Short-Segment Heart Sound Classification Using an Ensemble of Deep Convolutional Neural Networks

Fuad Noman[1], Chee-Ming Ting[1,2], Sh-Hussain Salleh[1], Hernando Ombao[2]

[1]School of Biomedical Engineering & Health Sciences, Universiti Teknologi Malaysia, Malaysia
[2]Biostatistics Group, King Abdullah University of Science and Technology, Saudi Arabia.

Motivation

Cardiac auscultation based on heart sound recordings or phonocardiogram (PCG) is a primary screening tool for diverse heart pathologies. Various algorithms have been developed for automated classification of normal and abnormal PCGs [1].

Challenges:

1. **Performance**: The classification accuracy of current methods is still far from being reliable for diagnostics in clinical or non-clinical settings.
2. **Noise**: One major challenge is to extract robust and discriminative features from the raw PCG recordings typically corrupted by various noise sources.
3. **Data length**: Short-segment PCG classification is a challenging task where most of the widely used feature maps can only be extracted from long recordings containing many cardiac cycles.

Contributions

We aim to classify normal and abnormal heart sounds based on short-segments of individual heart beats (single cardiac cycle)

1. We propose a deep CNN for classification of pathology in PCG of a single heart beat.
2. We design a new architecture called time-frequency ensemble CNN (TF-ECNN) that combines a 1D-CNN and a 2D-CNN to learn multiple levels of representations respectively from the time-domain raw PCG signals and time-frequency MFCC features as inputs.
3. The proposed TF-ECNN shows improved classification performance over strong state-of-the-art baseline classifiers and feature sets.

Database

We evaluate our method on a large heart sound dataset from PhysioNet CinC challenge 2016 [2]

Table 1 Distribution of train and test set

	Train		Test	
	normal	abnormal	normal	abnormal
Recordings	1150	284	1150	288
Heartbeats	32574	8170	32582	8177

Heart Sound Classification

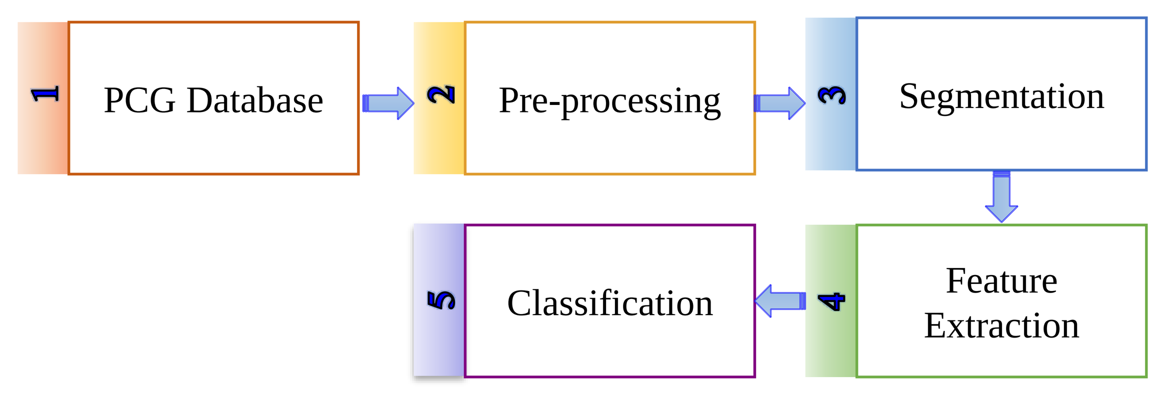

Fig. 1 Overview of PCG classification system

Pre-processing and Segmentation

- **Down-sampling** of heart sound signals to 1000 Hz
- **Band-pass filtering** at 25 - 400 Hz to eliminate unwanted low-frequency artifacts (e.g., baseline drift) and high-frequency noise (e.g., background noise)
- **Normalization** by mean subtraction and division by its standard deviation
- **Segmentation** of each recording into individual heartbeats (from beginning of atrial activity to end of ventricular activity) based on expert annotations

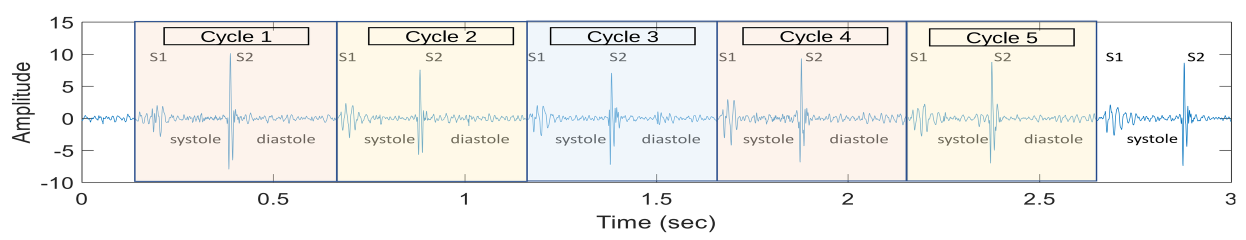

Fig. 2. Segmentation of PCG into cardiac cycles

Feature Extraction

The proposed TF-ECNN classifier accepts combination of 1D and 2D time-frequency features as inputs:

1. **1D-CNN**: accepts one-dimensional PCG time series data (i.e., the raw heartbeat signal)
2. **2D-CNN**: accepts two-dimensional time-frequency feature maps of Mel-frequency cepstral coefficients (MFCCs) and time-varying autoregressive (TV-AR) coefficients

References

[1] Clifford et al. Recent advances in heart sound analysis. *Physiol. Meas. 2017*

[2] Liu et al. An open access database for the evaluation of heart sound algorithms. *Physiol. Meas. 2016*

[3] Noman et al. A Markov-switching model approach to heart sound segmentation and classification. *arXiv:1809.03395 2018*

Time-Frequency Ensemble CNN Classifier

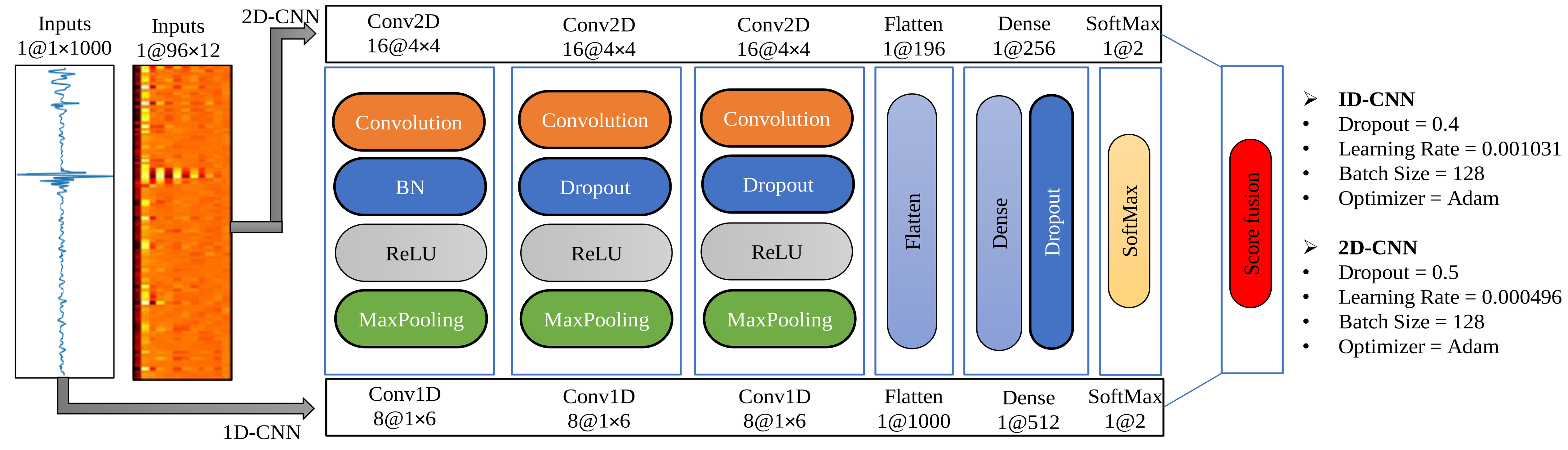

Fig. 3. TF-ECNN model architecture combining 1D-CNN and 2D-CNN, with inputs of raw signals and time-frequency feature maps, respectively. BN: Batch-normalization layer. ReLU: rectified linear unit activation function.

Hierarchical Time-Frequency Features

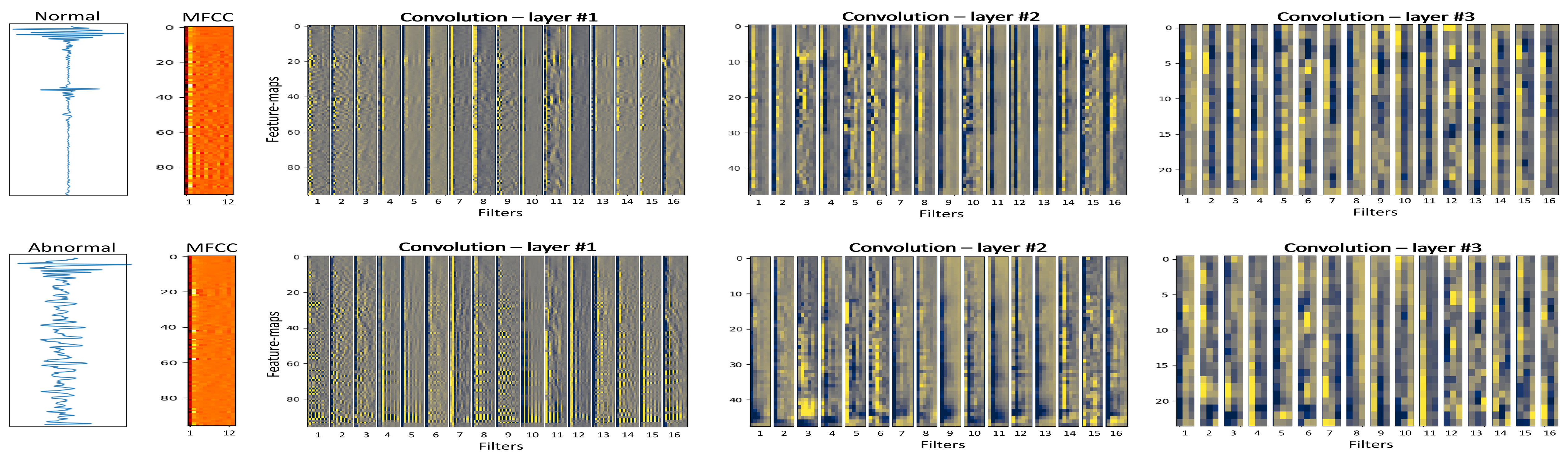

Fig. 4. Visualization of feature maps of the convolutional layers of the 2D-CNN learned from MFCC inputs for normal and abnormal heartbeats.

Results

Table 2. Performance comparison of different classifiers on the test set

Classifier	Features	Accuracy (%)	Sensitivity (%)	Specificity (%)	MAcc (%)
SVM	Time & Freq	84.87 (85.09)	85.82 (94.09)	81.09 (48.95)	83.46 (71.52)
Tree Ensemble	Time & Freq	86.20 (86.23)	**90.55** (94.25)	68.84 (54.26)	79.70 (74.26)
HMM	MFCC	87.07 (n/a)	85.97 (n/a)	91.45 (n/a)	88.71 (n/a)
1D-CNN	Raw (zero-pad)	86.34 (85.63)	87.80 (95.11)	80.32 (46.41)	84.06 (70.76)
	Raw (norm-dur)	87.23 (87.52)	87.57 (91.51)	85.84 (71.64)	86.71 (81.58)
2D-CNN	TVAR	86.41 (86.91)	88.85 (91.79)	76.69 (67.45)	82.77 (79.62)
	MFCC	87.18 (89.30)	86.08 (92.49)	**91.55** (76.61)	**88.82** (84.55)
ECNN	Raw (norm-dur) + MFCC	**89.22** (89.58)	89.94 (93.07)	86.35 (75.68)	88.15 (84.37)

Numbers in parentheses correspond to performance before applying weight cost for imbalanced classes