# 1 Editor

2 Reviewer 1

Issue 0:

> In truth, the purpose of the work is not precisely defined. Reading the article, one has the impression that one will take an ECG signal, apply a wavelet transform to de-noise the signal, then count the STFT to get a spectrogram and finally apply a CNN to classify the signal with and without heart disease. To do this using the well-known tools available in python. Where is the novelty here?

Answer: Thank you for this issue. The novelty in our work is representation of data to convolutional neural network model in well understood i.e., Spectrograms. Moreover, we applied the frequency filtration through STFT with purpose of reducing data at possible limit at which a simple architecture of CNN model showed high accuracy rather than to utilize the complex model which ultimately required more momeory and computational power.
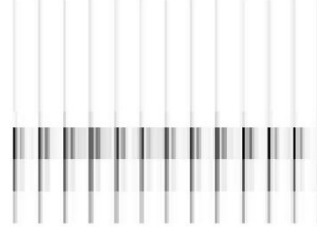
Issue 1:

> The work concerns the classification of ECG signals, and in particular "Anteroseptal myocardial infarction (ASMI) considered in this study is one of the life-threatening heart diseases that occurs due to the rupture of the volatile atherosclerotic plaque in the left anterior descending artery [3]." In the entire paper, there is not a single example of an ECG signal waveform with this cardiovascular disease, nor is it shown how the spectrogram of the ECG signal differs between a healthy person and one with ASMI or other cardiovascular disease.

Answer: Thank you for this issue. We considered Lead V1 data for the ECG signals to differentiate between NORMAL and ASMI sample. The difference can be find in the base of the signals. As asked, the ECG signals and Spectrogram of both cateogories are presented in below figure. One can see a difference between NORM and ASMI spectorgram at signal base. The shadow is spreaded in ASMI spectrogram while it is condensed in NORMAL spectrogram.
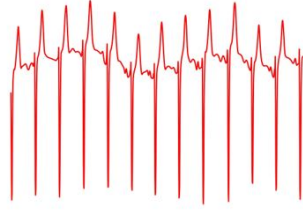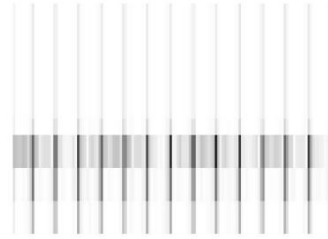
(a) Normal ECG Signal (Lead V1)

(b) Normal ECG Signal Spectrogram
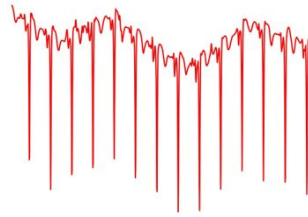
(c) ASMI ECG Signal (Lead V1)

(d) ASMI ECG Signal Spectrogram

Issue 2:

> The Short Time Fourier Transform (STFT) is one of the time-frequency transformation method existing in the signal processing. What other time-frequency transform methods have been considered?

Answer: Thanks for asking. In this study, we utilized the STFT for time-frequency transformation because it is comparatively faster and has good precision as highlighted in [46]. The other possible time-frequency transformation methods are Fast Fourier Transformation [45], synchrosqueezed transform [46] and Continuous Wavelet Transformation [47].

Issue 3:

> Quite a few modern (recently published) methods of noise reducing in the ECG signal have been presented, so why was the wavelet transform used, which has been known for more than a dozen years and is known not to work effectively in all cases, which is important when analyzing ECG signals with heart disease.

Answer: Thank you for raising this issue. The main reason for preference of wavelet transformation for signal denoising is that it offers range of wavelets i.e., db, haar, coif, sym and bior. Therefore, there are number of options available under single approach to find the best matching wavelet for the presented problem.

Issue 4:

Section 3.1.2 "Frequency Filtration". There is a lack of mathematical formulas for the steps involved in implementing the proposed algorithm. There is a description of the algorithm (Algorithm 2), but it is insufficient for a correct representation of the filtration. Why is the Sampling Frequency value calculated in step 3? Is it a different frequency than the one at which the signal was recorded? In the next step (del (...)) the numeric value of 2 was used. Where does this value come from? What is its meaning? This operation suggests cutting all frequency components above a certain frequency value in the frequency spectrum. It is worth showing the frequency spectrum of the analyzed signal after such an operation and the signal in the time domain. How does this operation affect the spectrograms obtained later (by STFT)? This operation suggests cutting all frequency components above a certain frequency value in the frequency spectrum. It is worth showing the frequency spectrum of the analyzed signal after such an operation and the signal in the time domain. How does this operation affect the spectrograms obtained later (by STFT)?

Issue 4a:

Section 3.1.2 "Frequency Filtration". There is a lack of mathematical formulas for the steps involved in implementing the proposed algorithm. There is a description of the algorithm (Algorithm 2), but it is insufficient for a correct representation of the filtration. Why is the Sampling Frequency value calculated in step 3? Is it a different frequency than the one at which the signal was recorded? In the next step (del (...)) the numeric value of 2 was used. Where does this value come from? What is its meaning? This operation suggests cutting all frequency components above a certain frequency value in the frequency spectrum. It is worth showing the frequency spectrum of the analyzed signal after such an operation and the signal in the time domain. How does this operation affect the spectrograms obtained later (by STFT)?

Answer: Thank you for this issue. The signal was decompose into time-frequency domain and then corresponding frequencies were calculated through the sampling frequency which were used in the next stage to apply the threshold for frequency filtration. As mentioned, the said sampling frequency is not different than at which signal was recorded but it is a type of variable in which corresponding frequencies were saved. Since, one of the study aim was to reduce the data, therefore, by analysing ECG signals visually and by experiments we set the threshold of 2 which means all the data points associated with frequency more than two i.e., high QRS peaks (downward in lead V1) will be discard because the signals can classify through the base patterns.

Issue 4b:

"Frequency Filtration". This operation suggests cutting all frequency components above a certain frequency value in the frequency spectrum. It is worth showing the frequency spectrum of the analyzed signal after such an operation and the signal in the time domain. How does this operation affect the spectrograms obtained later (by STFT)?

Answer: Thank you for this issue. The figure has been added.

(a) Before Freq. Filtration   (b) After Freq. Filtration
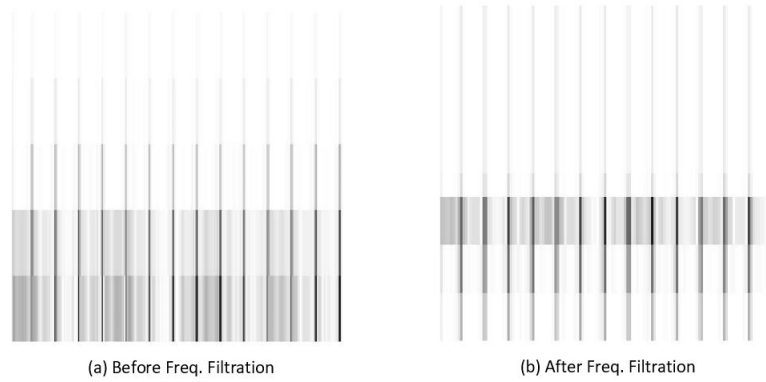
In the above figure, (a) is showing spectrogram before frequency filtration while (b) is showing after operation. One can see the black shadow reduced in fig (b) due to frequency filtration operation.

Issue 5:

What are the characteristics of the PTB-XL base signals used in this study (e.g., sampling rate, signal length)? What is the case count in each class of training, test and validation sets?

Issue 5a:

What are the characteristics of the PTB-XL base signals used in this study (e.g., sampling rate, signal length)?

Answer: Thank you for raising this issue. The PTB-XL signal were originaly calculated at sampling rate of 100 Hz and 500 Hz. However, we choose the 100 Hz sampling rate, in which each signal measured for 10 seconds of the length.

Issue 5b:

What is the case count in each class of training, test and validation sets?

Answer: Thank you for comment. The complete dataset contains 21,837 samples, out of which majority belongs to Normal class. However, we divide the data into two classes i.e., NORM and ASMI. The NORM class contains 18,954 (15,564 training and 3,390 validation). The ASMI class has 2,283 samples, therefore, we applied two data augmentation approaches on spectrograms i.e., horizontal flip and contrast to handle the class imbalance. We also performed some sample duplication in training set of ASMI class to achieve the equal figure.

Issue 6:

> Section 3.3 "Model Architecture". What criterion was used to select the structure of the CNN network? Why was the four-layer neural network model chosen?

Answer: Thank you for asking. Firstly, we choose the six number of layers randomly and attempted model training with different hyperparameters. However, the accuracy was not good as it is. Therefore, we reduce the number of layers and re-tune the hyperparameters. The best accuracy achieved at four layers and with the mentioned hyperparameters.

Issue 7:

> "Learning Rate". Figure 5. It is conventionally assumed that numerical values are presented in ascending order (in line 338 it is the reverse). Unfortunately, this can lead to erroneous conclusions when analyzing the waveforms presented in Figure 5. I would suggest using a logarithmic scale on the X axis and an ascending order of learning rate values.

Answer: Thank you for pointing out. The presented learning rates are in ascending order both at line 338 and in the figure X axis.

Issue 8:

> From description of PTB-XL database given at "PTB-XL, a large publicly available electrocardiography dataset v1.0.2 (physionet.org)" (https://physionet.org/content/ptb-xl/1.0.2/) it follows that "The waveform files are stored in WaveForm DataBase (WFDB) format with 16 bit precision at a resolution of 1microV/LSB and a sampling frequency of 500Hz (records500/). For the user's convenience we also release a downsampled versions of the waveform data at a sampling frequency of 100Hz (records100/)." What is the purpose of oversampling the ECG signal? From Figure 6, it can be seen that the highest accuracy is achieved just for the sampling frequency of 100 Hz? Any other frequency means a decrease in accuracy, because

where to get additional information about the signal, when physically this information is not there. The results achieved are just the result of interpolation. Such a study would make sense if the original sampling frequency of the ECG signal was high enough, such as 1 kHz or 2kHz. Another issue is connecting the points on the graph with 6 straight lines, after all, no calculations were made for 159 Hz.

Issue 8a:

From description of PTB-XL database given at "PTB-XL, a large publicly available electrocardiography dataset v1.0.2 (physionet.org)" (https://physionet.org/content/ptb-xl/1.0.2/) it follows that "The waveform files are stored in WaveForm DataBase (WFDB) format with 16 bit precision at a resolution of 1microV/LSB and a sampling frequency of 500Hz (records500/). For the user's convenience we also release a downsampled versions of the waveform data at a sampling frequency of 100Hz (records100/)." What is the purpose of oversampling the ECG signal?

Answer: Thank you for this issue. Our preposition was to increase and decrease the sampling rate gradually w.r.t. actual learning rate with the purpose to test whether we can achieve the equivalent accuracy w.r.t. original sampling rate or not.

Issue 8b:

From Figure 6, it can be seen that the highest accuracy is achieved just for the sampling frequency of 100 Hz? Any other frequency means a decrease in accuracy, because where to get additional information about the signal, when physically this information is not there. The results achieved are just the result of interpolation. Such a study would make sense if the original sampling frequency of the ECG signal was high enough, such as 1 kHz or 2kHz. Another issue is connecting the points on the graph with 6 straight lines, after all, no calculations were made for 159 Hz.

Answer: Thank you for this issue. Yes, the highest accuracy achieved at 100 Hz sampling rate which is the original by the author of the PTB-XL dataset. Although, we did not get any success in this experiment because of data loss in downsampling and adding noise when performed upsampling but this way we filled a research hole. Similarly, when we realized that there is no good accuracy coming up, we limit the sampling rate to 1 kHz. The said experiment reflects that the idea of data agumentation through upsampling or downsampling will not be a good step. However, we are open to remove this part from our study upon your suggestion after the comment response.

Issue 9:

Table 3. Accuracy is not the only parameter used to assess classification, and such a high value (without specifying other factors that assess classification) may be indicative of overtraining. I suggest giving values such as PPV, Sensitivity (recall) or F1-score.

Answer: Thanks for the suggestion. Parameters like sensitivity, specficity and precision has been added to the table in contrast with other authors obtained metrics.

Issue 10:

10. In Table 1, two times reference [25] exists at different authors

Answer: Thanks for pointing out. By mistake, the second was written as 25 but it is 26 in real and has been rectified.

Issue 11:

The given link is inactive https://github.com/mfarhan166/ECG-Signals-and-Spectrograms

Answer: Thanks for checking the status of the link. The repository has the implementation code and that's why we set it as Private at this stage. We will change the status to Public as soon we will receive the acceptace from the journal. The repository is working fine at us.