Achieving Cancellability in End-to-End Deep Biometrics with the Secure Triplet Loss

João Ribeiro Pinto [1,2]
joao.t.pinto@inesctec.pt
Miguel V. Correia [1,2]
mcorreia@fe.up.pt
Jaime S. Cardoso [1,2]
jaime.cardoso@inesctec.pt

[1] INESC TEC
Porto, Portugal
[2] Faculdade de Engenharia
Universidade do Porto
Porto, Portugal

Abstract

The literature on biometric recognition shows a chasm between the methods focused on high performance and the works focused on template security. To build a connection between these two worlds, this work describes the Secure Triplet Loss to achieve template cancellability within end-to-end deep learning models. Evaluated for off-the-person electrocardiogram authentication, the proposed methodology resulted in effective cancellability, irreversibility, and improved performance. Despite the high linkability, this shows that it is possible to combine the high performance of deep learning with adequate template security.

1 Introduction

As biometric recognition technologies quickly conquer a place of relevance in our society, the duality of performance *versus* security is yet to be adequately addressed [11]. This duality relates to how the research field of biometrics is currently composed of two 'worlds' apart, and while both work towards the same goal of improving human recognition systems, they have been following largely unconnected and uncoordinated research lines.

On the one hand, a substantial part of the literature in biometrics is focused on performance, following well-known and successful methodologies in computer vision tasks. These use mostly end-to-end convolutional neural networks (CNNs) that consider biometric recognition as a general classification problem [14], and have achieved outstanding levels of accuracy and robustness in challenging scenarios. However, since stored data protection is rarely addressed, these algorithms are incomplete and unfit for real biometric applications.

On the other hand, several algorithms have been proposed to protect personal data stored in biometric systems [5, 8]. These commonly use cryptography and information theory concepts to ensure stored biometric templates verify the essential properties of irreversibility, cancellability, and non-linkability. Nevertheless, being based on separate processes means these methodologies are not applicable to state-of-the-art end-to-end methods without significant negative impacts on performance. This is a relevant problem, since many biometric traits (including the electrocardiogram, ECG) rely on end-to-end CNNs to offer acceptable accuracy and robustness to challenging scenarios [9].

This work aims to bring the two aforementioned research lines together by answering the following question: *if deep learning models have successfully learnt so many different things, why not template security?* The proposed method is an adaptation of the triplet loss [2], which aims to achieve template irreversibility and cancellability on end-to-end CNNs while preserving recognition accuracy. This methodology is used to train a competitive end-to-end model for ECG biometric recognition [9] and evaluated on the off-the-person UofTDB database [13]. Thus, this work addresses the challenge of template protection on end-to-end networks for ECG and biometrics in general, contributing towards a synergy between performance and security in biometric recognition.

2 The Secure Triplet Loss

The triplet loss [2] is used to train models to determine whether or not two samples belong to the same class [3, 4, 9]. The model receives a triplet of inputs: an anchor (x_A of class i_A), a positive sample (x_P of class $i_P = i_A$), and a negative sample (x_N of class $i_N \neq i_A$). Considering the case of biometric recognition, the samples are biometric trait measurements and the classes are identities.

The model will output an embedding y for each input (*e.g.*, $y_A = f(x_A)$ for the anchor). Two embeddings can be compared through a metric of distance or dissimilarity $d(y_1, y_2)$ which can be used to determine if the respective inputs belong to the same class. The model can be trained through the triplet loss

$$l = \max\left(0, \alpha + d(y_A, y_P) - d(y_A, y_N)\right), \tag{1}$$

which will promote the maximisation of $d(y_A, y_N)$ and the minimisation of $d(y_A, y_P)$, grouping samples of the same class into compact clusters, at least α from other classes in the embedding space.

Although the triplet loss has been successfully applied to several pattern recognition problems, including biometric authentication, it does not address the important issue of template cancellability. Typically, this is performed separately, binding a subject-specific key k with the template after it is generated: changing k invalidates any compromised templates bound with other keys.

Here, we adapt the triplet loss to perform subject key binding with the template within the end-to-end model. Besides the biometric samples x_A, x_P, and x_N, the model will receive two keys, k_1 and k_2. Sample x_A is bound with k_1 and x_P and x_N are bound with each of the two keys, resulting in five embeddings: $y_A = f(x_A, k_1)$, $y_{P1} = f(x_P, k_1)$, $y_{P2} = f(x_P, k_2)$, $y_{N1} = f(x_N, k_1)$, $y_{N2} = f(x_N, k_2)$. From these, four distances are computed: $d_{SP} = d(y_A, y_{P1})$ (with matching identities and keys), $d_{DP} = d(y_A, y_{P2})$ (with matching identities but different keys), $d_{SN} = d(y_A, y_{N1})$ (with different identities but matching keys), and $d_{DN} = d(y_A, y_{N2})$ (with non-matching identities and keys).

Since d_{SP}, which corresponds to matching identities and keys, should be minimised, while the others should be maximised, the Secure Triplet Loss is computed through:

$$l = \max\left(0, \alpha + d_{SP} - \min(\{d_{SN}, d_{DP}, d_{DN}\})\right). \tag{2}$$

As with the triplet loss, α will enforce a margin between positive and negative distances. By minimising the loss in Eq. (2), the model learns to deal with the intrasubject and intersubject variability of the biometric trait and becomes able to recognise when the keys do not match, even if the identity is the same. Hence, if the stored templates become compromised, they can easily be invalidated through a key change.

3 Experimental Settings

The proposed training methodology was evaluated to off-the-person ECG-based biometric authentication. The University of Toronto ECG Database (UofTDB) [13], including 1019 identities, was used. Signals were divided into five-second segments. Data from the last 100 subjects were used for training (90 000 triplets) and validation (10 000 triplets), while the data from the remaining 918 subjects were reserved for testing (10 000 triplets). Keys were randomly generated as unidimensional arrays of 100 binary values.

The authentication model is adapted from [9] (see Fig. 1). Samples are bound with keys before the first dense layer. The vector of flattened feature maps ($s(x)$) is concatenated with a key k (after its normalisation to unit $l2$ norm). The last dense layer outputs the respective representation $y = f(s(x), k)$, which is then used in dissimilarity score computation using the Euclidean and normalised Euclidean distance, respectively, for training and testing. The model was trained using the Adam optimizer with an initial learning rate of 0.0001, for a maximum of 500 epochs, with early stopping based on validation loss (patience of 20 epochs).

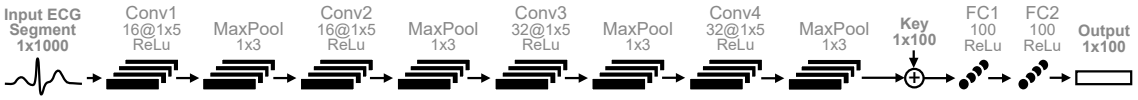

Figure 1: Architecture of the model trained for ECG-based authentication.

4 Results and Discussion

After training, the model's authentication performance was evaluated through the analysis of false acceptance (FAR), rejection rates (FRR), and equal error rates (EER) [10]. As presented in Fig. 2, the model trained with the Secure Triplet Loss achieved lower EER than with the original triplet loss (10.63% versus 12.55%). This is an important aspect of the proposed method, since security measures generally lead to a five-fold average increase in authentication error [8]. The proposed method is able to achieve this by retaining the capabilities of end-to-end networks and optimising for accuracy and cancellability simultaneously.

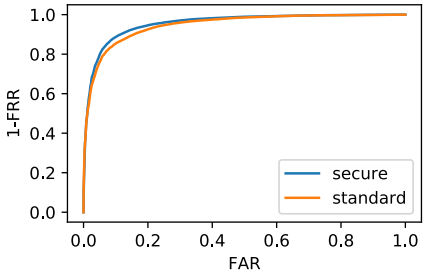

Figure 2: Receiver-operating characteristic curves of the model trained with the Secure Triplet Loss and with the original triplet loss.

The security of the templates output by the model was evaluated using the standard literature measures of privacy leakage rate, secrecy leakage, and secret key rate, through nearest-neighbour entropy estimation methods [1, 6, 7]. The model offered near-perfect privacy rate results, which means the biometric templates are irreversible as desired. This very useful property may be a consequence of using CNNs, which have been observed to present minimal mutual information between inputs and outputs when appropriately optimised [12]. Secrecy leakage also rendered perfect result (0) which may also be related to the nature of deep neural networks. At last, the proposed method offered 103.73 bits of secret key rate (output entropy) versus 14.20 bits for the original triplet loss, which means it will be harder to successfully attack the model trained with the proposed method.

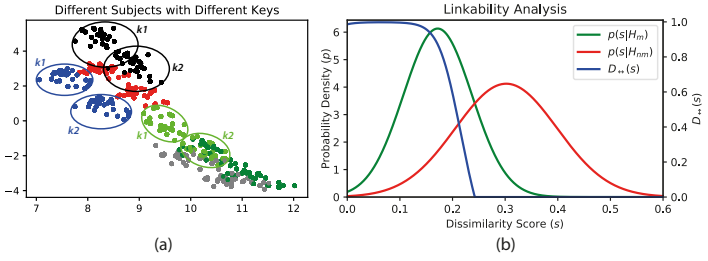

Figure 3: Results of the cancellability (a) and linkability (b) evaluation.

Regarding template cancellability, Singular Value Decomposition (SVD) was used to visualise the template distribution in the output space. Fig. 3 (a) shows the Secure Triplet Loss promotes the clustering of class samples when keys match. However, when the key is changed, the cluster is shifted on the output space in order to distance itself from (and effectively invalidate) the templates corresponding to cancelled keys. At last, template non-linkability was evaluated as established by Gomez-Barrero *et al.* [5] (see Fig. 3 (b)). The proposed secure triplet loss model offered $D_{\leftrightarrow}^{sys} = 0.67$, making it semi- to fully linkable. This is the main shortcoming of the Secure Triplet Loss, as it would be relatively easy for an attacker to discover whether two samples with different keys belong to the same subject. The desired behaviour would be for d_{DP}, d_{DN}, and d_{SN} to assume similar values greater than d_{SP}. Future research endeavours should focus on adapting the network to avoid template linkability.

5 Conclusion

This work proposes the Secure Triplet Loss, an adaptation of the triplet loss to promote biometric template cancellability in end-to-end deep models. Biometric templates are bound with subject-specific keys within the end-to-end model, without separate processes, and can be easily cancelled through a key change. The proposed loss proved successful when evaluated for ECG-based authentication, offering cancellability and improved performance.

While cancellability and irreversibility have been achieved, an important shortcoming regarding template linkability has been unveiled. Hence, further efforts should be devoted to achieve non-linkability alongside cancellability. Nevertheless, this study has shown it is possible to achieve template security within end-to-end deep biometric models, paving the path to a synergy between performance and security in 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". The authors wish to thank the administrators of the UofTDB database used in this work.

References

[1] P. Brodersen. Entropy estimators, 2017. URL `https://github.com/paulbrodersen/entropy_estimators`.

[2] G. Chechik, V. Sharma, U. Shalit, and S. Bengio. Large scale online learning of image similarity through ranking. *JMLR*, 11, 2010.

[3] W. Chen, X. Chen, J. Zhang, and K. Huang. Beyond triplet loss: A deep quadruplet network for person re-identification. In *CVPR*, 2017.

[4] D. Cheng, Y. Gong, S. Zhou, J. Wang, and N. Zheng. Person Re-Identification by Multi-Channel Parts-Based CNN With Improved Triplet Loss Function. In *CVPR*, 2016.

[5] M. Gomez-Barrero, C. Rathgeb, J. Galbally, C. Busch, and J. Fierrez. Unlinkable and irreversible biometric template protection based on bloom filters. *Information Sciences*, 370-371:18–32, 2016.

[6] L. F. Kozachenko and N. N. Leonenko. Sample estimate of the entropy of a random vector. *Problemy Peredachi Informatsii*, 23, 1987.

[7] A. Kraskov, H. Stögbauer, and P. Grassberger. Estimating mutual information. *Phys. Rev. E*, 69:066138, Jun 2004.

[8] K. Nandakumar and A. K. Jain. Biometric Template Protection: Bridging the performance gap between theory and practice. *IEEE Signal Processing Magazine*, 32(5):88–100, 2015.

[9] J. R. Pinto and J. S. Cardoso. A end-to-end convolutional neural network for ECG-based biometric authentication. In *BTAS*, 2019.

[10] J. R. Pinto, J. S. Cardoso, and A. Lourenço. Evolution, Current Challenges, and Future Possibilities in ECG Biometrics. *IEEE Access*, 6:34746–34776, 2018.

[11] J. R. Pinto, J. S. Cardoso, and M. V. Correia. Secure Triplet Loss for End-to-End Deep Biometrics. In *IWBF*, April 2020.

[12] N. Tishby and N. Zaslavsky. Deep learning and the information bottleneck principle. In *ITW*, April 2015.

[13] S. Wahabi, S. Pouryayevali, S. Hari, and D. Hatzinakos. On Evaluating ECG Biometric Systems: Session-Dependence and Body Posture. *IEEE TIFS*, 9(11):2002–2013, Nov. 2014.

[14] M. Wang and W. Deng. Deep face recognition: A survey. *arXiv*, 2018. 1804.06655.