}^{N-1} n[t]^2}, \tag{12}$$

where $h[t]$ represents harmonic amplitudes, $n[t]$ is the noise estimate and N is the total number of sample points in each segment) [32].

7) Spectral Peaks: The power spectral analysis of the segmented signal reveals the relation of frequency spectrum to time series. Welch's power spectral density (PSD) estimate is used to investigate the spectral contents of the segmented signal. Subsequently, the position and amplitude of six largest peaks in the density spectrum of the signal were localized. These features, in total 12 contributors (6 amplitude and 6 corresponding frequency value), are used to describe the distribution of power.

8) Spectral Power Features: In addition to the *Spectral Peak* features, we computed the total spectral power of the signal frame in different frequency bands as [0.5-1.5] Hz, [1.5-5] Hz, [5-10] Hz, [10-15] Hz, and [15-20] Hz in both accelerometer and gyroscope segmented signals independently for each sensor channel/axis. A periodogram function – also known as a nonparametric estimate of the PSD – was used to calculate spectral contents (in total 5 features) of the signal segments.

9) Median Amplitude Spectrum: In continuation of the spectral analysis, we calculate the median amplitude spectrum (MAS) by dividing each signal segment into G small subframes or bins (each of length $L = 100$ ms) and calculating MAS in each bin by FFT. MAS is a frequency at which the signal power spectrum is divided into two regions with equal amplitude and as a half of the total power (dividing the total power area into two equal parts) [33]. The definition of MAS is given by

$$P_g(f) = \frac{1}{2L}\sum_{f=1}^{L-1} |FFT[s_g]|, \tag{13}$$

where $P_g(f)$ is a vector of total amplitude spectrums for g^{th} subframe/bins, accordingly MAS is defined as

$$MAS = Median(P_g). \tag{14}$$

10) Covariance Features: A further characterization of the cardiac signals derived from inertial sensors includes autocorrelation features that are defined as the *height of main peak*, *position*, and *amplitude* of the first side peak. We calculate autocorrelation in time-domain and localize the corresponding main peak and first side peak.

11) Approximate Entropy and Sample Entropy: Approximate entropy (ApEn) is a well-known approach in computing the complexity of time series. $ApEn$ is a self-similarity parameter that quantifies the unpredictability of fluctuations in a time series [34], [35]. With calculating $ApEn$, time series containing regular patterns such as sinus rhythm are expected to have a relatively small $ApEn$, while a less predictable or irregular pattern signal (e.g. $AFib$) should have a higher $ApEn$ index [34]. Sample entropy (SampEn) is also a modified version of ApEn which is considered to assess the complexity or dynamics of physiological time-series. SampEn is calculated as the the negative natural logarithm of the conditional probability that a short episode of data is repeated during the time series [49].

To estimate the $ApEn$ parameter, we perform a two-step all-against-all matching procedure for each frame. In the first step, we compute the number of matches between the sub-segment s_i (with the length N) against all other sub-segment s_j within the same signal ($i \neq j$). The number of matches is updated in counter C where $\max(|s_i - s_j|)$ is less or equal than 0.2 (tolerance) times the overall standard deviation (SD) of the signal. We count similar patterns using distance calculation (with an embedded dimension $ED = 1$ and a mode level $M = EM + iterationsize - 1$) between the samples as

$$\Phi_M^{ED} = \frac{\sum \Phi_M^{ED} + log(C)}{N}, \tag{15}$$

where Φ_M^{ED} is a logarithmic counter updated at each instance of the subsampled signal. In the second iteration, similar procedure is executed for all-against-all matching test except that the number of samples, i.e. M, is incremented by one. Finally, the $ApEn$ is computed from Φ_1^{ED}- Φ_2^{ED}.

12) Spectral Entropy: Spectral entropy (SpEn) has been known as a measure of uncertainty or randomness of a time series. SpEn is a tool to quantify the spectral complexity of a signal [36]. SpEn relies on PSD analyses $P(f)$ which aims to show the distribution of power as a function of frequency [36].

The power spectrum is computed by first taking FFT and then squaring to obtain an approximation of PSD $P(f)$, where f is limited to the frequency band [5-80 Hz]. In order to minimize the effect of frequent low frequency components, an estimated noise floor is discarded from $P(f)$ by filtering frequencies with energy amplitude $P(f) < \frac{1}{6} \times \max(P(f))$. The resulting spectrum is then normalized to unit probability as

$$P_t(f) = P(f) / \sum_f P(f). \tag{16}$$

This normalization leads the frequency spectrum to a probability distribution prior to computing of spectral entropy. Finally, the spectral entropy of the signal $P_t(f)$ is computed as

$$SpEn = -\sum_f P_t(f) \times \log(P_t(f)). \tag{17}$$

The cardiogenic signals containing more random frequency components, e.g. $AFib$, are expected to yield greater $SpEn$ than samples with less randomness, e.g. sinus rhythm [25].

13) Shannon Entropy: Shannon Energy Entropy (SEE) is obtained by calculating the energy for each sample point(n) $\neq$ 0 of the original signal. Therefore, the SEE of each segment is defined as

$$SEE_j = -\sum_{n=1}^{N} [s_j(n)]^2 \times \log[s_j(n)^2], \tag{18}$$

where $s_j(n)^2$ is the energy of the n^{th} sample in j^{th} segment.

14) Turning Point Ratios: Turning point ratio (TPR) is another non-parametric statistical approach to determine the randomness of the signals. We define operator $RD(s_j)$ to calculate the total number of consecutive increasing and decreasing runs in signal segment s_j. TPR of s_j is therefore defined as

$$TPR(s) = \frac{RD(s_j)}{N-2}, \tag{19}$$

where N is the segment length. We consider TPR in the input signal segment s denoted as $TPR(s)$. In general, the number of runs in the input signal indicates waveform complexity, which can be considered as an indicator of AFib behavior in that segment.

15) Heart Rate Variability Features: Heart rate variability (HRV), the changes in the beat-to-beat heart rate, is a key indicator of the cardiac rhythm status [37]. We estimate cardiac time intervals using a previously developed method based on autocorrelation [25] and subsequently employ three HRV parameters serving as potential features for differentiating

AFib from SR. We consider the root mean square of the successive difference of NN interval durations (RMSSD) which is the most frequently used measure extracted from interval differences as

$$RMSSD = \sqrt{(\frac{\sum_{n=1}^{NP}(PP_j - PP_{j-1})^2}{NP})}, \quad (20)$$

where NP is the total number of estimated cardiac cycles in a segment and PP is the successive peak to peak distance or interbeat time interval. Furthermore, we calculate the median difference of successive beat-to-beat intervals (HRV_{median}) obtained in a segment as

$$HRV_{median} = \text{Median}(|PP_j - PP_{j-1}|). \quad (21)$$

The spectral density of the the beat-to-beat variations (HRV_{power}) is further computed as a potential feature for AFib analysis.

16) Higuchi Fractal Dimensions: Higuchi fractal dimension (HFD) is a measure of signal complexity which is often used for analyzing non-periodic or irregular time series [38]. The HFD is calculated on each signal segment as follows. Given a finite set of time series, e.g. biological samples, $s[1], s[2], \ldots, s[N]$, we construct a new time series S_k^m as:

$$S_k^m = \{s[m], s[m+k], \ldots, s[m + \lfloor\frac{N-m}{k}\rfloor k]\}, \quad (22)$$

for $m = 1, 2, \ldots, k$ where m is initial time; k is the time interval, $k = 1, \ldots, k_{max}$; k_{max} is a free parameter and $\lfloor \frac{N-m}{k} \rfloor$ calculates the integer part. The length of the curve L_k^m is defined as

$$L_m^k = \frac{1}{k}[\sum_{i=1}^{\frac{N-m}{k}} |s[m+ik] - s[m+(i-1)k]|.\frac{N-1}{\lfloor(N-m)/k\rfloor}], \quad (23)$$

where N is the length of original segment/frame and $\frac{N-1}{\lfloor(N-m)/k\rfloor}]$ is a normalization factor. L_m^k is the average value over k sets of L_m^k as

$$L(k) = \frac{\sum_{m=1}^{K} L_m^k}{k}. \quad (24)$$

A vector of average values $L(k)$ is obtained and the HFD can be estimated as the slope of least squares linear best fit from the plot of $\ln(L(k))$ against $\ln(1/k)$

$$HFD = \frac{\ln(L(k))}{\ln(1/k)}. \quad (25)$$

17) Hjorth Parameters: Hjorth parameters are quantitative features in description of statistical characteristics in time series [39]. Three parameters are defined in Hjorth method as:

i) *Activity*, which is quantified by the variance of the amplitude (squared standard deviation of the signal) representing signal power as

$$Activity = var(s(t)). \quad (26)$$

ii) *Mobility*, which measures the standard deviation of the slope with reference to the standard deviation of the amplitude in a time series. In other words, it is the square root of variance of the first derivative of the signal s(t) divided by variance of the signal s(t) as

$$Mobility = \sqrt{\frac{var(\frac{d}{dt}s(t))}{var(s(t))}}. \quad (27)$$

iii) *Complexity*, which represents changes in frequency domain by measuring the signal's similarity to a pure sine wave as

$$Complexity = \frac{Mobility(\frac{d}{dt}s(t))}{Mobility(s(t))}, \quad (28)$$

where *complexity* value converges to 1 if the signal is sinusoidal.

18) Kurtosis and Skewness: Other specifications of the cardiac signals can be obtained by *skewness* and *kurtosis*. Skewness is measure of symmetry or lack of symmetry in distribution of the sample points in a time frame. A distribution is symmetric if samples (e.g. in a histogram) are equally distributed to the left and right of the center point. Skewness in normal distribution is zero.

Kurtosis is a measure of the tailedness of the probability distribution and considers whether distribution of sample points is heavy-tailed or light-tailed with respect to a Gaussian normal distribution. Standard normal distribution has a kurtosis of zero, while positive and negative kurtosis values refer to heavy-tailed and light-tailed distributions, respectively.

19) Mean and RMS Level: The *mean value* as well as the *root-mean-square (RMS)* level of the segmented signal are also considered as potential features in the characterization of cardiogenic signals.

Fig. 3 presents a concise view of the time-frequency characteristics of the smartphone derived signals during AFib and SR episodes.

C. Feature Matrix Generation

We employ described features presented in Section III-B in a binary classifier setting to distinguish between AFib and SR. A total of 46 features are extracted from the segmented subsignals throughout the accelerometer and gyroscope sensor data. We extract these features from both SSA-filtered and signals' envelops separately, and subsequently incorporate them into a vector to build our feature matrix. Additionally, we consider feature matrices in three different formats: i) 6-axis combination of SCG-GCG features ($2 \times 6 \times 46 \times ns$), ii) 3-axis SCG features ($2 \times 3 \times 46 \times ns$), and iii) 3-axis GCG features ($2 \times 3 \times 46 \times ns$), where ns is the number of segments (ns changes depending on the segment length which varies from 10s to 60s) and number 2 refers to SSA plus envelope attribute-fusion. This arrangement allows us to assess the capability of attributes in the detection of AFib, when combined from both sensors (i) against the case where they are independently considered (ii and iii).

As described in Sec III-A.3, we segment signals with 6 different windowing lengths, i.e. 10, 20,…,60s, each with 75% overlap and subsequently construct a set of rows and columns, each representing a segment-wise array of features for each channel. Since each sensor's channel is divided into

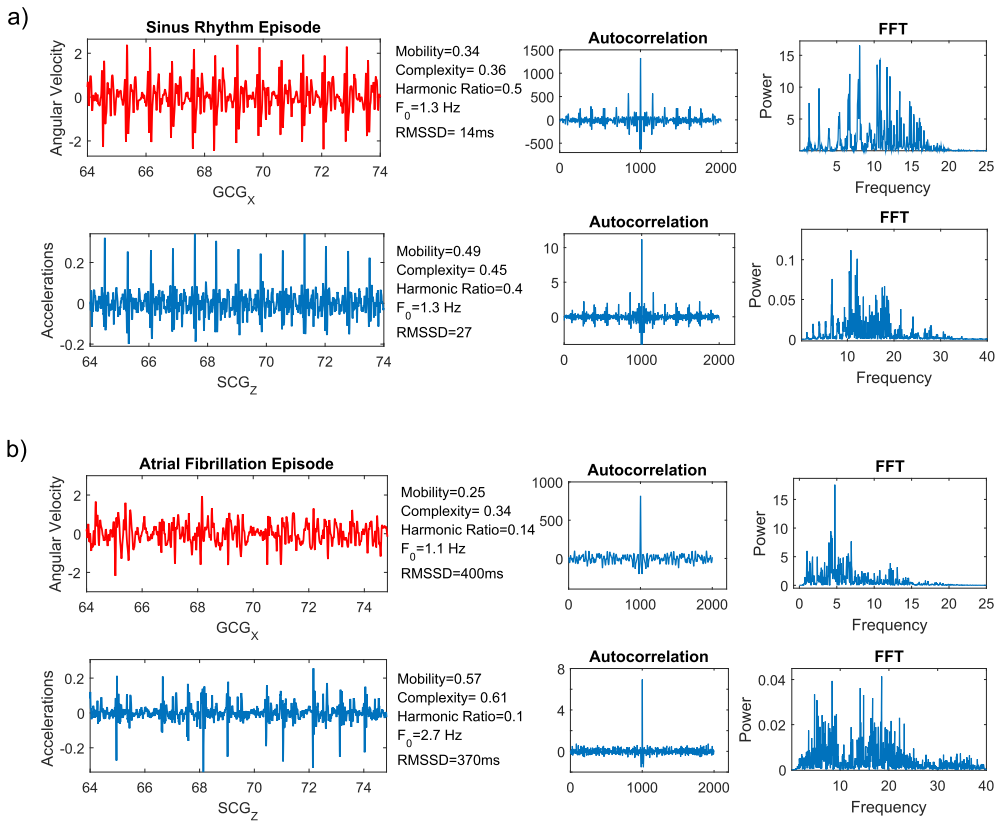

Fig. 3. Spectral and temporal characteristics of SCG and GCG signals during SR (panel a) and AFib (panel b) episodes.

variable number of segments — and in order to minimize the effect of noise and movement artifacts — we calculate the *median* value of each feature (column-wise) across all the segments [40]. With *median* averaging we aim at reshaping the feature matrix to a single vector per subject. Assuming we split cardiogenic signals into p number of segments per channel, there will be p rows and 46 columns in a feature matrix. A column-wise *median* will therefore result in a vector consisting of only 46 averaged features per channel (including both SSA and envelope). Later, we horizontally concatenate the resulting vector arrays to build a single feature vector for each subject. Finally, our feature matrix includes, for example, k rows (each corresponds to a study subject). See Fig. 1 part ii and iii in which feature matrix construction process is illustrated.

D. Classification and Performance Metrics

Three different learning algorithms were tested by their capability in differentiating AFib and SR using the previously explained features. We trained a support vector machine (SVM) with radian basis function kernels, a random forest (RF) classifier based on Bootstrap-aggregated (bagged) decision trees, and Robust Boosting (RB) which is an ensemble of classification trees and tolerant to noise [41]. The RB algorithm is an effective classification approach by which results from many weak learners, e.g. decision trees, are melded into one high-quality ensemble predictor.

A majority voting learner (MV) was also defined which combines predictions made by the other three classifiers and resulting in a new classifier that can potentially improve predictions. Our majority voting classifier combines predictions as

$$C_h(X) = mode\{P_1(x), P_2(x), P_3(x)\}, \tag{29}$$

where P_is are binary predictions made by SVM, RF, and RB. This approach allows the combination of classifiers by calculating the most frequently occurring value in the predictions [42].

We evaluate the performance of the above classifiers in the cross-validation phase and report the results using four classic performance measures, i.e. sensitivity (SE), specificity (SP), positive predictive value (PPV) or precision (PR), accuracy (ACC) and F_1 score, for each classifier independently. Additionally, we consider Cohen's Kappa coefficient which is a measure of agreement between the classification predictions and visual interpretation of telemetry ECG recordings.

E. Cross-Validation

We used a leave-one-person-out cross-validation (LOOCV) to evaluate the generalization performance of our AFib detection algorithm ($n = 300$) on different training sets, each representing certain segment length (10, 20,…, 60s). Additionally, we utilized three independent classifiers (SVM, RF, RB) to determine the presence of AFib in each single study subject, and further employed a majority voting learner over

TABLE II

CLASSIFICATION RESULTS OF THE CROSS-VALIDATION STUDY WITH DIFFERENT SEGMENTATION AND FEATURE GROUPS.

Feature matrix	6-axis SCG-GCG						3-axis SCG						3-axis GCG					
Segmentation (best classifier per group)	SE	SP	PR	kappa	F_1	ACC	SE	SP	PR	kappa	F_1	ACC	SE	SP	PR	kappa	F_1	ACC
10 sec (RF/RF/MV)	**98.6**	**95**	**95**	**0.94**	**97**	**97**	97.3	94.7	94.8	0.92	94	96	97.3	93	93	0.9	94.5	95
20 sec (MV/MV/MV)	98.6	93	93	0.92	96	96	98	94	94.2	0.92	96	96	96	93	93	0.87	94	94
30 sec (MV/MV/MV)	97.3	92.7	93	0.90	95	95	97.3	94.7	94.8	0.92	95.6	96	96.7	92	92.4	0.88	94	94.3
40 sec (MV/MV/MV)	98	93	93	0.90	94.5	95	98	92	92	0.9	95.8	95	97.3	93.3	93.5	0.91	95	95.3
50 sec (MV/RF/RF)	97.3	92	92.4	0.89	94.5	95	96	93.3	93.5	0.89	93	94.7	97.3	92.7	93	0.9	94.5	95
60 sec (RF/MV/MV)	97.3	92.7	93	0.90	95	95	98	93.3	93.6	0.91	95.5	96	96	93.3	93.5	0.89	94.6	95

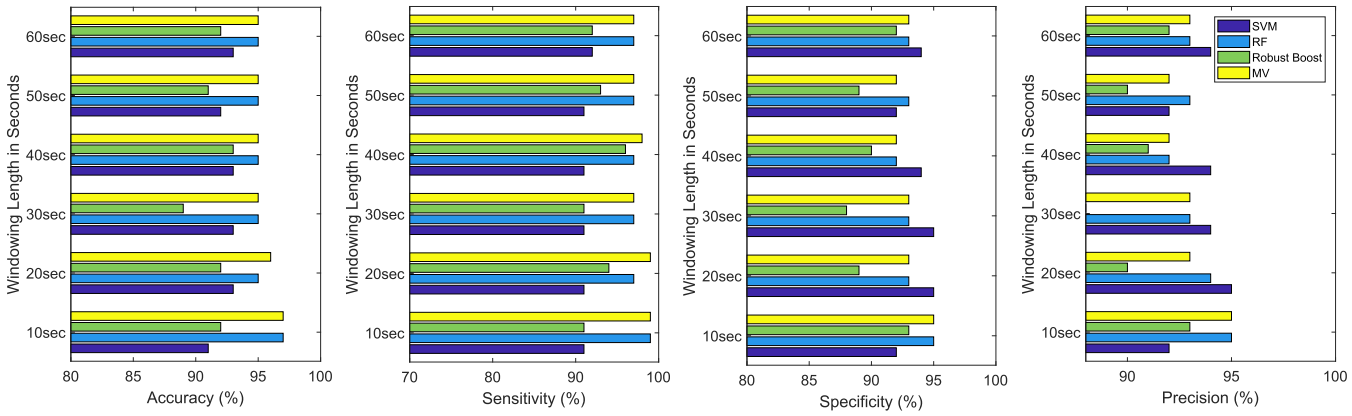

Fig. 4. Classification performance in the cross-validation study with three classifiers and majority voting (MV_{CV}).

all classifiers (MV_{CV} for cross validation and MV_{CD} for cross-database subsets). We report the results obtained with 3 different feature groups: i) 6-axis combination of SCG-GCG features (SSA + signal envelope), ii) tri-axial SCG features (SSA + signal envelope), and iii) tri-axial GCG feature (SSA + signal envelope).

IV. RESULTS

A. Cross-Validation Study

Table II presents the achieved LOOCV results where three sets of classification quality metrics are reported, each corresponds to a feature group. The table provides detailed results obtained for each segmentation scheme along with the best performing classifier for each corresponding feature group. Our criterion to select the best classifier was the F_1 score and Kappa coefficient obtained from each learning method. As can be seen from this table, the 6-axis feature combination delivers high discriminative performance than the two other feature groups, i.e. SCG/GCG derived features. Accordingly, we can see that the highest ACC (97%), SE (98.6%), SP (95%), and PPV (95%) were achieved with RF and 10s segments. The Cohen's kappa coefficient of the SCG-GCG features was 0.94 indicating a near-perfect agreement in rhythm classification between the smartphone algorithm and visual interpretation of telemetry ECG recordings. The F_1score was 97% showing a fairly excellent performance for the smartphone algorithm. A closer inspection of the table also reveals that although SCG and GCG report relatively close results to the 6-axis feature set, neither SCG nor GCG could alone succeed in achieving this high classification performances as 6-axis feature combination. The differences between the feature matrices and different segment lengths are highlighted. This suggests that with the minimum segment length, fusion of both SCG and GCG derived features, and a suitable classifier it is possible to accurately determine the presence of AFib. It is worth mentioning that the reported results are based on concatenating features extracted from SSA and signal envelope, while our experiments along with these findings showed poorer performances without concatenation, i.e. learning via either SSA or signal envelope features.

Fig. 4 presents the summary of classification performances obtained by different classifiers and segment lengths with the concatenated 6-axis feature matrix. As shown in this figure, the highest classification performance was achieved by the RF classifier with 10s signal segments, while a slightly poorer performance obtained when the segment length was 30s or more. By further evaluation of classifiers used in this study, we noted a remarkable performance with RF which positively affected the MV_{CV} results.

Fig. 5 shows the diagnostic ability of KSVM, RF, and RB using receiver operator characteristic (ROC) curves obtained for discriminating AFib from SR from 10s segmentation 6-axis SCG-GCG data. As shown, the area under the curve (AUC) obtained by RF is superior (0.987-0.991) as compared to KSVM (0.972-0.983), and RB (0.968-0.976) in all cases.

B. Cross-Database Blinded Experiment

We utilized the presented approach to determine the presence of AFib in 135 cases including 40 AFib subjects in

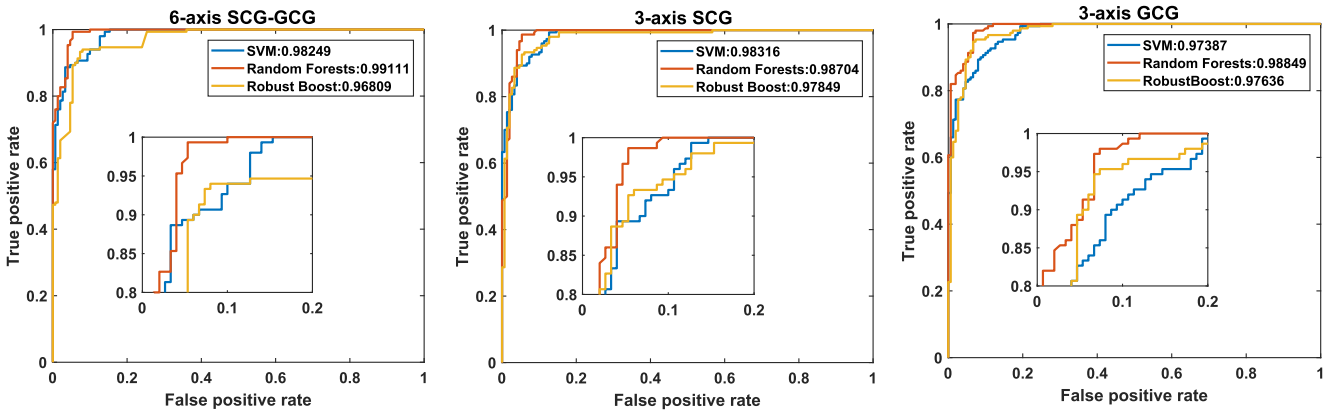

Fig. 5. ROC curves and AUC values obtained with all classifiers for 10s segmentation.

TABLE III
CLASSIFICATION RESULTS OF THE CROSS-DATABASE TEST STUDY WITH DIFFERENT SEGMENTATION AND FEATURE GROUPS

Feature matrix	6-axis SCG-GCG						3-axis SCG						3-axis GCG					
Segmentation (classifier per group)	SE	SP	PR	kappa	F_1	ACC	SE	SP	PR	kappa	F_1	ACC	SE	SP	PR	kappa	F_1	ACC
10 sec (MV/RF/MV)	**90**	**96.8**	**92.3**	**0.88**	**96**	**94.8**	90	96.8	92.3	0.88	95.6	94.8	88	97.9	94.6	0.87	96.3	94.8
20 sec (MV/RF/RF)	92.5	95.8	90.2	0.88	95	94.8	92.5	88.4	77.5	0.77	91	90	90	92.6	83.7	0.8	94	91.9
30 sec (MV/MV/MV)	85	94.7	87.2	0.79	94	91.85	87.5	92.6	83.3	0.79	92.5	91.1	82.5	97.9	94.3	0.82	95.4	92.6
40 sec (MV/MV/MV)	85	93.70	85	0.79	93.7	91.1	82.5	92.6	82.5	0.75	92.6	89.6	82.5	91.6	80.5	0.73	92.6	89.6
50 sec (MV/MV/MV)	92.5	92.6	84.1	0.82	95	92.6	85	90.5	79	0.7	92	88.9	90	93.7	85.7	0.82	92	92.6
60 sec (RF/RF/RF)	90	89.3	78.2	0.75	91.5	89.5	90	88	77	0.75	89	89	90	88	77	0.74	92	89

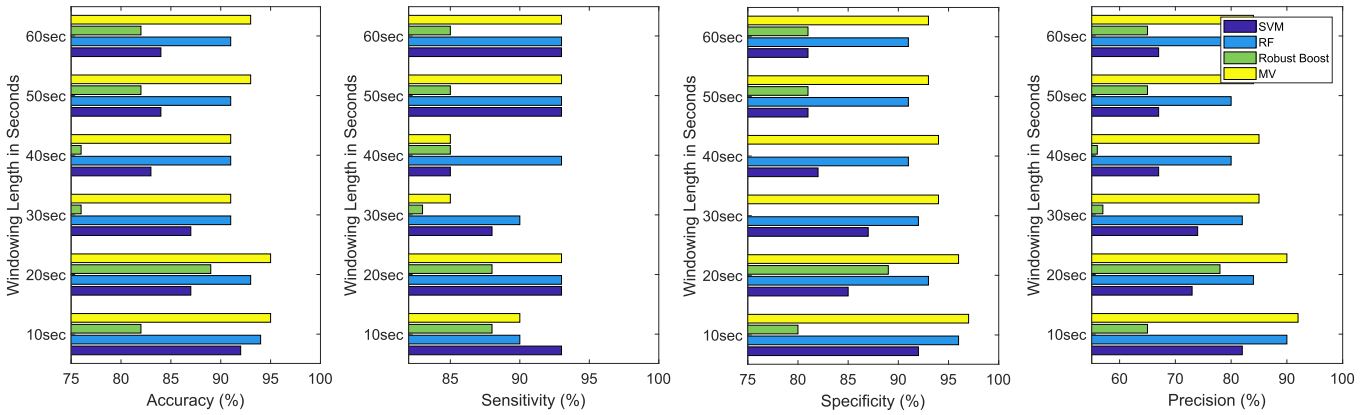

Fig. 6. Classifications with KSVM, RF, RB, and MV_{CD} classifiers and combined SCG-GCG features.

a blinded manner. We trained our classifiers, i.e. KSVM, RF, RB, and MV, to predict AFib from SR in this unseen data with similar segmentation scheme, i.e. 10s,…, 60s. The classifiers were trained using 300 subjects and subsequently tested with 135 cases.

Table III shows the overall performance of the presented algorithms while tested with different segmentation lengths. The algorithms demonstrate 94.8% accuracy and 92.3% precision with 10s segmentation using 6-axis feature SCG-GCG combination and MV learner. A sensitivity and specificity of 90% and 96.8% were also obtained. The kappa coefficient was 0.88 indicating a fairly perfect agreement. The F_1 score was 96% yielding balanced precision and recall obtained by MV learning. Fig. 6 shows the performance of classifiers in details (classification interpretations with respect to segmentation lengths).

Taken together, the experimental results obtained in cross-validation and cross-database studies suggest that presented framework can detect AFib with a remarkable performance. Tables II and III are quite revealing in several ways. Both tables demonstrate that 10s segmentation is superior. Combination of SCG and GCG derived features has resulted in greater certainty in discriminating AFib from SR. Moreover, according to the classification rates reported in Tables II and III, MV learner was able to effectively separate AFib from SR with most of the segmentation cases. This suggests that majority

voting learner is a potentially useful approach in screening AFib as it resembles aggregation of two or more virtual experts in disease classification.

V. Discussion

We developed a machine learning approach for AFib classification using smartphone MCG. The developed algorithm was cross-validated with a large clinical dataset (n = 300) demonstrating 98.6% sensitivity and 95% specificity in discriminating AFib from SR. Furthermore, we tested our approach in a cohort of 135 subjects and obtained a sensitivity and specificity up to 92.5% and 96.8%, respectively. In general, an accuracy of 97% and 95% was achieved during cross-validation and test studies, respectively. We also performed rigorous analyses in order to find the most sufficient signal segmentation paradigm, and determined the best combination of dual-sensor features for achieving excellent discriminative performance. Accordingly, results indicated that 10s segmentation yields the best performance ($F_1 = 97\%$ and Kappa = 0.94 with the CV and $F_1 = 95\%$ and Kappa = 0.88 with the CD study) for AFib detection with smartphone.

The findings of this study suggest that mobile SCG-GCG based AFib detection has impressive clinical implications. In 2014 AHA/ACC/HRS guidelines on the treatment of AFib, no recommendations on the screening of AFib are provided other than stating that efforts for early detection of AFib (e.g. frequent monitoring of ECG) may be needed in individuals at risk for stroke [43]. The 2016 guideline on the management of AFib by the European Society of Cardiology gives special recommendation for opportunistic screening of AFib at healthcare contacts in those aged 65 years or older by pulse palpation followed by the recording of a 12-lead ECG when irregularity is detected to confirm the diagnosis of AFib [2]. The obvious shortcoming of the approach proposed in the guideline is that healthcare contacts usually occur irregularly and infrequently, thus providing a very limited chance to diagnose paroxysmal AFib. Unsurprisingly, studies have shown that repeated ECG recordings have superior yields in comparison to both single and 24-hour continuous ECG recordings in AFib detection [44], [45]. These findings suggest that intermittent and frequent screening is required to increase the chance of detecting AFib, creating a demand for self-operated rhythm screening tools.

Smartphones are quickly becoming ubiquitous, even among the elderly adults, and the number of smartphone users is also increasing at a rapid rate in developing countries, where low-cost healthcare solutions are particularly crucial. The readily available smartphones provide a unique opportunity for cost-effective screening of AFib, if they can be harnessed to reliably detect the arrhythmia. Several single-lead ECG-applications for smartphones and portable stand-alone ECG devices have been developed and studied for the diagnosis of AFib, although currently only AliveCor has US Food and Drug Administration approval. Variable and even suboptimal sensitivities and specificities for automated algorithm detection of AFib with single-lead ECG applications have been presented in different studies (71.4-98.5% and 91.4-99.4%, respectively) [13], [46]–[48]. Thus, the performance of our MCG-based algorithm compares favorably to the various ECG-based techniques. Additionally, single-lead ECG tracings are often of insufficient quality and visual diagnosis of AFib may prove difficult even to cardiologists as suggested by the results presented by Lau *et al.* [13]. Furthermore, all ECG approaches, unlike MCG, always require additional hardware, which renders the method impractical considering large-scale screening efforts.

PPG is another method, which may be utilized to detect AFib with a smartphone, while no additional hardware is needed. PPG recordings are obtained by placing a fingertip on the camera and flash of the device and extracting a pulse wave signal from a video recording. In previous smaller studies, PPG algorithms have demonstrated 90-96% sensitivity and 85-97.7% specificity in detecting AFib as compared to physician-interpreted concurrent single-lead ECG recordings [20]. Although our MCG algorithm and the recent PPG algorithms have comparable accuracy, MCG has several advantages over PPG. Holding a fingertip precisely against the camera lens of a smartphone and maintaining a proper contact for several minutes is difficult and may prove infeasible for some elderly adults, as acknowledged by Krivoshei *et al.* [20]. Additionally, motion artifacts caused by finger movement during recording has been reported to cause false positive results, which is a major limitation considering the difficulty of maintaining a static finger position [48]. MCG recordings, however, are not affected by this problem, as the subject is merely required to passively seat or lay supine for the duration of recording, while potential motion artifacts are removed by the algorithm. Furthermore, while our MCG algorithm retained optimal performance even when the duration of the analyzed section of recording was reduced to 10 seconds, PPG seems to require longer recording durations for best performance as demonstrated by Krivoshei *et al.* [20].

Passive PPG screening methods by using a smart watch as the recording device are being developed [8], but fundamental challenges regarding motion artifacts and battery consumption remain to be resolved. There are also several possible approaches for passive MCG screening of AFib, such as utilizing smart watches, most of which also contain accelerometer and gyroscope. Any passive screening approach will, however, require further algorithm development and validation. Low battery consumption is a crucial advantage of MCG in passive rhythm screening, enabling long recording periods.

A limitation of this work is that the design of the current study might somewhat overestimate the accuracy of our algorithm as this study considered MCG recordings obtained by a single investigator. Self-recording and large-scale population screening studies are still needed to further evaluate the precision of our modality in diagnosing AFib. Efforts to create a passive approach to screening are also warranted. Even though MCG recordings in this study were obtained by only Sony Xperia smartphones, pre-study investigations did not show any considerable difference in terms of signal quality between the other brands and models for both Android and IOP operated phones.

The concept of measuring the mechanical movement of the heart muscle offers an entirely new and innovative method to evaluate cardiac rhythm and function [22], [23]. We believe

that the modality has considerable potential to offer new and useful tools for clinicians treating cardiac patients. Further studies on different MCG applications are currently under way.

VI. Conclusion

Smartphone mechanocardiography reliably detects AFib without any additional hardware and provides a new easy-to-use and accessible concept for AFib screening. Moreover, due to the wide availability of smart devices with embedded inertial sensors, the proposed methods could potentially scale to other fields such as wearable or embedded body sensor networks. The obtained results in this study are encouraging and warrant further investigation into this topic. It should be stated that in this work the focus was on classification between AFib and SR; the effect of other arrhythmias or cardiac disorders to the quality of the classification needs to be considered in future works.

Acknowledgments

The authors declare no conflict of interest.

References

[1] V. Fuster *et al.*, "ACC/AHA/ESC guidelines for the management of patients with atrial fibrillation: Executive summary: A report of the American College of Cardiology/American Heart Association task force on practice guidelines and the European society of cardiology committee for practice guidelines and policy conferences," *J. Amer. College Cardiol.*, vol. 38, no. 4, pp. 1231–1265, 2001.

[2] P. Kirchhof *et al.*, "2016 ESC guidelines for the management of atrial fibrillation developed in collaboration with EACTS," *Eur. Heart J.*, vol. 37, no. 38, pp. 2893–2962, 2016.

[3] A. J. Camm, P. Kirchhof, G. Y. Lip, and E. A. Schotten, "Guidelines for the management of atrial fibrillation: The task force for the management of atrial fibrillation of the European Society of Cardiology (ESC)," *Eur. Heart J.*, vol. 31, no. 19, pp. 2369–2429, 2010.

[4] S. M. Vaziri, M. G. Larson, E. J. Benjamin, and D. Levy, "Echocardiographic predictors of nonrheumatic atrial fibrillation. The Framingham heart study," *Circulation*, vol. 89, no. 2, pp. 724–730, 1994.

[5] D. H. Lewis, G. W. Deitz, J. D. Wallace, and J. R. Brown, Jr., "Intracardiac phonocardiography in man," *Circulation*, vol. 16, no. 5, pp. 764–775, 1957.

[6] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiol. Meas.*, vol. 28, no. 3, p. R1, 2007.

[7] J. Lee, B. A. Reyes, D. D. McManus, O. Mathias, and K. H. Chon, "Atrial fibrillation detection using an iPhone 4S," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 1, pp. 203–206, Jan. 2013.

[8] A. Carpenter and A. Frontera, "Smart-watches: A potential challenger to the implantable loop recorder?" *Europace*, vol. 18, no. 6, pp. 791–793, 2016.

[9] O. Inan, M. Etemadi, R. Wiard, L. Giovangrandi, and G. Kovacs, "Robust ballistocardiogram acquisition for home monitoring," *Physiol. Meas.*, vol. 30, no. 2, p. 169, 2009.

[10] A. G. Bonomi *et al.*, "Atrial fibrillation detection using photoplethysmography and acceleration data at the wrist," in *Proc. Comput. Cardiol. Conf. (CinC)*, 2016, pp. 277–280.

[11] T. Conroy, J. H. Guzman, B. Hall, G. Tsouri, and J.-P. Couderc, "Detection of atrial fibrillation using an earlobe photoplethysmographic sensor," *Physiol. Meas.*, vol. 38, no. 10, p. 1906, 2017.

[12] J.-P. Couderc *et al.*, "Detection of atrial fibrillation using contactless facial video monitoring," *Heart Rhythm*, vol. 12, no. 1, pp. 195–201, 2015.

[13] J. K. Lau *et al.*, "iPhone ECG application for community screening to detect silent atrial fibrillation: A novel technology to prevent stroke," *Int. J. Cardiol.*, vol. 165, no. 1, pp. 193–194, 2013.

[14] T. Hendrikx, M. Rosenqvist, P. Wester, and H. Sandström, and R. Hörnsten, "Intermittent short ECG recording is more effective than 24-hour Holter ECG in detection of arrhythmias," *BMC Cardiovascular Disorders*, vol. 14, no. 1, p. 41, Apr. 2014.

[15] R. Tieleman *et al.*, "Validation and clinical use of a novel diagnostic device for screening of atrial fibrillation," *Europace*, vol. 16, no. 9, pp. 1291–1295, 2014.

[16] P. M. Barrett *et al.*, "Comparison of 24-hour Holter monitoring with 14-day novel adhesive patch electrocardiographic monitoring," *Amer. J. Med.*, vol. 127, no. 1, pp. 95.e11–95.e17, Jan. 2014.

[17] J. A. Walsh, E. J. Topol, and S. R. Steinhubl, "Novel wireless devices for cardiac monitoring," *Circulation*, vol. 130, no. 7, pp. 573–581, 2014.

[18] J. M. Bumgarner *et al.*, "Smartwatch algorithm for automated detection of atrial fibrillation," *J. Amer. College Cardiol.*, vol. 71, no. 21, pp. 2381–2388, 2018.

[19] S. Nemati *et al.*, "Monitoring and detecting atrial fibrillation using wearable technology," in *Proc. IEEE 38th Annu. Int. Conf. Eng. Med. Biol. Soc. (EMBC)*, Aug. 2016, pp. 3394–3397.

[20] L. Krivoshei *et al.*, "Smart detection of atrial fibrillation," *EP Europace*, vol. 19, no. 5, pp. 753–757, 2017.

[21] B. S. Bozhenko, "Seismocardiography—A new method in the study of functional conditions of the heart," *Terapevticheskii Arkhiv*, vol. 33, pp. 55–64, Sep. 1961.

[22] M. J. Tadi *et al.*, "Gyrocardiography: A new non-invasive monitoring method for the assessment of cardiac mechanics and the estimation of hemodynamic variables," *Sci. Rep.*, vol. 7, no. 1, p. 6823, 2017.

[23] O. Lahdenoja *et al.*, "Atrial fibrillation detection via accelerometer and gyroscope of a smartphone," *IEEE J. Biomed. Health Inform.*, vol. 22, no. 1, pp. 108–118, Jan. 2017.

[24] C. Bruser, J. Diesel, M. D. Zink, S. Winter, P. Schauerte, and S. Leonhardt, "Automatic detection of atrial fibrillation in cardiac vibration signals," *IEEE J. Biomed. Health Inform.*, vol. 17, no. 1, pp. 162–171, Jan. 2013.

[25] T. Hurnanen *et al.*, "Automated detection of atrial fibrillation based on time–frequency analysis of seismocardiograms," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 5, pp. 1233–1241, Sep. 2017.

[26] O. T. Inan *et al.*, "Ballistocardiography and seismocardiography: A review of recent advances," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 4, pp. 1414–1427, Jul. 2015.

[27] J. Jaakkola *et al.*, "Mobile phone detection OF atrial fibrillation: The mode-AF study," *Circulation*, vol. 71, no. 11, p. A410, 2018.

[28] N. Golyandina and A. Zhigljavsky, *Singular Spectrum Analysis for Time Series*. Berlin, Germany: Springer-Verlag, 2013.

[29] F. J. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," *Proc. IEEE*, vol. 66, no. 1, pp. 51–83, Jan. 1978.

[30] D. Fry, *The Physics of Speech (Cambridge Textbooks in Linguistics)*. Cambridge, MA, USA: Cambridge Univ. Press, 1979.

[31] J. Jezewski, D. Roj, J. Wrobel, and K. Horoba, "A novel technique for fetal heart rate estimation from Doppler ultrasound signal," *Biomed. Eng. OnLine*, vol. 10, no. 1, p. 92, 2011.

[32] P. J. Murphy, "Periodicity estimation in synthesized phonation signals using cepstral rahmonic peaks," *Speech Commun.*, vol. 48, no. 12, pp. 1704–1713, 2006.

[33] A. Phinyomark, S. Thongpanja, H. Hu, P. Phukpattaranont, and C. Limsakul, "The usefulness of mean and median frequencies in electromyography analysis," in *Computational Intelligence in Electromyography Analysis—A Perspective on Current Applications and Future Challenges*. Rijeka, Croatia: InTech, 2012.

[34] K. K. Ho *et al.*, "Predicting survival in heart failure case and control subjects by use of fully automated methods for deriving nonlinear and conventional indices of heart rate dynamics," *Circulation*, vol. 96, no. 3, pp. 842–848, Aug. 1997.

[35] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *Amer. J. Physiol. Heart Circulatory Physiol.*, vol. 278, no. 6, pp. H2039–H2049, 2000.

[36] I. A. Rezek and S. J. Roberts, "Stochastic complexity measures for physiological signal analysis," *IEEE Trans. Biomed. Eng.*, vol. 45, no. 9, pp. 1186–1191, Sep. 1998.

[37] Electrophysiology and Task Force of the European Society of Cardiology the North American Society of Pacing, "Heart rate variability: Standards of measurement, physiological interpretation, and clinical use," *Circulation*, vol. 93, no. 5, pp. 1043–1065, 1996.

[38] T. Higuchi, "Approach to an irregular time series on the basis of the fractal theory," *Phys. D, Nonlinear Phenom.*, vol. 31, no. 2, pp. 277–283, 1988.

[39] B. Hjorth, "EEG analysis based on time domain properties," *Electroencephalogr. Clin. Neurophysiol.*, vol. 29, no. 3, pp. 306–310, 1970.

[40] S. P. Shashikumar, A. J. Shah, Q. Li, G. D. Clifford, and S. Nemati, "A deep learning approach to monitoring and detecting atrial fibrillation using wearable technology," in *Proc. IEEE EMBS Int. Conf. Biomed. Health Inform. (BHI)*, Feb. 2017, pp. 141–144.
[41] Y. Freund. (2009). "A more robust boosting algorithm." [Online]. Available: https://arxiv.org/abs/0905.2138
[42] G. James, "Majority vote classifiers: Theory and applications," Ph.D. dissertation, Dept. Statist., Stanford Univ., Stanford, CA, USA, 1998.
[43] C. T. January *et al.*, "2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation," *Circulation*, vol. 130, no. 23, pp. e199–e267, 2014.
[44] E. Svennberg, J. Engdahl, F. Al-Khalili, L. Friberg, V. Frykman, and M. Rosenqvist, "Mass screening for untreated atrial fibrillation the STROKESTOP study," *Circulation*, vol. 131, no. 25, pp. 2176–2184, 2015.
[45] P. D. Sobocinski, E. A. Rooth, V. F. Kull, M. von Arbin, H. Wallén, and M. Rosenqvist, "Improved screening for silent atrial fibrillation after ischaemic stroke," *EP Europace*, vol. 14, no. 8, pp. 1112–1116, 2012.
[46] N. Lowres *et al.*, "Feasibility and cost-effectiveness of stroke prevention through community screening for atrial fibrillation using iPhone ECG in pharmacies," *Thromb Haemost*, vol. 111, no. 6, pp. 1167–1176, 2014.
[47] J. Orchard *et al.*, "Screening for atrial fibrillation during influenza vaccinations by primary care nurses using a smartphone electrocardiograph (iECG): A feasibility study," *Eur. J. Preventive Cardiol.*, vol. 23, no. 2, pp. 13–20, 2016.
[48] P.-H. Chan *et al.*, "Diagnostic performance of a smartphone-based photoplethysmographic application for atrial fibrillation screening in a primary care setting," *J. Amer. Heart Assoc.*, vol. 5, no. 7, p. e003428, 2016.
[49] D. E. Lake, J. S. Richman, M. P. Griffin, and J. R. Moorman, "Sample entropy analysis of neonatal heart rate variability," *Amer. J. Physiol.-Regulatory, Integrative Comparative Physiol.*, vol. 283, no. 3, pp. R789–R797, 2002.

Mojtaba Jafari Tadi, photograph and biography not available at the time of publication.

Saeed Mehrang, photograph and biography not available at the time of publication.

Matti Kaisti, photograph and biography not available at the time of publication.

Olli Lahdenoja, photograph and biography not available at the time of publication.

Tero Hurnanen, photograph and biography not available at the time of publication.

Jussi Jaakkola, photograph and biography not available at the time of publication.

Samuli Jaakkola, photograph and biography not available at the time of publication.

Tuija Vasankari, photograph and biography not available at the time of publication.

Tuomas Kiviniemi, photograph and biography not available at the time of publication.

Juhani Airaksinen, photograph and biography not available at the time of publication.

Timo Knuutila, photograph and biography not available at the time of publication.

Eero Lehtonen, photograph and biography not available at the time of publication.

Tero Koivisto, photograph and biography not available at the time of publication.

Mikko Pänkäälä, photograph and biography not available at the time of publication.