Long-Term Performance of a Convolutional Neural Network for ECG-Based Biometrics

Gabriel Lopes[1,2]

João Ribeiro Pinto[1,2]
joao.t.pinto@inesctec.pt

Jaime S. Cardoso[1,2]
jaime.cardoso@inesctec.pt

Ana Rebelo[1,3]
ana.m.rebelo@inesctec.pt

[1] INESC TEC
Porto, Portugal

[2] Faculdade de Engenharia da Universidade do Porto
Porto, Portugal

[3] Universidade Portucalense
Porto, Portugal

Abstract

Recent studies have stressed the severity of accuracy decay over time in electrocardiogram-based biometric recognition. However, none have addressed this problem in end-to-end deep learning approaches, which have shown increased robustness to noise and variability of ECG signals. In this work, a convolutional neural network proposed in the literature was implemented and tested for long-term performance. Model update based on fine-tuning was applied to the end-to-end network, and first-in-first-out template update was applied to an adaptation of the network using k-Nearest Neighbours (kNN) as a classifier. Results with 24h Holter recordings from the E-HOL-24h database show the performance decay is also felt for end-to-end deep networks, although update techniques are successful in improving overall performance, especially when using the kNN classifier. Further efforts are needed to make ECG-based biometric algorithms reliable for long-term recognition.

1 Introduction

The Electrocardiogram (ECG) is a medical signal that has, for the last two decades, been studied as a trait for biometric recognition [5]. When compared with more common traits like the face or fingerprints, the ECG offers clear advantages for spoofing prevention. Although the literature typically reports high accuracy for ECG-based biometric algorithms, these results are mostly unrealistic due to inadequate performance evaluation settings [4].

Among these problems, one of the gravest is the decay of performance over time. This was first studied by Labati *et al.* [1, 2, 3], who tested their proposed methodology in the E-HOL-24h dataset (comprised of long Holter ECG recordings), and reported a noticeable decay in authentication performance just two hours after enrollment.

Recently, Lopes *et al.* [4] have implemented several ECG-based biometric algorithms to study the general effect of long-term variability in identification accuracy, using the E-HOL-24h dataset. Aiming for a highly realistic performance assessment, the authors have found that, although the performance is relatively high immediately after enrollment, a sharp decay is felt by all algorithms. Common template update techniques were able to improve the results, but not enough to make the algorithms reliable for real applications.

Such performance degradation is likely justified by the intrasubject variability of the ECG signals. Nevertheless, it poses a great threat to ECG-based biometrics. The high results reported in the literature cannot be trusted, and ECG-based recognition systems will likely behave poorly in real conditions, with few data from each enrolled user, whose signals will vary widely over relatively short periods (24h).

Some more sophisticated methods, such as the convolutional neural network (CNN) recently proposed by Pinto *et al.* [6], have shown increased robustness to noise and variability in off-the-person signals. Such algorithms could be able to offer better initial performance (after enrollment) and slower performance decay over time, and be useful to overcome the problem of poor long-term performance in ECG biometrics.

This paper aims to contribute to the mitigation of long-term performance decay in ECG-based identification through the combined study of deep learning and template update. The experiments of Lopes *et al.* are replicated for the deep learning model proposed by Pinto *et al.*, to evaluate the long-term performance decay and the influence of template update techniques in the model. The goal is to confirm if deep learning approaches could bring performance advantages in realistic long-term applications.

2 Methodology

2.1 Identification Model

The identification algorithm receives a biometric trait measurement and performs the adequate processing in order to enable a correct identity assignment. In this work, the algorithm implemented for biometric identification is drawn from the work of Pinto *et al.* [6]. It consists of an end-to-end unidimensional convolutional neural network that receives five-second blindly-segmented z-score normalised ECG segments (see Fig. 1).

The feature extraction part of the model is composed of four convolutional layers (24, 24, 36, and 36 filters) interleaved with three max-pooling layers. All layers have filter/pooling size 1×5, and convolutional layers include ReLU activation units. The classification part of the network is composed of a single fully-connected layer, with N neurons (where N is the number of enrolled identities), and softmax activation units.

Following the findings of Lopes *et al.* [4] that template update is most effective with nearest neighbour classifiers, additional experiments were performed where the features output by the feature extraction part of the model were fed to a k-Nearest Neighbour (kNN) model responsible for the identity decisions.

2.2 Template or Model Update Techniques

Template or model update techniques are used to adapt the biometric template galleries or the recognition models to the continuous evolution of the user's biometric traits. In this work, two techniques were implemented and applied to the identification model: fine-tuning and first-in-first-out. Fine-tuning was applied to the end-to-end CNN model, while the FIFO technique was studied for the kNN using the convolutional network for feature extraction. In both techniques, test samples are used for update if the respective scores output by the model meet a set threshold.

In the fine-tuning technique, the CNN model is briefly optimised with the samples accepted for update, using the predicted labels. The model retains knowledge of the users' supervised training samples, as it was trained using their enrollment samples, but is slightly adapted to the new personal patterns of the users' signals.

In the first-in-first-out (FIFO) technique, the system stores a number N_k of samples from each subject k. Initially, all these samples originate from the enrollment phase. The oldest samples of the sets are the first to be replaced by test samples at each update step. This means the system gradually updates the biometric gallery of each subject to keep up with their variability.

3 Experimental Settings

The experiments used data from the E-HOL-24h database, available at the Telemetric and Holter ECG Warehouse of the University of Rochester[1]. The provided database includes Holter signals of 201 subjects, with up to 24h of continuous three-lead ECG chest ECG recordings at 200 Hz. Data from thirteen subjects were discarded due to inadequate leads or unacceptable signal quality. Only one lead was used for each user, the one that most closely resembled Lead I signals.

[1] THEW, E-HOL-03-0202-003 database. Available on: `http://thew-project.org/Database/E-HOL-03-0202-003.html`.

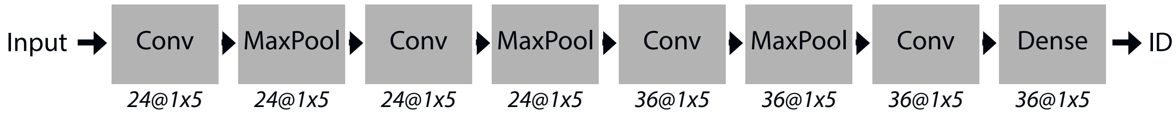

Figure 1: Architecture of the neural network used in this work, proposed by Pinto *et al*. [6].

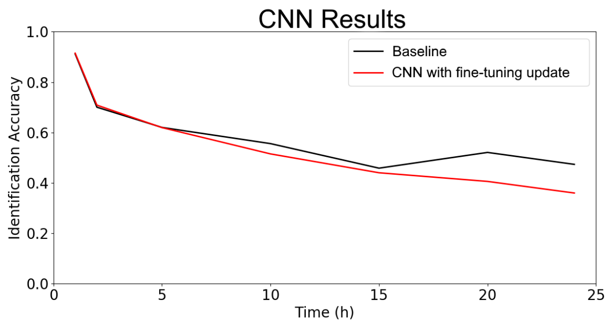

Figure 2: Performance results over time for the implemented end-to-end model with fine-tuning-based model update.

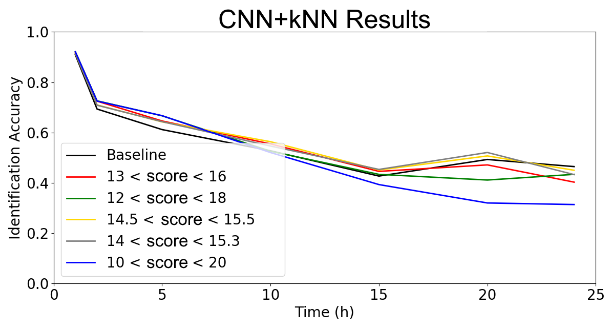

Figure 3: Performance results over time for the implemented model with kNN decision and FIFO template update, for several threshold criteria (samples with scores between the presented values are accepted for update).

The network was trained with fifteen minutes of data from each of twenty-five subjects reserved for this purpose, following the protocol described by Pinto *et al*. [6]. The recordings from the remaining 163 subjects were solely used for testing. For each of these subjects, the last 30 seconds of the first recording hour were divided into five-second segments, with four-second overlap, that were used as enrollment samples. From the remainder of the signals, 15 minute segments were used as test points over the 24h period (at 1, 2, 3, 5, 10, 15, 20, and 24 hours). Each test point was divided into several five-second segments.

4 Results and Discussion

The results for the implemented end-to-end convolutional neural network are presented in Fig. 2. Model update offered a small improvement in performance in the first test point (91.48% *versus* 91.15 without update). However, the model experiences sharp performance decay and, after the fifth hour, the model update is unable to improve identification accuracy. In fact, model update caused a decrease in identification rate, which is coherent with the findings regarding update with multilayer perceptron classifiers reported by Lopes *et al*. [4].

Different results were obtained when the fully-connected layer of the network was replaced by a kNN classifier (see Fig. 3). Although the results with kNN are slightly worse than those of the end-to-end network (90.89% *vs.* 91.15% for the first test point), the template update technique is more successful and is able to offer performance improvements for almost all test points.

When compared with the results reported by Lopes *et al*., the CNN (either end-to-end or with kNN classification) offers the best performance in the first test points after enrollment. However, it gradually loses that advantage as the time passes, even with template update.

Likely, the network will need more data with more variability during the first training phase. Increasing the training data to thirty or even a few hours per subject would enable the network to better learn the common variability patterns of the ECG. This should not only increase the initial performance, immediately after enrollment, but also reduce the performance decay over time.

5 Conclusion

This work studied the effect of ECG long-term variability in biometric identification performance of an end-to-end convolutional neural network. Although the deep learning algorithm is better than traditional methods immediately after enrollment, it offers slightly worse performance as time progresses.

Template update was able to offer performance improvements, especially when the fully-connected layer of the CNN was replaced by a kNN classifier. However, the obtained results in these more realistic settings show that the performances reported in the literature would likely not be verified upon real application. Hence, much work is still needed on template update techniques and deep learning methodologies to offer good and consistent performance in ECG biometrics.

Acknowledgements

This work was financed by the ERDF – European Regional Development Fund through the Operational Programme for Competitiveness and Internationalization - COMPETE 2020 Programme and by National Funds through the Portuguese funding agency, FCT – Fundação para a Ciência e a Tecnologia within project "POCI-01-0145-FEDER-030707", and within the PhD grant "SFRH/BD/137720/2018". Data used for this research was provided by the Telemetric and Holter ECG Warehouse of the University of Rochester (THEW), NY.

References

[1] R. D. Labati, R. Sassi, and F. Scotti. ECG biometric recognition: Permanence analysis of QRS signals for 24 hours continuous authentication. In *2013 IEEE International Workshop on Information Forensics and Security (WIFS)*, pages 31–36, Nov 2013. doi: 10.1109/WIFS.2013.6707790.

[2] R. D. Labati, V. Piuri, R. Sassi, F. Scotti, and G. Sforza. Adaptive ECG biometric recognition: a study on re-enrollment methods for QRS signals. In *2014 IEEE Symposium on Computational Intelligence in Biometrics and Identity Management (CIBIM)*, pages 30–37, Dec 2014. doi: 10.1109/CIBIM.2014.7015440.

[3] R. D. Labati, E. Muñoz, V. Piuri, R. Sassi, and F. Scotti. Deep-ECG: Convolutional Neural Networks for ECG biometric recognition. *Pattern Recognition Letters*, 2018. doi: 10.1016/j.patrec.2018.03.028. (in press).

[4] G. Lopes, J. R. Pinto, and J. S. Cardoso. Don't You Forget About Me: A Study on Long-Term Performance in ECG Biometrics. In *Proceedings of Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA)*, July 2019. doi: 10.1007/978-3-030-31321-0_4.

[5] J. R. Pinto, J. S. Cardoso, and A. Lourenço. Evolution, current challenges, and future possibilities in ecg biometrics. *IEEE Access*, 6: 34746–34776, 2018. doi: 10.1109/ACCESS.2018.2849870.

[6] J. R. Pinto, J. S. Cardoso, and A. Lourenço. Deep Neural Networks For Biometric Identification Based On Non-Intrusive ECG Acquisitions. In K. V. Arya and R. S. Bhadoria, editors, *The Biometric Computing: Recognition and Registration*. CRC Press, 2019. ISBN 978-0-8153-9364-1. (in press).