8266: A Unified Joint Contrastive Triplet Loss with Temporal and Frequency Signal Fusion for Diagnosing Heart Murmurs

by:

Ayushi Pal, Arka Roy, and Udit Satija

Department of Electrical Engineering

Indian Institute of Technology Patna, Bihar, 801106

Content

Introduction

- Cardiovascular disorders (CVDs)- group of disorders present in the heart and the blood vessels [1].
  - **Cause:** hypertension, hyperlipidemia, lack of physical activity, over usage of alcohol.
  - **General symptoms:** angina, dyspnea, fatigue, dizziness.
  - **Different types:** coronary artery diseases, arrhythmia, peripheral artery diseases, heart valve disorders [1].
- According to WHO, CVDs are the leading cause of death globally. An estimated 17.9 million people died from CVDs in 2019, representing 32% of all global deaths [1].

[1] https://my.clevelandclinic.org/21493-cardiovascular-disease

- Heart valve disorders (HVDs)- disfunctionality of heart valves at the open and closure time [2].
- Heart consists of four valves named as AV (Aortic valve), TV (Tricuspid valve), MV (Mitral valve), and PV (Pulmonary valve).
- Due to the improper function of these valves, the following HVDs occur:
  - Atresia: valve opening is absent.
  - Regurgitation: leakage is present.
  - Stenosis: too small valve opening.

Introduction

- **Modalities used:** catheterization, echocardiography, and X-ray [3].
- **Limitations:** discomfort, painful, long process.
- **Why any other modality??**
  - Early detection.
  - Low cost non-invasive modality.
- A phonocardiogram (PCG) signal records the vibrations produced in the heart [4].

Table 1: Different events occurring in heart sound PCG signal [5].

S1	S2	S3	S4
Closing of mitral and tricuspid valve	Closing of aortic and pulmonary valve	Occurs 120 to 180 ms after onset of S2	Occurs 90 ms before S1
40-200 Hz	50-250 Hz	50-90 Hz	50-80 Hz
70-150 ms	60-120 ms	40-100 ms	40-80 ms

Infant HR= 70-190 bpm, adult HR= 60 to 100 bpm, cardiac cycle=800 ms, murmur freq= 200-400 Hz.

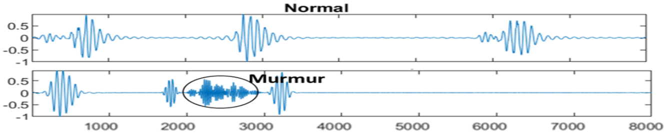

Figure 1: Fig. showing PCG signal (a) normal, and (b) with murmur.

Motivation & contribution

Motivation

- Heart murmurs are critical indicators of HVDs and require early detection for timely intervention.
- Existing methods primarily use time or freq domain analysis, mainly on the adult age group, which struggles with generalization across all patient age groups.
- Phonocardiogram (PCG) signals are complex and prone to noise, necessitating a robust diagnostic approach.

Contributions

- **Joint Time-Frequency Fusion:** A novel framework integrating both time and frequency domain representations for improved murmur classification.
- **Contrastive Triplet Loss Optimization:** Introduction of a joint contrastive triplet loss to enhance intra-domain and cross-domain feature learning.
- **Lightweight Deep Learning Architecture:** Utilization of depthwise convolution and modified inception blocks (MIBs) for efficient embedding extraction.
- **Superior Performance:** Achieved **89.87% classification accuracy**, surpassing existing methods on the CirCor DigiScope dataset.

Database description

Database used: circor heart sound database [6]

- A large pediatric and adult heart sound dataset.
- Data collection: two mass screening campaigns (Brazil, 2014 & 2015).
- Total 1568 patients of different age groups.
- Total 5282 heart sound recordings.
- Age groups: 0.1 to 356.1 months.
- Time range: 4.8 to 80.4 seconds.
- Sampling frequency: 4000 Hz.
- Contains vast information about timing, pitch, grading, quality, and so on.
- Used for three classes classification as heart murmur present (HMP), heart murmur absent (HMA), and heart murmur unknown (HMU).

Time and frequency domain representation of PCG signals

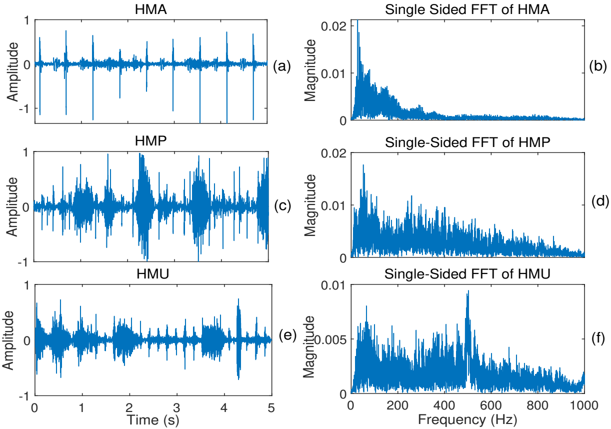

Figure 2: Demonstrates the three different types of HM present in PCG signal of 5-second with corresponding single-sided FFT spectrum. Fig. (a), (c), and (e) show the time domain PCG signal while Fig. (b), (d), and (f) show the single-sided spectrum of FFT calculated for each signal.

Block diagram

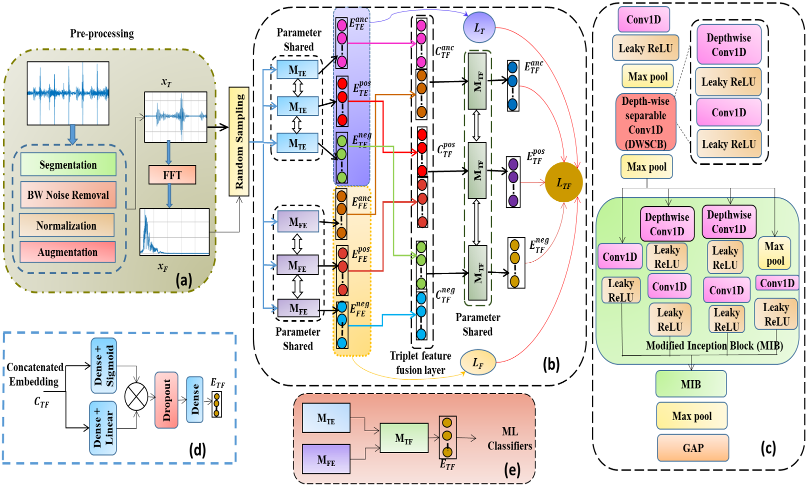

Figure 3: Block diagram of the proposed method (a) shows the pre-processing done on the resampled PCG signals, (b) shows the embedding extraction for the time and frequency domain and then fusing them to get the frozen embedding, (c) shows the detailed model architecture of the triplet network M_{TE} and M_{FE} used for time and frequency embedding respectively, (d) shows the detailed model of the network M_{TF} used for cross-domain embedding extraction, and (e) shows the overall proposed methodology used for three class classification of HMs.

Stage 1: preprocessing

Steps include:

- Resampled to 2KHz.
- Segmented the signal into 5-sec segment with 50% overlapping window using

$$HS_q[m] = HS_r[b \cdot q \cdot (1-s) + m], \qquad m = 0, 1, 2, \ldots M-1, \tag{1}$$

  where $HS_q[m]$ denotes the q^{th} segmented frame of the resampled PCG signal with M number of samples in each segmented signal.
- Baseline Wander (BW) noise is removed using Discrete Fourier Transform (DFT) [7].
- Normalization is done in the range of [-1,1].
- Augmentation is done to overcome data imbalancing.
- Embedding extraction is done using
  - Time domain embedding using the segmented time domain signal.
  - Frequency domain embedding using the fast Fourier transform (FFT) of each segment.

Stage 2- proposed joint TF triplet loss and the trainable parameters breakdown

- The final loss function (L) can be formulated as:

$$L = L_{TF} + \lambda \cdot L_T + (1 - \lambda) \cdot L_F, \tag{2a}$$

$$L_T = \max\{0, \mathcal{D}(E_{TE}^{anc}, E_{TE}^{pos}) - \mathcal{D}(E_{TE}^{anc}, E_{TE}^{neg}) + \zeta\}, \tag{2b}$$

$$L_F = \max\{0, \mathcal{D}(E_{FE}^{anc}, E_{FE}^{pos}) - \mathcal{D}(E_{FE}^{anc}, E_{FE}^{neg}) + \zeta\}, \tag{2c}$$

$$L_{TF} = \max\{0, \mathcal{D}(E_{TF}^{anc}, E_{TF}^{pos}) - \mathcal{D}(E_{TF}^{anc}, E_{TF}^{neg}) + \zeta\}, \tag{2d}$$

  where $\lambda, and\ \zeta$ denote the loss scaling factor and margin parameter.
- Ensures robust feature representation by aligning intra-domain and cross-domain embeddings.
- Frozen embeddings are fed to ML classifiers like SVM and KNN to detect murmurs.
- The parameter breakdown of the proposed methodology is shown in table below:

Model name	Parameters
Time Encoder ((M_{TE}))	78,336
Frequency Encoder ((M_{FE}))	78,336
Time-freq Network ((M_{TF}))	33,280
Total Trainable Parameters	**189,952**

Results

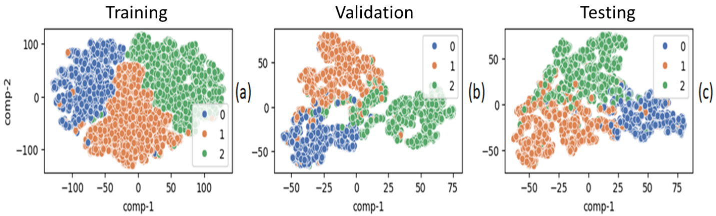

Figure 4: Illustartes the t-SNE plots for (a) embedding before training the network, (b) validation instances embedding, and (c) testing instances embedding.

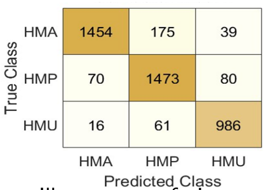

Figure 5: Illustrates confusion matrix obtained using the proposed method.

Table 2: Performance measures obtained using the proposed method.

Classifier	acc (%)	prec (%)	rec (%)	F1-score (%)
SVM	89.02	89.30	89.02	89.05
KNN	89.87	90.08	89.87	89.89

Performance comparison with existing methods

Table 3: Comparative Analysis for Three classes of HMs using Circor Database.

S. No.	Authors	Signal	# of Subjects	Method	Results Obtained (%)
1	Singstad et al. [8]	Time domain	All	1D-CNN	w.acc- 59.3, acc-49.7, F1-score- 39.8
2	Chen et al. [9]	Time domain	942	PCG-AResnet	acc- 94.58, F1-score- 92.22
3	Walker et al. [10]	Mel-spectrograms	942	Dual bayesian ResNet	w.acc- 77.1
4	Niizumi et al. [11]	Time domain	All	Self supervised learning masked modeling duo	w.acc- 83.2, rec- 71.3
5	Zhang et al. [12]	Mel-spectrograms	All	Parallel attentive model (self -attention and CNN)	acc- 79.8, F1-score - 65.1
6	**Proposed Method**	**Time and frequency domain**	**All**	**Supervised contrastive embedding learning based triplet loss network**	**w.acc- 87.88, acc- 89.87, prec- 90.08, rec- 89.87, F1-score - 89.89**

Conclusion

- A novel deep learning framework for diagnosing heart murmurs by fusing time and frequency domain embeddings.
- The proposed contrastive learning-based triplet loss approach effectively captures both intra-domain and cross-domain features, improving classification accuracy.
- Our method achieves state-of-the-art performance in terms of accuracy (89.87%) with less number of parameters, surpassing all existing works.
- Future work includes real-time implementation and validation on larger datasets.

Scan for code and more information

References I

Cardiovascular diseases World Health Organization. Available at: https://www.who.int/health-topics/cardiovascular-diseasestab=tab1.

Mubarak, Q., Akram, M. U., Shaukat, A., Hussain, F., Khawaja, S. G., Butt, W. H. (2018). Analysis of PCG signals using quality assessment and homomorphic filters for localization and classification of heart sounds. Computer Methods and Programs in Biomedicine, 164, 143-157. https://doi.org/10.1016/j.cmpb.2018.07.006.

Martel, J. (no date) Heart valve disorders: Causes, symptoms, and diagnosis, Healthline. Available at: https://www.healthline.com/health/heart/valve-disorderstreatments.

Yaseen, Gui-Young Son, and Soonil Kwon. 2018. "Classification of Heart Sound Signal Using Multiple Features" Applied Sciences 8, no. 12: 2344. https://doi.org/10.3390/app8122344.

Liu, Chengyu et al. "An open access database for the evaluation of heart sound algorithms." Physiological measurement vol. 37,12 (2016): 2181-2213. doi:10.1088/0967-3334/37/12/2181

Oliveira, J., Renna, F., Costa, P., Nogueira, M., Oliveira, A. C., Elola, A., Ferreira, C., Jorge, A., Bahrami Rad, A., Reyna, M., Sameni, R., Clifford, G., Coimbra, M. (2022). The CirCor DigiScope Phonocardiogram Dataset (version 1.0.2). PhysioNet. https://doi.org/10.13026/qcpr-xg17.

Satija, U., Ramkumar, B. and Manikandan, M.S., 2017. Real-time signal quality-aware ECG telemetry system for IoT-based health care monitoring. IEEE Internet of Things Journal, 4(3), pp.815-823.

B. -J. Singstad, A. M. Gitau, M. K. Johnsen, J. Ravn, L. A. Bongo and H. Schirmer, "Phonocardiogram Classification Using 1-Dimensional Inception Time Convolutional Neural Networks," 2022 Computing in Cardiology (CinC), Tampere, Finland, 2022, pp. 1-4, doi: 10.22489/CinC.2022.108

References II

Y. Chen, W. Zhu, J. Xu, J. Zhang, Z. Zhu and L. Wang, "A Heart Sound Classification Method Based on Residual Block and Attention Mechanism," 2022 IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), Wuhan, China, 2022, pp. 1060-1065, doi: 10.1109/TrustCom56396.2022.00145.

Walker, B., Krones, F., Kiskin, I., Parsons, G., Lyons, T., Mahdi, A. (2023). Dual Bayesian ResNet: A Deep Learning Approach to Heart Murmur Detection. ArXiv. https://doi.org/10.22489/cinc.2022.355

Niizumi, D., Takeuchi, D., Ohishi, Y., Harada, N., Kashino, K. (2024). Exploring Pre-trained General-purpose Audio Representations for Heart Murmur Detection. ArXiv. https://arxiv.org/abs/2404.17107.

Z. Zhang, T. Pang, J. Han and B. W. Schuller, "Intelligent Cardiac Auscultation for Murmur Detection via Parallel-Attentive Models with Uncertainty Estimation," ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Seoul, Korea, Republic of, 2024, pp. 861-865, doi: 10.1109/ICASSP48485.2024.10446246.

Thank You