PATIENT-ADAPTED AND INTER-PATIENT ECG CLASSIFICATION USING NEURAL NETWORK AND GRADIENT BOOSTING

L. Zaorálek, *J. Platoš*, *V. Snášel**

Abstract: Heart disease diagnosis is an important non-invasive technique. Therefore, there exists an effort to increase the accuracy of arrhythmia classification based on ECG signals. In this work, we present a novel approach of heart arrhythmia detection. The model consists of two parts. The first part extracts important features from raw ECG signal using Auto-Encoder Neural Network. Extracted features obtained by Auto-Encoder represent an input for the second part of the model, the Gradient Boosting and Feedforward Neural Network classifiers. For comparison purposes, we evaluated our approach by using MIT-BIH ECG database and also following recommendations of the Association for the Advancement of Medical Instrumentation (AAMI) for ECG class labeling. We divided our experiment into two scenarios. The first scenario represents the classification task for the patient-adapted paradigm and the second one was dedicated to the inter-patient paradigm. We compared the measured results to the state-of-the-art methods and it shows that our method outperforms the state-of-the art methods in the Ventricular Ectopic (VEB) class for both paradigms and Supraventricular Ectopic (SVEB) class in the inter-patient paradigm.

Key words: *ECG, AAMI, VEB, SVEB, GBM, ANN*

1. Introduction

The electrocardiographic (ECG) recordings offer a non-invasive method to analyse heart activity [5, 36]. The analysis can reveal cardiovascular diseases based on heart abnormalities. The ECG analysis brings the necessary manual review of each heartbeat and classify it. For that reason the analysis process is a very tedious and time-consuming task for humanbeings. Also there is a chance of human errors because the heart beat activity obtained by an ECG device, such as the ECG holter, has hours or even days of the ECG record. Over the last two decades, several approaches have been developed in order to automatically classify each beat and

---
*Lukáš Zaorálek – Corresponding author; Jan Platoš; Václav Snášel; Department of Computer Science, FEECS, VŠB – Technical Univesity of Ostrava, 17. listopadu 2172/15, Ostrava 708 33, Czech republic, E-mail: lukas.zaoralek@vsb.cz, jan.platos@vsb.cz, vaclav.snasel@vsb.cz

detect heart abnormalities in the ECG recordings. The advantage of the automated aproach is reducing human errors, low cost of the analysis and significantly decreased time to analyse. The heart beat can be classified as normal or abnormal depending on the heart activity. In the case of the abnormal heart beat activity, the heart has irregular rhythm. These rhythm disturbances are known as arrhythmias. There are several different irregular heart rhythms or arrhythmias that exist, for example extra beats, supraventricular tachycardias, ventricular arrhythmias, and bradyarrhythmias. In this paper, we present a novel approach in the classification task followed by [1]. There are three paradigms that exist on how to evaluate the automated system for classification of the ECG recordings:

- intra-patient

- inter-patient

- patient-adapted/patient-specific

The first one is the intra-patient paradigm. The intra-patient paradigm is focused on the classification of the ECG label. It means that the ECG recordings are randomly divided into training and test subsets. Therefore in this case, the ECG record belonging to one patient can occur in both of subsets. The evaluation of such an automated system has optimistic results, see [29]. The second paradigm is called inter-patient and brings out a more realistic scenario on how to evaluate the automated systém [17]. Specifically, the ECG recording of one patient can be in training or testing subset but not in both subsets simultaneously. The last one is patient-specific or patient-adapted paradigm [8, 16]. In this case, the first 2-5 minutes of each ECG test recording are evaluated manually by cardiologists.

In recent years, several classification methods have been used in the task of arrhythmia detection, for example [9, 19, 27] or [34]. The classification methods are composed of four basic steps [2, 13]: ECG signal preprocessing, heartbeat segmentation, feature extraction and classification. Also in this paper we follow these steps, which are described in the Section 3. The preprocessing and hearbeat segmentation were described in [25, 30, 35]. The feature extraction step was widely explored by multiple sources . Almost all works in the context of the feature extraction used R-R interval (distance between two R-peaks of heartbeats), for example [11]. Another commonly used technique for the features extraction is QRS interval [22], wavelet transformation [15, 24] or reduction of dimensionality [6, 21, 40]. There are four popular algorithms that exist which are used for the classification step: Support Vector Machines, Linear Discriminant Analysis, Artificial Neural Network and Decision Tree, such as [3, 23, 28, 31, 37, 39].

2. Dataset

In this work we used the MIT-BIH arrhythmia database [32]. This database is publicly available on Physionet. There are 48 two-lead recordings obtained from 25 men aged 32 to 89 years, and 22 women aged 23 to 89 years. There is a total of 109 000 annotated beats in the database. There are 15 possible annotations in MIT-BIH database. The annotation of each beat was corrected by two cardiologists

who worked independently on their corrections. For the purpose of comparison of our proposed methods with other works we used recommendations from the Association for the Advancement of Medical Instrumentation (AAMI). The AAMI designs the standard approach on how to evaluate methods and algorithms for ECG classification task. Specifically, we selected four to five annotations based on AAMI recommendations as shown in the Tab. I. In addition, the work [7] follows the AAMI recommendations and applies them to the MIT-BIH database. They divided MIT-BIH database into two subsets: DS1 and DS2.Specifically, the DS1 subset has following records: 101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223, 230. The DS2 includes 100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, 234 records. The following Tab. II summarizes the AAMI scheme for the MIT-BIH database. The recordings from DS1 are used for the training phase of the classification and the DS2 subset is dedicated to test and evaluate the performance of the proposed approach. Also other publications such as [8,12,13,16,18,19,27,28] and our own methods used the MIT-BIH database, AAMI scheme and de Chazal et al. s concept in order to compare performance of the ECG classifiers.

Group name	Shortcut	Symbol	Number of beats
Normal	N	N, e, j, L, R	90125
Ventricular Ectopic	V	V, E	2781
Supraventricular Ectopic	S	A, a, J, S	7009
Fusion	F	F	803
Unknown	Q	U	15

Tab. I *AAMI annotation groups for* MIT-BIH *database.*

Datasets	N	S	V	F	Q	Total
DS1	45866	944	3788	415	8	51021
DS2	44259	1837	3221	388	7	49712
Total	90125	2781	7009	803	15	100733

Tab. II *The* AAMI *scheme for* MIT-BIH *database.*

3. Methodology

Fig. 1 shows the scheme of our proposed approach. First, we removed noise from the ECG signal by using wavelet decomposition. After reconstructing the ECG signal without noise we normalized the signal. Next, we segmented each heartbeat and associated its annotation. In order to select the most important features from each ECG hearbeat, we performed a feature selection by using a type of Neural Network, namely the Auto-Encoder model. Finally, we trained proposed methods and measured the performance compared to other works. The following statements describe preprocessing, feature selection and classifier models in more detail.
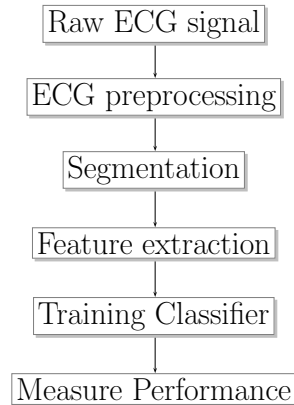
Raw ECG signal

ECG preprocessing

Segmentation

Feature extraction

Training Classifier

Measure Performance

Fig. 1 *The classification schema.*

3.1 ECG Preprocessing

In the preprocessing phase of the experiments, we removed a noise from each ECG signal. There are exists numerous approach that exist on how to denoise the ECG signal, such as an impulse response (FIR), a wavelet transformation or (Extended) a Kalman Filter. We chose discrete wavelet transformation as a denoising method based on promising results from [26, 33]. First of all, we calculated the discrete wavelet transformation (DWT) just as described by the Eq. (1),

$$\text{DWT}(m, k) = \frac{1}{\sqrt{a_0^m}} \sum_n x(n) \Psi\left(\frac{k - n b_0 a_0^m}{s}\right), \qquad (1)$$

where $\Psi(\cdot)$ is the mother wavelet, $a = a_0^m$ and $b = n b_0 a_0^m$ are the scaling and translation parameters, respectively [20, 33]. As the mother wave for DWT we selected "db4". After the DWT computation we performed thresholding in the wavelet domain. Next, we shrank the wavelet coefficients by thresholding. Finally, we reconstructed the ECG signal from thresholded wavelet coefficients using inverse DWT transformation. See Fig. 2 for details on how the ECG signal looks before and after removing high frequency noise. After the signal denoising procedure it is necessary to rescale the ECG recordings to normal range from -1 to 1 because we used neural network (NN) for the next step. The neural network cannot work properly if the range of the ECG signal is different than the normal range from -1 to 1.

3.2 Segmentation

For the classification purpose we divided each preprocessed ECG waveform into segments and created a segmented dataset, such as $(X_1, Y_1), \ldots, (X_p, Y_p)$, where p is the number of all segments created from preprocessed recordings, and X_i represents the vector of samples belonging to one segment and Y_i is associated annotation located nearby each R-peak in the original MIT-BIH database. For
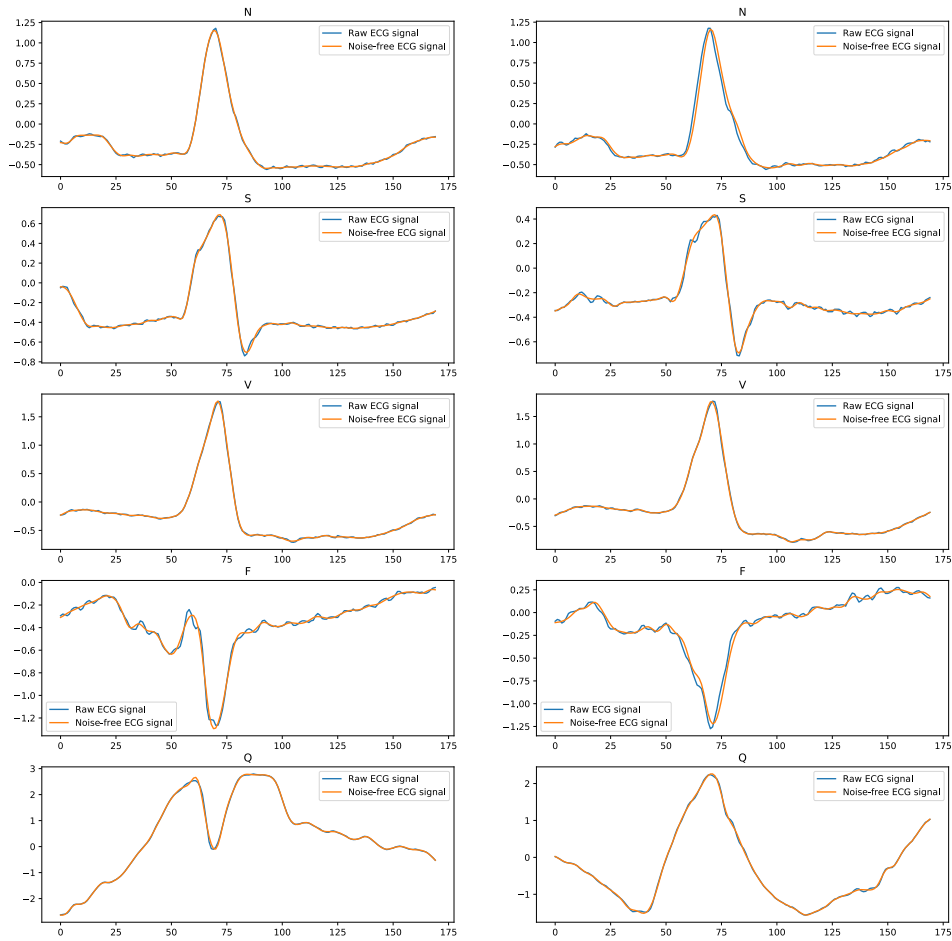
Fig. 2 *Comparison ECG signals before and after applied DWT.*

example, [14] uses 90 samples on either side of the R-peak. However, based on our experiences and results it is sufficient to take 70 samples before and 20 samples after the R-peak location in the case of the MIT-BIH database. Note that, the last segment of a record is avoided from the dataset if it does not have 20 samples after the R-peak.

3.3   Feature selection

There are many approaches that exist on how to select important features from the ECG signals for the classification task. The feature selection can be divided into three general groups: temporal, morphological and statistical. In our methods, we employed the AE model for extracting important features from the ECG waveform as morphology features. Also we included temporal features for the following classification step: R-R intervals from last 30 heartbeats (including the current one).

Therefore the first 30 heartbeats of the dataset are avoided. The training algorithm of the AE model is the same as described in the next section, Neural-network. The algorithm remains the same, but the output of the AE model is a reflection of the input vector. Therefore, the AE model is a self-supervised regression technique in which the desired output vector is the input vector. The effeciency of the AE model in the task of the feature selection has been proved in several works, such as [2, 4, 10, 38].

We utilized grid search method to find optimal hyperparameters of the AE model. Concretely, we built the AE model with 3 dense layers. The first and last dense layer has 90 neurons and the second dense layer ("latent layer") has 32 neurons. The Tab. III shows details of the AE model. Note that, $I(n)$ means an input layer, $D(n)$ is a dense layer, BN is a batch normalization layer, $O(n)$ is an output layer and n is the number of output neurons from each layer. The AE transforms each segment X_i into 32 dimensional latent spaces.

Parameter	Value/Type
Topology	[I(90), D(90), BN, D(32), BN, D(90), O(90)]
Loss function	RMSE
Optimizer	RMSPROP
Learning rate	0.01

Tab. III *The* AE *setup and hyperparameters.*

3.4 Classification

After feature selection the dataset was prepared to train the classifiers. The training phase includes two independent scenarios. In the first scenario, we performed classification tasks for patient-adapted paradigm. This means that the first 5 minutes of each DS2 record were added to DS1 subset. For the second scenario, based on the performance evaluation obtained from first scenario we chose better classifiers and measured performance within inter-patient paradigm. There were two classifiers in our experiments. The First was the Gradient Boosting Model (GBM) and the second one was the Feedforward Neural Network (FNN) model. See Tabs. IV and V for concrete setup and hyperparameters of the GBM and FNN models. The following statements describe how FNN and GBM model work in more detail.

Parameter	Value/Type
Topology	[I(62), D(32), BN, D(16), BN, O(5)]
Loss function	Cross-Entropy
Optimizer	RMSPROP
Learning Rate	0.01

Tab. IV *The* FNN *setup and hyperparameters.*

Parameter	Value/Type
Depth	10
Eta	0.01
Estimators	200
Boost Round	200
Min child weight	1
Subsample ratio	1

Tab. V *The Gradient Boosting setup and hyperparameters.*

3.4.1 Neural network

The Artificial Neural Network (ANN) is a model inspired by biological neural network. The ANN model is composed in layers. Each layer has a number of neurons. The topology of the ANN model defines how neurons in the layer are connected with neurons from the next layer. Note that it is possible to create connections with neurons from previous layer. This ANN model is called the recurrent NN model. However, in our experiment, we focused only on the forward ANN model without recurrent connections between neurons. The one connection between k-th neuron from $(l-1)$ layer and j-th neuron from l-th layer is defined as weight $w_j k^l$. Each neuron has internal potencial z_j^l defined as weighted input for the activation function σ. The internal potential and the activation function for the j-th neuron in the l-th layer is shown as (4) and (3) where b_j^l is a neuron bias.

Let's suppose ECG dataset $(X_1, Y_1), \ldots, (X_p, Y_p)$, where $X_j \in \mathbb{R}^n$ is a vector for multilead ECG signal and $Y_j \in \mathbb{R}^m$ is a vector for associated annotation label[1]. We would like to train the ANN model to predict label Y_j based on the input vector X_j. In other words, the goal is to minimize a loss function, such as a root mean square error (RMSE), between the actual output of the ANN model $a^L(X_j)$ based on the sample X_j, and the desired output vector Y_j as shown in Eq. (2), where L is the number of layers

$$C = \frac{1}{2n} \sum_{j=1}^{p} (Y_j - a^L(X_j))^2, \tag{2}$$

$$a_j^l = \sigma(z_j^l), \tag{3}$$

$$z_j^l = \sum_k w_{jk}^l a_k^{l-1} + b_j^l. \tag{4}$$

The way to improve neuron weights in order to minimize loss function is through the backpropagation algorithm. The backpropagation algorithm calculates the value of the loss function C and is based on the gradient composed by Eqs. (7), (8), (9), (10) updates each weight and neuron bias of the ANN model as shown (5), (6)

---

[1] Annotation should be represented by one-hot encoding.

$$w_{jk}^l = w_{jk}^l - \eta \frac{\partial C}{\partial w_{jk}^l}, \tag{5}$$

$$b_j^l = b_j^l - \eta \frac{\partial C}{\partial b_j^l}, \tag{6}$$

$$\frac{\partial C}{\partial w_{jk}^l} = a_k^{l-1} \delta_j^l, \tag{7}$$

$$\frac{\partial C}{\partial b_j^l} = \delta_j^l, \tag{8}$$

$$\delta_j^L = \frac{\partial C}{\partial a_j^L} \sigma'(z_j^l), \tag{9}$$

$$\delta_j^l = \sum_k w_{jk}^{l+1} \delta_k^{l+1} \sigma'(z_j^l). \tag{10}$$

3.4.2   Gradient Boosting Model

The GBM is a type of machine learning technique, where the model is composed from weak predictors (models). Typically, a weak predictor is a Decision Tree (like CART tree) but it could be Ordinary Linear Regression (OLR), Ridge Penalized Linear Regression (RPLR) or other predictors. The idea behind the GBM is to minimize loss function by iteratively adding additional weak predictors into the model based on expected values of the training dataset. Note that a previously added weak predictor does not change in the next learning iteration. Let's suppose that N datapoints pair X_j, Y_j of a dataset, where $X_j = \{x_1, \ldots, x_d\}$ is a multivariate input variable of the dataset and Y_j is an univariate output variable of the dataset. The following Eqs. (11) describe the learning phase of GBM classifier. The function $\bar{f} : \mathbb{R}^d \to \mathbb{R}$ is a minimization of expectations of the loss function as it shows in Eq. (11):

$$\bar{f}(X) = argmin_{f(X)} E_{y,x}[L(Y, f(X))], \tag{11}$$

where $\bar{f}$ is the additive model with minimized loss function for the multivariate input variable X and univariate output variable Y, L is the loss function and f is the additive model. Following Eq. (12) shows how GBM model represents by the function $\bar{f}(X)$ is composed from weaker models f_i:

$$\bar{f}(X) = \sum_{i=1}^M \beta_i f_i(X) = \sum_{i=1}^M \beta_i h_i(X; \theta_i), \tag{12}$$

where M is the number of algorithm iterations and f_i is the weak model and β_i is the expansion coefficient. The function h_i represents parameterized decision tree of weak model f_i, where $\theta_i = \{A_i, \gamma_i\}$ is parameter set of the tree. The parameter set θ_i has the region set A_i and the coefficients γ_i. The γ_i represents all outputs for all examples X associated to the region set A_i. The Eq. (11) does not have closed-form sulution. Instead of finding a solution in function space the training

algorithm can use the steepest gradient descent approach because the loss function L is differentiable. Eq. (13) shows the first step of the gradient descent update:

$$g_{ij}(X_j, Y_j, f) = \frac{\partial L[Y_j, f(X_j)]}{\partial f(X_j)},$$ (13)

where X_j is an j-th mutlivariate input variable, Y_j is an univariate output variable associated with the X_j variable, L is loss function and f is the function of additive model at step $i - 1$. The second step of the gradient descent update is the gradient originating from Eq. (13) transforms to construct i-th decision tree. The step seems to very hard task because it means find the solution for the boost increment in function space. The solution of this particular problem is a "proxy", where instead of working in function space, one can construct new additive model that is most correlated with $-g_i(x)$ with step-size ρ:

$$(\theta_i, \rho_i) = argmin_{\theta, \rho} = \sum_{j=1}^{N} (\rho h(X_j; \theta) - g_{ij}(X_j, Y_j, \bar{f}_{i-1}))^2,$$ (14)

where N is number of samples in the dataset, θ_i is found new tree parameters at i-th iteration, $h(X_j, \theta)$ is parameterized decision tree. Also the expansion coefficients can be estimated by Eq. (15):

$$\beta_j = argmin_{\beta} \sum_{j=1}^{N} L(Y_j, f_{i-1}(X_j) + \beta h(X_i; \theta_i)).$$ (15)

Finally, Eq. (16) defines construction of the new model with (β_i, θ_i) by using least squares fit method obtained from Eqs. (14) and (15) as follows:

$$\bar{f}_i = \bar{f}_{i-1} + \beta_i h(X, \theta_i).$$ (16)

4. Results

The following statements present the results for patient-adapted and inter-patient paradigms followed by the AAMI standard. First, we introduce the results for patient-adapted paradigma, where we used all five classes described in Section 2. Next we show the classification performance of three classes for inter-patient paradigms. In both cases, we compared our proposed methods with other state-of-the-art algorithms. Note that the classifiers were always trained on DS1 subset and performance was measured on DS2 subset as it was defined in Section 2 and mentioned in Section 3. The performance evaluation was calculated based on the sensitivity (Se), positive predictivity (Pp), specificity (Sp), and overall accuracy (Oa) for the class VEB (V) and the class SVEB (S) in the same way as it was proposed by [7].

4.1 Patient-adapted paradigma

Tab. VI shows summary beat-by-beat results of the GBM classifier for DS2 subset. There are 5 classes (N, S, V, F, and Q) recommended by AAMI standard. The

algorithms achieved higher true positive ratio for the class V in comparison to the class S. On the other hand, the proposed algorithm has problems recognizing class Q. The reason why Q was not recognized by the algorithm is that the class Q has very low ocurrence (see Tab. I). The total accuracy of the proposed method was 97.9 %.

		Classifier					
		N	S	V	F	Q	Total
	N	43886	140	39	123	0	44188
	S	525	1297	13	2	0	1837
Truth	V	172	10	3013	25	0	3220
	F	70	1	32	285	0	388
	Q	4	0	3	0	0	7
	Total	44657	1448	3100	435	0	49640

Tab. VI *The confusion matrix of the* GBM *classifier for DS2 dataset.*

In case of the FNN classifier, the algorithm has worse accuracy then the GBM classifier. Concretely, the FNN classifier correctly recognized only 249 S beats of the total 1837 S beats and 1898 V beats of total 3220 V beats. The classifier was unable to recognize any Q beat due to a problem with the number of Q beats.

		Classifier					
		N	S	V	F	Q	Total
	N	44158	1	28	1	0	44188
	S	1433	249	155	0	0	1837
Truth	V	1320	2	1898	0	0	3220
	F	308	0	48	32	0	388
	Q	5	0	2	0	0	7
	Total	47224	252	2131	33	0	49640

Tab. VII *The confusion matrix of the* FNN *classifier for DS2 dataset.*

Also we compare our proposed algorithms with three existing methods [8,16,18, 19]. Note that these methods also comply with AAMI standards. See Tab. VIII, where Se is sensitivity, Pp is positive predictivity, Sp is specificity and Oa is overall accuracy. Our methods clearly outperformed other mentioned methods. As we can see, our method has a higher positive predictive ratio and specificity for both classes S and V. Also, the sensitivity of the class V has the highest value, 94.0 %. Only in the case of the S class, the proposed method with GBM classifier has a lower sensitivity ratio. But, in the whole context of the measured statistics, such as sensitivity, positive predictivity and specificty, the GBM classifier for class S also achieved better performance than other methods. Mainly, the positive predictivity of the GBM classifier was 89.6 %. The second classifier which we proposed, the FNN classifier, achieved lower performance than the GBM classifier in both classes S and V.

Method	SVEB				VEB			
	Se	Pp	Sp	Oa	Se	Pp	Sp	Oa
de Chazal and Reilly [8]	87.7	47	N/A	95.9	94.3	96.2	N/A	99.4
Hu et al. [16]	N/A	N/A	N/A	N/A	78.9	75.8	96.8	94.8
Ince et al. [18]	81.8	63.4	98.5	96.1	90.3	92.2	98.8	97.9
Jiang and Kong [19]	74.9	78.8	98.8	97.5	94.3	95.8	99.4	98.8
Proposed (GBM)	**70**	**89.6**	**99.7**	**98.6**	**94.3**	**98.3**	**99.9**	**99.5**
Proposed (FNN)	13.6	98.8	99.9	96.8	58.9	91.2	99.6	96.9

Tab. VIII *Comparison patient-specific results for* DS2 *(In percent).*

4.2 Inter-patient paradigma

In this section, we present performance evaluation and comparison of the best proposed method from Patient-adapted paradigm. The chosen classifier based on the results from Tab. VIII was the GBM classifier. The GBM classifier we compared with other state-of-the-art methods within the class S and V. Note that, we ommit measurement of the class Q and F in this measurement evaluation due to very low occurence of the Q and F beat samples. Tab. IX shows confusion matrix of the GBM classifier on the DS2 subset. The GBM classifier recognized 43876 N beats and 2950 V beats but it had a lower predictivity rate of the S beats. Concretely, it predicted only 113 S beats of total 1836 S beats.

		Classifier			
		N	S	V	Total
Truth	N	43876	209	107	44192
	S	1616	113	107	1836
	V	255	15	2950	3220
	Total	45747	337	3164	49248

Tab. IX *The confusion matrix of the* GBM *classifier for inter-patient paradigm and DS2 subset.*

See also Tab. X where there is a comparison of the GBM classifier with other state-of-the-art methods, where Se is sensitivity, Pp is positive predictivity, Sp is specificity and Oa is overall accuracy. The GBM model has significantly out-performed other methods in case of the class V. Concretely, the GBM classifier achieved 91.6 % for sensitivity, 93.2 % for positive predictivity, 99.5 % for speci-

ficity and 99.0 % accuracy for the class V. But in case of the class S, the GBM classifier has only 6 % sensitivity because it predicted only 113 S beats of total 1836 S beats.

Method	SVEB				VEB			
	Se	Pp	Sp	Oa	Se	Pp	Sp	Oa
Lin & Yang [27]	81.0	31.0	N/A	N/A	86.0	73.0	N/A	N/A
Llamedo & Martínez [28]	77.0	39.0	95.3	94.7	81.0	87.0	99.0	97.7
Garcia et. al, 2016 [12]	30	26	98.3	95.0	85	66	96.2	95.4
Garcia et. al, 2017 [13]	62	53	97.9	96.6	87.3	59.4	95.9	95.4
Proposed (GBM)	6.0	33.5	99.5	96.0	91.6	93.2	99.5	99.0

Tab. X *Compare inter-patient results for* DS2 *(In percent).*

5. Conclusion

In this paper, we proposed a novel approach to the task of arrhythmia classification for patient-specific and inter-patient paradigms. We measured the performance of our approach on the MIT-BIH database and also compare the results with other state-of-the-art methods. Concretely, the denoised and normalized ECG signals we transformed to reduced feature space, and we show the performance of trained NN and GBM models in the reduced space. For the purpose of comparison, results obtained by our approach with other works, we used five classes defined by the AAMI standard for patient-adapted paradigm and three classes for inter-patient paradigm. We compared our approach to state-of-the-art methods based on statistical and machine learning methods. The results show that our proposed method outperforms state-of-the-art methods for VEB heartbeats in both paradigms, and outperforms the prediction of the SVEB heartbeat in case of the patient-adapted paradigm. The focus of the future work will be increasing the accuracy of Supraventricular Ectopic heartbeats, mainly in the context of the inter-patient paradigm.

References

[1] AAMI standard (ANSI/AAMI. Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms. American National Standards Institute, Inc. (ANSI), Association for the Advancement of Medical Instrumentation (AAMI). ANSI/AAMI/ISO EC57, 1998-(R)2008 (2008).)

[2] AL RAHHAL M.M., BAZI Y., ALHICHRI H., ALAJLAN N., MELGANI F., YAGER R.R. Deep learning approach for active classification of electrocardiogram signals. *Information Sciences*. 2016, 345, pp. 340–354.

[3] BAZI Y., ALAJLAN N., ALHICHRI H., MALEK S. Domain adaptation methods for ecg classification. In: *Computer Medical Applications (ICCMA), 2013 International Conference on* (2013), IEEE, pp. 1–4.

[4] BENGIO Y., ET AL. Learning deep architectures for ai. *Foundations and trends® in Machine Learning.* 2009, 2(1), pp. 1–127.

[5] BESTERMAN E., CREESE R. Waller–pioneer of electrocardiography. *British Heart Journal.* 1979, 42(1), p. 61.

[6] CASTELLS F., LAGUNA P., SÖRNMO L., BOLLMANN A., ROIG J.M. Principal component analysis in ecg signal processing. *EURASIP Journal on Advances in Signal Processing.* 2007, 2007 (1), p. 074580.

[7] DE CHAZAL P., O'DWYER M., REILLY R.B. Automatic classification of heartbeats using ecg morphology and heartbeat interval features. *IEEE transactions on biomedical engineering.* 2004, 51(7), pp. 1196–1206.

[8] DE CHAZAL P., REILLY R.B. A patient-adapting heartbeat classifier using ecg morphology and heartbeat interval features. *IEEE Transactions on Biomedical Engineering.* 2006, 53(12), pp. 2535–2543.

[9] DE LANNOY G., FRANÇOIS D., DELBEKE J., VERLEYSEN M. Weighted svms and feature relevance assessment in supervised heart beat classification. In: *International Joint Conference on Biomedical Engineering Systems and Technologies* (2010), Springer, pp. 212–223.

[10] DEL TESTA D., ROSSI M. Lightweight lossy compression of biometric patterns via denoising autoencoders. *IEEE Signal Processing Letters.* 2015, 22(12), pp. 2304–2308.

[11] DOQUIRE G., DE LANNOY G., FRANÇOIS D., VERLEYSEN M. Feature selection for interpatient supervised heart beat classification. *Computational intelligence and neuroscience.* 2011, 2011, p. 1.

[12] GARCIA G., MOREIRA G., LUZ E., MENOTTI D. Improving automatic cardiac arrhythmia classification: Joining temporal-vcg, complex networks and svm classifier. In: *Neural Networks (IJCNN), 2016 International Joint Conference on* (2016), IEEE, pp. 3896–3900.

[13] GARCIA G., MOREIRA G., MENOTTI D., LUZ E. Inter-patient ecg heartbeat classification with temporal vcg optimized by pso. *Scientific Reports.* 2017, 7(1), p. 10543.

[14] GHONGADE R., ET AL. A brief performance evaluation of ecg feature extraction techniques for artificial neural network based classification. In: *TENCON 2007-2007 IEEE Region 10 Conference* (2007), IEEE, pp. 1–4.

[15] GÜLER İ., ÜBEYLI E. D. Ecg beat classifier designed by combined neural network model. *Pattern recognition.* 2005, 38(2), pp. 199–208.

[16] HU Y.H., PALREDDY S., TOMPKINS W.J. A patient-adaptable ecg beat classifier using a mixture of experts approach. *IEEE transactions on biomedical engineering.* 1997, 44(9), pp. 891–900.

[17] HUANG H., LIU J., ZHU Q., WANG R., HU G. A new hierarchical method for inter-patient heartbeat classification using random projections and rr intervals. *Biomedical engineering online.* 2014, 13(1), p. 90.

[18] INCE T., KIRANYAZ S., GABBOUJ M. A generic and robust system for automated patient-specific classification of ecg signals. *IEEE Transactions on Biomedical Engineering.* 2009, 56(5), pp. 1415–1426.

[19] JIANG W., KONG S.G. Block-based neural networks for personalized ecg signal classification. *IEEE Transactions on Neural Networks.* 2007, 18(6), pp. 1750–1761.

[20] KIM C.H., AGGARWAL R. Wavelet transforms in power systems. part 1: General introduction to the wavelet transforms. *Power Engineering Journal.* 2000, 14(2), pp. 81–87.

[21] KIM J., SHIN H.S., SHIN K., LEE M. Robust algorithm for arrhythmia classification in ecg using extreme learning machine. *Biomedical engineering online.* 2009, 8(1), p. 31.

[22] KORÜREK M., DOĞAN B. Ecg beat classification using particle swarm optimization and radial basis function neural network. *Expert systems with Applications.* 2010, 37(12), pp. 7563–7569.

[23] KUMAR R.G., KUMARASWAMY Y. Investigation and classification of ecg beat using input output additional weighted feed forward neural network. In: *Signal Processing Image Processing & Pattern Recognition (ICSIPR), 2013 International Conference on* (2013), IEEE, pp. 200–205.

[24] KUTLU Y., KUNTALP D. Feature extraction for ecg heartbeats using higher order statistics of wpd coefficients. *Computer methods and programs in biomedicine.* 2012, 105(3), pp. 257–267.

[25] LI C., ZHENG C., TAI C. Detection of ecg characteristic points using wavelet transforms. *IEEE Transactions on biomedical Engineering.* 1995, 42(1), pp. 21–28.

[26] LIN C.-H., DU Y.-C., CHEN T. Adaptive wavelet network for multiple cardiac arrhythmias recognition. *Expert Systems with Applications.* 2008, 34(4), pp. 2601–2611.

[27] LIN C.-H., DU Y.-C., CHEN T. Nonlinear interpolation fractal classifier for multiple cardiac arrhythmias recognition. *Chaos, Solitons & Fractals.* 2009, 42(4), pp. 2570–2581.

[28] LLAMEDO M., MARTÍNEZ J.P. Heartbeat classification using feature selection driven by database generalization criteria. *IEEE Transactions on Biomedical Engineering.* 2011, 58(3), pp. 616–625.

[29] LUZ E., MENOTTI D. How the choice of samples for building arrhythmia classifiers impact their performances. In: *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE* (2011), IEEE, pp. 4988–4991.

[30] LUZ E.J.D. S., SCHWARTZ W.R., CÁMARA-CHÁVEZ G., MENOTTI D. Ecg-based heartbeat classification for arrhythmia detection: A survey. *Computer methods and programs in biomedicine.* 2016, 127, pp. 144–164.

[31] MAHESH V., KANDASWAMY A., VIMAL C., SATHISH B. Ecg arrhythmia classification based on logistic model tree. *Journal of Biomedical Science and Engineering.* 2009, 2(06), p. 405.

[32] MOODY G.B., MARK R.G. The mit-bih arrhythmia database on cd-rom and software for use with it. In: *Computers in Cardiology 1990, Proceedings.* (1990), IEEE, pp. 185–188.

[33] PETEREK T., ZAORÁLEK L., DOHNÁLEK P., GAJDOS P. Recognition of pathological beats in ecg signals based on singular value decomposition of wavelet coefficients and support vector machine. In: *Telecommunications and Signal Processing (TSP), 2015 38th International Conference on* (2015), IEEE, pp. 1–5.

[34] ROY U.D., GHORAI S., MUKHERJEE A. Kernel-based feature extraction for patient-adaptive ecg beat classification. In: *Systems in Medicine and Biology (ICSMB), 2016 International Conference on* (2016), IEEE, pp. 144–147.

[35] SAYADI O., SHAMSOLLAHI M.B. Ecg denoising and compression using a modified extended kalman filter structure. *IEEE Transactions on Biomedical Engineering.* 2008, 55(9), pp. 2240–2248.

[36] SPACH M. S., KOOTSEY J. M. The nature of electrical propagation in cardiac muscle American Journal of Physiology-Heart and Circulatory Physiology. 1983, 244, pp. H3–H22.

[37] ÜBEYLI E.D. Combining recurrent neural networks with eigenvector methods for classification of ecg beats. *Digital Signal Processing.* 2009, 19(2), pp. 320–329.

[38] XIONG P., WANG H., LIU M., ZHOU S., HOU Z., LIU X. Ecg signal enhancement based on improved denoising auto-encoder. *Engineering Applications of Artificial Intelligence.* 2016, 52, pp. 194–202.

[39] YE C., KUMAR B.V., COIMBRA M.T. Heartbeat classification using morphological and dynamic features of ecg signals. *IEEE Transactions on Biomedical Engineering.* 2012, 59(10), pp. 2930–2941.

[40] YU S.-N., CHOU K.-T. Selection of significant independent components for ecg beat classification. *Expert Systems with Applications.* 2009, 36(2), pp. 2088–2096.