Deep Learning for Pulse Detection in Out-of-Hospital Cardiac Arrest Using the ECG

Andoni Elola[1], Elisabete Aramendi[1], Unai Irusta[1], Artzai Picón[2], Erik Alonso[1],
Pamela Owens[3], Ahamed Idris[3]

[1] University of the Basque Country, Bilbao, Bizkaia, Spain
[2] TECNALIA, Derio, Bizkaia, Spain
[3] University of Texas Southwestern Medical Center, Dallas, Texas, USA

Abstract

Pulse detection during out-of-hospital cardiac arrest is necessary to identify cardiac arrest and detect return of spontaneous circulation. Currently, carotid pulse checking and checking for signs of life are the most widespread procedures for pulse detection, but both have been proven inaccurate and time consuming. Automatic methods that could be integrated in Automated External Defibrillators (AEDs) are needed. In this study we propose a deep neural network classifier to detect pulse using exclusively the ECG. We extracted 3914 segments of 4 s from 279 patients, all of them with an organized rhythm. They were annotated as pulsed rhythm or pulseless rhythm based on clinical information. A total of 2372 pulsed rhythms and 1542 pulseless rhythms were included in the study. To train and test the model 3038 (223 patients) and 876 segments (56 patients) were used, respectively. The model was evaluated in terms of Sensitivity (Se) and Specificity (Sp) for pulse detection. The network showed a Se/Sp of 89.4%/97.2% during training process and 91.7%/92.5% for the test set.

1. Introduction

Sudden cardiac death is one of the leading causes of death in the industrialized world. Despite progress in different fields, survival rates in the out-of-hospital settings remain low, around 10%. The detection of pulse is crucial for the recognition of both, cardiac arrest and the Return of Spontaneous Circulation (ROSC) [1].

Palpation of carotid has been long used to detect pulse, but it has been proven to be inaccurate and time consuming [2,3]. Current Resuscitation guidelines [1] recommend looking for signs of life in the patient, which has not been proven to be more accurate [4].

Several automatic methods have been proposed using the ECG and the thoracic impedance recorded by the defibrillation pads [5–7]. Availability and low resolution of the impedance signal compromise the applicability of those methods. A more universal approach, usable in any Automated External Defibrillator (AED), is to use only the ECG to detect pulse. This paper presents a new approach based on a deep neural network.

Deep learning techniques showed good accuracies in physiological signal classification tasks [8]. In this paper we propose a novel deep network to classify an organized ECG into pulsed rhythm (PR) or pulseless electrical activity (PEA).

2. Materials

The data used for this study were a subset of a large Out-of-Hospital Cardiac Arrest (OHCA) database recorded by the DFW center for resuscitation research (UTSW, Dallas). All episodes were recorded using the Philips HeartStart MRx device, including ECG with a sampling frequency of 250 Hz and a resolution of 1.03μV per least significant bit.

A total of 1015 episodes containing concurrent ECG and impedance recordings were considered and separated into two groups (ROSC/no-ROSC) using the ROSC and time of ROSC (t_{rosc}) annotations made by the clinicians in the scene. ROSC episodes had no chest compressions or shocks after t_{rosc}. In the no-ROSC group, we discarded patients transported to hospital and episodes with sustained organized rhythms once chest compressions were suspended, because such actions are associated with patients in ROSC.

Five second segments presenting an organized rhythm were automatically extracted during intervals without chest compression artifacts [9]. These segments were labelled as PR and PEA for classification. PR segments were extracted in the ROSC episodes after t_{rosc} with a minimum interval between segments of 30 s. PEA segments were extracted in the no-ROSC group with a minimum interval between segments of 1 s.
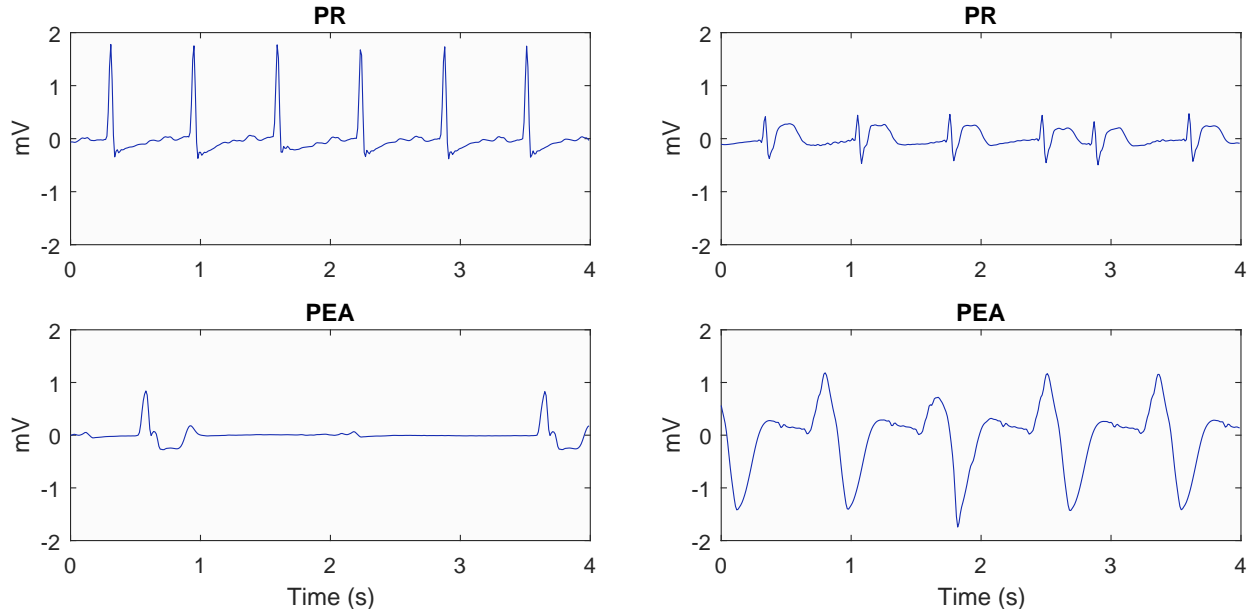
Figure 1. Pulsed rhythm (PR) and Pulseless Electrical Activity (PEA) examples.

The final dataset contained 279 episodes (134 ROSC/145 no-ROSC), and a total of 3914 segments, 2372 PR and 1542 PEA. Figure 1 shows examples of PR and PEA segments, and shows that narrower QRS complexes and higher heart rates are associated with PR. Data were divided in patient-wise training and test sets. The first one is composed by 3038 segments from 223 patients (1871 PR and 1167 PEA) and the second one by 876 segments from 56 patients (501 PR and 375 PEA).

3. Methods

ECG data were downsampled to 100 Hz and bandpass filtered between 0.5-30 Hz. The analysis window was 4 s, so 400 samples were input to the deep neural network classifier. The network was implemented using Keras with Tensorflow backend [10, 11].

3.1. Network design

Figure 2 shows the overall scheme of the deep neural network applied in this proposal. A total of 4 layers constitute the final solution (in blue) and the extra layers were used to train the model (in red).

The first layer adds gaussian noise (mean zero and $\sigma = 0.03$) to the input signal to avoid overfitting in the training process.

The second layer is a convolutional layer, which applies temporal convolution to the input data. The n-th sample of the ℓ-th filtered signal is obtained following the next equation:

$$x_\ell[n] = f\left(\sum_{i=0}^{m-1} w_{i\ell} x[n-i] + b_{i\ell} \right) \qquad (1)$$

where $f(\bullet)$ is the linear rectifier and $m = 3$ was taken. A total of 6 filters ($\ell = 1, \ldots, 6$) were applied, and the weigths, w_{il}, and the biases, b_{il}, were optimized in the training process.

The third layer, the pooling layer, downsamples the filtered signals by a ratio of 3 and a maximum criteria.

The fourth layer was a recurrent layer, used to exhibit the temporal behaviour of the data. Gate Recurrent Unit (GRU) is a simpler version of Long Short-Term Memory (LSTM), which was designed to learn long-term dependencies [12], but its accuracy is similar [13]. In our solution two GRU layers were applied, backwards and forwards, of 10 units each.

The final stage, includes a single neuron as classification stage with the sigmoid as activation function. The output, y, is the likelihood of the segment being PR.

In the training process we adopted some methods to avoid overfitting. Besides adding gaussian noise, dropout layers were added between third and fourth layer and between the last two layers. These kind of layers drops out units under a certain probability [14], shown between layers in Figure 2. According to [14], when using dropout probabilities it is specially useful to constrain the norm of the weight vector at each layer, i.e. the layers are optimized under the constraint $||\boldsymbol{w}|| < \gamma$. In our solution $\gamma = 2$ and the constraint was applied to convolutional and recurrent
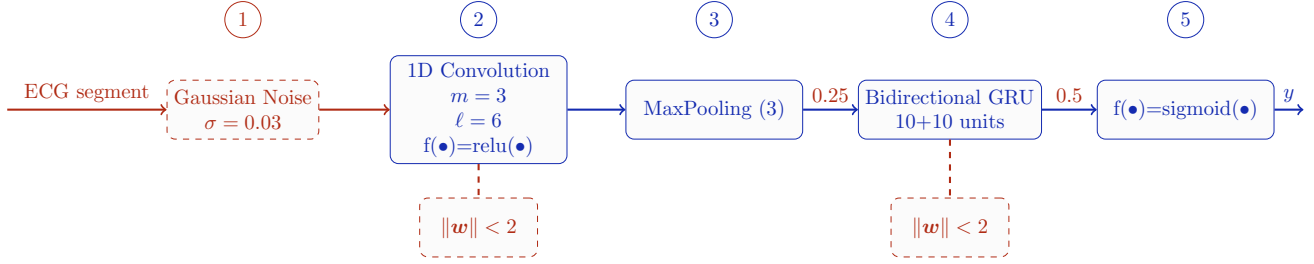
Figure 2. Overall architecture of the deep learning solution

layers. All weights were initialized using Xavier normal distribution and we used binary cross-entropy as loss function to optimize the network. All patients were weighted equally to compute the loss function and we used the RMSprop optimizer. Finally, training data were shuffled at the beginning of each epoch to change the batches. The batch size was set to 4 and the number of epochs was 200.

3.2. Performance metrics

The unweighted accuracy (UAcc) of the network proposed was evaluated for the training and the test set at each epoch in order to analyse the convergence of the network.

The overall performance of the method was given in terms of Sensitivity (Se), proportion of correctly classified PRs; Specificity (Sp), proportion of correctly classified PEAs and Balanced Accuracy (BAC), the mean value of Se and Sp. To compute the performance metrics each patient was weighted equally.

4. Results

Figure 3 shows the unweighted accuracy of the network at each epoch for the training and testing sets. UAcc converges for epoch 100 with a difference below 4% in the end between training and test sets.

At the final epoch, the network showed a Se/Sp of 89.4%/97.2% for the training set and 91.7%/92.5% for the test set. A BAC reduction of 1.2% was measured for the test set.

5. Discussion and conclusions

Pulse detection remains challenging during OHCA for both experts and laypeople, so there is a clear need of accurate automatic methods. This is, for the best of our knowledge, the first study that proposes using a deep learning approach to detect pulse using only the ECG signal. This allows the universal use of the algorithm, as recording the ECG signal is necessary to analyse the patient's rhythm.

Our solution is based only in 4 layers, and several regularization techniques (adding gaussian noise, constraining the weights and dropout) were used to avoid overfitting.
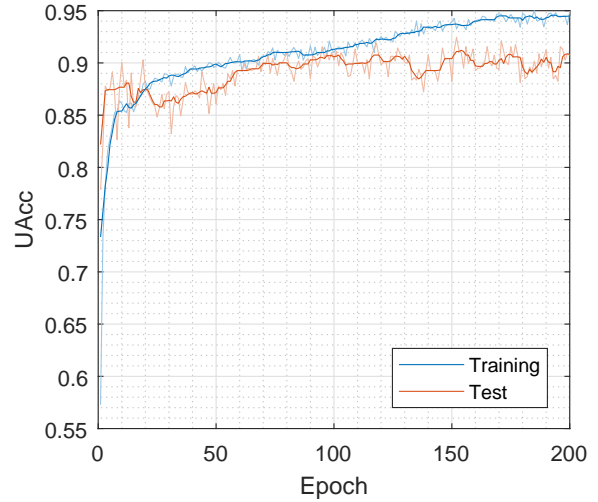

Figure 3. Unweighted accuracy of the network at each epoch of the training process.

Figure 3 shows the UAcc of the training and test sets per epoch. UAcc of the training set improves along epochs, but the test set's UAcc does not improve from epoch 100 to the end. However, it does not fall, showing that the adopted regularization techniques were good enough. We also trained and tested the network without regularization techniques, and achieved a Se/Sp of 98.7%/97.0% during training, but 94.6%/87.8% for the test set. The difference in BAC is 6.7 points, well above from 1.2 points achieved using regularization techniques. Further adjustment of the regularization parameters, using data augmentation and more data could further reduce the gap between training and test performances.

The algorithm showed a Se/Sp of 91.7%/92.5% respectively. Those scores are similar to other solutions that require both ECG and impedance [5, 6]. Our solution based exclusively on the ECG permits the universal use of the algorithm in any AED.

References

[1] Soar J, Nolan JP, Böttiger BW, Perkins GD, Lott C, Carli P, Pellis T, Sandroni C, Skrifvars MB, Smith GB, et al. European resuscitation council guidelines for resuscitation 2015: section 3. adult advanced life support. Resuscitation 2015; 95:100–147.

[2] Eberle B, Dick W, Schneider T, Wisser G, Doetsch S, Tzanova I. Checking the carotid pulse check: diagnostic accuracy of first responders in patients with and without a pulse. Resuscitation 1996;33(2):107–116.

[3] Tibballs J, Russell P. Reliability of pulse palpation by healthcare personnel to diagnose paediatric cardiac arrest. Resuscitation 2009;80(1):61–64.

[4] Ruppert M, Reith MW, Widmann JH, Lackner CK, Kerkmann R, Schweiberer L, Peter K. Checking for breathing: evaluation of the diagnostic capability of emergency medical services personnel, physicians, medical students, and medical laypersons. Annals of Emergency Medicine 1999; 34(6):720–729.

[5] Risdal M, Aase SO, Kramer-Johansen J, Eftesol T. Automatic identification of return of spontaneous circulation during cardiopulmonary resuscitation. IEEE Transactions on Biomedical Engineering 2008;55(1):60–68.

[6] Alonso E, Aramendi E, Daya M, Irusta U, Chicote B, Russell JK, Tereshchenko LG. Circulation detection using the electrocardiogram and the thoracic impedance acquired by defibrillation pads. Resuscitation 2016;99:56–62.

[7] Ruiz JM, de Gauna SR, González-Otero DM, Saiz P, Gutiérrez JJ, Veintemillas JF, Bastida JM, Alonso D. Circulation assessment by automated external defibrillators during cardiopulmonary resuscitation. Resuscitation 2018; 128:158–163.

[8] Faust O, Hagiwara Y, Hong TJ, Lih OS, Acharya UR. Deep learning for healthcare applications based on physiological signals: a review. Computer Methods and Programs in Biomedicine 2018;.

[9] Alonso E, Ruiz J, Aramendi E, González-Otero D, de Gauna SR, Ayala U, Russell JK, Daya M. Reliability and accuracy of the thoracic impedance signal for measuring cardiopulmonary resuscitation quality metrics. Resuscitation 2015;88:28–34.

[10] Chollet F, et al. Keras. https://keras.io, 2015.

[11] Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, Corrado GS, Davis A, Dean J, Devin M, Ghemawat S, Goodfellow I, Harp A, Irving G, Isard M, Jia Y, Jozefowicz R, Kaiser L, Kudlur M, Levenberg J, Mané D, Monga R, Moore S, Murray D, Olah C, Schuster M, Shlens J, Steiner B, Sutskever I, Talwar K, Tucker P, Vanhoucke V, Vasudevan V, Viégas F, Vinyals O, Warden P, Wattenberg M, Wicke M, Yu Y, Zheng X. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. URL https://www.tensorflow.org/. Software available from tensorflow.org.

[12] Hochreiter S, Schmidhuber J. Long short-term memory. Neural computation 1997;9(8):1735–1780.

[13] Chung J, Gulcehre C, Cho K, Bengio Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint arXiv14123555 2014;.

[14] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: A simple way to prevent neural networks from overfitting. The Journal of Machine Learning Research 2014;15(1):1929–1958.

Address for correspondence:

Andoni Elola
Engineering school of Bilbao
andoni.elola@ehu.eus