IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received February 19, 2017, accepted March 26, 2017, date of publication March 31, 2017, date of current version May 17, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2690380

Suppression of Intensive Care Unit False Alarms Based on the Arterial Blood Pressure Signal

PETRE LAMESKI[1], (Student Member, IEEE), EFTIM ZDRAVEVSKI[1], SASO KOCESKI[2], ANDREA KULAKOV[1], (Member, IEEE), AND VLADIMIR TRAJKOVIK[1]
[1]Faculty of Computer Science and Engineering, Ss. Cyril and Methodius University in Skopje, 1000 Skopje, Macedonia
[2]Faculty of Informatics, Goce Delčev University in Štip, 2000 Štip, Macedonia

Corresponding author: P. Lameski (lameski@gmail.com)

This work was supported in part by the Faculty of Computer Science and Engineering, Ss. Cyril and Methodius University in Skopje, Macedonia, and in part by the Goce Delčev University in Štip, Macedonia, under Grant 0307-252/24.

ABSTRACT Patient monitoring in intensive care units requires collection and processing of high volumes of data. High sensitivity of sensors leads to significant number of false alarms, which cause alarm fatigue. Reduction of false alarms can lead to better reaction time of medical personnel. This paper aims to develop a method for false alarm suppression and evaluate it on a publicly available data set with manually annotated alarms. First, an automated feature engineering was performed using the signal for arterial blood pressure (ABP) and a processed signal that contained the times of each heartbeat from the ABP signal. Next, support vector machines, random forest, and extreme random trees classifiers were trained to create classification models. The best suppression performance was achieved for the extreme tachycardia alarm, for which 90.3% of the false alarms were suppressed, while only 0.54% of the true alarms were incorrectly suppressed. This paper demonstrates that alarm suppression can be achieved with high accuracy using an automated feature engineering coupled with machine learning algorithms. The proposed approach can be utilized as aid to medical personnel and experts, allowing them to be more productive and to respond to alarms in a more timely manner.

INDEX TERMS Machine learning, medical expert systems, signal processing, patient monitoring, time series analysis, pattern recognition, data preprocessing.

I. INTRODUCTION

The Intensive Care Units (ICU) monitors, record and monitor patients parameters. There are many available commercial data acquisition systems [1] with very sensitive sensors needed for accurate identification of 100% of the dangerous levels of some parameters. Nonetheless, these systems currently have limited ability to extract and process the raw data. As a result, the decision making during diagnostics significantly relies on the acquired raw data. However, because of the high sensitivity of sensors, a significant number of false alarms occur. According to [2], 72% to 99% of the alarms are false, which in turn leads to alarm fatigue for the medical personnel. The increased number of alarms reduces the reaction speed of medical personnel, who sometimes even consciously neglect the alarms.

To aid the research related to the suppression of false alarms from ICU monitors, a subset from MIMIC II dataset [1], [3] was processed and annotated. This represents the "gold standard" dataset where the critical alarms obtained from the ICU monitors were validated and the false alarms were marked [4]. The subset was selected based on two criteria. First, a critical ECG arrhythmia alarm needed to be issued. Second, ABP waveform and at least one ECG channel needed to be available at the time of the alarm. Initial results for the false alarms suppression were also reported. Five types of alarms were annotated: Asystole, Extreme tachycardia, Extreme bradycardia, Ventricular tachycardia and Ventricular fibrillation.

The scientific community has been trying to attain a general method for data fusion and a general model for decision support based on large volumes of medical data, especially from data obtained from ICUs [5]. One of the areas where such methods and models are needed is the false alarm suppression from the ICU.

One of the approaches that can reduce the false alarms is by applying signal processing and data mining techniques on the raw ICU measurements. According to [6], data mining methods can be used for false alarm detection. Furthermore, this study empirically demonstrated that the false alarm rate can be reduced.

Another approach used game theory to select the most important features from wavelet transformation of the input time series signals [7]. The evaluation was performed with Bayes Net classifier trained on the 30 best features selected by the proposed approach.

In [8], authors use several signal quality indices (SQI) obtained from the ECG segments before the alarm [8]. An average accuracy of 94.6% for false alarm prediction using SVM classifiers trained on the SQIs on 4050 alarms from the MIMIC II database was reported in this study.

The significance of the automated alarm suppression has motivated the community to organize ''Physionet: Computing in cardiology challenge 2015'' [9]. Competitors implemented many existing and proposed novel approaches for improving the suppression rate of false alarms.

All algorithms that are applied to suppress false alarms from ICU monitors must perform the processing relatively fast. An alarm must be raised no later than 10s from the event that caused the alarm [10]. This is not always a trivial task because the volumes of data that need to be processed require considerable computational time. This is why features need to be reduced before making predictions or classifications. It would ensure that the processing is done faster and that the results from the processing are more stable and robust.

In this paper we present a novel framework that incorporates automatic feature engineering and selection, as well as machine learning algorithms that can automatically tune for false alarm detection and suppression. The main scientific contribution of our approach is that the feature engineering from the data is automated and does not require any manual work by a domain expert. Moreover, the generated features and used machine learning algorithm are computationally efficient, so they can be used in real time even on devices with limited hardware resources.

II. METHODS AND ANALYSIS

The standard approach for applying machine learning to the problem of false alarm detection was performed. First we used statistical and other types of transformations of the time series from the dataset to generate the features. After that, we used feature selection process to improve the classification accuracy, to reduce overfitting and to speed up the machine learning algorithms [11]. Finally, when the feature selection process finished, we applied several machine learning algorithms to generate the classification models. The dataset, the feature extraction and selection process and the used machine learning algorithms are described in more detail in this section.

A. DATASET

The annotated alarms from the MIMIC II waveform database[1] were used in this study. The manual annotation of alarms means that an expert manually analyzed the readings prior each alarm and labeled them as true or false. The patients with annotated alarms, according to [4], are adult patients with ages ranging from 18 to more than 90 years with mean of 68.3 years. The age and the health state of the patients was not taken into consideration in our experiments because new data is added to the database continuously, and at the time of the experiments, the age and health data was not available for many patients. The alarms that contained an ABP signal were considered, and for each of the alarms, 6 time windows with length of 17 seconds were selected per the recommendations in [10]. This window length was selected so more than 10 seconds before the alarm are available and to provide enough time for processing after the alarm occurs. The time windows started 11 to 16 seconds prior to the alarm with a 1 second shift between consecutive windows (i.e. 16 seconds overlap between adjacent windows). This guarantees that the response time will be within 10 seconds after the alarm, which is enough to notify the personnel in a timely manner. At the same time, this provides enough time for the alarm signal to be processed by the proposed system. For each of these windows, the open source heartbeat detector [12] was used for extracting the beats.

The input data for the feature generation and selection consisted of the ABP signal measured with 125Hz frequency and the Heartbeats signal that is consisted of zeros and ones (one for the moment where the beat starts, zero for the rest of the beat). For each of the time windows, for the alarm there are two time series with length of 2125 measurements, which correspond to 17 seconds. These two arrays were generated for each annotated alarm that contained a valid ABP signal during the time intervals of the measurements.

TABLE 1. Distribution of alarm types in the dataset.

Alarm type	True alarm	False alarm	Total
ASYSTOLE	378	4992	5370
BRADY	3414	2028	5442
TACHY	12659	3588	16247
V-FIB/TACH	780	1703	2483
V-TACH	7902	7734	15636
All	25133	20045	45178

From all annotated alarms, 7531 contained a valid ABP signal. The total distribution of alarms per alarm type and label (true or false alarm) is presented in Table 1. Having at most 6 ABP and heartbeats time series for each alarm, there are in total 45178 instances in the dataset that is used for modeling and evaluation of the proposed system. As it can be observed, for most of the alarm types, the dataset is not balanced. For each alarm type, there are more false alarms than true alarms labelled in the data, which is consistent with the real life conditions where most of the alarms are false. This, however, introduces additional challenge towards building a robust alarm suppression model, because most of the machine learning approaches require balanced datasets where the labels are uniformly distributed (equal number of true and false alarms).

[1] MIMIC II waveform database, https://www.physionet.org/pn5/mimic2db

B. FEATURE ENGINEERING

Due to the density of the available data collected by the ICU unit and the nature of the time series data, one of the most important tasks that needs to be performed when analyzing it, is data representation by feature extraction. Additional challenge for building robust features is dealing with drift in the data over time, which can be the result of: data generated by diverse sensors, difference in distribution of patient demographics, data collected by different people, diversity of sensor manufactures, amortization of sensors over time, etc. Ideally, such variations would have little to no effect on the trained models.

Fig. 1 represents the algorithm that combines feature extraction and selection. Initially the system generates new time series (TS) in the Time Series Generation block. Then, it extracts various features from each time series in the Feature Engineering block. The dataset is processed one instance at a time, where an instance refers to a sliding window segment of 17 seconds. After that, feature ranking and selection is performed, which identifies which feature extraction transformations are useful. First, a fast feature selection identifies and discards individually non-informative features. Then it analyzes which time series are valuable in relation to the target variable or class. Random Forest classifier [13], [14] with high number of features is able to estimate the informativeness of individual features and of various subsets of features, while estimating of how predictive the whole problem is.

The most cumbersome data source in all health care systems are the sensors which generate time series (TS) data. Generating features from them is a challenging task, but using a systematic approach, we could generate a variety of features that can robustly describe a dataset. A recent data mining competition for posture recognition of firefighters [15] inspired different feature engineering (FE) approaches that are very effective. Based on those ideas, new time series were generated, hoping that the features extracted from them will potentially improve the robustness of the system and enhance the predictive performance.

- Frequency domain features obtained with Fast Fourier Transformation (FFT). Spectral centroid is one value, but additionally we obtain new series with amplitudes, frequencies and magnitudes.
- Delta series which calculate the relative deviations of the average value of original readings.
- Sliding window time series. From a time series with n samples we generate several smaller time series with sliding windows that might overlap. The idea behind this approach is that we want to identify if some parts of the time series are more informative than others (e.g. the most recent values).

Then, from each original and newly generated time series the algorithm generates the following types of features, where n is the number of values in the corresponding sliding window segment:

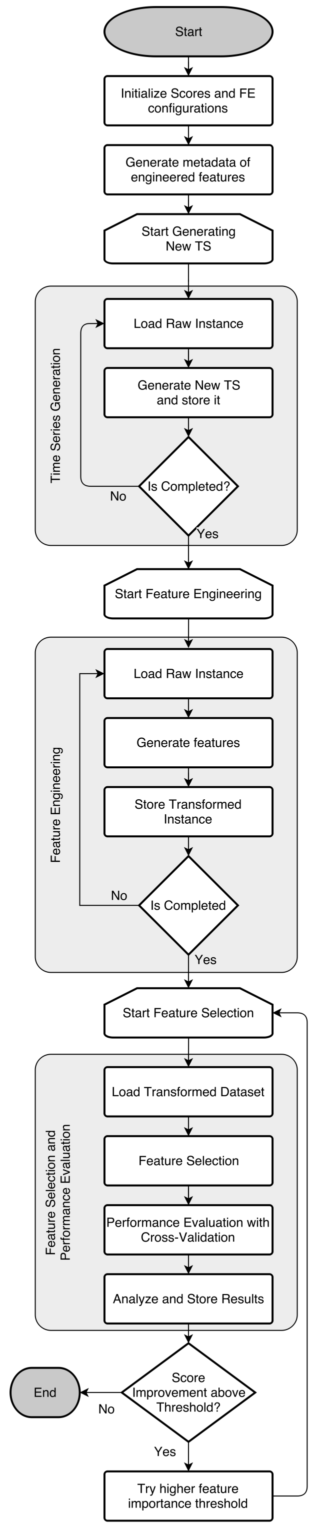


FIGURE 1. Algorithm for feature engineering, feature selection and evaluation.

- Basic statistics (i.e. minimum, maximum, range, arithmetic mean, harmonic mean, geometric mean, standard deviation, variance, skewness, kurtosis, signal-to-noise ratio, energy, average energy per sample, and mode.)

- Linear and quadratic curve fitting parameters.
- Equal-width histogram with number of intervals calculated as $\lceil log_2 n + 1 \rceil$, where n is number of values in the interval, based on the Struges rule [16].
- Percentile based features: first quartile, median, third quartile, inter-quartile range, amplitude and some other percentiles.
- Auto-correlation of the signal (classical auto-correlation, Pearson, Spearman and Kendall correlation).
- Inter-correlations between each pair of raw time series values using the aforementioned types of correlation coefficients. The obtained values are useful for predictive purposes as features, but they are equally important in the process of understanding the associations between the various quantities in the system.

The purpose of the variety of time series and features is describe the signals in a variety of ways and to have a rich feature set that will be later filtered during the feature selection phase. Identifying predictive features is important and can lead to design of specialized algorithms that can be embedded in ICU units to suppress false alarms. In a production system, only the essential features for optimal alarm suppression rate would be computed.

C. CLASSIFICATION MODEL

After the process of feature selection, we experimented with three classification algorithms for generating the classification model. The first algorithm is Random Forest [13], [14]. Random forest (RF) is an efficient algorithm that generates multiple decision trees [17] by randomly sampling training instances from the dataset and also randomly selecting m features from each sample, where M is the total number of features per sample and $m < M$. The sampling is random but consistent while growing a single tree. The multiple decision trees are trained on the training data independently. The tree branching is performed by finding the best split from the m features on each node. In the process of classification each tree votes for the class and the majority class is chosen.

The Extremely Randomized Trees (ERT) algorithm [18] similar to RF, also generates an ensemble of trees, however, unlike in RF, ERT algorithm chooses the split from the attributes randomly. This increases the training speed because the number of calculations per node is decreased. Both trees give excellent results for classification and are able to train models on very large datasets very fast.

We also used the Support Vector Machine (SVM) classifier to generate another model for false alarm detection. We used SVM classifier with Gaussian kernel that needs optimization of two parameters: the cost value and the gamma value. We calculated the best value for cost and gamma by using the grid search method [19].

We generated classification models using the RF, ERT and SVM algorithms and we performed 10-fold regular and stratified cross-validation (CV) for evaluating the classification performance. In stratified 10-fold CV, the training and test subset keep the class distribution the same, and in regular CV, the 10-fold split is random, so the class distributions for the training and the test set may not be equal in general. For each alarm type we trained and evaluated a separate classification model.

TABLE 2. Top 10 most informative features average for all alarm types.

Series transformation	Value transformation
delta add ABP	quantile 60
ABP	spectral centroid
delta add ABP	hist range [2.5761448;35.9510081)
delta add ABP	median
sliding window [1062-2125) ABP	quantile 95
sliding window [1062-2125) ABP	auto-correlation Pearson t=64
sliding window [1062-2125) ABP	mode
sliding window [1062-2125) ABP	SNR
delta add ABP	auto-correlation Pearson t=64
sliding window [0-1062) ABP	SNR

III. RESULTS AND DISCUSSION

The framework we propose also evaluates the informativeness of the generated features. The best results presented below were obtained by training each classifier on the top 152 features from the total of 3035 generated features. The results differed in the fourth decimal when the classification was performed using 379, 304, 228 and 152 features. The 152 features were selected as the smallest number that did not cause any significant degradation in the accuracy of the classification. Table 2 lists the top 10 features that were found to be most informative for false alarm classification on average for all alarm types.

For each alarm type we generated a separate classification model using the three classification algorithms described in subsection II-C.

When evaluating the classification accuracy with stratified and regular 10-fold CV we obtained very similar results. The best average accuracy over all alarm types was 94.04% for both CV schemes. The results were slightly better or comparable to the ones reported in [6]. The sensitivity and specificity obtained when using 10-fold stratified split for each alarm type and for each classifier used are presented in Table 3. The reported sensitivity was calculated as:

$$\frac{PositiveTrueAlarms}{(PositiveTrueAlarms + NegativeFalseAlarms)}$$

And the specificity is calculated as:

$$\frac{PositiveFalseAlarms}{(PositiveFalseAlarms + NegativeTrueAlarms)}$$

The table contains the results for each alarm type, where ASYSTOLE stands for the Asystole alarm, TACHY for the Extreme tachycardia alarm, BRADY for the Extreme bradycardia alarm, V-TACH for the Ventricular tachycardia alarm and V-FIB/TACH for the Ventricular fibrillation alarm.

The percentage of suppressed false and true alarms by the proposed approach for each alarm type is shown in Fig. 2. For the ASYSTOLE, BRADY, V-TACH and V-FIB/TACH we could achieve from 95.5% to 99.9% false alarm suppression,

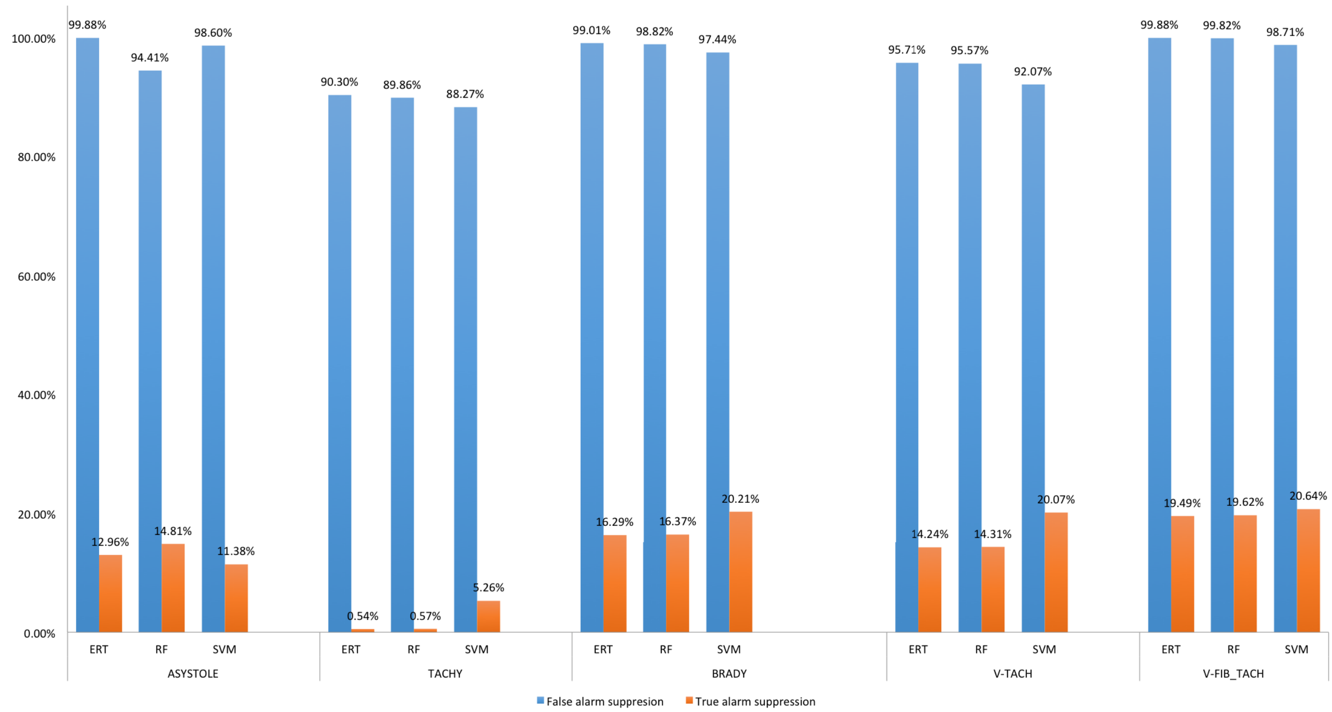


FIGURE 2. False alarm suppression and True alarm suppression for each alarm type and each classification algorithm.

TABLE 3. Sensitivity and Specificity for 10-fold stratified cross-validation.

	ERT		RF		SVM	
	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
ASYSTOLE	0.8624	0.9988	0.8439	0.9982	0.8862	0.9860
BRADY	0.8371	0.9882	0.8360	0.9877	0.7979	0.9758
TACHY	0.9946	0.9030	0.9943	0.8986	0.9474	0.8827
V-FIB/TACH	0.8051	0.9988	0.8038	0.9982	0.7936	0.9871
V-TACH	0.8570	0.9578	0.8569	0.9557	0.7993	0.9207

albeit with unacceptably high true alarm suppression of up to 20%. Nonetheless, for the TACHY alarm a false alarm suppression of 90% was obtained while experiencing a very low true alarm suppression of about 0.5%. This performance is very encouraging and could be used an inspiration for tuning models for other alarm types by modifying the decision threshold so the true alarm suppression is improved, at the cost of reducing the false alarm suppression too.

From Fig. 3, it is evident that the training time needed to generate the model for ERT and RF was significantly shorter than the time needed to generate the SVM model. Note that for SVM the presented time does not include the time needed to perform parameter tuning.

The main advantage of the proposed approach is that the process of feature extraction and selection is automated. The selection of most informative features reduced the size of the feature vector and the time needed to train the model. It also provided insight to the nature of the signals. For instance, initially we considered using other signals besides the ABP signal, however the framework discovered that too many missing values causing all features derived from the other signals to be with low informativeness. Additionally it identified the transformations that produced informative features, such as the ones listed in Table 2. Apparently, the second half of the whole window is more informative than the first half, which is intuitively expected as it more recent relative to the occurrence of the alarm. Furthermore, we discovered that using the ABP signal directly is less informative than using the Delta series (which calculated the relative deviations from the ABP signal). A similar finding is that features based on histograms and percentiles were more informative than basic statistics, such as mean, standard deviation, minimum, maximum, etc. Another peculiar discovery was that the most informative feature originating from the heartbeats time series was ranked 184-th out of the 972 features that were generated. In fact, during the feature selection all heartbeat features were discarded, meaning that manual extraction of the heartbeats signal from the ABP signal is not required because the feature engineering already is able to generate more informative features.

Using a separate classification model for each alarm type allowed the framework to identify features which are more

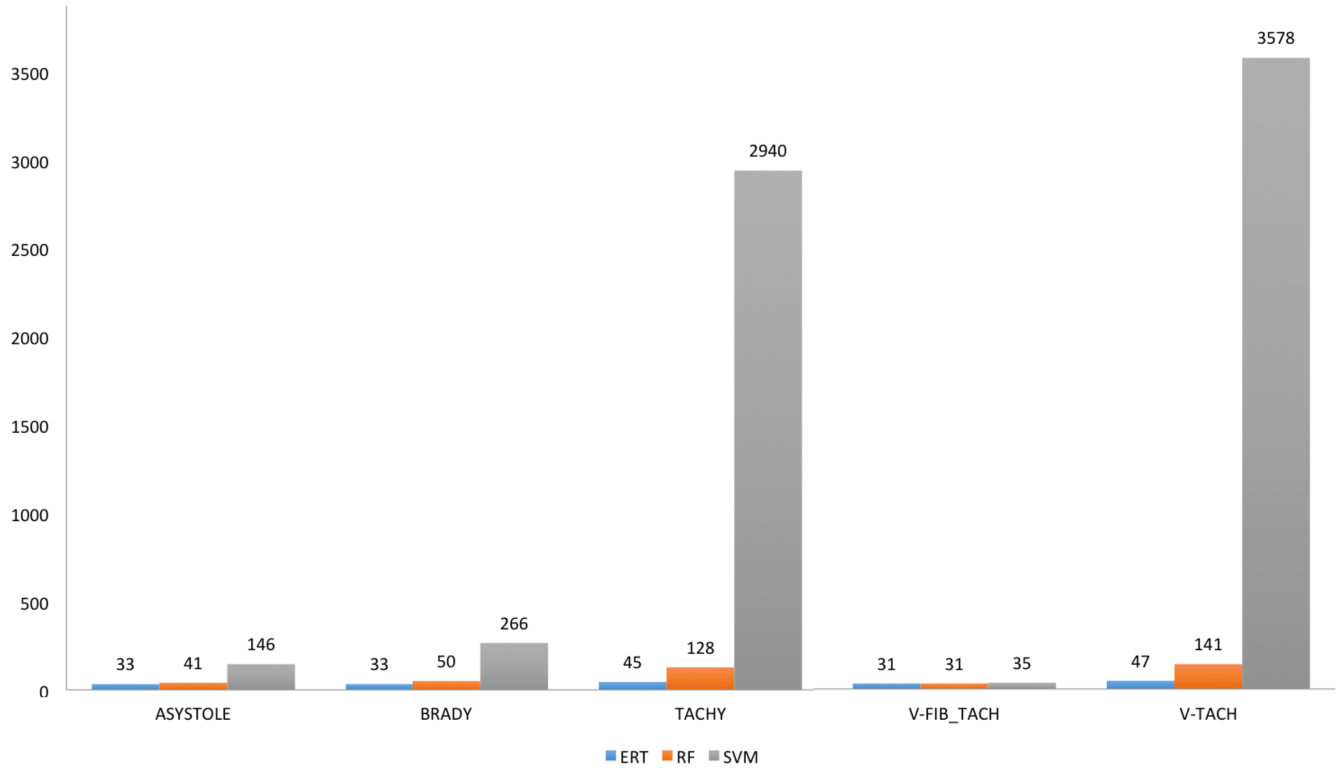


FIGURE 3. Classification time [s] for the best classifiers per alarm type (RF=Random Forest, ERT=Extremely Randomized Trees, SVM=Support Vector Machines).

relevant for particular alarms. This is a very convenient functionality because it can be used for defining cut-off values for different parameters and alarms. One feature could be very relevant for the detection of one alarm type, but not so informative for another. For example, the first feature in Table 2 is 10 times more important for the detection of the V-FIB/TACH alarm than for the BRADY alarm.

This valuable feedback provided by the system can be manually observed by an expert and some specifics of the nature of the alarms can be more easily discovered. Additionally, is intrinsically used in the system for selecting only the informative features that are both predictive and robust for each alarm type separately.

We compared three classification algorithms for their ability to accurately classify the false alarms. Although all the algorithms showed similar performance in the alarm suppression rates, the suppression rate of SVMs was slightly worse than of the ERT and the RF classifiers, except for the ASYSTOLE alarm where it showed lowest rate of true alarm suppression. With the used dataset, the ERT classifier had the overall best performance in both alarm suppression performance and in training time. In terms of suppression of false alarms, the best results were obtained for the suppression of TACHY alarm where the percentage of suppressed true alarms was the lowest 0.54% and the percentage of suppressed false alarms was 90.3%. Further improvements are needed, however, because the percentage of suppressed true alarms needs to be 0% in order to make the approach useful for practical usage in hospitals. Any suppression of true alarms could be fatal for the patients in the ICU.

In [20] and [21] authors report better results in terms of true alarm suppression rates and worse results in term of false alarm suppression rate, however they use custom tailored approaches for each alarm type or an ensemble of custom tailored classifiers. On the other hand, we use a general framework that does not require any prior knowledge of the nature of the signals that are processed. With additional tuning, our approach allows usage of any labeled time series data and building robust models without prior understanding of the nature of the problem at hand.

Based on the obtained results we can conclude that our approach is highly accurate for false alarm suppression based on the ABP signal and the heartbeats signal. To increase the performance in future work, we plan to include other types of available signals such as ECG signals. Additional tuning of the algorithm will be required to allow penalizing of true alarm suppression since in real live system that would be used for alarm suppression there should be zero tolerance for true alarms suppression. Our best result is very close to this level of performance, because it suppresses over 90% of the false alarms while the true alarm suppression rate stays very low.

The signal quality index (SQI), proposed in [22], can be used to assess the quality of the signal. When the system is deployed in a production environment, SQI can be used as an initial validation step. Namely, if the quality of the signal is poor, the usage of the proposed classification model would be disabled in order to prevent incorrect classifications. On the other hand, if the quality is sufficient, the proposed system can be further used to process the signal by extracting features of it and making classifications. Given that the analyzed dataset is manually validated, the SQI index had no effect on it.

Additional patient data could be beneficial for the quality of a general model for false alarm suppression. The current performance estimation with 10-fold regular and stratified cross-validation provides insights on how the proposed framework will perform when analyzing time series and predicting if the alarm is false or true. It is likely that since the ABP signal differs for different ages and different cases, the proposed approach would decrease its accuracy for patient groups that were not included in the training data. To alleviate this, the framework would benefit from additional nominal features, such as age, gender, etc. Such features could stratify the patients into more uniform groups, so the performance does not degrade. Unfortunately only a very small subset of patients, for which we had valid labeled alarms and ABP signals, had available data about gender and age.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interest regarding the publication of this paper.

REFERENCES

[1] M. A. De Georgia, F. Kaffashi, F. J. Jacono, and K. A. Loparo, "Information technology in critical care: Review of monitoring and data acquisition systems for patient care and research," *Sci. World J.*, vol. 2015, 2015, Art. no. 727694.

[2] S. Sendelbach and M. Funk, "Alarm fatigue: A patient safety concern," *AACN Adv. Critical Care*, vol. 24, no. 4, pp. 378–386, 2013.

[3] G. B. Moody and R. G. Mark, "A database to support development and evaluation of intelligent intensive care monitoring," in *Proc. Comput. Cardiol.*, Sep. 1996, pp. 657–660.

[4] A. Aboukhalil, L. Nielsen, M. Saeed, R. G. Mark, and G. D. Clifford, "Reducing false alarm rates for critical arrhythmias using the arterial blood pressure waveform," *J. Biomed. Informat.*, vol. 41, no. 3, pp. 442–451, Jun. 2008.

[5] G. D. Clifford, W. J. Long, G. B Moody, and P. Szolovits, "Robust parameter extraction for decision support using multimodal intensive care data," *Philos. Trans. Roy. Soc. London A, Math. Phys. Sci.*, vol. 367, no. 1887, pp. 411–429, 2009.

[6] B. Baumgartner, K. Rödel, and A. Knoll, "A data mining approach to reduce the false alarm rate of patient monitors," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2012, pp. 5935–5938.

[7] F. Afghah, A. Razi, S. M. R. Soroushmehr, S. Molaei, H. Ghanbari, and K. Najarian, "A game theoretic predictive modeling approach to reduction of false alarm," in *Smart Health* (Lecture Notes in Computer Science), vol 9545, X. Zheng, D. Zeng, H. Chen, and S. Leischow, Eds. Cham, Switzerland: Springer, 2016.

[8] J. Behar, J. Oster, Q. Li, and G. D. Clifford, "ECG signal quality during arrhythmia and its application to false alarm reduction," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 6, pp. 1660–1666, Jun. 2013.

[9] G. D. Clifford *et al.*, "The PhysioNet/computing in cardiology challenge 2015: Reducing false arrhythmia alarms in the ICU," in *Proc. Comput. Cardiol. Conf. (CinC)*, 2015, pp. 273–276.

[10] A. for the Advancement of Medical Instrumentation *et al.*, "Cardiac monitors, heart rate meters, and alarms [American national standard]," Amer. Nat. Standard, Arlington, VA, USA, Tech. Rep. ANSI/AAMI EC13: 2002, 2002.

[11] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, Jan. 2003.

[12] W. Zong, T. Heldt, G. B. Moody, and R. G. Mark, "An open-source algorithm to detect onset of arterial blood pressure pulses," in *Proc. Comput. Cardi.*, Sep. 2003, pp. 259–262.

[13] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[14] A. Liaw and M. Wiener, "Classification and regression by randomforest," *R News*, vol. 2, no. 3, pp. 18–22, 2002.

[15] M. Meina, A. Janusz, K. Rykaczewski, D. Ślęzak, B. Celmer, and A. Krasuski, "Tagging firefighter activities at the emergency scene: Summary of AAIA'15 data mining competition at knowledge pit," in *Proc. Federated Conf. Comput. Sci. Inf. Syst. (FedCSIS)*, Sep. 2015, pp. 367–373.

[16] H. A. Sturges, "The choice of a class interval," *J. Amer. Statist. Assoc.*, vol. 21, no. 153, pp. 65–66, 1926.

[17] J. R. Quinlan, "Induction of decision trees," *Mach. Learn.*, vol. 1, no. 1, pp. 81–106, 1986.

[18] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Mach. Learn.*, vol. 63, no. 1, pp. 3–42, 2006.

[19] C. Staelin, "Parameter selection for support vector machines," Hewlett-Packard Company, Palo Alto, CA, USA, Tech. Rep. HPL-2002-354R1, 2003.

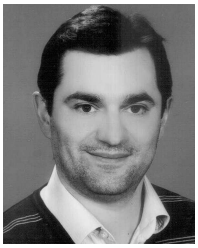
[20] R. Rodrigues and P. Couto, "Detection of false arrhythmia alarms with emphasis on ventricular tachycardia," *Physiol. Meas.*, vol. 37, no. 8, p. 1326, 2016. [Online]. Available: http://stacks.iop.org/0967-3334/37/i=8/a=1326

[21] S. Fallet, S. Yazdani, and J.-M. Vesin, "False arrhythmia alarms reduction in the intensive care unit: A multimodal approach," *Physiol. Meas.*, vol. 37, no. 8, p. 1217, 2016. [Online]. Available: http://stacks.iop.org/0967-3334/37/i=8/a=1217

[22] W. Zong, G. B. Moody, and R. G. Mark, "Reduction of false arterial blood pressure alarms using signal quality assessement and relationships between the electrocardiogram and arterial blood pressure," *Med. Biol. Eng. Comput.*, vol. 42, no. 5, pp. 698–706, 2004.

PETRE LAMESKI (S'08) was born in Kavadarci in 1985. He received the degree (Hons.) from the Faculty of Electrical Engineering and Information Technologies, Ss. Cyril and Methodius University in Skopje, in 2008, and the master's degree (Hons.) from the Faculty of Electrical Engineering and Information Technologies, Ss. Cyril and Methodius University in Skopje, in 2010, with a thesis in robotics. From 2008 to 2011, he was an Assistant with the Faculty of Electrical Engineering and Information Technologies, Ss. Cyril and Methodius University in Skopje. He is currently a Teaching and Research Assistant with the Faculty of Computer Science and Engineering, Skopje, Macedonia. He has participated in several national and international projects and has authored over 30 conference and journal papers. His research interests include intelligent robotics and computer systems, machine intelligence, machine vision, big data, decision support systems, cognitive systems and cognitive robotics, time series data analysis, wireless sensor, and actor networks.

EFTIM ZDRAVEVSKI received the master's degree in data mining from the Faculty of Electrical Engineering and Information Technologies, Ss. Cyril and Methodius University in Skopje, Macedonia, in 2010. During his studies, he was continuously rewarded for exceptional performance. He is currently pursuing the Ph.D. degree with the Faculty of Computer Science and Engineering, Skopje. From 2008 to 2015, he was a Developer and a Researcher with NI TEKNA–Intelligent technologies, where he designed and implemented information systems and data warehouses, and involved in fraud detection and other machine learning related projects. Since 2010, he has been a Teaching and Research Assistant with the Faculty of Computer Science and Engineering. He has authored over 30 papers in international conferences and journals. His active fields of research interests are big data, machine learning, data mining, cloud computing, expert systems, knowledge-based systems, decision support systems, intelligent information systems, data bases and data warehouses, processing of sensory data, and parallel algorithms. He has been an MENSA Member since 2004.



SASO KOCESKI received the Ph.D. degree in robotics and artificial intelligence from the University of L'Aquila, Italy, in 2009. He is currently an Associate Professor with the Faculty of Computer Science, Goce Delčev University of Štip, Macedonia, and also the Head of the Institute of Computer Science. He has authored or co-authored over 100 refereed journal and conference papers, 15 book chapters, and three books. His current research interests are focused in the field of artificial intelligence, robotics, bioengineering, HCI, and medical imaging. He has been involved in many international and national scientific and applicative projects, including HORIZON 2020, FP7, EU-TEMPUS, EU-COST, and NATO Science for Peace, as a Project Coordinator and a Participant. He is a member of several professional societies. He serves as an Editor and a Reviewer in several SCI journals.

ANDREA KULAKOV received the Dipl.Eng. degree in 1995, the M.Sc. degree in cognitive science in 1998, and the Ph.D. degree in computer science in 2006. He is currently an Associate Professor with the Computer Science and Engineering Department, Ss. Cyril and Methodius University in Skopje. He is also a Professor at the following courses, such as Artificial Intelligence, Machine Vision, Introduction to Cognitive Science, and Introduction to Robotics at undergraduate level, and Advanced Topics in Artificial Intelligence and Sensor-Robotic Systems at postgraduate level. He has authored over 70 scientific papers in journals and conference proceedings. He participated and led several national and international projects. In 2005, for a part of his Ph.D. research, he received the Walter Karplus Summer Research Grant by the IEEE Computational Intelligence Society.

VLADIMIR TRAJKOVIK received the Ph.D. degree. He is currently a Full Professor with the Faculty of Computer Science and Engineering, Ss. Cyril and Methodius University in Skopje. His research interests include information systems analyses and design, distributed systems, educational systems, ICT-based collaboration systems, and mobile services with a special focus on educational and health-care systems.

He has published four books as an Author or an Editor, over ten chapters in different books published by international publishers, and over 150 journal and conference papers.

He has participated in more than 30 national and international ICT projects related to providing novel ICT-based services. He has coordinated over ten of those projects.

He has done management consultancy in five ICT companies. He has been a Mentor of five ICT companies in the last 15 years. He is a Co-Founder and a Partner in four ICT companies.

He has organized or participated as an Invited Lecturer, a Mentor, and a Judge on more than 20 events related to promotion of novel ICT services and Entrepreneurship in the last five years.