Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 120 (2017) 268–275

Procedia Computer Science

www.elsevier.com/locate/procedia

9th International Conference on Theory and Application of Soft Computing, Computing with Words and Perception, ICSCCW 2017, 24-25 August 2017, Budapest, Hungary

Cardiac arrhythmia detection using deep learning

Ali Isin[a,*] Selen Ozdalili[b]

[a,*] Department of Biomedical Engineering, Near East University, P.O.BOX:99138, Nicosia, North Cyprus, Mersin 10 Turkey
[b] Department of Innovation and Knowledge Management, Near East University, P.O.BOX:99138, Nicosia, North Cyprus, Mersin 10 Turkey

Abstract

An electrocardiogram (ECG) is an important diagnostic tool for the assessment of cardiac arrhythmias in clinical routine. In this study, a deep learning framework previously trained on a general image data set is transferred to carry out automatic ECG arrhythmia diagnostics by classifying patient ECG's into corresponding cardiac conditions. Transferred deep convolutional neural network (namely AlexNet) is used as a feature extractor and the extracted features are fed into a simple back propagation neural network to carry out the final classification. Three different conditions of ECG waveform are selected from MIT-BIH arrhythmia database to evaluate the proposed framework. Main focus of this study is to implement a simple, reliable and easily applicable deep learning technique for the classification of the selected three different cardiac conditions. Obtained results demonstrated that the transferred deep learning feature extractor cascaded with a conventional back propagation neural network were able to obtain very high performance rates. Highest obtained correct recognition rate is 98.51% while obtaining testing accuracy around 92%. Based on these results, transferred deep learning proved to be an efficient automatic cardiac arrhythmia detection method while eliminating the burden of training a deep convolutional neural network from scratch providing an easily applicable technique.

Peer-review under responsibility of the scientific committee of the 9th International Conference on Theory and application of Soft Computing, Computing with Words and Perception

Keywords: biomedical signal processing; deep learning; ECG classification; convolutional neural networks; cardiac arrhythmia; tranfer learning

* Ali Işın. Tel.: +90 0392 22 84 267; fax: +90 0392 22 88 407.
E-mail address: ali.isin@neu.edu.tr *

1877-0509 © 2018 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the 9th International Conference on Theory and application of Soft Computing, Computing with Words and Perception.
10.1016/j.procs.2017.11.238

1. Introduction

The Electrocardiogram (ECG) is an established technique in cardiology for the analysis of cardiac condition of the patients. In its basic definition, ECG is the electrical representation of the contractile activity of the heart, and can be recorded fairly easily by using surface electrodes on the limbs or chest of the patient. The ECG is one of the most recognized and used biomedical signal in the field of medicine. The rhythm of the heart in terms of beats per minute (bpm) can be easily calculated by counting the R peaks of the ECG wave during one minute of recording (see Fig.1 for a single ECG waveform). More importantly, rhythm and the morphology of the ECG waveform is altered by cardiovascular diseases and abnormalities such as the cardiac arrhythmias (Rangayyan, 1999), which their automatic detection and classification is the main focus of this paper.

In current medical routine, careful study of the ECG by expert cardiologists is necessary for the diagnosis of life threatening cardiac arrhythmias. However, automatic classification of cardiac arrhythmias can both provide objective diagnostic results and save time for the cardiologists. These advantages have provided considerable commercial interests in the computer aided classification and diagnosis of the ECG signals in hospital and health community. The interpretation of the ECG signal is an application of pattern recognition. The purpose of pattern recognition is to automatically categorize a system into one of a number of different classes. An expert cardiologist can easily diagnose various cardiac arrhythmias just by looking at the ECG waveforms printout. In some specific cases, sophisticated ECG analysers can achieve a higher degree of accuracy than that of cardiologist, but at present there remains a group of ECG waveforms that are difficult to identify by computers. However, the use of computerized analysis of easily obtainable ECG waveforms can considerably reduce the cardiologist's workload. Some analysers can assist the cardiologist by producing a ready diagnosis while others can provide a limited number of parameters by which the cardiologist can make his own diagnosis.

The aim of this paper is to develop such a computer aided diagnostic system which assists expert cardiologists by providing intelligent, cost effective and time saving ECG arrhythmia diagnostics. To achieve this goal, conventional ECG signal processing techniques along with the state of the art deep learning methods are implemented to the task of ECG arrhythmia pattern recognition. At its current state, the proposed system can specifically distinguish and classify cardiac arrhythmias known as Right Bundle Branch Blocks (RBBB) from Paced Beats and Normal (Healthy) Beats. Where, Normal beats are healthy adult human ECG waveform; Paced beats are artificial beats from the device called pacemaker; and RBBB is an arrhythmia that is frequently associated with ischemic, hypertensive, rheumatic and pulmonary heart disease, right ventricular hypertrophy and some drug intoxication which has a ECG waveform with QRS duration between 0.10 and 0.11 sec (incomplete RBBB) or 0.12sec or more (complete RBBB), prolonged ventricular activation time or QR interval (0.03sec or more) and right axis deviation (Milliken, 1971) (see Fig 2 for sample waveforms). However, apart from these the system can also be easily adapted to further classify other various similar cardiac arrhythmias.

The rest of the paper is organized as follows: First, recent related deep learning based methods for ECG arrhythmia detection are briefly reviewed in section 2. Then, in section 3, proposed method is explained in detail and in section 4, experiments and results are discussed. Finally, in conclusions, the current state of the proposed method is assessed and future directions for development are provided.

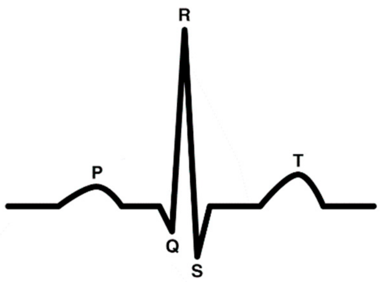

Fig. 1. Normal/healthy ECG waveform. Where, P-wave, QRS-complex and T-wave represent the contraction/depolarization of atria, contraction/depolarization of ventricles and repolarization of ventricles respectively.

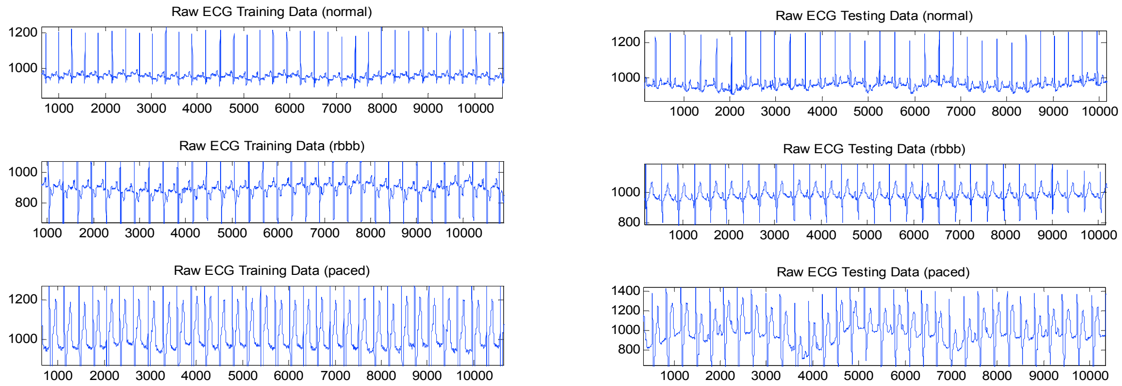

Fig. 2. An example of three different ECG recordings used in this work. From top to bottom: Normal beats, Rbbb beats and Paced beats. Three recordings on the left section represent part of the training data and three recordings on the right section represent part of the testing data. (X-axis represents samples while y-axis is amplitude in mV)

2. Related Work

Although most of the conventional pattern recognition techniques have been previously applied successfully to the ECG arrhythmia detection tasks, recent state-of-the-art performances obtained by deep learning methods, particularly Convolutional Neural Networks (CNNs), in popular pattern/object recognition challenges (Krizhevsky et al., 2012), encouraged researchers to implement these techniques to the field of medical image and signal processing. Application of deep learning methods even to the most complex medical pattern recognition tasks proved to be very promising, obtaining state-of-the-art results (Işın et al., 2016). In general, extracting highly representative features from the data in hand has the most significant impact on the performance of computerized classification/recognition systems. However this is a time-consuming process which requires expert knowledge and often selected features fail to be robust with respect to the variations in the data. When compared to conventional classification methods, CNNs automatically learn representative complex features directly from the data itself, thus the need for handcrafting features to represent the data effectively is eliminated.

One of the recent studies proposed a deep neural network to recognize premature ventricular contraction (PVC) beats from ECG recordings (Jun et al., 2016). A deep neural network with six hidden layers is trained by feeding six different features extracted from ECGs to carry out the classification between normal and PVC beats. Although, a deep neural network is used authors still preferred handcrafting their own features from the ECG data.

In Pourbabaee et al. (2016), a deep convolutional neural network is trained to extract features directly from raw ECG signals and to carry out the classification between two cardiac conditions, namely; normal beats and paroxysmal atrial fibrillation (PAF). One of the drawbacks of training a deep convolutional neural network from scratch is the need for a quite large labeled training dataset size to obtain improved network performance by making the network deeper (i.e. increasing the number of convolutional layers). Even if a very large dataset is available, increasing the depth of the network will also increase the computation cost during training due to the increase in complex convolutional operations throughout deep convolutional layers. Thus, powerful GPU powered computers are required for such training tasks.

Transfer learning and its variation, transfer learning with fine tuning (Tajbakhsh et al., 2016), provide a useful solution for these drawbacks enabling a deep learning application where limited training data, not enough expertise for training and moderate computer resources are available. In transfer learning, a pre-trained deep CNN is imported to the desired task in hand and used as an automatic feature extractor. For example, a CNN pre-trained on a general image dataset can be imported for a medical imaging task, like pathology classification, and be used for automatically extracting features that can be used as an input to a classifier, for carrying out the final classification. Additionally, in fine tuning one or more layers of the pre-trained network can be re-trained (i.e fine-tuned) using the data for the task in hand.

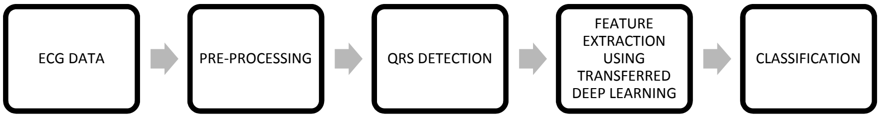

Fig. 3. Block diagram of the proposed ECG pattern recognition and classification system.

3. Method

In this work, four major steps are applied to solve the ECG pattern recognition and classification problem (Fig. 3). These steps include, signal pre-processing, QRS detection, ECG feature extraction using transferred deep learning and ECG signal classification using a conventional Artificial Neural Network (ANN).

3.1. ECG Data Acquisition

For this study, the ECG data is obtained from MIT-BIH Arrhythmia database (MIT-BIH ECG database, 2017). MIT-BIH Arrhythmia database is an online database formed using a set of over 4000 long-term ECG Holter recordings. Nearly 60% of these recordings were obtained from patients. The whole database contains 23 records (numbered from 100 to 124, some numbers missing) chosen at random from this set, and 25 records (numbered from 200 to 234, again some numbers missing) selected from the same set to include a variety of rare but clinically important phenomena. Each of these 48 records is slightly over 30 minutes long. All the waveforms present in these recordings are annotated by expert cardiologists. For the proposed system, ECG records 100, 118 and 217 (normal, rbbb and paced beats respectively) are used in the training and records 101, 231 and 107 (normal, rbbb and paced beats respectively) are used in testing process (see Fig.2).

Each of these records has a sample frequency of 360Hz and contains two leads (two signals recorded from different angles on chest). Only one minute long sections of each record is extracted obtaining a total of six recordings (3 for training and 3 for testing), each containing 21600 samples and approximately 60-80 waveforms depending on heart rate and class (normal, rbbb or paced). As a final step, one of the channels is removed and only one channel for each recording is used for the rest of the system (channel MLII). In the end of the Data acquisition part a total of 214 waveforms (as 3 separate recordings each representing a normal, rbbb and paced waveform class) for training and 202 waveforms for testing are prepared ready for next step which is signal pre-processing.

3.2. Signal Pre-processing

Although ECG data obtained from MIT-BIH database is not expected to contain as much disruptive noise as an ECG data obtained directly from a patient, it still contains some noise which requires attention to improve the subsequent steps of the system. Therefore, signal pre-processing step is focused on removing noise from ECG recordings.

As a first step, mean removal is applied to remove the dc noise present in the ECG signals. By subtracting the mean of the ECG recording from every sample point, the unwanted dc component is removed and the signal baseline amplitude is pulled back to level zero.

Almost all of the ECG recordings also contain high and low frequency noise that are imposed by several factors like, the contraction of the muscles, respiratory movements, poor electrode contact and presence of other external devices. To remove high frequency noise, mostly caused by patients muscle contractions during recording, a 10- point moving average filter which passes low frequencies but attenuates high frequencies is chosen and the signals are filtered. After the removal of high frequency noise, the next step is to remove low frequency noise components. This low frequency noise shows itself as baseline wandering that is caused mostly by the respiration of the patient. To remove this low frequency noise, a derivative based (high pass) filter, which passes high frequencies but attenuates low frequencies, is applied to all ECG recordings.

Power line interference is another type of noise present in the ECG data and is caused by the electricity current flowing in wires and power lines. Power line interference that is present in the ECG signals consists of 60Hz pickup and harmonics. Since frequency of 60Hz overlaps with ECG frequency range (which is generally between 0.05-150 Hz), a filter should be applied to suppress only 60Hz frequency components and its harmonics without disturbing the frequencies around. To achieve this, a comb filter, which is a band-stop filter which attenuates a certain band of frequencies and their harmonics, is used and 60Hz power line interference with its harmonics is removed from the ECG signals.

All of the above steps are applied to all training and testing ECG records and filtered ECG signals are obtained ready for the next QRS detection step.

3.3. QRS detection

QRS complex is the most striking waveform within the ECG. Since it reflects the electrical activity within the heart during the ventricular contraction, the time of its occurrence as well as its shape provide valuable information about the current state of the heart. In that sense, QRS detection provides the fundamentals for almost all automated ECG analysis algorithms. Additionally, previous research (Ozbay and Karlik, 2001) proved that in automated ECG classification, focusing on the morphology of the R-T interval (of ECG) is very representative for certain important cardiac conditions, which is also true for rbbb and paced beats. In order to extract R-T interval from the ECG, R-peaks should be detected effectively using QRS detection.

A well-known Pan-Tompkins algorithm (Pan and Tompkins, 1985) is applied to carry out the QRS detection. The algorithm includes a series of methods that perform derivative, squaring, integration, adaptive thresholding and search procedures for the detection of R-peaks of the ECG signal.

3.4. ECG feature extraction using transferred deep learning

As the ECG data in hand is investigated, it can be observed that the features that clearly distinguish each class (normal, rbbb and paced) lies between the R-T interval. Also it can be easily observed that each member of a class shows same form of pattern in this interval. In this regard, 200 sample points after each R-peak (approximately this amount of samples corresponds to the R-T interval with sampling frequency of 360Hz) are extracted (see Fig. 4) and fed as an input to the transferred deep learning based feature extractor. All other parts of the ECG waveforms are then removed.

AlexNet (Krizhevsky et al., 2012) is a deep learning system, based on a convolutional neural network, which is trained on the 1.2 million high-resolution RGB images of the ImageNet dataset (Deng et al., 2009). This system can classify those images into 1000 different categories with a state-of-the-art performance.

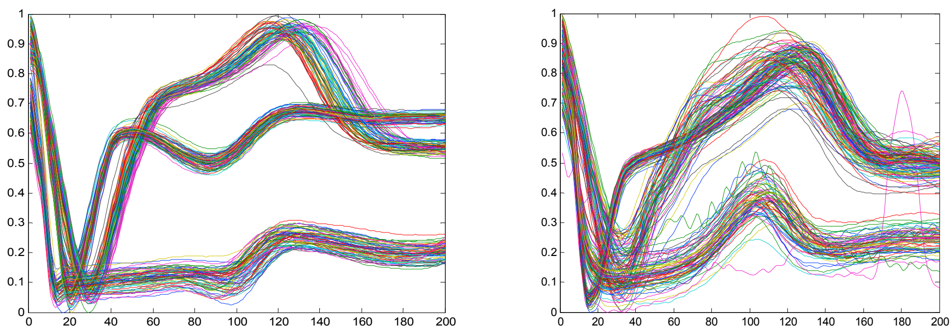

Fig. 4. Left: graphical representation of R-T intervals of all training ECG data, Right: graphical representation of R-T intervals of all testing ECG data. On the left figure, top cluster represents paced beats, middle cluster represents rbbb beats and bottom cluster represents normal beats. (X-axis represents samples while y-axis is normalized amplitude)

The architecture of the AlexNet CNN contains total of eight layers, five convolutional and three fully-connected layers, which are trained on the generic images of the ImageNet. As stated in previous research (Yosinski et al. 2014), deeper layers of a deep learning framework trained on a very large annotated dataset, like ImageNet, can be generic enough to be transferred and impelemented for another image classification task. In this regard, for the proposed ECG classification system, the outputs of the deeper layers of the pre-trained AlexNet can be extracted as features for the given ECG inputs. One of the main challenges here is that the AlexNet accepts 256x256 sized RGB images as input where the remaining ECG waveforms after the previously explained steps are discrete signals with 200 sample points. To solve this, all training and testing data (as shown in Fig. 4.), are converted into 256x256 sized binary images where background and the signal is represented by white and black pixels respectively. Each image representing a single ECG waveform is then reproduced three times to imitate an RGB image.

After the ECG signals are converted into 256x256x3 sized images (total of 214 image for training and 202 image for testing) representing R-T intervals, they are fed into the pre-trained AlexNet and the outputs of the 6th and 7th fully connected layers of the AlexNet are extracted as features representing our R-T segment images for the three different cardiac conditions. Since the fully connected layers of the AlexNet contains 4096 neurons each, this procedure will result in 4096 features for each input ECG waveform that will enable accurate ECG arrhythmia classification. All these obtained features are then converted into feature vectors with 4096 elements (F_v=f_1,f_2,f_3,…,f_{4096}) ready for the next classification step.

3.5. ECG classification

Simple conventional multi-layer feed-forward neural networks that have 1 input 1 hidden and 1 output layers are designed to carry out the classification between three different cardiac conditions. These networks are trained using scaled conjugate gradient back-propagation algorithm.

Since the number of features extracted from the 6th and 7th fully connected layers of the AlexNet contains very large number of features (4096 features each) for each ECG R-T segment, we applied principal component analysis (PCA) to reduce the feature vector dimension, decreasing the computational load and risk of over fitting.

After the PCA based feature vector dimension reduction is applied, three different networks are designed and trained. The first network (N-Fc6) takes the output of the 6th fully connected layer of AlexNet as inputs, the second one (N-Fc7) takes the output of the 7th layer, and the third network (N-tst) takes 200 samples of the R-T intervals directly as inputs without any deep learning feature extraction. Additionally, each of these networks are trained using different number of hidden layer neurons and the best performing hidden layer neuron numbers are fixed for the final tests. Output layer neuron number is also fixed as three neurons.

As the final step, output vectors are formed as shown in Table 1 and the training of the networks are carried out using the default settings of the Matlab`s neural network pattern recognition application. Accompanying each record in the MIT-BIH database there is an annotations file in which each heartbeat has been identified by expert cardiologist annotators. This annotated information is employed for designing the corresponding target vector for each R-T segment and also used for evaluating the classifier performance.

Feature vectors along with their corresponding 214x3 output target vector is fed into the designed networks for training with back propagation algorithm. Training is continued until error goal is achieved or maximum epoch is reached. After the training is finished, networks outputs are compared with output target vector and correct training recognition rates and accuracies are recorded. Correct recognition is counted when the same output neuron shows

Table 1. Output target vector representations. 214x3 and 202x3 output vectors are formed for training and testing respectively.

Cardiac condition	Corresponding Output Vector
Normal (n)	[1 0 0]
Right Bundle Branch Block (rbbb)	[0 1 0]
Paced Beats (p)	[0 0 1]

the maximum value both for actual output and desired output. Accuracy is found by subtracting networks actual output of the neuron that should show the maximum value (that should be classified) from the desired output which is always 1. During training, %15 of the training data and %15 of the testing data is combined and used for validation purposes.

After the training of all three networks are completed, networks are tested using the testing dataset and its corresponding 202x3 output vector, and the performances are recorded for evaluation.

All the mentioned procedures are coded and applied in Matlab environment using the built-in signal & image processing toolbox, neural network toolbox, deep learning toolbox and AlexNet toolbox on a PC with Intel i5-6500 3.20 GHz CPU, NVidia Geforce GTX 1060 GPU and 16 GBs of RAM.

4. Results and Discussion

ECG Data obtained from MIT-BIH database are pre-processed, QRS complexes are detected and features in R-T intervals are extracted using three different methods. After all these steps, three different networks that use these different features as inputs are designed, trained, tested and evaluated for pattern recognition and classification of ECG signals for three different cardiac conditions. Table 2 summarizes the evaluation results obtained for the three different networks.

When all of the tested networks are evaluated it is found that networks based on transferred deep learning feature extraction (N-Fc6 and N-Fc7) obtained almost %100 recognition rates and accuracies above %96 in training phase. For this reason it is not found necessary to change any parameters like maximum epochs, error limits or learning rates. When compared, the network that is not based on deep learning (N-tst) obtained %92 recognition rate and accuracy of around %90 in training phase.

During testing phase, N-Fc7 obtained 98.51% testing recognition rate and an average accuracy around 92.4% in its best run. N-Fc6's numbers were 97.53% recognition rate and 91.2% accuracy. N-tst obtained 91.58% recognition rate but only 85% accuracy.

As the obtained results clearly indicate, ECG arrhythmia detection methods based on transferred deep learning feature extraction outperformed the non-deep learning approach by a good margin. Additionally, the results obtained by the networks that use features extracted from deep full convolutional layers (6th and 7th) of the AlexNet proved that the deeper layers of a deep convolutional neural network trained on a very large annotated data set can be generic enough to be transferred and implemented for ECG arrhytmia classification task.

Table 2. Comparison of the performances of the three different networks designed for ECG arrhythmia detection in their best runs

Network	Correct recognition				Accuracies			
	Training		Testing		Training		Testing	
	Correct recognition/total patterns	Percentage	Correct recognition/total patterns	Percentage	Value	Percentage	Value	Percentage
N-Fc7	213/214	99.53	199/202	98.51	0.9735	97.4	0.9244	92.4
N-Fc6	210/214	98.13	195/202	97.53	0.9608	96.1	0.9121	91.2
N-tst	197/214	92.06	185/202	91.58	0.9015	90.1	0.8501	85

5. Conclusions

In clinical routine, computer aided diagnosis of cardiac arrhythmias can reduce the workload of cardiologists dramatically enabling them to focus more on treatment rather than diagnostics. In this paper, an efficient transferred deep learning based ECG classification system is realized to carry out automatic ECG arrhythmia diagnostics by classifying patient ECG’s into corresponding three different cardiac conditions; normal, paced or right bundle branch block.

After the ECG records are acquired from the online MIT-BIH arrhythmia database, they are filtered from noises and QRS waves are detected to extract R-T segments of the ECG. Pre-trained AlexNet, is transferred and used as a feature extractor for the ECG classification task in hand. The extracted features are fed into a simple back propagation neural network to classify the input ECG R-T segments into one of the three cardiac conditions.

Obtained very high correct recognition rate (98.51%) and testing accuracy (around 92%) of the proposed system demonstrated that the transferred deep learning feature extractor cascaded with a simple back propagation neural network is an efficient automatic cardiac arrhythmia detection method. Additionally, transferring a pre-trained deep CNN eliminates the need for high expertise and computational power required for training a deep convolutional neural network from scratch.

With the recent state-of-the-art performances of deep learning based medical image and signal processing methods, biomedical scientists are coming one step closer to the effective realization of a marketable computer aided diagnostic system that will be well established in clinical routine to assist clinicians and patients alike. In this regard, in the very near future, it won’t be too surprising to see state-of-the-art performances from deep learning applications not only in medical signals and imaging diagnostics but also in other popular sub-fields of biomedical imaging and signals (Wang, 2016) (Işın et al., 2017).

References

Deng, J. et al., 2009. ImageNet: A Large-Scale Hierarchical Image Database. CVPR 09.

Işın, A., Direkoğlu, C., Şah, M., 2016. Review of MRI-based Brain Tumor Image Segmentation Using Deep Learning Methods. Procedia Computer Science 102, 317-324.

Işın, A. et al., 2017. Monte Carlo simulation of PET/MR scanner and assessment of motion correction strategies, JINST 12, C03089.

Jun, T. J. et al., 2016. Premature ventricular contraction beat detection with deep neural networks. 15th IEEE International Conference on Machine Learning and Applications, 859-864.

Krizhevsky, A., Sutskever, I., Hinton, G. E., 2012. Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 1097–1105.

Milliken, J. A. et al., 1971.Validity of computer interpretation of the electrocardiogram. C.M.A. Journal 105, 1147-1150.

MIT-BIH ECG database, 2017. http://www.physionet.org/physiobank/database/html/mitdbdir/mitdbdir.htm.

Ozbay, Y., Karlik, B., 2001. Recognition of ECG Arrhythmias Using Artificial Neural Networks. Proceedings 23rd Annual Conference IEEE/EMBS.

Pan, J., Tompkins, W. J., 1985. A Real Time QRS Detection Algorithm. IEEE Transactions on Biomedical Engineering 32, 230-236.

Pourbabaee, B., Roshtkhari, M. J., Khorasani, K., 2016. Feature Learning with Deep Convolutional Neural Networks for Screening Patients with Paraoxysmal Atrial Fibrillation. IEEE International Joint Conference on Neural Networks, 5057-5064.

Rangayyan, R. M., 1999. Biomedical Signal Analysis: A Case-Study Approach. IEEE Press.

Tajbakhsh, N. et al., 2016. Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?. IEEE Trans. Med. Imaging 35, 1299-1312.

Wang, G., 2016. A Perspective on Deep Learning. IEEE Access 4, 8914-8924.

Yosinski, J. et al., 2014. How transferable are features in deep neural networks? Advances in Neural Information Processing Systems, 3320–3328.