ECG-based Human Authentication using High-level Spectro-temporal Signal Features

Sara S. Abdeldayem and Thirimachos Bourlai
Lane Department of Computer Science and Electrical Engineering
West Virginia University, Morgantown, WV 26506 USA
ssa0006@mix.wvu.edu, Thirimachos.Bourlai@mail.wvu.edu

Abstract*—Electrocardiography (ECG) is the process of recording the electrical activity of the human heart over time using electrodes that are placed over the skin. While the primary usage of electrocardiograms, the recorded signals, has been focused on the check of signs of heart-related diseases, recent studies have moved also toward their usage for human authentication. Thus, an ECG signal can be unique enough to be used independently as a biometric modality. In addition to its inherent liveness detection, it is easy to collect and can be easily captured either via sensors attached to the human body (fingertips, chest, wrist) or even passively using wireless sensors. In this paper, we propose a novel approach that exploits the spectro-temporal dynamic characteristics of the ECG signal to establish personal recognition system using both short-time Fourier transform (STFT) and generalized Morse wavelets (CWT). This process results in enriching the information extracted from the original ECG signal that is inserted in a 2D convolutional neural network (CNN) which extracts higher level and subject-specific ECG-based features for each individual. To validate our proposed CNN model, we performed nested cross-validation using eight different ECG databases. These databases are considered challenging since they include both normal and abnormal heartbeats as well as a dynamic number of subjects. Our proposed algorithms yield superior performance when compared to other state-of-art approaches discussed in the literature, i.e. the STFT-based one achieves an average identification rate, equal error rate (EER), and area under curve (AUC) of 97.86%, 0.0268, and 0.9933 respectively, whereas the CWT achieves comparable to STFT results in 97.5%, 0.0386, and 0.9882 respectively.

***Keywords*-Electrocardiogram (ECG); Biometrics; Human Identification; Spectro-temporal Features; Convolutional neural networks.**

I. Introduction

Automatic, well founded and accurate human recognition is required in various fields such as civilian applications, surveillance, healthcare and financial information protection in order to allow the confidentiality, integrity and access to only legitimate users. Traditional procedures involved ID cards, tokens, or passwords. However, these procedures are vulnerable to identification theft and circumvention. Biometric-based systems on the other hand utilize the intrinsic properties of the individual, which is not easy to falsify depending on the biometric modality we use. Thus, biometric-based approaches have gained popularity in the last decades utilizing different traits such as fingerprints, face, iris, hand geometry, ear patterns, or speech.

Electrocardiography (ECG) is a unique biometric trait and has been studied for human identity recognition before [1], [2]. The advantage of this trait is that the ECG signals are more difficult to counterfeit when compared to other modalities. For instance, an iris recognition system can be compromised by using different types of attacks including the usage of fake iris images or lenses [3]. Although the intrinsic physiology of the heart, and accordingly the ECG signal is difficult to be mimicked, some work in ECG system attacks has been emerged such as Eberz et al. work [4]. The ECG signal also provides additional information about an individual health status (e.g. normal or arrhythmic beats as well as mental and emotional status), and hence, it can be used also for automated or semi-automated diagnosis of heart condition using machine learning algorithms [5].

Nevertheless, there are some factors that can affect the intra-variability of the ECG biometric system which can be considered as a challenge. These factors are either long-term or short-term. The long-term factors include the age and health condition, whereas the short-term factors include the emotional and mental conditions.

A. Related Work

Several researches have been proposed in the ECG biometrics. These researches can be categorized into three types depending on the features extracted from the ECG signal: fiducial-based approaches, non-fiducial approaches, and hybrid approaches.

The fiducial-based approaches employ the characteristics of the fiducial points of the ECG signal as features in the recognition system. This can be done by extracting either all or part of the characteristic points, which might be prone to noise. The features involved may include the amplitude, angle, or the duration of the wave points. For instance, Irvine et al. [2] studied the effect of the emotional and mental state variation on the ECG identification system performance. The involved fiducial features included the L'P', S'T', and QT intervals yielding a total of fifteen feature. Their results showed that ECG can be used in the biometric process. Using the same database, Israel et al.

Table I. The proposed approach compared to the literature work.

Authors	Features	Feature Selection	Classifier	Leads Count	Abnormality	Databases Count
Bassiouni et al. [6]	Temporal &Spectral	Yes	NN	1	No	1
Shen et al. [7]	Temporal	No	DBNN	1	No	1
Dar et al. [8]	Spectral	Yes	RF	1	No	3
Agrafioti et al. [9]	Temporal	Yes	KNN	1	No	3
Venkatesh et al. [10]	Temporal	No	KNN+FLDA	1	No	1
Camara et al. [11]	Spectral& Temporal	Yes	KNN	2	No	1
Fatemian et al. [12]	Temporal	Yes	Template matching	1	No	1
Palaniappan et al. [13]	Temporal	No	NN+SFA	2	No	1
Tantawi et al. [14]	Spectral	Yes	RBF NN	1	No	3
Sidek et al. [15]	Temporal	No	BN,MLP,RBF,KNN	1	Yes	3
Poree et al. [16]	Temporal	No	Template matching	12	No	1
Zhang et al. [17]	Spectral	No	1D-CNN	1	Yes	8
Our Approach	Spectro-Temporal	No	2D-CNN	1	Yes	8

[18] used fifteen fiducial features that were extracted with respect to the R peaks. Their algorithm achieved 82% and 79% heartbeat identification rates using different ECG lead location (i.e. base of the neck and fifth intercostal spacing), and average accuracies of 80.1% and 64.5% using different anxiety states. Zhang and Wei [19] on the other hand, used fiducial features that included the ECG characteristic points' amplitude and duration as well as fiducial intervals such as QRS and PR intervals achieving accuracies of 79%, and 85.3% using different lead configurations.

In the non-fiducial-based approaches, the features extracted can be either from the time domain or the frequency domain. For instance, Hejazi et al. [20] proposed an ECG-based biometric system that utilizes the autocorrelation along with linear dimension reduction. The cohort of the study involved 52 subjects. The signals are first denoised using the discrete wavelet transform (DWT). Several feature reduction methods including linear discriminant analysis and principle component analysis (PCA), as well as kernel principle component analysis (KPCA) were tested. The reduced features are then used to feed an support vector machine (SVM) classifier to identify the subjects. The results showed that using the Gaussian KPCA as a feature reduction method for the autocorrelation coefficients achieved a lowest false non-matching rate of 4.83% in case of one session recording and 2.297% in case of two sessions. Camara et al. [11] used a dataset of 18 subjects' ECG signals for identification. The Hadamard transform [21], [22] was applied to the two leads' ECG signals, where only 24 coefficients were used per lead. A K-nearest neighborhood classifier (KNN) was used in the classification process for simplicity, where the average accuracy obtained for the 48 coefficient features was 94.2% and 96.6% by incorporating the two entropies proposed in their work. Moreover, Zhang et al. [17] used deep learning in their HeartID ECG identification system. The wavelet transform is done on two-second ECG segments and the autocorrelation coefficients were obtained for the transformed segments. A one-dimensional convolution neural network was then used to identify the individuals. Different wavelet components combinations were tested on the identification rate, and an average rate of 93.5% was reported. There have been also other approaches that utilize the discrete cosine transform or autocorrelation coefficients [6], [23], [24].

On the other hand, the hybrid approaches involve the usage of both fiducial and non-fiducial features. For instance, Wang et al. [25] utilized the P, Q, R, S and T positions and amplitudes as fiducial features. Whereas the non-fiducial features involved the autocorrelation coefficients and discrete cosine transform.

We believe that prior work in the literature work does not address the utilization of deep learning and the spectro-temporal features of the ECG signal as illustrated in Table I. Although Camara et al. [11] and Bassiouni et al. [6] utilized the frequency and time domain features, they only used separate features from each domain. In other words, they studied the variations in the frequency and the time domains, but not the variations of frequency with time. On the other hand, Zhang et al. [17] used both deep learning and wavelets-based approach, but different combination of scales had to be selected before the classification takes place base on the value of the final identification rate. Moreover, Pouryayevali et al. [26] utilized the wavelet of the ECG signal, however, linear discriminant analysis procedure was used for feature selection [26].

To the best of our knowledge, there is no approach that utilizes the high-level features of the spectro-temporal information of the ECG signal. In this paper, we propose a biometric-based approach that utilizes the ECG signal for human identification. We enrich the extracted features by investigating the short-time Fourier transform (STFT) and

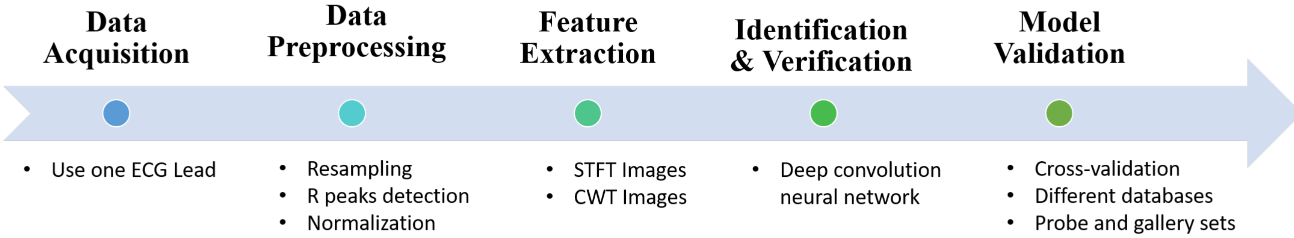


Figure 1. Workflow of the proposed approach.

continuous wavelet transform, namely generalized Morse wavelet, (CWT) of the extracted ECG segments. Then, the high-level features are extracted from the spectro-temporal domain by utilizing a 2D deep convolution neural network (CNN). At the end of the study we determine which approach is the most efficient in terms of a set of performance metrics including rank-1 identification rate, equal error rate (EER), and area (AUC) under the receiver operating curve (ROC).

B. Contributions

Based on our previous work that utilized the spectro-temporal texture in arrhythmia detection [5], we propose an ECG-based biometric approach for human identification. The contributions of this work are as follows:

- We utilize the spectro-temporal domain features using short-time Fourier transform and generalized Morse wavelets to gain more information about the ECG signal patterns by investigating the changes in the frequency components along the time.
- Unlike the literature work that uses deep networks to classify subjects, we employ a deep learning-based approach by using a 2D convolutional neural network model to extract the intermediate layers' high-level features of the spectro-temporal images of the ECG signals. This results in a more discriminative features that can be further used in an open-set biometric system.
- The size of the high-level feature template vector for each subject is a main concern in terms of the memory that is needed in the database or the matching time. We provide a feature compression by employing a distinctive feature vector of size 100 to represent a heartbeat of size 200 sample point. Therefore, instead of storing the whole template ECG signal for each subject, we only store the feature vectors. This will consume less memory and matching time and complexity.
- For model's robustness and generalization assessment, we test our approach on eight ECG databases. These databases encompass both normal and abnormal ECG beats, where the abnormality includes different heart conditions such as atrial or ventricular fibrillation, ventricular flutter, myocardial infarction, etc. Moreover, one of the databases have large number of subjects (i.e. 290 subjects) compared to other databases (i.e. average of 25 subject). To the best of our knowledge no other papers provided the same size of test databases, where the closest paper is [17] as they use the same number of databases. However, we propose larger number of subjects specially in PTB database which can mimic the real-world scenario of having a large-scale system.

The rest of the paper is organized as follows: Section 2 discusses the details of our proposed approach. Then, we analyze the experimental results in Section 3, and conclude in Section 4.

II. Methodology

The workflow of the ECG biometric system entails five main steps: 1) data acquisition, 2) data preprocessing, 3) feature extraction, 4) human identification, 5) model validation and testing. Our workflow of the proposed approach is illustrated in figure 1. What follows is a discussion on each step of the proposed approach starting from data acquisition.

A. Data Acquisition

In the diagnostic electrocardiogram, the ECG signals are usually obtained from 12-lead configuration system including the bipolar limb leads (Leads I, II, and III) that are measured by placing the electrodes on the patients' limbs, the augmented unipolar limb leads (aVF, aVL, and aVR) using the voltage difference between two of the previously mentioned limb leads and the third one, and the precordial leads (V1,V2,.. to V6) that are placed on the front and left side of the chest. Although employing more ECG leads will increase the accuracy of identification as proved in [27], this configuration needs time to setup and a prior knowledge of the appropriate leads' positions. Therefore, in biometric ECG systems, it is better to use a smaller number of leads (i.e. one or two leads) such that the biometric system is practical. Thus, less time and complexity in the enrollment and authentication processes. This can be achieved by utilizing the modified limb leads (i.e. lead I or lead II). Another lead configuration can be found in [28] where the subject holds two electrodes using his/her thumb.

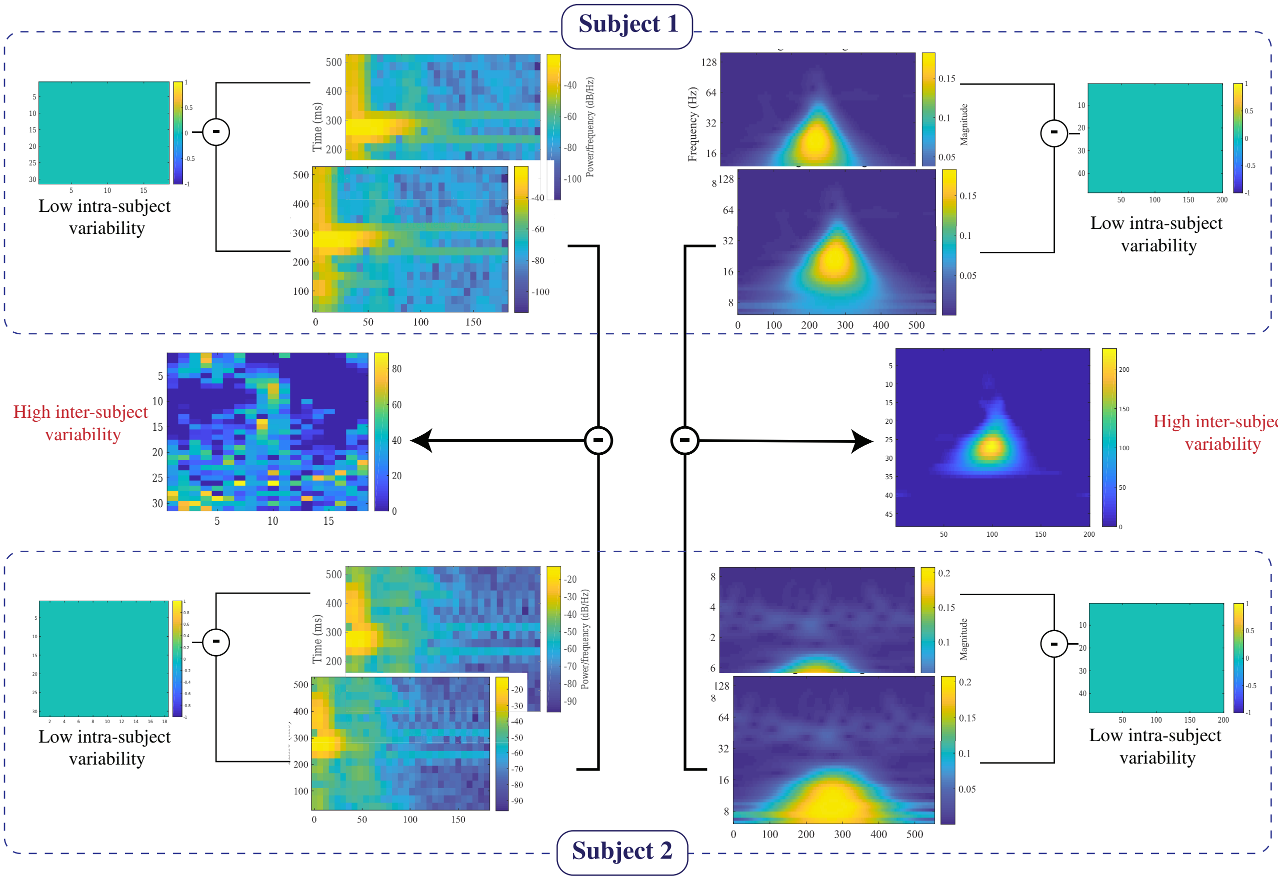


Figure 2. The intra- and inter-subject variability of both STFT images on the left and CWT images on the right of two subjects from the MITDB. The low intra-subject is shown in the green images, while inter-subject has high variability. (Better seen in the electronic version)

In this paper, since we utilize different databases that were obtained using different lead configurations, we employ the modified limb lead (Lead II) for obtaining the ECG signal used in the identification. It is also worth mentioning that in the databases we used, subjects have given a consent, and all the records were anonymized which comply with the general data protection regulations (GDPR). Another property of the ECG biometric system is that the data acquisition time should be long enough to obtain a good identification rate, and short enough not to become an inconvenient and time-consuming experience for the users. Therefore, we adapted in our approach an identification model that is designed based on a short-time duration ECG signal (i.e. order of minutes to have sufficient data for training purposes, and order of seconds to use during identification or verification purposes).

B. Data Preprocessing

The acquisition of the ECG signal is susceptible to many factors of noise such as the muscle movement and respiration as well as electric powerline interference.

In this work, since we use multiple databases, the ECG signal is resampled at 360 Hz. Then, we apply a bandpass filter to remove any frequencies other than 0.05-40 Hz. A notch filter is then used to remove the 60 Hz power line noise. Afterwards, the R peaks are identified using the Pan and Tompkins detector [29]. The ECG beat is then extracted using 200 sample points (99 before and 100 points after the R peak). This accounts for around 0.56 seconds, which is adequate to capture most of the heart cycle [30]. Finally, a signal mean normalization is done.

C. Feature Extraction

In the biometric system, the extracted features that represent the individual should exhibit both uniqueness and circumvention to discriminate the individual effectively and prohibit spoofing. The challenge with the ECG is that it is a 1D signal. Therefore, the features can be extracted from the temporal domain. However, by investigating the features in the spectro-temporal domain, we can enrich the information extracted from the ECG signal. This can be supported by the fact that the ECG signal is a non-stationary signal. Thus, the signal's parameters can change over time. Therefore, applying either Fourier or wavelet transform over the whole signal will not reveal the transitions and changes in the spectral contents over time. These characteristics of the signal support the idea of studying the time variation of the frequency components (i.e. Fourier coefficients).

We employ two different spectro-temporal domain representation of the ECG segments: the short-time Fourier transform, and wavelet transform.

1) Short-time Fourier Transform (STFT): The main idea of the standard Fourier transform (FT) is to decompose the ECG signal into the frequencies that compose it. This can be done by performing an inner product of a family of basis functions with the signal. These family basis are the complex oscillations $exp(i\omega t)$, where ω is the frequency parameter. Therefore, the Fourier transform of a signal $x(t)$ can be given as follow:

$$\mathcal{F}\{x(t)\}(\omega) = \int_{-\infty}^{\infty} e^{-i\omega\tau}\, x(\tau)d\tau \tag{1}$$

This transform can be interpreted as the frequency domain representation of the original signal. However, due to the fact that the ECG signal is a non-stationary signal, we can study the behavior, namely the spectral behavior, of the signal in short enough time window such that we can assume stationarity. This can be achieved by utilizing the STFT. The STFT follows the same idea of the standard Fourier and adds the time dimension to the base function by integrating a window of the complex exponential for the sake of localization to be $\omega(t-t_0)exp(i\omega t)$. Therefore, STFT of the ECG signal represents the Fourier coefficients of a fixed-size window contained within the signal. This can be achieved by sliding a window through the ECG time signal with a control amount of overlapping between the consecutive windows. This overlap that is performed during the windowing is provided to avoid any discontinuities or artifacts in the windowing process. The mathematical formula of the STFT is as follow:

$$STFT\{x(t)\}(\tau,\omega) = \int_{-\infty}^{\infty} \omega(t-\tau)e^{-i\omega\tau}\, x(\tau)d\tau \tag{2}$$

where $\omega(t)$ is the window function. Different types of windows can be used such as rectangular, Hann, or Hamming window. However, in this work we utilize the Hamming window as it has better attenuation lobes.

The result of this transform is a time varying spectrum image, where one axis is the time and the other is the frequency. This image can be interpreted as the changes of relative energy content in the frequency over time.

2) Continuous Wavelet Transform (CWT): Although the STFT gives a good representation of the time-frequency characteristics of the signal, it provides a fixed resolution in the frequency domain, which is not ideal in some situations. For instance, the uncertainty of 10 Hz component would be tolerable around 1000 Hz but not around 50 Hz. Contrarily, the wavelet transform provides a variable window size that adapts to the frequency on the expense of the time resolution.

The main idea of the wavelets is to have a base function that has both localization and oscillation properties while having a zero mean such that the integral over the space is zero. So, if the wavelet was scaled in the time, the oscillation frequency changes too. Therefore, the wavelet transform can be given as follow:

$$W_\psi(t,s) = \int_{-\infty}^{\infty} \frac{1}{s^n}\psi\left(\frac{\tau-t}{s}\right)x(t)d\tau \tag{3}$$

where $x(t)$ is the time signal, s is the scale, and n is the scale normalization.

In this work, we utilize the generalized Morse wavelet [31], [32]. This wavelet can be considered as a superfamily for all commonly used analytic wavelets. Therefore, to overcome the burden of spending time and effort of choosing with wavelet is better for which specific application, this wavelet family has two main parameters that controls the time and the frequency domains. The generalized Morse wavelet is represented as follow:

$$\Psi_{\beta,\gamma}(w) = \int_{-\infty}^{\infty} \psi_{\beta,\gamma}(t)e^{-i\omega t}dt = U(w)a_{\beta,\gamma}\omega^{\beta}e^{-\beta\gamma} \tag{4}$$

where ω is the frequency, β and γ are controlling parameters, $U(w)$ is the unit step function, and $a_{\beta,\gamma}$ is a normalization constant.

By changing the controlling parameters, β and γ, the shape of the wavelet family is changed. The γ controls the symmetry or skewness of the wavelet in time, or the high frequency decay in the frequency domain, while β controls the behavior near zero in the time domain. Note that we have also the scale parameter that controls the compression or dilation of the wavelet.

The application of this transform on the 1D ECG signal yields a 2D image where one of the axes is the frequency and the other is the time.

Figure 2 shows the intra- and inter-subject variability of the spectro-temporal images. We randomly picked two subjects (i.e. subject 1 and subject 2) from the MITDB database, then the spectro-temporal images using STFT and CWT are created for an average of 10 heartbeats. This is done two times for each subject. Thus, a total of eight images are created, four for each subject, where two are the STFT of two different heartbeats, and the other two are for the CWT of the same heartbeats. As shown in figure 2, the intra-subject variability can be interpreted by subtracting the STFT or CWT images of the same subject, while the inter-subject variability can be interpreted by subtracting the STFT or CWT of the two different subjects.

Inspired by the small intra-subject variability and high inter-subject variability as illustrated in figure, we can utilize

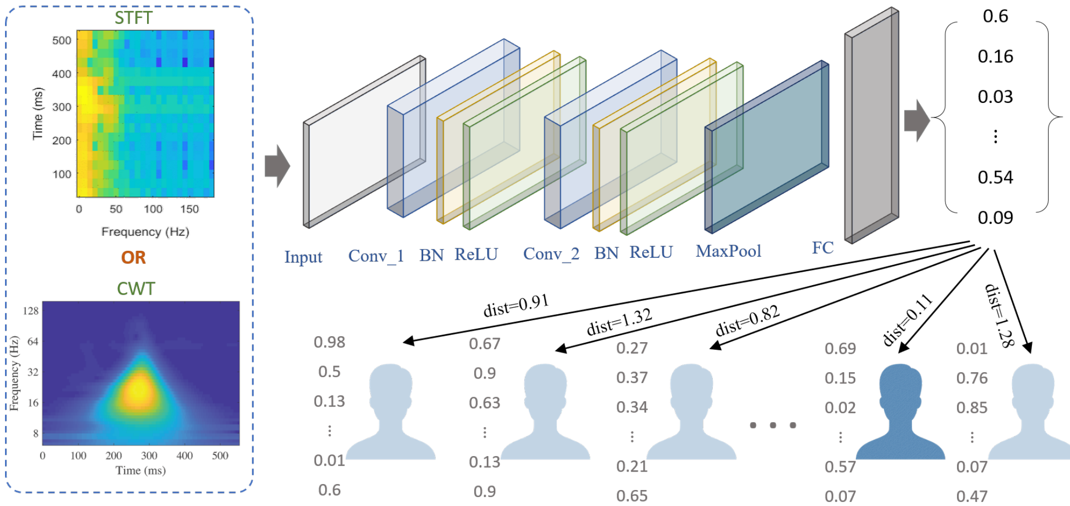


Figure 3. The architecture of the proposed CNN using the spectral image as input and the feature vector from the fully connected layer to be compared with other subjects for identification.

these images as input to the convolutional neural network in our ECG biometric system.

D. Individual Identification and Verification

Deep learning allows the extraction of high-level features that discriminate the individual with high accuracy. In this work, we utilize a 2D deep convolutional neural network to extract those features from the spectro-temporal domain images. The network is first trained using the STFT or CWT images as input and the subjects' IDs as labels.

The architecture of the proposed network is illustrated in figure 3. The input of the network is the STFT or the CWT image of the heartbeat. The convolution layers then allow the extraction of the high-level features, whereas the batch normalization layer will serve in scaling the activations in the convolution layer to make sure that there is no activation becomes too high or too low. A dropout layer is then added to avoid the over-fitting of the model followed by a fully connected layer of size equal to the number of classes to support a final softmax layer. The sizes of each layer of the CNN are set empirically and are shown in Table II.

Table II. The proposed CNN layers information.

Layer ID	Layer Size
Input Layer	128x128
Conv_1	5x5@16
Conv_2	5x5@32
MaxPool	2x2
FC	100
Softmax	Variable

The size of the ID-vectors in this architecture is 100. This compression in representation, as we represent 200 data points with 100 features, allows database constriction and usage of less memory.

E. Model Validation

Model validation is an essential step to assess the robustness, generalizability, and validity of the system. In this work, we validate the model on eight public ECG databases. In each database, we divide the images into training and testing sets. The training set is used to train the CNN using both the images and the labels. Whereas in the testing set, we subdivide it into gallery and probe sets. The gallery serves as a reference template of the individual that is stored in the database, while the probe is the data that is obtained from the subject and being tested against the gallery. In this work, the gallery and probe vectors are high-level features that are extracted from the full connected layer of the network. Thus, the size of each vector is 100. A distance metric is then used to assess the similarity between the vectors (i.e. gallery and probe vectors). Nevertheless, In order to account for heartbeat variability, we average M heartbeats' feature vector and use it as gallery. To avoid any bias or over-fitting in the validation step, the training and testing of the system is done on the base of 5-fold cross-validation to account for any outliers in the database. In addition, the selection of the gallery and probe is repeated 10 times, each time the instances in each category is picked randomly.This yields a total of 5x10=50 result records. Therefore, the results that we will show later is the average of those 50 iterations.

In this paper, we test the proposed approach on both identification and verification scenarios. We use different performance measures to assess the proposed system's performance. These measures include:

1) Identification Rate (Accuracy): the portion of correctly identified subjects.
2) The Cumulative Match Curve (CMC): the plot of the rank-k accuracy of the system.
3) The Receiver Operating Characteristics Curve (ROC):

the plot of the false acceptance rate (FAR) in the x-axis versus the genuine acceptance rate in the y-axis.

4) Area Under Curve (AUC): this is the area under the ROC curve. It measures the discrimination ability of the system and it is equal to the probability that the system will rank a randomly chosen matching instance higher than a randomly chosen non-matching one.

III. Experimental Results

In this work, we tested the system on eight ECG databases that have both normal and abnormal ECG beats (i.e. ventricular flatter, left or right bundle block, atrial fibrillation, etc.) as illustrated in table III. We utilized only 15 minutes of each recording for both training and testing. The challenges in these particular databases ensembles some cardiac diseases that are hard to extract the QRS complex from and that the normal ECG beats are distorted. Moreover, the challenge in the PTBDB is the number of subjects. So, unlike the other databases that have at most 50 subjects, PTBDB has 290 individuals, which is a more realistic scenario in real-world applications.

Table III. ECG databases used to validate the proposed algorithm

Type	Database	Subj. #	Freq. (Hz)
N	CEBSDB [33], [34]	20	5000
	NSRDB [35]	18	128
	Fantasia [36]	40	250
N/AbN	MITDB [37]	47	360
	STDB [38]	28	360
	AFDB [39]	23	250
	VFDB [40]	22	250
	PTBDB [41]	290	1000

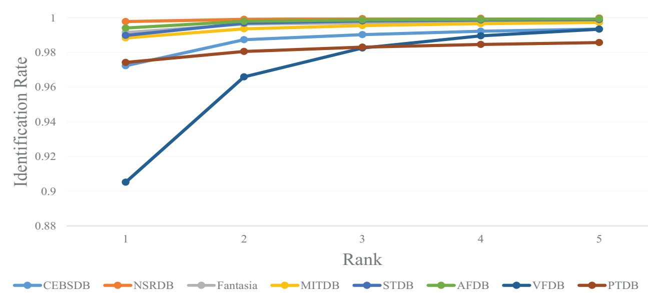
N= Normal, AbN=Abnormal, Sub. #= Subjects Number, Freq.= Sampling frequency.

The ECG signal for each subject is first resampled to 360 Hz. The R peaks were then extracted, and the heartbeats were defined using the 200 number of sample points. Afterwards, each signal is filtered then normalized.

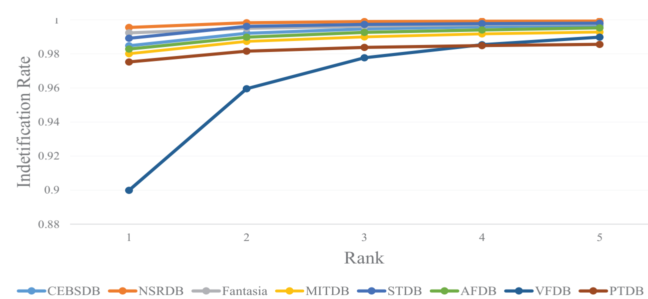
Each database is then divided into 80% and 20%, where the former is used to train the model using a learning rate of 0.002, while the latter is used to test and includes the gallery and probe sets.

The gallery of each subject is created by averaging a randomly selected 10 vectors from the testing set, whereas the probe set represent another 100 randomly records (non-overlapped). The Euclidean distance was then utilized in measuring the distance between any two vectors, having the closest vector as the identified subject.

In the identification scenario, the CMC curves of both approaches are illustrated in figure 4. In both algorithms, the VFDB scored the least average identification rate at rank-1 with a value of approximately 90%. However, it exceeded 97% in rank-2. On the other hands, all other databases' identification rates are above 97%. While the performance of the two approach are close, the CWT achieved higher identification rate in the CEBSDB database with a difference of 0.8%. Therefore, both approaches provide a good performance in the identification procedure, while having the STFT outperformed the CWT overall the databases.



(a) STFT



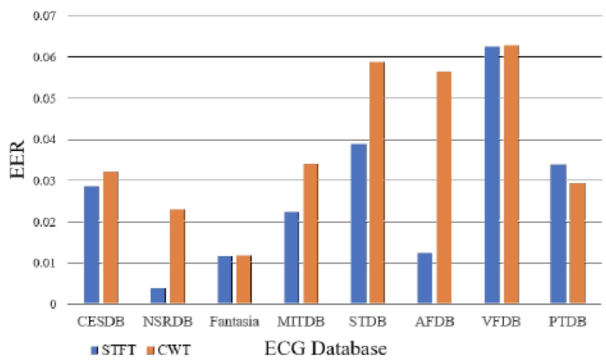
(b) CWT

Figure 4. The CMC curve of the proposed approach on the validation databases using CWT and STFT images as input, where the y axis starts at 0.88.

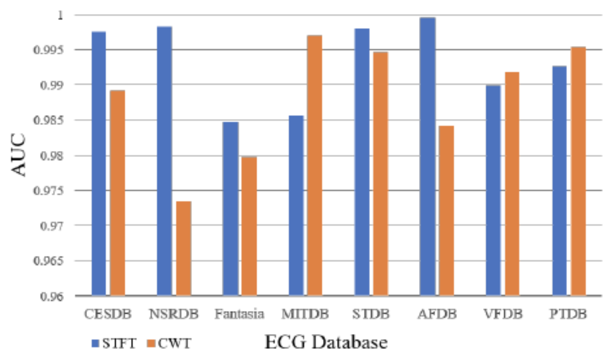
In the verification scenario, the average EER of the STFT is lower than the CWT in five databases, namely MITDB, STDB, NSRDB, AFDB, and CEBSDB, while it is equal in the VFDB, and greater in PTBDB as shown in figure 5 (a). While the average of the STFT and CWT EERs are 0.0267, and 0.03859 respectively, the maximum EER achieved by both approaches is 0.0629 at the VFDB. Therefore, the performance of the STFT is slightly better than the CWT in terms of EER.

The ROC curves of both approaches are illustrated in figure 6. The performance of the databases is high in both approaches. The VFDB showed the least performance as mentioned before.

To gain more insight into the ROC curves, we studied the AUC as shown in figure 5 (b). The STFT has a better performance than the CWT in the normal databases, while this performance varies in the abnormal databases. The STFT achieved an average AUC of 0.993, while the CWT achieved 0.988. This means that the STFT has a better performance on average when compared to CWT. This can be explained based on our previous research [5], in which we proved that CWT sustain the discriminative ability of the system to detect arrhythmia heartbeats. This in turn

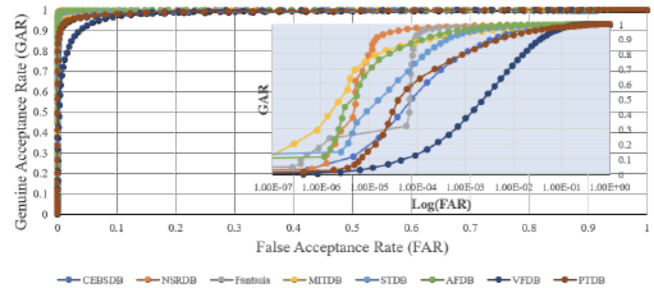


(a) EER

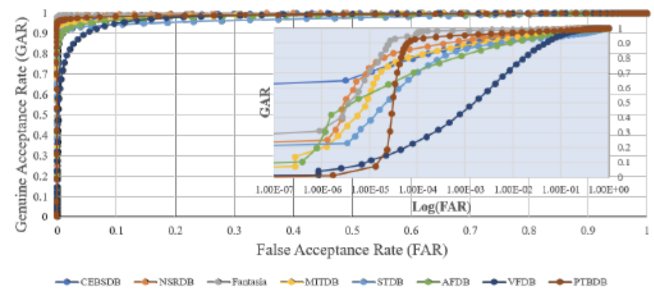


(b) AUC

Figure 5. The equal error rate (EER) and area under curve (AUC) of both proposed approaches on the eight ECG databases.



(a) STFT



(b) CWT

Figure 6. The receiver operating curves of both STFT and CWT using the proposed approach including the log_{10} scale cropped at 10^{-7} for more illustration.

Table IV. The performance of the ECG databases using the proposed system.

Database	STFT			CWT		
	IDR	*EER*	*AUC*	*IDR*	*EER*	*AUC*
CEBSDB	99.942	0.022	0.998	98.479	0.034	0.989
NSRDB	99.778	0.039	0.998	99.554	0.059	0.973
Fantasia	97.992	0.063	0.985	99.236	0.0630	0.980
Avg. N	99.237	0.041	0.994	99.090	0.060	0.981
MITDB	98.834	0.004	0.986	98.012	0.023	0.997
STDB	98.978	0.012	0.998	98.918	0.0118	0.995
AFDB	99.405	0.012	0.999	98.284	0.056	0.984
VFDB	90.515	0.034	0.990	89.9864	0.029	0.992
PTBDB	97.423	0.029	0.993	97.529	0.0322	0.995
Avg. N/AbN	97.031	0.018	0.993	96.546	0.0306	0.993
Avg. All	97.859	0.0268	0.993	97.500	0.0386	0.988

N= Normal, AbN=Abnormal, Avg.=Average, IDR= Identification Rate.

leads to higher intra-subject variability, and therefore lower discriminative power in the biometric system.

Table IV shows the performance of the proposed approach on the eight databases in terms of the identification rate, EER, and AUC. The normal databases achieved an average identification rate of 99.23% with and EER of 0.041 using the STFT which is comparable with CWT in terms of the accuracy but is lower in terms of ERR, and higher in terms of AUC.

In contrast, the abnormal databases showed equal values of AUC, while keeping the high performance of the STFT over the CWT. For instance, VFDB achieved an average accuracy of 90.65% in both approaches. This comes from the fact that the subjects in the VFDB database encompass sustained ventricular tachycardia, ventricular flutter, and ventricular fibrillation, where the ECG signal appears as a sine wave pattern and suffers from unclear QRS and T waves. In case of abnormality, the difference in the average identification rate between the STFT and CWT is higher, around 0.5% compared to 0.2% in normal case. Overall the databases, the EER for the STFT is lower compared to CWT. The STFT approach achieved a total of 97.85% overall the eight databases, which is also close to the CWT. However, it achieved lower average EER, and higher AUC.

Table V shows the performance comparison with other approaches in the literature. Both of the proposed approaches achieved the highest performance when compared to the literature work. Although validating using less dataset may increase the performance, the proposed approach using spectro-temporal features achieved 97.9% average accuracy on different eight databases using the STFT images as input to the convolutional neural network. The utilization of abnormal heartbeats is an advantage since it impersonate the real-world scenario as the subject might suffer from any cardiac condition. Moreover, for a fair comparison with Zhang et al. [17], the proposed approaches were trained

Table V. Comparison of the average performance of proposed approach with the literature work.

Authors	Databases (count) Type	Average IDR
Yue et al. [42]	(3) N/AbN	85.1
Dar et al. [8]	(3) N	93.2
Agrafioti et al. [9]	(3) N	96.2
Tantawi et al. [14]	(3) N	95.9
Zhang et al. [17]	(8) N/AbN	93.5
Proposed Approach (STFT)	(8) N/AbN	97.85
Proposed Approach (CWT)	(8) N/AbN	97.5

N=Normal, AbN=Abnormal, IDR= Identification Rate.

using randomly selected 250 heartbeats and tested on another 250 heartbeats. The average identification rates achieved in that scenarios are 96% and 94.72% for STFT and CWT respectively.

IV. Conclusion

Electrocardiogram signals has gained the attention of the researchers in the fields of biometrics since it offers embedded liveness detection and additional information about the subject's health conditions. In this work, we studied the utilization of the deep convolutional neural network and the spectro-temporal changes of the ECG signal. The short-time Fourier transform (STFT) as well as generalized Morse wavelets (CWT) were studied in our proposed approach. The models' validation was done using different eight ECG databases that include both normal and abnormal heartbeats. The spectro-temporal features showed an enhancement over other methods as we incorporate more information about the signal. Moreover, the deep network allowed the extraction of more discriminative features as the weights are learned during the training. The two approaches showed a close performance. However, the overall performance of the STFT was better than the CWT since the latter has high abnormality discriminative ability. In addition, among the databases used in validation, the abnormal databases that have ventricular anomalies such as VFDB have low identification rate. This might be due to the fact that the shape of the QRS complex is strongly affected by the disease. The SFTF system achieved an average accuracy of 97.85% over the eight studied databases, having EER and AUC of 0.0268 and 0.993 respectively.

Acknowledgment

This work was supported by the Fulbright Foreign Student Program and the U.S. Department of State.

References

[1] L. Biel, O. Pettersson, L. Philipson, and P. Wide, "ECG analysis: a new approach in human identification," *IEEE Transactions on Instrumentation and Measurement*, vol. 50, no. 3, pp. 808–812, 2001.

[2] J. M. Irvine, S. A. Israel, M. D. Wiederhold, and B. K. Wiederhold, "A new biometric: human identification from circulatory function," in *Joint Statistical Meetings of the American Statistical Association, San Francisco*, 2003.

[3] J. Galbally, J. Fierrez, and J. Ortega-García, "Vulnerabilities in biometric systems: Attacks and recent advances in liveness detection," *Database*, vol. 1, no. 3, pp. 1–8, 2007.

[4] S. Eberz, N. Paoletti, M. Roeschlin, M. Kwiatkowska, I. Martinovic, and A. Patané, "Broken hearted: How to attack ecg biometrics," in *The Network and Distributed System Security Symposium (NDSS)*, 2017.

[5] S. Abdeldayem and T. Bourlai, "Automatically detecting arrhythmia-related irregular patterns using the temporal and spectro-temporal textures of the ECG signal," *International Conference of Pattern Recognition*, 2018.

[6] M. Bassiouni, W. Khaleefa, E. El-Dahshan, and A.-B. M. Salem, "A machine learning technique for person identification using ECG signals," *International Journal of Applied Physics*, vol. 1, pp. 37–41, 2016.

[7] T.-W. Shen, W. Tompkins, and Y. Hu, "One-lead ECG for identity verification," in *IEEE Engineering in medicine and biology, biomedical engineering society EMBS/BMES conference*, vol. 1, 2002, pp. 62–63.

[8] M. N. Dar, M. U. Akram, A. Shaukat, and M. A. Khan, "ECG based biometric identification for population with normal and cardiac anomalies using hybrid HRV and DWT features," in *IEEE International Conference on IT Convergence and Security (ICITCS)*, 2015, pp. 1–5.

[9] F. Agrafioti and D. Hatzinakos, "ECG biometric analysis in cardiac irregularity conditions," *Signal, Image and Video Processing*, vol. 3, no. 4, p. 329, 2009.

[10] N. Venkatesh and S. Jayaraman, "Human electrocardiogram for biometrics using DTW and FLDA," in *International conference on Pattern recognition (ICPR)*, 2010, pp. 3838–3841.

[11] C. Camara, P. Peris-Lopez, and J. E. Tapiador, "Human identification using compressed ECG signals," *Journal of medical systems*, vol. 39, no. 11, p. 148, 2015.

[12] S. Z. Fatemian and D. Hatzinakos, "A new ECG feature extractor for biometric recognition," in *IEEE 16th International Conference on Digital Signal Processing*, 2009, pp. 1–6.

[13] R. Palaniappan and S. M. Krishnan, "Identifying individuals using ECG beats," in *IEEE International Conference on Signal Processing and Communications*, 2004, pp. 569–572.

[14] M. Tantawi, K. Revett, A.-B. Salem, and M. F. Tolba, "ECG based biometric recognition using wavelets and RBF neural network," in *Proceedings of 7th European Computing Conference (ECC)*, 2013, pp. 100–105.

[15] K. A. Sidek, I. Khalil, and H. F. Jelinek, "ECG biometric with abnormal cardiac conditions in remote monitoring system," *IEEE Transactions on systems, man, and cybernetics: systems*, vol. 44, no. 11, pp. 1498–1509, 2014.

[16] F. Porée, G. Kervio, and G. Carrault, "ECG biometric analysis in different physiological recording conditions," *Signal, image and video processing*, vol. 10, no. 2, pp. 267–276, 2016.

[17] Q. Zhang, D. Zhou, and X. Zeng, "Heartid: a multiresolution convolutional neural network for ECG-based biometric human identification in smart health applications," *IEEE Access*, vol. 5, pp. 11 805–11 816, 2017.

[18] S. A. Israel, J. M. Irvine, A. Cheng, M. D. Wiederhold, and B. K. Wiederhold, "ECG to identify individuals," *Pattern recognition*, vol. 38, no. 1, pp. 133–142, 2005.

[19] Z. Zhang and D. Wei, "A new ECG identification method using Bayes' theorem," in *IEEE region 10 conference, Tencon*, 2006, pp. 1–4.

[20] M. Hejazi, S. A. R. Al-Haddad, Y. P. Singh, S. J. Hashim, and A. F. A. Aziz, "ECG biometric authentication based on non-fiducial approach using kernel methods," *Digital Signal Processing*, vol. 52, pp. 72–86, 2016.

[21] W. K. Pratt, J. Kane, and H. C. Andrews, "Hadamard transform image coding," *Proceedings of the IEEE*, vol. 57, no. 1, pp. 58–68, 1969.

[22] J. Whelchel Jr and D. Guinn, "The fast fourier-hadamard transform and its use in signal representation and classification," MELPAR FALLS CHURCH VA, Tech. Rep., 1968.

[23] M. Hejazi, S. Al-Haddad, S. J. Hashim, A. F. A. Aziz, and Y. P. Singh, "Non-fiducial based ECG biometric authentication using one-class support vector machine," in *Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA), 2017*. IEEE, 2017, pp. 190–194.

[24] R. Tan and M. Perkowski, "Toward improving electrocardiogram (ECG) biometric verification using mobile sensors: A two-stage classifier approach," *Sensors*, vol. 17, no. 2, p. 410, 2017.

[25] Y. Wang, F. Agrafioti, D. Hatzinakos, and K. N. Plataniotis, "Analysis of human electrocardiogram for biometric recognition," *EURASIP journal on Advances in Signal Processing*, vol. 2008, no. 1, p. 148658, 2007.

[26] S. Pouryayevali, "ECG biometrics: new algorithm and multimodal biometric system," Ph.D. dissertation, University of Toronto, Department of Electrical and Computer Engineering, 2015.

[27] F. Agrafioti and D. Hatzinakos, "Fusion of ecg sources for human identification," in *IEEE 3rd International Symposium on Communications, Control and Signal Processing ISCCSP*, 2008, pp. 1542–1547.

[28] A. D. Chan, M. M. Hamdy, A. Badre, and V. Badee, "Wavelet distance measure for person identification using electrocardiograms," *IEEE transactions on instrumentation and measurement*, vol. 57, no. 2, pp. 248–253, 2008.

[29] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE transactions on biomedical engineering*, no. 3, pp. 230–236, 1985.

[30] I. Assadi, A. Charef, T. Bensouici, and N. Belgacem, "Arrhythmias discrimination based on fractional order system and KNN classifier," in *2nd IET International Conference on Intelligent Signal Processing*, 2015, pp. 1–6.

[31] S. C. Olhede and A. T. Walden, "Generalized Morse wavelets," *IEEE Transactions on Signal Processing*, vol. 50, no. 11, pp. 2661–2670, 2002.

[32] J. M. Lilly and S. C. Olhede, "Generalized morse wavelets as a superfamily of analytic wavelets," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 6036–6041, 2012.

[33] M. A. García-González, A. Argelagós-Palau, M. Fernández-Chimeno, and J. Ramos-Castro, "A comparison of heartbeat detectors for the seismocardiogram," in *IEEE Computing in Cardiology Conference (CinC)*, 2013, pp. 461–464.

[34] M. Garcia-Gonzalez, A. Argelagós, M. Fernández-Chimeno, and J. Ramos-Castro, "Differences in QRS locations due to ECG lead: relationship with breathing," in *XIII Mediterranean Conference on Medical and Biological Engineering and Computing*, 2014, pp. 962–964.

[35] A. Golberger, L. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. Mark, J. Mietus, G. Moody, P. Chung-Kan, and H. Stenley, "Physiobank, physiotoolkit, and physionet: Component of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[36] N. Iyengar, C. Peng, R. Morin, A. L. Goldberger, and L. A. Lipsitz, "Age-related alterations in the fractal scaling of cardiac interbeat interval dynamics," *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, vol. 271, no. 4, pp. R1078–R1084, 1996.

[37] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, 2001.

[38] P. Albrecht, "ST segment characterization for long term automated ECG analysis," Ph.D. dissertation, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, 1983.

[39] G. Moody, "A new method for detecting atrial fibrillation using RR intervals," *Computers in Cardiology*, pp. 227–230, 1983.

[40] S. D. Greenwald, "The development and analysis of a ventricular fibrillation detector," Ph.D. dissertation, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, 1986.

[41] R. Bousseljot, D. Kreiseler, and A. Schnabel, "Nutzung der ekg-signaldatenbank cardiodat der ptb über das internet," *Biomedizinische Technik/Biomedical Engineering*, vol. 40, no. s1, pp. 317–318, 1995.

[42] C. Ye, M. T. Coimbra, and B. V. Kumar, "Investigation of human identification using two-lead electrocardiogram (ECG) signals," in *Fourth IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS)*, 2010, pp. 1–8.