

Article

# Human Identification by Cross-Correlation and Pattern Matching of Personalized Heartbeat: Influence of ECG Leads and Reference Database Size

Irena Jekova <sup>1,\*</sup>, Vessela Krasteva <sup>1</sup> and Ramun Schmid <sup>2</sup>

<sup>1</sup> Institute of Biophysics and Biomedical Engineering, Bulgarian Academy of Sciences, Acad. G. Bonchev Str. Bl 105, 1113 Sofia, Bulgaria; vessika@biomed.bas.bg

<sup>2</sup> Schiller AG, Signal Processing, 6341 Baar, Switzerland; ramun.schmid@schiller.ch

\* Correspondence: irena@biomed.bas.bg; Tel.: +359-2-9793-631

Received: 17 January 2018; Accepted: 24 January 2018; Published: 27 January 2018

**Abstract:** Human identification (ID) is a biometric task, comparing single input sample to many stored templates to identify an individual in a reference database. This paper aims to present the perspectives of personalized heartbeat pattern for reliable ECG-based identification. The investigations are using a database with 460 pairs of 12-lead resting electrocardiograms (ECG) with 10-s durations recorded at time-instants T1 and T2 > T1 + 1 year. Intra-subject long-term ECG stability and inter-subject variability of personalized PQRST (500 ms) and QRS (100 ms) patterns is quantified via cross-correlation, amplitude ratio and pattern matching between T1 and T2 using 7 features × 12-leads. Single and multi-lead ID models are trained on the first 230 ECG pairs. Their validation on 10, 20, ... 230 reference subjects (RS) from the remaining 230 ECG pairs shows: (i) two best single-lead ID models using lead II for a small population RS = (10–140) with identification accuracy  $AccID = (89.4–67.2)\%$  and aVF for a large population RS = (140–230) with  $AccID = (67.2–63.9)\%$ ; (ii) better performance of the 6-lead limb vs. The 6-lead chest ID model—(91.4–76.1)% vs. (90.9–70)% for RS = (10–230); (iii) best performance of the 12-lead ID model—(98.4–87.4)% for RS = (10–230). The tolerable reference database size, keeping  $AccID > 80\%$ , is RS = 30 in the single-lead ID scenario (II); RS = 50 (6 chest leads); RS = 100 (6 limb leads), RS > 230—maximal population in this study (12-lead ECG).

**Keywords:** electrocardiogram; ECG based biometrics; person identification; correlation; template matching; PQRST pattern; QRS pattern

## 1. Introduction

Biometric identity recognition systems using human body measurement data are vital for the security of financial transactions, access control, travel, etc. The various biometrics that are currently being adopted in identity recognition scenarios such as fingerprint, iris, face and speech recognition have operational trade-offs in terms of performance, measurability, circumvention and liveness detection [1]. These limitations have led to the development of smart solutions such as: (i) improving voice recognition by sophisticated pre-processing, followed by heuristical selection of significant features in time and frequency domains [2]; (ii) integrating face- and body-related soft biometric traits [3]; (iii) providing security and flexibility by simultaneous application of numeric password data and biometric fingerprint information [4]; (iv) assuring liveness detection via optical sensors, which establish the presence of pulse, the variations of optical characteristics caused by pressure changes and skin reaction to illumination with different wavelengths [5].

For more than a decade the electrocardiogram (ECG) has been intensively investigated in the context of biometric applications. The attention is attracted by its ability to provide an unambiguous

sign of liveness and to exhibit several personalized traits, reflecting the underlying physiological properties of the heart such as relative timings of the P, QRS, T waves and differences in beat geometry. An additional precondition for the use of biomedical data for identification purposes is the rapid development of multidimensional sensors making it possible to efficiently acquire such data [6,7]. Recently, the strategic application of biometric systems has been extended towards:

- remote healthcare monitoring [8,9] with biosensors integrated into mobile devices [10];
- secure wireless body area sensor networks [11,12];
- continuous authentication applications with adaptive strategies to follow individual beat variations in 24 h ECG recordings [13];
- continuous human recognition in driving settings [14].

Similar to any other biometric system, the operation of those, embedding ECG analysis, could be spitted into two phases [15]:

- Enrolment or registration phase, when a reference database from subjects with known identity is collected;
- Matching phase with two possible scenarios:
  - Person verification, when the system accepts or rejects someone's identity based on a single comparison to the data for a given ID in the reference database (one-to-one scenario);
  - Person identification, when the system is finding the identity of an individual based on a search in a reference database for the subject whose data best matches the input data (one-to-many scenario).

In terms of feature extraction methods, identification techniques based on ECG analysis could be assigned into three categories:

- Fiducial based approaches, which analyse:
  - Temporal, amplitude, area and angle features over single lead [9] or 12-lead ECG [16];
  - Only temporal features measured in a single lead that are hypothesized to be invariant to sensor placement [12,17,18].

The accuracy of these methods strongly depends on the correct localization of the waves' boundaries within the P-QRS-T segment.

- Non-fiducial methods based on calculation of:
  - Autocorrelation functions and their post-processing via discrete cosine transform [14,19];
  - Linear discriminant analysis [20] for feature space reduction and further identification;
  - Assessment of correlation coefficients between beats with known and unknown identity [21–23];
  - Pattern matching and estimation of normalized Euclidian distance between the tested and the target heartbeat [24];
  - Features describing the dynamic characteristics and chaotic behaviour in the ECG time series [25].
- Hybrid human identification methods combining both fiducial and non-fiducial strategies, which include:
  - Hierarchical identification scheme for temporal and amplitude measurements based on fiducial points detection and estimation of waveforms' similarity via Euclidean distance [26];
  - Fiducial-based measurements for reduction of the matching candidates' number followed by wavelet transform and application of the coefficients for template matching [8].

Numerous studies are focused on general-purpose practical applications limited to the use of a single ECG lead, usually from the limbs [10,17,18,23,24]. Others aim at effective channel combination schemes by involving 12-lead ECG [16,20–22] into the analysis in view of feasible applications in healthcare scenarios. The merit of multi-lead ECG is controversial, considering: on the one hand the identification accuracy improvement by increasing the analysed ECG leads in a fiducial independent application [22] and on the other hand the same level of identification accuracy for all 12 leads, only limb, only chest and only lead I in a fiducial-based application [16].

Despite that the majority of the studies have used public [9,10,19,20,23,26] or proprietary [10,16–18,21–24] clinical databases, some authors have addressed the application of different non-invasive portable single-channel ECG sensors which support the acquisition of ECG signals for human identification purposes. The reported identification accuracies are:

- ~99% for 30 ECG signals acquired from the individual via a biosensor integrated into mobile devices [10];
- ~90% for 30 ECG test patterns acquired from the palms of 9 subjects via a two poles self-developed ECG signal acquisition system [25];
- ~94% achieved for ECG recordings acquired from the fingers of 16 individuals via self-developed ECG sensor [27].

The use of such ECG sensors facilitates the practical application of ECG biometrics in areas such as cyber-security [25,27], cloud data security or remote healthcare systems [10].

The majority of the cited methods report high identification accuracy. However, certain limitations are observed such as:

- Small population size (<30 individuals) [16–22,25–27] with ECG recordings acquired in a very short temporal interval [16,21,22] or even during the same session [9,17,18,20,26];
- Application of one and the same database for both training and testing.

This is far from the realm of biometrics. Significant accuracy degradation could be expected [28] with these methods, due to the influence of different unpredictable factors such as long-term ECG changes [13], noisy environment [14], electrode misplacement, overtraining of the identification model, etc.

The aim of this paper is to present the 12-lead ECG as a biometric modality by describing three methods for comparison between input-to-template waveforms of personalized beat patterns based on cross-correlation, amplitude relations and a new method for binary pattern matching. We further investigate their potential for human identification within an uncommonly large population with one-year distant collection of the reference and test datasets. Unbiased statistical analysis on independent validation set is presented for different perspectives of the human identification problem, considering: (i) the choice of the optimal single and multi-lead ECG set; (ii) the influence of the reference database size, providing evidence about the reasonable size of the population that could be tolerated in ECG biometric studies.

## 2. Materials and Methods

### 2.1. ECG Database

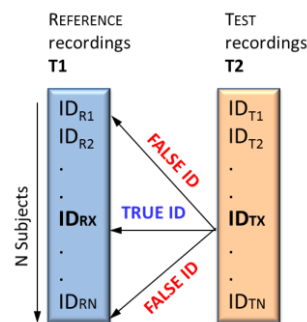
The study is using a proprietary clinical ECG database (Schiller AG, Baar, Switzerland) provided for research purposes and for the investigation of ECG's potential for human biometrics on a large population observed over time. It contains two 10-s sessions of standard 12-lead resting ECGs from 460 non-cardiac patients (235/225 male/female, 18–106 years old) admitted in the emergency department of the Basel University Hospital between 2004 and 2009. The ECGs were recorded via a commercial ECG device (SCHILLER AT-110 (500 Hz, 2.5  $\mu$ V/LSB, 0.05 Hz–150 Hz bandwidth), Schiller AG, Baar, Switzerland) at different time-points, i.e. reference point T1, corresponding to the enrolment phase of

the authentication task and remote point  $T2 > T1 + 1$  year, corresponding to the matching/login phase. The 460 ECG couples were equally divided into two datasets from non-overlapping populations for independent training and validation of the designed human identification models:

- Training dataset, including the ECG couples from 1 to 230;
- Validation dataset, including the ECG couples from 231 to 460.

## 2.2. Methods

The identification concept (Figure 1) is based on a comparison between a tested subject with an unknown identity ( $ID_T$ ) from dataset (T2) and a set of  $N$  subjects from a previously recorded reference dataset (T1) with known identities ( $ID_R$ ). The purpose in the identification task is to answer the question: “Who is this person?”



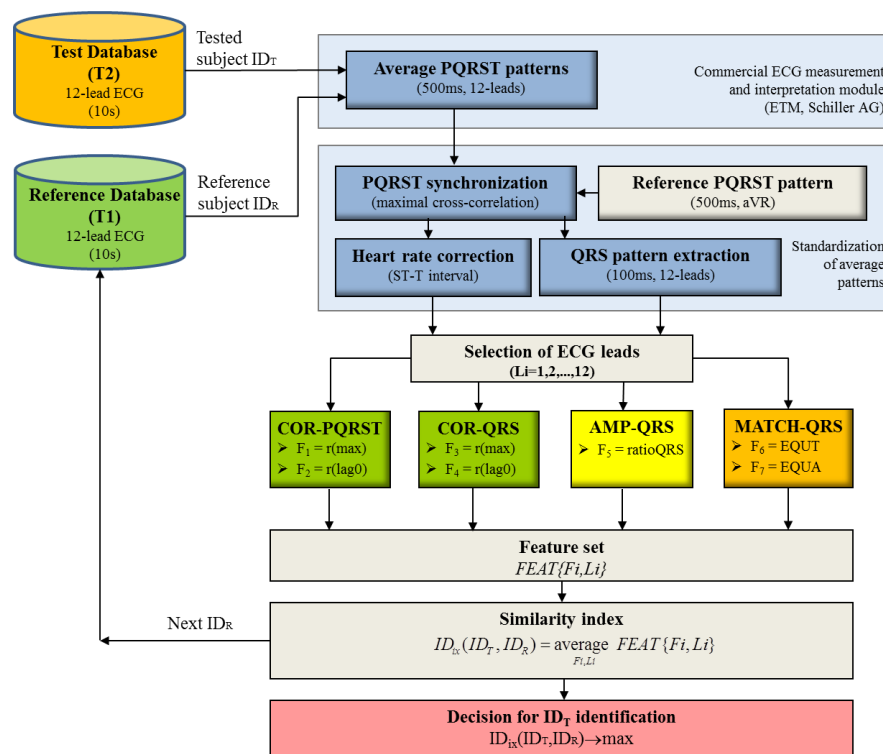
**Figure 1.** Identification concept scheme illustrating the comparison of the tested recording ( $ID_{TX}$ ) in T2 to all reference recordings in T1 (from  $ID_{R1}$  to  $ID_{RN}$ ).

The identification accuracy ( $AccID$ ) is calculated as follows:

$$AccID = 100 \cdot \frac{TrueID}{NSubjects} (\%) \quad (1)$$

where the true identifications ( $TrueID$ ) consider the number of all correctly decided cases for which the identity of  $ID_T(T2)$  exactly corresponds to the identity of  $ID_R(T1)$  and  $NSubjects$  is the number of tested subjects in the reference database.

The block-diagram of the module designed for human identification is presented in Figure 2. The ECG pre-processing and the methods for cross-correlation and amplitude features extraction are inherited from our recent study on ECG-based human verification [29]. The novelty in this identification model is generally in the designed brand-new method for QRS pattern matching and the decision making, based on features' common similarity index. We further describe the methods in detail to facilitate their reproduction.



**Figure 2.** Block-diagram of the identification process representing ECG pre-processing for the extraction of personalized 12-lead average PQRST pattern, features calculation and decision making.

### 2.2.1. ECG pre-processing

All ECG recordings are processed by a certified commercial ECG measurement and interpretation module (ETM, Schiller AG, Baar, Switzerland) for the extraction of a 12-lead average PQRST pattern with duration of 500 ms. This pattern, further referred as personalized pattern, provides higher signal-to-noise ratio and greater robustness against respiration-induced morphology changes in ECG as compared to single heartbeats.

Inter-subject standardization of 12-lead PQRST patterns is applied by:

- Detrending (removal of DC offset and linear trend);
- Synchronization of the cardiac depolarization process by time-alignment of the personalized PQRST pattern to a reference pattern by maximal cross-correlation in lead aVR. The reference pattern was selected at the beginning of the identification study. It represents PQRST segment in lead aVR with normal morphology (negative P, R, T waves) belonging to a healthy subject from the reference database (T1).
- Extraction of QRS pattern (100 ms), synchronously in all 12 leads within a window of 30 ms before and 70 ms after the fiducial point, aligned to the R-peak of the reference pattern. The additional investigations of this short pattern are motivated by the findings in our previous study [30] which highlight the biometric potential of the R and S-waves and reject the verification capability of the P, ST, T parts due to their low intra-subject reproducibility and low inter-subject variability.
- Calculation of heart rate corrected ST-T interval by means of Bazett's formula  $QTc = QT / \sqrt{RR}$ , where  $QTc$  is the corrected  $QT$  interval and  $RR$  is the RR interval.

### 2.2.2. Features extraction

Comparison between the waveforms of the personalized PQRST and QRS patterns of a tested vs. reference subject ( $ID_T$  vs.  $ID_R$ ) is performed independently for each of the 12 leads by three feature

extraction methods—cross-correlation analysis of QRS/PQRST and amplitude measurements over the QRS (applied also in [29]) and a brand-new method for assessment of QRS pattern matching between  $ID_T$  and  $ID_R$ .

(1) Cross-correlation analysis of QRS and PQRST patterns (COR-QRS and COR-PQRST) based on the cross-correlation function  $r$ :

$$r_{ID_T, ID_R}(lag) = \frac{\sum_{i=1}^{PD} Pattern_{ID_T}(i) Pattern_{ID_R}(i + lag)}{\sqrt{\sum_{i=1}^{PD} Pattern_{ID_T}(i)^2 \sum_{i=1}^{PD} Pattern_{ID_R}(i)^2}} \quad (2)$$

where *Pattern* denotes QRS or PQRST with pattern duration  $PD = 100$  ms or  $500$  ms, respectively. The *lag* value is changed in the range  $[-PD; PD]$ .

Two correlation coefficients are calculated in the normalized scale  $[0; 100]\%$ :

*Maximum correlation*  $r(\max)$  is representative for the best matching of the waveforms:

$$r(\max) = \max_{lag \in [-PD, PD]} r_{ID_T, ID_R}(lag) * 100(\%) \quad (3)$$

*Zero-lag correlation*  $r(lag0)$  is representative for the non-synchronized similarity between the ECG patterns:

$$r(lag0) = r_{ID_T, ID_R}(lag = 0) * 100(\%) \quad (4)$$

Reduced values of both correlation coefficients are expected for subjects with different ID ( $ID_T \neq ID_R$ ) due to the different spatio-temporal dynamics of the cardiac vector between subjects.

(2) Amplitude measurements of the QRS patterns (AMP-QRS) estimating the feature:

*Ratio of the minimal-to-maximal QRS amplitude* representative of the peak-to-peak amplitude equality between the QRS patterns in the studied lead, comparing  $ID_T$  and  $ID_R$  recordings.

$$ratio_{QRS} = \frac{\min_{ID=ID_T, ID_R} QRS_{ID}}{\max_{ID=ID_T, ID_R} QRS_{ID}} * 100(\%) \quad (5)$$

A ratio close to 100% is expected for equal identity subjects  $ID_{T(T1)} = ID_{R(T2)}$  considering the same recording conditions in T1 and T2 sessions (e.g. posture, electrode placement, recording ECG device, etc.) and therefore similar QRS amplitudes.

(3) QRS pattern matching (MATCH-QRS) applying amplitude normalization; time-amplitude approximation; and extraction of features for assessment of QRS pattern matching between  $ID_T$  and  $ID_R$ .

#### Amplitude normalization

The amplitudes of the QRS patterns in any pair ( $ID_T, ID_R$ ) and any lead are linearly scaled to fit in the range  $[-1; 1]$  in order to use one and the same computational range for all individuals and all leads regardless of their signal amplitudes. This is achieved by dividing each sample of the QRS patterns in a particular ECG lead to the maximal absolute amplitude observed in this lead among  $QRS_{ID_T}$  and  $QRS_{ID_R}$ .

#### Time-amplitude approximation

The QRS waveforms of  $ID_T$  and  $ID_R$  are approximated by the following binary transform ( $binQRS_{ID}$ ), which converts the QRS pattern into a 2-dimensional matrix:

$$binQRS_{ID}(ti, aj) = \begin{cases} 1 & \text{if } A(aj) \in [QRS_{ID}(ti) \pm 2\Delta a] \text{ or } A(aj) \in [QRS_{ID}(ti \pm \Delta t)] \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where

- $ti = [1, 2, \dots, 100]$  is the index of the time scale  $T(ti) = [1, 1 + \Delta t, 1 + 2\Delta t, \dots, 100]$  ms with resolution  $\Delta t = 1$  ms.
- $aj = [1, 2, \dots, 80]$  is the index of the normalized amplitude scale  $A(aj) = [-1, -1 + \Delta a, -1 + 2\Delta a, \dots, 1]$  with resolution  $\Delta a = 0.025$ .

The introduced tolerances  $ti \pm 1$  ms and  $A(aj) \pm 0.05$  manage small intra- and inter-subject variations of the cardiac depolarization process. Larger values of  $\Delta t$  and  $\Delta a$  would lead to a rougher approximation of the QRS patterns and potential loss of waveform details while smaller values support finer approximation at the cost of increased computational resources.

#### Extraction of pattern matching features

Two MATCH-QRS features are calculated for the couple  $(ID_T, ID_R)$ , normalized in the scale  $[0; 100]\%$ :

*Measure of the equality in the time scale (EQU T)* representing the time of overlapping of both QRS patterns after binary element-wise multiplication (AND operation):

$$EQU T = \frac{\Delta t}{100ms} \sum_{ti=1}^{100} binQRS_{ID_T}(ti, aj) \wedge binQRS_{ID_R}(ti, aj) * 100(\%) \text{ for } aj = (180) \quad (7)$$

*EQU T* ranges between 0% (null time coincidence, i.e. patterns do not overlap for any *binQRS* entry over the complete pattern length) and 100% (full-time coincidence, i.e. patterns overlap for all entries over the complete pattern length). The largest *EQU T* values are expected for equal identity subjects  $ID_T = ID_R$ .

*Measurement of the equality in the amplitude scale (EQU A):*

$$EQU A = 100 - DIFA(\%) \quad (8)$$

where *DIFA* (difference in the amplitude scale) represents the area enclosed between the non-overlapping amplitudes of both QRS patterns after binary element-wise multiplication and inversion (NAND operation):

$$DIFA = \frac{\Delta t}{100ms} \Delta a \sum_{ti=1}^{100} \sum_{aj=amin}^{amax} binQRS_{ID_T}(ti, aj) \bar{\wedge} binQRS_{ID_R}(ti, aj) * 100(\%) \quad (9)$$

The integration interval in the amplitude scale is enclosed between the minimal and maximal QRS amplitudes among  $ID_T$  and  $ID_R$  patterns measured at each specific time index  $ti$ , i.e.,  $[amin(ti) = \min_{ID=ID_R, ID_T} binQRS_{ID}(ti); [amax(ti) = \max_{ID=ID_R, ID_T} binQRS_{ID}(ti)]$ . *DIFA* ranges between 0% (full-amplitude coincidence, i.e., patterns overlap for all *binQRS* entries over the complete pattern length) and 100% (null amplitude coincidence, corresponding to the largest pattern differences that cover the full amplitude range). Therefore, minimal *DIFA* and inversely maximal *EQU A* values are expected for equal identity subjects  $ID_T = ID_R$ .

The redundant input feature space is a matrix  $FEAT\{Fi, Li\}$ , where:

- $Fi = (1-7)$  is the index in the feature vector  $\{r(\max)_{PQRST}, r(\text{lag}0)_{PQRST}, r(\max)_{QRS}, r(\text{lag}0)_{QRS}, ratio_{QRS}, EQU T, EQU A\}$ , including the option to enable/disable any feature from analysis and thus to test ECG identification scenarios with different feature sets.
- $Li = (1-12)$  is the index of the lead including the option to enable/disable any lead from analysis and thus to test different single- and multi-lead configurations for ECG identification.



### 2.2.3. Decision Making

For the purpose of human identification, the best matching between  $ID_T$  and  $ID_R$  is found by maximization of the similarity index:

$$ID_{ix} = \underset{Fi, Li}{\text{average FEAT}}\{Fi, Li\} \quad (10)$$

Optimal identification (ID) models with non-redundant features are trained by forward stepwise feature selection until maximization of accuracy  $AccID$  on the training dataset. The observed accuracy on the independent validation set could be considered as unbiased assessment of the human identification models.

The software package Matlab (The Mathworks Inc., Natick, MA, USA) was used for the management of the signal processing and statistical study. This includes training and validation of the forward stepwise ID models (reported as mean  $\pm$  confidence interval (CI)) as well as for comparison of continuous feature distributions (represented as mean  $\pm$  standard deviation (std)) via paired Student's *T*-Test. A value of  $p \leq 0.05$  is considered statistically significant.

## 3. Results

An example of a real human identification scenario based on 12-lead ECG is presented in Figure 3. In all leads we observe close matching between the personalized PQRST and QRS patterns of the same individual at two different time points (red vs. blue trace). In contrast, apparent deviances of the waveforms across different subjects are visible (red vs. grey traces) regardless of the overlapping of certain waveforms. Further in section Results we have presented the training and validation of ID models which deal with inter-subject waveform differences and long-term intra-subject waveform stability in 12-lead ECGs of a large population.

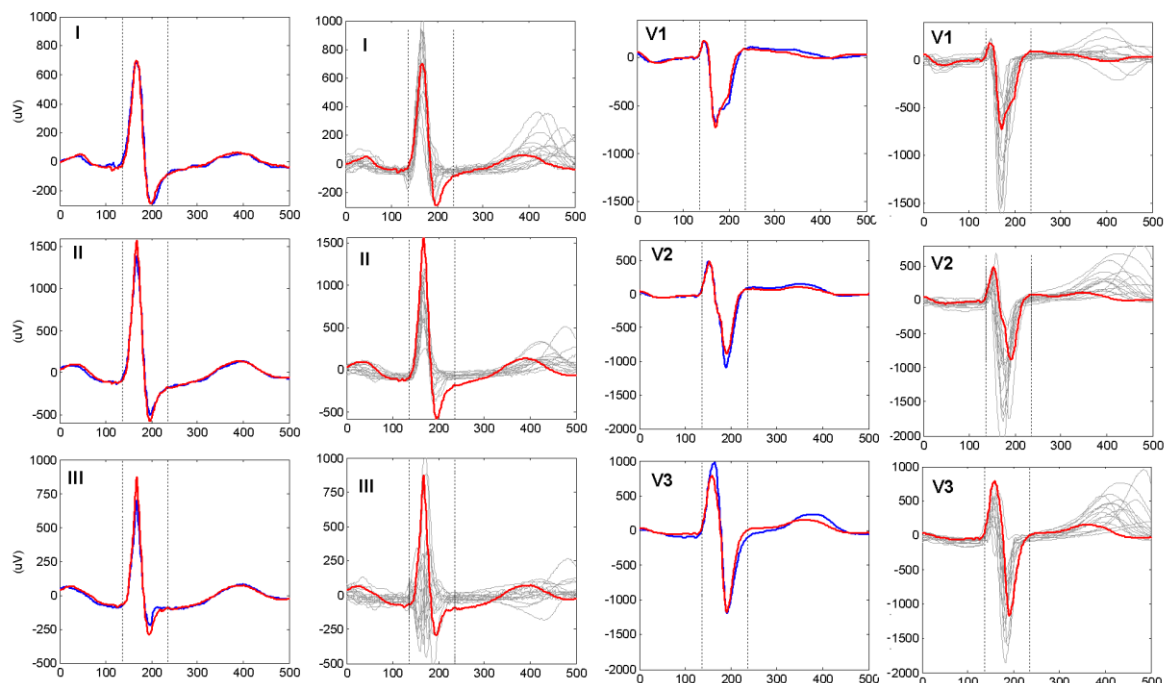
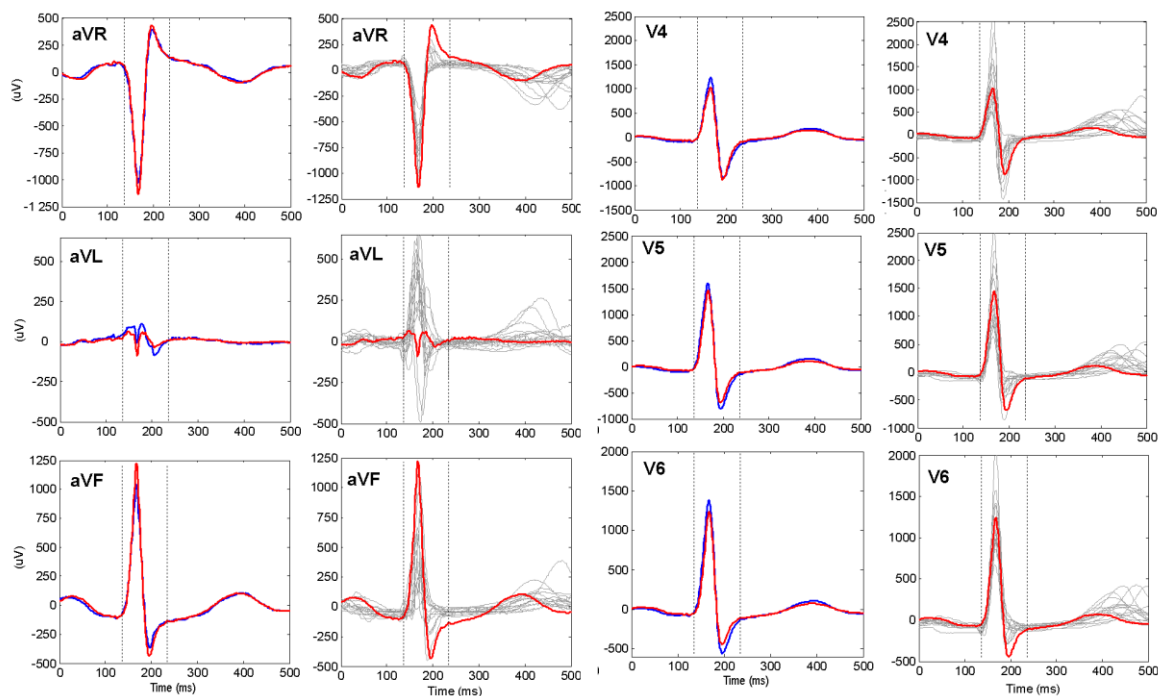


Figure 3. Cont.





**Figure 3.** Example of personalized 12-lead PQRST patterns analysed during the ID process. The identity of the tested subject  $ID_T$  (red line) is examined via comparison to a reference subject  $ID_R$  with the same identity (1st and 3rd columns, blue line) and 15 different identities (2nd and 4th columns, grey traces). The vertical dotted lines enclose the extracted QRS pattern (100 ms) which is used for the calculation of COR-QRS, AMP-QRS and MATCH-QRS features. COR-PQRST features are estimated over the entire 500 ms segment.

### 3.1. Statistical Study

The distributions of all features are estimated on the training dataset for 230 equal ( $ID_T = ID_R$ ) and  $230 \times 229 = 52670$  different ( $ID_T \neq ID_R$ ) identity pairs (Table 1). They prove the biometric potential of all features with significantly higher mean values for the groups of equal vs. different subjects (82–96% vs. 60–86%,  $p < 0.0001$ ). This is a precondition to expect distinguishable similarity index  $ID_{ix}$  involving all or a selected set of features used in the next phase of training of optimal ID models.

**Table 1.** Statistical distributions (Mean  $\pm$  std) of the features' values averaged for 12-lead ECG in the groups of equal ( $ID_T = ID_R$ ) and different ( $ID_T \neq ID_R$ ) individuals. Statistically significant differences are observed for all features.

Feature Extraction Method	Feature	$ID_T = ID_R$ (%)	$ID_T \neq ID_R$ (%)	<i>p</i> -Value
COR-PQRST	$r(\max)$	$93.81 \pm 2.92$	$81.10 \pm 5.93$	$<0.0001$
	$r(\text{lag}0)$	$92.2 \pm 3.69$	$70.01 \pm 10.74$	$<0.0001$
COR-QRS	$r(\max)$	$96.14 \pm 2.28$	$85.99 \pm 6.05$	$<0.0001$
	$r(\text{lag}0)$	$93.77 \pm 3.90$	$72.0 \pm 12.59$	$<0.0001$
AMP-QRS	$\text{ratioQRS}$	$83.44 \pm 4.87$	$70.25 \pm 7.62$	$<0.0001$
MATCH-QRS	$EQU_T$	$82.26 \pm 5.54$	$59.54 \pm 7.73$	$<0.0001$
	$EQU_A$	$92.74 \pm 2.49$	$80.59 \pm 4.22$	$<0.0001$

### 3.2. Training of ID Models on 230 Subjects ( $ID = 1$ to 230)

In order to assess the intrinsic identification capability of each feature, the simplest ID models involve one feature in one lead ( $FEAT\{Fi, Li\}$ , where  $Fi = \{1\}, \{2\}, \dots, \{7\}$ ;  $Li = \{1\}, \{2\}, \dots, \{12\}$ ). The results in Table 2 indicate poor  $AccID$  if only one feature is included in the ID model (not exceeding 41%).

The best identification capability is observed for COR-PQRST and COR-QRS single features in limb leads, while the worst performance (<1%) is observed for AMP-QRS features in all leads.

**Table 2.** Training *AccID* (%) of each feature in each lead. The colour map distinguishes five *AccID* levels for visual ranking of the feature's ability for human identification: the worst ≤5%, is given in red; [5–15]%, in orange; [15–25]%, in yellow; [25–35]%, in light green; and the highest, >35%, in dark green.

Lead	COR-PQRST		COR-QRS		AMP-QRS	MATCH-QRS	
	<i>r</i> (max)	<i>r</i> (lag0)	<i>r</i> (max)	<i>r</i> (lag0)	<i>ratio</i> QRS	EQUIT	EQUA
I	20	25.7	15.2	23.0	<1	10.0	23.9
II	23.5	29.6	18.7	24.3	<1	11.3	24.3
III	27.0	37.4	25.2	34.8	<1	5.7	22.2
aVR	15.2	13.0	10.9	12.2	<1	4.3	10.0
aVL	34.8	40.9	31.7	34.8	<1	4.8	24.8
aVF	33.0	39.1	23.5	38.3	<1	8.7	26.5
V1	15.7	21.3	13.0	13.9	<1	1.3	16.1
V2	16.5	20.4	11.7	19.6	<1	6.5	17.8
V3	12.6	19.6	7.0	11.7	<1	1.7	13.9
V4	14.3	20	7.4	12.2	<1	2.6	12.6
V5	11.7	15.7	7.4	8.3	<1	1.7	10.0
V6	13.0	13.5	5.2	7.4	<1	<1	6.5

More complex ID models involve all features from the same feature extraction method ( $FEAT\{Fi, Li\}$ , where  $Fi = \{1, 2\}, \{3, 4\}, \{5\}, \{6, 7\}$ ;  $Li = \{1, 2, \dots, 12\}$ ), to compare the identification capability of COR-PQRST, COR-QRS, AMP-QRS, MATCH-QRS. Table 3 indicates the best identification capability for MATCH-QRS with 15 features (83%), followed by COR-PQRST with 18 features (77%), COR-QRS with 9 features (74%) and the worst AMP-QRS with 7 features (24%). The non-redundant feature sets in the ID models include features that are equally selected from limb and chest leads.

**Table 3.** Optimal ID models for different feature extraction methods using 12-lead ECG. The non-redundant features and leads included in each model are marked with '+'. The training *AccID* (mean ± CI) of each model is reported in the last row. The applied colour map highlights the features involved in a model and the achieved *AccID* as follows: light green for COR-PQRST model; yellow for COR-QRS model; orange for AMP-QRS model and dark green for MATCH-QRS model.

Lead	COR-PQRST		COR-QRS		AMP-QRS	MATCH-QRS	
	<i>r</i> (max)	<i>r</i> (lag0)	<i>r</i> (max)	<i>r</i> (lag0)	<i>ratio</i> QRS	EQUIT	EQUA
I	+	+	+	+	+	+	+
II	+	+		+	+		+
III	+				+		+
aVR	+			+		+	+
aVL	+	+				+	
aVF	+	+		+			+
V1	+	+	+	+	+	+	+
V2	+	+			+		+
V3	+				+	+	
V4		+					+
V5		+	+		+		+
V6	+		+				+
<i>AccID</i> (%)	77.4 ± 5.5		73.5 ± 5.8		23.9 ± 5.5	83.5 ± 4.9	

The next type of ID models involves all features in a single lead set (1 lead, 6 limb leads, 6 chest leads, 12 leads, i.e.  $FEAT\{Fi, Li\}$ , where  $Fi = \{1, 2, \dots, 7\}$ ;  $Li = \{1\}, \{2\}, \dots, \{12\}, \{1, 2, \dots, 6\}, \{7, 8, \dots, 12\}, \{1, 2, \dots, 12\}$ ). The purpose is to assess the intrinsic identification capability of that specific lead set (Table 4). The identification capability of single limb leads is the lowest in III, aVR (about 47% with up to 5 features) and the highest in I, II (about 57–58% with 5 features). It is increased by 20% when

6 limb leads are used in the ID model (about 77% with 19 features). The identification capability of single chest leads is the lowest in V3 (about 27% with 3 features) and the highest in V1 (about 44% with 4 features). It is increased by 20% when 6 chest leads are used in the ID model (about 64% with 18 features). The most complex ID model with 12-leads and 27 features improves *AccID* to 91.3%, which is about 14% and 27% points better than 6 limb and 6 chest leads, respectively.

**Table 4.** Optimal ID models for single- and multi-lead sets. The non-redundant features included in the models are marked: '+' for single leads, '\*' for limb leads, 'o' for chest leads, '#' for 12-lead ECG. Additionally, the last 3 rows belonging to the multi-lead models show the number of leads per feature, involved in the model. The training *AccID* (mean  $\pm$  CI) of each model is reported in the last column. The following colour map is applied: the features involved in the single lead models and the achieved *AccID* are highlighted in yellow; the features involved in the model using the 6 chest leads and the respective *AccID* are highlighted in orange; the features involved in the model using the 6 limb leads and its *AccID* are highlighted in light green; and the features involved in the 12-leads model and the achieved *AccID* are highlighted in dark green.

Lead Set	COR-PQRST		COR-QRS		AMP-QRS	MATCH-QRS		<i>AccID</i> (%)	
	<i>r</i> (max)	<i>r</i> (lag0)	<i>r</i> (max)	<i>r</i> (lag0)	<i>ratio</i> QRS	EQUIT	EQUA		
I	*	#	+	*	#	+	*	#	57.8 $\pm$ 6.4
II	+	*	#	+	*	#	+	*	57.4 $\pm$ 6.4
III	+	*	#	+	*	#	+	*	46.5 $\pm$ 6.5
aVR	+	*	#	+	*	#	+	*	47.0 $\pm$ 6.6
aVL	+	*	#	+	*	#	+	*	50.4 $\pm$ 6.5
aVF	+	*	#	+	*	#	+	*	55.2 $\pm$ 6.5
V1	+	o	#	+	o	#	+	o	44.3 $\pm$ 6.4
V2	+	o	#	+	o	#	+	o	34.8 $\pm$ 6.2
V3	+	o	#	+	o	#	+	o	27.4 $\pm$ 5.8
V4	+	o	#	+	o	#	+	o	30.0 $\pm$ 6.0
V5	+	o	#	+	o	#	+	o	34.8 $\pm$ 6.2
V6	+	o	#	+	o	#	+	o	37.4 $\pm$ 6.3
6 Limb	3	5	3	4		2	2		77.4 $\pm$ 5.5
6 Chest	4	3	4	3		4	4		63.9 $\pm$ 6.2
12-leads	7	5	4	2		3	6		91.3 $\pm$ 3.7

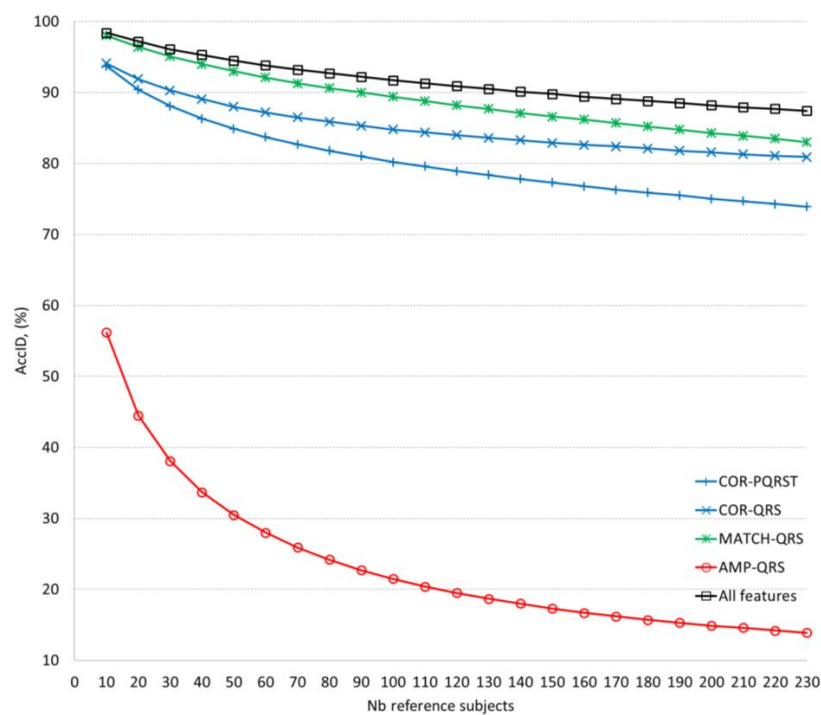
### 3.3. Validation of ID Models on 230 Subjects (ID = 231 to 460)

The identification accuracy strongly depends on the number of reference subjects, i.e. The more the subjects there are, the higher the random match probability is, i.e. to find a subject with better similarity than his own. Therefore, *AccID* worsening is expected for reference database with larger size. We observe this trend for the independent validation set in order to find the number of reference subjects that could be managed by the trained ID models within different *AccID* tolerance ranges. For this purpose, we consider different number of reference subjects and test the ID models with all possible combinations of 10, 20, 30, . . . , 210, 220, 230 subjects within the total validation set of 230 subjects.

Table 5 and Figure 4 compare the validation performance of the optimal ID models for different feature extraction methods using 12-lead ECG. Accepting a confident *AccID* threshold (>90%), we found that it is satisfied for 140, 90, 30, 20 reference subjects by the ID models based on all features, MATCH-QRS, COR-QRS, COR-PQRST, respectively. The AMP-QRS model is limited to maximum 56% achieved just for 10 reference subjects.

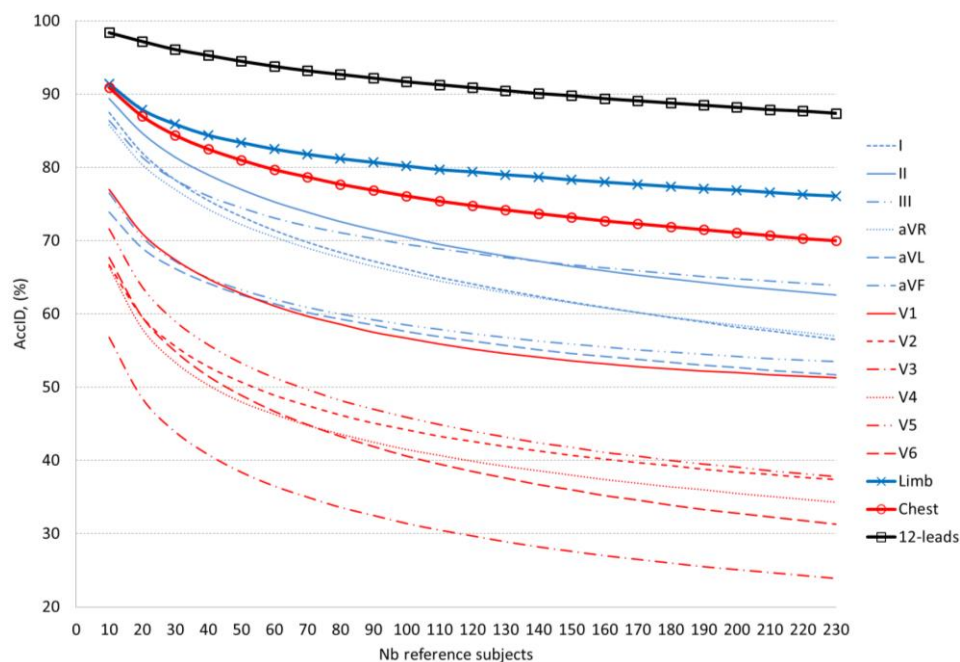
**Table 5.** Validation *AccID* (mean  $\pm$  CI) of the optimal ID models for different feature extraction methods using 12-lead ECG (Table 3). *AccID* is evaluated for different number of reference subjects (10 to 230 subjects included in the validation set). The colour map is provided for visual identification of the ID models' ability to keep the mean *AccID* within different tolerance ranges:  $\geq 90\%$  in dark green,  $[80-90)\%$  in light green,  $[70-80)\%$  in yellow,  $[50-70)\%$  in orange,  $<50\%$  in grey.

Nb Subjects	COR-PQRST	COR-QRS	AMP-QRS	MATCH-QRS	All Features
10	93.7 $\pm$ 2.1	94.1 $\pm$ 2.4	56.2 $\pm$ 4.7	98.0 $\pm$ 0.8	98.4 $\pm$ 0.8
20	90.4 $\pm$ 2.9	91.9 $\pm$ 2.9	44.5 $\pm$ 5.0	96.4 $\pm$ 1.4	97.2 $\pm$ 1.3
30	88.1 $\pm$ 3.3	90.3 $\pm$ 3.2	38.1 $\pm$ 4.9	95.1 $\pm$ 1.9	96.1 $\pm$ 1.7
40	86.3 $\pm$ 3.6	89.1 $\pm$ 3.4	33.7 $\pm$ 4.8	94.0 $\pm$ 2.2	95.3 $\pm$ 2.0
50	84.9 $\pm$ 3.9	88.0 $\pm$ 3.6	30.5 $\pm$ 4.8	93.0 $\pm$ 2.5	94.5 $\pm$ 2.2
60	83.7 $\pm$ 4.1	87.2 $\pm$ 3.8	28.0 $\pm$ 4.7	92.1 $\pm$ 2.7	93.8 $\pm$ 2.4
70	82.7 $\pm$ 4.2	86.5 $\pm$ 3.9	25.9 $\pm$ 4.7	91.3 $\pm$ 2.8	93.2 $\pm$ 2.6
80	81.8 $\pm$ 4.4	85.9 $\pm$ 4.1	24.2 $\pm$ 4.6	90.6 $\pm$ 3.0	92.7 $\pm$ 2.7
90	81.0 $\pm$ 4.5	85.3 $\pm$ 4.2	22.7 $\pm$ 4.6	90.0 $\pm$ 3.2	92.2 $\pm$ 2.9
100	80.2 $\pm$ 4.6	84.8 $\pm$ 4.3	21.5 $\pm$ 4.5	89.4 $\pm$ 3.3	91.7 $\pm$ 3.0
110	79.6 $\pm$ 4.7	84.4 $\pm$ 4.3	20.4 $\pm$ 4.5	88.8 $\pm$ 3.4	91.3 $\pm$ 3.1
120	78.9 $\pm$ 4.8	84.0 $\pm$ 4.4	19.5 $\pm$ 4.5	88.2 $\pm$ 3.6	90.9 $\pm$ 3.3
130	78.4 $\pm$ 4.9	83.6 $\pm$ 4.5	18.7 $\pm$ 4.5	87.7 $\pm$ 3.7	90.5 $\pm$ 3.4
140	77.8 $\pm$ 5.0	83.3 $\pm$ 4.6	18.0 $\pm$ 4.5	87.1 $\pm$ 3.8	90.1 $\pm$ 3.5
150	77.3 $\pm$ 5.1	82.9 $\pm$ 4.6	17.3 $\pm$ 4.5	86.6 $\pm$ 3.9	89.8 $\pm$ 3.6
160	76.8 $\pm$ 5.1	82.6 $\pm$ 4.7	16.7 $\pm$ 4.5	86.2 $\pm$ 4.0	89.4 $\pm$ 3.7
170	76.3 $\pm$ 5.2	82.4 $\pm$ 4.7	16.2 $\pm$ 4.4	85.7 $\pm$ 4.2	89.1 $\pm$ 3.8
180	75.9 $\pm$ 5.3	82.1 $\pm$ 4.8	15.7 $\pm$ 4.4	85.2 $\pm$ 4.3	88.8 $\pm$ 3.9
190	75.5 $\pm$ 5.4	81.8 $\pm$ 4.9	15.3 $\pm$ 4.4	84.8 $\pm$ 4.4	88.5 $\pm$ 3.9
200	75.0 $\pm$ 5.5	81.6 $\pm$ 4.9	14.9 $\pm$ 4.5	84.3 $\pm$ 4.5	88.2 $\pm$ 4.0
210	74.7 $\pm$ 5.5	81.3 $\pm$ 5.0	14.6 $\pm$ 4.5	83.9 $\pm$ 4.6	87.9 $\pm$ 4.1
220	74.3 $\pm$ 5.6	81.1 $\pm$ 5.0	14.2 $\pm$ 4.5	83.5 $\pm$ 4.7	87.7 $\pm$ 4.2
230	73.9 $\pm$ 5.7	80.9 $\pm$ 5.1	13.9 $\pm$ 4.5	83.0 $\pm$ 4.9	87.4 $\pm$ 4.3



**Figure 4.** Mean value of the validation *AccID* in function of the number of reference subjects (data in Table 5). The graphs are presented in respect of the feature extraction methods.

Table 6 and Figure 5 compare the validation performance of the optimal ID models for different lead sets. The accepted *AccID* threshold (>90%) is not satisfied for any single lead, even for just 10 reference subjects. The most prominent single limb and chest leads fulfil reduced requirements for *AccID* thresholds. The top-ranked limb lead II achieves *AccID* > 80% for 30 reference subjects and *AccID* > 70% for 100 reference subjects. The top-ranked chest lead V1 achieves only the minor threshold *AccID* > 70% for 20 reference subjects. The multi-lead sets are more powerful and they satisfy higher *AccID* requirements. Identification accuracy *AccID* > 90% is achieved for 10 reference subjects using either 6 chest leads or 6 limb leads and 140 reference subjects when using 12-lead ECG. A lower threshold of *AccID* > 80% is satisfied for 50 reference subjects using 6 chest leads, 100 reference subjects using 6 limb leads and more than 230 reference subjects (maximal population in this study) using 12-lead ECG.



**Figure 5.** Mean value of the validation *AccID* in function of the number of reference subjects (data in Table 6). The graphs are presented in respect of the ECG lead set implemented in the model: Blue lines: Single limb leads (dashed lines) and all limb leads (solid line); Red lines: Single chest leads (dashed lines) and all chest leads (solid line); Black lines: 12-leads.

**Table 6.** Validation *AccID* (mean  $\pm$  CI, %) of the optimal ID models for single- and multi-lead sets (Table 4). *AccID* is evaluated for different number of reference subjects (10 to 230 subjects included in the validation set). The colour map is provided for visual identification of the ID models' ability to keep the mean *AccID* within different tolerance ranges:  $\geq 90\%$  in dark green,  $[80\text{--}90)\%$  in light green,  $[70\text{--}80)\%$  in yellow,  $[50\text{--}70)\%$  in orange,  $<50\%$  in grey. The bolded values correspond to the top ranked *AccID* across limb leads (II for 1–140 subjects, aVF for 140–230 subjects), chest leads (V1) and multi-lead sets (12-leads).

Nb Subjects	I	II	III	aVR	aVL	aVF	V1	V2	V3	V4	V5	V6	Limb	Chest	12-Leads
10	87.5 $\pm$ 3.1	<b>89.4 <math>\pm</math> 2.9</b>	76.5 $\pm$ 4.5	85.8 $\pm$ 3.4	73.9 $\pm$ 4.9	86.4 $\pm$ 3.4	<b>77.0 <math>\pm</math> 4.5</b>	66.7 $\pm$ 5.1	<b>56.8 <math>\pm</math> 5.3</b>	66.4 $\pm$ 4.9	<b>71.6 <math>\pm</math> 4.6</b>	67.7 $\pm$ 4.9	91.4 $\pm$ 2.8	90.9 $\pm$ 2.7	<b>98.4 <math>\pm</math> 0.8</b>
20	82.0 $\pm$ 3.8	84.7 $\pm$ 3.6	70.5 $\pm$ 5.1	80.4 $\pm$ 4.1	69.0 $\pm$ 5.3	81.4 $\pm$ 4.1	<b>71.0 <math>\pm</math> 5.0</b>	59.6 $\pm$ 5.5	48.5 $\pm$ 5.5	58.0 $\pm$ 5.4	63.7 $\pm$ 5.1	59.6 $\pm$ 5.3	87.9 $\pm$ 3.5	87.0 $\pm$ 3.5	<b>97.2 <math>\pm</math> 1.3</b>
30	<b>78.3 <math>\pm</math> 4.3</b>	<b>81.4 <math>\pm</math> 4.0</b>	67.2 $\pm$ 5.4	77.0 $\pm$ 4.5	66.2 $\pm$ 5.5	78.3 $\pm$ 4.5	<b>67.4 <math>\pm</math> 5.2</b>	55.6 $\pm$ 5.6	43.9 $\pm$ 5.5	53.4 $\pm$ 5.6	59.0 $\pm$ 5.4	54.9 $\pm$ 5.4	85.9 $\pm$ 3.9	84.4 $\pm$ 3.9	<b>96.1 <math>\pm</math> 1.7</b>
40	75.5 $\pm$ 4.6	<b>79.0 <math>\pm</math> 4.3</b>	64.9 $\pm$ 5.6	74.3 $\pm$ 4.7	64.2 $\pm$ 5.6	76.1 $\pm$ 4.8	<b>64.8 <math>\pm</math> 5.4</b>	52.8 $\pm$ 5.7	40.8 $\pm$ 5.5	50.3 $\pm$ 5.7	55.8 $\pm$ 5.5	51.5 $\pm$ 5.5	84.4 $\pm$ 4.2	82.5 $\pm$ 4.2	<b>95.3 <math>\pm</math> 2.0</b>
50	73.3 $\pm$ 4.8	<b>77.0 <math>\pm</math> 4.5</b>	63.3 $\pm$ 5.7	72.2 $\pm$ 4.9	62.6 $\pm$ 5.7	74.5 $\pm$ 5.0	<b>62.8 <math>\pm</math> 5.5</b>	50.7 $\pm$ 5.7	38.4 $\pm$ 5.5	48.0 $\pm$ 5.7	53.3 $\pm$ 5.6	48.9 $\pm$ 5.5	83.4 $\pm$ 4.3	81.0 $\pm$ 4.4	<b>94.5 <math>\pm</math> 2.2</b>
60	71.4 $\pm$ 4.9	<b>75.3 <math>\pm</math> 4.7</b>	62.0 $\pm$ 5.8	70.5 $\pm$ 5.1	61.4 $\pm$ 5.7	73.1 $\pm$ 5.1	<b>61.1 <math>\pm</math> 5.6</b>	48.9 $\pm$ 5.8	36.5 $\pm$ 5.4	46.3 $\pm$ 5.8	51.3 $\pm$ 5.7	46.7 $\pm$ 5.5	82.5 $\pm$ 4.5	<b>79.7 <math>\pm</math> 4.5</b>	<b>93.8 <math>\pm</math> 2.4</b>
70	69.8 $\pm$ 5.1	<b>73.9 <math>\pm</math> 4.9</b>	60.9 $\pm$ 5.8	69.0 $\pm$ 5.2	60.2 $\pm$ 5.8	72.0 $\pm$ 5.3	<b>59.7 <math>\pm</math> 5.7</b>	47.5 $\pm$ 5.8	35.0 $\pm$ 5.4	44.8 $\pm$ 5.8	49.7 $\pm$ 5.7	44.9 $\pm$ 5.6	81.8 $\pm$ 4.6	78.7 $\pm$ 4.7	<b>93.2 <math>\pm</math> 2.6</b>
80	68.4 $\pm$ 5.2	<b>72.6 <math>\pm</math> 5.0</b>	60.0 $\pm$ 5.9	67.7 $\pm$ 5.3	59.3 $\pm$ 5.9	71.1 $\pm$ 5.4	<b>58.6 <math>\pm</math> 5.8</b>	46.2 $\pm$ 5.8	33.6 $\pm$ 5.4	43.6 $\pm$ 5.8	48.2 $\pm$ 5.8	43.3 $\pm$ 5.6	81.2 $\pm$ 4.7	77.7 $\pm$ 4.8	<b>92.7 <math>\pm</math> 2.7</b>
90	67.2 $\pm$ 5.3	<b>71.5 <math>\pm</math> 5.1</b>	59.2 $\pm$ 5.9	66.5 $\pm$ 5.4	58.5 $\pm$ 5.9	70.3 $\pm$ 5.4	<b>57.5 <math>\pm</math> 5.9</b>	45.1 $\pm$ 5.9	32.5 $\pm$ 5.4	42.5 $\pm$ 5.8	47 $\pm$ 5.8	41.9 $\pm$ 5.6	80.7 $\pm$ 4.7	76.9 $\pm$ 4.9	<b>92.2 <math>\pm</math> 2.9</b>
100	66.1 $\pm$ 5.4	<b>70.5 <math>\pm</math> 5.3</b>	58.5 $\pm$ 6.0	65.5 $\pm$ 5.5	57.6 $\pm$ 6.0	69.5 $\pm$ 5.5	<b>56.7 <math>\pm</math> 6.0</b>	44.2 $\pm$ 5.9	31.4 $\pm$ 5.4	41.5 $\pm$ 5.8	45.9 $\pm$ 5.8	40.6 $\pm$ 5.6	80.2 $\pm$ 4.8	76.1 $\pm$ 5.0	<b>91.7 <math>\pm</math> 3.0</b>
110	65.0 $\pm$ 5.5	<b>69.5 <math>\pm</math> 5.4</b>	57.9 $\pm$ 6.0	64.5 $\pm$ 5.6	56.9 $\pm$ 6.0	68.9 $\pm$ 5.6	<b>55.9 <math>\pm</math> 6.0</b>	43.3 $\pm$ 5.9	30.5 $\pm$ 5.4	40.7 $\pm$ 5.9	44.9 $\pm$ 5.9	39.5 $\pm$ 5.7	79.7 $\pm$ 4.9	75.4 $\pm$ 5.1	<b>91.3 <math>\pm</math> 3.1</b>
120	64.1 $\pm$ 5.6	<b>68.7 <math>\pm</math> 5.5</b>	57.3 $\pm$ 6.1	63.7 $\pm$ 5.7	56.3 $\pm$ 6.0	68.3 $\pm$ 5.6	<b>55.2 <math>\pm</math> 6.1</b>	42.6 $\pm$ 6.0	29.7 $\pm$ 5.4	39.9 $\pm$ 5.9	44.0 $\pm$ 5.9	38.5 $\pm$ 5.7	79.4 $\pm$ 4.9	74.8 $\pm$ 5.2	<b>90.9 <math>\pm</math> 3.3</b>
130	63.2 $\pm$ 5.7	<b>67.9 <math>\pm</math> 5.5</b>	56.8 $\pm$ 6.1	62.9 $\pm$ 5.7	55.7 $\pm$ 6.1	67.7 $\pm$ 5.7	<b>54.6 <math>\pm</math> 6.1</b>	41.9 $\pm$ 6.0	28.9 $\pm$ 5.4	39.2 $\pm$ 5.9	43.2 $\pm$ 5.9	37.6 $\pm$ 5.7	79.0 $\pm$ 5.0	74.2 $\pm$ 5.3	<b>90.5 <math>\pm</math> 3.4</b>
140	62.4 $\pm$ 5.7	<b>67.2 <math>\pm</math> 5.6</b>	56.3 $\pm$ 6.2	62.2 $\pm$ 5.8	55.1 $\pm$ 6.1	67.2 $\pm$ 5.8	<b>54.1 <math>\pm</math> 6.2</b>	41.3 $\pm$ 6.0	28.2 $\pm$ 5.4	38.6 $\pm$ 5.9	42.4 $\pm$ 6.0	36.7 $\pm$ 5.7	78.7 $\pm$ 5.0	73.7 $\pm$ 5.3	<b>90.1 <math>\pm</math> 3.5</b>
150	61.6 $\pm$ 5.8	66.5 $\pm$ 5.7	55.9 $\pm$ 6.2	61.5 $\pm$ 5.9	54.6 $\pm$ 6.2	66.7 $\pm$ 5.8	<b>53.6 <math>\pm</math> 6.2</b>	40.7 $\pm$ 6.0	27.6 $\pm$ 5.4	38.0 $\pm$ 5.9	41.8 $\pm$ 6.0	36.0 $\pm$ 5.8	78.3 $\pm$ 5.1	73.2 $\pm$ 5.4	<b>89.8 <math>\pm</math> 3.6</b>
160	60.9 $\pm$ 5.9	65.9 $\pm$ 5.8	55.5 $\pm$ 6.2	60.8 $\pm$ 5.9	54.2 $\pm$ 6.2	66.3 $\pm$ 5.9	<b>53.2 <math>\pm</math> 6.3</b>	40.2 $\pm$ 6.1	27.0 $\pm$ 5.4	37.4 $\pm$ 6.0	41.1 $\pm$ 6.0	35.2 $\pm$ 5.8	78.0 $\pm$ 5.1	72.7 $\pm$ 5.5	<b>89.4 <math>\pm</math> 3.7</b>
170	60.2 $\pm$ 6.0	65.3 $\pm$ 5.9	55.1 $\pm$ 6.3	60.2 $\pm$ 6.0	53.8 $\pm$ 6.3	65.9 $\pm$ 5.9	<b>52.8 <math>\pm</math> 6.3</b>	39.7 $\pm$ 6.1	26.5 $\pm$ 5.4	36.9 $\pm$ 6.0	40.6 $\pm$ 6.1	34.6 $\pm$ 5.8	77.7 $\pm$ 5.2	72.3 $\pm$ 5.5	<b>89.1 <math>\pm</math> 3.8</b>
180	59.5 $\pm$ 6.0	64.8 $\pm$ 5.9	54.8 $\pm$ 6.3	59.6 $\pm$ 6.1	53.4 $\pm$ 6.3	65.5 $\pm$ 6.0	<b>52.5 <math>\pm</math> 6.3</b>	39.3 $\pm$ 6.1	26.0 $\pm$ 5.4	36.4 $\pm$ 6.0	40.0 $\pm$ 6.1	33.9 $\pm$ 5.8	77.4 $\pm$ 5.3	71.9 $\pm$ 5.6	<b>88.8 <math>\pm</math> 3.9</b>
190	58.9 $\pm$ 6.1	64.3 $\pm$ 6.0	54.5 $\pm$ 6.3	59.0 $\pm$ 6.1	53.0 $\pm$ 6.3	65.1 $\pm$ 6.0	<b>52.2 <math>\pm</math> 6.4</b>	38.8 $\pm$ 6.1	25.5 $\pm$ 5.4	36.0 $\pm$ 6.0	39.5 $\pm$ 6.1	33.3 $\pm$ 5.9	77.1 $\pm$ 5.3	71.5 $\pm$ 5.7	<b>88.5 <math>\pm</math> 3.9</b>
200	58.2 $\pm$ 6.2	63.8 $\pm$ 6.1	54.2 $\pm$ 6.4	58.5 $\pm$ 6.2	52.7 $\pm$ 6.4	64.8 $\pm$ 6.1	<b>52.0 <math>\pm</math> 6.4</b>	38.4 $\pm$ 6.2	25.1 $\pm$ 5.5	35.5 $\pm$ 6.1	39.1 $\pm$ 6.2	32.8 $\pm$ 5.9	76.9 $\pm$ 5.4	71.1 $\pm$ 5.7	<b>88.2 <math>\pm</math> 4.0</b>
210	57.7 $\pm$ 6.3	63.4 $\pm$ 6.1	53.9 $\pm$ 6.4	58.0 $\pm$ 6.3	52.3 $\pm$ 6.4	64.5 $\pm$ 6.1	<b>51.7 <math>\pm</math> 6.4</b>	38.1 $\pm$ 6.2	24.7 $\pm$ 5.5	35.1 $\pm$ 6.1	38.6 $\pm$ 6.2	32.3 $\pm$ 5.9	76.6 $\pm$ 5.4	70.7 $\pm$ 5.8	<b>87.9 <math>\pm</math> 4.1</b>
220	57.1 $\pm$ 6.3	63.0 $\pm$ 6.2	53.7 $\pm$ 6.4	57.5 $\pm$ 6.3	52.0 $\pm$ 6.4	64.2 $\pm$ 6.2	<b>51.5 <math>\pm</math> 6.4</b>	37.7 $\pm$ 6.2	24.3 $\pm$ 5.5	34.7 $\pm$ 6.1	38.2 $\pm$ 6.2	31.8 $\pm$ 6.0	76.3 $\pm$ 5.5	70.3 $\pm$ 5.9	<b>87.7 <math>\pm</math> 4.2</b>
230	56.5 $\pm$ 6.4	62.6 $\pm$ 6.3	53.5 $\pm$ 6.5	57.0 $\pm$ 6.4	51.7 $\pm$ 6.5	63.9 $\pm$ 6.2	<b>51.3 <math>\pm</math> 6.5</b>	37.4 $\pm$ 6.3	23.9 $\pm$ 5.5	34.3 $\pm$ 6.2	37.8 $\pm$ 6.3	31.3 $\pm$ 6.0	76.1 $\pm$ 5.5	70.0 $\pm$ 5.9	<b>87.4 <math>\pm</math> 4.3</b>



#### 4. Discussion

Although being a part of one general biometric task, the aims of human “identification” and human “verification” are quite different, i.e. to recognize an individual in a database by comparing a single input sample to many stored templates and to verify/reject person’s identity by comparing a single input sample to a single stored template. The same features for comparison between input-to-template waveforms could be used in both sub-tasks, however, the decision-making process is different. Usually, the decision for “identification” is performed via the best similarity assessment approach, while the “verification” relies on threshold based techniques. In our recent study [29] we solve the “verification” problem via 6 cross-correlation and 2 amplitude features of QRS and PQRS patterns, involved in a linear discriminant threshold decision function. In this study, we use the same large population to solve the “identification” task adopting 4 cross-correlation and 1 amplitude features and adding 2 more QRS pattern features that measure the similarity in the time and amplitude scales. They are calculated with a brand-new method for binary template matching of short-duration QRS patterns (100 ms). The presented identification model applies a normalized scale for representation of all 7 features (0–100%), where 0% corresponds to the least matching and 100% to the best matching between input and template waveforms. The optimization of the feature selection scheme involved in the calculation of the common similarity index is presented by iterative maximization of the identification accuracy *AccID*.

The questions of greatest concern for the human identification task are: “*Is there an individual whose personalized beat pattern would be more similar to another than to his own in a long-term basis?*”, “*What is the probability to find such an individual within an increasing population of reference subjects?*”

A quantitative answer to these questions is presented in this study following a straightforward list of training and validation tests which aim to reveal specific aspects of the current human ID investigation.

(1) *Is there a statistical justification for using the defined features for waveform similarity?*

Table 1 presents the statistics of the features’ distributions in the training dataset, proving that all of them exhibit significantly higher similarity of the waveforms measured across the same individuals in a long-term basis (>1 year) than what can be seen between different individuals (mean value in the range about 82–96% vs. 60–86%,  $p < 0.0001$ ). This proves that the similarity index  $ID_{ix}$  defined for finding the best match between subjects could include all features on the same basis.

(2) *Which are the most reliable QRS and PQRS features for human ID?*

This aspect is investigated by training optimal ID models with different configurations of features using a large dataset with 230 individuals. The simplest model (one-feature in one-lead, Table 2) distinguishes  $r(lag0)$  in leads (aVL, aVF, III) with maximal *AccID* in the range of 35 to 41%. It is about 10% points better than the other distinguished feature  $r(max)$  where *AccID* ranges between 24 and 35%. The advance of both correlation coefficients is due to different spatial dynamics of the cardiac vector between different subjects. This dissimilarity is enhanced in the zero-lag correlation which additionally represents the temporal desynchronization of the depolarization-repolarization process in different individuals. We suggest that the latter effect is compensated in aVR by using this lead for synchronization of each subject’s average beat to the reference pattern and this explains the remarkably low *AccID* in aVR (11–15%) in respect to the other limb leads. The joint analysis of  $r(lag0)$  and  $r(max)$  in 12-lead ECG increases *AccID* by about 40% reaching 73.5% (COR-QRS) and 77.4% (COR-PQRS)—see Table 3. These results indicate that the QRS pattern (100 ms) carry the essential subject-specific information, while P and T waves in the PQRS pattern (500 ms) contribute only for a slight ID accuracy improvement by about 4 percentage points.

The analysis of AMP-QRS shows definitively that the QRS peak-to-peak range across individuals is an indistinguishable feature. It provides *AccID* < 1% for all single leads (Table 2), which increases up



to 24% for the optimal ID model (Table 3) including the 7 ECG leads that are the most distinguishable in amplitude (I, II, III, V1, V2, V3, V5).

The MATCH-QRS method is not powerful when its features are analysed in single leads (maximal accuracy of 26.5% for EQUA in aVF, Table 2). This is suggested by the methodological approach for QRS waveform approximation in *binQRS* matrix [ $100 \times 80$ ], which gives indistinguishably high waveform similarity for many subjects whose QRS waveforms are matching within the pre-set time and amplitude tolerances of  $\pm 1$  ms and  $\pm 5\%$ , respectively. The accumulated similarity over 12 leads, however stays high for the same identity subjects and decreases for different identity subjects. Thus, 15 MATCH-QRS measures are selected in the best ID model (*AccID* = 83.5%, Table 3), which represents the best similarity for 10 leads in the amplitude scale (I, II, III, aVR, aVF, V1, V2, V4, V5, V6) and for 5 leads in the time scale (I, aVR, aVL, V1, V3).

### (3) Which are the most reliable leads for human ID?

This aspect is investigated by training optimal ID models with different configurations of leads on a large dataset with 230 individuals (Table 4). In a single lead ID scenario, we highlight limb leads I and II (maximal *AccID* of 58%) followed by aVF (55%) and aVL (50%) including 3 to 5 features from all feature extraction methods. Compared to the best limb leads (I and II), the single chest leads are from 14 to 30% less powerful for the purposes of human ID. They are ranked in the following order: V1 (maximal *AccID* of 44%) followed by V6, V5, V2 (37–35%) and V3, V4 (27–30%). Such a drop in accuracy, especially in anterior leads, could be explained by V1–V6 proximity to the heart, so that small electrode misplacements across different recording sessions of the same individual result in considerable ECG morphology changes seen in chest leads.

Multi-lead ID models gain considerable raise in accuracy (about 20%) comparing 6 limb leads to the best II (77% vs. 58%) and 6 chest leads to the best V1 (64% vs. 44%). The optimal ID models are built with a prevalent selection of features from leads I, II (6 features), V1, V6 (5 features), aVR, V5 (4 features). The complete ID model, selecting features from all 12-leads, reaches maximal *AccID* of 91.3%. This is better than 6 limb, 6 chest leads and the best single lead (II) with about 14%, 27% and 33% points, respectively.

### (4) What is the tolerated reference database size for searching human ID?

This question arises from our previous experience on human identification [23] showing that the time elapsed between the collection of the reference and the test dataset as well as the number of enrolled subjects in the reference dataset have a severe impact on the ID performance. There, *AccID* has been found to drop by 15% (from 93% to 78%) while the population increases by 35 subjects (from 14 to 49). The last is in agreement with [28], reporting *AccID* drop by 10% (from 99% to 89%) for population increase by 40 subjects (from 10 to 50). Naturally, *AccID* worsening could be expected for reference database with larger size, since the more the subjects there are, the higher the random match probability is, i.e. to find a subject with better similarity than his own. According to a comparative study [1], the majority of the published investigations on human ID have been conducted on small populations (about a few dozens of subjects). Considering the missing research information about the influence of the reference database size on the ID accuracy, we present consistent validation of our optimal ID models on an independent dataset by increasing its size from 10 to 230 subjects. Our validation results definitively confirm the expected trend for *AccID* drop with the increase of the number of the reference subjects which however is non-linear and depends on the particular ID model (Figures 4 and 5). We consider as good models those with higher *AccID* and lower *AccID* drop in respect of the number of reference subjects.

The validation of the ID models applying different feature extraction methods (Table 5, Figure 4) closely reproduces the *AccID* ranges on the training dataset with 230 reference subjects (Table 3). The observed *AccID* differences (training vs. validation) are: COR-PQRST (77% vs. 74%), COR-QRS (74% vs. 81%), AMP-QRS (24% vs. 14%), MATCH-QRS (84% vs. 83%), all features (91% vs. 87%).

The differences are relatively small which is a sign of confident training of the ID models with 12-lead ECG capable of adequately evaluating independent database without bias. We observed that the minimal set of 2 MATCH-QRS features is capable of providing human ID with almost the same accuracy ( $AccID > 90\%$ ) as the maximal set of all 7 features. However, this is valid for limited reference datasets of up to 100 subjects. More subjects are best identified by the complete ID model including all features which maintains relatively stable performance with a drop of 11% (from 98% to 87%) for population increase by 220 subjects (from 10 to 230). These results outperform the cited in [23,28] which report the same level of  $AccID$  drop (10–15%) for about 6 times smaller population increase (35–40 subjects). The best performance of our complete ID model could be explained by the fact that larger datasets increase the probability to find similar characteristics across different subjects and thus, more and more robust features are required to distinguish their individuality.

In a single lead ID scenario (Figure 5, Table 6) we can distinguish two best ID models based on analysis of lead II and aVF for small (10–140) and large (140–230) number of reference subjects, respectively. Overall, even for the best limb leads (II, aVF), we observe a limited ID capability, not exceeding 70% for more than 100 reference subjects and 80% for more than 30 reference subjects. Single chest leads are not usable for human ID. The majority of them (V2–V6) do not exceed 50% for more than 50 subjects. Only V1 can be distinguished with  $AccID$  above 70% for less than 20 subjects. We recommend multi-lead ID scenarios for management of human search in large populations. Both ID models based on 6 chest and 6 limb leads pass the  $AccID > 70\%$  threshold for the maximal reference dataset (230 subjects) and  $AccID > 80\%$  for up to 100 subjects (limb leads) and up to 50 subjects (chest leads). As expected, 12-lead ECG provides the most robust beat patterns for human ID with  $AccID > 90\%$  for up to 140 subjects, reaching 98.4% for 10 subjects.

The performance of our best ID model (12-leads, 27 features, trained in Table 4, validated in Tables 5 and 6; Figures 4 and 5) is compared to other published studies on human identification (Table 7, Figure 6).

All published studies do not report training and validation of their ID models on independent datasets (as used in this study), therefore biased  $AccID$  is probable due to overtraining. For example, such an effect is suspected for studies with accuracy close to 100%, especially on small sized databases [9,16,18–20,26]. Overall, the ID studies with enrolment and test phase in a single day report higher  $AccID$  [9,10,18,26,28]. However, they could not consider longitudinal ECG morphology changes due to physiological or technical sources (e.g. electrode displacement across different sessions). There are two ID studies using multiple sessions in different days on large populations with 74 [24] and 89 subjects [10] that report less than 5 percentage points better performance than our model. This might be due either to superior model algorithm or to the lack of independent validation, considering that the differences between our training and validation accuracies are in the same range. Figure 6 illustrates validation results of this study which extends the knowledge on the ID model performance over the largest span from 10 to 230 reference subjects. The observed decrease of  $AccID$  could be explained with the appearance of new subjects in the reference dataset who present PQRST waveforms similar to already existing ones. Thus, some of the subjects that match exactly their own pattern in the reference dataset of e.g. 100 subjects show higher correlation and pattern matching with different subject in the reference dataset that contains 230 individuals. This could be due to time-related ECG changes and possible small differences in the electrode placement during the acquisition of the test dataset. Possible approach to keep high  $AccID$  for larger reference datasets is to include additional robust features in the identification model. However, this should be done very carefully and the designed models should be validated on independent dataset in order to avoid overtraining. Another option to prevent from the assignment of wrong identity to the tested subject is to keep the size of the reference dataset as small as possible for the particular application and to update it with actual PQRST patterns on a regular basis.



containing two ECGs per patient recorded between 1 to 2 years apart. The benefits from the use of such database are its representativeness for physiologically related long-term ECG changes and effects of possible multi-session technical differences. The investigations demonstrate the identification ability of cross-correlation, amplitude and pattern matching, applied on a personalized heartbeat pattern in single and multi-lead ECG scenarios. The detailed analysis of the identification accuracy highlights the ID model based on cross-correlation and pattern matching of 12-lead ECG as the best one, providing 91.3% identification accuracy over a large training dataset of 230 subjects. The independent validation of this ID model on datasets with different sizes (10–230 subjects) confirms its reliability for human identification with accuracy between 98.4% and 87.4% and extends the knowledge on the ID model performance over a large span of reference subjects.

## 6. Future Work

Up to this moment, our experience in human verification/identification involves the application of fiducial-based and fiducial independent approaches with databases collected with clinical ECG devices. Our future work would revolve around the design of hybrid methods for human verification/identification and their implementation in an identification ECG device prototype, working with a pre-selected minimal lead set.

**Acknowledgments:** This study is supported by Bulgarian National Science Fund, grant No. T02/11.

**Author Contributions:** Conceived and designed the experiments: Irena Jekova (I.J.), Vessela Krasteva (V.K.). Performed the experiments: I.J., V.K. Analyzed the data: I.J., V.K. Contributed materials/analysis tools: Ramun Schmid (R.S.). Wrote the paper: I.J., V.K., R.S.

**Conflicts of Interest:** The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript and in the decision to publish the results.

## References

- Odinaka, I.; Lai, P.H.; Kaplan, A.; O’Sullivan, J.; Sirevaag, E.; Rohrbaugh, J. ECG biometric recognition: A comparative analysis. *IEEE Trans. Inf. Forensics Secur.* **2012**, *7*, 1812–1824. [[CrossRef](#)]
- Fong, S.; Lan, K.; Wong, R. Classifying human voices by using hybrid SFX time-series pre-processing and ensemble feature selection. *Biomed. Res. Int.* **2013**, *2013*, 720834. [[CrossRef](#)] [[PubMed](#)]
- Arigbabu, O.; Ahmad, S.; Adnan, W.; Yussof, S. Integration of multiple soft biometrics for human identification. *Pattern Recognit. Lett.* **2015**, *68*, 278–287. [[CrossRef](#)]
- Ju, S.; Seo, H.; Han, S.; Ryou, J.; Kwak, J. A study on user authentication methodology using numeric password and fingerprint biometric information. *Biomed. Res. Int.* **2013**, *2013*, 427542. [[CrossRef](#)] [[PubMed](#)]
- Drahansky, M.; Dolezel, M.; Vana, J.; Brezinova, E.; Yim, J.; Shim, K. New optical methods for liveness detection on fingers. *Biomed. Res. Int.* **2013**, *2013*, 197925. [[CrossRef](#)] [[PubMed](#)]
- Procházka, A.; Schätz, M.; Vyšata, O.; Vališ, M. Microsoft Kinect Visual and Depth Sensors for Breathing and Heart Rate Analysis. *Sensors* **2016**, *16*, 996. [[CrossRef](#)] [[PubMed](#)]
- Charvátová, H.; Procházka, A.; Vaseghi, S.; Vyšata, O.; Vališ, M. GPS-based Analysis of Physical Activities Using Positioning and Heart Rate Cycling Data. *Signal Image Video Process.* **2017**, *11*, 251–258. [[CrossRef](#)]
- Nawal, M.; Purohit, G. ECG Based Human Authentication: A Review. *Int. J. Emerg. Eng. Res. Technol.* **2014**, *2*, 178–185.
- Tantawi, M.; Revett, K.; Salem, A.; Tolba, M. Fiducial feature reduction analysis for electrocardiogram (ECG) based biometric recognition. *J. Intell. Inf. Syst.* **2013**, *40*, 17–39. [[CrossRef](#)]
- Tan, R.; Perkowski, M. Toward improving electrocardiogram (ECG) biometric verification using mobile sensors: A two-stage classifier approach. *Sensors* **2017**, *17*, 410. [[CrossRef](#)] [[PubMed](#)]
- Pirbhulal, S.; Zhang, H.; Mukhopadhyay, S.C.; Li, C.; Wang, Y.; Li, G.; Wu, W.; Zhang, Y.T. An efficient biometric-based algorithm using heart rate variability for securing body sensor networks. *Sensors* **2015**, *15*, 15067–15089. [[CrossRef](#)] [[PubMed](#)]
- Peter, S.; Reddy, B.P.; Momtaz, F.; Givargis, T. Design of Secure ECG-Based Biometric Authentication in Body Area Sensor Networks. *Sensors* **2016**, *16*, 570. [[CrossRef](#)] [[PubMed](#)]

13. Labati, R.D.; Piuri, V.; Sassi, R.; Sforza, G.; Scotti, F. Adaptive ECG biometric recognition: A study on re-enrollment methods for QRS signals. In Proceedings of the IEEE Workshop on Computational Intelligence in Biometrics and Identity Management (CIBIM'2014), Orlando, FL, USA, 9–12 December 2014; pp. 30–37.
14. Pinto, J.R.; Cardoso, J.S.; Lourenço, A.; Carreiras, C. Towards a continuous biometric system based on ECG signals acquired on the steering wheel. *Sensors* **2017**, *17*, 2228. [[CrossRef](#)] [[PubMed](#)]
15. Ortega-Garcia, J.; Bigun, J.; Reynolds, D.; Gonzalez-Rodriguez, J. Authentication gets personal with biometrics. *IEEE Signal Process. Mag.* **2004**, *3*, 50–62. [[CrossRef](#)]
16. Biel, L.; Pettersson, O.; Philipson, L.; Wide, P. ECG analysis: A new approach in human identification. *IEEE Trans. Instrum. Meas.* **2001**, *50*, 808–812. [[CrossRef](#)]
17. Kyoso, M.; Uchiyama, A. Development of an ECG identification system. In Proceedings of the 23rd Annual EMBS International Conference, Istanbul, Turkey, 25–28 October 2001.
18. Israel, S.; Irvine, J.; Cheng, A.; Wiederhold, M.; Wiederhold, B. ECG to identify individuals. *Pattern Recognit.* **2005**, *38*, 133–142. [[CrossRef](#)]
19. Plataniotis, K.; Hatzinakos, D.; Lee, J. ECG biometric recognition without fiducial detection. In Proceedings of the 2006 Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference, Baltimore, MD, USA, 19–21 September 2006.
20. Agrafioti, F.; Hatzinakos, D. Fusion of ECG sources for human identification. In Proceedings of the 3rd International Symposium on Communications, Control and Signal Processing, St Julians, Malta, 12–14 March 2008; pp. 1542–147.
21. Poree, F.; Bansard, J.; Kervio, G.; Carrault, G. Stability analysis of the 12-lead ECG morphology in different physiological conditions of interest for biometric applications. *Comput. Cardiol.* **2009**, *36*, 285–288.
22. Poree, F.; Gallix, A.; Carrault, G. Biometric identification of individuals based on the ECG. Which conditions? *Comput. Cardiol.* **2011**, *38*, 761–764.
23. Jekova, I.; Bortolan, G. Personal verification/identification via analysis of the peripheral ECG leads. Influence of the personal health status on the accuracy. *Biomed. Res. Int.* **2015**, *2015*, 135676. [[CrossRef](#)] [[PubMed](#)]
24. Wubbeleer, G.; Stavridis, M.; Kreiseler, D.; Bousselet, R.D.; Elster, C. Verification of humans using the electrocardiogram. *Pattern Recognit. Lett.* **2007**, *28*, 1172–1175. [[CrossRef](#)]
25. Chen, C.; Lin, C.; Chio, Y. Individual identification based on chaotic electrocardiogram signals. *Conf. Ind. Electron. Appl.* **2011**, *6*, 1765–1770.
26. Wang, Y.; Agrafioti, F.; Hatzinakos, D.; Plataniotis, K. Analysis of human electrocardiogram for biometric recognition. *EURASIP J. Adv. Signal Process.* **2008**, *2008*, 148658. [[CrossRef](#)] [[PubMed](#)]
27. Lourenco, A.; Silva, H.; Fred, A. Unveiling the biometric potential of finger-based ECG signals. *Comput. Intell. Neurosci.* **2011**, *2011*, 720971.
28. Zokaee, S.; Faez, K. Human identification based on electrocardiogram and palmprint. *Int. J. Electr. Comput. Eng.* **2012**, *2*, 261–266.
29. Krasteva, V.; Jekova, I.; Abächerli, R. Biometric verification by cross-correlation analysis of 12-lead ECG patterns: Ranking of the most reliable peripheral and chest leads. *J. Electrocardiol.* **2017**, *50*, 847–854. [[CrossRef](#)] [[PubMed](#)]
30. Jekova, I.; Krasteva, V.; Leber, R.; Schmid, R.; Twerenbold, R.; Müller, C.; Reichlin, T.; Abächerli, R. Intersubject variability and intrasubject reproducibility of 12-lead ECG metrics: Implications for human verification. *J. Electrocardiol.* **2016**, *49*, 784–789. [[CrossRef](#)] [[PubMed](#)]

