Myocardial Infarction Detection Using a Compression-Based Approach

João M. Carvalho, Susana Brás and Armando J. Pinho

IEETA - Dept. of Electronics Telecommunications and Informatics, University of Aveiro, Portugal

{joao.carvalho|susana.bras|ap}@ua.pt

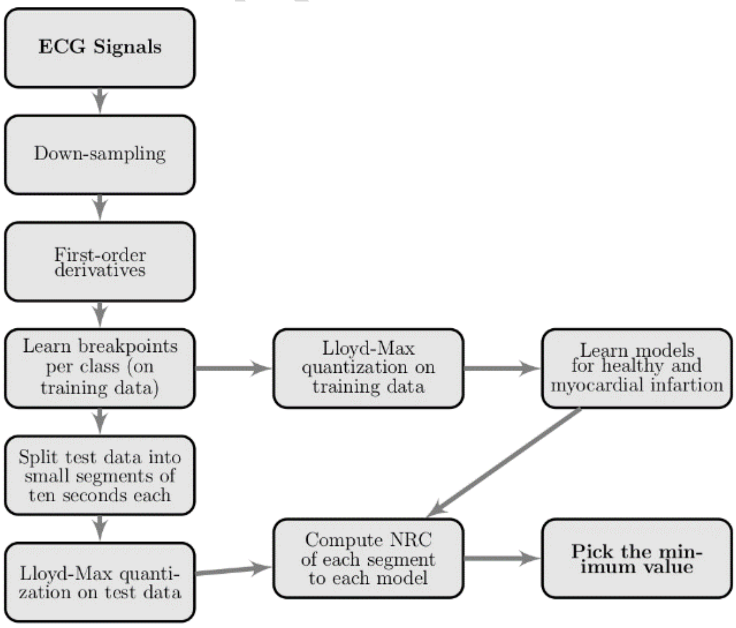
The automatic classification of myocardial infarction (commonly known as heart attack) is, undoubtedly, a useful tool for hospitals. There are several studies comparing different approaches for this task on the literature.

In this study, we propose an alternative approach to solve the same task -- a compression-based non-fiducial method, that uses a measure of similarity, called the Normalized Relative Compression (NRC). The NRC is a measure related to the Kolmogorov complexity of strings. which has proven useful on different pattern recognition tasks, such as like image pattern recognition [1], authorship attribution [4], DNA analysis [6] and biometric identification using electrocardiographic signals [8][9].

Our method uses extended-alphabet finite-context models (*xaFCM*s [8]) on the quantized first-order derivative of the signal. An interesting fact is that no feature extraction is needed to use this method, neither does the classification task at hand influence anything besides a few tweaks on the parameters, at least in theory -- more tests need to be done to claim this fact.

As a testbed, we used the PTB database [10], collected at 1000Hz, and publicly available at Physionet [11]. Since the classes of this database are highly unbalanced, we decided to use only 80 ECG recordings for each class (healthy and myocardial infarction) to obtain some preliminary results.

The method used in this work was similar to the one described in [9] and a general overview can be seen in the following picture:



In order to reproduce our results, the parameters used for pre-processing of the signal were: Butter-Worth low-pass filter of 6th order at 40Hz, down-sampling to 40Hz afterwards and quantization to an alphabet size of 15 on the consecutive differences. Regarding the *xaFCM*s, a context size k=4 was used, and a depth d=2. The parameter alpha was chosen automatically, according to what is defined in [8].

To evaluate the performance of this method, we performed random sampling of the data and have chosen 90% for training data and the other 10% for test data. The test data was then split into smaller segments of just 10 seconds and the classification test is done on this small segments. All this process is repeated 100 times, in order to obtain a more trustworthy accuracy metric. Using this evaluation scheme, we obtained an accuracy of around 85.6% on the described classification task.

Even though this does not beat the current state of the art, it shows that compression-based methods can be suitable to this kind of problem and future work should be done to see how much these preliminary results can be improved.

Acknowledgments - This work was funded by national funds, European Regional Development Fund, FSE through COMPETE2020, through FCT, in the scope of the framework contract foreseen in the numbers 4, 5 and 6 of the article 23, of the Decree-Law 57/2016, of August 29, changed by Law 57/2017, of July 19; and in the scope of the projects UID/CEC/00127/2013 (IEETA/UA), CMUP-ERI/FIA/0031/2013 (VR2Market). J. M. Carvalho acknowledges the Doctoral Grant from FCT, ref. SFRH/BD/136815/2018.

References

1. Pinho, A.J., Ferreira, P.: Finding unknown repeated patterns in images. 19th Eur. Signal Process. Conf. (EUSIPCO 2011) (2011)

2. Pinho, A., Ferreira, P.: Image similarity using the normalized compression distance based on finite context models. 18th IEEE International Conference on Image Processing (2011)

3. Pinho, A., Pratas, D., Ferreira, P.: A new compressor for measuring distances among images. Image Analysis and Recognition 1 (2014) 30–37

4. Pinho, A.J., Pratas, D., Ferreira, P.J.S.G.: Authorship attribution using compression distances. In: Data Compression Conference. (2016)

5. Pratas, D., Pinho, A.J.: A conditional compression distance that unveils insights of the genomic evolution. In: 2014 Data Compression Conference, IEEE (mar 2014) 421–421

6. Pratas, D., Pinho, A.J., Silva, R.M., Rodrigues, J., Hosseini, M., Caetano, T., Ferreira, P.: Falcon: a method to infer metagenomic composition of ancient dna. bioRxiv (jan 2018)

7. Brás, S., Pinho, A.J.: Ecg biometric identification: A compression based approach. In: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). (aug 2015) 5838–5841

8. Carvalho, J.M., Brás, S., Pratas, D., Ferreira, J., Soares, S.C., Pinho, A.J.: Extended-alphabet finite-context models. Pattern Recognition Letters 112 (sep2018) 49–55

9. Carvalho, J.M., Brás, S., Pinho, A.J.: Compression-based classification of ecg using first-order derivatives. In: Proceedings of the 10th International Conference on Intelligent Technologies for Interactive Entertainment. (2018)

10. Bousseljot R, Kreiseler D, S.A.N.: Ekg-signaldatenbank cardiodat. Biomedizinische Technik (1995)

11. Goldberger, A.L., Amaral, L.A.N., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.K., Stanley, H.E.: Physiobank, physiotoolkit, and physionet. Circulation 101(23) (2000) e215–e220