Joint Training of Hidden Markov Model and Deep Neural Network for Heart Sound Segmentation

Francesco Renna[1], Miguel L. Martins[2], Miguel T. Coimbra[2]

[1]Instituto de Telecomunicações, Porto, Portugal
[2]INESC TEC, Porto, Portugal

Aims: Phonocardiogram (PCG) segmentation is a fundamental step in automatic heart sound analysis that allows to identify the position of possible extra sounds (murmurs, clicks, etc.) and to further process and analyze individual components, thus providing precious information regarding the mechanical activity of the heart. State-of-the-art approaches for PCG segmentation have recently shifted from the application of Markov models to deep neural networks, due to the enhanced discriminative power of the latter. On the other hand, deep networks can be prone to overfitting and do not encapsulate *per se* information regarding the quasi-periodic nature of heart sounds.

The proposed method aims to combine the advantages of both these approaches via the use of a hybrid model-based/data-driven framework.

Methods: The proposed model embeds prior information into deep neural networks via the combination of an underlying hidden Markov model (HMM), which regulates the dynamics of state transitions in the PCG, and a convolutional neural network (CNN) that is used to evaluate the HMM emission distributions. The overall model is trained end-to-end via gradient-based optimization and using a maximum mutual information criterion.

Results: The proposed approach was tested over the PhysioNet 2016 dataset (10-fold cross-validation), assessing performance in terms of average sample accuracy (ACC), sensitivity (SE), and positive predictive value (PPV). The proposed method, which combined a 4-state HMM with a vanilla CNN (3 convolutional layers and 2 fully connected layers), obtained the following results: ACC 92.8%, SE 94.6%, PPV 94.5%, slightly outperforming the state-of-the-art, U-Net-based method by Renna et al. (2019): ACC 91.8%, SE 92.5%, PPV 93.8%.
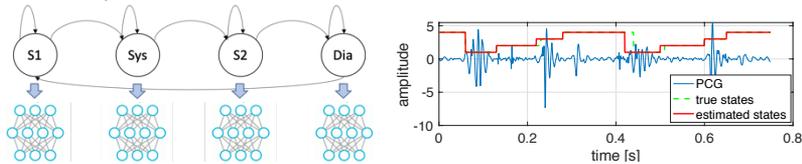

Fig. 1: Scheme of the proposed hybrid HMM-CNN model and segmentation example.

Conclusions: The reported results, though not conclusive, show that joint training of a simple CNN with an underlying HMM can potentially provide segmentation performance in line with (or even superior to) current, more sophisticated CNN architectures.