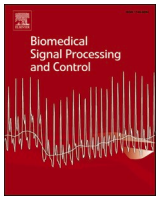
Biomedical Signal Processing and Control 99 (2025) 106844

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc



Arrhythmogenic sites identification in post-ischemic ventricular tachycardia electrophysiological studies by explainable deep learning

Andrea Pitzus [a,1], Giulia Baldazzi [a,*,1], Marco Orrù [b,a], Luigi Raffo [a], Graziana Viola [c], Petar M. Djurić [d], Danilo Pani [a]

[a] Department of Electrical and Electronic Engineering (DIEE), University of Cagliari, Cagliari, Italy
[b] Department of Informatics, Bioengineering, Robotics and Systems Engineering (DIBRIS), University of Genoa, Genoa, Italy
[c] Clinical and Interventional Cardiology Unit, Santissima Annunziata Hospital, Sassari, Italy
[d] Department of Electrical and Computer Engineering, Stony Brook University, Stony Brook, NY, USA

ARTICLE INFO

Keywords:
Abnormal ventricular potentials
Arrhythmogenic sites
Late potentials
Explainable deep learning
Ventricular tachycardia

ABSTRACT

Background and objective: Abnormal ventricular potentials (AVPs) in intracardiac electrograms (EGMs) are frequently considered as markers of arrhythmogenic sites in post-ischemic ventricular tachycardia (VT) during electroanatomic mapping (EAM) procedures. Their detection is strongly operator-dependent and time-consuming. This work explores the adoption of explainable deep learning to support the discrimination between physiological EGMs and AVPs.

Methods: Three convolutional neural networks were trained to discriminate the target signals based on their time–frequency representations by synchrosqueezed wavelet transform. The efficacy of the method was assessed on 2561 real bipolar EGMs collected from nine post-ischemic VT patients.

Results: The proposed approach achieved high performance, with accuracy levels reaching up to 89%. It also demonstrated coherent localization of the arrhythmogenic sites with respect to conventional voltage and local activation time maps. Moreover, by using saliency maps, AVPs discriminant signatures were highlighted at high frequencies (i.e., in the 103–125 Hz band, which was generally relevant for every network), in line with prior evidence.

Conclusion: For the first time, deep learning has been successfully applied and robustly evaluated in the field. The proposed approach paves the way to the development of effective AI-driven systems. These systems will enable a faster, trustworthy and operator-independent identification of AVPs in VT EAM procedures. Furthermore, even without injecting prior knowledge in the adopted models, the analysis of saliency maps revealed that CNNs are prone to autonomously select time–frequency ranges of the EGMs in agreement with the current knowledge.

1. Introduction

Out-of-hospital cardiac arrest is a leading cause of death in Europe [1], with an incidence between 67 and 170 per 100,000 inhabitants per year. Among these, ventricular tachycardia (VT) in ischemic cardiomyopathy is responsible for 40% of cases [2]. Abnormal impulse propagation pathways, and specifically scar-related reentrant circuits, represent the main arrhythmogenic mechanism of VT [3,4].

Catheter ablation represents an effective therapeutic option for different forms of VT [5–8]. Nowadays, catheter ablation may be guided by substrate mapping during electroanatomic mapping (EAM) procedures through the recording of local electrical activations called electrograms (EGMs). This approach has proved to be effective in reducing the arrhythmia recurrence [9,10] and the incidence of implantable cardioverter-defibrillator therapy [11]. Substrate-guided mapping may be based on the identification of low-voltage, slow-conducting areas, in sinus rhythm, which become targets for catheter ablation to inhibit electrical conduction through the arrhythmogenic tissue [12]. Several EGM characteristics are frequently exploited for the identification of the arrhythmogenic substrate [13,14], such as the presence of abnormal

* Corresponding author at: Piazza d'Armi, 09123 Cagliari, Italy.
E-mail addresses: andrea.pitzus@unica.it (A. Pitzus), giulia.baldazzi@unica.it (G. Baldazzi), marco.orru96@unica.it (M. Orrù), luigi.raffo@unica.it (L. Raffo), petar.djuric@stonybrook.edu (P.M. Djurić), danilo.pani@unica.it (D. Pani).
[1] These authors equally contributed.

https://doi.org/10.1016/j.bspc.2024.106844
Received 27 March 2024; Received in revised form 31 July 2024; Accepted 6 September 2024
Available online 15 September 2024

ventricular potentials (AVPs). After arrhythmogenic site identification, different substrate-based ablation strategies can be chosen by electrophysiologists [15,16], with a variable impact on the healthy myocardial substrate. However, both EAM and ablation are entirely manual and time-consuming procedures, and their outcome significantly depends on clinicians' expertise and level of attention.

Due to the relevance of accurately detecting VT arrhythmogenic substrates to reduce recurrences, various supporting tools, strategies for mapping [17–21], and detection algorithms [22–25] have been introduced.

In this regard, artificial intelligence (AI) may represent an important tool in cardiac electrophysiology, providing effective insights and innovative perspectives in the field [26]. AI tools, such as machine learning (ML) and deep learning (DL), have been widely applied to the study of cardiac arrhythmias [27,28] by considering different data types for training according to the specific application. For instance, AI on the 12-lead electrocardiograms (ECG) have been deeply investigated, both for detection of arrhythmic events [29–32], and their treatment, as in the case of scar localization in the myocardium [33]. Interestingly, AI has been also adopted in EGM analysis for computer-aided localization of ablation targets for atrial fibrillation (AF) treatment [34–36]. Besides applications based on noninvasive electrophysiological signals [33,37,38] and imaging [39,40], the use of ML has been recently investigated on EGMs for automatic detection of arrhythmogenic substrate in VT [41–43]. Indeed, this kind of approach has proven to be useful in distinguishing between physiological potentials and AVPs by extracting multiple features from different domains. Furthermore, in a very preliminary investigation, the use of DL for the detection of VT arrhythmogenic sites has been reported [44]. However, in such an explorative study, only the performance of a single convolutional neural network (CNN) has been paradigmatically presented, without any assessment of the models and parameterizations.

In this work, DL has been extensively investigated for the detection of arrhythmogenic sites in post-ischemic VT, by using the information enclosed in the morphology of EGMs and their spectral characteristics, leveraging the synchrosqueezed wavelet transform (SSWT) [45] to represent the signals by embedding both time and frequency information. To achieve this, three well-established pre-trained CNN architectures for image recognition have been investigated by transfer learning. Furthermore, the Gradient-Weighted Class Activation Mapping (Grad-CAM) [46] technique was used to perform the analysis of model explainability, pointing out the most informative time–frequency ranges for each EGM class.

The significance of this study lies in the ability of the proposed DL models to achieve highly accurate classification performance by leveraging only the time and frequency-domain signatures of the EGMs routinely acquired during EAM procedures, therefore without performing any manual feature extraction. To the best of the authors' knowledge, this study constitutes the first robust investigation into the use of DL in targeting VT arrhythmogenic sites.

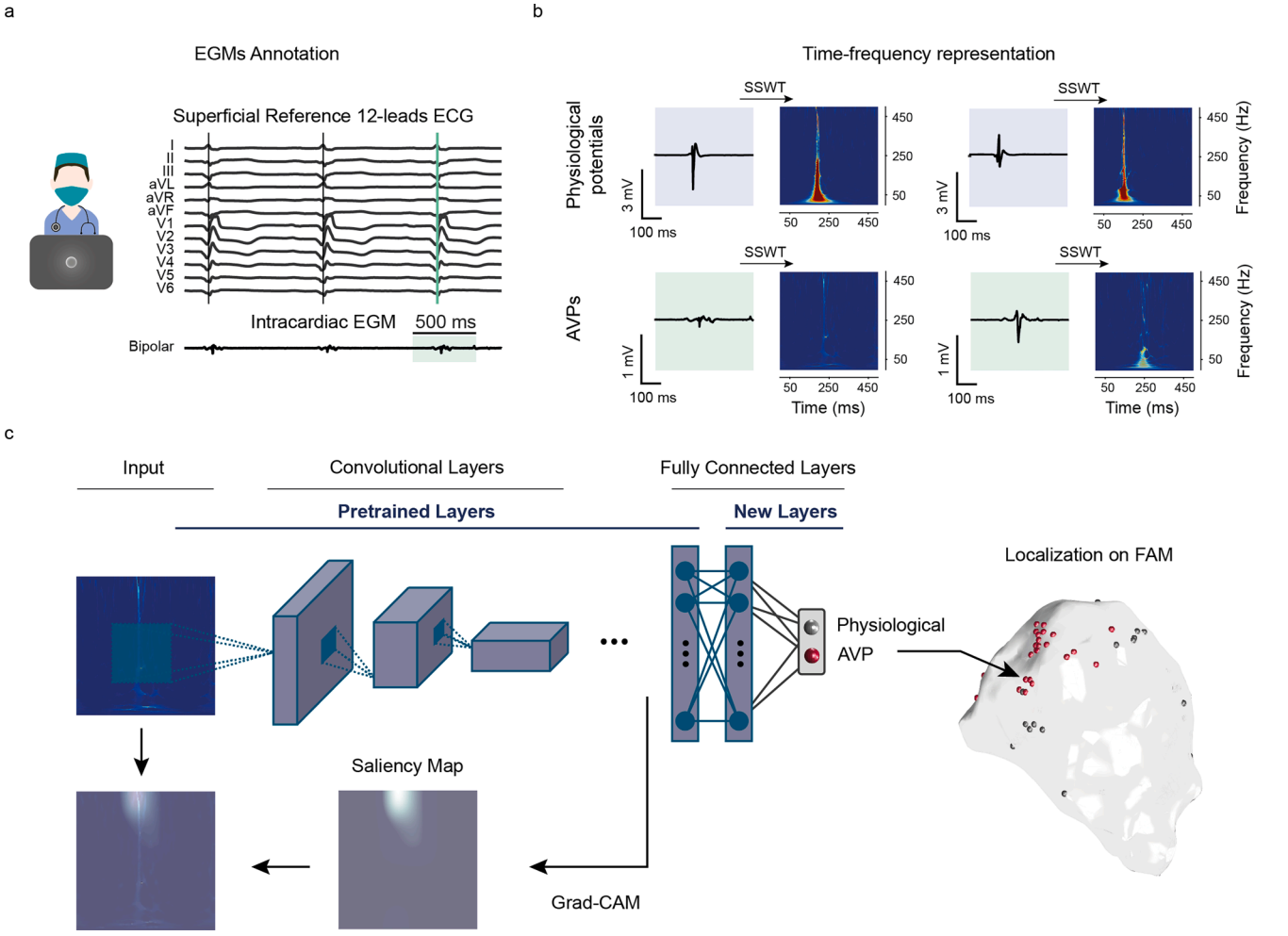


Fig. 1. Workflow. On the top left (a), an example of a multi-channel recording presented to the electrophysiologist during the annotation procedure is represented, involving superficial ECG leads and intracardiac EGMs; in the latter, the 500-ms window of interest is highlighted for subsequent time–frequency analysis. On the right (b), the image creation process is depicted. Finally, on the bottom (c), a visual explanation of the proposed approach for AVP detection is depicted: here, the transfer learning strategy scheme, along with a saliency map generation and a visualization of the CNN predictions onto a fast anatomical map (FAM), are reported.

2. Methods

The overview of the proposed approach is represented in Fig. 1. Automatic AVPs identification by CNN was modelled as a binary classification problem between AVPs and physiological potentials. To overcome the limitations associated with the small size and variety of instances in the used dataset, transfer learning has been adopted on three popular pre-trained CNNs (AlexNet [47], GoogLeNet [48] and VGG16Net [49]). Specifically, transfer learning involved updating the weights in pre-trained models, by submitting new training instances. A voting procedure was also carried out on the predicted labels to investigate the effect of the joint use of the three CNNs on the classification performance. In the following subsections, we outline the steps involved in the proposed methodology, including the preprocessing pipeline by using the SSWT, the training procedure of the CNNs, and the model explainability aimed at pointing out the most informative time–frequency ranges.

All the processing steps were performed in MATLAB R2022a (MathWorks Inc., MA, USA), and the DL training and test processes were carried out on a 64-GB NVIDIA Tesla P100 GPU cluster.

2.1. Dataset acquisition, labelling, and partition

The EAM procedures were carried out on nine post-ischemic VT patients between 2017 and 2018 at the San Francesco Hospital (Nuoro, Italy), by using the CARTO®3V6 mapping system. The study on anonymised data was approved by the Independent ATS (Azienda Tutela Salute Sardegna) Ethical Committee (Prot. n. 351/2021/CE, date of approval: 13/07/2021) and performed following the principles outlined in the 1975 Helsinki Declaration, as revised in 2000. All patients provided their informed consent.

Bipolar EGMs were acquired by PentaRay™ (Biosense Webster, Inc.) 2–6-2 mm by exploiting only the 2-mm spaced electrode pairs, as well as by ThermoCool SmartTouch® and ThermoCool SmartTouch® SF (Biosense Webster, Inc.) catheters. All the signals were recorded during sinus rhythm, using a sampling frequency of 1000 Hz, and were band-pass filtered between 16 and 500 Hz. Once exported, the EGMs were retrospectively blindly annotated by a single experienced electrophysiologist, using an ad-hoc MATLAB graphical user interface, as in [41]. Specifically, the electrophysiologist was unaware of the clinical case, the 3D EA maps, and the location of the EGM on the EA map during the annotation, so that the EGM labelling was carried out by relying on the morphological characteristics of the signal only. The annotation was performed by visually inspecting each single EGM along with the simultaneous surface ECG recordings and labelling as AVPs all abnormal potentials occurring during or after the corresponding QRS complex, as in [50]. Since the electrophysiologist was provided with the EGM and the corresponding surface ECG leads, to be able to identify the earliest and latest QRS activations, the annotation completely relied on the EGM morphology and its latency w.r.t. the surface QRS. This ensured any physiological early occurrences with high-frequency components (e.g., Purkinje potentials) to not be mistakenly labeled as AVP.

A different number of EGMs was collected and labelled as useful for each patient, according to both the signal quality and the specific procedure. Specifically, only the EGMs for which the electrophysiologist could confidently assign the label were included, while ambiguous or noisy traces were discarded. This preliminary data pruning ensured the investigation to be focused on reliable EGM morphologies only, at the expense of a reduced map density. Therefore, from the nine EAM procedures, a dataset composed of 1809 physiological EGMs and 752 AVPs (i.e., consisting of 84 ± 96 AVPs and 201 ± 222 physiological EGMs per patient, mean ± standard deviation) was obtained. Table 1 reports the numerosity of AVPs and physiological EGMs for each patient.

To evaluate the performance of the proposed approach, we conducted a stratified leave-one-subject-out (LoSO) cross-validation. Despite the limited dataset size, this condition is considered closer to a real application scenario, wherein a trained classifier is used on a completely unknown patient.

Table 1
Dataset information.

	Patient ID								
	1	2	3	4	5	6	7	8	9
Number of AVPs	83	99	67	84	324	36	41	13	5
Number of physiological EGMs	69	287	40	700	375	84	32	65	157
Sex (F/M)	F	M	M	M	M	M	M	F	M
Age (y.o.)	68	68	83	50	78	61	60	57	69
Ejection fraction (%)	30	32	42	30	27	31	21	28	24

To provide a balanced training set, at each cross-validation step (hence, for each patient to be tested according to the LoSO), all EGMs from eight out of nine patients were randomly downsampled to obtain an equal number of physiological potentials and AVPs. Therefore, the balanced training set A_i^{tr} used while testing the i_{th} patient has been obtained by the union of all the balanced EGMs subsets from the remaining participants, as:

$$A_i^{tr} = \bigcup_{j \neq i} A_j \tag{1}$$

$$A_j = \begin{cases} A_j^{AVP} \cup \widehat{A}_j^{Phy} \text{ if } N_{Phy} > N_{AVP} \\ \widehat{A}_j^{AVP} \cup A_j^{Phy} \text{ if } N_{Phy} < N_{AVP} \end{cases} \text{with } \mathrm{size}(A_j) = 2 \times \min\left(N_{Phy}, N_{AVP}\right) \tag{2}$$

Where $\widehat{A}_j^c$ with $c = \{AVP, Phy\}$ represents the downsampled subset of EGMs belonging to the c_{th} class for the j_{th} patient, N_{Phy} the number of physiological potentials and N_{AVP} the number of AVPs.

This balanced set A_i^{tr} was further divided into 80% for training and 20% for validation. Remarkably, the validation set was only used for the prevention of overfitting, by stopping the model training according to the early stopping approach, as detailed in **Section 2.3**. As such, this process resulted in a different number of training samples per each tested patient, starting from the available samples (both physiological EGMs and AVP) presented in Table 1, i.e., 1013, 965, 1060, 989, 605, 1066, and 1072 EGMs in the training sets when considering patient 1, 2, 3, 4, 5, 6 and 7 as testing sets, respectively, and 253, 241, 264, 247, 151, 266, and 268 EGMs in the validation sets of the same patients.

Conversely, all EGMs of the tested patient in the LoSO procedure were included in the test set only (so neither in the training set nor in the validation one) even if the numerosity of the instances in the two classes was unbalanced. However, to provide a more robust assessment, only patients with more than 30 potentials in the AVP and in the physiological classes were tested:

$$A_i^{test} = A_i^{AVP} \cup A_i^{Phy} \text{ if size}\left(A_i^{AVP}\right) > 30 \text{ and size}\left(A_i^{Phy}\right) > 30 \tag{3}$$

Due to this constraint, the patients with ID8 and ID9 in Table 1 were not considered in the test sets. This choice was made to obtain classification metrics on a reliable statistical sample, even though typically imbalanced, as no balancing was pursued on the tested patient.

2.2. Image generation

The SSWT [45] was adopted to enhance the time–frequency localization of relevant features on signals with rapidly changing or localized frequency contents. In fact, the SSWT is recognized as one of the most effective techniques for producing robust time–frequency representations, by highlighting the presence of a distinctive frequency signature [50], while simultaneously providing accurate information from both

the time and frequency domains. It has been extensively and successfully applied in various biomedical contexts due to its ability to yield sharper and precise time–frequency representations through the reassignment method, outperforming simpler methods like traditional wavelet transforms [45]. Specifically, in SSWT, the continuous wavelet transform (CWT) is firstly computed as [51]:

$$W_x(a,b) = \int x(t) a^{-1/2} \psi^* \left(\frac{t-b}{a} \right) dt \tag{4}$$

where x is the signal, a the scale, b the time shift, and ψ is an appropriately chosen mother wavelet. In this study, the analytic Morlet wavelet was selected to emphasize the fragmented peaks of the signals. In the second step, the reassignment rule is applied through synchrosqueezing, which exploits the phase information embedded in the wavelet coefficients W_x. Synchrosqueezing is a process used to estimate instantaneous frequency by relying on the phase derivative. This information is then used to redistribute the wavelet coefficients' energy, i.e., transferring the information from the time-scale to the time–frequency domain. Indeed, from the CWT, the instantaneous frequency $\omega_x(a,b)$ can be computed as [52]:

$$\omega_x(a,b) = -i(W_x(a,b))^{-1} \frac{\partial}{\partial b} W_x(a,b) \tag{5}$$

and finally, the synchrosqueezed transform may be computed as:

$$T_x(\omega_x, b) = \int_{A(b)} W_x(a,b) a^{-3/2} \delta(\omega_x(a,b) - \omega) \, da \tag{6}$$

where $A(b)$ identifies the domain in which the wavelet coefficients $W_x(a, b)$ are different from zero.

As such, for the input image generation, firstly we extracted a 500-ms window around the reference annotation for each EGM, considering 200 ms before and 300 ms after such a reference. Then, we generated the RGB images by computing the modulus of the SSWT, to have a representation invariant with respect to the polarity of the signal, plotting the results setting the "jet" colormap and removing the axes (see Fig. 1**(b)**). Specifically, since "jet" is a conventional multi-gradient colormap that exploits a wide range of colors, it might reduce redundancy between layers of the RGB image, thus potentially supporting the identification of distinct and meaningful patterns and features. While recognizing that using a different colormap might have an impact on the image generation and then on the models' performance, no optimization was performed in this regard, to maintain greater generalizability of the models. Next, two images per signal were saved, with slightly different size according to the input layer of the different CNNs (i.e., 227x227 pixels for AlexNet, 224x224 pixels for GoogLeNet and VGG16Net).

During the image generation, the saturation limit of the colormap was fixed using a value obtained in a preliminary analysis, aimed at optimizing the AVPs detection rate (see **Fig. S1** in Supplementary Material for additional information).

2.3. Network training

The AlexNet, GoogLeNet, and VGG16Net DL models, pre-trained on ImageNet [53], were used in this study, adapting them to a binary classification problem to distinguish between physiological potentials and AVPs. As such, the final classification layers (i.e., the last fully connected, the softmax, and the classification output layers) have been changed in each CNN. In particular, the fully connected layer was replaced by a new fully connected layer of the same size, with randomly initialized weights obtained by using the Glorot uniform initialization method [54], while the size of both softmax and classification layers was set equal to two, for the binary classification (see Fig. 1**(c)**). To enable a faster learning in the new fully connected layer, we increased both the learning rate for the weights and bias by a factor of 10, compared to the transferred layers.

The decision to adopt transfer learning was supported by a series of preliminary empirical evaluations. A preliminary comparative analysis between the classification performance achieved through transfer learning and that obtained by training the models from scratch using random weights initialization across the three selected CNNs indicated a significant improvement with transfer learning. This improvement was accompanied by a reduction in computational costs (see Section II in Supplementary Material).

The models were trained using the stochastic gradient descent with momentum algorithm and the binary cross-entropy as loss function, by considering a pre-established set of values for each hyperparameter undergoing optimization. The hyperparameters associated with the model training, i.e., learning rate, learning rate decay factor, momentum, and batch size, were tuned through a conventional grid search (see Section II in Supplementary Material for further details). Besides hyperparameter optimization, we decided to prevent the overfitting by introducing an early stopping condition as training option. In this regard, by considering a trade-off between the computational time needed for the training and the best model identification, a validation frequency of 50 was chosen. Furthermore, by leveraging a preliminary analysis of the loss function, the validation patience was set equal to 9 steps, while the maximum number of epochs was fixed to 1000, to accommodate a large number of iterations if a long training is needed.

2.4. Performance assessment

The CNNs performance was assessed in terms of accuracy (Acc), true positive rate (TPR), true negative rate (TNR), precision (or positive predictive value, PPV) and F1 score, by using a balanced classification threshold at 0.5, representing the case where equal risk is assigned to both classes. Hereinafter AVPs were associated with the positive class whereas the physiological potentials to the negative class. Moreover, receiver operating characteristic (ROC) and the area under the curve (AUC) were analyzed during performance assessment. Specifically, a global ROC was computed for each CNN by averaging the patient-specific ROCs after interpolation (see Fig. 3 **(a)**). This approach was chosen as it provides a better understanding of how the model performs across different patients while avoiding the potential bias introduced by patients with a larger number of samples.

In addition to training and testing the three CNNs separately, we also explored the performance of a multiple classifier based on a majority voting on the CNNs' predictions. Finally, the spatial localization of both correct and incorrect classifications was visually inspected on the fast anatomical map (FAM) colored according to the voltage and the local activation time, by considering the predictions of the CNN exhibiting the best performance.

2.5. Explainability and saliency maps analysis

The Grad-CAM technique [46] was used to study the underpinnings behind the predictions provided by the three CNNs in their best settings. This aspect is relevant to build a more trustable AI-based system in terms of explainability, and to compare the model outcomes with current knowledge on EGM morphology and spectral characteristics.

By exploiting the class-specific gradient information $\frac{\partial y^c}{\partial A^k}$, where y^c is the score computed before the softmax for the class c, and A^k is the k-th layer of feature map A in the final convolutional layer, the Grad-CAM produces a saliency map that provides a coarse localization of the most informative regions in the input image, as follows:

$$L^c_{Grad-CAM} = ReLu\left(\sum_k \alpha^c_k A^k \right) \tag{7}$$

where α^c_k represents the importance of each feature map layer k, A^k, for

the class c, and it is computed as:

$$\alpha_k^c = \frac{1}{Z}\sum_i\sum_j \frac{\partial y^c}{\partial A_{ij}^k} \tag{8}$$

where Z is the number of pixels in A^k, and i and j are the indexes for the x and y direction in A^k.

To assess a $L^c_{Grad-CAM}$ map for each class, the following algorithm was implemented. Indeed, for the n-th CNN architecture and the p-th patient, we considered the common occurrences t that were correctly classified by all the CNNs and computed a set of maps $L^c_{n,p,t}$. This step allowed considering only the maps representing useful patterns for the overall correct prediction. At first, by linear interpolation, we resized every saliency map $L^c_{n,p,t}$ to 500 × 500 pixels. Such a resizing procedure allows studying the maps, which are referred to SSWT images, with a unitary resolution on both frequency and time axes (i.e., 1 Hz and 1 ms, respectively), according to the sampling frequency and the duration of the signals. Since Grad-CAM is sensitive to the amplitude of the pixels composing the input image, we normalized each saliency map $L^c_{n,p,t}$ by its standard deviation, thus unbiasing the resulting maps. Next, we performed an intra-subject weighted ensemble average on the normalized maps $L^c_{n,p,t}$ corresponding to each class c, for the n-th CNN architecture. This process results in a patient-specific map for each class $L^c_{n,p}$. At this stage, weighting was performed exploiting the softmax output, which gives the model probability ρ_t^c for each instance t to be assigned to the class c, as:

$$L^c_{n,p} = \sum_t L^c_{n,p,t}\rho_t^c \tag{9}$$

This intra-subject weighting was introduced to favor the images in which the models were more certain about the prediction, thus highlighting the saliency maps where the model was strongly confident about the classification, while hindering those whose outcome was doubtful.

Finally, to obtain a single saliency map L_n^c for each class c and for the n-th CNN architecture, we computed the weighted average of $L^c_{n,p}$ across the patients. This inter-subject weighting was performed according to the patient's true prediction rate of each class TR_p^c (i.e., the TPR and the TNR for the AVPs and the physiological potentials maps, respectively), as:

$$L_n^c = \sum_p L^c_{n,p} TR_p^c \tag{10}$$

This inter-subject weighting was introduced to identify the time–frequency regions that resulted to be more effective in predicting the right outcome for the n-th CNN architecture, by emphasizing the patient-related salience maps $L^c_{n,p}$ with high performance.

In each weighting (i.e., intra- and inter-subject), each resulting map was normalized by its standard deviation. A flow-chart of the whole algorithm followed for the saliency maps generation is presented in Fig. 2.

Finally, to further define the ranges of interest in both time and frequency domains, we included two additional plots alongside the saliency maps. These plots represent the average contribution across frequencies and across time, respectively. The ranges of interest were identified by applying a threshold equal to the median of the values on each side-plot.

3. Results

Considering the higher performance when transfer learning is adopted in the training stage (see **Fig. S2** in Supplementary Material), the Tables 2 and 3 provide the results for each of the three CNNs with transfer learning and hyperparameter optimization, as well as for the voting of their predictions, in terms of overall results and patient-related evaluations.

As can be seen from Table 2, the three networks showed Acc values above 84.5%, up to 89% for AlexNet, even if AlexNet exhibited higher variability (i.e., with standard deviation of 10.6%) compared to the other CNNs.

Overall, TPR was lower than TNR for all the CNNs, with values ranging from 80.6% by GoogLeNet to 88.6% by AlexNet. Moreover, in accordance with the other metrics, AlexNet exhibited the best performance also in terms of F1 (0.87). Conversely, the VGG16Net outperformed the other CNNs in terms of TNR and PPV (90.3% and 89.5%, respectively).

Contrarily to expectations, the voting approach failed in leading to better classification performance. Indeed, although yielding to reasonable efficacy, it did not outperform the AlexNet performance, except for TNR.

On the other hand, when examining the patient-specific performance in Table 3, it is evident that the proposed approaches showed good results on all tested subjects, with the exception of the one identified by ID 5. There, a significant performance drop emerged for TNR regardless of the chosen CNN, thus leading to lower Acc and to large standard deviation values in Table 2.

The results reported in Tables 2 and 3 are further confirmed by the ROCs and AUC values showed in Fig. 3**(a)**. Indeed, AlexNet and VGG16Net exhibited sharper ROC curves, which closely approach an ideal classification performance and higher AUC values (0.96 and 0.95, respectively), compared to GoogLeNet.

Fig. 4 presents the distribution of AVPs and physiological potentials on the local activation time (LAT) (Fig. 4**(a)**) and the greyscale voltage maps (Fig. 4**(b)**) of three exemplary patients (i.e., ID 1, 2 and 5). Fig. 4 **(b)** also shows the predicted class by AlexNet for every EGM. As both the scientific literature and the clinical practice show that thresholds on EGMs amplitude are still highly valued to guide ablation procedures, the Pearson's correlation coefficient between the probability prediction scores of the three CNNs and the EGMs amplitude was computed to investigate possible biases in the prediction due to the amplitude information. This investigation led to a correlation coefficient of −0.04,

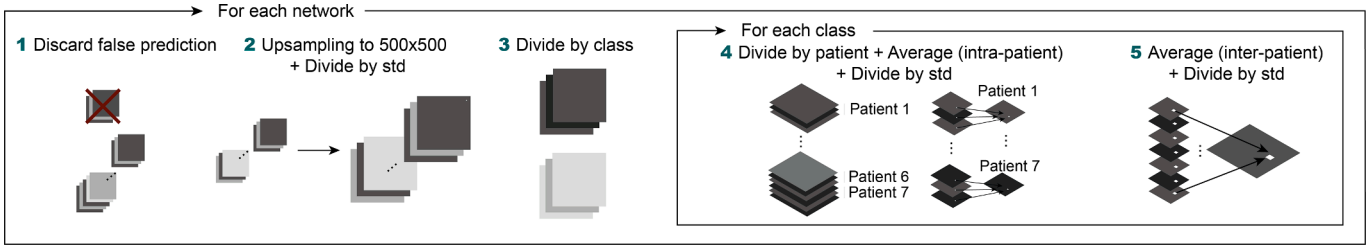


Fig. 2. Flow diagram for saliency maps computation. From left to right, in step 1, the saliency maps related to the false prediction in at least one CNN were discarded. Then, in step 2, images were resized to obtain a unitary resolution in time (1 ms) and frequency (1 Hz) domains. In step 3, the maps were grouped by class (i.e., AVP or physiological). After this, for each class, the maps were grouped per subject, and a weighted ensemble average was performed using the model probability score (step 4). Finally, at step 5, on the patient-related maps, a weighted ensemble average was performed using TPR (for AVPs) and TNR (for physiological class).

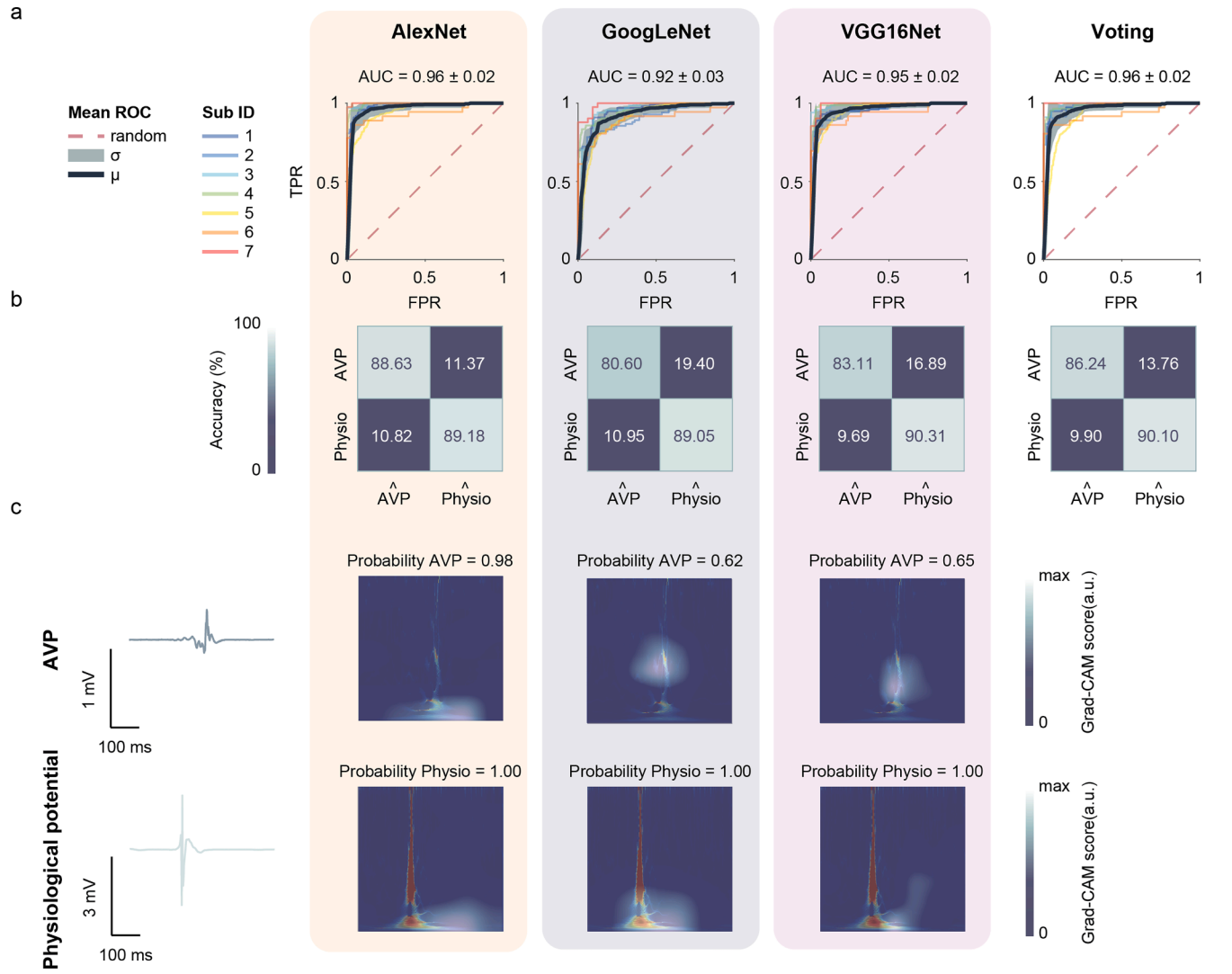


Fig. 3. Classification performance by ROCs, AUCs, confusion matrices and Grad-CAM maps. At the top (a), ROC curves for the different patients (sub) and CNNs. The patient-specific and average ROC curves are depicted, along with their dispersion. In the middle (b), confusion matrices are reported for the different CNNs and the voting of their predictions. Finally at the bottom (c), selected instances from the testing dataset are illustrated with Grad-CAM maps. These maps highlight the regions in the input images that contribute the most to the classification decisions made by the models.

Table 2
Overall classification results. Results are reported as mean ± standard deviation across the different patients.

	Acc (%)	TPR (%)	TNR (%)	PPV (%)	F1
AlexNet	89.0 ± 10.6	88.6 ± 9.4	89.2 ± 22.8	89.2 ± 15.9	0.87 ± 0.07
GoogLeNet	84.8 ± 5.9	80.6 ± 11.6	89.1 ± 11.6	81.6 ± 14.8	0.80 ± 0.06
VGG16Net	88.5 ± 9.0	83.1 ± 16.8	90.3 ± 18.6	89.5 ± 14.4	0.84 ± 0.09
Voting	88.9 ± 9.1	86.2 ± 13.6	90.1 ± 19.1	89.1 ± 14.4	0.86 ± 0.07

−0.07, and −0.04 for AlexNet, GoogLeNet, and VGG16Net, respectively, thus excluding any potential bias in this regard. Indeed, some false AVPs exhibited quite high peak-to-peak amplitude, whereas physiological signals tend to reveal low amplitudes; in contrast, signal fractionation could have been led to false AVP detections, possibly due to its signature in higher frequency band.

Although the performance of the different CNNs was quite comparable, the training time was dramatically different. Indeed, on a high-end 64-GB NVIDIA Tesla P100 GPU, training time lasted (mean ± standard deviation): 70 ± 10 s for AlexNet, 191 ± 34 s for GoogLeNet, and 333 ± 36 s for VGG16Net, with a number of performed training iterations equal to 714 ± 107, 686 ± 131, and 607 ± 73, respectively.

3.1. Saliency maps analysis results

To better understand the learning outcome and the decision process of the different CNNs, we examined the saliency maps L_n^c for both classes c. In this regard, Fig. 5 shows the range of highest saliency identified in the input images, as highlighted by the side-plot. For physiological EGMs, the most useful contributions were coherently localized by all the CNNs in the lower part of the spectrum, without any component above 223 Hz (Fig. 5 **(d)-(f)**). Conversely, the range of interest for AVPs was less consistent across the CNNs, but generally including a higher part of the spectrum too, compared to the physiological potentials (up to over 440 Hz). This aspect is particularly evident for GoogLeNet and VGG16Net in Fig. 5 (i.e., comparing **(b)** vs. **(e)**, and **(c)** vs. **(f)**). Overall, the frequency ranges of interest were limited to a compact interval around a single maximum frequency, leading to a unimodal curve in the side-plot for all the CNNs (Fig. 5 **(b), (c)**) except for AlexNet, which

Table 3
Patient-specific classification results.

		Patient ID						
		1	2	3	4	5	6	7
Acc (%)	**AlexNet**	88.8	91.7	86.9	98.1	66.8	93.3	97.3
	GoogleNet	80.9	85.5	79.4	94.1	78.3	85.0	90.4
	VGG16Net	80.9	90.9	93.5	96.7	72.5	87.5	97.3
	Voting	82.2	92.2	90.7	97.2	71.8	90.8	97.3
TPR (%)	**AlexNet**	79.5	96.0	79.1	90.5	100	77.8	97.6
	GoogleNet	65.1	90.9	68.7	84.5	92.6	72.2	90.2
	VGG16Net	65.1	92.9	89.6	76.2	99.7	58.3	100
	Voting	67.5	96.0	85.1	85.7	100	69.4	100
TNR (%)	**AlexNet**	100	90.2	100	99.0	38.1	100	96.9
	GoogleNet	100	83.6	97.5	95.3	65.9	90.5	90.6
	VGG16Net	100	90.2	100	99.1	49.1	100	93.8
	Voting	100	90.9	100	98.6	47.5	100	93.8
PPV (%)	**AlexNet**	100	77.2	100	91.6	58.3	100	97.6
	GoogleNet	100	65.7	97.9	68.3	70.1	76.5	92.5
	VGG16Net	100	76.7	100	91.4	62.8	100	95.4
	Voting	100	78.5	100	87.8	62.2	100	95.4
F1	**AlexNet**	0.89	0.86	0.88	0.91	0.74	0.88	0.98
	GoogleNet	0.79	0.76	0.81	0.76	0.80	0.74	0.91
	VGG16Net	0.79	0.84	0.95	0.83	0.77	0.74	0.98
	Voting	0.81	0.86	0.92	0.87	0.77	0.82	0.98

showed two peaks above the threshold for the AVPs class (Fig. 5 **(a)**). Furthermore, from the time-domain perspective, the AVP class was characterized by a lower dispersion compared to the physiological class.

4. Discussion

In this work, the development of a DL approach to support clinicians in the automatic identification of AVPs and physiological potentials in EGMs acquired during EAM in sinus rhythm from post-ischemic VT patients has been introduced and deeply analyzed.

To generate detailed and robust time–frequency representation for each EGM to be provided as input to the CNNs, the SSWT was used because of its higher performance compared to other simpler transforms. In fact, the short-time Fourier transform is constrained by the uncertainty principle due to its fixed window duration, while the CWT, despite mitigating some of these issues, still results in blurred or smeared signal representations in the time–frequency domain and exhibits significant edge effects [55], which can be overcome by the SSWT through the reassignment method along the instantaneous frequency components.

Since an in-target pre-training was hampered by the unavailability of large amount of annotated data from VT EAM procedures, in this work we investigated the adoption of transfer learning from a generic domain. As also demonstrated in this work (see Supplementary Material), transfer learning can improve performance when employing pre-trained CNN architectures on a different target domain, regardless of their complexity and depth. This is because certain knowledge may be shared across domains, thus enhancing generalization [56]. As such, we choose CNNs trained on large general-purpose image datasets, being able to recognize different types of patterns, that might be difficult to learn from scratch. This prior knowledge can advantage the extraction of features that might be highly relevant to the detection of the AVPs. Indeed, in our work, the pretrained CNNs led to an increased accuracy and TPR with respect to the use of non-pretrained models, but contrarily affected the TNR. Despite this trend, the selectivity remains a less relevant performance metric in our clinical application compared to sensitivity and

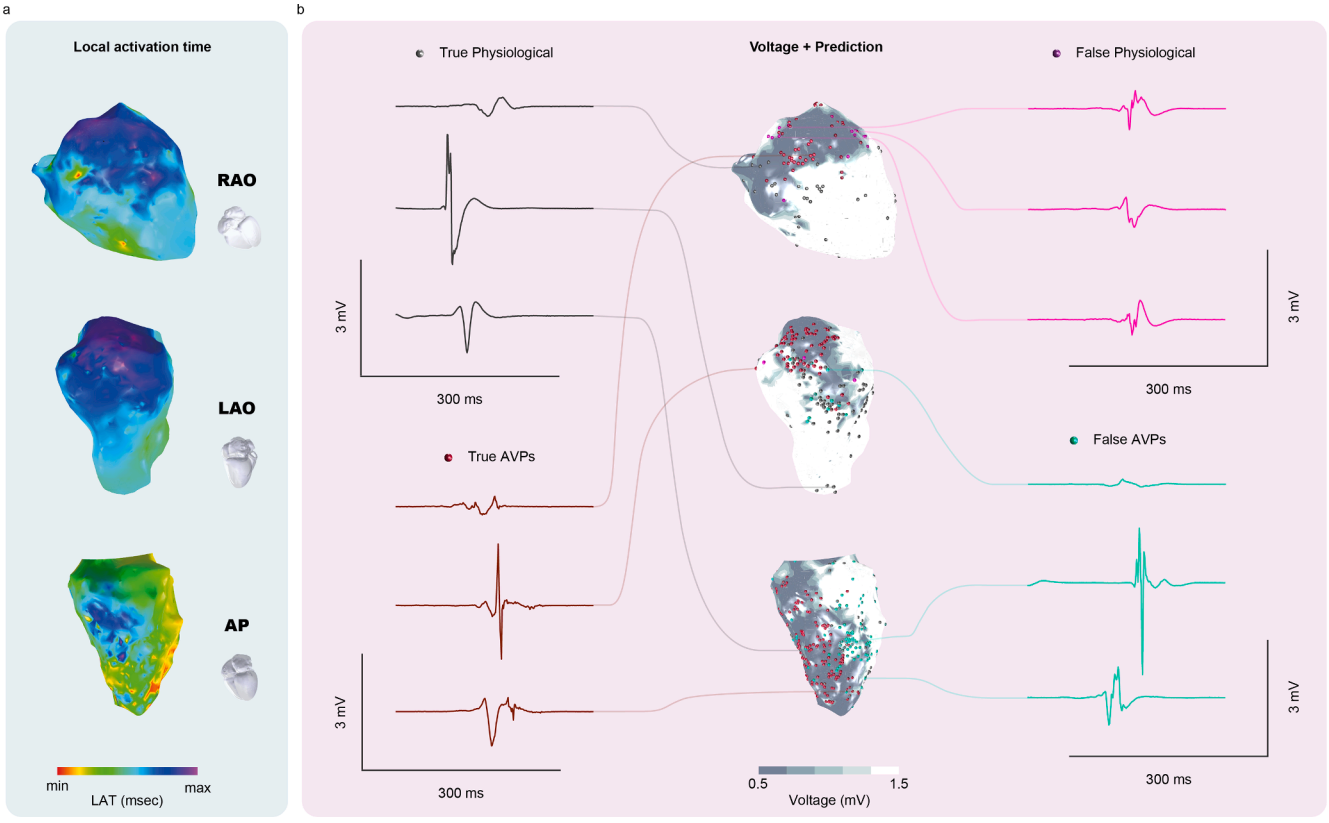


Fig. 4. EAM and classification maps for three different patients (i.e., from left to right ID 1, 2 and 5). From top to bottom, the local activation time (LAT) maps (a) and grey-scale voltage maps (b) with the localization of the predicted labels (both correct and incorrect predictions) by the best-performing CNN (i.e., AlexNet) are reported. Some examples of correctly classified and misclassified signals for both the physiological and AVP classes are presented.

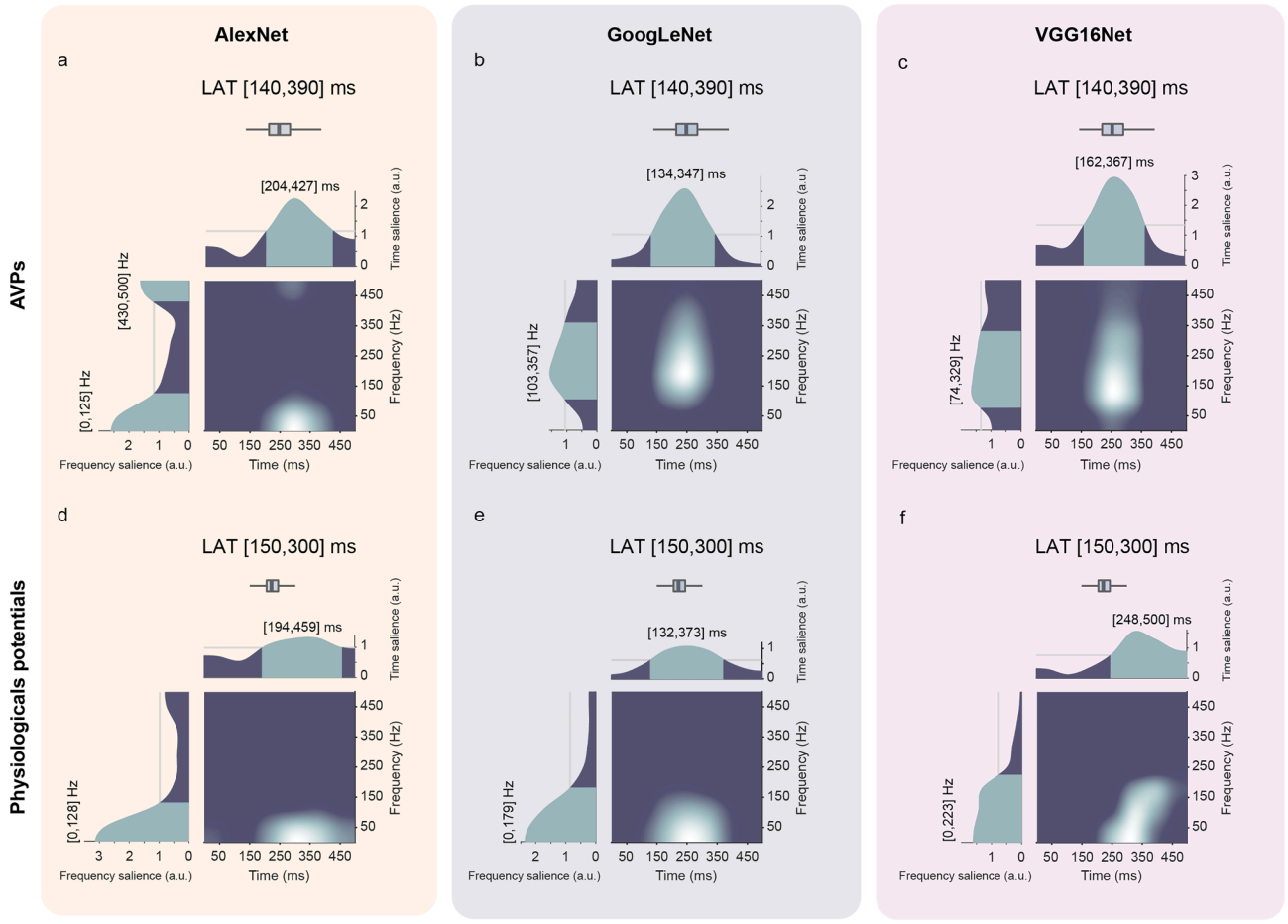


Fig. 5. Results of the saliency maps analysis. Saliency maps for the AVPs (a-c) and physiological potentials (d-f) by the three different CNNs with their best settings. On the top of each map, a side-plot represents the average contribution across frequencies, allowing to identify the ranges of interest (in squared brackets) in the time domain; similarly, on the left of each map, a side-plot represents the average contribution across time, to identify the ranges of interest (in squared brackets) in the frequency domain.

accuracy. Considering the aim of the study, which is the recognition of VT arrhythmogenic sites identified by AVPs, the network reporting the highest TPR, would be preferable. Indeed, a higher TPR indicates that the network reliably identified AVPs, while a few AVPs were missed, ensuring effective alerts to the clinicians, while a low TNR would reflect a high number of physiological EGMs misrecognized as AVPs, resulting in more frequent false alarms. Therefore, after testing both pre-trained and not pre-trained models and evaluating their pros and cons, we concluded that pretrained models are preferable, since led to a substantial increase in the metrics mostly useful in this application scenario. Furthermore, the adoption of transfer learning significantly reduced the number of training iterations, regardless of the CNN (see **Table S1** in Supplementary Material), thus saving computation time.

For the same reasons, we identified AlexNet as the best performing network in this scenario, since it reported the highest TPR and accuracy among the three examined pre-trained CNNs. Remarkably, the classification threshold was not optimized in this study, even though its tuning would have allowed to a further increase in TPR (AUC = 0.96, see Fig. 3 **(a)**) at the expense of a reduced generalizability, also considering the limited dataset size.

Overall, the results revealed higher accuracies in the physiological potential recognition than for the AVPs (i.e., TNR > TPR for the three CNNs). A possible explanation is that physiological potentials exhibit a quite invariant signature in the SSWT representation across different patients, whereas the same is not true for AVPs. In fact, according to the literature, AVPs can be categorized in different sub-classes as per their onset and duration [57], which introduces further variability that was not considered in this work, nor stratified for different folds and patients in the cross-validation.

Moreover, high standard deviations could reflect the different number of instances for each patient, in terms of number of EGMs acquired in the single EAM procedure and in terms of number of EGMs labelled as AVPs or physiological, that led to slight performance differences according to the tested patient, as reported in Table 3. Indeed, different patients could exhibit a different number and type of labelled AVPs, along with intrinsic differences in terms of scar morphology, deepness, and localization onto the ventricle. This fact is particularly visible for patient ID 5, for which an overall performance drop can be seen for all the CNNs (see Table 3), especially in terms of TNR (i.e., in the recognition of physiological EGMs). Despite the aim of the study is to provide an effective detection of arrhythmogenic sites, and, therefore, TPR plays a relevant role, this patient deserves further investigation. Remarkably, patient ID 5 exhibited a high (and approximately balanced) number of labelled EGMs (see Table 1), and an outstanding number of AVPs with respect to the other examined clinical cases. Accordingly, when this patient was tested in the LoSO cross-validation, the numerosity of the training samples dramatically dropped to 605 EGMs (see **Section 2.1**), being that patient the main contributor of AVPs. Such a drop in the number of AVP instances (and so of physiological EGMs too) used for training could have affected the generalization capabilities of that models. Furthermore, by looking at the related voltage map in Fig. 4 **(b)**, it is clear that this patient also showed a very extended scar with a strongly complex morphology: therefore, most of its EGMs have been acquired in highly damaged myocardial areas. This aspect could have

possibly led to patient-specific characteristic morphologies in its physiological EGMs, mainly related to highly scarred areas. Furthermore, looking at the classification results shown in Fig. 4**(b)**, in this case false detections mainly occurred in the most irregular zones, i.e., the boundaries of scar. This hypothesis was also supported by the related saliency map, which showed different ranges of interest in the frequency domain for this patient, when classifying physiological EGMs (data not shown). Interestingly, upon analyzing the saliency maps by merging the ranges resulted by the three CNNs, it was discovered that the lower spectral contents (below 128 Hz) provided the most informative cues for identifying the physiological potentials. Conversely, for the AVP identification, the CNNs found higher frequency contributions more relevant. In particular, the frequency range between 103 Hz and 125 Hz was generally highly considered by all the CNNs, which is in agreement with the literature in the field [50], describing the main frequency contribution for AVPs at higher frequencies (i.e., between 40 Hz and 320 Hz) compared to physiological potentials (i.e., below 20 Hz). Although the discriminating frequency ranges highlighted by the saliency maps overlap, the CNNs prioritized also lower frequency components for physiological EGMs compared to AVPs, since for the latter the frequencies below 103 Hz were not considered significant for classification. However, AlexNet and VGG16Net also focused on higher frequencies to identify AVPs, up to over 440 Hz. Furthermore, looking at the same results in the time domain, some interesting aspects could be also highlighted. As shown in Fig. 5, AlexNet used to focus on the latest portion of the EGM, detecting eventual signatures after the far field activation (see Fig. 5 **(a)** and **(d)**). Conversely, both GoogLeNet and VGG16Net generally focused more on the far field activation, with a single sporadic exception (see Fig. 5 **(f)**). This aspect contributes to the explanation of the different frequency ranges considered by AlexNet, since such model considered different portions of the EGMs in the discrimination, which might have been associated with different spectral contents.

Unfortunately, the saliency maps themself do not elucidate the influence of signal amplitude on the classification process. Interestingly, a complete lack of correlation between the AVPs predictions and scar areas, characterized by low-amplitude EGMs, was found. This was further confirmed by a closer examination of the reported EGMs (see Fig. 4.**b**), in which both AVPs and physiological EGMs exhibited either high or low amplitudes.

This study further proves the effectiveness of AI-based solutions in cardiac electrophysiology. However, a point performance comparison with pre-existent literature cannot be easily performed, according to the specific application. Indeed, some AI-based solutions were proposed so far for the localization of ablation targets for AF electrophysiological procedures, as [34], in which the deep neural network achieved 95% of accuracy in distinguishing rotational from non-rotational images in AF, outperforming other ML techniques tested on the same dataset. Furthermore, similar results have been obtained in [35], when using CNNs to identify ablation targets by spatiotemporal dispersion in atrial recordings, and in [36], where a commercial ML tool called VX1 (Volta Medical, Marseille, France) was exploited to localize the spatiotemporal dispersion site in the atria from the analysis of the EGM, exhibiting high effectiveness in ablation sites localization for AF termination occurrence (with accuracies of 92% and 83% in patients belonging to the primary and satellite centers, respectively). Focusing on post-ischemic VTs, some studies investigated the identification of arrhythmogenic areas by defining cut-off thresholds on features or criteria and evaluated the performance of their algorithms by assessing the proportion of identified areas w.r.t. scar areas [22,23], successful ablation sites [19] or areas with LAVA, VT isthmuses or late potentials [17]. Moreover, some studies (e.g., [22,25,23]) validated their strategies on VT patients and a non-VT control group. However, these approaches are intrinsically not comparable with ours, whose aim is to distinguish between AVPs and physiological EGMs in the same post-ischemic VT population. Furthermore, some other works aimed at detecting sites of VT origin [20] or VT isthmuses [22–24,17]. However, the different research objective of these studies further hampers the comparison with our DL approach, which aims at detecting AVPs to assist clinicians in determining arrhythmogenic sites (which are not strictly limited to VT isthmuses nor coincide with sites of VT origin).

Regarding more similar approaches, a previous study [41] reported a comparison of three different ML methods for supervised classification of AVPs and physiological potentials in bipolar EGMs, exploiting multi-domain features. In such a work, an accuracy of 84.7% was reported when using the best performing model in a LoSO cross-validation, with a different efficacy in AVPs and physiological potential recognition (TPR = 79.4%, TNR = 88.1%). Nonetheless, the CNN models presented in this work, demonstrated higher performance compared to such ML-based methods exploiting simpler temporal and spectral features: indeed, by considering the best-performing approach in both studies, despite an improvement in accuracy is slightly perceived (+1.5%), a stronger impact is seen on TPR (+13.5%), but not in TNR (−5%). Indeed, since this study aims at providing a computer-aided system based on DL to detect arrhythmogenic sites, these findings are very encouraging, especially in TPR. Interestingly, in this previous study [41], the most valuable spectral features (highlighted by the feature selection) covered most of the 20-Hz sub-bands within the 80–240 Hz range (albeit using different methodologies for spectral estimation), which partially aligns with the ranges identified by our saliency maps.

Although the reported performance of our study is promising, the restricted dataset size is a limitation that suggest taking some care in generalizing the achieved results. In fact, despite the adoption of transfer learning, which reduces the impact of the small number of instances provided for training, a larger dataset would allow for a better generalization. In this sense, a larger dataset would guarantee an improved robustness, which should lead to more comparable results between different subjects. Additionally, a larger dataset would allow to have a separate cohort of patients for the validation set when applying the early-stopping criterion, to further avoid any bias for the model even during the training phase. We also acknowledge that the presented DL approach has not been validated using external dataset. However, currently there are no freely accessible datasets on post ischemic VT in the literature that would allow for a comparison. Indeed, most studies use nonpublic datasets, making comparative analysis between different algorithms impossible.

Another limitation is the labelling of the available dataset by a single cardiologist, albeit done in a blinded manner with respect to algorithm development and tuning. This labeling process could potentially introduce bias to the results. Remarkably, these limitations are shared with the most similar studies in the field.

5. Conclusions

This work introduced an AI-driven tool to support clinicians in identifying AVP during EAM procedures for post-ischemic VT. To achieve this goal, we proposed a DL approach that explores three different pre-trained CNNs and using information from EGMs, as they are routinely acquired during EAM procedures. Our approach leverages the time–frequency signatures of EGMs by the SSWT, resulting in a high accuracy and AVP detection rates, both close to 90%, on the available dataset. In particular, AlexNet provided the best performance, with better results also compared to a majority voting procedure performed on the prediction of all the trained CNNs. Remarkably, to investigate the reasons behind the models' prediction, we exploited the saliency maps, which highlighted the most important regions of interest for the CNNs. Since this approach seems to offer precise detection of AVPs, and so arrhythmogenic sites, this achievement paves the way to the introduction of DL for VT procedures, with potential implications in the EAM procedure acceleration and success.

Future studies that reveal the underpinnings of the reentrant mechanisms, or the use of masks on the feature maps based on the saliency

maps results, may enhance the predictive capabilities of the proposed AI tools. These studies could provide additional information to inject into the model, which alongside with signal localization on EA map, could exploit the electrical properties of the substrate.

CRediT authorship contribution statement

Andrea Pitzus: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Giulia Baldazzi:** Writing – review & editing, Writing – original draft, Visualization, Project administration, Methodology, Formal analysis, Data curation, Conceptualization. **Marco Orrù:** Writing – review & editing, Writing – original draft, Visualization, Software, Formal analysis, Data curation. **Luigi Raffo:** Writing – review & editing, Resources. **Graziana Viola:** Resources, Investigation, Data curation. **Petar M. Djurić:** Writing – review & editing, Supervision, Methodology, Formal analysis, Conceptualization. **Danilo Pani:** Writing – review & editing, Writing – original draft, Supervision, Resources, Methodology, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The authors wish to thank Mr. A. Valdes Rey for his contributions in the preliminary development of this study.

This work has been carried out in the framework of the Ablation Reinforcement by computer-aided Guidance and Optimization (ARGO) study, approved by Azienda Tutela Salute Sardegna (ATS Sardegna). The research leading to these results has received funding from the European Union - NextGenerationEU through the Italian Ministry of University and Research under PNRR - M4C2-I1.3 Project PE_00000019 “HEAL ITALIA” to G. Baldazzi, CUP F53C22000750006. The views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bspc.2024.106844.

References

[1] J.-T. Gräsner, et al., European resuscitation council guidelines 2021: epidemiology of cardiac arrest in Europe, Resuscitation 161 (2021) 61–79, https://doi.org/10.1016/j.resuscitation.2021.02.007.
[2] A. Timmis, et al., European Society of Cardiology: cardiovascular disease statistics 2021, Eur. Heart J. 43 (8) (2022) 716–799, https://doi.org/10.1093/eurheartj/ehab892.
[3] A. Natale, et al., Venice chart international consensus document on ventricular tachycardia/ventricular fibrillation ablation, J. Cardiovasc. Electrophysiol. 21 (3) (2010) 339–379, https://doi.org/10.1111/j.1540-8167.2009.01686.x.
[4] E.J. Ciaccio, et al., Structure and function of the ventricular tachycardia isthmus, Heart Rhythm 19 (1) (2022) 137–153, https://doi.org/10.1016/j.hrthm.2021.08.001.
[5] R. Tung, et al., Freedom from recurrent ventricular tachycardia after catheter ablation is associated with improved survival in patients with structural heart disease: an International VT Ablation Center Collaborative Group study, Heart Rhythm 12 (9) (2015) 1997–2007, https://doi.org/10.1016/j.hrthm.2015.05.036.
[6] B.K. Martinez, et al., Systematic review and meta-analysis of catheter ablation of ventricular tachycardia in ischemic heart disease, Heart Rhythm 17 (1) (2020) e206–e219, https://doi.org/10.1016/j.hrthm.2019.04.024.
[7] J.L. Sapp, et al., Ventricular tachycardia ablation versus escalation of antiarrhythmic drugs, N. Engl. J. Med. 375 (2) (2016) 111–121, https://doi.org/10.1056/NEJMoa1513614.
[8] S.G. Priori, et al., 2015 ESC Guidelines for the management of patients with ventricular arrhythmias and the prevention of sudden cardiac death: The Task Force for the Management of Patients with Ventricular Arrhythmias and the Prevention of Sudden Cardiac Death of the European Society of Cardiology (ESC) Endorsed by: Association for European Paediatric and Congenital Cardiology (AEPC), Eur. Heart J. 36 (41) (2015) 2793–2867, https://doi.org/10.1093/eurheartj/ehv316.
[9] E.M. Cronin, et al., 2019 HRS/EHRA/APHRS/LAHRS expert consensus statement on catheter ablation of ventricular arrhythmias, EP Europace 21 (8) (2019) 1143–1144, https://doi.org/10.1093/europace/euz132.
[10] S.R. Dukkipati, J.S. Koruth, S. Choudry, M.A. Miller, W. Whang, V.Y. Reddy, Catheter ablation of ventricular tachycardia in structural heart disease, J. Am. Coll. Cardiol. 70 (23) (2017) 2924–2941, https://doi.org/10.1016/j.jacc.2017.10.030.
[11] V. Y. Reddy, M. Taborsky, J. N. Ruskin, “Prophylactic Catheter Ablation for the Prevention of Defibrillator Therapy,” *n engl j med*, p. 9, 2007.
[12] Y. Komatsu, Substrate-based approach for ventricular tachycardia in structural heart disease: tips for mapping and ablation, J. Arrhythmia 30 (4) (2014) 272–282, https://doi.org/10.1016/j.joa.2014.04.014.
[13] R. Tung, Substrate mapping in ventricular arrhythmias, Cardiac Electrophysiol. Clin. 11 (4) (2019) 657–663, https://doi.org/10.1016/j.ccep.2019.08.009.
[14] P. Santangeli, F.E. Marchlinski, Substrate mapping for unstable ventricular tachycardia, Heart Rhythm 13 (2) (2016) 569–583, https://doi.org/10.1016/j.hrthm.2015.09.023.
[15] G.S. Guandalini, J.J. Liang, F.E. Marchlinski, Ventricular tachycardia ablation, JACC: Clin. Electrophysiol. 5 (12) (2019) 1363–1383, https://doi.org/10.1016/j.jacep.2019.09.015.
[16] D.F. Briceño, et al., Substrate ablation of ventricular tachycardia: late potentials, scar Dechanneling, local abnormal ventricular activities, core isolation, and homogenization, Cardiac Electrophysiol. Clin. 9 (1) (2017) 81–91, https://doi.org/10.1016/j.ccep.2016.10.014.
[17] C.A. Martin, et al., Use of novel electrogram ‘lumipoint’ algorithm to detect critical isthmus and abnormal potentials for ablation in ventricular tachycardia, JACC: Clin. Electrophysiol. 5 (4) (2019) 470–479, https://doi.org/10.1016/j.jacep.2019.01.016.
[18] V. Luther, et al., A prospective study of ripple mapping the post-infarct ventricular scar to guide substrate ablation for ventricular tachycardia, Circ: Arrhyth. Electrophysiol. 9 (6) (2016), https://doi.org/10.1161/CIRCEP.116.004072.
[19] H. Launer, T. Clark, T. Dewland, C.A. Henrikson, B. Nazer, An automated fractionation mapping algorithm for mapping of scar-based ventricular tachycardia, Pacing Clin. Electrophysiol. 42 (8) (2019) 1133–1140, https://doi.org/10.1111/pace.13758.
[20] M. Orini, et al., Evaluation of the reentry vulnerability index to predict ventricular tachycardia circuits using high-density contact mapping, Heart Rhythm 17 (4) (2020) 576–583, https://doi.org/10.1016/j.hrthm.2019.11.013.
[21] Z. Aziz, et al., Targeted Ablation of ventricular tachycardia guided by wavefront discontinuities during sinus rhythm: a new functional substrate mapping strategy, Circulation 140 (17) (2019) 1383–1397, https://doi.org/10.1161/CIRCULATIONAHA.119.042423.
[22] C.-Y. Lin, et al., Simultaneous amplitude frequency electrogram transformation (SAFE-T) mapping to identify ventricular tachycardia arrhythmogenic potentials in sinus rhythm, JACC: Clin. Electrophysiol. 2 (4) (2016) 459–470, https://doi.org/10.1016/j.jacep.2016.01.013.
[23] B. Campos, M.E. Jauregui, F.E. Marchlinski, S. Dixit, E.P. Gerstenfeld, Use of a novel fragmentation map to identify the substrate for ventricular tachycardia in postinfarction cardiomyopathy, Heart Rhythm 12 (1) (2015) 95–103, https://doi.org/10.1016/j.hrthm.2014.10.002.
[24] K. Kuroki, et al., New substrate-guided method of predicting slow conducting isthmuses of ventricular tachycardia: preliminary analysis to the combined use of voltage limit adjustment and fast-fourier transform analysis, Circ: Arrhyth. Electrophysiol. 11 (4) (2018), https://doi.org/10.1161/CIRCEP.117.005705.
[25] D. Gupta, M.H. Shariat, M. Baetz-Dougan, J. Hashemi, S. Akl, D. Redfearn, Novel automated paced fractionation detection algorithm for ablating ventricular tachycardia, JBiSE 09 (10) (2016) 488–500, https://doi.org/10.4236/jbise.2016.910044.
[26] S. Khan, et al., Artificial Intelligence and Machine Learning in Electrophysiology—a Short Review, Curr. Treat Options Cardio. Med. 25 (10) (2023) 443–460, https://doi.org/10.1007/s11936-023-01004-4.
[27] A.K. Feeny, et al., Artificial intelligence and machine learning in arrhythmias and cardiac electrophysiology, Circ: Arrhyth. Electrophysiol. 13 (8) (2020), https://doi.org/10.1161/CIRCEP.119.007952.
[28] V.D. Nagarajan, S.-L. Lee, J.-L. Robertus, C.A. Nienaber, N.A. Trayanova, S. Ernst, Artificial intelligence in the diagnosis and management of arrhythmias, Eur. Heart J. 42 (38) (2021) 3904–3916, https://doi.org/10.1093/eurheartj/ehab544.
[29] U.R. Acharya, et al., A deep convolutional neural network model to classify heartbeats, Comput. Biol. Med. 89 (2017) 389–396, https://doi.org/10.1016/j.compbiomed.2017.08.022.
[30] A. Picon, et al., Mixed convolutional and long short-term memory network for the detection of lethal ventricular arrhythmia, PLoS One 14 (5) (2019) e0216756.

[31] A.Y. Hannun, et al., Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network, Nat. Med. 25 (1) (2019) 65–69, https://doi.org/10.1038/s41591-018-0268-3.
[32] Z.I. Attia, et al., Screening for cardiac contractile dysfunction using an artificial intelligence–enabled electrocardiogram, Nat. Med. 25 (1) (2019) 70–74, https://doi.org/10.1038/s41591-018-0240-2.
[33] N. Gumpfer, D. Grün, J. Hannig, T. Keller, M. Guckert, Detecting myocardial scar using electrocardiogram data and deep neural networks, Biol. Chem. 402 (8) (2021) 911–923, https://doi.org/10.1515/hsz-2020-0169.
[34] M. I. Alhusseini *et al.*, "Machine learning to classify intracardiac electrical patterns during atrial fibrillation: machine learning of atrial fibrillation," *Circ: Arrhythmia and Electrophysiology*, vol. 13, no. 8, Aug. 2020, doi: 10.1161/CIRCEP.119.008160.
[35] A. Ghrissi, D. Almonfrey, F. Squara, J. Montagnat, V. Zarzoso, Identification of spatiotemporal dispersion electrograms in atrial fibrillation ablation using machine learning: a comparative study, Biomed. Signal Process. Control 72 (2022) 103269, https://doi.org/10.1016/j.bspc.2021.103269.
[36] J. Seitz *et al.*, "Artificial intelligence software standardizes electrogram-based ablation outcome for persistent atrial fibrillation," *Cardiovasc electrophysiol*, p. jce.15657, Sep. 2022, doi: 10.1111/jce.15657.
[37] X. Fagan, K. Ivanko, and N. Ivanushkina, "Detection of Ventricular Late Potentials in Electrocardiograms Using Machine Learning," in *Advances in Computer Science for Engineering and Education III*, Z. Hu, S. Petoukhov, I. Dychka, and M. He, Eds., Cham: Springer International Publishing, 2021, pp. 487–497.
[38] N. Pilia, et al., Non-invasive localization of the ventricular excitation origin without patient-specific geometries using deep learning, Artif. Intell. Med. 143 (2023) 102619, https://doi.org/10.1016/j.artmed.2023.102619.
[39] R.C. Lozoya, B. Berte, H. Cochet, P. Jais, N. Ayache, M. Sermesant, Model-based feature augmentation for cardiac ablation target learning from images, I.E.E.E. Trans. Biomed. Eng. 66 (1) (2019) 30–40, https://doi.org/10.1109/TBME.2018.2818300.
[40] M. Sermesant, H. Delingette, H. Cochet, P. Jaïs, N. Ayache, Applications of artificial intelligence in cardiovascular imaging, Nat. Rev. Cardiol. 18 (8) (2021) 600–609, https://doi.org/10.1038/s41569-021-00527-2.
[41] G. Baldazzi, M. Orrù, G. Viola, D. Pani, Computer-aided detection of arrhythmogenic sites in post-ischemic ventricular tachycardia, Sci. Rep. 13 (1) (Apr. 2023) 6906, https://doi.org/10.1038/s41598-023-33866-w.
[42] 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 (CinC), Singapore, Dec. 2019, pp. 1–4. doi: 10.22489/CinC.2019.343.
[43] G. Baldazzi, M. Orrù, M. Matraxia, G. Viola, D. Pani, "Supervised Classification of Ventricular Abnormal Potentials in Intracardiac Electrograms," presented at the 2020 Computing in Cardiology (CinC), Rimini, Italy, Dec. 2020, pp. 1–4. doi: 10.22489/CinC.2020.397.
[44] A. Pitzus *et al.*, "Exploring transfer learning for ventricular tachycardia electrophysiology studies," presented at the 2022 Computing in Cardiology (CinC), Tampere, Finland, Dec. 2022, pp. 1–4. doi: 10.22489/CinC.2022.382.
[45] I. Daubechies, J. Lu, H.-T. Wu, Synchrosqueezed wavelet transforms: an empirical mode decomposition-like tool, Appl. Comput. Harmon. Anal. 30 (2) (2011) 243–261, https://doi.org/10.1016/j.acha.2010.08.002.
[46] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: visual explanations from deep networks via gradient-based localization," in: *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice: IEEE, Oct. 2017, pp. 618–626. doi: 10.1109/ICCV.2017.74.
[47] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, Commun. ACM 60 (6) (2017) 84–90, https://doi.org/10.1145/3065386.
[48] C. Szegedy *et al.*, "Going Deeper with Convolutions," Sep. 16, 2014, *arXiv*: arXiv: 1409.4842. Accessed: May 31, 2022. [Online]. Available: http://arxiv.org/abs/1409.4842.
[49] K. Simonyan, A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv:1409.1556 [cs]*, Sep. 2014, Accessed: Jun. 23, 2019. [Online]. Available: http://arxiv.org/abs/1409.1556.
[50] G. Baldazzi, M. Orrù, G. Solinas, M. Matraxia, G. Viola, D. Pani, Spectral characterisation of ventricular intracardiac potentials in human post-ischaemic bipolar electrograms, Sci. Rep. 12 (1) (2022) 4782, https://doi.org/10.1038/s41598-022-08743-7.
[51] C. Torrence, G.P. Compo, A practical guide to wavelet analysis, Bull. Amer. Meteor. Soc. 79 (1) (1998) 61–78, https://doi.org/10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2.
[52] I. Daubechies, S. Maes, "A nonlinear squeezing of the continuous wavelet transform based on auditory nerve models," in *Wavelets in Medicine and Biology*, 1st edition., New York: CRC Press, 1996, pp. 527–546. [Online]. Available: https://doi.org/10.1201/9780203734032.
[53] J. Deng, W. Dong, R. Socher, L.-J. Li, Kai Li, Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," in: *2009 IEEE Conference on Computer Vision and Pattern Recognition*, Miami, FL: IEEE, Jun. 2009, pp. 248–255. doi: 10.1109/CVPR.2009.5206848.
[54] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," presented at the Proceedings of the thirteenth international conference on artificial intelligence and statistics, 2010, pp. 249–256.
[55] H.-T. Wu, Current state of nonlinear-type time–frequency analysis and applications to high-frequency biomedical signals, Curr. Opin. Syst. Biol. 23 (2020) 8–21, https://doi.org/10.1016/j.coisb.2020.07.013.
[56] S.J. Pan, Q. Yang, A survey on transfer learning, IEEE Trans. Knowl. Data Eng. 22 (10) (2010) 1345–1359, https://doi.org/10.1109/TKDE.2009.191.
[57] J. Silberbauer *et al.*, "Noninducibility and late potential abolition: a novel combined prognostic procedural end point for catheter ablation of postinfarction ventricular tachycardia," *Circ: Arrhythmia and Electrophysiology*, vol. 7, no. 3, pp. 424–435, Jun. 2014, doi: 10.1161/CIRCEP.113.001239.