

# DEEP NEURAL NETWORKS FOR BIOMETRIC IDENTIFICATION BASED ON NON-INTRUSIVE ECG ACQUISITIONS

JOÃO RIBEIRO PINTO, JAIME S. CARDOSO, ANDRÉ LOURENÇO

## Abstract

The electrocardiogram has gained traction as a biometric trait due to its outstanding combination of universality, permanence, and measurability, with a hidden nature that makes it harder to steal or counterfeit. The state-of-the-art mostly consists of pipeline algorithms, composed of separate stages of denoising, segmentation, feature extraction, and decision. However, Convolutional Neural Networks (CNNs), possess the tools to integrate all phases of processing, from acquisition to decision, in a single model. This integration replaces separate, step-by-step tuning with an holistic optimisation process, synergically adapting the model to attain the best performance possible. In this chapter, we introduce and explore the capabilities of convolutional neural networks for biometric identification using non-intrusive ECG signal acquisitions, and propose a CNN architecture for the complete integration of traditional pipeline stages in a single accurate and robust model. The method was evaluated on the highly complete and challenging UofTDB collection, and has shown promising results on identification tasks when compared with recent and successful state-of-the-art methods.

## 1. Introduction

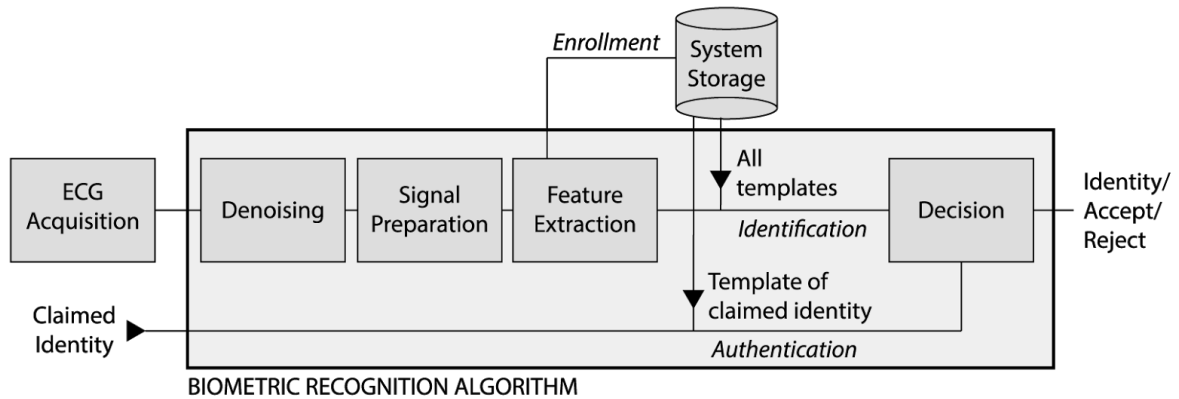
The benefits of biometric systems are well-known: they avoid the risk of loss, copy, or theft of credentials by using the user itself as the credential, and make the user authenticable at all times, without the need to carry physical credentials or remember codes or passwords (Jain et al. 2011, Kaur et al 2014). The merit of biometrics has been increasingly recognised in all fields of industry, as biometrics quickly replace traditional authentication methods in smartphones, computers, building entrances, and airports.

Biometric recognition is largely dominated by four traits: face, iris, fingerprints, and voice. These, in 2015, represented 75% of the biometrics market (Mani and Nadeski 2015), and their state of development is significantly more advanced than other traits [such as retinal scans, signature, gait, or keystroke \(Adeoye 2010, Chauhan et al. 2010\)](#). Nevertheless, the techniques for circumvention of such systems, based on recordings of the traits, are constantly evolving (Belgacem et al. 2012, Fratini et al. 2015), and require the system to become more sophisticated to ensure the liveness of the acquired trait.

Medical [or physiological](#) biometrics, such as the electrocardiogram (ECG), have inherent liveness information. The electrocardiogram is the most promising, [on par with the most common traits \(Kaur et al. 2014\)](#), due to its combination of universality, permanence, and measurability (Li and Narayanan 2010, Agrafioti et al. 2012), along with increased acceptability and comfort when acquired on non-intrusive off-the-person settings, on the hands of the subjects using ungelled electrodes. Furthermore, there are very few and recent circumvention techniques for ECG biometrics (Eberz et al.

2017), and present many constraints (like continued contact with both hands of the subject) that reduce the likelihood of success.

The algorithms for biometric recognition based on electrocardiographic signals have the goal to output an identity or match score, based on a given short segment of ECG and the data stored by the biometric system (Bolle et al. 2004, Jain et al. 2011). The existing algorithms can generally be divided into four stages, according to their purpose, as illustrated in Fig. 1.



**Figure 1.** Schema of the typical structure of an ECG-based biometric system, with emphasis for the four stages of an ECG biometric recognition algorithm.

First, we have the denoising stage, that generally consists of Butterworth filters, line-fitting smoothing procedures, and other preprocessing methods that aim towards the attenuation of noise in the received signal. Then, we have the preparation stage, where minor preparations of the denoised signal are applied to ease the extraction of meaningful features. These preparation usually include fiducial detection (the most frequent is R-peak localisation), heartbeat segmentation, time and amplitude normalisation, and outlier removal (Odinaka et al. 2010, Lourenço et al. 2011, Pinto et al. 2017). The processes included in the denoising and signal preparation have also usually been grouped in one larger stage, designated as signal preprocessing stage (Bolle et al. 2004, Jain et al. 2011). However, as the field evolves towards more comfortable and acceptable acquisitions, the importance and complexity of such processes increases to face the growing influence of noise and variability.

The third stage is feature extraction, that aims to enhance the information that is linked with the identity and individual characteristics of each individual, and remove useless information that can harm the process of recognition. The fourth and last stage is decision, where a classifier can be used to compare all stored data templates to the received signal and return an identity (identification), or to compare the received signal with the claimed identity's template and, based on a measure of similarity or dissimilarity, return a decision of acceptance or rejection (authentication).

Deep learning has multiple times shown its potential to integrate different functionalities into a single model (LeCun et al. 2015). The flexibility of the convolutional and fully-connected layers, together with the techniques of regularisation and data augmentation make it able to autonomously learn the most fitted features for the classification task at hand, while keeping the ability to generalise and be robust against high variability and noise dominance over the signal (LeCun et al. 2015, Zhang et al. 2017). The integration of all the often separate stages into a single, unified model, means that the

optimisation is not individual and partitioned but, instead, holistic and synergic: all stages are optimised as a whole towards the best model possible.

Here, we propose a novel algorithm for biometric recognition based on convolutional neural networks (CNNs). The proposed method was evaluated on the most complete and challenging off-the-person collection currently available, and replaces the entire traditional pipeline stages from denoising to decision. Furthermore, we studied how the progressive integration of traditional stages on a CNN can benefit the process of biometric identification using ECG signals acquired in non-intrusive settings, and the algorithm previously proposed in Pinto et al. (2017) was improved and adapted to be used as the baseline algorithm for direct comparison.

Besides this introduction, this chapter presents a brief but comprehensive overview of the state-of-the-art in ECG-based biometrics and some recent applications of deep learning in biometrics and with signal acquisitions, in section 2; a description of the proposed convolutional neural network architecture and its constituent parts, regularisation and data augmentation techniques, in section 3; the baseline algorithm adapted from Pinto et al. (2017), in section 4; the results of the evaluation of the proposed architecture, the study of progressive integration of pipeline stages, and the comparison with the baseline and state-of-the-art algorithms, in section 5; and final remarks and conclusions on the work performed and results obtained, in section 6.

## 2. State-of-the-Art Overview

The research topic of ECG-based biometrics is fairly recent. Despite being prophesied by Forsen et al. (1977), the first studies specifically aiming to use the ECG for individual recognition were only published by the start of the millenium (Biel et al. 1999, Kyoso et al. 2000), after important studies on the inter and intrasubject variability of the signal (Hoekema et al. 1999, Schijvenaars 2000). On the other hand, deep learning methodologies, specifically Convolutional Neural Networks, were first used by LeCun et al. (1998), but have only become trendy in the last years, due to exceptional computational requirements that could only be efficiently met with recent technology. Below, we present a more detailed overview of both the state-of-the-art in ECG biometrics and the use of deep learning for biometrics and signal analysis.

### 2.1. ECG-Based Biometrics

One of the most successful algorithms is that of Plataniotis et al. (2006) and Agrafioti and Hatzinakos (2008), that extracted normalised autocorrelation features from fixed-length ECG segments, and applied dimensionality reduction with linear discriminant analysis. Without requiring fiducial detection, they used a nearest neighbour classifier for classification, achieving 100% IDR in a dataset with 27 subjects. The algorithm was then adapted for continuous identification by Matta et al. (2011). Odinaka et al. (2010) used Log-STFT ([Short-Time Fourier Transform](#)) Spectrograms of heartbeat segments to train Gaussian Mixture Models, and log-likelihood maximisation was used to output an identity, attaining 99% identification rate (IDR or accuracy) in a private dataset with 269 subjects.

Wang et al. (2013) used representation elements from max-pooling of sparse coding coefficients of ECG segments without denoising, fed to a nearest neighbour classifier, obtaining 99.5% IDR with 100 subjects of the public PTB database. Brás and Pinho (2015) also evaluated their algorithm on the PTB

database. They proposed feature extraction using Kolmogorov-based normalised relative compression, from unsegmented signals denoised using a combination of moving average, notch, and lowpass filters, and reported 99.9% IDR with 52 subjects.

More recently, Carreiras et al. (2016) used heartbeats segmented after denoising with a 5-20 Hz bandpass filter, and obtained 84.4% IDR with 618 subjects using a nearest neighbour classifier. Tan et al. (2017) fed temporal, amplitude, and angle fiducials into a Random Forest (RF) classifier, and used Discrete Wavelet Transform (DWT) coefficients as features on a nearest neighbour classifier using wavelet distance, and fused the two classifiers at the score level. It resulted in 99.5% IDR with recordings of 184 subjects gathered from several public datasets.

The performance results reported by these works is quite encouraging, with identification rates equal or very near 100%. Nevertheless, it is important to recall the nature of the signal acquisitions used: these early initiatives focused mainly on on-the-person ECG acquisitions: intrusive recordings using several gelled electrodes, characterised by high signal quality and low influence of noise.

The performance degrades when we consider more acceptable and comfortable acquisitions, on non-intrusive off-the-person settings. These are characterised by the use of few ungelled electrodes placed on the hands or fingers, with increased influence of noise due to free movements and variability due to unstable electrode placement and contact with the skin.

Using off-the-person signals, Lourenço et al. (2011) proposed the use of averaged heartbeats, passed through processes of filtering, time and amplitude normalisation. Using a nearest neighbour classifier in identification tasks, the method attained 94.3% IDR with 16 subjects. More recently, Pinto et al. (2017) selected Discrete Cosine Transform (DCT) and Haar coefficients from ensemble average heartbeats, extracted from five-second segments denoised by a combination of Savitzky-Golay with a moving average filter. With data acquired from six drivers, during normal unconstrained driving activity using a conductive leather steering wheel cover, the method attained 94.9% IDR.

Wieclaw et al (2017) proposed the use of a multilayer perceptron (MLP) on individual heartbeats extracted from segments denoised using a bandpass filter, including methods for the rejection of noisier heartbeats, achieving 88.97% IDR with ECG signals acquired in off-the-person settings from 18 subjects. These last three works illustrate the significantly negative effect of higher noise and variability of off-the-person acquisitions on the performance of ECG biometric algorithms, despite the relatively small number of subjects in the datasets used.

As for the trending deep learning methodologies, few have yet ventured into their application on ECG biometrics, and they are yet to be explored to the fullest extent of their potential. Zhang et al. (2017) aimed to replace only the classification stage with a convolutional neural network that separately receives and processes (on the convolutional layers) the autocorrelation of the approximation and detail coefficient sets of the wavelet transform of two-second ECG segments, after component selection. The information on those levels is only unified at the classification level (on the fully-connected layers), and resulted in an average 93.5% accuracy for separate datasets with 18-47 subjects. Eduardo et al. (2017) used autoencoders to replace only the denoising and feature extraction stages of the biometric system. The autoencoders learned lower-dimensional representations of heartbeats, used on a k-Nearest-Neighbours classifier, which, for a dataset with 706 subjects in medical settings, rendered 0.91% identification error.

Salloum and Kuo (2017) used Recurrent Neural Networks (RNN) with Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU) with a sequence of segmented heartbeats, and obtained 100% IDR with 90 subjects of the public ECG-ID database. At last, Luz et al. (2018) proposed the substitution of the feature extraction and classification stages, through the fusion of two separate convolutional networks at the score level, one receiving segmented heartbeats and the other receiving the heartbeats' spectrograms. Tested for authentication tasks, their method achieved 14.27% Equal Error Rate (EER) on the UofTDB database.

Unlike these recent works, our aim is to more completely take advantage of the potential of deep learning: to integrate all stages (denoising, preparation, feature extraction, and classification) of processing into a convolutional neural network, to use it for biometric identification purposes, and to achieve competitive performance on the largest and most complete public dataset of off-the-person signals acquired from 1019 individuals.

## *2.2. Deep Learning for Signals and Biometrics*

Besides the recent efforts towards deep ECG-based biometric systems, deep learning has been extensively explored for biometric applications with other traits, and for various other purposes. In the realm of biometric recognition, Taigman et al. (2014) have proposed the well-known DeepFace algorithm for face biometric authentication. Trained with a large dataset with more than four thousand identities, the algorithm was able to learn representations that proved robust to unconstrained settings, even when transferred to different datasets. Raja et al. (2015) explored the extraction of deep sparse filtering features, that proved robust for biometric recognition based on unconstrained smartphone iris photographs. Nogueira et al. (2016) proposed the use of a CNN for liveness detection (the important process of determining if the measured trait is real or fake) in fingerprint biometric recognition, and was awarded for the high accuracy achieved.

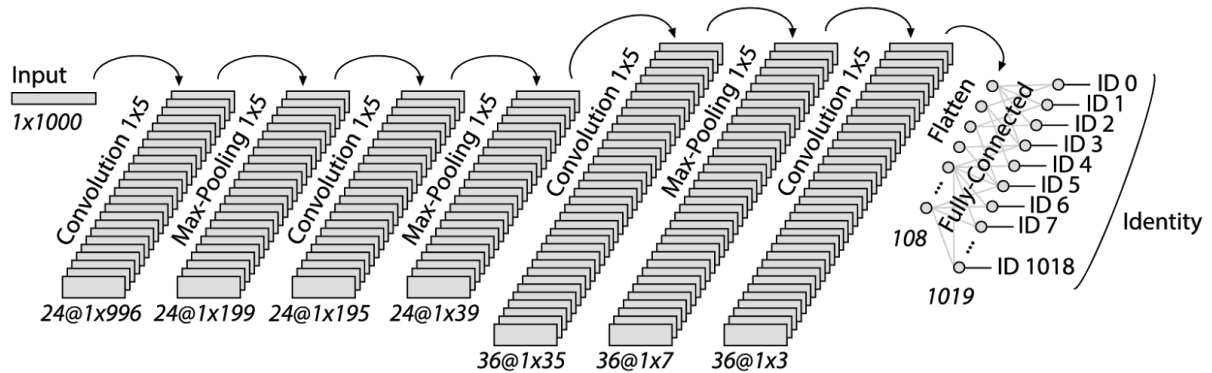
Deep learning has also been used with signal acquisitions for diverse purposes, resulting in significant performance benefits. Rajpurkar et al. (2017) proposed a 34-layer Convolutional Neural Network for detection of several types of arrhythmias on single-lead ECG acquisitions, and reported performance superior to that of professional cardiologists. Hannun et al. (2014) aimed to discard overly engineered processing pipelines for speech recognition, instead using a much simpler Recurrent Neural Network (RNN), that proved robust against common problems like speaker variation and background noise, without needing to be personally tuned for that purpose. Um et al. (2017) explored CNNs for classification of the motor state of Parkinson's disease patients, along with several unidimensional data augmentation techniques to work around dataset size limitations, predominant noise, and high variability, which resulted in significant performance improvement.

With these and many other examples, deep learning has impressed, not only for the outstanding results it achieved, but also for its robustness against noise and variability, and its adaptability to a very broad range of tasks and their specificities. Thus, deep learning presents itself as a very promising alternative to common pattern recognition, image and signal processing, and computer vision techniques applied in biometric recognition. For ECG-based biometrics, it could be the way to overcome current limitations and increase competitiveness between ECG and more advanced traits like the face and fingerprints.

### 3. A Convolutional Neural Network for ECG Biometrics

#### 3.1. General Structure Overview

The proposed methodology follows the typical structure of a convolutional neural network: the first layers are convolutional, along with max-pooling, to allow the network to learn the most advantageous representation of the input signal segment for classification, but also quickly reduce the dimensionality to keep the number of parameters low, and thus control the computational cost and training time.



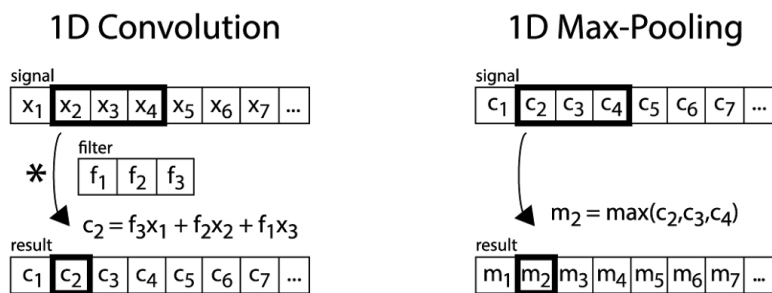
**Figure 2.** Architecture of the proposed CNN model (the number of neurons on the fully connected layer refer to the entire dataset with 1019 possible identities).

The proposed convolutional neural network (cf. Fig. 2) was developed to integrate all common pipeline stages into a single, end-to-end model, receiving raw five-second ECG segments and delivering the corresponding identity. In the following subsections, each part of the convolutional networks and its purpose is presented, along with the baseline algorithm used for comparison and to complement them in some settings.

#### 3.2. Convolutional and Pooling Layers

The convolutional and pooling layers compose the first part of a convolutional neural network (LeCun et al. 2015). Each convolutional layer (cf. Fig. 3) is composed of feature maps, that are connected to the previous layer or input by sets of weights called filter banks. The filter banks are used for discrete convolution to obtain higher representations of the inputs at each convolutional layer (Zhang et al. 2017). Each unit of the feature map of a convolutional layer will correspond to a local patch of the previous layer or input, and these local weighted sums are passed through an activation function (LeCun et al. 2015). The dimensions of the local patches and the feature maps can be controlled by tuning the size of the filters, the stride of the convolution, and by using pooling layers. In the proposed architecture, the activation function used was the Rectified Linear Unit (ReLU) (Krizhevsky et al. 2012), the size of the filters was set at 5, the stride was set at 1, through empirical tuning.

Pooling layers (cf. Fig. 3) have the main goal of reducing the dimension of the feature maps (and thus avoid unnecessary or harmful information, and control the computational costs), by merging similar features into one (LeCun et al. 2015). In the proposed CNN, Max-Pooling was used, and it works by scanning along each feature map, and keeping only the maximum value in each local neighbourhood considered. Here, the pooling stride was set at 1, and the pool size was set at 5.



**Figure 3.** Illustration of the behaviour of convolutional and max-pooling layers in unidimensional convolutional neural networks.

### 3.3. Fully-Connected Layers

The output of the last convolutional layer is flattened (or concatenated) in order to transform the feature maps into a single unidimensional vector of features, that will be the input to the fully-connected layers, the classification structure of the CNN. The fully-connected layers act as a multilayer perceptron, where each neuron is a function of the outputs of the neurons of the previous layer and the weights of the respective connections (Rumelhart et al. 1986), which means the features will be appropriately weighted and combined at each neuron, in order for the classifier to output the expected scores.

In the proposed architecture, a single fully-connected layer is used, composed of  $N$  neurons (where  $N$  is the number of subjects), with softmax activations to output a normalised distribution over the identity labels (Taigman et al. 2014). The neuron that outputs the highest value will correspond to the predicted identity.

### 3.4. Optimisation and Regularization

#### 3.4.1. Optimiser and Loss

At each moment of the training phase, based on a batch of train samples fed to the network, a measure of loss is computed by comparing the output of the network with the true labels of the batch. Based on that loss, the weights/parameters that compose the neural network are adjusted to reduce the loss, through error backpropagation (Rumelhart et al. 1986, LeCun et al. 1998).

An optimiser is a function that controls the way the weights are updated. In this work, the optimiser Adam was used. Adam is a first-order gradient-based optimization method for stochastic functions, and has been widely used for the simplicity of implementation, hyper-parameter tuning, and effectiveness in diverse tasks (Kingma and Ba 2015). The learning rate was empirically adjusted to each situation, inside the range  $[0.01, 0.001]$ , without decay, in order to allow for a quick and stable optimisation. Being a biometric identification task, a classification task with several labels, the loss measure selected was the Sparse Categorical Cross-entropy, to more effectively work with large quantity of labels and a single identity for each object.

#### 3.4.2. Dropout

In order to reach the loss minimum on the train set, neural networks will tend to memorise the train samples, i.e., overfit by learning overly specific patterns on the train data, that do not correspond with the validation data. Dropouts are used to avoid this situation (Krizhevsky et al. 2012, Srivastava et al. 2014). They are placed between two layers on the convolutional network and act upon the connections between them, setting the corresponding input to zero. In the proposed method, dropouts are used on the connections between the flattened vector of features and the fully-connected layer, effectively blocking the access of the classifier to a part of the features, and requiring it to become less specific to the training set, and more robust to unexpected variability and noise.

### 3.4.3. Data Augmentation

Dropout is effective in keeping the balance between train and test performance and avoiding an overly specific neural network. However, it sometimes falls short, and data augmentation is used to avoid overfitting and attain a more robust classifier. Data augmentation works by applying small transformations/changes to the train samples, while protecting the integrity underlying label of each sample, to simulate larger datasets, and ensure the network is robust to such variabilities (Krizhevsky et al. 2012, Chatfield et al. 2014).

Like deep learning in general, data augmentation techniques are significantly more frequent in 2D networks (for images) than in 1D (signals). Nevertheless, based on the recent work of Um et al. (2017), and taking into account the unique characteristics of the electrocardiographic signals, we propose and explored seven different types of data augmentation for 1D convolutional neural networks, presented below and illustrated on Fig. 4.

*Baseline Wander* - Based on the typical ECG noise with the same name, this technique of data augmentation simulates a periodic ondulation on the signal, by adding a sinusoidal wave with frequency near 1 Hz;

*Cropping* - This technique mutates an original train sample by taking a smaller, contiguous subsegment out of it, and resampling to match the original length. In the case of ECG signals, this technique effectively simulates slower cardiac frequencies;

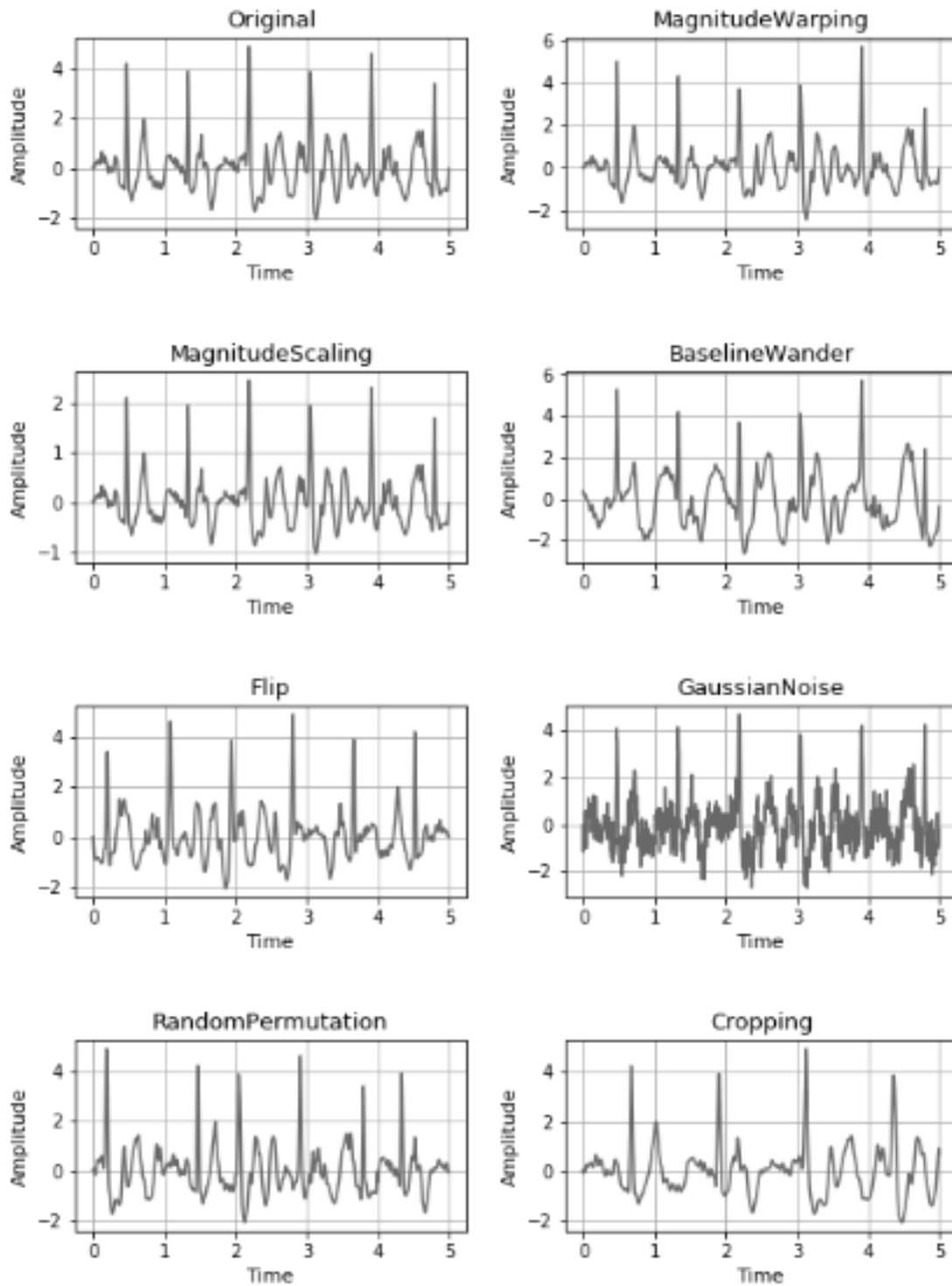
*Flip* - Consists on the inversion of the signal along the time axis: the first sample becomes the last, and similarly for all samples of the segment. This technique causes the inversion of the waveforms P, Q, R, S, and T of the heartbeats, and the inversion of their relative locations;

*Gaussian Noise* - Gaussian noise (with mean zero and standard deviation about ten times lower than the signal amplitude) is introduced to the signal to cause high frequency distortions, similar to movement artifacts and powerline interference noise on the ECG signal;

*Magnitude Scaling* - This technique consists on the rescaling of the original train sample by the multiplication of the amplitude by a factor inferior or superior (but close) to 1;

*Magnitude Warping* - Similar to the previous technique, this type of data augmentation rescales the signal in a non-uniform fashion, using a sinusoidal wave instead of a fixed factor, so that certain parts of the signal will have their amplitude shrunk, and others will see their amplitude expanded;





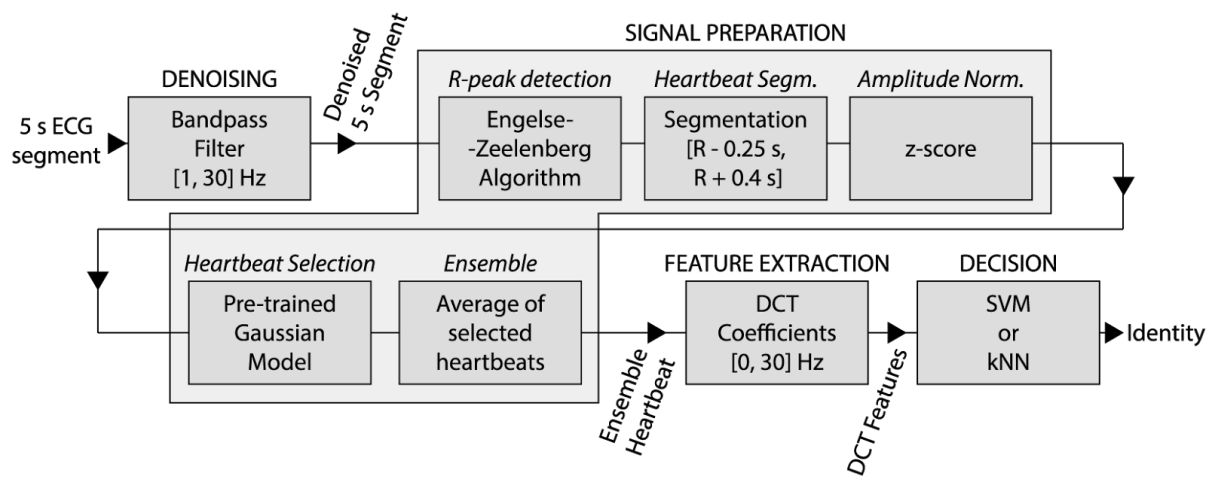
**Figure 4.** Illustration of the effects of the different data augmentation techniques on an example five-second ECG segment.

*Random Permutations* - The signal is divided into  $N$  contiguous subsegments with similar length, and their order is randomly changed. This may cause discontinuities in the heartbeats and their constituent waveforms, simulating sensor faults or abrupt segment terminations.

These data augmentation techniques were implemented on an online data generator for unidimensional data, that applied them, randomly, to the samples of each batch before feeding them to the neural network.

#### 4. Baseline Algorithm

Throughout the development of the proposed Convolutional Neural Network, an algorithm, adapted and improved from the prior work of Pinto et al. (2017) and Pinto (2017), was used for direct comparison. As a pipeline algorithm, it is composed by stages of denoising, preparation, feature extraction, and decision, that are presented below and in Fig. 5.



**Figure 5.** Schema of the baseline algorithm adapted from **Pinto et al. (2017)** and **Pinto (2017)**.

*Denoising stage* - A bandpass filter with cutoff frequencies 1 and 30 Hz is used to clean the five-second ECG segments from both low and high frequency noise such as baseline wander ( $> 1$  Hz) and powerline interference (60 Hz for the Canadian UofTDB signals);

*Preparation stage* - The Engelse-Zeelenberg algorithm is used to locate the R-peaks in the five-second ECG segments. Heartbeat segments are extracted by cropping the original segment 0.25 s before and 0.4 s after the R-peak locations. The heartbeats are then z-score normalised, subtracting the mean amplitude and dividing by the standard deviation. Among the heartbeats extracted from a five-second segment, the less noisy are selected using a Gaussian model, previously trained with a selection of cleaner heartbeats from train data of fifty subjects. The selected heartbeats are then averaged to build a single ensemble heartbeat that represents the five-second segment;

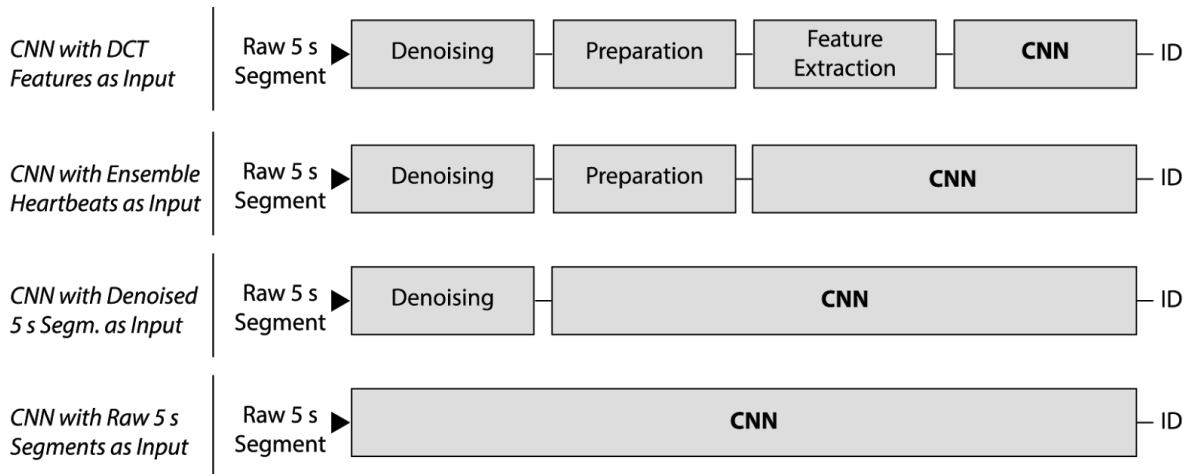
*Feature extraction stage* - The Discrete Cosine Transform (DCT) is applied to each ensemble heartbeat. As most useful information on the ECG signal is included in the frequency range [1, 30], the DCT coefficients that correspond to this range are selected and used as features on the decision stage;

*Decision stage* - The features are fed to a Support Vector Machine (SVM) classifier, the best classifier in cross-validation settings as shown by Pinto et al. (2017), that will output a predicted identity. The second best classifier, k-Nearest Neighbours (kNN) was also used.

## 5. Results and Benchmarking

The performance of the proposed convolutional neural network architecture, as previously described, was evaluated on electrocardiographic recordings of the University of Toronto ECG Database (UofTDB) (Wahabi et al. 2014). This signal collection includes signals acquired from a total of 1019 subjects, using ungelled electrodes placed on their fingertips. The subjects were measured in five different postures (supine, tripod, exercise, standing, and sitting), on up to six occasions over a time period of six months, with a sampling frequency of 200 Hz.

Besides the entire database of 1019 subjects, two subsets were also used, with 25 and 100 subjects, to evaluate the performance in smaller datasets. The datasets were divided, 70% of the data for training and 30% for testing. Also, besides the proposed method, the baseline algorithm as described here, and the state-of-the-art algorithm based on autoencoders proposed by Eduardo et al. (2017), and the algorithm based on AC/LDA features by Matta et al. (2011) were also evaluated in the same conditions. For these, it was occasionally necessary to perform adaptations in order to attend to different sampling frequencies.

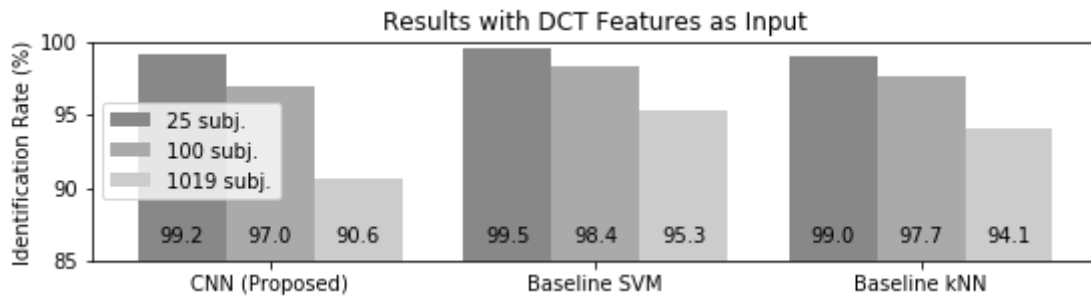


**Figure 6.** Illustration of the progressive phases of integration of the traditional pipeline stages into the CNN architecture, as explored in this chapter. For further details on each pipeline stage, cf. Fig. 5.

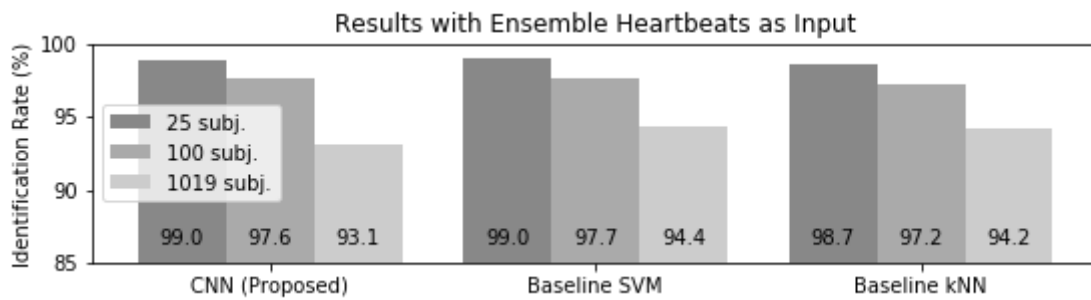
The proposed method also suffered slight adaptations to allow the study of the progressive integration of the traditional pipeline stages into the CNN model (cf. Fig. 6). Thus, besides the proposed end-to-end version that receives raw five-second ECG segments, three other variants were evaluated: the first receives denoised five-second segments (integrating all stages but denoising in the CNN); the second receives ensemble average heartbeats (integrating the stages of feature extraction and decision); and the third one receives DCT features (the CNN replaces only the decision stage). In these cases, the stages not integrated on the CNN correspond to the baseline algorithm (cf. section 4).

Analysing the results of the proposed algorithm with DCT features as input (cf. Fig. 7), it is possible to verify that its performance is similar to that of the baseline algorithm for 25 subjects. However, with the increase of subjects on the dataset (with 100 and 1019 subjects), the proposed algorithm falls behind. We hypothesise this may be caused by the very concise information that the input carries, that is fitted for typical pipeline algorithms as the baseline, but not for the proposed deep learning method.

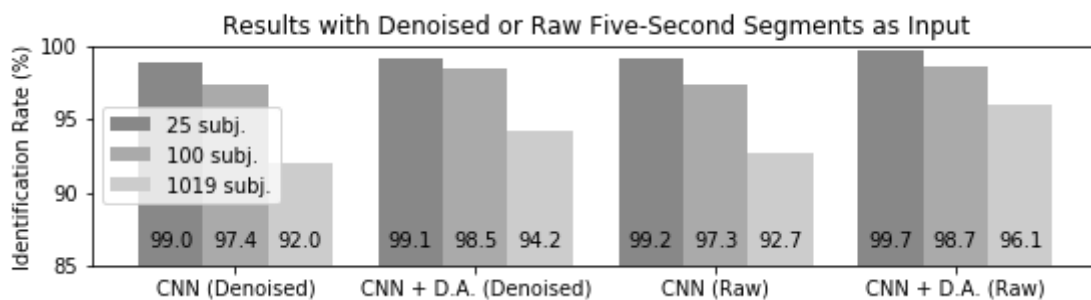
The results of the evaluation without a feature extraction stage, using ensemble heartbeats as input (cf. Fig. 8), support this hypothesis, as the performance increases and approaches that of the baseline methods, and even surpasses that of the kNN classifier on the two smaller datasets.



**Figure 7.** Results of the proposed and baseline algorithms, when using DCT features as input.



**Figure 8.** Results of the proposed and baseline algorithms, when using ensemble heartbeats as input.

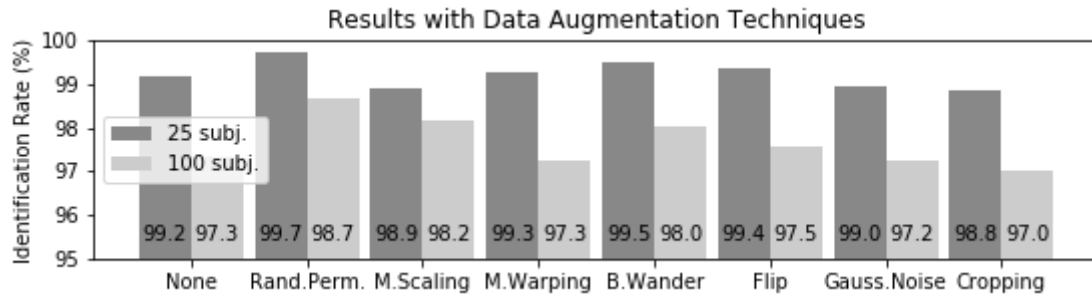


**Figure 9.** Results of the proposed and baseline algorithms, when using five-second ECG segments as input, raw or after the denoising stage.

Integrating the stages of decision, feature extraction, signal preparation, and even denoising, on the deep learning model allows us to simplify the traditional pipeline structure, and use longer signal segments as inputs (in this case, five seconds). This means an increase of complexity of the input, which can harm the performance of the convolutional neural network, but also an increase of available information and variability, that can allow for a more robust model.

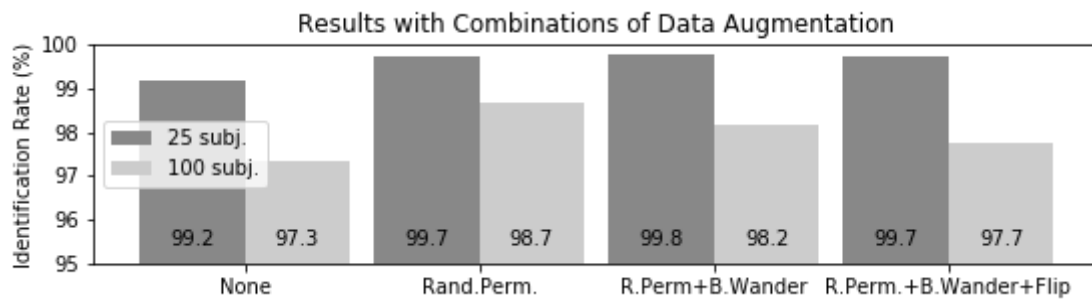
The results of the use of denoised or raw five-second segments (cf. Fig. 9) show that tradeoff between signal complexity and the increase of robustness due to extra information and variability, as the results are similar to those of the CNN receiving ensemble heartbeats. One other interesting observation is

that, in general, the results of the CNN with raw segments surpassed those of the CNN with denoised segments, which likely result from the benefit of increased variability.



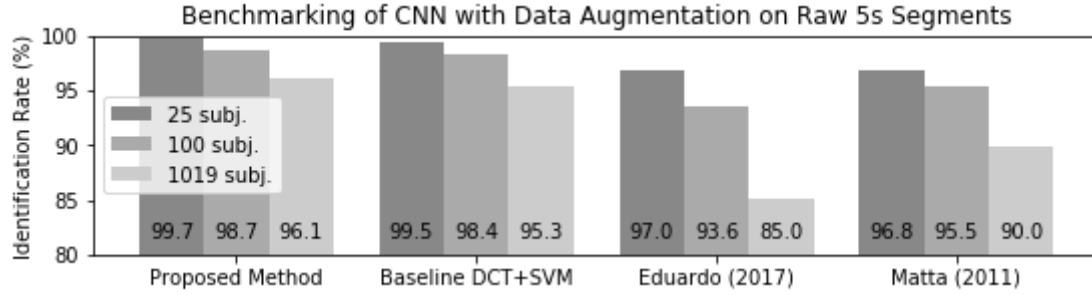
**Figure 10.** Results of the proposed algorithm, receiving raw five-second segments as input, with each technique of data augmentation, on the datasets of 25 and 100 subjects.

Increased variability is, in turn, the goal of data augmentation (D.A.). The aforementioned techniques were separately tested on the datasets of 25 and 100 subjects, and the results are presented in Fig. 10. As presented on the figure, most of the techniques of data augmentation bring improvements to the algorithm in the form of increased identification rates. The exceptions were magnitude scaling in the 25-subject dataset, cropping and magnitude warping in the 100-subject dataset, and Gaussian noise in both datasets. Knowing that the highest risk of data augmentation is the corruption of the underlying labels, especially sensitive in biometric tasks, this is the likely cause of the performance decay with these techniques.



**Figure 11.** Results of the proposed algorithm, receiving raw five-second segments as input, with combinations of the proposed data augmentation techniques.

Analysing the remaining techniques, the most promising were random permutations (that excelled in both datasets), baseline wander, and flip. These were evaluated in groups, by order of performance, to assess if the combination of two or three techniques would be beneficial to the performance of the algorithm. The results (cf. Fig. 11) show that the sole use of random permutations is the best option, despite the fact the combinations also caused the improvement of identification performance.



**Figure 12.** Direct benchmarking between the proposed architecture - CNN with Random Permutations as data augmentation and raw five-second segments as input, with the best baseline algorithm, and two implemented state-of-the-art algorithms.

We compared the proposed and baseline algorithms with state-of-the-art algorithms. As they were implemented and tested in the same context, the algorithms of Eduardo et al. (2017) and Matta et al. (2011) can be used for a direct benchmarking (cf. Fig. 12). The proposed method presents better results than all alternatives and a slightly slower decay with the increase of the number of subjects, which can denote better scalability to larger populations. The state-of-the-art algorithms likely suffer from using nearest neighbour classifiers, prone to overfit, as the results of the baseline algorithm with kNN were also consistently worse than with SVM. The method of Eduardo et al. (2017), despite showing remarkably good results in the denoising of signals during our experiments (using the entire encoder-decoder), falls short in these settings.

**Table I.** Comparison of the proposed and baseline algorithms with recent state-of-the-art methods.

Authors	Brief description	Dataset	IDR
Proposed Method	Raw segments + CNN with data augment.	UofTDB (off-the-person) - 1019 subjects	96.1%
Baseline	DCT features + SVM	UofTDB (off-the-person) - 1019 sub.	95.3%
Salloum et al. (2017)	LSTM-GRU RNN	ECG-ID (on-the-person) - 90 subj.	100%
Zhang et al. (2017)	Multiscale CNN	Several (on-the-person) - 18-47 subj.	93.5%
Wieclaw et al. (2017)	Heartbeats + MLP	Private (off-the-person) - 18 subj.	89.0%
Tan et al. (2017)	Fiducials + RF fused with DWT + WDIST kNN	Several (on-the-person) - 184 subj.	99.5%
Carreiras et al. (2016)	Heartbeats + kNN	Private (on-the-person) - 618 subj.	84.4%
Brás and Pinho (2015)	Kolmogorov-based compression	PTB (on-the-person) - 52 subj.	99.9%
Wang et al. (2013)	Max-pooling of sparse coding coefficients	PTB (on-the-person) - 100 subj.	99.5%

Finally, recalling the state-of-the-art works previously described, we can also compare the results of the proposed and baseline algorithms with the results reported by the most recent prior art works. The results of this comparison can be analysed in Table I. The IDR of the proposed and baseline algorithms may pale in comparison with some results reported in some of the considered prior works, but it is important to consider the evaluation settings. Only Wieclaw et al. (2017) used an off-the-person database as well, as opposed to the much cleaner signals of on-the-person databases still used by most researchers. Also, the UofTDB collection allowed us to evaluate our algorithm with a much larger set of subjects than any other identification method.

However, it is important to recall that deep learning both requires and benefits greatly from large datasets, where each class is represented by a large amount of samples. While, as visible on the results presented here, data augmentation attenuates the prejudicial effects of scarce data, it is difficult to acquire sufficient ECG signals from each subject to compensate for the increased noise and variability of off-the-person settings. In the datasets used, each subject was represented, in average, by just 170 five-second ECG segments, which is arguably too few to train a convolutional neural network to robustly discriminate between 1019 individuals. Considering this, with future efforts devoted to adequately deal with scarce data, deep learning methodologies could see their potential for ECG biometrics be better harnessed and place themselves as clearly better alternatives to traditional pipeline algorithms.

## 6. Conclusion

A method for biometric identification based on non-intrusive electrocardiogram acquisitions was proposed in this chapter. The method is based on the common structure of convolutional neural networks, that have proven to be able to completely and adequately replace traditional pipeline algorithms. With this goal in mind, the proposed method was evaluated when integrating more and more stages of the traditional ECG biometric pipeline, including a complete substitution by the CNN architecture, that received raw five-second ECG segments and output a decision on the corresponding identity.

Besides this study on the progressive integration of stages on the traditional pipeline, seven data augmentation techniques for unidimensional signals were explored and their individual and collective impact on the algorithm's performance was assessed. Furthermore, a previously proposed algorithm was adapted and improved and evaluated as a baseline algorithm, along with some promising state-of-the-art methods.

The proposed CNN, the baseline, and the implemented state-of-the-art algorithms were evaluated on the University of Toronto ECG Collection, currently the most complete and challenging ECG database for biometrics, using the entire population of 1019 subjects and subsets of 25 and 100 individuals. The results have shown that the total integration of traditional pipeline processes in the CNN architecture was successful, as the results of the proposed CNN with data augmentation and receiving raw five-second segments surpassed, in all settings, those of the baseline and state-of-the-art algorithms in direct benchmarking. Among other recent state-of-the-art methods, considering the diverse dataset characteristics, the proposed method has also shown success as an accurate and robust biometric identification algorithm.

## Acknowledgment

This work was partially funded by the Project “NanoSTIMA: Macro-to-Nano Human Sensing: Towards Integrated Multimodal Health Monitoring and Analytics/NORTE-01-0145-FEDER-00001” financed by the North Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, and through the European Regional Development Fund (ERDF). The authors also wish to acknowledge the access to the University of Toronto ECG Database, gracefully granted by Professor Dimitris Hatzinakos and the team of Biometrics Security Laboratory, that provided a challenging environment that significantly improved the quality of our work.

## References

- Adeoye, O. S. 2010. A Survey of Emerging Biometric Technologies. *International Journal of Computer Applications*, 9 (10): 8887-8899.
- Agrafioti, F., and Hatzinakos, D. 2008. ECG Based Recognition Using Second Order Statistics. *CNSR 2008: 6th Annual Communication Networks and Services Research Conference*, pages 82–87.
- Agrafioti, F., Bui, F. M., and Hatzinakos, D. 2012. Secure Telemedicine: Biometrics for Remote and Continuous Patient Verification. *Journal of Computer Networks and Communications*, 2012: 924791.
- Belgacem, N., Nait-Ali, A., Fournier, R., and Bereksi-Reguig, F. 2012. ECG Based Human Authentication Using Wavelets and Random Forests. *International Journal on Cryptography and Information Security (IJCIS)*, 2 (2): 1-11.
- Biel, L., Pettersson, O., Philipson, L., and Wide, P. 1999. ECG analysis: a new approach in human identification. *IMTC/99: Proceedings of the 16th IEEE Instrumentation and Measurement Technology Conference*, 1: 557–561.
- Bolle, R. M., Connell, J. H., Pankanti, S., Ratha, N., and Senior A. W. 2004. *Guide to Biometrics*. New York: Springer Science+Business Media, 1<sup>st</sup> edition.
- Brás, S., and Pinho, A. J. 2015. ECG biometric identification: A compression based approach. *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 5838–5841.
- Carreiras, C., Lourenço, A., Silva, H., Fred, A., and Ferreira, R. 2016. Evaluating Template Uniqueness in ECG Biometrics. *Informatics in Control, Automation and Robotics: 11th International Conference, ICINCO 2014*, pages 111–123.
- Chatfield, K., Simonyan, K., Vedaldi, A., and Zisserman, A. 2014. Return of the devil in the details: Delving deep into convolutional nets. *Proceedings of the British Machine Vision Conference 2014 (BMVC)*.
- Chauhan, S., Arora, A. S., Kaul, A. 2010. A survey of emerging biometric modalities. *Procedia Computer Science*, 2: 213-218.
- Eberz, S., Paoletti, N., Roeschlin, M., Patané, A., Kwiatkowska, M., and Martinovic, I. 2017. Broken hearted: How to attack ECG biometrics. *Network and Distributed System Security Symposium 2017*.
- Eduardo, A., Aidos, H., and Fred, A. 2017. ECG-based Biometrics using a Deep Autoencoder for Feature Learning: An Empirical Study on Transferability. *Proceedings of the INSTICC International Conference on Pattern Recognition Applications and Methods - ICPRAM*, 1: 463-470.



- Fratini, A., Sansone, M., Bifulco, P., and Cesarelli, M. 2015. Individual identification via electrocardiogram analysis. *BioMedical Engineering OnLine*, 14 (1): 78.
- Forsen, G. E., Nelson, M. R., and Staron Jr., R. J. 1977. *Personal Attributes Authentication Techniques*. Pattern Analysis and Recognition Corporation, Rome Air Development Center.
- Hannun, A., Case, C., Casper, J., et al. 2014. Deep speech: Scaling up end-to-end speech recognition. *arXiv preprint arXiv:1412.5567*.
- Hoekema, R., Uijen, G. J. H., & van Oosterom, A. 1999. Geometrical aspects of the inter-individual variability of multilead ECG recordings. *Computers in Cardiology*, 1999: 499–502.
- Jain, A. K., Ross, A. A., and Nandakumar, K. 2011. *Introduction to Biometrics*. Springer Publishing Company, Incorporated, 1<sup>st</sup> edition.
- Kaur, G., Singh, G., and Kumar, V. 2014. A Review on Biometric Recognition. *International Journal of Bio-Science and Bio-Technology*, 6: 69–76.
- Kingma, D. P., and Ba, J. L. 2015. Adam: A Method for Stochastic Optimization. *ICLR 2015 - International Conference on Learning Representations*.
- Krizhevsky, A., Sutskever, I., and Hinton, G. E. 2012. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 1: 1097-1105.
- Kyoso, M., Ohishi, K., and Uchiyama, A. 2000. Development of ECG Identification System. *Japanese Journal of Medical Electronics and Biological Engineering*, 38 (Supplement): 392.
- LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86 (11): 2278-2324.
- LeCun, Y., Bengio, Y., and Hinton, G. 2015. Deep learning. *Nature*, 521: 436.
- Li, M., and Narayanan, S. 2010. Robust ECG Biometrics by Fusing Temporal and Cepstral Information. *20th International Conference on Pattern Recognition (ICPR)*, 2010: 1326–1329.
- Lourenço, A., Silva, H., and Fred, A. 2011. Unveiling the Biometric Potential of Finger-based ECG Signals. *Computational Intelligence and Neuroscience*, 2011: 720971.
- Luz, E., Moreira, G., Oliveira, L. S., Schwartz, W. R., and Menotti, D. 2019. Learning Deep Off-the-person Heart Biometrics Representations. *IEEE Transactions on Information Forensics and Security*, 13 (5): 1258-1270.
- Mani, A., and Nadeski, M. 2015. *Processing solutions for biometric systems*. Texas Instruments Incorporated.
- Matta, R., Lau, J. K. H., Agrafioti, F., and Hatzinakos, D. 2011. Real-time continuous identification system using ECG signals. *24th Canadian Conference on Electrical and Computer Engineering (CCECE)*, pages 001313–001316.
- Nogueira, R. F., Lotufo, R. A., and Machado, R. C. 2016. Fingerprint Liveness Detection Using Convolutional Neural Networks. *IEEE Transactions on Information Forensics and Security*, 11 (6): 1206-1213.
- Odinaka, I., Lai, P. H., Kaplan, A. D. et al. 2010. ECG biometrics: A robust short-time frequency analysis. *2010 IEEE International Workshop on Information Forensics and Security*, pages 1–6.
- Pinto, J. R. 2017. *Continuous Biometric Identification on the Steering Wheel*. Master's thesis. Faculdade de Engenharia da Universidade do Porto.
- Pinto, J. R., Cardoso, J. S., Lourenço, A., and Carreiras, C. 2017. Towards a Continuous Biometric System Based on ECG Signals Acquired on the Steering Wheel. *Sensors*, 17 (10): 2228.

- Plataniotis, K. N., Hatzinakos, D., and Lee, J. K. M. 2006. ECG Biometric Recognition Without Fiducial Detection. *2006 Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference*, pages 1–6.
- Raja, K.B, Raghavendra, R., Vemuri, V. K., and Busch C. 2015. Smartphone based visible iris recognition using deep sparse filtering. *Pattern Recognition Letters*, 57: 33-42.
- Rajpurkar, P., Hannun, A. Y., Haghpanahi, M., Bourn, C., and Ng, A. Y. 2017. Cardiologist-level arrhythmia detection with convolutional neural networks. *arXiv preprint arXiv:1707.01836*.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. 1986. Learning representations by back-propagating errors. *Nature*, 323: 533-536.
- Salloum, R., and Kuo, C.-C. J. 2017. ECG-based biometrics using recurrent neural networks. *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2062-2066.
- Schijvenaars, R. J. A. 2000. *Intra-individual Variability of the Electrocardiogram: Assessment and exploitation in computerized ECG analysis*. Erasmus University Rotterdam.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. 2014. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, 15: 1929-1958.
- Taigman, Y., Yang, M., Ranzato, M., and Wolf, L. 2014. DeepFace: Closing the Gap to Human-Level Performance in Face Verification. *2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1701-1708.
- Tan, R., Perkowski, M. 2017. Toward Improving Electrocardiogram (ECG) Biometric Verification using Mobile Sensors: A Two-Stage Classifier Approach. *Sensors*, 17 (2): 410.
- Um, T. T., Pfister, F. M. J., Pischler, D., et al. 2017. Data Augmentation of Wearable Sensor Data for Parkinson's Disease Monitoring using Convolutional Neural Networks. *ICMI 2017 - Proceedings of the 19th ACM International Conference on Multimodal Interaction*, pages 216-220.
- Wahabi, S., Pouryayevali, S., Hari, S., and Hatzinakos, D. 2014. On evaluating ECG biometric systems: session-dependence and body posture. *IEEE Transactions on Information Forensics and Security*, 9 (11): 2002-2013.
- Wang, J., She, M., Nahavandi, S., and Kouzani, A. 2013. Human Identification From ECG Signals Via Sparse Representation of Local Segments. *IEEE Signal Processing Letters*, 20 (10): 937–940.
- Wieclaw, L., Khoma, Y., Falat, P., Sabodashko, D., and Herasymenko, V. 2017. Biometric identification from raw ECG signal using deep learning techniques. *2017 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, 1: 129-133.
- Zhang, Q., Zhou, D., and Zeng, X. 2017. HeartID: A Multiresolution Convolutional Neural Network for ECG-based Biometric Human Identification in Smart Health Applications. *IEEE Access* 5: 11805-11816.