Myocardial Ischemia Detection Using Body Surface Potential Mappings and Machine Learning

James N Brundage[1,2], Vai Suliafu[2], Jake A Bergquist[2,3,4], Brian Zenger, Lindsay C Rupp[2,3,4], Bao Wang[2,5], Rob MacLeod[2,3,4]

[1] School of Medicine, University of Utah, SLC, UT, USA
[2] Scientific Computing and Imaging Institute, University of Utah, SLC, UT, USA
[3] Nora Eccles Cardiovascular Research and Training Institute, University of Utah, SLC, UT, USA
[4] Department of Biomedical Engineering, University of Utah, SLC, UT, USA
[5] Department of Mathematics, University of Utah, SLC, UT, USA

Abstract

Recent improvements in detecting acute myocardial ischemia via noninvasive body surface recordings have been driven by modern machine learning. While extensive research has been done using single and 12 lead ECGs, almost no models have incorporated body surface potential mappings. We created two contrasting machine learning models, logistic regression and XGBoost Classifier, and trained them on experimentally acquired body surface mappings with ground truth ischemia measurements recorded from within the heart. These models achieved a mean accuracy of 96.46% and 97.63%, as well as a mean AUC of 0.9927 and 0.9972 for the Logistic Regression and XGBoost classifiers, respectively. The anatomical location and relative contribution of each electrode were visualized and ranked. Then, new models were trained using data from only the top 12, 8, and 3 electrodes. These models trained on only a subset of the electrodes still exhibited relatively high accuracy and AUC, although at much faster training times.

1. Introduction

Myocardial ischemia is caused by an imbalance between cardiac perfusion and metabolic demand, which causes changes to cardiac cellular and tissue physiology. These changes, if untreated, can lead to both short- and long-term consequences, including sudden cardiac death [1]. Serious adverse events and even death can occur within a matter of minutes; thus, early detection of acute myocardial ischemia plays a key role in the management of at-risk patients.

The 12-lead electrocardiogram (ECG) is one of the most common tools used for diagnosis of myocardial ischemia, although it suffers from poor sensitivity and specificity between $50 - 70\%$ and $70 - 90\%$, respectively [2, 3]. Body surface potential mapping (BSPM), which captures signals from hundreds of electrodes across the entire torso surface, has been shown to be more sensitive to detecting myocardial ischemia[4, 5]. However, adoption of BSPM as a screening tool has been hindered by the difficulty of interpreting data from hundreds of electrodes rather than just 12. One way to remedy this issue is to develop computational tools that simplify analysis.

Machine learning (ML) is a computational approach that can be used to simplify the analysis of electrocardiographic signals. These tools have been well explored with 12-lead ECG analysis, where they have been used to predict a variety of pathologies and estimate relevant lab values [6, 7]. Very little has been published exploring how ML approaches can be used with BSPMs. However, recently published studies show that ML is a promising tool in the context of BSPM [8, 9] Many of the 12-lead ECG ML publications rely on "deep" learning models, which require millions of parameters [7, 10]. In contrast, "shallow" machine learning models require fewer parameters, are computationally cheaper, and can be less prone to overfitting. Furthermore, shallow machine learning models often provide interpretable predictions. We hypothesized that because BSPM contain more information than ECGs, and have greater redundancy, shallow learning tools may perform well for BSPM analysis. Success of shallow learning tools in this domain would mean computationally simpler models.

Here, shallow learning techniques are used to detect acute ischemia from BSPMs recorded from a large animal model of acute myocardial ischemia. In this study, we tested two commonly used shallow learning architectures i.e., logistic regression (LR) and XGBoost (XGB).

The success of shallow learning BSPM interpretation tools could ease the burden of BSPM interpretation at a reasonable computational cost, allowing clinicians to more easily detect, treat, and impede the progression of acute myocardial ischemia.

2.   Methods

2.1.   Machine Learning

Two "shallow" learning models were trained and evaluated: logistic regression and an ensembled forest model called XGBoost. Input data were BSPMs labeled as ischemic or nonischemic. Data acquisition and the labeling process are described in the next section. In preparation for training, body surface recordings from multiple sensors were flattened into a single vector containing concatenated signals for a single heartbeat from each sensor.

The data were split into train/test splits of size $30,000/4,200$, respectively, before training a logistic regressor implemented in Pytorch and an XGBoost Classifier from the XGBoost package. The logistic regressor consisted of one fully connected layer of size $N*D$ where N is the number of leads (in this instance 96), and D is the number of time points (1000) connecting to the sigmoid function. The logistic regressor model was trained in minibatches of size 600 using the binary cross-entropy loss and the Adam Optimizer. The XGBoost model was trained without changing the default settings. Experiments were performed on a 16 AMD Opteron 8360 SE 2.5GHz cores (HT disabled) CPU with 96GB of RAM running OpenSUSE Leap 15.2 from the Scientific Computing and Imaging Institute (SCI).

2.2.   Experimental Data Collection

The dataset used for training consisted of signals simultaneously recorded from within the myocardial wall and on the torso surface in an anaesthetised large animal. Methods were used that allow for the accurate control of ischaemic stress in the heart. The procedure is described in detail by Zenger *et al.* [11]. In brief, recordings were collected from 20-30 needle electrodes in the heart wall and 96 body surface electrodes simultaneously at a sampling rate of 1kHz. Pacing of the heart and blood supply to the heart were controlled, allowing for manipulation of the ischemic condition of the heart. The tight control over ischemic stress made it possible to create a dataset that contains a broad variety of ischaemic conditions ideal for training ML models. A recording was labelled as ischemic if the needle recordings showed ST40 potentials at or above 3 mV on at least 5 electrodes. Data from 4 animals were included in the dataset.

2.3.   Reporting Metrics

To evaluate the performance of the ML models at the task of ischemia detection, we use accuracy and receiver operating characteristic area under the curve (ROC-AUC). Each model outputs a number between 0 and 1. The receiver operating characteristic (ROC) plots the false positive rate against the true positive rate. The area under the ROC curve is a metric commonly used to evaluate how well a model can discriminate two outcomes in a binary classification task. We also evaluate the training time of each model, which is the duration of the training process in seconds.

2.4.   Weights Analysis

To determine which leads contributed significantly to model performance, the model weights were saved from each training process. Each model had 96,000 weights, each one correlating to an electrode position and time point. Individual electrode importance was determined by comparing the sum of all 1000 weights associated with each electrode. Weights were also visualized on torso models using SCIRun [12] in order to identify possible anatomical areas of significance. The most highly ranked electrodes were then used to train new LR and XGB models. Along with determining important electrodes, we also investigated which time points along the trace contributed most to model decisions. First, the root mean squared trace was calculated for representative ischemic and nonischemic beats. Because each recording has 1000 time points, 96 weights are associated with each time instance. The sum of those weights was taken, and the top 100 contributing time points were plotted on top of representative ischemic and nonischemic traces to determine which portions of the signal contributed most to model decision-making.

2.5.   Model Reduction

In order to further decrease computation time and model complexity, we trained new models using only the most highly weighted subsets of leads for each model. New LR and XGB models were trained that utilized only the 3, 8 and 12 most important electrodes as described in the weights analysis section. For example, the models trained on the Top 12 electrodes were trained using only data from the top 12 most highly weighted electrodes for each model. The ML techniques and metrics are the same as previously described.

3.   Results

3.1.   Full Dataset Training

First, the logistic regression and XGBoost models were trained on a dataset containing data from all 96 recording electrodes. The logistic regression model achieved an accuracy of 96.46% and a ROC-AUC of 0.9927, whereas the XGBoost model achieved an accuracy of 97.63% and a ROC-AUC of 0.9972. The ROC curves of both models are also seen in Figure 1. The models were fit with a train time of 24.9 and 13.9 minutes for logistic regression and XGBoost, respectively. The weights of each model were extracted, ranked by the sum of weights for each electrode, and visualized on a representative torso anatomy to investigate regions of common anatomical significance (Figure 2). The LR relied on electrodes from the top of the chest, whereas the XGB model relied on a few, scattered electrodes. Model weights were also assessed in the time domain by plotting the top 100 relevant time points along a root mean square trace of representative beats (Figure 2). LR relies primarily on information from the T-wave, although important weights are also located at the onset of the recording. The XGB model utilizes more information from the QRS and the T-wave.
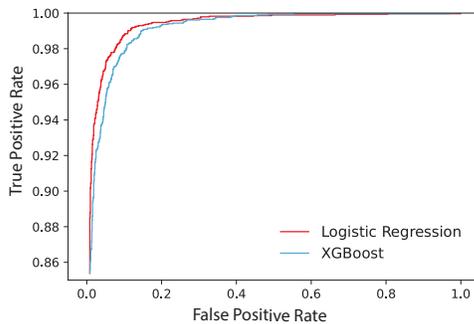

Figure 1. Receiver operating characteristic (ROC) curve for the logistic regression and XGBoost models trained on data from all 96 electrodes.

3.2.   Limited Electrode Modeling

The results of training on the reduced number of electrodes can be viewed in Table 1. The ROC-AUC decreased as the number of electrodes decreased for both LR and XGB models, although this value remains greater than 0.90 even when using only 3 electrodes. The training time of the limited top 12 electrode models was much smaller than the training time for the full model for both architectures. Following that reduction, the training time did not change much for logistic regression but continued to drop off for XGB.
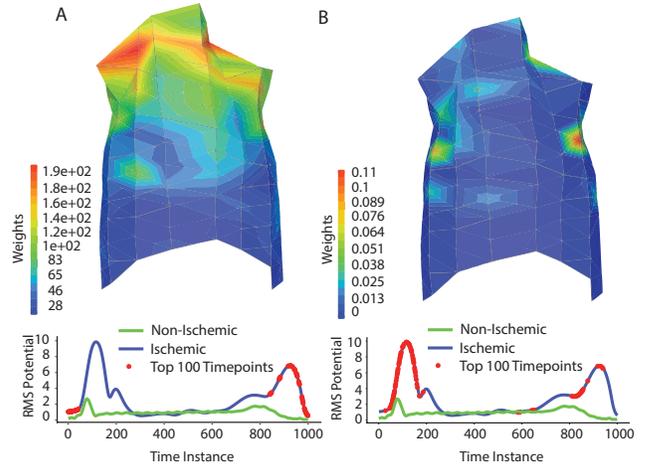

Figure 2. Weight visualization: The weights of the LR (A) and XGB (B) are visualized on a representative torso and representative root mean square (RMS) traces. The red dots along the RMS traces represent the top 100 most highly weighted time points.

Model	ROC-AUC	Fit Time
LR Full	0.9972	24.9
LR Top 12	0.9727	11.4
LR Top 8	0.9590	10.98
LR Top 3	0.9475	10.86
XGBoost Full	0.9927	13.9
XGBoost Top 12	0.9484	3.43
XGBoost Top 8	0.9448	1.93
XGBoost Top 3	0.9321	1.85

Table 1. Training outcomes for models trained on all electrodes (Full), or the most highly weighted 12, 8 and 3 electrodes.

4.   Discussion and Conclusions

This project aimed to assess the efficacy of shallow learning tools at detecting acute myocardial ischemia from BSPM signals. The results demonstrate that these models exhibit excellent performance detecting acute myocardial ischemia from BSPMs (Figure 1). The different models tested utilized slightly different anatomical and temporal features for decision-making, although both models utilize upper lateral leads and T-wave-associated time points (Figure 2). Models trained on subsets of electrodes had decreased but still relatively good performance (Table 1).

In training an ML model at this task, we hope that models will use signals from known relevant anatomic and temporal features. In this instance, the different model architectures rely on slightly different features. Although both models clearly rely on upper lateral electrodes as well as T-wave features, the LR model relies on more diffuse electrode placement than the XGB model, and the XGB model relies on more time points from around the QRS peak. The drop in ROC-AUC from the full model to the top 12 model

was larger for the LR than for the XGB model, which may indicate that intermediately weighted electrodes and time points may be more valuable to the XGB than the LR model. Further research characterising which features are most easily detected by specific models, and how that may affect model performance under differing ischemic conditions, merits further evaluation.

This study is limited by a number of important factors. The dataset is comprised of large animal recordings, and it is important to consider how well these approaches would function with human data. Some evidence indicates that models trained on simulated BSPMs and fine tuned on a small human dataset perform well [8]. Similar transfer learning approaches may help bridge the gap between models trained on animal data and human outcomes. The dataset was slightly unbalanced, with a ratio of 2.08 ischemic to nonischemic data points, which may have biased the classifier toward ischemic rather than nonischemic classifications. Both the training and testing data came from the same experimental paradigm, and we do not know whether these models will perform well for other datasets. Assessment of generalizability will be essential to determine how the models tested will translate to datasets with differently acquired signals.

These results suggest that the ML tools tested here are good candidates for clinical implementation in the detection of acute myocardial ischemia. BSPM has already been shown to have higher sensitivity and specificity than 12-lead ECGs, although interpretation of hundreds of leads is much less practical in the clinical setting. The ML tools tested here simplify interpretation to a probability score that could aid clinical decision-making, making BSPM more feasible for common use. Although this paper demonstrates the specific utility of detecting acute ischemia, shallow ML tools may perform well for other clinically relevant questions. Future studies should examine other possible implementations of these approaches.

Acknowledgments

References

[1] McCarthy B, Beshansky J, D'Agostino R, Selker H. Missed diagnoses of acute myocardial infarction in the emergency department: results from a multicenter study. Annals of Emergency Medicine 1993;22:579–582. ISSN 0196-0644.

[2] Stern S. State of the art in stress testing and ischaemia monitoring. Card Electrophysiol Rev September 2002;6(3):204–208.

[3] Knuuti J, Ballo H, Juarez-Orozco LE, Saraste A, Kolh P, Rutjes AWS, Jüni P, Windecker S, Bax JJ, Wijns W. The performance of non-invasive tests to rule-in and rule-out significant coronary artery stenosis in patients with stable angina: A meta-analysis focused on post-test disease probability. Europ Heart J September 2018;39(35):3322–3330. ISSN 15229645.

[4] Hoekstra JW, O'Neill BJ, Pride YB, Lefebvre C, Diercks DB, Peacock WF, Fermann GJ, Gibson CM, Pinto D, Giglio J, Chandra A, Cairns CB, Konstam MA, Massaro J, Krucoff M. Acute detection of st-elevation myocardial infarction missed on standard 12-lead ecg with a novel 80-lead real-time digital body surface map: primary results from the multicenter occult mi trial. Annals of Emergency Medicine 2009;54. ISSN 1097-6760.

[5] Ornato JP, Menown IBA, Peberdy MA, Kontos MC, Riddell JW, Higgins GL, Maynard SJ, Adgey J. Body surface mapping vs 12-lead electrocardiography to detect st-elevation myocardial infarction. The American Journal of Emergency Medicine 9 2009;27:779–784. ISSN 1532-8171.

[6] Trayanova NA, Popescu DM, Shade JK. Machine Learning in Arrhythmia and Electrophysiology, 2021.

[7] Rim B, Sung NJ, Min S, Hong M. Deep learning in physiological signal data: A survey. Sensors Basel Switzerland 2 2020;20.

[8] Giffard-Roisin S, Delingette H, Jackson T, Webb J, Fovargue L, Lee J, Rinaldi CA, Razavi R, Ayache N, Sermesant M. Transfer learning from simulations on a reference anatomy for ECGI in personalized cardiac resynchronization therapy. IEEE Transactions on Biomedical Engineering 2018;66(2):343–353.

[9] Zenger B, Good WW, Bergquist JA, Tate JD, Sharma V, Macleod RS. Electrocardiographic comparison of dobutamine and BRUCE cardiac stress testing with high resolution mapping in experimental models. IEEE Computers in Cardiology September 2018;45:1–4.

[10] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015 5217553 5 2015;521:436–444. ISSN 1476-4687.

[11] Zenger B, Good W, Bergquist J, Burton B, Tate J, Berkenbile L, Sharma V, MacLeod R. Novel experimental model for studying the spatiotemporal electrical signature of acute myocardial ischemia: a translational platform. J Physiol Meas Feb 2020;41(1):015002.

[12] Institute S, 2015. SCIRun: A Scientific Computing Problem Solving Environment, Scientific Computing and Imaging Institute (SCI), Download from: http://www.scirun.org.

Address for correspondence:

James Brundage
University of Utah
72 Central Campus Dr, Salt Lake City, UT 84112
james.brundage@hsc.utah.edu