Deep Neural Network Approach for Artifact Detection in Raw ECG

Tanja Boljanić, Jovana Malešević, Goran Kvaščev

***Abstract*—Electrocardiography is a non-invasive technique for monitoring the electrical activity of the heart, and its analysis can detect and then prevent many health problems. Alterations that are not related to cardiac electrical activity represent artifacts in signal and should be minimized in order to correctly interpretate the signal. This is of great importance in wearable systems for electrophysiological monitoring that have numerous applications in healthcare and fitness. This paper presents how to build a classification model to detect artifacts in electrocardiogram (ECG) signal using deep neural network. The Long Short-Term Memory (LSTM) network was proposed for classifying 10-s single-channel ECG segments as *Valid* and *Artifact.* Data set consists of 10,231 raw ECG samples. The results show that the proposed method can classify the data with the accuracy of 90.1%, i.e., efficiently deal with acceptance of good (93.8%) and rejection of poor (80.1%) ECG quality.**

***Index Terms*—ECG; Deep Neural Network; LSTM, Classification Model**

I. Introduction

ELECTROCARDIOGRAPHIC (ECG) artifacts are alterations that are not related to cardiac electrical activity. The artifacts could be caused by the unexpected motion intensity, loss of electrode-skin contact or movements of different part of the system such as cables. Additionally, some of the common noises that appear in the ECG signal are electromyogram (EMG) noise and baseline drift due to breathing or sudden movement, and such noises can be easily removed using various filtering approaches. The problem occurs if the artifacts last too long, and completely compromise the shape of the signal. Then, the current physical state of the subject could be misinterpreted. That is why it is of great importance to identify these types of artifacts and ignore the parts of the signal in which they appear. This is especially referring to the wearable systems for electrophysiological monitoring that are used for healthcare or fitness where health condition of the subjects is further decided based on the parameters extracted from the ECG signal. Since the subjects perform various physical activity, the quality of the ECG signal even more decreases. As predictions of the physical state are given in the real time, the artifacts detection as well as the other algorithms should be performed automatically.

Tanja Boljanić and Goran Kvaščev are with the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: tanja.boljanic@gmail.com)

Tanja Boljanić and Jovana Malešević are with the Tecnalia Serbia Ltd, Deligradska 9, 11000 Belgrade, Serbia

Lui et al. [1] developed a wearable system for early detection of cardiovascular diseases and used machine learning algorithm for classifying ECG segments as acceptable and unacceptable for further analysis. Using Support Vector Machine (SVM), they could exclude unacceptable segments with an accuracy of 96.4%.

Neural networks are widely used for classification of different types of artifacts or arrhythmias in ECG signal [2-7]. Saadatnejad at al. [2] proposed a method consisting of wavelet transform and multiple recurrent neural networks for classifying arrhythmias in continuous cardiac monitoring on wearable devices. Deep network with wavelet sequences as input was used for classification of five heartbeat signals, resulting in high recognition performance [3]. Six common types of urgent arrhythmias are classified using deep neural network with an overall accuracy of 81% [4]. Deep learning algorithms were also used to classify shockable versus non-shockable rhythms in the presence and absence of cardiopulmonary resuscitation (CPR) artifact for automated external defibrillators [5]. Chauhan and Vig used deep recurrent neural network architecture with Long Short Term Memory (LSTM) units to detect abnormal and normal signals in ECG data [6]. The data included four different types of abnormal beats and the proposed detection system provided 96.5% performance.

The aim of the presented work is to develop a system that can automatically identify artifacts in ECG signals. We propose deep learning method for classification of unwanted artifacts and ECG signal that could be further processed, as we believe that these differences in signal could be reliably detected by a properly trained neural network.

Section II contains the method, including data preparation, and an explanation of used algorithm with the configuration of its parameters. The results are presented and discussed in section III, while the conclusion is attached in the final section.

II. The Method

The proposed algorithm is implemented in the Matlab R2019b software installed in a Windows 10 Pro platform, using the Signal Processing and Deep Learning toolboxes. The computer that was used is equipped with NVIDIA GeForce RTX 3060 graphics processing unit.

A. Dataset and Implementation

Dataset that was used in this study was collected by Tecnalia Serbia during the field trials within SIXTHSENSE

 ISBN 978-86-7466-930-3

project [8] in Bormio, Italy. The sensing module based on a multi-electrode array (MEA) for ECG recording was placed below the left major pectoralis. The MEA contains two recording and one referent electrode. The module provides conditioning and A/D conversion of the signal, and it is connected to the sensor for data acquisition via the flat cable. Described prototype was developed for research purposes within the project. This dataset comprises one-lead ECG recordings from 12 mountain rescuers that were performing the rescue task (male/female, 11/1, age (mean ± std), 31.7 ± 6.0 year). The length of the ECG signals ranges from 90 to 200 min with the average of 150 min, and the sampling rate is 1000 Hz. The signal contains motion artifacts due to the unexpected motion intensity and motion state, and the noises due to the change of relative displacement between electrode and skin, as well as the artifacts due to the cable movements. These noises have typical characteristics as transient high amplitude impulse and signal saturation. In order to utilize the automatic identification of the artifacts in signal, the ECG was visually inspected and we manually selected segments that are extremely noisy.

All signals were divided into 10-s segments, with total of 7,517 valid ECG and 2,714 segments labeled as signal with artifacts (7,517 + 2,714 = 10,231). Artifact segments are selected so that more than a half of the segment contain pure noise resulting in visually undetectable QRS complexes. An example of signal's both classes is presented in Fig. 1.

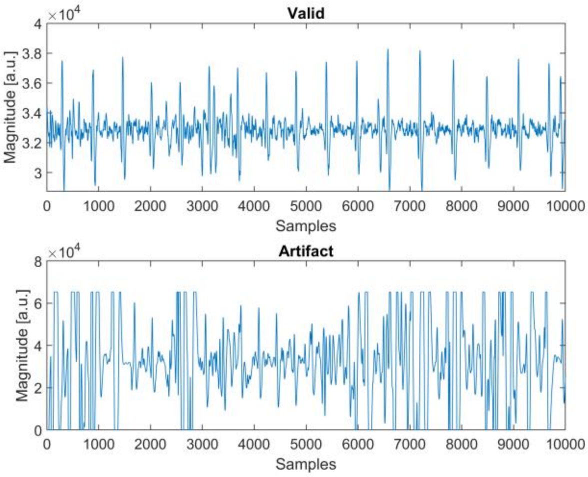


Fig. 1. The example of 10-s ECG segments: a segment without artifacts (*Valid*, top) and a segment with artifacts (*Artifact*, bottom)

The raw data segments were divided into two sets: 90% and 10% for the training and testing, respectively. Both classes were randomly divided into these two sets. Since 73.5% of the dataset are valid segments, a classifier would learn that it can achieve a high accuracy simply by classifying all signals as *Valid.* To avoid this bias, the *Artifact* signal was augmented so that there is the same number of *Valid* and *Artifact* signals. This is one form of data augmentation used in deep learning, known as oversampling [9]. *Artifact* signals were augmented after splitting data into two sets, hence the data from the test set are not included in the training set. At the end, the distribution between *Valid* and *Artifact* signals was evenly balanced, as showed in Table I.

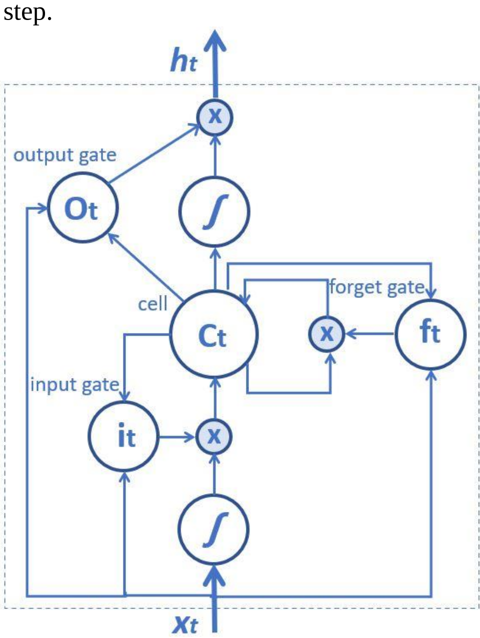
B. *Proposed Deep Neural Network*

Deep Neural Networks (DNN) are one type of model for machine learning that is subfield of artificial inteligence (AI) [10]. The appropriate deep learning algorythm depends on the task and the available data. Long short-term memory (LSTM) networks are the most commonly used variation of recurrent neural networks (RNN) that are well situated to study sequence and time-series data [11].

TABLE I
THE DIVISION OF THE ENTIRE DATASET INTO TRAINING AND TESTING SETS

Classes	The original number of 10-s segments	Number of 10-s segments after leveling	
		Training	Test
Valid	7,517	6,765	752
Artifact	2,714	6,765	271

The LSTM network can effectively learn long-term relationships between time steps of a sequence. It consists of an input gate, forget gate, output gate and cell unit. The cell remembers values over arbitrary time intervals, and the gates regulate the flow of information into and out of the cell. Input gate protects the unit from irrelevant input events, while the forget gate controls when to forget previous memory contents. The output gates controls the output flow. Graphical representation of an LSTM unit is presented in Fig. 2. The LSTM network can look at the time sequence in the forward direction and in both forward and backward directions, which is than called bidirectional LSTM (BiLSTM). This is useful when there is need to learn from the complete time series at each time step.



 ISBN 978-86-7466-930-3

Fig. 2. LSTM block diagram. The variables are: x_t – input vector to the LSTM unit, i_t – input gate's activation vector, f_t – forget gate's activation vector, o_t – output gate's activation vector, c_t – cell state vector, h_t – output vector of the LSTM unit.

C. *Network Parameter Configuration*

LSTM network was proposed for ECG classification. The network parameters were selected experimentally or following the other studies [9,12,13]. The bidirectional LSTM (BiLSTM) with 200 hidden units was used, as it looks at the sequence in both directions – backwards and forward, which is important when network should learn from full-time series at each time step. Sequence input was set to one dimension, because the input signal is only the amplitude of raw ECG. Two classes were specified by including a fully connected layer of size 2. The last two layers were softmax and classification layer with cross-entropy loss function. The adaptive moment estimation algorithm (ADAM) was used as the optimization method. Architecture of a network is summarized in Table II. The layer information from the table includes the layer type, the size and format of the layer activations, and the size of learnable parameters.

TABLE II
THE DETAILED INFORMATION FOR EACH LAYER OF THE PROPOSED NETWORK MODEL

Name	Type	Activations	Learnables
input	Sequence input	1	-
BiLSTM	BiLSTM	400	InputWeights 1600x1 RecurrentWeights 1600x200 Bias 1600x1
fc	Fully Connected	2	Weights 2x400 Bias 2x1
softmax	Softmax	2	-
output	Classification Output	-	-

The BiLSTM layer has the following parameters: initial learning rate = 0.01, mini-batch size = 150, epoch = 10, gradient threshold = 1, sequence length = 1000, dropout = 0. Number of epochs in the number of passes through the training data, and increasing this number wasn't resulting in better classification accuracy. Mini-batch size is the number of signals that neural network looks at a time, while the signal is broken into smaller sequences (sequence length) so that the computer does not run out of the memory. Dropout layer was not used as the previous studies showed that it did not increase the network generalization ability [9,12].

III. RESULTS AND DISCUSSION

The ECG data were classified using the LSTM network and performance measures of the model were evaluated using a confusion matrix. Confusion matrixes that are obtained after the training and testing process are presented in Fig. 4-5.

A row-normalized row summary represents the percentages of correctly and incorrectly classified observations for each true class, while a column-normalized column summary represents the same thing but for each predicted class. In order to report performance results for binary classification of *Valid* and *Artifact* ECG, five statistical metrics are extracted from the confusion matrix and presented in Table III: accuracy, sensitivity, specificity, precision and F1-score.

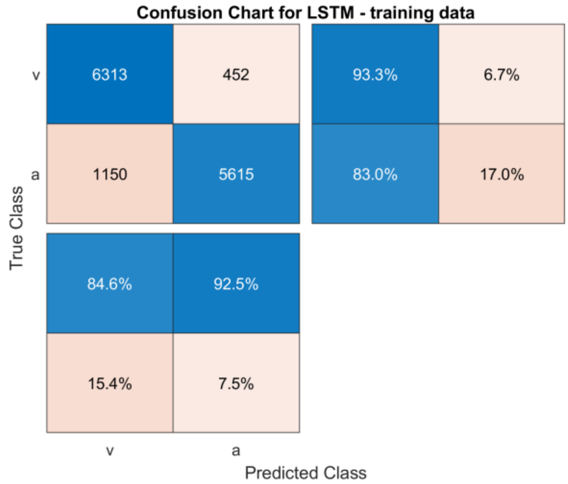


Fig. 4. Confusion matrix for the training set of the LSTM with a raw ECG. The axes labels represent the class labels, *Valid* - "v" and *Artifact* - "a".

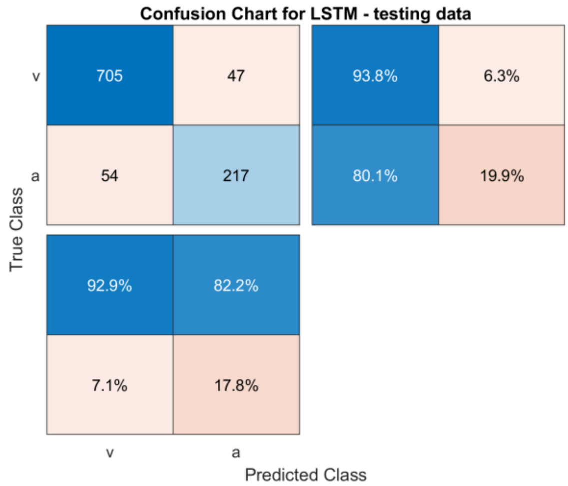


Fig. 5. Confusion matrix for the test set of the LSTM with a raw ECG. The axes labels represent the class labels, *Valid* - "v" and *Artifact* - "a".

The LSTM accuracy for the training set was 88.2%, while for the testing set it was 90.1%. The accuracy is the proportion of correctly classified ECG segments of all ECG segments. Sensitivity calculates the number of correctly classified valid ECG segments out of the total samples in the class, while the specificity calculates the number of correctly classified artifact ECG segments out of the total samples in the mentioned class. The precision calculates the number of true positives out of the positive classified classes. Finally, F1-score is the harmonic mean of both the precision and sensitivity measures and it is used as an overall score on how well the model is performing.

 ISBN 978-86-7466-930-3

TABLE III
EVALUATION METRICS FOR A CLASSIFICATION TEST

Accuracy	90.1%
Sensitivity	93.8%
Specificity	80.1%
Precision	92.9%
F1-score	93.3%

The time consumption for training was 53 min, which is acceptable, considering the large database that included more than 10,000 ECG segments (a 10-s duration for each segment). Segments of 10 seconds was used, because most of the ECG monitors display and analyze such signal duration, and Hajeb-M et al. [5] reported that 8 s segments is the best choice for classification accuracy. As mentioned in the method section, we selected the segments of ECG signal that are extremely noisy as *Artifact* signals. In practical applications, it would be of interest to observe not only the signals that are incredibly noisy, but rather signals with various degrees of noise, and it will be considered in the future work. Network parameters that are used are the optimal one for this type of dataset. Other options did not help the network to improve the classification accuracy. Some of the changes that were performed are decreasing the learning rate and the mini-batch size, and increasing the number of epochs.

The overall accuracy of 90.1% indicates that the proposed model could provide accurate prediction on a raw ECG data. By observing the sensitivity and specificity values, it can be seen that performed model recognizes valid ECG segments better than artifacts.

Similar study [9] that was also using LSTM network on raw ECG data showed the accuracy of 70.8%. Chen et al. [4] developed a classification model for six types of urgent arrhythmias combining CNN (Convolutional Neural Network) and LSTM with accuracy of 81.0%, sensitivity of 82.0% and specificity of 97.0%. Combination of CNN and LSTM was also used by Liang et al. [14], who verified the classification accuracy on three different datasets of raw ECG signals and obtained F1-scores of 85.0%, 80% and 82.6%. For detection of shockable rhythms in the presence and absence of cardiopulmonary resuscitation (CPR) artifact, Hajeb-M et al. applied deep-learning algorithm using convolutional layers, residual networks and BiLSTM [5]. The sensitivity, specificity, accuracy and F1-score were 95.2%, 86.0%, 88.1% and 83.5%, respectively.

IV. CONCLUSION

Automatic detection of artifacts in ECG signal is important goal in wearable monitoring systems in order to accurate determine subject's physical state. This work proposes to use deep learning technique for classifying ECG signal. LSTM network was used on a raw ECG signal that was divided into 10-s segments. The classification accuracy of 90.1% indicate that the proposed model shows promising results. The future work will include more data to improve the training of the neural network. Also, ECG signals with various degrees of noise will be considered. One of the possible applications of this research could be within the SIXTHSENSE project, to improve the existing algorithm for determining the heart rate signal from the ECG of the first responders (mountain rescuers and firefighters).

ACKNOWLEDGMENT

The authors would like to acknowledge that this research was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia (contract 451-03-68/2022-14/200103).

The work in this study was also performed within the SIXTHSENSE project, which has received funding by European Union's Horizon 2020 research and innovation programme under grant agreement No 883315.

REFERENCES

[1] Liu, C., Zhang, X., Zhao, L., Liu, F., Chen, X., Yao, Y., & Li, J. (2018). Signal quality assessment and lightweight QRS detection for wearable ECG SmartVest system. *IEEE Internet of Things Journal*, *6*(2), 1363-1374.

[2] Saadatnejad, S., Oveisi, M., & Hashemi, M. (2019). LSTM-based ECG classification for continuous monitoring on personal wearable devices. *IEEE journal of biomedical and health informatics*, *24*(2), 515-523.

[3] Yildirim, Ö. (2018). A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification. *Computers in biology and medicine*, *96*, 189-202.

[4] Chen, Y. J., Liu, C. L., Tseng, V. S., Hu, Y. F., & Chen, S. A. (2019, May). Large-scale classification of 12-lead ECG with deep learning. In *2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)* (pp. 1-4). IEEE.

[5] Hajeb-M, S., Cascella, A., Valentine, M., & Chon, K. H. (2021). Deep Neural Network Approach for Continuous ECG-Based Automated External Defibrillator Shock Advisory System During Cardiopulmonary Resuscitation. *Journal of the American Heart Association*, *10*(6), e019065.

[6] Chauhan, S., & Vig, L. (2015, October). Anomaly detection in ECG time signals via deep long short-term memory networks. In *2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)* (pp. 1-7). IEEE.

[7] Attia, Z. I., Kapa, S., Lopez-Jimenez, F., McKie, P. M., Ladewig, D. J., Satam, G., ... & Friedman, P. A. (2019). Screening for cardiac contractile dysfunction using an artificial intelligence–enabled electrocardiogram. *Nature medicine*, *25*(1), 70-74.

[8] https://sixthsenseproject.eu/

[9] Kłosowski, G., Rymarczyk, T., Wójcik, D., Skowron, S., Cieplak, T., & Adamkiewicz, P. (2020). The use of time-frequency moments as inputs of lstm network for ecg signal classification. *Electronics*, *9*(9), 1452.

[10] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.

[11] Patterson, J., & Gibson, A. (2017). *Deep learning: A practitioner's approach*. " O'Reilly Media, Inc.".

[12] Gao, J., Zhang, H., Lu, P., & Wang, Z. (2019). An effective LSTM recurrent network to detect arrhythmia on imbalanced ECG dataset. *Journal of healthcare engineering*, *2019*.

[13] Kim, B. H., & Pyun, J. Y. (2020). ECG identification for personal authentication using LSTM-based deep recurrent neural networks. *Sensors*, *20*(11), 3069.

[14] Liang, Y., Yin, S., Tang, Q., Zheng, Z., Elgendi, M., & Chen, Z. (2020). Deep learning algorithm classifies heartbeat events based on electrocardiogram signals. *Frontiers in Physiology*, 1255.

 ISBN 978-86-7466-930-3