Meta-learning based Active Learning Approach for Computer-Assisted Pace-Mapping

Pradeep Bajracharya[1], Dylan O'Hara[1], Casey Meisenzahl[1], Karli Gillette[2], Anton J Prassl[2], Gernot Plank[2], John L Sapp[3], Linwei Wang[1]

[1] Rochester Institute of Technology, Rochester, NY, USA
[2] Gottfried Schatz Research Center, Division of Medical Physics and Biophysics, Medical University of Graz, Graz, Austria
[3] QEII Health Sciences Centre, Dalhousie University, Halifax, Nova

Abstract

Cardiac diseases are the leading cause of death worldwide, with ventricular tachycardia (VT) being a major contributor. Successful VT intervention requires precise localization of the abnormal activation source, typically achieved through pace-mapping i.e. pacing different sites in the heart and comparing the resulting ECGs to the clinical VT ECG. Pace-mapping is invasive and time-consuming, motivating the search for more efficient alternatives. Active learning methods using Gaussian process (GP) surrogate models have shown promise in reducing data requirements, but they rely on initial labeled data and cannot transfer knowledge between tasks. We propose a meta-learning–based neural active surrogate approach that transfers the knowledge learned from one pacing localization task to subsequent tasks without requiring any seed labeled data. Evaluated across two conditions (sinus and infarcted) in a heart geometry, our method achieved an 94% and 70% reduction in localization steps compared to random search and GP-based Bayesian optimization, respectively.

1. Introduction

Cardiac diseases are the leading cause of death worldwide [1] with ventricular tachycardia (VT) being a major contributor (more than half [2]) to these statistics. VT is characterized by abnormally rapid heartbeats due to irregular electrical activation originating in the ventricles, the lower chamber of the heart [3]. This results in a disruption of blood flow to the rest of the body leading to sudden cardiac death [2]). Successful intervention requires localization of the source of abnormal activation. A common approach is to stimulate various sites in the heart and then analyze the resulting ECG, comparing with the observed faulty ECG [4]. This approach is called *pacemapping*.

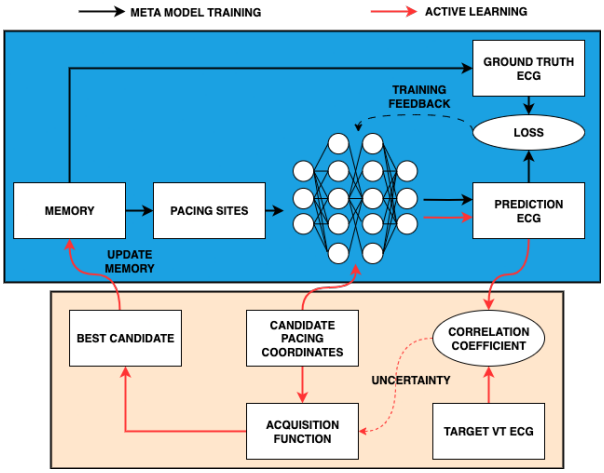

Figure 1. Block Diagram of the Proposed Method

However, this method is largely invasive and relies on trial-and-error approaches that depend on rapid interpretation and informed assessments by trained cardiologists.

There have been deep learning and machine learning based approaches that estimate the origin of VT signals. Deep learning approaches [5–7] train a surrogate model to map the relation between ECG and pacing sites by training on a large dataset from diverse cohort of patients where trained models are then applied to new patients. These methods are data intensive and rely on an assumption that new patients would lie in the same distribution as the training data. Contrary to these, there are active-learning (AL) based approaches that learns a patient-specific surrogate model, such as Gaussian processes (GP), to capture the relationship between pacing sites and the target ECG (i.e. error metrics between predicted and target ECG signal[8]) for a specific heart. The uncertainty estimates provided by the GP guide the exploration of the search space, selecting subsequent pacing sites via a Bayesian optimization

strategy that balances exploration and exploitation. This process minimizes the amount of pacemapping required to obtain sufficient training data for accurate localization.

A key limitation of this approach is that the surrogate model is tied to mapping relation to a specific target ECG. As a result, each new VT to localize (which we refer to as *tasks* hereafter), characterized by their own distinct target ECG, requires the development of a separate surrogate model. This prevents the reuse of a trained surrogate across tasks, thereby limiting transfer of knowledge.

In this work, our goal is to avoid starting the active learning process from scratch for every new task. We achieve this through three key innovations. First, instead of approximating a function tied to a specific target ECG, we design a surrogate neural network model that learns a function shared between tasks by representing ECGs as a function of pacing sites. Second, we adapt active learning to train the surrogate model efficiently, requiring only a small amount of data for each task. Finally, to ensure that the surrogate model retains knowledge across tasks, we propose a meta-learning-based active surrogate modeling framework that integrates neural surrogates with active learning to enable efficient knowledge transfer in pacemapping. By combining continual meta-learning with uncertainty-driven sampling, the framework reduces the number of required pacing sites while maintaining accuracy, thus improving the overall efficiency of the pacemapping process.

2. Method

In this section, we define three main components of our proposed method namely the deep neural network surrogate model, active strategy and meta-learning approach to promote transfer of knowledge.

2.1. Deep Neural Network Surrogate

A deep neural network (DNN) surrogate meta-model is used to map the relationship between the pacing sites and their corresponding ECG signals. The model uses a multilayer perceptron (MLP) architecture with three hidden layers of 128 nodes each, followed by ReLU activations. The output layer consists of two heads that predict the mean (m) and standard deviation (s), with the final prediction sampled from a normal distribution $N(m, s)$.

If $y_q = [y_{q_1}, y_{q_2}, ..., y_{q_t}]$ denotes the q^{th} predicted ECG signal of length t for some input $x_q = (U_q, V_q, W_q)$ where t is the total length of the signal.

$$y_q = T(x_q) = T(U_q, V_q, W_q) \tag{1}$$

2.2. Active Learning

2.2.1. Acquisition Function

To guide the selection of new pacing sites, we define an acquisition function (AF), in terms of the neural surrogate model introduced in Section 2.1. Specifically, at any pacing site x, we predict an ECG signal using the surrogate model, which is compared to the target VT ECG y_{target} using the correlation coefficient (CC) as a similarity metric. Bayesian optimization with the expected improvement (EI) function is then used, balancing the exploration and exploitation based on the predicted mean $\mu(\mathbf{x})$ and uncertainty $\sigma(\mathbf{x})$ of this CC.

$$\begin{aligned} \mathrm{EI}(\mathbf{x}) \quad &= (\mu(\mathbf{x}) - \mathbf{f}^{+})\Phi\left(\tfrac{\mu(\mathbf{x})-\mathbf{f}^{+}}{\sigma(\mathbf{x})}\right) + \\ &\quad \sigma(\mathbf{x})\phi\left(\tfrac{\mu(\mathbf{x})-\mathbf{f}^{+}}{\sigma(\mathbf{x})}\right), \end{aligned} \tag{2}$$

where, f^* is the best observed CC so far, and Φ and ϕ are the CDF and PDF of the standard normal distribution, respectively.

2.2.2. Correlation Coefficient and Uncertainty

The $\mu(\mathbf{x})$ and $\sigma(\mathbf{x})$ in Eq. 2 relies on the outputs of the neural surrogate model. For a candidate pacing site x_d, we obtain a predicted ECG signal from the surrogate model. As the prediction includes high-frequency noise, it is first passed through a band-pass filter to obtain a denoised version y_d, where the standard deviation std of the removed noise is also computed. Using y_d, the mean and uncertainty of the CC with respect to the target VT ECG are estimated as follows:

$$\begin{aligned} \mu(x) \quad &= CC(y_d, y_{target}) \\ CC_{upper} \quad &= CC(y_d + std, y_{target}) \\ CC_{lower} \quad &= CC(y_d - std, y_{target}) \\ \sigma(x) \quad &= CC_{upper} - CC_{lower} \end{aligned} \tag{3}$$

2.2.3. Stopping Criteria

The newly acquired pacing location and its respective ECG signal is then used to update the neural surrogate model. The process is repeated until the correlation coefficient between the predicted ECG and the target ECG is greater or equal to 0.97. This means that the localized pacing sites are within $5mm$ of the target pacing site that resulted in the VT signal.

2.3. Meta Learning

To enable pacing site localization across multiple tasks or patients, we introduce a meta-learning strategy based on Meta-Experience Replay (MER). To ensure continual adaptation across tasks $h \in H$, we maintain a compact memory buffer M of fixed size m, initially empty, which stores previously acquired labeled data and is updated throughout training. We also track g, that records the age of the memory, i.e. the total number of memory updates performed.

2.3.1. Model and Memory Initialization

For the first task, we initialize the memory M with the available labeled data L. The surrogate meta-model is then trained, and active learning is applied iteratively, following the procedure described in Section 2.2.1. In the multitask setting, the key distinction is that after each active learning iteration, memory M is updated to incorporate newly acquired labeled data using reservoir sampling as outlined in Algorithm 1. This ensures that the meta-model retains the knowledge across tasks while continuing to adapt to new localization problems.

Algorithm 1 Memory Update

Require: Newly acquired data pair $(X_{\text{new}}, Y_{\text{new}})$, memory buffer M, memory size limit m, current age g

if size$(M) < m$ **then**
  $M \leftarrow M \cup \{(X_{\text{new}}, Y_{\text{new}})\}$
else
  $index \leftarrow \text{random}(0, g)$
  if $index <$ size(M) **then**
    $M[index] \leftarrow (X_{\text{new}}, Y_{\text{new}})$
  end if
end if

2.3.2. Subsequent Tasks and Model Updates

Algorithm 2 outlines the iterative active learning and meta-learning process. For each task h, the memory buffer M and meta-model T are carried forward. At the beginning of each active learning iteration, snapshots T_{prior} and T_{post} of the meta-model are saved. The new data point is combined with memory samples to update T_{post} over e epochs, after which T is updated based on the difference between T_{post} and T_{prior}. Finally, the memory is then updated according to Algorithm 1. This cycle repeats for each subsequent tasks until the stopping criteria is reached.

Algorithm 2 Meta-Learning with Active Learning and Memory Update for task $h \in H$

Require: Memory buffer M, meta model T, age g, target ECG y_{target}, candidate data U_h, learning rate α, meta learning rate β

repeat
  $(X_{\text{new}}, Y_{\text{new}}) \leftarrow \text{ActiveLearning}(T, U_h)$
  $T_{\text{prior}} \leftarrow \text{copy}(T)$
  $T_{\text{post}} \leftarrow \text{copy}(T_{\text{prior}})$
  for e in epochs **do**
    $L_h = (X_{\text{new}}, Y_{\text{new}}) \cup \text{shuffle}(M)$
    $\theta_{T_{\text{post}}} \leftarrow \theta_{T_{\text{post}}} - \alpha \cdot \nabla_{\theta_{T_{\text{post}}}} \text{Loss}(L_h)$
    $\theta_T \leftarrow \theta_{T_{\text{prior}}} + \beta \cdot (\theta_{T_{\text{post}}} - \theta_{T_{\text{prior}}})$
  end for
  $\text{MemoryUpdate}(M, (X_{\text{new}}, Y_{\text{new}}), g)$
until $\text{CC}(y_{\text{target}}, Y_{\text{new}}) > 0.97$

3. Experiment and Results

3.1. Data

We consider a single human biventricular model obtained from the Experimental Data and Geometric Analysis Repository (EDGAR) [9]. The model includes a healthy (sinus) condition and an infarcted condition. For each condition, 12-lead ECGs were calculated for 924 pacing sites under healthy conditions and 687 pacing sites under infarcted condition. ECG signals are processes such that each lead signal is of length 170.

3.2. Baselines

We compare our results against two baselines namely: *Random* and *GP-based Bayesian Optimization (GP BO)*. In the *Random*, candidate pacing sites are randomly selected from the coordinate search space without considering model uncertainty. In contrast, the *GP BO* leverages the uncertainty of a task-specific surrogate model, modeled for individual target VT ECG for each patient, to guide the selection of the next candidate site. Importantly, neither baseline incorporates knowledge transfer, as a new surrogate model is created independently for each task.

3.3. Metrics

The number of candidate pacing sites required for the respective predicted ECG to reach CC of 97% with the target VT ECG signal is measured.

3.4. Experiment Setup

For the initialization (first) task, the meta-model is seeded with a total of five labeled data points, which are

(A) SINUS CONDITION

METHODS	TASK 1	TASK 2	TASK 3	TASK 4	TASK 5	TASK 6
RANDOM	78.80+/-18.60	62.80+/-33.95	78.60+/-12.65	78.00+/-21.00	69.60+/-26.27	74.60+/-23.83
GP BO	15.80+/-4.66	12.30+/-7.23	11.20+/-4.89	7.40+/-1.43	8.90+/-3.75	9.20+/-2.14
OURS	**6.80+/-1.83**	**3.70+/-1.73**	**3.40+/-2.62**	**1.90+/-1.45**	**4.60+/-3.72**	**1.80+/-0.87**
METHODS	TASK 7	TASK 8	TASK 9	TASK 10	TASK 11	TASK 12
RANDOM	85.00+/-0.00	63.20+/-33.21	70.80+/-28.62	68.60+/-26.25	75.60+/-20.63	85.00+/-0.00
GP BO	7.90+/-1.51	7.10+/-0.94	10.40+/-4.76	20.50+/-11.11	15.70+/-5.98	18.30+/-7.29
OURS	**2.70+/-1.00**	**2.30+/-1.27**	**2.40+/-2.20**	**4.10+/-2.02**	**3.70+/-1.42**	**5.30+/-2.49**

(B) INFARCTED CONDITION

METHODS	TASK 1	TASK 2	TASK 3	TASK 4	TASK 5	TASK 6
RANDOM	74.30+/-22.64	77.30+/-23.10	75.30+/-20.60	75.30+/-19.20	85.00+/-0.00	81.30+/-11.10
GP BO	7.00+/-1.61	9.70+/-1.68	6.60+/-0.80	10.60+/-6.37	7.30+/-1.00	14.20+/-12.41
OURS	**6.70+/-0.64**	**4.30+/-3.07**	**2.10+/-0.70**	**1.90+/-0.94**	**1.50+/-0.50**	**4.90+/-2.81**
	TASK 7	TASK 8	TASK 9	TASK 10	TASK 11	TASK 12
RANDOM	58.00+/-34.33	71.30+/-23.51	70.50+/-27.69	85.00+/-0.00	85.00+/-0.00	80.10+/-14.70
GP BO	16.40+/-8.44	11.60+/-3.44	6.50+/-0.81	41.00+/-24.09	7.50+/-0.81	30.80+/-28.75
OURS	**6.30+/-3.35**	**7.00+/-6.57**	**2.50+/-1.69**	**6.90+/-3.11**	**1.00+/-0.00**	**3.40+/-1.43**

Figure 2. Localization steps (mean +/- std) for (A) Sinus and (B) Infarct. Results shows our method (bold with green) reaches target site with minimum number of steps.

stored in the memory M. The memory size is fixed at $m = 10$, and the model is trained for 100 epochs. We conducted experiments across 12 tasks within the same heart condition (i.e., sinus or infarcted), where each experiment was repeated across 10 random seeds. For the *Random* baseline, we limit the number of acquisition steps to 85.

4. Results and Discussion

Figure 2 report the number of acquisition steps (candidate sites) required to pace-map given a target ECG in different tasks, expressed as mean steps $\pm$ std. The *Random* baseline generally requires a large number of steps and exhibit high variability, indicating substantial fluctuation in the steps needed to identify VT source. The *GP BO* baseline performs better, requiring fewer mean steps and a lower standard deviation than *Random*.

In contrast, our method *OURS* shows a significantly lower number of acquisition steps, both in terms of mean and standard deviation, compared to both baselines. This demonstrates the effectiveness of our appraoch in modeling a task-independent surrogate model in combination with meta-learning for knowledge transfer across tasks. Statistical analysis showed 95% confidence that our method identifies VT source in significantly fewer steps, with an improvement of over 94% over *Random* and 70% over *GP BO*. Unlike the baselines that must retrain surrogate models from scratch, *OURS* leveraged knowledge transferred across tasks, eliminating the need for initial labeled samples. Clinically, this is crucial since patients arrive sequentially, and reusing surrogates can greatly reduce intervention time and effort.

5. Conclusion

In conclusion, we present a meta-learning–based active surrogate modeling framework that effectively transfers knowledge across pace-mapping tasks. By combining continual and active learning, it reduces the number of pacing sites needed for accurate VT localization and avoids retraining from scratch, enhancing both efficiency and clinical applicability for faster, more reliable VT source identification.

Acknowledgments

This study was supported by funding provided by NIH Award No: R01HL145590 and NSF Award No: OAC-2212548.

References

[1] Di Cesare M, Perel P, Taylor S, Kabudula C, Bixby H, Gaziano TA, McGhie DV, Mwangi J, Pervan B, Narula J, et al. The heart of the world. Global heart 2024;19(1).

[2] Koplan BA, Stevenson WG. Ventricular tachycardia and sudden cardiac death. In Mayo clinic proceedings, volume 84. Elsevier, 2009; 289–297.

[3] Samuel M, Elsokkari I, Sapp JL. Ventricular tachycardia burden and mortality: association or causality? Canadian Journal of Cardiology 2022;38(4):454–464.

[4] Josephson ME, Callans DJ. Using the twelve-lead electrocardiogram to localize the site of origin of ventricular tachycardia. Heart Rhythm 2005;2(4):443–446.

[5] Sapp JL, Bar-Tal M, Howes AJ, Toma JE, El-Damaty A, Warren JW, MacInnis PJ, Zhou S, Horáček BM. Real-time localization of ventricular tachycardia origin from the 12-lead electrocardiogram. JACC Clinical Electrophysiology 2017;3(7):687–699.

[6] Zhou S, AbdelWahab A, Sapp JL, Warren JW, Horáček BM. Localization of ventricular activation origin from the 12-lead ecg: a comparison of linear regression with non-linear methods of machine learning. Annals of biomedical engineering 2019;47:403–412.

[7] Missel R, Gyawali PK, Murkute JV, Li Z, Zhou S, AbdelWahab A, Davis J, Warren J, Sapp JL, Wang L. A hybrid machine learning approach to localizing the origin of ventricular tachycardia using 12-lead electrocardiograms. Computers in biology and medicine 2020;126:104013.

[8] Meisenzahl C, Gillette K, Prassl AJ, Plank G, Sapp JL, Wang L. Boatmap: Bayesian optimization active targeting for monomorphic arrhythmia pace-mapping. Computers in Biology and Medicine 2024;182:109201.

[9] Aras K, Good W, Tate J, Burton B, Brooks D, Coll-Font J, Doessel O, Schulze W, Potyagaylo D, Wang L, et al. Experimental data and geometric analysis repository—edgar. Journal of electrocardiology 2015;48(6):975–981.

Address for correspondence:

Pradeep Bajracharya
1 Lomb Memorial Drive, Rochester, NY, USA
pb8294@rit.edu