

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/285493884>

Convolutional Neural Networks for Patient-Specific ECG Classification

Conference Paper · August 2015

DOI: 10.1109/EMBC.2015.7318926

CITATIONS

68

READS

2,691

4 authors:



[Serkan Kiranyaz](#)

Qatar University

247 PUBLICATIONS 4,553 CITATIONS

[SEE PROFILE](#)



[Turker Ince](#)

Izmir University of Economics

95 PUBLICATIONS 2,943 CITATIONS

[SEE PROFILE](#)



[Hamila Ridha](#)

165 PUBLICATIONS 972 CITATIONS

[SEE PROFILE](#)



[Moncef Gabbouj](#)

Tampere University

867 PUBLICATIONS 12,764 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



video stabilization [View project](#)



Fast and energy-efficient video coding [View project](#)

Convolutional Neural Networks for Patient-Specific ECG Classification

Serkan Kiranyaz, Turker Ince, Ridha Hamila and Moncef Gabbouj

Abstract— We propose a fast and accurate patient-specific electrocardiogram (ECG) classification and monitoring system using an adaptive implementation of 1D Convolutional Neural Networks (CNNs) that can fuse feature extraction and classification into a unified learner. In this way, a dedicated CNN will be trained for each patient by using relatively small common and patient-specific training data and thus it can also be used to classify long ECG records such as Holter registers in a fast and accurate manner. Alternatively, such a solution can conveniently be used for real-time ECG monitoring and early alert system on a light-weight wearable device. The experimental results demonstrate that the proposed system achieves a superior classification performance for the detection of ventricular ectopic beats (VEB) and supraventricular ectopic beats (SVEB).

I. INTRODUCTION

A wide range of methods have been proposed in the past for generic and fully automatic ECG classification based on signal processing techniques such as frequency analysis [1], wavelet transform [2], [3], and filter banks [4], statistical [5] and heuristic approaches [6], hidden Markov models [7], support vector machines [8], artificial neural networks (ANNs) [9], and mixture-of-experts method [10]. Generally speaking they have not performed well in practice due to the inter-patient variations of the ECG signals and thus they usually exhibit a common drawback of having an inconsistent performance when, for instance, classifying a new patient's ECG signal. This makes them unreliable to be widely used clinically or in practice, and they tend to have high variations in their accuracy and efficiency for larger databases, [11], [15]. To address this major drawback, few methods have recently been proposed with a patient-specific design, [10], [15], [18], [19] and [20] and they have in particular demonstrated significant performance improvements over generic ECG classification methods. They have a common approach with two major operations: feature extraction and classification over the extracted features. However, using such fixed and hand-crafted features may not

represent the characteristics of the underlying signal in an optimal way and obviously this is against the philosophy of a “patient-specific” approach since the *same* set of features will be used for any patient regardless whether or not that feature extraction method is capable of characterizing the ECG signal. Therefore, a true “patient-specific” solution indeed requires the design of the best possible features for each individual ECG data. Moreover, extracting several features, especially in the transform domains along with the post-processing methods such as PCA may significantly increase the computational complexity of the overall process and this may hinder them from the usage in lightweight applications such as wearable heart monitoring devices.

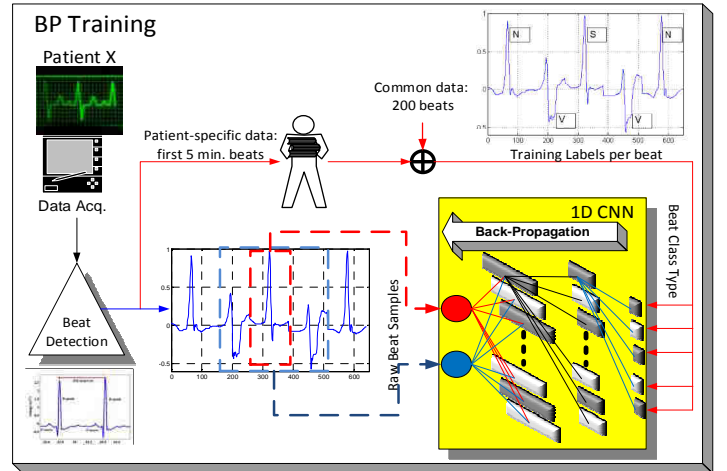


Figure 1: Overview of the proposed approach in BP training.

To address such deficiencies and to design a true patient-specific method with a superior classification performance, in this paper we propose a novel ECG classification approach based on adaptive 1D Convolutional Neural Networks (CNNs). CNNs are hierarchical neural networks whose convolutional layers alternate with subsampling layers, reminiscent of simple and complex cells in the human visual cortex that are now commonly used for the “deep learning” tasks such as object recognition in large image archives whilst achieving the state-of-the-art performances [12]–[14]. To our knowledge this is the first work where they are used over 1D signals, in particular for the purpose of ECG classification and anomaly detection. With certain modifications and adaptations over the traditional 2D CNNs, the proposed system can be tuned to classify directly the raw data of the heart beats in any sampling rate, therefore, voiding the need for any manual feature extraction and pre- or

S. Kiranyaz is with Electrical Engineering, College of Engineering, Qatar University, Qatar; e-mail: Serkan.Kiranyaz@gmail.com.

T. Ince is with the Electrical & Electronics Engineering Department, Izmir University of Economics, Turkey; e-mail: turker.ince@izmirekonomi.edu.tr.

R. Hamila is with Department of Electrical Engineering, Qatar University; Doha, Qatar, email: hamila@qu.edu.qa

M. Gabbouj is with the Department of Signal Processing, Tampere University of Technology, Finland; e-mail: Moncef.gabbouj@tut.fi.

post-processing.

With the proper training the convolutional layers of CNNs can learn to extract *patient-specific* features over which the MLP layers can learn to produce the final class vectors of each beat. With the same limited training data as proposed in [10], [17]-[20], we shall demonstrate that simple CNNs with only 3 hidden layers will suffice to achieve a superior classification performance rather than the complex ones that are commonly used for deep learning tasks. We shall further show that the proposed 1D CNNs are easier to train with only few dozens of back-propagation (BP) epochs and can thus perform the classification task with an utmost speed (requiring only few hundreds of 1D convolutions). The data used for training the individual patient's classifier consists of two parts: global (common to each patient) and local (patient-specific) training patterns. While patient-specific data contains the first 5 minute segment of each patient's ECG record and is used as part of the training data to perform patient adaptation, the global data set contains a relatively small number of representative beats randomly chosen from each class in the training files and helps the classifier learn other arrhythmia patterns that are not included in the patient-specific data. An illustration of the BP training of the proposed approach is shown in Figure 1.

II. ADAPTIVE 1D CONVOLUTIONAL NEURAL NETWORKS

Adaptive 1D CNNs are used for both feature extraction and classification of the raw ECG data from each individual patient in the database. To simplify the CNN analogy, to have the freedom of any input layer dimension independent from the CNN parameters and also to have the ability of a 'CNN-only' design – without the MLP layers, neurons of the hidden CNN layers are extended to enable both convolution and sub-sampling operations as shown in Figure 2.

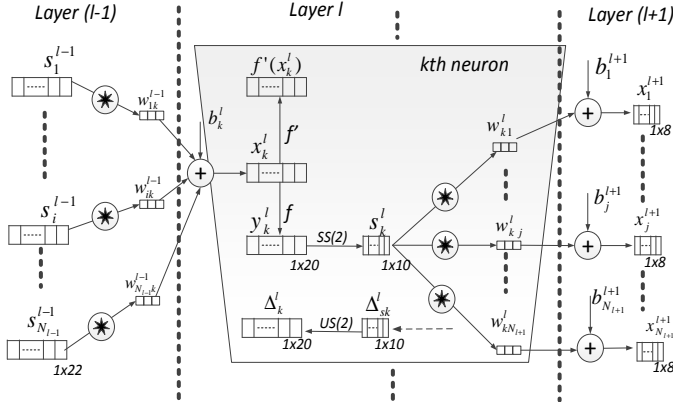


Figure 2: The adaptive 1D CNN implementation.

The 1D forward propagation (FP) can be expressed as,

$$x_k^l = b_k^l + \sum_{i=1}^{N_{l-1}} \text{conv1D}(w_{ik}^{l-1}, s_i^{l-1}) \quad (1)$$

where x_k^l is the input, b_k^l is the bias of the k^{th} neuron at layer l , and s_i^{l-1} is the output of the i^{th} neuron at layer $l-1$. w_{ik}^{l-1} is the

kernel from the i^{th} neuron at layer $l-1$ to the k^{th} neuron at layer l . With such a design it is aimed that the number of hidden CNN layers can be set to any number. This ability is possible in this implementation because the sub-sampling factor of the output CNN layer (the hidden CNN layer just before the first MLP layer) is set to the dimensions of its input map, e.g., in Figure 2 if the layer $l+1$ would be the output CNN layer, then the sub-sampling factors for that layer is automatically set to $ss = 8$ since the input map dimension is 8 in this sample illustration. Besides the sub-sampling, note that the dimension of the input maps will gradually decrease due to the convolution without zero padding, i.e., in Figure 2, the dimension of the neuron output is 22 at the layer $l-1$ that is reduced to 20 at the layer l . As a result of this, the dimension of the input maps of the current layer is reduced by $K-1$ where K is the size of the kernel.

We shall now briefly formulate the back-propagation (BP) steps while skipping the detailed derivations due to the space limitations. The BP of the error starts from the MLP output layer. Let $l=1$ and $l=L$ be the input and output layers, respectively. The error in the output layer can be written as,

$$E = E(y_1^L, \dots, y_{N_L}^L) = \sum_{i=1}^{N_L} (y_i^L - t_i)^2 \quad (2)$$

For an input vector p , and its corresponding output vector, $[y_1^L, \dots, y_{N_L}^L]$, we are interested to find out the derivative of this error with respect to an individual weight (connected to that neuron, k) w_{ik}^{l-1} , and bias of the neuron k , b_k^l , so that we can perform gradient descent method to minimize the error accordingly. Once all the delta errors in each MLP layer are determined by the BP, then weights and bias of each neuron can be updated by the gradient descent method. Specifically, the delta of the k^{th} neuron at layer l , Δ_k^l will be used to update the bias of that neuron and all weights of the neurons in the previous layer connected to that neuron, as:

$$\frac{\partial E}{\partial w_{ik}^{l-1}} = \Delta_k^l y_i^{l-1} \quad \text{and} \quad \frac{\partial E}{\partial b_k^l} = \Delta_k^l \quad (3)$$

So from the input MLP layer to the output CNN layer, the regular (scalar) BP is simply performed as,

$$\frac{\partial E}{\partial s_k^l} = \Delta s_k^l = \sum_{i=1}^{N_{l+1}} \frac{\partial E}{\partial x_i^{l+1}} \frac{\partial x_i^{l+1}}{\partial s_k^l} = \sum_{i=1}^{N_{l+1}} \Delta_i^{l+1} w_{ki}^l \quad (4)$$

Once the first BP is performed from the next layer, $l+1$, to the current layer, l , then we can further back-propagate it to the input delta, Δ_k^l . Let zero order up-sampled map be: $us_k^l = up(s_k^l)$, then one can write:

$$\Delta_k^l = \frac{\partial E}{\partial y_k^l} \frac{\partial y_k^l}{\partial x_k^l} = \frac{\partial E}{\partial us_k^l} \frac{\partial us_k^l}{\partial y_k^l} f'(x_k^l) = up(\Delta s_k^l) \beta f'(x_k^l) \quad (5)$$

where $\beta = (ss)^{-1}$ since each element of s_k^l was obtained by averaging ss number of elements of the intermediate output, y_k^l . The inter-BP (among CNN layers) of the delta error ($\Delta s_k^l \leftarrow \sum \Delta_i^{l+1}$) can be expressed as,

$$\Delta s_k^l = \sum_{i=1}^{N_{l+1}} \text{conv1Dz}(\Delta_i^{l+1}, \text{rev}(w_{ki}^l)) \quad (6)$$

where $\text{rev}(\cdot)$ reverses the array and $\text{conv1Dz}(\cdot, \cdot)$ performs full convolution in 1D with $K-1$ zero padding. Finally, the weight and bias sensitivities can be expressed as,

$$\frac{\partial E}{\partial w_{ki}^l} = \text{conv1D}(s_k^l, \Delta_i^{l+1}) \quad \frac{\partial E}{\partial b_k^l} = \sum_n \Delta_k^l(n) \quad (7)$$

III. EXPERIMENTAL RESULTS

The beats are represented using 128 samples centered around the R-peak point. We purposefully used a simple 1D CNN in all experiments with only 3 CNN layers and 2 MLP layers, in order to achieve an utmost computational efficiency for both training and particularly for real-time classification. On top of this we aim to demonstrate that deep learners are not indeed needed to achieve a superior ECG classification performance. The 1D CNN used in all experiments has 32 and 16 neurons in the first and second hidden CNN layers and 10 neurons in the hidden MLP layer. The output (MLP) layer size is 5 which is the number of beat classes and the input (CNN) layer size is 2. For 128 sample beat representations, the kernel sizes are set to $K=15$, and the sub-sampling factors are set to $ss=6$. As a result the proposed adaptive CNN implementation, the sub-sampling factor for the last CNN layer is adaptively set to 5.

For all experiments we employ a shallow training: the maximum number of BP iterations is set to 50 and another stopping criterion is the minimum train classification error level that is set to 3% to prevent over-fitting. Therefore, the training will terminate if either of the criteria is met. We initially set the learning factor, ϵ , as 0.001 and applied a global adaptation during each BP iteration: if the train MSE decreases in the current iteration we slightly increase ϵ by 5%; otherwise, we reduce it by 30%, for the next iteration. As BP is a deterministic gradient-descent optimization technique, which makes it quite dependent on the initial (random) setting of the network parameters (kernels, weights and biases), we performed 10 individual BP runs for each patient in the database and the average classification performance is reported.

We performed classification experiments on 44 records of the MIT/BIH arrhythmia database, which includes a total of 100389 beats to be classified into five heartbeat types following the AAMI convention [16]. For the classification experiments in this paper, the common part of the training data set contains a total of 245 representative beats, including 75 from each type N, -S, and -V beats, and all (13) type-F and (7) type-Q beats, randomly sampled from each class from the first 20 records (picked from the range 100 to 124) of the MIT/BIH database. The patient-specific training data includes the beats from the first 5 min of the corresponding patient's ECG record. Patient-specific CNNs are trained with a total of 245 common training beats and a variable number of patient-specific beats depending

on the patient's heart rate, so only less than 1% of the total beats are used for training. The remaining beats (25 min) of each record, in which 24 out of 44 records are completely new to the classifier, are used as test patterns for performance evaluation.

Classification performance is measured using the four standard metrics found in the literature [10]: classification accuracy (Acc), sensitivity (Sen), specificity (Spe), and positive predictivity (Ppr). While accuracy measures the overall system performance over all classes of beats, the other metrics are specific to each class and they measure the ability of the classification algorithm to distinguish certain events (i.e. VEBs or SVEBs) from nonevents (i.e. non-VEBs or non-SVEBs). The respective definitions of these four common metrics using true positive (TP), true negative (TN), false positive (FP), and false negative (FN) are as follows: *Accuracy* is the ratio of the number of correctly classified patterns to the total number of patterns classified, $Acc = (TP+TN)/(TP+TN+FP+FN)$; *Sensitivity* is the rate of correctly classified events among all events, $Sen = TP/(TP+FN)$; *Specificity* is the rate of correctly classified nonevents among all nonevents, $Spe = TN/(TN+FP)$; and *Positive Predictivity* is the rate of correctly classified events in all detected events, $Ppr = TP/(TP+FP)$. Since there is a large variation in the number of beats from different classes in the training/testing data (i.e. 39465/50354 type-N, 1277/5716 type-V, and 190/2571 type-S beats), sensitivity, specificity, and positive predictivity are more relevant performance criteria for medical diagnosis applications.

The classification performance of the proposed system is compared with the four existing algorithms, [10], [17], [18], and [19], all of which comply with the AAMI standards. For a comparative performance evaluation, the problem of VEB and SVEB detection is considered individually and results are summarized in Table I. The benchmark database is partitioned into three evaluation datasets. For VEB detection the dataset 1 contains 11 test recordings (200, 202, 210, 213, 214, 219, 221, 228, 231, 233, and 234) and for SVEB detection, comparison results are based on 14 common recordings (with the addition of records 212, 222, and 232). Dataset 1 is common for all competing methods. Dataset 2 is the test partition of the benchmark database containing 24 records (200 and onwards). Three methods (the proposed, [18], and [19]) are tested on this dataset. Finally, the dataset 3 is the entire database with all records over which two methods, the proposed and [19] are tested.

Several interesting observations can be made from these results. First, for SVEB detection, sensitivity and positive predictivity rates are comparably lower than VEB detection, while a high specificity performance is achieved. The reason for the worse classifier performance in detecting SVEBs is that SVEB class is under-represented in the training data and hence more SVEB beats are misclassified as normal beats and vice versa. In overall, for both test datasets (1 and 2), the performance level of the proposed approach in both VEB and SVEB detection is comparable or better than the competing methods for most of the measures. Particularly for VEB classification, the highest classification accuracies, and sensitivities are consistently achieved over all datasets by the proposed approach. Over the entire dataset (dataset 3) it has the highest performance measures for both VEB and SVEB classification except the SVEB sensitivity.

We have implemented the adaptive 1D CNN with C++ programming over MS Visual Studio 2013 in 64bit. This is a non-GPU implementation and the experiments are performed on a computer with I7-4700MQ at 2.4GHz CPU. Specifically

for the single-CPU implementation, the total time for a FP of a single beat to obtain the class vector is about 0.74 *msec* which is more than 1000x faster than the real-time requirement.

Table I: VEB and SVEB classification performance of the proposed method and comparison with the four major algorithms from the literature (best results are highlighted).

Methods	VEB				SVEB			
	<i>Acc</i>	<i>Sen</i>	<i>Spe</i>	<i>Ppr</i>	<i>Acc</i>	<i>Sen</i>	<i>Spe</i>	<i>Ppr</i>
Hu et al. [10] ¹	94.8	78.9	96.8	75.8	N/A	N/A	N/A	N/A
Chazal et al. [17] ¹	96.4	77.5	98.9	90.6	92.4	76.4	93.2	38.7
Jiang and Kong [18] ¹	98.8	94.3	99.4	95.8	97.5	74.9	98.8	78.8
Ince et al. [19] ¹	97.9	90.3	98.8	92.2	96.1	81.8	98.5	63.4
Proposed¹	98.9	95.9	99.4	96.2	96.4	68.8	99.5	79.2
Jiang and Kong [18] ²	98.1	86.6	99.3	93.3	96.6	50.6	98.8	67.9
Ince et al. [19] ²	97.6	83.4	98.1	87.4	96.1	62.1	98.5	56.7
Proposed²	98.6	95	98.1	89.5	96.4	64.6	98.6	62.1
Ince et al. [19] ³	98.3	84.6	98.7	87.4	97.4	63.5	99.0	53.7
Proposed³	99	93.9	98.9	90.6	97.6	60.3	99.2	63.5

¹The comparison results are based on 11 common recordings for VEB detection and 14 common recordings for SVEB detection.

²The VEB and SVEB detection results are compared for 24 common testing records only.

³The VEB and SVEB detection results of the proposed system for all training and testing records.

IV. CONCLUSIONS

In this work, we proposed a patient-specific ECG heartbeat classifier with an adaptive implementation of 1D Convolutional Neural Networks (CNNs) that are able to fuse the two major blocks of the traditional ECG classification into a single *learning* body: feature extraction and classification. Such a compact implementation for each patient over a simple CNN not only negates the necessity to extract hand-crafted manual features, or any kind of pre- and post-processing, also makes it a primary choice for a real-time implementation of heart monitoring and anomaly detection. Besides the amazing speed and computation efficiency achieved, the proposed method *only* requires 1D convolutions (multiplications and additions) that make any hardware implementation simpler and cheaper. In addition to that once a dedicated CNN is trained for an individual patient, it can solely be used to classify his/her long ECG records such as Holter registers in a fast and accurate manner.

The results of the classification experiments, which are performed over the benchmark MIT/BIH arrhythmia database show that the proposed approach can achieve the highest classification performance in most of the measures and among the previous state-of-the-art methods that were tested according to the AAMI recommendations. As a result, the proposed approach achieves the main design objectives, i.e. maintaining a fast, robust and patient-specific system with a superior classification performance.

REFERENCES

- [1] K. Minami, H. Nakajima, and T. Toyoshima, "Real-Time discrimination of ventricular tachyarrhythmia with Fourier-transform neural network," *IEEE Trans. Biomed. Eng.*, vol. 46, no. 2, pp. 179–185, Feb. 1999.
- [2] L. Y. Shyu, Y. H. Wu, and W. C. Hu, "Using wavelet transform and fuzzy neural network for VPC detection from the holter ECG," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 7, pp. 1269–1273, Jul. 2004.

- [3] Inan, et al., "Robust neural-network based classification of PVCs using wavelet transform and timing interval features," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pp. 2507–2515, Dec. 2006.
- [4] X. Alfonso and T. Q. Nguyen, "ECG beat detection using filter banks," *IEEE Trans. Biomed. Eng.*, vol. 46, no. 2, pp. 192–202, Feb. 1999.
- [5] J. L. Willems and E. Lesaffre, "Comparison of multigroup logistic and linear discriminant ECG and VCG classification," *J. Electrocardiol.*, vol. 20, pp. 83–92, 1987.
- [6] J. L. Talmon, *Pattern Recognition of the ECG*. Berlin, Germany: Akademisch Proefschrift, 1983.
- [7] D. A. Coast, R. M. Stern, G. G. Cano, and S. A. Brillner, "An approach to cardiac arrhythmia analysis using hidden Markov models," *IEEE Trans. Biomed. Eng.*, vol. 37, no. 9, pp. 826–836, Sep. 1990.
- [8] S. Osowski, L. T. Hoai, and T. Markiewicz, "Support vector machine based expert system for reliable heartbeat recognition," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 4, pp. 582–589, Apr. 2004.
- [9] Y. H. Hu, W. J. Tompkins, J. L. Urrusti, and V. X. Afonso, "Applications of artificial neural networks for ECG signal detection and classification," *J. Electrocardiol.*, pp. 66–73, 1994.
- [10] Y. Hu, S. Palreddy, and W. J. Tompkins, "A patient-adaptable ECG beat classifier using a mixture of experts approach," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 9, pp. 891–900, Sep. 1997.
- [11] S. C. Lee, "Using a translation-invariant neural network to diagnose heart arrhythmia," in *IEEE Proc. Conf. Neural Information Processing Systems*, Nov. 1989.
- [12] D. C. Ciresan, U. Meier, L. M. Gambardella, and J. Schmidhuber, "Deep big simple neural nets for handwritten digit recognition," *Neural Computation*, vol. 22, no. 12, pp. 3207–3220, 2010.
- [13] D. Scherer, A. Muller, and S. Behnke, "Evaluation of pooling operations in convolutional architectures for object recognition," in *Int. Conf. on Artificial Neural Networks*, 2010.
- [14] A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet classification with deep convolutional neural networks", In *Advances in Neural Information Processing Systems (NIPS)*, 2012.
- [15] P. de Chazal and R. B. Reilly, "A patient-adapting heartbeat classifier using ECG morphology and heartbeat interval features," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pp. 2535–2543, Dec. 2006.
- [16] "Recommended practice for testing and reporting performance results of ventricular arrhythmia detection algorithms," Association for the Advancement of Medical Instrumentation, Arlington, VA, 1987.
- [17] P. de Chazal, M. O'Dwyer, and R. B. Reilly, "Automatic classification of heartbeats using ECG morphology and heartbeat

- interval features,” *IEEE Trans. Biomed. Eng.*, vol. 51, no. 7, pp. 1196–1206, Jul. 2004.
- [18] W. Jiang and S. G. Kong, “Block-based neural networks for personalized ECG signal classification,” *IEEE Trans. Neural Networks*, vol. 18, no. 6, pp. 1750–1761, Nov. 2007.
- [19] T. Ince, S. Kiranyaz, and M. Gabbouj, “A Generic and Robust System for Automated Patient-specific Classification of Electrocardiogram Signals”, *IEEE Transactions on Biomedical Engineering*, vol. 56, issue 5, pp. 1415-1426, May 2009.
- [20] S. Kiranyaz, T. Ince and M. Gabbouj, *Multi-dimensional Particle Swarm Optimization for Machine Learning and Pattern Recognition*, Book: Springer, 383 pages, Aug. 2013.