A Convolutional Neural Network Based Approach to QRS Detection

Marko Šarlija
Faculty of Electrical
Engineering and Computing
University of Zagreb
Zagreb, Croatia
Email: marko.sarlija@fer.hr

Fran Jurišić
Faculty of Electrical
Engineering and Computing
University of Zagreb
Zagreb, Croatia
Email: fran.jurisic@fer.hr

Siniša Popović
Faculty of Electrical
Engineering and Computing
University of Zagreb
Zagreb, Croatia
Email: sinisa.popovic@fer.hr

Abstract—In this paper we present a QRS detection algorithm based on pattern recognition as well as a new approach to ECG baseline wander removal and signal normalization. Each point of the zero-centred and normalized ECG signal is a QRS candidate, while a 1-D CNN classifier serves as a decision rule. Positive outputs from the CNN are clustered to form final QRS detections. The data is obtained from the 44 non-pacemaker recordings of the MIT-BIH arrhythmia database. Classifier was trained on 22 recordings and the remaining ones are used for performance evaluation. Our method achieves a sensitivity of 99.81% and 99.93% positive predictive value, which is comparable with most state-of-the-art solutions. This approach opens new possibilities for improvements in heartbeat classification as well as P and T wave detection problems.

Keywords—Electrocardiogram (ECG), QRS complex detection, convolutional neural networks (CNN), clustering

I. Introduction

Electrocardiogram (ECG) reflects electrical activity of the heart, comprising different waveforms that represent either polarization or depolarization of the myocardium. QRS complex, the most prominent one, reflects ventricular contraction. Due to its specific shape and the fact that other ECG waveforms can be small, the QRS complex provides most significant information in the automatic detection of heartbeats. Improvements in quality of automatic heartbeat detection are important for features that greatly benefit from accurate heartbeat localization, like measures derived from heart rate variability [1]. Accurately aquired heart rate and heart rate variability have broad clinical importance, including applications in the area of e-health [2], as well as in psychophysiological research [3] and affective computing [1], [4].

The detection of QRS complexes is not always an easy task due to intersubject and physiological variability of the QRS, as well as the presence of artefacts like baseline drift, power-line interference and muscle activity. Review and comparison of traditional QRS detection methods can be found in [5], while a more recent revision of QRS detection algorithms can be found in [6]. Modern approaches include robust and well optimised linear methods [7], [8], compression of ECG morpholology [9], adaptive mathematical morphology [10], artifical neural networks [11], [12] and even exploration of various processing steps combinations [13]. Neural network approach to QRS detection stagnated in terms of accuracy in the past few years, with some exceptions in the form of patient-specific ECG classification by 1-D convolutional neural networks [14].

In this work, we aim to show that 1-D CNNs can be very accurate at subject-independent QRS complex detection. The use of 1-D CNNs enables detection of various heart beat morphologies without the hand crafted features or decision rules, which was demonstrated in [14] in subject-specific manner. Novel adaptations of existing mathematical methods in the ECG preprocessing step will be presented, together with the 1-D CNN classifier for QRS detection. Performance on a subject-independent test set and processing time will be reported.

II. Methods

A. Data

Data used comes from the MIT-BIH arrhythmia database [15], which includes both normal clinical recordings and recordings containing various complex arrhythmias. This database is a common reference as it has been extensively used to evaluate the performance of numerous QRS detection, heartbeat classification, and P or T wave detection algorithms. The database is publicly available [16] and the nature of the data is well described in [17]. Originally, the database contains 48 recordings from 47 different individuals. Signals are sampled at 360Hz. All recordings including modified lead II electrode placement (in most recordings lead A is lead MLII), that contain no paced beats are used, making a total of 44 used recordings.

Most of the ECG classification studies can be divided into 2 groups, "class-based" and "subject-based". As this is a QRS detection study, a "subject-based" approach is employed taking all of the AAMI recommended [18] types as unified beat annotations: all normal beat annotations (N, L, R, e, j), supraventricular ectopic beats (A, a, J, S), ventricular ectopic beats (V, e), fusion beats (F) and unknown beats (f, Q). This heartbeat type mapping is adopted from [19], with the exclusion of paced beats in the unknown beats class. Descriptions of the original 15 heartbeat subclasses are available online.[1]

B. Preprocessing

Traditionally, the preprocessing and filtering of ECG signals based on domain knowledge is used to reduce various

---

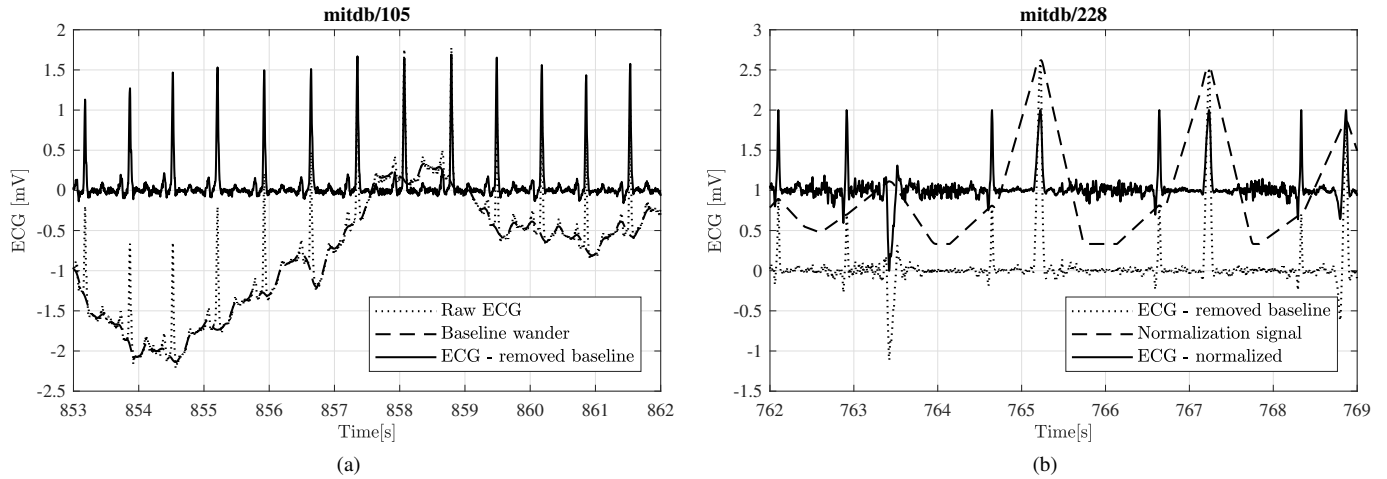[1]https://physionet.org/physiobank/annotations.shtml

Figure 1: (a) Demonstration of the baseline wander removal algorithm. (b) Signal normalization algorithm. A +1mV DC value is added to the normalized signal for clarity.

forms of noise, remove artefacts, or to limit the signal to a specifically chosen bandwidth. Such approach can give good results, but inherently leads to information loss. In this study, the preprocessing step is conducted solely to ensure the data is zero-centred and normalized for the classifier training and prediction step. The idea is to have a well extracted ECG morphology, as first proposed by [19], with as little information loss as possible.

1) Baseline wander removal: The baseline wander removal is based on a robust locally weighted linear regression algorithm first presented and explained in [20]. ECG signal is passed to the iterative algorithm (implementation is available in Matlab's Curve Fitting Toolbox) that assigns lower weight to outliers in the regression. That way, the relatively short-lasting QRS complexes, P waves, and T waves become the outliers in the regression, and the algorithm converges to the solution that well represents the sole baseline of the ECG signal. Due to the complexity of the algorithm, it is conducted locally on 4s windows of the ECG signal, with each window being subsampled to 200 points. The windowing and subsampling reduces the algorithm execution time without significant information loss, as such rigorous subsampling will remove most of the information contained in the higher frequency components of the signal that are considered to contain the outliers of the regression curve, which is also the baseline wander curve. This makes the subsampling rate justifiable, and the method relatively robust to selection of its hyperparameters (subsampling rate, window size, $rloess^2$ regression span). The algorithm performance on an example ECG signal slice from the database is shown in Figure 1a.

2) Signal normalization: The amplitude of the measured ECG signal is susceptible to various external disturbances that do not directly reflect cardiac activity. Most prominent are the effects of respiration on ECG: respiratory sinus arrhythmia (RSA) is the effect of respiration on heart rate, while R-peak amplitude modulation (RPA) represents the effect of respiration on the amplitude of the measured ECG signal. These effects are so strong that ECG has been long since observed

as a reliable source of respiratory information, and is therefore used to estimate respiratory activity [21]–[23]. Skin-electrode impedance also has a strong effect on the recording of cardiac potentials [24], [25], particularly in the sense of amplitude of the measured signal. To address the described problem, a simple algorithm that successfully normalizes the ECG signal is presented. It is a locally conducted normalization based on the local minima/maxima values of the ECG signal with no baseline wander. Equation 1 shows the simplified relation between the samples of the non-normalized ECG signal f and the normalization signal g.

$$g(i) = \max(|\max(f(j))|, |\min(f(j))|),$$
$$j \in [i - \epsilon, i + \epsilon] \tag{1}$$

The window size is the number of samples that corresponds to the time analogue of $2\epsilon = 0.5s$. Normalization signal is additionally smoothed with a moving average filter. Averaging window size is the same as in the previous step to preserve the peak values after filtering. The window size parameter ϵ is chosen to be small enough to cover the inter-beat variations in QRS amplitudes. Inherently, the value of g will be very small in parts of the ECG signal with the absence of QRS complex. Point-wise signal normalization $f_{norm} = f/g$ would actually be noise amplification in those parts. Therefore, the normalization signal is additionally clipped to be bounded from below with half the mean value of the original normalization signal. Figure 1b shows the normalization algorithm performance in a case of strong variations in both inter-beat interval times and QRS amplitudes.

C. Classification

First significant heartbeat classification algorithm using ECG morphology was presented in [19]. More recent reports of ECG morphology exploitation in the context of QRS detection [10] and heartbeat classification [9] exist as well. ECG morphology features in the most basic form are the amplitude values of the ECG signal in a close neighbourhood of a signal point. The QRS detection strategy is to determine whether

---

[2]a smoothing method implemented in Matlab's *smooth* function

a candidate point is a beat or not, based on morphological properties of the ECG signal around the observed point.

1) Model input: The strategy behind choosing the right sampling window is to cover as much of the heart cycle's morphology information (including the P wave, QRS complex, and T wave) that can be used to classify the given point as a heartbeat or non-heartbeat point. In [19], various sampling methods for extracting ECG morphology features were used and the feature selections were divided into several groups, with sampling windows varying in size from 150ms to 500ms. The authors of [9] chose to take 300 sampling points, 100 before the beat location and 200 after it, as the sample length. Considering the sampling rate, the 300 samples are equivalent to 833ms. In this work, driven by findings of a relatively recent QT interval duration study [26], we choose a sampling window size of 100ms before and 300ms after a given candidate point. Each candidate point of the ECG signal is therefore described with a 1-D array of 145 neighbouring points. Each value in the array is the amplitude value of the corresponding point of the zero-centred and normalized ECG signal.

2) Train data and labelling: The 44 recordings mentioned earlier were divided into two datasets with each dataset containing 22 recordings. This data division strategy is adopted from [19][3]. Both datasets contain the same approximate proportion of beat types as well as a mixture of routine clinical and complex arrhythmia recordings. The first dataset was used to train the classification model, while the second dataset served solely for final performance evaluation. The ECG data from train set was transformed in a way suitable for classifier training. Signals were centred and normalized as described in subsection II-B, and every point of the signal was described by a sample of 145 neighbouring points, as described in II-C1. This resulted in a total of 14296832 data samples. Each data sample is labelled positive if the candidate point is inside a ±40ms distance of an original positive beat annotation. This criterion is chosen as it is approximately accordant with what is considered a normal QRS complex duration. All other points are labelled as negative, with the dataset stratification of 11.06% positive data samples. The morphology of a single beat, together with the labelling strategy, is shown in Figure 2.

3) 1-D Convolutional neural network: Besides classical filtering and rule based QRS detection approaches (a common

---

[3]Train set comprises data from recordings 101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223 and 230. Test set comprises data from recordings 100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233 and 234
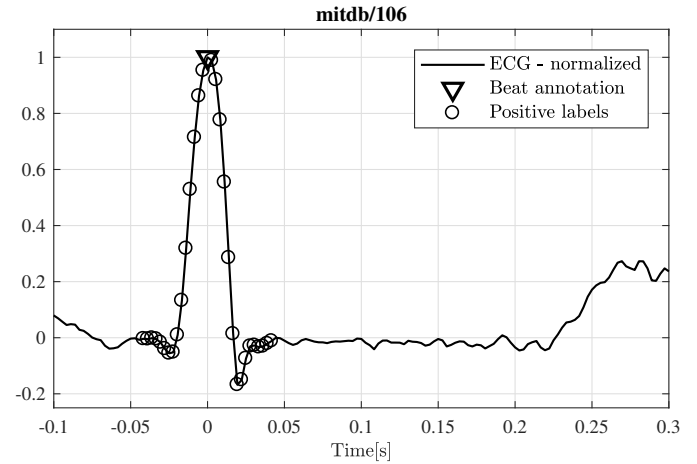

Figure 2: The neighbourhood of a beat annotation describing its morphology. All other positively labelled signal points are emphasized as well. Time axis is relative to the observed beat annotation.

one being the Pan-Tompkins algorithm [27]), various neural network based approaches have been proposed as well [28]–[30]. The proposed network architectures were mainly simple fully connected feed-forward neural networks or multi-layer perceptrons. Convolutional neural networks are composed of successive convolutional (filtering) and pooling (subsampling) layers, often followed by one or more fully connected layers. Currently, they are predominantly used in image recognition, where the network inputs are (usually raw) 2-D images. However, input of the CNN can be a one-dimensional signal, where that dimension is time. Relatively recent 1-D CNN applications are in the area of automatic sleep stage scoring from raw single-channel EEG data [31], and signal processing in speech [32]. A 1-D CNN application in the area of ECG classification uses patient-specific approach [14], for real-time monitoring and alerts during longer wearable ECG recordings. Conversely, efficient offline heart rate computation from short (e.g., 5-60 minutes) multi-subject recordings in psychophysiological experiments may benefit from subject-independent QRS detection approach, in which training has been conducted in advance on standard ECG databases.

Our CNN architecture, shown in Table I, besides a 1-D input layer, consists of two convolutional layers (C1-C2) with a max-pooling layer between them (P1), two fully-connected layers (F1-F2) and a softmax classification layer. All convolutional and fully-connected layers have a dropout probability of 0.5 to reduce overfitting on training data and

Table I: CNN architecture.

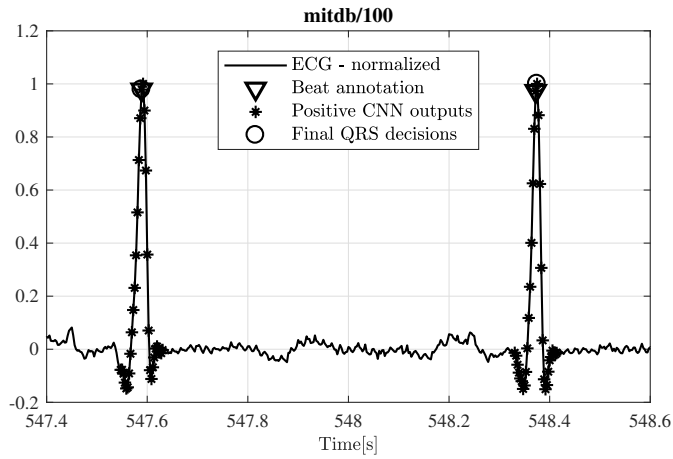| Layer	Layer Type	#Channels	Dropout	Unit type	Kernel size	Stride	Output size
Input		1					(145,1,1)
C1	convolutional	32	0.5	ReLU	(5,1)	(1,1)	(143,1,32)
P1	pooling			max	(3,1)	(2,1)	(71,1,32)
C2	convolutional	32	0.5	ReLU	(5,1)	(1,1)	(66,1,32)
F1	fully-connected		0.5	ReLU			1024
F2	fully-connected		0.5	ReLU			512
Output	softmax			logistic			2

Figure 3: Clusters of multiple positive CNN outputs forming final QRS detections.

make the trained model less sensitive to partial deformations of ECG morphology.

Stochastic gradient descent with 0.9 momentum and an initial learning rate of 0.005 was used to train the network. Mini-batch size was 128 samples, and the training was limited to 3 epochs. Input layer also included a zero-centring step (subtracting the training set mean from every sample). That option does not affect the results much as ECG normalization is conducted on a signal level in the preprocessing step (II-B).

4) Clustering: Data samples for performance testing were obtained using the same preprocessing steps as in the training step. Each feasible point of the remaining 22 test set recordings was described with the corresponding neighbourhood and fed through the trained CNN. As we expect multiple positive CNN outputs in a ±40ms distance of a QRS occurrence (ideally 30 points, taking into account the sampling frequency of 360Hz), CNN decisions have to be clustered to form final QRS detections. We perform hierarchical group-average agglomerative clustering upon all CNN decisions of a single recording, with clustering criterion being the temporal euclidean distance with a 80ms cut-off. Final QRS detection is the mean of all CNN detections within the same cluster. Clusters with less than 5 elements were elected to be too small and were therefore ignored. Since the complexity of agglomerative clustering is $\mathcal{O}(n^2 \log n)$, to reduce the execution time the clustering was performed in batches. This batch division is possible in our case as we *a priori* know the strict order of possible clusters along the one and only dimension - time. Two positively classified signal points that are separated by more than the 80ms cut-off cannot belong to the same cluster. Performance of the CNN detection and final QRS decision making by clustering is illustrated in Figure 3.

III. RESULTS

The performance of the proposed detection algorithm was evaluated with sensitivity (Se) and positive predictive value ($+P$), determined as follows:

$$Se = \frac{TP}{TP + FN} \quad (2)$$

$$+P = \frac{TP}{TP + FP} \quad (3)$$

Table II: Performance of the proposed QRS complex detection method on 22 MIT/BIH arrhythmia test recordings.

Recording	TP	FN	FP	Se	+P
100	2272	1	0	0.9996	1
103	2084	0	0	1	1
105	2550	22	22	0.9914	0.9914
111	2120	4	2	0.9981	0.9991
113	1794	1	0	0.9994	1
117	1534	1	0	0.9993	1
121	1862	1	0	0.9995	1
123	1518	0	0	1	1
200	2595	6	0	0.9977	1
202	2130	6	2	0.9972	0.9991
210	2629	21	2	0.9921	0.9992
212	2747	1	0	0.9996	1
213	3248	3	0	0.9991	1
214	2258	4	0	0.9982	1
219	2154	0	0	1	1
221	2424	3	0	0.9988	1
222	2475	8	0	0.9968	1
228	2048	5	4	0.9976	0.9981
231	1571	0	0	1	1
232	1780	0	1	1	0.9994
233	3075	4	0	0.9987	1
234	2752	1	0	0.9996	1
Total	49620	92	33	0.9981	0.9993

To determine whether a detection is a true positive (TP), a ±75ms tolerance window is used, which is accordant with [18]. If the detection is in the ±75ms range from a beat annotation, it is considered a TP, otherwise it is considered false positive (FP). Similarly, no detection of QRS complex in the same range around an annotation is considered false negative (FN). Results of the proposed QRS detection algorithm on the test set are shown in Table II. Processing time of the algorithm takes 1.02 seconds per 1 minute of ECG signal length, making the method suitable for real-time implementation as well[4].

The reported results are comparable with the ones obtained with a method based on exploration of processing steps combinations of several QRS detection algorithms [13], whose authors found that the combination of Sun Yan's MMF or MMD methods [33] with Elgendi's algorithm [8] works best for QRS detection (Se=99.77%, +P=99.72% on MIT/BIH arrhythmia database).

IV. CONCLUSION

This paper has presented a contemporary convolutional neural network approach to QRS detection, which was demonstrated to work very well, while avoiding hand crafted features or decision rules. A new adaptation of an existing mathematical method has been applied for baseline wander removal in ECG, and a simple signal normalization has been presented as well.

---

[4]the processing time is reported as average on all test set signals, executed in Matlab with Intel Xeon Processor E3-1275 @ 3.60GHz, Nvidia Quadro M2000

The detection performance is evaluated on MIT/BIH arrhythmia database, making the detection algorithm's performance comparable with existing methods.

Our method achieved sensitivity of 99.81% and a 99.93% positive predictive value on unseen data, which is comparable to state-of-the-art solutions, and the best subject-independent method based on neural networks, to the best of our knowledge. As future work, the presented methods hyper-parameters can be optimized through cross-validation over more databases, additionally improving and increasing the significance of these preliminary performance measures. This approach also has potential to be used in the task of heartbeat classification and can even be considered for P and T wave detection, opening a number of additional future work possibilities.

REFERENCES

[1] P. Melillo, M. Bracale, and L. Pecchia, "Nonlinear Heart Rate Variability features for real-life stress detection. Case study: students under stress due to university examination," *Biomedical engineering online*, vol. 10, no. 1, p. 96, 2011.

[2] M. Elgendi, A. Mohamed, and R. Ward, "Efficient ECG Compression and QRS Detection for E-Health Applications," *Scientific Reports (Nature Publisher Group)*, vol. 7, p. 1, 2017.

[3] I. C. Christie and P. J. Gianaros, "Physioscripts: An extensible, open source platform for the processing of physiological data," *Behavior research methods*, vol. 45, no. 1, pp. 125–131, 2013.

[4] M. Nardelli, G. Valenza, A. Greco, A. Lanata, and E. P. Scilingo, "Recognizing emotions induced by affective sounds through heart rate variability," *IEEE Transactions on Affective Computing*, vol. 6, no. 4, pp. 385–394, 2015.

[5] G. M. Friesen, T. C. Jannett, M. A. Jadallah, S. L. Yates, S. R. Quint, and H. T. Nagle, "A comparison of the noise sensitivity of nine QRS detection algorithms," *IEEE Transactions on biomedical engineering*, vol. 37, no. 1, pp. 85–98, 1990.

[6] M. Elgendi, B. Eskofier, S. Dokos, and D. Abbott, "Revisiting QRS detection methodologies for portable, wearable, battery-operated, and wireless ECG systems," *PloS one*, vol. 9, no. 1, p. e84018, 2014.

[7] J. Kim and H. Shin, "Simple and robust realtime QRS detection algorithm based on spatiotemporal characteristic of the QRS complex," *PloS one*, vol. 11, no. 3, p. e0150144, 2016.

[8] M. Elgendi, "Fast QRS detection with an optimized knowledge-based method: Evaluation on 11 standard ECG databases," *PloS one*, vol. 8, no. 9, p. e73557, 2013.

[9] S. Chen, W. Hua, Z. Li, J. Li, and X. Gao, "Heartbeat classification using projected and dynamic features of ECG signal," *Biomedical Signal Processing and Control*, vol. 31, pp. 165–173, 2017.

[10] S. Yazdani and J.-M. Vesin, "Extraction of QRS fiducial points from the ECG using adaptive mathematical morphology," *Digital Signal Processing*, vol. 56, pp. 100–109, 2016.

[11] R. Benali, F. B. Reguig, and Z. H. Slimane, "Automatic classification of heartbeats using wavelet neural network," *Journal of medical systems*, vol. 36, no. 2, pp. 883–892, 2012.

[12] B. Abibullaev and H. D. Seo, "A new QRS detection method using wavelets and artificial neural networks," *Journal of medical systems*, vol. 35, no. 4, pp. 683–691, 2011.

[13] K. Friganovic, A. Jovic, D. Kukolja, M. Cifrek, and G. Krstacic, "Optimizing the detection of characteristic waves in ECG based on exploration of processing steps combinations," in *EMBEC & NBC 2017*. Springer, 2017, pp. 928–931.

[14] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ECG classification by 1-D convolutional neural networks," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp. 664–675, 2016.

[15] R. Mark and G. Moody, "MIT-BIH arrhythmia database 1997," *URL http://ecgmit. edu/dbinfo. html*, 1997.

[16] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[17] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, 2001.

[18] A. ECAR, "Recommended practice for testing and reporting performance results of ventricular arrhythmia detection algorithms," *Association for the Advancement of Medical Instrumentation*, 1987.

[19] P. De Chazal, M. O'Dwyer, and R. B. Reilly, "Automatic classification of heartbeats using ECG morphology and heartbeat interval features," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 7, pp. 1196–1206, 2004.

[20] W. S. Cleveland, "LOWESS: A program for smoothing scatterplots by robust locally weighted regression," *The American Statistician*, vol. 35, no. 1, pp. 54–54, 1981.

[21] G. B. Moody, R. G. Mark, A. Zoccola, and S. Mantero, "Derivation of respiratory signals from multi-lead ECGs," *Computers in cardiology*, vol. 12, no. 1985, pp. 113–116, 1985.

[22] D. Cysarz, R. Zerm, H. Bettermann, M. Frühwirth, M. Moser, and M. Kröz, "Comparison of respiratory rates derived from heart rate variability, ECG amplitude, and nasal/oral airflow," *Annals of biomedical engineering*, vol. 36, no. 12, pp. 2085–2094, 2008.

[23] C. Orphanidou, S. Fleming, S. A. Shah, and L. Tarassenko, "Data fusion for estimating respiratory rate from a single-lead ECG," *Biomedical Signal Processing and Control*, vol. 8, no. 1, pp. 98–105, 2013.

[24] M. S. Spach, R. C. Barr, J. W. Havstad, and E. C. Long, "Skin-electrode impedance and its effect on recording cardiac potentials," *Circulation*, vol. 34, no. 4, pp. 649–656, 1966.

[25] A. S. Berson and H. V. Pipberger, "Skin-electrode impedance problems in electrocardiography," *American heart journal*, vol. 76, no. 4, pp. 514–525, 1968.

[26] I. Goldenberg, A. J. Moss, W. Zareba *et al.*, "QT interval: how to measure it and what is "normal"," *Journal of cardiovascular electrophysiology*, vol. 17, no. 3, pp. 333–336, 2006.

[27] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE transactions on biomedical engineering*, no. 3, pp. 230–236, 1985.

[28] C. García-Berdonés, J. Narváez, U. Fernández, and F. Sandoval, "A new QRS detector based on neural network," *Biological and Artificial Computation: From Neuroscience to Technology*, pp. 1260–1269, 1997.

[29] Q. Xue, Y. H. Hu, and W. J. Tompkins, "Neural-network-based adaptive matched filtering for QRS detection," *IEEE Transactions on Biomedical Engineering*, vol. 39, no. 4, pp. 317–329, 1992.

[30] M. A. Hasan, M. Ibrahimy, and M. Reaz, "NN-based R-peak detection in QRS complex of ECG signal," in *4th Kuala Lumpur International Conference on Biomedical Engineering 2008*. Springer, 2008, pp. 217–220.

[31] O. Tsinalis, P. M. Matthews, Y. Guo, and S. Zafeiriou, "Automatic Sleep Stage Scoring with Single-Channel EEG Using Convolutional Neural Networks," *arXiv preprint arXiv:1610.01683*, 2016.

[32] D. Palaz, R. Collobert, and M. M. Doss, "Estimating phoneme class conditional probabilities from raw speech signal using convolutional neural networks," *arXiv preprint arXiv:1304.1018*, 2013.

[33] Y. Sun, K. L. Chan, and S. M. Krishnan, "Characteristic wave detection in ECG signal using morphological transform," *BMC cardiovascular disorders*, vol. 5, no. 1, p. 28, 2005.