Electromechanical Wave Imaging with machine learning for automated isochrone generation

Lea Melki, Melina Tourni and Elisa E. Konofagou, *Member, IEEE*

***Abstract*— Standard Electromechanical Wave Imaging isochrone generation relies on manual selection of zero-crossing (ZC) locations on incremental strain curves for a number of pixels in the segmented myocardium for each echocardiographic view and patient. When considering large populations, this becomes a time-consuming process, that can be limited by inter-observer variability and operator bias. In this study, we developed and optimized an automated ZC selection algorithm, towards a faster more robust isochrone generation approach. The algorithm either relies on heuristic-based baselines or machine learning classifiers. Manually generated isochrones, previously validated against 3D intracardiac mapping, were considered as ground truth during training and performance evaluation steps. The machine learning models applied herein for the first time were: i) logistic regression; ii) support vector machine (SVM); and iii) Random Forest. The SVM and Random Forest classifiers successfully identified accessory pathways in Wolff-Parkinson-White patients, characterized sinus rhythm in humans, and localized the pacing electrode location in left ventricular paced canines on the resulting isochrones. Nevertheless, the best performing classifier was proven to be Random Forest with a precision rising from 89.5% to 97%, obtained with the voting approach that sets a probability threshold upon ZC candidate selection. Furthermore, the predictivity was not dependent on the type of testing dataset it was applied to, contrary to SVM that exhibited a 5% drop in precision on the canine testing dataset. Finally, these findings indicate that a machine learning approach can reduce user variability and considerably decrease the durations required for isochrone generation, while preserving accurate activation patterns.**

***Index Terms*— Automated isochrone generation, Electromechanical Wave Imaging, machine learning, Random Forest classifier, zero-crossing selection.**

I. Introduction

HEART disease stands as the leading cause of mortality and morbidity worldwide, accounting for 30% of all deaths [1], [2]. By 2030, the number of fatalities are expected to increase to over 24 million a year globally [3], [4]. In addition, cardiovascular disease poses a severe economic burden and is projected to cost the U.S. healthcare system more than $800 million yearly by 2030 [5], [6]. Early detection of cardiac conduction malfunctions such as arrhythmias would definitely help reduce the burden cardiovascular diseases have on public health and overcome the current clinical challenges [7], [8], [9]. The imaging techniques currently available to clinicians for cardiac activation sequence mapping are invasive, ionizing, time-consuming and costly [10], [11], [12]. Thus, there is an undeniable, urgent need for a non-invasive and reliable imaging tool, which could play a crucial role in the early diagnosis of conduction diseases and allow physicians to choose the best course of action [13], [14].

Echocardiography-based imaging techniques are low-cost, non-invasive, and do not require the use of potentially harmful ionizing radiation [15]. Ultrasound is therefore a highly desirable imaging modality from a patient perspective. It enables visualizing the mechanics and measuring the dynamics of the heart through the cardiac cycle, yielding diagnostic and prognostic information [16]. Furthermore, the portability and ease-of-use of ultrasound systems have led to widespread adoption among cardiologists, establishing ultrasound as the most frequently used modality in clinical cardiology [17].

Electromechanical Wave Imaging (EWI) is a high frame-rate ultrasound-based functional modality that can non-invasively map the electromechanical activation of the heart, i.e., the transient deformations immediately following the electrical activation, estimated in the time-domain on the radiofrequency signals in 1D [18], [19], [20]. The type of infinitesimal strain measured, whether longitudinal or radial, does not hinder the activation map generation process. In fact, since EWI is an angle independent technique, the only key factor is the location of the sign change of this strain, surrogate for the electrical activation timing [30]. Nonetheless, there is room for potential improvements in the activation map generation process at different steps of the EWI processing pipeline: from the high frame-rate ultrasound acquisition sequence itself [15], [21], to the beamforming algorithm [22]- [25], the type of displacement and strain estimators [26], [27], or even at the final processing step of zero-crossing (ZC) selection on the incremental strain curves [28], [29]. In this paper, we will focus exclusively on enhancing and accelerating that last ZC selection task.

In previous clinical studies from our group [31], [32], we demonstrated that in conjunction with catheter-based electroanatomical mapping and 12-lead electrocardiogram (ECG), EWI could be a viable assisting tool for non-invasive diagnosis, treatment-planning and monitoring of ventricular and atrial arrhythmias in the clinic. Nevertheless, as explained in these papers, standard isochrone generation in the past relied

This work was supported in part by the National Institutes of Health under grant NIH R01 HL-140646-01.

Lea Melki and Melina Tourni are with the Department of Biomedical Engineering, Columbia University, New York, NY 10032 USA (e-mails: lm3088@columbia.edu and mt3393@columbia.edu).

Elisa E. Konofagou is with the Biomedical Engineering and Radiology Departments, Columbia University, New York, NY 10032 USA (e-mail: ek2191@columbia.edu).

The final version of record is available at http://dx.doi.org/10.1109/TMI.2021.3074808

on manual selection of the ZC locations on the incremental axial strain curves for about 150 pixels in the segmented myocardium mask for each echocardiographic view and each patient. When considering large patient populations, this can become a time-consuming process. Our group had introduced a semi-automated algorithm, used previously in the study by Bunting et al. [33], to automatically identify and select the ZCs corresponding to the electromechanical activation for each point in the tissue. However, this heuristic algorithm relied on hard thresholding and empirical conditions that were not further optimized nor validated against manually generated isochrones in other EWI studies. Improving and accelerating the isochrone generation process would bring EWI one step closer to integration in the clinic.

Due to the plethora of images typically acquired from patients in echocardiography (thousands per second), machine learning (ML) and deep learning have gained a lot of interest in the ultrasound field over the past few years [34], not only for i) image reconstruction and beamforming [35], [36] with the ADMIRE model among others [37], [38]; but also for ii) tumor detection [39]; and iii) segmentation of echocardiographic views with neural networks [40], [41]. ML has also been applied to myocardial strain estimates derived from strain rate imaging by Tissue Doppler or Speckle Tracking Echocardiography [42]- [44] for various applications: quality assurance for reliability assessment [45], [46], strain curve classification into physiological or artifactual categories [47], as well as extracting spatiotemporal characteristics of echocardiographic deformation curves for infarct classification purposes [48].

In this study, we sought to improve some aspects of the standard EWI ZC selection processing pipeline: reducing inter-observer variability and possible bias of the manually generated isochrones, while also decreasing the time needed to generate the activation maps, towards real-time implementation. To achieve this goal, we developed an automated zero-crossing detection algorithm relying either on heuristic-based baselines or machine learning classifiers. We then investigated how training machine learning models to automatically generate isochrones with manually selected zero-crossing features from past validated cases could lead to more robust activation maps. Finally, after optimization of the classifier, we determined what was the most suitable supervised machine learning algorithm resulting in a faster overall isochrone generation process, while also preserving the accurate electromechanical activation pattern from manual generation.

II. Materials and Methods

A. Standard Electromechanical Wave Imaging

Electromechanical Wave Imaging relies on a 2-second high frame-rate ultrasound sequence composed of a single diverging wave at 2000 Hz, acquired in four standard echocardiographic apical views on a Vantage 256 research scanner (Verasonics Inc., Kirkland, WA, USA) with a 2.5 MHz phased-array transducer (ATL P4-2, Philips, Andover, Massachusetts). The displacement estimation is performed on the radiofrequency signals from each element of the ultrasound probe with 1D axial cross-correlation (10 wavelength window, 90% overlap) [49], followed by a least-squares estimator to derive the axial strains (5-mm kernel) [50].

The wavefront of the electromechanical activation is defined as the time point at which the incremental strain value changes from relaxation to contraction. In the apical views, since the ultrasound beam is aligned with the ventricular myocardial wall that is shortening during systole, this corresponds to a positive-to-negative flip or downward ZC of the incremental axial strain. For most strain curves, a single negative peak exists during systole and in these cases, selection of the ZC location is straightforward. However, incremental strain curves sometimes exhibit more than one negative peak after the onset of the QRS complex or p-wave, for the ventricular or atrial isochrones respectively. In these instances, instead of one clear ZC, multiple ZC candidates are identified.

Usually, in the standard EWI isochrone generation process, the myocardium of interest is first segmented on the B-mode image. For about 150 randomly chosen points in the segmented mask, the operator manually selects on the incremental axial strain curves the timing of first sign change after the QRS onset for the ventricles, respectively p-wave onset for the atria. A Delaunay triangulation–based cubic interpolation is then applied to the 2D scattered activation time ZC values in order to achieve a continuous isochrone pattern throughout the entire myocardium mask grid. The activation timings in milliseconds are finally color-coded on a 2D map or isochrone, with red being early activation and blue late. Once the maps have been generated in the four views, a 3D-rendering algorithm automatically detects the longitudinal left ventricular (LV) median axis on the four apical isochrones [51]. The multi-2D views are then co-registered around that axis and linear interpolation of the activation times is performed around the circumference. This leads to 3D-rendered activation maps. The full standard manual EWI isochrone generation processing pipeline can take up to 90 minutes and runs in MATLAB.

Further details on the imaging sequence, beamforming, standard EWI processing steps and associated parameters can be found in previous publications from our group [30]-[33]. Furthermore, to address the cases with multiple ZCs and choose the best candidate consistently without observer bias from the standard manual approach, we developed the automated isochrone generation algorithm described below.

B. Developing the automated algorithm

Prior to feature detection and ZC selection on the incremental axial strain curves, a search window was set to 200 ms after the QRS origin picked on our single lead ECG, outside of which we do not look for ZC candidates (vertical black dotted line on Fig.1). This assumption is based on the fact that we expect the ventricles to have fully activated within that 200 ms range, because in this study we included healthy subjects with no history of ischemia, infarct or other co-morbidities, as explained in the ensuing ground truth datasets description subsection. In the case of pathological cases, the search window would need to be extended to a longer time interval to make sure we aren't missing any ZC candidate in the abnormal myocardial tissue. Zero-crossing detection and

feature extraction from strain curves have been used by other groups for various ultrasound applications [52], [53].

In this study, we are interested in automatically collecting the following features, as illustrated on Figure 1:

1) $(\mathbf{X_i}, \mathbf{Y_i})$: spatial coordinates of each point within the segmented myocardial mask, with $i \in [1 : k]$; e.g. a limit of 150 pixels $(k = 150)$ can be applied when comparing directly the automated results to manually generated isochrones.
2) $\mathbf{t_{ZC_{i,n}}}$: all positive to negative ZC time-points (activation timings in ms), with n index of the ZC (e.g., $n = 1$ if only single ZC possible or $n = 2$ on Fig.1).
3) $\mathbf{slope_{ZC}}$: slope of the strain curve at each ZC location (red dotted tangent line on Fig.1).
4) $\epsilon_{\mathbf{max}}$ and $\epsilon_{\mathbf{min}}$: maximum positive and minimum negative strain values surrounding each ZC.
5) $\mathbf{t}_{\epsilon_{\mathbf{max}}}$ and $\mathbf{t}_{\epsilon_{\mathbf{min}}}$: time-point locations of the maximum and minimum strain values (in ms).

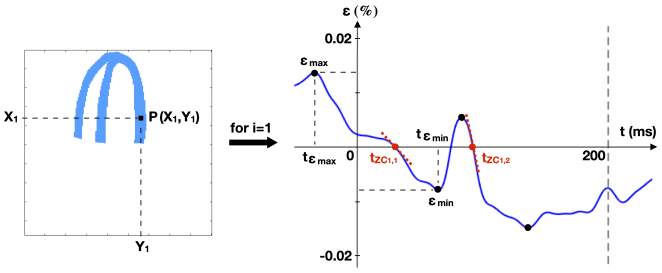


Fig. 1. Incremental strain curve and feature collection for a given pixel $\boldsymbol{P(X_1, Y_1)}$ in the apical four-chamber view ventricular mask.

From here onward, we will refer to the points with single ZC occurrences (n = 1) as "stable ZCs", compared to "ZC candidates" when multiple options are available (n > 1). In the 35 ground truth datasets included in this study (section II.D.), stable ZCs represented $\approx 40\%$ of the cases, while ZC candidates were present $\approx 60\%$ of the time.

C. Defining heuristic-based automated approaches as a reference baseline

Once the incremental axial strain curve features were collected, we considered two heuristic-based automated approaches and compare the resulting isochrones for each one of them to the manually generated ground truth maps. The approaches were implemented as follows in MATLAB:

1) Naive approach: always select the first positive to negative ZC immediately following the QRS complex onset and occurring within the 200 ms search window.
2) Rule-based approach: select the first positive to negative ZC that satisfies the following empirical conditions:

$$\epsilon_{min} \geq 20\%\ min(\epsilon_{min}\ \forall i) \quad (1)$$

$$\mid slope_{ZC} \mid\ \geq 20\%\ max(\mid slope_{strain\ curve} \mid) \quad (2)$$

These ZCs, defined based on conditions on the negative ZC peak amplitude and the steepness of the strain curve slope at the crossing, were previously introduced in [33] and referred to as "clear" ZCs.

The second approach tries to mimic the observer's reasoning when there is more than one ZC candidate to choose from. In fact, these two conditions (1) and (2) on the strain curve's slope at the location of the ZC and on the amplitude of the peak negative strain are common biases experienced by operators upon their manual ZC selection process.

D. Ground truth datasets

To be able to compare the performance of the different automated approaches, we need to establish and confirm which method yields the most accurate physiological ZCs, and if the resulting isochrones preserve the accurate manually generated activation pattern or maybe even outperform it. Towards that goal, we have to test the algorithm on some well characterized datasets. Manually generated isochrones can only be considered ground truth data if they have previously been validated by electroanatomical mapping (EAM) [54].

First, we considered twenty-four (N=24) Wolff-Parkinson-White (WPW) patients taken from our group's pediatric (14/24) and adult (10/24) clinical studies, [31] and [32], respectively, all previously validated against intracardiac mapping prior to their ablation procedures. The patients had ventricular pre-excitation, but were known to be otherwise healthy with normal cardiac anatomy and function (e.g., no ischemia as it could have an effect on the strain curves due to tethering for instance or reduced strain amplitudes). We also included five (N=5) human cases in sinus rhythm, from patients post successful WPW ablation.

In addition, we included data from two prior open-chest canine experiments [32] with EAM validation on the epicardial surface of the LV using the clinical EnSite mapping system (Abbott Medical, St. Paul, MN, USA). Contrary to patient data, this allowed us to achieve exact LV wall co-registration between the surface mapped and the wall portions being imaged in the apical views. This was easily accomplished after the thoracotomy by probing the epicardium with the mapping catheter, ensuring visualization on the corresponding B-mode before acquiring the views, and placing the associated labels on the electroanatomic maps. One canine was paced at a single antero-lateral LV location, while the other animal had 5 pacing electrodes positioned across its LV anterior surface; leading to a total of six (N=6) different LV-paced canine datasets.

TABLE I
PATIENT DEMOGRAPHICS IN GROUND TRUTH DATASETS

Wolff-Parkinson-White patients (N=24)		
Gender	Male	n=13 (54%)
	Female	n=11 (46%)
Age (yrs)	Range	7-67
	Median	17
Patients in sinus rhythm (N=5)		
Gender	Male	n=2 (40%)
	Female	n=3 (60%)
Age (yrs)	Range	7-17
	Median	12

Therefore, we ended up with a total of 35 datasets with zero-crossings validated thanks to the manual selection ground

truths: 35 cases x 4 apical views x 150 spatial points = 21 000 incremental strain curves. These 21 000 strain curves were all annotated by a single expert.

All datasets were previously acquired with a Vantage 256 research scanner. The human subject study was conducted under a protocol approved by the institutional review board (IRB) of Columbia University and was carried out in accordance with IRB guidelines and regulations. Informed consent was obtained from all participants. In parallel, the animal study protocol was approved by the Institutional Animal Care and Use Committee (IACUC) at Columbia University, New York, NY, USA, and was compliant with the Public Health Service Policy on Humane Care and Use of Laboratory Animals.

E. Machine learning classification models

A supervised learning model was subsequently built to classify the ZC candidates. Supervised learning consists of building a model that approximates a function trained with ground truth samples (i.e., labeled data). For each sample, this learnt function takes features as an input, and returns a value. In our case, we need a binary classifier that determines whether a ZC candidate is the correct one. In order to do so, we started by labelling the data: for a given pixel (X_i, Y_i) in the segmented myocardium mask, the ZC candidate with the activation time $t_{ZC_{i,n}}$ closest to the manually selected ZC location was assigned the label = 1 (i.e., true location), while the others $t_{ZC_{i,j}}$ with $j \neq n$ were assigned the label = 0. The feature engineering step was performed on the above-mentioned ground truth datasets.

Pre-processing of the dataset is required prior to training the classification model. On top of the automatically collected numerical strain curve features listed above, we also included the following categorical features: the apical view type the strain curves were retrieved from (4-chamber, 3.5-chamber, 2-chamber or 3-chamber view), as well as the type of acquisition used (i.e., field of view depth and pixel resolution). The list of all features is provided in Table II with a total of: 16 features = 8 categorical features + 8 numerical features.

TABLE II
FEATURES USED FOR THE MACHINE LEARNING ALGORITHM

	Categorical features	Numerical features
1	Apical 4-chamber	X_i coordinate
2	Apical 3.5-chamber	Y_i coordinate
3	Apical 2-chamber	$t_{ZC_{i,n}}$ ZC location
4	Apical 3-chamber	$slope_{ZC}$ slope at ZC location
5	14-cm deep field of view	ϵ_{max} peak positive strain before ZC
6	20-cm deep field of view	ϵ_{min} peak negative strain after ZC
7	$\lambda/16$ pixel axial resolution	$t_{\epsilon_{max}}$ peak location before ZC
8	$\lambda/8$ pixel axial resolution	$t_{\epsilon_{min}}$ peak location after ZC

We then performed "one-hot encoding" on the categorical features, which consists of converting categorical variables into a binary form that can be provided as a feature to train the ML algorithm. Finally, the last pre-processing step was to standardize the numerical features in order to convert them to a common scale with an average of zero and a standard deviation of one. This was achieved with z-normalization: (X_i, Y_i) standardization was performed by patient and by view for consistent distance measurement purposes across cases, while the other numerical features were standardized across patients prior to training.

The developed machine learning algorithm is a binary classifier. For each $P(X_i, Y_i)$ with no stable ZC, the algorithm has to vote for the best ZC candidate. Therefore, we want to learn a function f that maps the features described above to a probability $p \in [0 : 1]$ of being a true ZC, and predict that the candidate with the highest probability is indeed the ZC, while the others are not. To find the most appropriate algorithm, we tried three classic machine learning approaches: Logistic Regression, Support Vector Machine (SVM) and Random Forest. Logistic regression is a widely used statistical generalized linear model, often chosen as a reference baseline in ML, as it is simple and straightforward to implement [55]. The logistic regression classifier used herein relied on an elastic net regularization, linearly combining L1 and L2 regularization techniques on the regression coefficients included in the penalty term of the loss function [56]. The corresponding hyperparameter in this case is called the L1-ratio; being equal to 1 if the penalty term is only an L1 regularization, while equal to 0 if it's only an L2 regularization (Table III). Besides, SVM, also known as support vector networks, is a supervised learning model used for classification, which defines a hyperplane to divide the two categories [57]. The SVM kernel defines what type of separation will be applied between the classes and allows to find the optimal hyperplane in a different dimensional space. Finally, Random Forest is an ensemble learning method for classification that operates by constructing a multitude of decision trees [58]. Random Forest is known to be good for dealing with heterogeneous features [59]. Each decision tree gives a predicted label to the samples, and the Random Forest ends up assigning the label that has the most votes out of all the decision trees. All models were implemented in Python with the Scikit-learn library to define the three classifiers and their associated hyperparameters (Table III).

Furthermore, how did these models vote for the correct ZCs for each $P(X_i, Y_i)$? Two voting options were possible in order to convert the probability values of the ZC candidates to binary labels and determine the predicted labels for each candidate:

1) Always select the highest probability ZC candidate for each pixel and assign it the label = 1:

$$\arg\max_n \left(p(t_{ZC_{i,n}})\right) \;\; \forall (X_i, Y_i) \tag{3}$$

2) Swipe through a set a probability thresholds to find the best cutoff value: highest precision given a condition set on the recall (also known as sensitivity or true positive rate). Then, pick the best ZC candidate only if its probability is higher than the cutoff threshold, and assign it the label = 1:

$$\arg\max_n \left(p(t_{ZC_{i,n}})\right) \;\; \textit{only if} \;\; p(t_{ZC_{i,n}}) > p_{thres} \tag{4}$$

These approaches were evaluated on the previously described and validated ground truth cases. Towards that goal, we first split the 24 WPW datasets into a training set, and a

TABLE III
MACHINE LEARNING MODELS, ASSOCIATED HYPERPARAMETERS AND LIST OF CORRESPONDING VALUES USED UPON ITERATIVE TUNING

ML model	Hyperparameters	Values used for iterative tuning
Logistic regression	elastic net, L1-ratio	[0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1]
SVM	kernel type	'linear', 'poly', 'rbf', 'sigmoid'
Random Forest	n estimators	[20, 50, 100, 200, 300, 400, 500]

Poly = polynomial, rbf = radial basis function, n estimators = number of estimators or trees in the forest.

validation set with a 70/30 ratio: 16 patients for training (9614 labelled strain curves) vs 8 for evaluating the hyperparameters (4453 strain curves). The hyperparameter tuning process was iterative: we fitted a model on the training set with a set of given hyperparameters, and evaluated its performance on the validation set. The approach taken was that of a grid search, where we iteratively swept through all possible combinations of hyperparameters for a given model within ranges we had set (Table III). We compared the performance of the different models based on evaluation metrics described in more details in the next paragraph, computed on the validation set. Once we found a machine learning algorithm and a set of hyperparameters that yielded good results on the validation set, we refitted the model with these settings on the whole dataset constituted of both the training and validation samples. This step allowed for a larger sample size in order to fit, hopefully, a better model. Finally, we tested the subsequent model on a held-out set called the test set. The latter included the remaining 11 ground truth datasets not used so far: 5 human sinus rhythm subjects and 6 LV paced canines. We used this test set only once to avoid leaking information to our hyperparameter tuning process and damaging the generalizability of our model by overfitting it to the test set. The resulting testing metrics are the true performance evaluation of our model, since they were obtained on a dataset that was neither used to fit the model, nor used to select hyperparameters (Table IV, b-c). Although generalizability is a central problem in machine learning, we can build some confidence regarding future performance of our predictive models by using different test datasets that illustrate various realistic tasks, such as a different set of patients, or transferability of our model's predictivity to canine subjects.

F. Evaluation metrics

In order to evaluate the performance of the automated machine learning algorithms, the metrics used were the following:

$$precision = \frac{TP}{TP + FP} \qquad and \qquad recall = \frac{TP}{TP + FN}$$

with TP = true positive, FP = false positive and FN = false negative. In our scenario, the precision represents what portion of our prediction is true, while the recall corresponds to how many true ZCs did we find overall.

III. RESULTS

Multi-2D isochrones on an illustrative example corresponding to the canine testing dataset are shown in Figure 2 for the ground truth manual zero-crossings, as well as for the two heuristic-based automated approaches. The anterior LV paced canine case satisfyingly exhibited a single early activated region on the ground truth manual isochrones, displayed in red on the LV anterior wall of the apical 2-chamber view (Fig.2-a). While the first occurring ZCs approach did catch the pacing electrode's (earliest activated) region, it also displayed several other early activated areas, even on the right ventricular (RV) free wall very distal from the location of the pacing electrode (Fig.2-b). Furthermore, the isochrones generated with the negative peak amplitude and slope conditions (1) and (2) were less noisy and more similar to the manual ground truth in terms of predominantly late activated blue color and overall activation pattern, despite missing the pacing electrode's location on the anterior wall of the apical two-chamber view (Fig.2-c). None of the two heuristic-based automated approaches were very convincing.

Regarding the machine learning classification models, the set of tuned hyperparameters that was found to yield the best results (highest recall or sensitivity) on the validation set was the following: logistic regression, elastic net L1-ratio = 0.5; SVM, kernel = radial basis function; and Random Forest, n estimators = 200. The precision-recall curves for the three models (logistic regression in red, SVM in blue and Random Forest in green) were evaluated on the validation, human test and canine test datasets, and are shown in Figure 3. SVM and Random Forest are shown to significantly outperform logistic regression on all three datasets, while the two former performing similarly and not falling below 88% precision even in the worst case. Additionally, the stars and values overlaid onto the curves represent the performances for the different voting approaches and threshold scenarios used to generate the EWI isochrones throughout this paper, as explained in details in Table IV. More specifically, we decided to take a closer look at two particular recall values: 70% and 40%. These recall percentages were selected based upon our assessment of the precision-recall curves for each machine learning model and each dataset. The two values were subsequently set as accepted thresholds that balanced between precision and recall rates, and generated isochrones with the best quality tradeoff. Given that we initially started with about 150 ground truth pixels $P(X_i, Y_i)$ for each apical view, a 70% recall would mean having at least 100 points with manually selected ZCs to choose from when generating our automated isochrones. However, setting a 40% recall (e.g., 60 pixels) would lead to spatial under sampling within the segmented myocardium mask and thus, drastically impact the isochrone activation pattern, which would then mostly result from interpolation and not actual selected activation times. Nevertheless, if we collect the features on about 2500 pixels per view, we would

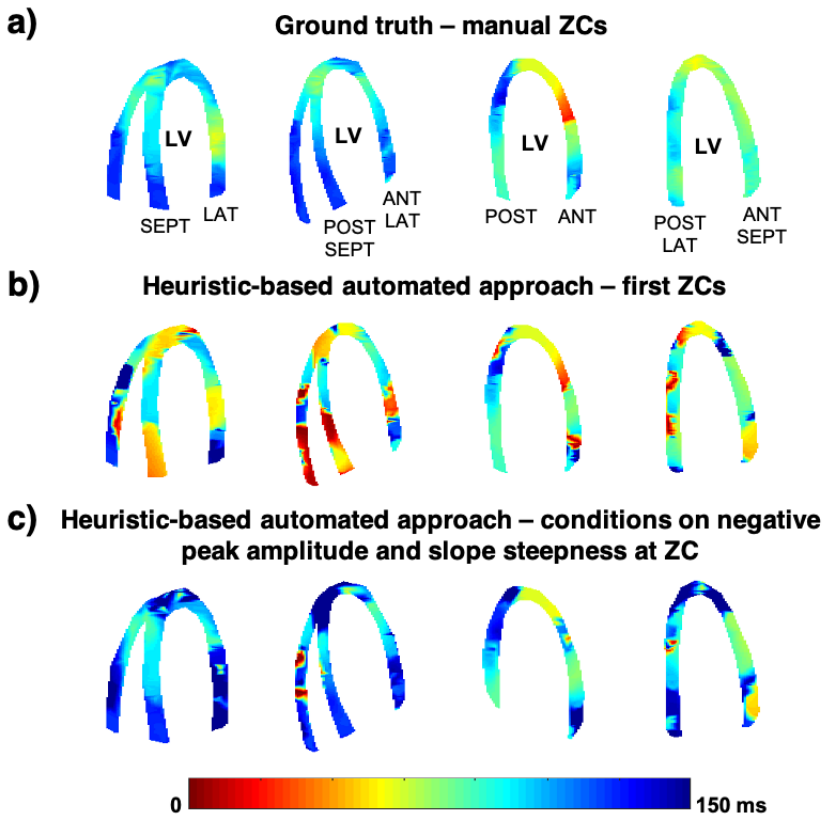


Fig. 2. Multi-2D isochrones generated with different heuristic-based automated approaches on the example of an anterior LV paced canine taken from the testing dataset and compared to the manually generated ground truth isochrones. The multi-2D isochrones are shown in the four apical views (from left to right: 4, 3.5, 2 and 3-chamber view) for: a) Manual ground truth; b) Naive heuristic-based automated approach selecting always the first occurring ZCs; c) Rule-based automated approach selecting the ZCs satisfying conditions (1) and (2). ANT=anterior; LAT=lateral, POST=posterior; SEPT=septum.

start with a larger pool of axial incremental strain curves. In fact, once the models have been trained, the ground truth and labelled ZCs are not required on all the points to perform the testing on new data. Therefore, even with a low recall of 40%, we would still end up with 1000 pixels, which corresponds to an order of magnitude higher than our current standard manual EWI isochrone generation process.

Quantitatively, Table IV details the ML models' performance evaluated on the entire validation dataset (Table IV-a), as well as on both test datasets: human (Table IV-b) and canine (Table IV-c). The two different voting options for the multiple ZC candidates are shown: the highest probability approach always voting for a ZC candidate at all times (3), and the probability threshold option only voting for the best ZC candidate that satisfies the probability threshold condition (4). It is worth noting that the precision and recall values for the first voting approach are always equal to one another. This is due to the fact that the highest probability models always make a ZC candidate selection for each pixel. In that case, because of that specific voting process FN=FP and thus, the evaluation metrics formulas written above end up being the same. On the other hand, for the second ZC voting approach (4), the table displays the best precision performance for two scenarios: i) when satisfying a recall $> 70\%$, and ii) when satisfying a recall $> 40\%$ respectively. Lastly, the corresponding probability thresholds p_{thres} are explicitly listed in the last column for each case. As expected with the precision-recall tradeoff, when the recall decreases from 70% to 40%, the precision improves by at least 2% and up to 5% for both SVM and Random Forest models depending on the dataset of interest. In addition, even though SVM initially performed very similarly to Random Forest in terms of precision, on the validation dataset as well as on the human testing dataset, SVM's precision exhibited a 5% drop on the canine testing set (Table IV-c). Meanwhile, Random Forest remained stable and barely witnessed any change in precision when applying the model to canine data.

Figure 4 illustrates the 3D-rendered isochrone results for the heuristic-based and ML-based automated approaches on the example of an antero-lateral LV paced canine, retrieved from the test dataset. The isochrone patterns of the heuristic naive baseline selecting the first ZCs and logistic regression automated approaches look very similar to one another (Fig.4-b, d), but are far from comparable to the manual ground truth result (Fig.4-a). Besides, the isochrone resulting from the rule-based heuristic approach does have a later blue activation pattern overall closer to the ground truth, despite missing the early activated region of interest from the LV pacing electrode in the antero-lateral region (Fig.4-c). Meanwhile, SVM and Random Forest (Fig.4-e, f, top row) clearly outperformed the three other approaches and had promising qualitative performance, despite

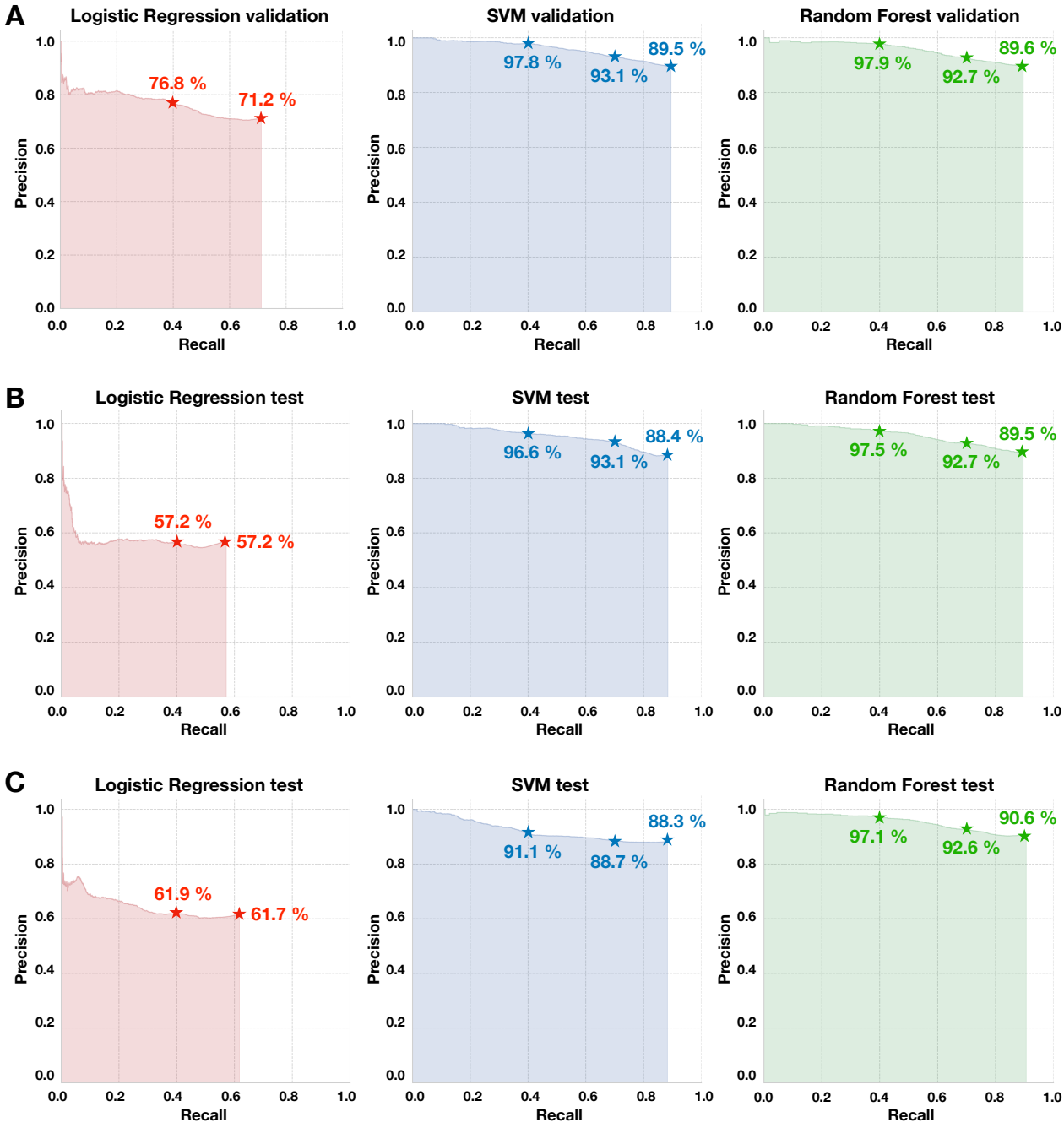


Fig. 3. Precision-recall curves for the three ML models (Left column: logistic regression in red, Center: SVM in blue, and Right column: Random Forest in green). The curves are displayed for performances evaluated on: a) the human validation dataset; b) the human test dataset; and c) the canine test dataset. The ML models performance are shown for their tuned hyperparameters: elastic net L1-ratio = 0.5; kernel = radial basis function and n estimators = 200.

revealing discontinuities in the activation pattern anteriorly compared to the manually generated isochrone. The recall here was maximal, since the models did not discard any pixel upon voting for the ZC candidates (3). In order to further increase the precision, we used the second voting option applying a probability threshold to the ZC candidates (4) and making sure the recall remained $> 70\%$ (Fig.4, middle row). In that case, the best performing algorithm was the Random Forest model. The latter is emphasized here by the blue dashed box. This was indeed previously quantified and confirmed by the higher performance evaluation metrics (Table IV). Furthermore, when setting the voting condition such that the recall was higher than 40%, while simultaneously considering a larger amount of pixels $(P(X_i, Y_i)$ with $i \in [1 : 2500])$ not previously used in the initial ground truth manual ZC selection (Fig.4-e, f, bottom row), the Random Forest precision increased even more. In fact, the antero-lateral LV pacing spot displayed in red became more localized in the longitudinal direction and did not spread towards the base anymore. Thus, the overall optimal automated algorithm that allowed the most successful

TABLE IV
MACHINE LEARNING MODELS PERFORMANCE EVALUATION

a)	Highest probability ZC		Probability threshold on ZCs		
	Precision	**Recall**	**Precision**	**Recall**	p_{thres}
Logistic regression	71.19 %	71.19 %	71.20 %	71.17 %	*0.254*
			76.81 %	40.02 %	*0.643*
SVM	89.51 %	89.51 %	93.14 %	70.18 %	*0.767*
			97.80 %	42.02 %	*0.938*
Random Forest	89.58 %	89.58 %	92.69 %	70.31 %	*0.720*
			97.86 %	41.05 %	*0.915*
b)	**Highest probability ZC**		**Probability threshold on ZCs**		
	Precision	**Recall**	**Precision**	**Recall**	p_{thres}
Logistic regression	57.16 %	57.16 %	57.16 %	57.16 %	*0.200*
			57.16 %	57.16 %	*0.200*
SVM	88.41 %	88.41 %	93.07 %	70.12 %	*0.775*
			96.61 %	40.18 %	*0.937*
Random Forest	89.53 %	89.53 %	92.74 %	70.06 %	*0.721*
			97.52 %	40.09 %	*0.903*
c)	**Highest probability ZC**		**Probability threshold on ZCs**		
	Precision	**Recall**	**Precision**	**Recall**	p_{thres}
Logistic regression	61.67 %	61.67 %	61.67 %	61.67 %	*0.221*
			61.91 %	40.43 %	*0.576*
SVM	88.26 %	88.26 %	88.68 %	70.03 %	*0.698*
			91.12 %	40.06 %	*0.919*
Random Forest	90.59 %	90.59 %	92.61 %	70.05%	*0.704*
			97.13 %	40.01 %	*0.912*

Performance evaluated on: a) the validation dataset; b) the human test dataset; and c) the canine test dataset.
The three models display the performance for their tuned hyperparameters: L1-ratio = 0.5; kernel = radial basis function and n estimators = 200.

localization of the antero-lateral LV pacing spot, as validated by the manual isochrones, is outlined by the green dashed box (Random Forest, recall $>$ 40% and 2500 pixels per view). Nevertheless, minor differences in the activation pattern remain visible at the basal level compared to the manual ground truth isochrone. These discrepancies are caused by the ZC selection process on the apical 3.5-chamber view for the Random Forest classifier when a probability threshold is applied to the ZC candidates upon voting, and further emphasized by the subsequent interpolation of the activation times in 2D. Representative strain curves are displayed on Figure 5 for four points located in the antero-lateral LV wall of the 3.5-chamber isochrone mask. The Random Forest voting approach with the recall conditions (both $>$ 70% and $>$ 40%) discarded some ZC points of lower confidence. This lead to a reduced spatial sampling in that particular wall region, which resulted here namely in P_2 falling outside of the interpolated isochrone mask at the base and therefore, explains the minor divergence from the manual isochrone.

Moreover, the comprehensive overview of the corresponding multi-2D EWI isochrone slices in all four apical views prior to 3D-rendering for the SVM and Random Forest models are included in Supplemental Figure 1, as well as the original ground truth manual 2D isochrones for that same LV paced canine. The 2D isochrones clearly demonstrate the improvements resulting from the second voting approach with the probability threshold, as well as the progress made when the model was applied to a larger number of pixels. The dashed circles outline where the precision was recovered when the recall decreased. Figure 6 and Supplemental Figure 2 further clarify the precision-recall tradeoff on this particular canine example. The pixel density is mapped spatially in each of the four apical views for all ZC voting scenarios (Supplemental Figure 2). This illustrates how the Random Forest classifier has a lower confidence in particular myocardial wall regions, but doesn't risk voting for a ZC candidate in these regions and would rather keep the precision metric higher than SVM by discarding more lower probability points, as previously shown and explained with P_2 in Figure 5. More specifically, Figure 6 displays representative examples of incremental strain curves for 4 points in the segmented myocardium mask within the RV free wall of the 3.5-chamber view this time. The presence of a second earlier ZC candidate on three out of the four curves, different from the manually selected ground truth ZC, informs the change in the Random Forest model's behavior between the voting approaches in the recovered precision region with the recall conditions (green dashed circle).

Finally, other examples of automated isochrones compared to manual ground truth are given in Fig.7 and Supplemental Figure 3, this time on a patient in sinus rhythm taken from the human testing dataset and a left lateral WPW case taken from the validation dataset, respectively. Once more, the Random Forest algorithm performed the best, clearly depicting the normal early activation of the atrioventricular node at the basal septum and the location of the accessory pathway on the left lateral wall of the LV.

IV. DISCUSSION

In this optimization study, we developed an automated zero-crossing detection algorithm to generate EWI isochrones in a

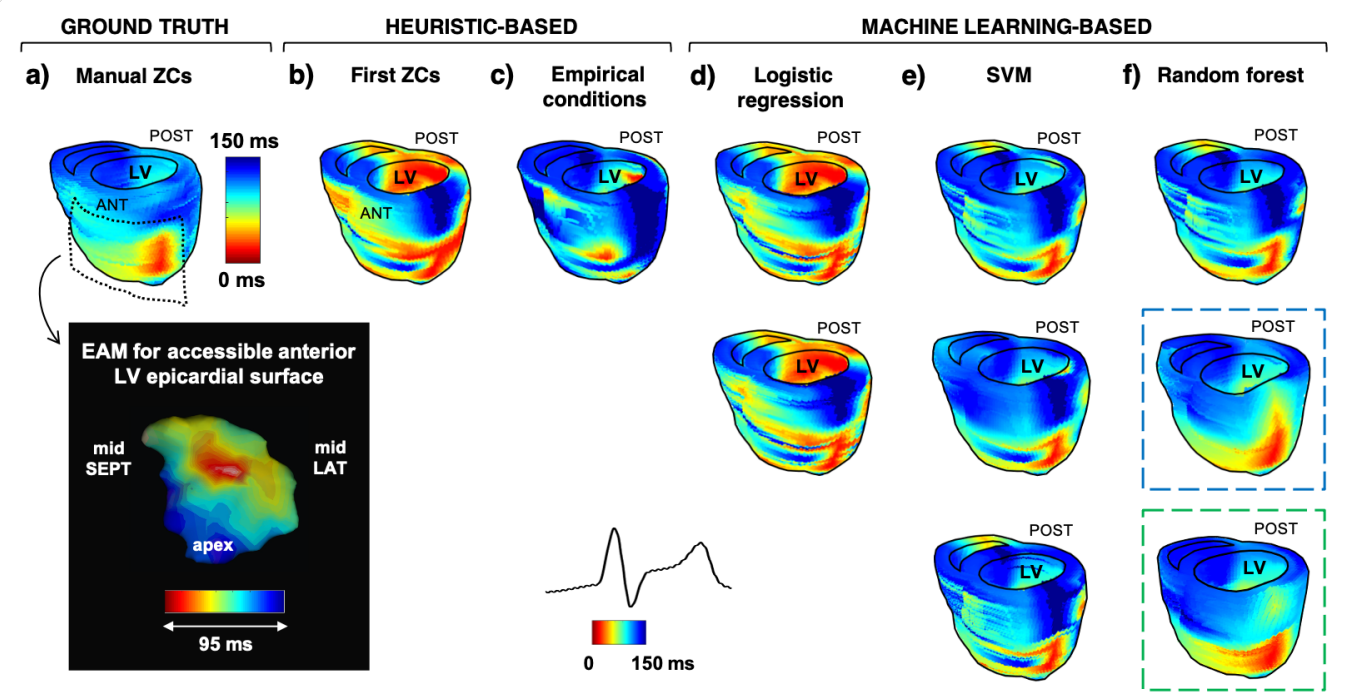


Fig. 4. 3D-rendered isochrones generated with different automated approaches on the example of an antero-lateral LV paced canine taken from the testing dataset compared to the manually generated isochrone. The 3D-rendered isochrones are shown for: a) Manual ground truth; b) Naive heuristic-based automated approach selecting always the first occurring ZCs; c) Rule-based heuristic automated approach selecting the ZCs satisfying conditions (1) and (2); d) ML model with logistic regression classifier; e) ML model with SVM classifier; and f) ML model with Random Forest classifier. The corresponding EnSite electroanatomic map (EAM) is included for a limited epicardial surface of the anterior LV, accessible with the mapping catheter in the chest cavity during the canine experiment. The middle row for (d-f) corresponds to the ML models results with the $\mathbf{2}^{nd}$ voting approach: probability threshold applied to the ZC candidates (4) and satisfying a recall $>$ 70%, while the bottom row for (e-f) sets the condition to a recall $>$ 40%. Another key difference for the very last row is the considerably larger amount of pixels the models were applied on ($\boldsymbol{P}(\boldsymbol{X_i}, \boldsymbol{Y_i})$ with $i \in [\mathbf{1:2500}]$), not previously used in the initial ground truth manual ZC selection. The ML model-based automated algorithms generated the isochrones with their tuned hyperparameters: elastic net L1-ratio = 0.5; kernel = radial basis function and n estimators = 200. ANT=anterior; LAT=lateral, POST=posterior; SEPT=septum.

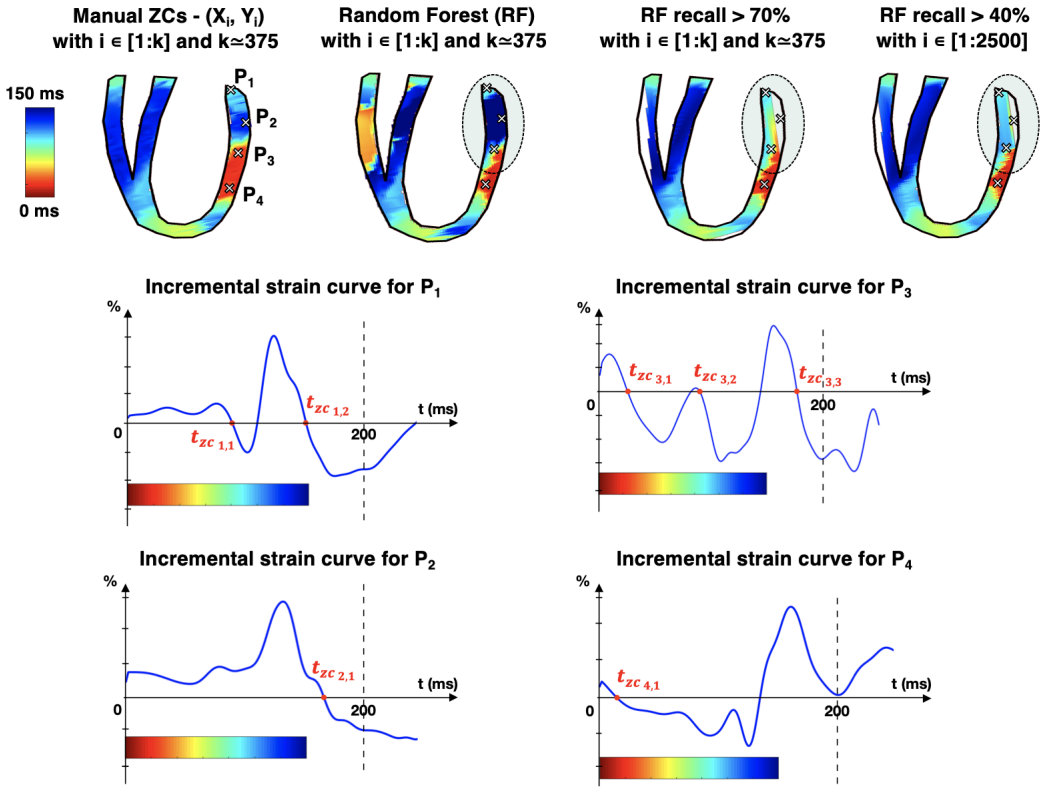
faster and more robust way, with no inter-observer variability. This study also reported the use of machine learning models for the first time for automatically selecting the best ZC candidates on axial incremental strain curves in WPW patients, sinus rhythm humans and LV paced canines. The three ML models (logistic regression, SVM and Random Forest) were trained with manually selected zero-crossing features from past validated cases, therefore, considered as ground truth.

First, we evaluated the models performance with precision-recall curves on the validation and testing datasets (Figure 3). Both SVM and Random Forest models clearly outperformed logistic regression on all three datasets. In fact, the precision of the logistic regression classifier collapsed very quickly below 65% on the testing datasets, as illustrated on the middle and bottom rows precision-recall curves in red. Table IV further described the models' performance for the different voting approaches and probability threshold scenarios used to generate the EWI isochrones in this study. The logistic regression model could not yield a respective recall higher than 57.2% and 61.7% in the human and canine test datasets, despite applying the probability threshold to the ZC candidates with the voting approach. This led once more to situations where the precision was equal to the recall, due to the maximum recall cut-off.

Additionally, even though SVM initially performed as well as Random Forest in terms of precision on the validation and human testing datasets, it ended up not being as generalizable nor transferable to other types of datasets, like in the canine example for which it exhibited a 5% drop in precision (Table IV). We can assume the SVM model was probably more overfitted to the training and validation human datasets. Meanwhile, Random Forest showed robustness and its predictivity was not impacted by the type of dataset. This allowed us to build confidence about future performance of the Random Forest model with respect to dealing with potential data it has not yet been exposed to. Thus, the Random Forest model with 200 estimators or trees was hereby proven to have the best performances overall no matter the dataset it was applied to. Setting a probability threshold upon ZC candidate voting improved the precision from 89.5% to 92.7%, and went as high as 97.5% on the testing datasets at the expense of a significantly lower recall however (40%).

Furthermore, automated operator independent 3D-rendered isochrone generation approaches such as SVM and Random Forest were shown capable of successfully identifying the accessory pathway in WPW patients, the normal early activated basal septum close to the atrioventricular node in sinus rhythm patients and the pacing electrode location in LV paced canines (Fig.4, Fig.5, Fig.7 and Supplemental Figures 1 & 3). The total processing time needed to generate the resulting 3D-rendered EWI activation maps decreased from the usual 90 minutes for

The final version of record is available at http://dx.doi.org/10.1109/TMI.2021.3074808



	Manual ZC	RF ZC	RF ZC recall > 70%	RF ZC recall > 40%
P_1	$t_{zc\ 1,1}$	$t_{zc\ 1,1}$	$t_{zc\ 1,1}$	$t_{zc\ 1,1}$
P_2	$t_{zc\ 2,1}$	$t_{zc\ 2,1}$	none, discarded	none, discarded
P_3	$t_{zc\ 3,1}$	$t_{zc\ 3,2}$	$t_{zc\ 3,2}$	$t_{zc\ 3,2}$
P_4	$t_{zc\ 4,1}$	$t_{zc\ 4,1}$	$t_{zc\ 4,1}$	$t_{zc\ 4,1}$

Fig. 5. Incremental axial strain curves for 4 points in the LV antero-lateral wall of the apical 3.5-chamber view isochrone from the LV paced canine (Fig.4). The circled region over the myocardial isochrone mask outlines the basal region of the antero-lateral wall which displayed a few discrepancies on the corresponding 3D-rendered isochrones (dashed boxes on Fig.4-f) compared to the manually generated one (Fig.4-a). This region also corresponds to the area of lower recall circled in red on Supplemental Figure 1 for the Random Forest model with the voting approaches that apply a probability threshold to the candidates, and consequently matches the area of lower spatial sampling in Supplemental Figure 2. The different zero-crossing candidates are displayed in red on the strain curves for the given 4 points ($\boldsymbol{P_i}$) and the chosen candidates are listed in the Table at the bottom for each of the manual selection and three Random Forest approaches (no voting threshold, recall $>$ 70% and recall $>$ 40%). The Random Forest classifier selected a different ZC candidate than the manual ground truth for $\boldsymbol{P_3}$ and discarded point $\boldsymbol{P_2}$ altogether due to low voting confidence in the case of both approaches with the recall conditions.

the standard manual approach to less than 5 minutes per patient with the machine learning based automated approaches. More particularly, the times required just for the ZC selection step itself are listed in Table V for the three classifiers on the entire human and canine test datasets.

Besides, Random Forest was deemed to have the best performance visually with the second voting approach while setting the condition on the recall evaluation metric (Fig.4-f and Fig.7-f), outperforming SVM with the same probability threshold voting approach (Fig.4-e and Fig.7-e). These findings were all the more true when considering a larger amount of pixels (2500 points) per apical view, not used in the initial ground truth manual ZC selection. Even if the incremental axial strain curves and associated features were no longer labeled, as they were not previously manually annotated, this allowed us to go as low as 40% for the recall value. In fact, at this point we were interested in the highest possible precision, no matter the recall. The resulting EWI isochrone activation patterns were very satisfying. Despite discarding about 60% of the pixels per view and only making ZC predictions on 1000 points out of the 2500 (still more than four times the amount of pixels currently used for the standard manually generated isochrones), the precision rose above 97%.

TABLE V
PROCESSING TIME REQUIRED FOR ZERO-CROSSING SELECTION

ML model	Human test dataset	Canine test dataset
Logistic regression	1.6 seconds	1.9 seconds
SVM	43.9 seconds	43.2 seconds
Random Forest	40.9 seconds	94.5 seconds

The processing times include both the fitting and ZC predicion times for each machine learning model on the entire datasets of annotated strain curves.

The final version of record is available at http://dx.doi.org/10.1109/TMI.2021.3074808

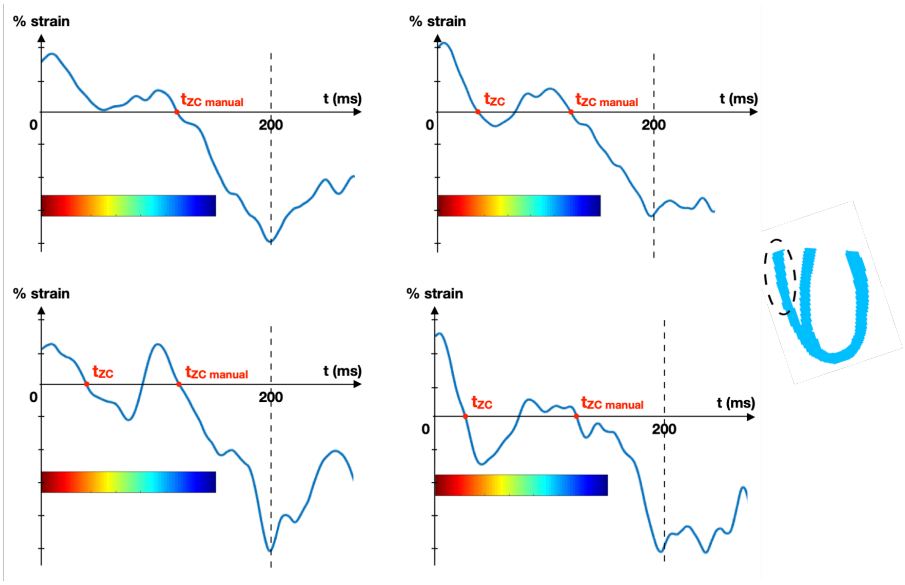


Fig. 6. Incremental axial strain curves for 4 points in the RV free wall (black dashed outline) from the apical 3.5-chamber view isochrone of the antero-lateral LV paced canine (Fig.4 and Fig.5). The different zero-crossing candidates are displayed in red on the curves. The circled region over the myocardial mask corresponds to the area of recovered precision from Supplemental Figure 1 for the Random Forest model with the 2nd voting approach: probability threshold applied to the ZC candidates while satisfying a given set recall value.

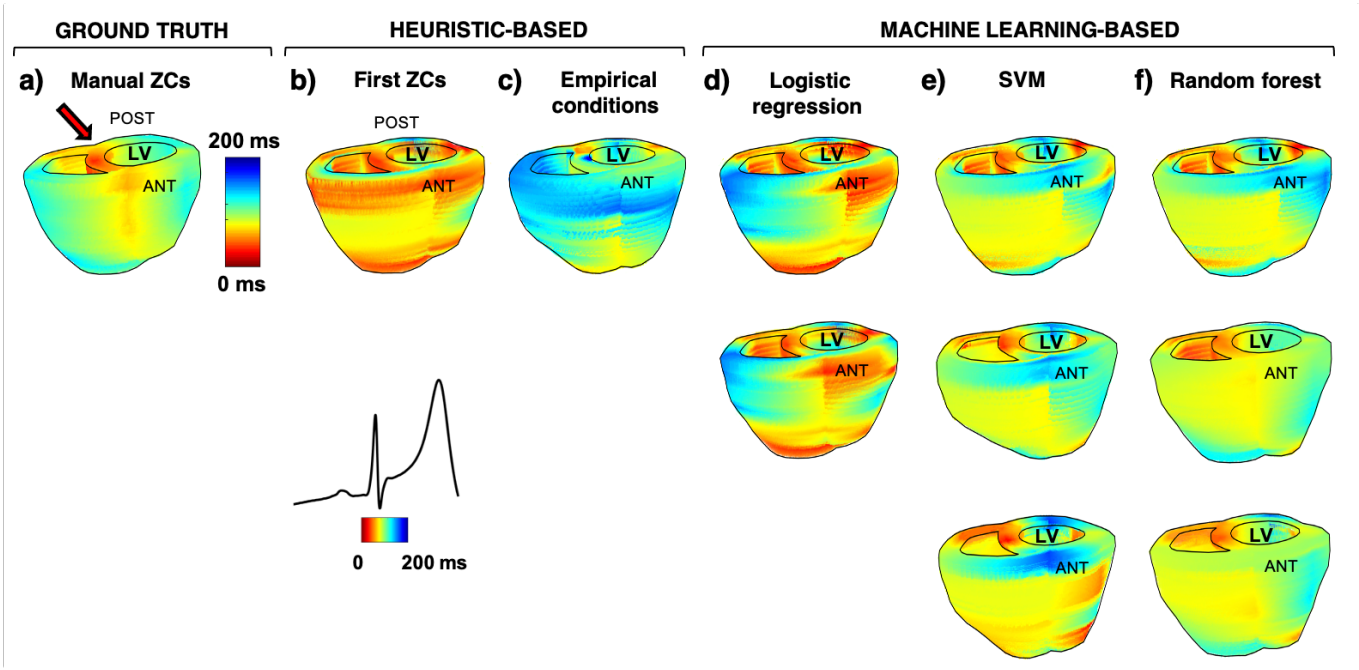


Fig. 7. 3D-rendered isochrones generated with different automated approaches on the example of a patient in sinus rhythm taken from the human testing dataset compared to the manually generated isochrone. The 3D-rendered isochrones are shown for: a) Manual ground truth; b) Naive heuristic-based automated approach selecting always the first occurring ZCs; c) Rule-based heuristic automated approach selecting the ZCs satisfying conditions (1) and (2); d) ML model with logistic regression classifier; e) ML model with SVM classifier; and f) ML model with Random Forest classifier. The middle row for (d-f) corresponds to the ML models results with the $\mathbf{2}^{nd}$ voting approach: probability threshold applied to the ZC candidates (4) and satisfying a recall $>$ 70%, while the bottom row for (e-f) sets the condition to a recall $>$ 40%. Another key difference for the very last row is the considerably larger amount of pixels the models were applied on $(\boldsymbol{P}(\boldsymbol{X_i}, \boldsymbol{Y_i})$ with $i \in [\mathbf{1:2500}])$, not previously used in the initial ground truth manual ZC selection. The ML model-based automated algorithms generated the isochrones with their tuned hyperparameters: elastic net L1-ratio = 0.5; kernel = radial basis function and n estimators = 200. ANT=anterior; LAT=lateral, POST=posterior; SEPT=septum.

Lastly, since the electromechanical activation constitutes a wavefront, it is expected to propagate continuously. Therefore, we can assume when having several ZC candidates to choose from, that the correct and physiological approach would be to select the ZC that gives the smoothest transition in the activation gradient in order to decrease the amount of sudden "jumps" or discontinuities in activation times. The incremental strain curves, multi-2D isochrones and corresponding pixel locations in the segmented masks, shown in Figure 6, Supplemental Figures 1 and 2 respectively, highlight that particular

The final version of record is available at http://dx.doi.org/10.1109/TMI.2021.3074808

interesting behavior the Random Forest model adopted when faced with such a scenario. Three out of the four curves displayed the presence of a second earlier ZC candidate, different from the manually selected ground truth ZC (Fig. 6). An operator manually generating this 3.5-chamber view isochrone would most probably have applied a "nearest neighbor" approach: inferring which ZC candidate was the correct physiological one by looking at the nearest single stable ZC and choosing accordingly between the multiple candidates in the neighboring points. It seems the Random Forest classifier intrinsically applied a similar reasoning. This could also be justified by the highest amplitude of the second ZC negative peak strain. In fact, as previously mentioned when introducing the heuristic-based automated isochrone generation approaches in the Methods, applying a condition on the negative peak amplitude is a common bias experienced by the operator upon his manual ZC selection process. Thus, it is possible that the training datasets included information regarding one zero-crossing candidate disappearing while another one appears with higher negative peak strain amplitudes, and the ML model might have been predisposed to identifying such relationships because they were embedded in the training features. Investigating and further characterizing the effect of single versus multi-expert ground truth annotation on training bias is a topic of ongoing and future investigation in our group. Moreover, future studies should consider implementing a nearest neighbour with smooth activation gradient method as a more sophisticated heuristic-based automated approach and comparing the results to machine learning approaches.

Nevertheless, there were a few limitations to this study. The automated algorithms were only applied to ventricular isochrone generation with positive-to-negative ZCs. Upward negative-to-positive ZCs for other types of views such as parasternal long-axis or atrial isochrones were not considered in the training process. Including negative-to-positive ZCs might help remove discontinuities in the isochrone activation patterns at the apex due to complex fiber orientation in that region. In addition, the ground truth datasets only included sinus rhythm cases or WPW patients that were otherwise known to be healthy with no other co-morbidities. These datasets also did not include other types of cardiac arrythmias. Besides, the limited number of cases presented in this paper might raise some concern on repeatability of the machine learning approaches. The present study examines automated EWI isochrone generation with machine learning for the first time and is an initial evaluation and optimization of the best classifier for zero-crossing selection. Future investigations still need to be conducted to further refine the Random Forest classifier and assess its reproducibility, by adding for instance more categorical features (e.g., type of cardiac rhythm) as inputs to the ML model to potentially improve its robustness and transferability even more. Considering the totality of the points contained in the segmented myocardium mask in each view: (X_i, Y_i) with $i \in [1 : k]$ and $k > 40000$, instead of a subset of pixels with $k \leq 2500$, could also improve the isochrone outcomes at low recall values and decrease the suboptimal portion of the results.

In the future, we could also explore the use of a different learning approach, e.g., semi-supervised learning to increase the number of labeled samples included in the training dataset. In addition, we could envision using a deep learning approach with no manual feature engineering step required, where the model would extract the features on its own. This could be achieved by directly feeding the incremental strain curve signals to a convolutional neural network (CNN) as the input data. We will exploit the capability of convolutional layers to extract useful knowledge and learn the internal representation of time-series data [60]. In fact, in CNNs, different features are extracted through convolution using filters whose weights are automatically learned during training [61], [62]. In our case, the automated isochrone generation process would consist of a regression task [63], in which the CNNs would need to find the correct ZC positions on the temporal strain curves, and not classify the strain curves themselves.

Finally, for real-time implementation of EWI, we can conceive a combination of automated followed by manually adjusted ZC selection for isochrone generation. In fact, an observer's input could help the ML classifier by manually labeling a limited number of samples for which the algorithm has the least confidence in (e.g., some of the discarded pixels with lower ZC voting probabilities when the recall condition was set to 40%). Through "active learning", we could therefore further improve the model's learning process.

V. Conclusion

Electromechanical Wave Imaging with machine learning was hereby illustrated to automatically detect zero-crossing time points on incremental axial strain curves for a faster, more robust, and less operator dependent isochrone generation process. The Random Forest classifier was identified as the best performing algorithm compared to logistic regression and SVM: capable of identifying accessory pathways as well as pacing locations in humans and canines respectively, while also resulting in the most precise isochrone activation patterns. Finally, these findings indicate that the standard manual processing pipeline required to obtain EWI activation maps can now considerably be abbreviated without a significant trade-off in accuracy, towards real-time implementation of EWI and its automated translation to the clinic.

Acknowledgements

The authors would like to thank Alexandre Yahi, PhD from the Department of Biomedical Informatics at Columbia, for sharing precious inputs on the feature engineering step and providing valuable insights on the machine learning models implementation. The authors also express their sincere appreciation to Julien Grondin, PhD, Vincent Sayseng, PhD, Jad El Harake, MS, and Stephen Lee, BS of the Ultrasound and Elasticity Imaging Laboratory, Department of Biomedical Engineering at Columbia, for all their helpful suggestions and discussions.

References

[1] J. A. Finegold, P. Asaria, and D. P. Francis, "Mortality from ischaemic heart disease by country, region, and age: statistics from World Health Organisation and United Nations", *International Journal of Cardiology*, vol. 168, no 2, pp. 934-945, 2013.

The final version of record is available at http://dx.doi.org/10.1109/TMI.2021.3074808

[2] D. Mozaffarian, et al. "Executive summary: heart disease and stroke statistics—2016 update: a report from the american heart association", *Circulation*, vol. 133, no 4, pp. 447-454, 2016.

[3] S. Kaptoge, et al. "World Health Organization cardiovascular disease risk charts: revised models to estimate risk in 21 global regions", *The Lancet Global Health*, vol. 7, no 10, pp. e1332-e1345, 2019.

[4] J. Mackay and G. A. Mensah, "The atlas of heart disease and stroke", *World Health Organization*, 2004.

[5] G. A. Roth, et al. "Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017", *The Lancet*, vol. 392, no 10159, pp. 1736-1788, 2018.

[6] P. A. Heidenreich, et al. "Forecasting the future of cardiovascular disease in the united states: a policy statement from the american heart association", *Circulation*, vol. 123, no 8, pp. 933-944, 2011.

[7] J. N. Cohn, et al. "Screening for early detection of cardiovascular disease in asymptomatic individuals", *American Heart Journal*, vol. 146, no 4, pp. 679-685, 2003.

[8] J. R. Smith, et al. "Principles for national and regional guidelines on cardiovascular disease prevention: a scientific statement from the World Heart and Stroke Forum", *Circulation*, vol. 109, no 25, pp. 3112-3121, 2004.

[9] K. J. Greenlund, N. L. Keenan, P. F. Clayton, D. K. Pandey, and Y. Hong, "Public health options for improving cardiovascular health among older Americans", *American Journal of Public Health*, vol. 102, no 8, pp. 1498-1507, 2012.

[10] C. Ramanathan, R. N. Ghanem, P. Jia, K. Ryu and Y. Rudy, "Noninvasive electrocardiographic imaging for cardiac electrophysiology and arrhythmia", *Nature Medicine*, vol. 10, no 4, pp. 422-428,2004.

[11] C. Knackstedt, P. Schauerte and P. Kirchhof, " Electro-anatomic mapping systems in arrhythmias", *Europace*, vol. 10, no suppl. 3, pp. iii28-iii34, 2008.

[12] E. H. Ibrahim, "Myocardial tagging by cardiovascular magnetic resonance: evolution of techniques–pulse sequences, analysis algorithms, and applications", *Journal of Cardiovascular Magnetic Resonance*, vol. 13, no 1, pp. 36, 2011.

[13] D. Duprez, "Early detection of cardiovascular disease-the future of cardiology", *European Society of Cardiology e-Journal of Cardiolology Practice*, vol. 4, no 19, pp. 1, 2006.

[14] P. M. Van Dam, T. F. Oostendorp, A. C. Linnenbank and A. Van Oosterom, "Non-invasive imaging of cardiac activation and recovery", *Annals of Biomedical Engineering*, vol. 37, no 9, pp. 1739-1756, 2009.

[15] M. Cikes, L. Tong, G. R. Sutherland and J. D'Hooge, "Ultrafast cardiac ultrasound imaging: technical principles, applications, and clinical benefits", *JACC: Cardiovascular Imaging*, vol. 7, no 8, pp. 812-823, 2014.

[16] S. Urheim, T. Edvardsen, H. Torp, B. Angelsen and O. A. Smiseth, "Myocardial strain by Doppler echocardiography: validation of a new method to quantify regional myocardial function", *Circulation*, vol. 102, no 10, pp. 1158-1164, 2000.

[17] A. L. Klibanov and J. A. Hossack , "Ultrasound in radiology: from anatomic, functional, molecular imaging to drug delivery and image-guided therapy", *Investigative radiology*, vol. 50, no 9, pp. 657, 2015.

[18] J. Provost, W. N. Lee, L. Fujikura and E. E. Konofagou, "Electromechanical wave imaging of normal and ischemic hearts in vivo", *IEEE Transactions on Medical Imaging*, vol. 29, no 3, pp. 625-635, 2009.

[19] J. Provost, V. Gurev, N. Trayanova and E. E. Konofagou, "Mapping of cardiac electrical activation with electromechanical wave imaging: an in silico–in vivo reciprocity study", *Heart Rhythm*, vol. 8, no 5, pp. 752-759, 2011.

[20] J. Provost, W. N. Lee, L. Fujikura and E. E. Konofagou, "Imaging the electromechanical activity of the heart in vivo", *Proceedings of the National Academy of Sciences*, vol. 108, no 21, pp. 8565-8570, 2011.

[21] M. Tanter and M. Fink, "Ultrafast imaging in biomedical ultrasound", *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 61, no 1, pp. 102-119, 2014.

[22] J. Synnevag, A. Austeng, and S. Holm, "Benefits of minimum-variance beamforming in medical ultrasound imaging", *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 56, no 9, pp. 1868-1879, 2009.

[23] M. A. Lediju Bell, R. Goswami, J. A. Kisslo, J. J. Dahl, and G. E. Trahey, "Short-lag spatial coherence imaging of cardiac ultrasound data: Initial clinical results", *Ultrasound in Medicine and Biology*, vol. 39, no 10, pp. 1861-1874, 2013.

[24] L. Tong, H. Gao, and J. D'hooge, "Multi-transmit beam forming for fast cardiac imaging-a simulation study", *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 60, no 8, pp. 1719-1731, 2013.

[25] A. Ortega, et al. "A comparison of the performance of different multiline transmit setups for fast volumetric cardiac ultrasound", *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 63, no 12, pp. 2082-2091, 2016.

[26] F. Viola and W. F. Walker, "A comparison of the performance of time-delay estimators in medical ultrasound", *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 50, no 4, pp. 392-401, 2003.

[27] G. F. Pinton, J. J. Dahl, and G. E. Trahey, "Rapid tracking of small displacements with ultrasound", *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 53, no 6, pp. 1103-1117, 2006.

[28] K. F. Kvale, et al. "Detection of regional mechanical activation of the left ventricular myocardium using high frame rate ultrasound imaging", *IEEE Transactions on Medical Imaging*, vol. 38, no 11, pp. 2665-119, 2675.

[29] K. F. Kvale, et al. "Comparison of two methods for mechanical activation detection using high frame rate ultrasound imaging", *2019 IEEE International Ultrasonics Symposium*, vol. 61, no 1, pp. 1410-1413, 2019.

[30] L. Melki, A. Costet, and E. E. Konofagou, "Reproducibility and angle independence of electromechanical wave imaging for the measurement of electromechanical activation during sinus rhythm in healthy humans", *Ultrasound in Medicine and Biology*,vol. 43, no 10, pp. 2256–2268, 2017.

[31] L. Melki, et al. "Localization of accessory pathways in pediatric patients with Wolff-Parkinson-White syndrome using 3d-rendered electromechanical wave imaging," *JACC: Clinical Electrophysiology*, vol. 5, no 4, pp. 427-437, 2019.

[32] C. S. Grubb, et al. "Noninvasive localization of cardiac arrhythmias using electromechanical wave imaging," *Science Translational Medicine*, vol. 12, no 536, 2020.

[33] E. Bunting, L. Lambrakos, P. Kemper, W. Whang, H. Garan, and E. E. Konofagou, "Imaging the propagation of the electromechanical wave in heart failure patients with cardiac resynchronization therapy", *Pacing and Clinical Electrophysiology*, vol. 40, no 1, pp. 35–45, 2017.

[34] R. J. G. Van Sloun, R. Cohen, and Y. C. Eldar, "Deep learning in ultrasound imaging", *Proceedings of the IEEE*, vol. 108, no 1, 2020.

[35] A. A. Nair, T. D. Tran, A. Reiter, and M. A. L. Bell, "A deep learning based alternative to beamforming ultrasound images", in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 3359–3363, 2018.

[36] A. C. Luchies and B. C. Byram, "Deep neural networks for ultrasound beamforming", *IEEE Transactions on Medical Imaging*, vol. 37, no 9, pp. 2010–2021, 2018.

[37] B. Byram and M. Jakovljevic, "Ultrasonic multipath and beamforming clutter reduction: a chirp model approach", *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 61, no 3, pp. 428-440, 2014.

[38] K. Dei and B. Byram, "A robust method for ultrasound beamforming in the presence of off-axis clutter and sound speed variation", *Ultrasonics*, vol. 89, pp. 34–45, 2018.

[39] L. R. Sultan, T. W. Cary, and C. M. Sehgal, "Machine learning to improve breast cancer diagnosis by multimodal ultrasound", in *2018 IEEE International Ultrasonics Symposium*, pp. 1–4, 2018.

[40] E. Smistad, I. M. Salte, A. Ostvik, S. Leclerc, O. Bernard, and L. Lovstakken, "Segmentation of apical long axis, four-and two-chamber views using deep neural networks", in *2019 IEEE International Ultrasonics Symposium*, pp. 8–11, 2019.

[41] S. Leclerc, et al. "Deep learning for segmentation using an open large-scale dataset in 2D echocardiography", *IEEE Transactions on Medical Imaging*, vol. 38, no 9, pp. 2198–2210, 2019.

[42] G.R. Sutherland, G. Di Salvo, P. Claus, J. D'Hooge, and B. Bijnens, "Strain and strain rate imaging: a new clinical approach to quantifying regional myocardial function", *Journal of the American Society of Echocardiography*, vol. 17, no 7, pp. 788-802, 2004.

[43] H. Geyer, et al. , "Assessment of myocardial mechanics using speckle tracking echocardiography: fundamentals and clinical applications", *Journal of the American Society of Echocardiography*, vol. 23, no 4, pp. 351-369, 2010.

[44] V. Mor-Avi, et al., "Current and evolving echocardiographic techniques for the quantitative evaluation of cardiac mechanics: ASE/EAE consensus statement on methodology and indications endorsed by the Japanese Society of Echocardiography", *European Journal of Echocardiography*, vol. 12, no 3, pp. 167-205, 2011.

The final version of record is available at http://dx.doi.org/10.1109/TMI.2021.3074808

[45] M. Tabassian, M. Alessandrini, R. Jasaityte, L. De Marchi, G. Masetti, and J. D'hooge, "Handling missing strain (rate) curves using k-nearest neighbor imputation", in *2016 IEEE International Ultrasonics Symposium*, pp. 1–4, 2016.

[46] M. Tabassian, O. Zulaicaiglesias, S. Unlu, J. Voigt, and J. D'hooge, "Machine learning for quality assurance of myocardial strain curves", in *2018 IEEE International Ultrasonics Symposium*, pp. 1–4, 2018.

[47] A. Yahav, G. Zurakhov, O. Adler, and D. Adam, "Strain curve classification using supervised machine learning algorithm with physiological constraints", *Ultrasound in Medicine and Biology*, 2020.

[48] M. Tabassian, et al. "Machine learning of the spatio-temporal characteristics of echocardiographic deformation curves for infarct classification", *The International Journal of Cardiovascular Imaging*, vol. 33, no 8, pp. 1159–1167, 2017.

[49] J. Luo and E.E. Konofagou, "A fast normalized cross-correlation calculation method for motion estimation", *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* , vol. 57, no 6, pp.1347-1357, 2010.

[50] F. Kallel and J. Ophir, "A least-squares strain estimator for elastography", *Ultrasonic imaging*, vol. 19, no 3, pp. 195-208, 1997.

[51] P. Nauleau, L. Melki, E. Wan, and E. E. Konofagou, "A 3-D rendering algorithm for electromechanical wave imaging of a beating heart", *Medical Physics*, vol. 44, no 9, pp. 4766–4772, 2017.

[52] J. Gao and J. M. Rubin, "Ultrasound strain zero-crossing elasticity measurement in assessment of renal allograft cortical hardness: a preliminary observation", *Ultrasound in Medicine and Biology*, vol. 40, no 9, pp. 2048–2057, 2014.

[53] M. V. Andersen, et al. "Quantitative parameters of high-frame-rate strain in patients with echocardiographically normal function", *Ultrasound in Medicine and Biology*, vol. 45, no 5, pp. 1197–1207, 2019.

[54] D. Bhakta and J. M. Miller, "Principles of electroanatomic mapping", *Indian pacing and electrophysiology journal*, vol. 8, no 1, pp. 32–50, 2008.

[55] D. R. Cox, "The regression analysis of binary sequences", *Journal of the Royal Statistical Society: Series B (methodological)*, vol. 20, no 2, pp. 215–232, 1958.

[56] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net", *Journal of the Royal Statistical Society: Series B (statistical methodology)*, vol. 67, no 2, pp. 301–320, 2005.

[57] C. Cortes and V. Vapnik, "Support-vector networks", *Machine Learning*, vol. 20, no 3, pp. 273–297, 1995.

[58] L. Breiman, "Random forests", *Machine Learning*, vol. 45, no 1, pp. 5–32, 2001.

[59] C. Kingsford and S. L. Salzberg, "What are decision trees?", *Nature Biotechnology*, vol. 26, no 9, pp. 1101–1013, 2008.

[60] H. I. Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P. Muller, "Deep learning for time series classification: a review", *Data Mining and Knowledge Discovery*, vol. 33, no 4, pp. 917–963, 2019.

[61] L. Sadouk, "CNN Approaches for Time Series Classification", in *Time Series Analysis-Data, Methods, and Applications. IntechOpen*, 2018.

[62] J. Yang, M. N. Nguyen, P. P. San, X. L. Li, and S. Krishnaswamy, "Deep convolutional neural networks on multichannel time series for human activity recognition", in *Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2015.

[63] M. Binkowski, G. Marti, and P. Donnat, "Autoregressive convolutional neural networks for asynchronous time series", in *International Conference on Machine Learning*, pp. 580–589, 2018.