Exploring Transfer Learning for Ventricular Tachycardia Electrophysiology Studies

Andrea Pitzus[1], Giulia Baldazzi[1,2], Marco Orrù[1], Alberto Valdes Rey[1], Graziana Viola[3], Luigi Raffo[1], Petar Djuric[4], Danilo Pani[1]

[1]Department of Electrical and Electronic Engineering (DIEE), University of Cagliari, Cagliari, Italy
[2]Department of Informatics, Bioengineering, Robotics and Systems Engineering (DIBRIS), University of Genova, Genova, Italy
[3]EP STAFF, Santissima Annunziata Hospital, Sassari, Italy
[4]Stony Brook University, Stony Brook, NY, USA

Abstract

Arrhythmogenic sites in post-ischemic ventricular tachycardia (VT) are usually identified by looking for abnormal ventricular potentials (AVPs) in intracardiac electrograms (EGMs). Unfortunately, the accurate recognition of AVPs is a challenging problem for different reasons, including the intrinsic variability in the AVP waveform. Given the high performance of deep neural networks in several scenarios, in this work, we explored the use of transfer learning (TL) for AVPs detection in intracardiac electrophysiology.

A balanced set of 1504 bipolar intracardiac EGMs was collected from nine post-ischemic VT patients. The time-frequency representation was generated for each EGM by computing the synchrosqueezed wavelet transform to be used in the re-training of the convolutional neural network.

The proposed approach allows obtaining high recognition results, above 90% for all the investigated performance indexes, demonstrating the effectiveness of deep learning in the recognition of AVPs in post-ischemic VT EGMs and paving the way for its use in supporting clinicians in targeting arrhythmogenic sites. In addition, this study further confirms the efficacy of the TL approach even in case of limited dataset sizes.

1. Introduction

One of the most frequent treatments for patients affected by post-ischemic ventricular tachycardia (VT) is silencing their myocardial arrhythmogenic sites by radiofrequency ablation. Typically, these regions are targeted during electroanatomic mapping procedures by visual inspection of the electrograms (EGMs) recorded locally, supported by some vendor-specific tools. During these clinical procedures, the cardiologists may aim at identifying target points for subsequent ablation, guided by the presence of so-called abnormal ventricular potentials (AVPs), thus leading to operator-dependent and time-consuming procedures.

To overcome these limitations, some recent studies exploited artificial intelligence tools in this field [1], [2]. Specifically, to the best of our knowledge, only few automatic approaches based on machine learning were proposed so far to identify the arrhythmogenic sites in the treatment of VT, both considering features extracted from different domains [2] and/or simply the time-series of the recorded EGMs [1]. However, the application of more complex artificial intelligence tools, such as deep neural networks, has not been investigated yet.

In this work, we propose an automatic approach exploiting a deep convolutional neural network (CNN) to support clinicians in targeting VT arrhythmogenic substrates in a fast, user-unbiased, and precise way. To this aim, we investigated the use of transfer learning (TL) for AVPs detection, by exploiting an annotated real dataset for both training and testing the CNN performance.

2. Materials and methods

A dataset composed of bipolar intracardiac EGMs collected from nine patients affected by post-ischemic VT during left ventricle electroanatomic mapping procedures at the San Francesco Hospital (Nuoro, Italy) has been used. A total of 1504 bipolar EGMs were included, equally divided into 752 physiological potentials and 752 AVPs. All the EGMs were annotated by an expert cardiologist using an ad-hoc MATLAB graphical user interface [1], [3].

The acquisition of all the bipolar EMGs was performed through the CARTO®3 mapping system (Biosense Webster, Inc., Diamond Bar, California). Given the adopted CARTO®3 recording settings (i.e., sampling rate equal to 1 kHz and band-pass filtering between 16 and 500 Hz), for each recorded EGM, only the reference beat given by the CARTO®3 was considered, thus ensuring a proper EGM acquisition when the catheter was in contact with the endocardium. Specifically, for each EGM, we considered only a 500-ms window around the reference annotation provided by the CARTO®3 system.

2.1 Image generation

In order to implement the CNN-based automatic approach, a first step for the generation of the input images was needed.

Specifically, since the presence of spectral signatures for the AVPs has been recently investigated [3], from each 500-ms EGM segment a time-frequency representation was obtained by means of the synchrosqueezed wavelet transform (SSWT) [4], thus providing the CNN with all the information about AVPs and physiological EGM instantaneous spectral contents.

From the SSWT we extracted the modulus, resulting in a positive defined range of the image values, with different maximum levels. In order to obtain an unbiased representation for all the EGM spectrograms, the same upper limit was imposed for the saturation of all the input images. In this regard, we computed the maximum value for each SSWT representation, and we chose the median value of the maxima distribution as upper saturation limit.

Furthermore, all images were generated by removing the time and frequency axes and setting the 'jet' colormap and the flat shading. Then, each instance was stored as a .jpg file with a spatial resolution corresponding to the input layer sizes of the adopted CNN (i.e., 227×227).

2.2 Adopted CNN and transfer learning approach

Among the different CNNs available in the literature, in this investigation, we chose the AlexNet deep CNN network [5] to perform the AVP and physiological EGM recognition, because of its small number of weights with respect to other CNNs, and its capability to maintain high performances on different types of images. As can be seen from Figure 1, AlexNet is able to manage input images with size of 227×227 by five convolutional layers and three fully connected layers (FC6, FC7, and output layer).

Indeed, AlexNet has been used in other scientific works [6], [7], demonstrating its versatility in various classification problems solved by TL. Due to the limited size of our dataset, in this work, a pre-trained version of

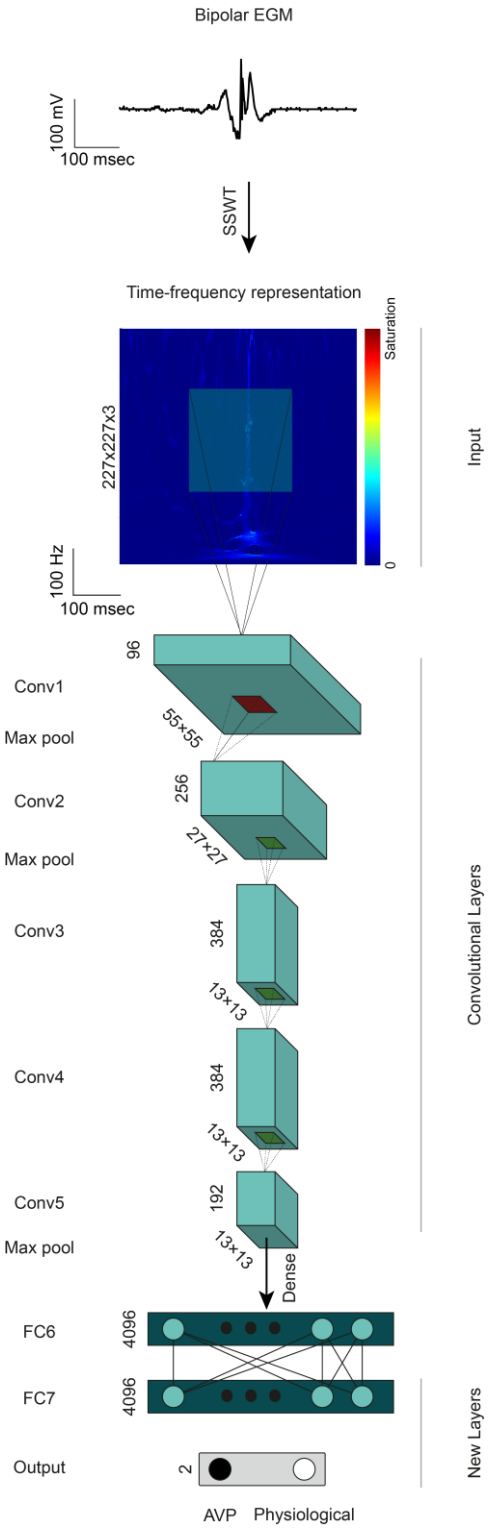

Figure 1. Pipeline for CNN-based AVPs automatic recognition approach proposed in this work, along with a detailed representation of the AlexNet architecture.

the AlexNet network has been used, in which the weights were already imposed by a training stage on more than a million images from the ImageNet database. Originally, the network was able to classify a wide pool of images into 1000 object categories (e.g., keyboard, mouse, pencil, and many animals). However, among them, no AVP spectrogram or time-frequency representation was included, thus making our classification problem a very challenging issue for this model.

Therefore, to retrain the model on our instances and adapt its behavior to binary classification, the final classification layers (i.e., the last fully connected, the softmax, and the classification output layers) have been modified. In particular, the fully connected layer has been substituted with a new fully connected layer of the same size with randomly initialized weights obtained by using the Glorot uniform initialization method, while the output size was modified from 1000 to binary (see Figure 1).

As regards the training parameterizations and settings, different choices were made. Specifically, to update network weights we used the stochastic gradient descent with momentum as training algorithm, and the binary cross-entropy as loss function to compute the gradient. For the TL, according to many studies in the scientific literature [7], [8], we imposed a learning rate equal to 10^{-4}, since the weights of the different layers were already tuned on different types of images, rather than randomly initialized, thus allowing also slow little re-train to obtain good results. Furthermore, a batch size of 10 instances per iteration was selected, to provide a more stable and reliable training, which can improve generalization of the network predictions [9], while the momentum was set to 0.9. Conversely, since the weights were randomly initialized in the new fully connected layer, the weight learning rate factor and the bias learning rate factor were increased by a factor of ten, to allow faster learning.

To avoid overfitting, an early-stop condition was imposed by evaluating the validation set every 50 training iterations and arresting the training after nine times in which the loss stopped decreasing.

All computations have been carried out in MATLAB 2022a.

2.3 Performance indexes

To assess the efficacy of the proposed AVPs automatic recognition approach by CNN, we evaluated its performance in a 10-fold cross-validation. Specifically, eight folds were considered for the training set, while the remaining ones were used as validation and test sets.

Performance evaluation was carried out by computing different metrics: accuracy (Acc), sensitivity (TPR), specificity (TNR), precision (PPV), and F1-score, as:

$$Acc = (TP + TN)/(P + N) \quad (1)$$
$$TPR = TP/P \quad (2)$$
$$TNR = TN/N \quad (3)$$
$$PPV = TP/(TP + FP) \quad (4)$$
$$F1 = 2\,(TPR \times PPV)/(TPR + PPV) \quad (5)$$

where P and N identify the total number of AVPs and physiological instances, TP and FP those EGMs correctly and erroneously classified as AVPs, respectively, and TN the EGMs properly recognized as physiological signals.

3. Results

Performance results are reported in Table 1 in terms of mean and standard deviation computed across the 10-fold cross-validation, while Figure 2 depicts the cumulative confusion matrix among the 10 folds, in which all the samples of the dataset have been tested. As can be seen, very high recognition performances, above 92%, were achieved both in terms of overall accuracy and true identification rates of the two classes (i.e., AVPs and physiological EGMs), with very high precision (mean F1 equal to 0.92). Compared to previous studies, the obtained results are very promising: indeed, although in [2] performances equal to ~94% were reported in all metrics, these were obtained by using a significantly smaller dataset (i.e., 86 intracardiac EGMs) compared to the one introduced in this study. Furthermore, in [1], where a larger dataset was used (i.e., 953 intracardiac EGMs), the reported results showed remarkably lower accuracy performance (i.e., ~79% in both approaches) with respect to those obtained in the presented work.

Despite the higher computation load related to the CNN, on a 64-GB NVIDIA Tesla P100 GPU cluster, the training time for each fold was fast, resulting equal to 91.9 ± 16.2 sec on average, with a 1295 ± 263 number of training iterations. Remarkably, the best model according to the validation loss was reached after 845 iterations, corresponding to 5.62 epochs, due to the imposed validation frequency and patience.

Table 1. Mean (μ) and standard deviation (σ) for all indexes computed in the 10-fold cross-validation.

Index	μ ± σ
Acc (%)	92.5 ± 2.5
TPR (%)	92.0 ± 4.0
TNR (%)	93.0 ± 3.3
PPV (%)	93.0 ± 3.0
F1-score	0.92 ± 0.03

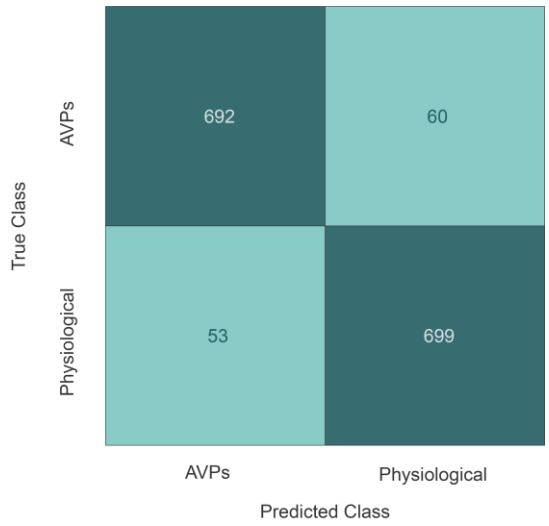

Figure 2. Cumulative confusion matrix across all the tested EGMs.

4. Conclusions

In this study, we proposed an automatic approach to target arrhythmogenic sites using TL, exploiting the intracardiac signal signatures in a time-frequency representation. As it represents the first explorative study on the feasibility of detecting AVPs using deep learning architectures, we did not perform any hyperparameter optimization, but we adopted a series of generally acceptable values for the hyperparameters in the context of the TL.

Despite the reported results are promising for the development of a computer-aided system for VT treatment by catheter ablation, it is important to underline that the proposed approach was not stressed by providing imbalanced test sets, as in typical real electroanatomic mapping scenarios, and that the dataset is limited in the number of entries and in the labeling bias due to the single annotator. As such, a deeper investigation on a larger dataset is needed, to demonstrate the generalizability and robustness of the proposed approach.

References

[1] G. Baldazzi, M. Orrù, M. Matraxia, G. Viola, and D. Pani, 'Supervised Classification of Ventricular Abnormal Potentials in Intracardiac Electrograms', presented at the 2020 Computing in Cardiology Conference, Dec. 2020. doi: 10.22489/CinC.2020.397.

[2] G. Baldazzi, M. Orrù, M. Matraxia, G. Viola, and D. Pani, 'Automatic Recognition of Ventricular Abnormal Potentials in Intracardiac Electrograms', presented at the 2019 Computing in Cardiology Conference, Dec. 2019. doi: 10.22489/CinC.2019.343.

[3] G. Baldazzi, M. Orrù, G. Solinas, M. Matraxia, G. Viola, and D. Pani, 'Spectral characterisation of ventricular intracardiac potentials in human post-ischaemic bipolar electrograms', *Sci Rep*, vol. 12, no. 1, p. 4782, Dec. 2022, doi: 10.1038/s41598-022-08743-7.

[4] I. Daubechies, J. Lu, and H.-T. Wu, 'Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool', *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 243–261, Mar. 2011, doi: 10.1016/j.acha.2010.08.002.

[5] A. Krizhevsky, I. Sutskever, and G. E. Hinton, 'ImageNet classification with deep convolutional neural networks', *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.

[6] S. Lu, Z. Lu, and Y.-D. Zhang, 'Pathological brain detection based on AlexNet and transfer learning', *Journal of Computational Science*, vol. 30, pp. 41–47, Jan. 2019, doi: 10.1016/j.jocs.2018.11.008.

[7] K. M. Hosny, M. A. Kassem, and M. M. Foaud, 'Classification of skin lesions using transfer learning and augmentation with Alex-net', *PLoS ONE*, vol. 14, no. 5, p. e0217293, May 2019, doi: 10.1371/journal.pone.0217293.

[8] G. Van Steenkiste, G. van Loon, and G. Crevecoeur, 'Transfer Learning in ECG Classification from Human to Horse Using a Novel Parallel Neural Network Architecture', *Sci Rep*, vol. 10, no. 1, p. 186, Dec. 2020, doi: 10.1038/s41598-019-57025-2.

[9] D. Masters and C. Luschi, 'Revisiting Small Batch Training for Deep Neural Networks'. arXiv, Apr. 20, 2018. Accessed: May 31, 2022. [Online]. Available: http://arxiv.org/abs/1804.07612

Address for correspondence:

Andrea Pitzus
Department of Electrical and Electronic Engineering,
University of Cagliari, Cagliari, Italy
andrea.pitzus@unica.it