COMPUTING IN CARDIOLOGY 2022

Exploring a Segmentation-Classification Deep Learning-based Heart Murmurs Detector

Daniel Enériz[1], Antonio J. Rodríguez-Almeida[2], Himar Fabelo[2,3], Samuel Ortega[4], Francisco Balea-Fernandez[2,5], Nicolás Medrano[1], Belén Calvo[1], Gustavo M. Callicó[2]

[1]Aragon Institute of Engineering Research (I3A), University of Zaragoza (Unizar), Zaragoza, Spain
[2]Institute for Applied Microelectronics (IUMA), University of Las Palmas de Gran Canaria (ULPGC), Las Palmas de G. C., Spain
[3]Fundación Canaria Instituto de Investigación Sanitaria de Canarias, Las Palmas de G. C., Spain
[4]Norwegian Institute of Food Fisheries and Aquaculture Research (Nofima), Tromsø, Norway
[5]Dept. of Psychology, Sociology and Social Work, ULPGC, Las Palmas de Gran Canaria, Spain

Introduction

- Heart Murmur Detection from Phonocardiogram Recordings: The George B. Moody PhysioNet Challenge 2022
- Two targets:
  - Murmur presence detection
  - Clinical outcome prediction
- Data: CirCor Digiscope dataset
  - PCG recordings
  - Different auscultation locations
  - Heart state annotations
  - Presence and outcome labels
  - Murmur description (loc., pitch, shape)
  - Demographics

Results

Segmentation similar to Renna *et al.* 2019

	Acc. (%)	Sens. (%)	PPV. (%)
Renna *et al.*	91.5±1.6	91.2±2.3	94.1±2.1
Ours	90.2±0.9	96.0±1.1	96.0±1.0

Classification

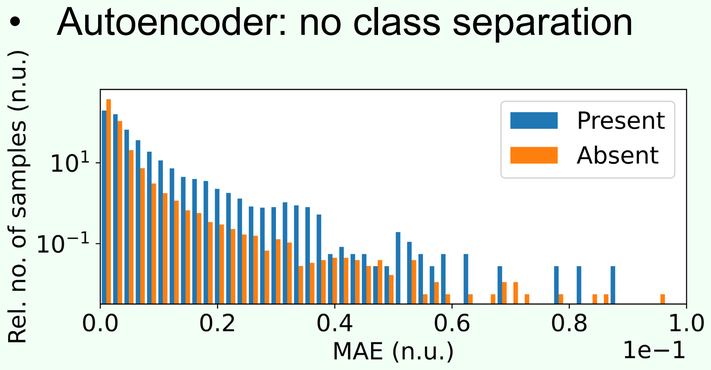
- Murmur presence

Model	Acc.	AUROC	Recall	Specifity
Original	0.58±0.02	0.81±0.02	0.46±0.03	0.89±0.01
Variation	0.56±0.02	0.80±0.02	0.41±0.03	0.91±0.01

- Outcome

Model	Acc.	AUROC	Recall	Specifity
Original	0.60±0.03	0.64±0.04	0.60±0.03	0.60±0.03
Variation	0.60±0.03	0.64±0.04	0.60±0.03	0.60±0.03

- Autoencoder: no class separation



Global rule (final architecture results)

- Presence

Model	AUROC	AUPRC	F-measure
Original	0.89±0.05	0.79±0.09	0.41±0.12
Variation	0.79±0.09	0.65±0.09	0.45±0.09

Model	Acc.	W. Acc.	Challenge cost
Original	0.79±0.07	0.54±0.14	22042±4082
Variation	0.80±0.06	0.58±0.10	21446±3777

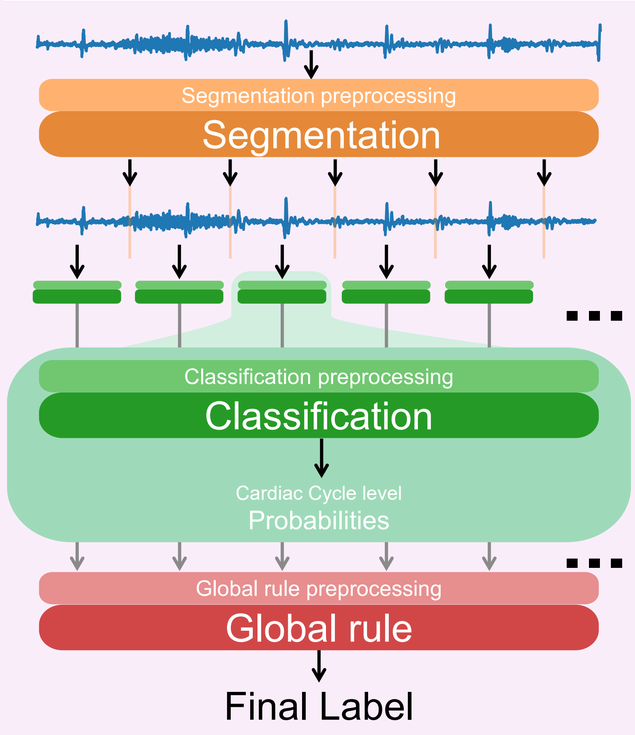
- Outcome

Model	AUROC	AUPRC	F-measure
Original	0.69±0.03	0.67±0.03	0.78±0.07
Variation	0.63±0.03	0.61±0.03	0.70±0.05

Model	Acc.	W. Acc.	Challenge cost
Original	0.79±0.06	0.74±0.09	10735±2208
Variation	0.70±0.05	0.64±0.06	12718±2176

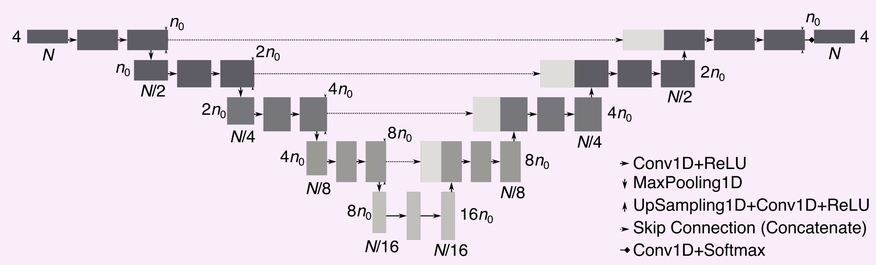
All evaluations are done with 10-fold cross validation over the publicly available data. Unfortunately, no evaluation over the hidden data could be obtained

Method



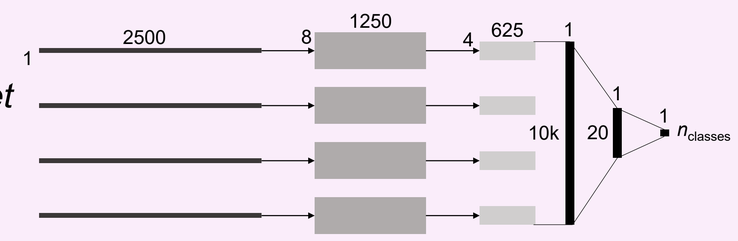
Segmentation (Renna *et al.* 2019)

- Preprocessing:
  - BP filter: 25 -40 Hz
  - Schmidt *et al.* spike removal
  - 4 envelograms: Hilbert, Homomorphic, PSD and Wavelet
  - Downsampling to 50 Hz
  - Windowing: N=64, τ=8
- Model: 1D U-Net
- Training:
  - 5 epochs in 2016 challenge data
  - 5 epochs in 2022 challenge data
  - Batch size = 1, learning rate = 10^{-4}



Classification (Three trial models)

- Preprocessing:
  - BP filter between 25 and 40 Hz, Schmidt *et al.* spike removal, downsampling to 1 kHz
  - 2.5 seconds cardiac cycle extraction with the segmentation. If shorter, zero padding
- *Model 1*: Potes *et al.* 2016 Original
  - 4 BP filtered signals: [25, 45], [45, 80], [80, 200] and [200, 400] in Hz
  - 4 independent 1D extractors: 2×Conv1D+ReLU+MaxPool(→)
  - MLP, 20 neurons in the hidden layer
  - Training: 100 epochs, bs =1024, lr = $7{\cdot}10^{-4}$
- *Model 2*: Potes *et al.* 2016 Variation
  - Non independent extractors. Smaller
  - Same training
- *Model 3*: Autoencoder, anomaly detection
  - Adapted from the Segmentation model (Tab.)
  - Compression in both time and channels
  - Trained only in the normal samples
  - Compute error between the input and the output
  - Threshold to separate classes



Block/Layer	# filters	# params.	Output shape
Input	-	-	(2496, 1)
Encoder_0	64	12480	(1248, 64)
Encoder_1	32	9216	(624, 32)
Encoder_2	16	2304	(312, 16)
Encoder_3	8	576	(156, 8)
Conv1D	4	96	(156, 4)
Conv1D	4	48	(156, 4)
Decoder_0	8	288	(312, 8)
Decoder_1	16	1152	(624, 16)
Decoder_2	32	4608	(1248, 32)
Decoder_3	64	18432	(2496, 64)
Conv1D	1	192	(2496, 1)

Global rule

- Preprocessing: Statical metrics extraction
  - Mean, Std. and Maximal per each class
- Model: MLP, hidden layer of 5 neurons
- Training: 10 epochs, bs =16, lr = 10^{-3}

Discussion and Conclusions

- Limited performance in all classifiers for both tasks, thus modest throughput in the final model
- Realistic scenario has
  - Intrinsic noises: respiration, frictions…
  - External noises: voices, cars…
- A more sophisticated preprocessing in the classification stage to address this noises' sources would improve results
- For the presence task, there is a high unbalance of classes at the cardiac cycle level, since patients with no murmur have longer PCGs. This can be the reason for the better performance shown in the outcome task.

Contact and more info

Daniel Enériz
eneriz@unizar.es

Antonio J. Rodriguez-Almeida
aralmeida@iuma.ulpgc.es

Instituto Universitario de Investigación en Ingeniería de Aragón
Universidad Zaragoza

ULPGC Universidad de Las Palmas de Gran Canaria
Instituto Universitario de Microelectrónica Aplicada
IUMA

This work has been supported by the "Ministerio de Ciencia e Innovación" through the projects PID2019-106570RB-I00 and PID2020-116417RB-C42; the Goverment of Aragón through the PhD grants BOA20201210014; the University of Zaragoza, the "Fundación Bancaria Ibercaja" and the "Fundación CAI" through the research stay grant IT 9/21; the "Agencia Canaria de Investigación, Innovación y Sociedad de la Información"; and the European Union through the NextGenerationEU/PRTR funds