34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Detection of Acute Myocardial Infarction From Serial ECG Using Multilayer Support Vector Machine

Akshay Dhawan,[⌋,‡] *Brian Wenzel, Ph.D.,*[‡] *Samuel George, M.D.,*[‡] *Ihor Gussak, M.D.,*[‡] *Bosko Bojovic, Ph.D.,*[‡] *Dorin Panescu, Ph.D.*[‡]

[⌋] School of Electrical and Computer Engineering, Cornell University, Ithaca, NY [‡]NewCardio, Inc., Santa Clara, CA

***Abstract* - Acute Myocardial Infarction (AMI) remains a leading cause of mortality in the United States. Finding accurate and cost effective solutions for AMI diagnosis in Emergency Departments (ED) is vital. Consecutive, or serial, ECGs, taken minutes apart, have the potential to improve detection of AMI in patients presented to ED with symptoms of chest pain. By transforming the ECG into 3 dimensions (3D), computing 3D ECG markers, and processing marker variations, as extracted from serial ECG, more information can be gleaned about cardiac electrical activity. We aimed at improving AMI diagnostic accuracy relative to that of expert cardiologists. We utilized support vector machines in a multilayer network, optimized via a genetic algorithm search. We report a mean sensitivity of 86.82%±4.23% and specificity of 91.05%±2.10% on randomized subsets from a master set of 201 patients. Serial ECG processing using the proposed algorithm shows promise in improving AMI diagnosis in Emergency Department settings.**

I. INTRODUCTION

Acute Myocardial Infarction (AMI) remains a leading cause of mortality in the United States and worldwide. As such, finding accurate and cost effective solutions for AMI diagnosis at Emergency Departments (ED) is vital [1].While a valuable tool, 12-lead electrocardiograms (ECG) have high interpretative variability and relatively low diagnostic accuracy. On average, cardiologists exhibit 51% sensitivity and 91% specificity in AMI detection based on first collected ECG of ED patients complaining of chest pain [2]. Computer-aided analysis and smart, adaptive learning based approaches can improve diagnostic accuracy, which is important for improving health and patient care [3].

The effectiveness of the ECG can be augmented by 3-dimensional (3D) vector analysis [4]. 3D ECGs provide additional information that may improve diagnostic accuracy [4] [5]. Along with a 3D approach, the use of information from consecutive or serial 12-lead ECG has been shown to increase sensitivity in the diagnosis of AMI [2]. While the aforementioned study concentrated on ST segment instability, we hypothesize that instability in other 3D ECG features, or markers, would indicate AMI. These 3D markers include angular, temporal, planarity, and ratio-metric parameters.

To test the diagnostic capability of serial ECGs, we extracted 3D ECG markers from a set of 201 patients (pts) who had presented to an ED with symptoms of chest pain. The AMI or non-AMI clinical diagnosis, as provided by the ED final medical records, constituted our gold standard. The changes in 3D ECG markers, as extracted from serial ECGs, were processed using support vector machines (SVM), which have been shown to be powerful tools for diagnosing heart disease using features of ECG [3]. By constructing an optimal separating hyperplane using the maximum margin between data points of separate classes, the SVM provides a reliable binary classification in high dimensional feature space [3].

To optimize the training data and feature space, we utilized a genetic algorithm search, which is an evolutionary algorithm search that operates on the principles of Darwinian evolution [6]. In the present study, the classification error rate was minimized with respect to a known subset of patients.

We present a multilayer of support vector machines with features, training data, and parameters optimized with genetic algorithms aimed at improved AMI detection accuracy. Our approach shows substantial sensitivity and specificity gains compared to cardiologists' average diagnosis.

II. MATERIALS AND METHODS

A. SECG Data Acquisition and Feature Extraction

A total of 201 pts, 65.25% male, 57.2 ± 13.2 years, experienced chest pain and presented to an urban ED (113 pts) or to a catheterization laboratory (88 pts). Of these, 112 pts had a final clinical diagnosis of AMI (52 STEMI, 60 NSTEMI) and 89 pts had no AMI. STEMI stands for ST Elevated Myocardial Infarction, whereas NSTEMI stands for Non-ST Elevated Myocardial Infarction. The medical records obtained at discharge from either location were used to establish our AMI/non-AMI gold standard. Two ECGs were taken for each patient between 10 – 60 min apart, and were transformed to 3D ECGs [4]. The pair of ECGs shall be referred to as serial ECGs.

The heart vector was computed by normalizing the weight of each of the leads from the 12-lead ECG. The normalization algorithm has been described before [4]. Briefly, the ECGs were converted into X, Y, Z components of the heart vector $\vec{H}$ using the inverse Dower matrix (ID) and the 12-lead voltage $\vec{V}$:

$$\vec{H} = ID \cdot \vec{V} \quad (1)$$

Individual attenuation factors ρ_i are calculated for each of the six precordial leads to minimize the least-squares difference between the actual and the derived ECG

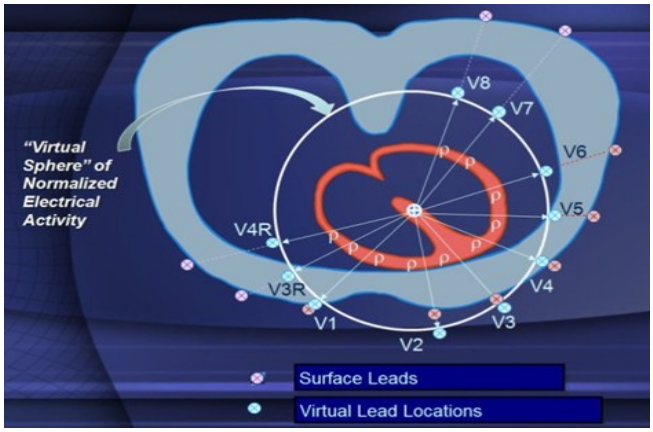


Figure 1. Calculation of a "virtual sphere" of normalized cardiac electrical activity [5].

waveforms, as calculated from the heart vector at that point. The normalized attenuation factor ρ is selected from the range of individual attenuation factors ρ_i. The time dependent voltage in any virtual leads can be written as:

$$V_d(t) = \vec{H}(t) \cdot \overrightarrow{L_\iota} \cdot \rho = \rho(L_x * X + L_y * Y + L_z * Z) \qquad (2)$$

where $\overrightarrow{L_\iota}$ is the unit lead vector.

Fig. 1 illustrates the concept of assigning equal weight to ECG leads to form a "virtual sphere" of normalized cardiac electrical activity. Based on the normalized heart vector, parameters were extracted as measured on the vector magnitude ECG and constituted our set of 3D ECG markers. Examples include QRS-T angles, planarity of QRS and T loops, directional changes in the ST vector, and ratio-metric markers, such as the relative change in the peak of the R wave with respect to the shift in the ST segment. Fig. 2 illustrates the computation of the QRS-T angle marker. In patients with a normal ECG, the QRS and T loops are expected to be coplanar and this angle yields low values. Conversely, in AMI patients, particularly in STEMIs, the QRS and T loop reside in planes that form angles that typically exceed 45°. Consequently, the QRS-T angle marker is expected to display larger values. Percent changes in 3D ECG marker values across each patient's serial ECG were also computed. Initially, a total of 227 3D ECG markers were extracted.

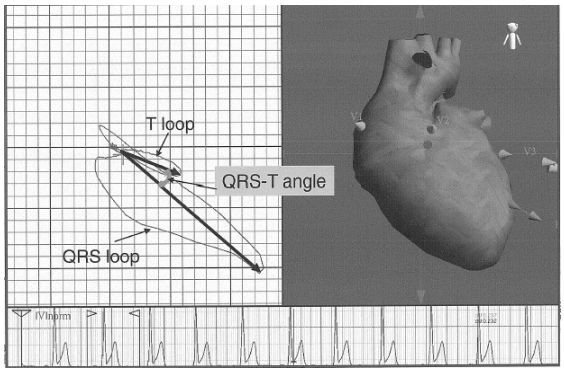


Figure 2: Calculation of QRS-T angle.

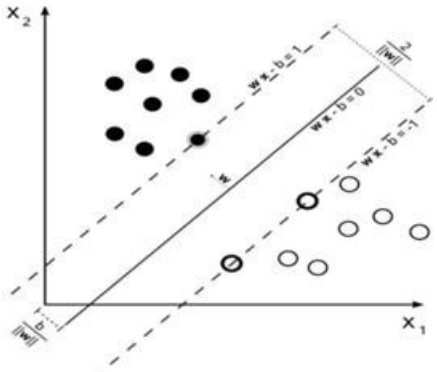


Figure 3. Basic Principle of Support Vector Machines. Two classes (black and clear dots) are separated linearly with the maximum amount of margin between them.

B. Multilayered Support Vector Machine

In the following section, let $\boldsymbol{x} \in R^n$ denote a set of features to be classified into $y = \pm 1$. Let $\{(x_i, y_i), i = 1,2, \dots, l\}$ denote a set of l training examples [7].

Review

In the case of linearly separable data, the SVM finds a linear decision function of the form:

$$f(\vec{x}) = \vec{w}^T \vec{x} + b \qquad (3)$$

that yields $f(\overrightarrow{x_\iota}) \geq 0$ for $y = 1$. and $f(\overrightarrow{x_\iota}) \leq 0$ for $y = -1$. [7]. A simple case of this classification is shown in Fig. 3, in which a linear separation maximizes the margin between two classes. For a nonlinear separation, such as our case, the decision function is modified to

$$f(x) = \vec{w}^T \varphi(\vec{x}) + b \qquad (4)$$

where $\varphi(\vec{x})$ is a nonlinear operator to map $\vec{x}$ to some higher-dimension space [7]. For the same conditions as the linear version, (4) provides a hyperplane that can separate 2 classes. The parameters of the hyperplane can be found by minimizing the following cost function:

$$\min J(\vec{w}, \xi) = \frac{1}{2}\|\vec{w}\|^2 + C \sum_{i=1}^{l} \xi_i \qquad (5)$$

subject to the following constraint:

$$y_i(\vec{w}^T \varphi(\vec{x}) + b) \geq 1 - \xi_i, \ \ \xi_i \geq 0, i = 1,2, \dots, l \qquad (6)$$

where ξ is a slack variable defining the relaxation of the separability in (4), and C is a regularization parameter [7].

Minimizing the cost function in (5) can be done using the method of Lagrange multipliers. This produces the weight function with constants a_i:

$$\vec{w} = \sum_{i=1}^{l} a_i y_i \varphi(\overrightarrow{x_\iota}) \qquad (7)$$

The vectors $\overrightarrow{x_\iota}$ are data for which the decision function is exactly ±1. These are called support vectors and represent the borderline examples from two classes. Plugging (7) into the decision function in (4) yields

$$f(x) = \sum_{i=1}^{l} a_i y_i \varphi^T(\overrightarrow{x_\iota}) \varphi(\vec{x}) + b = \sum_{i=1}^{l} a_i y_i K(\overrightarrow{x_\iota}, \vec{x}) + b \qquad (8)$$

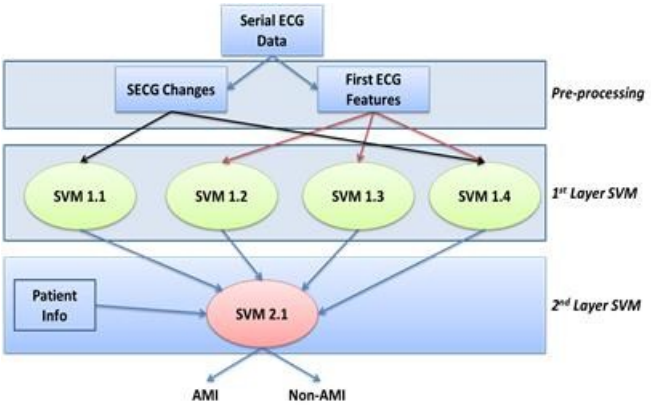


Figure 4: Block Diagram of multi-layer SVM algorithm.

where $K(\vec{x_i}, \vec{x})$ is a kernel function that transforms the problem to a high dimensional feature space. The Gaussian radial basis function (RBF) is a common kernel that takes the following form

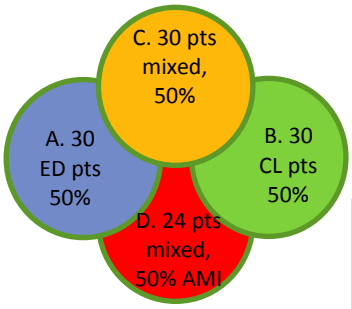
$$K(\vec{x}, \vec{y}) = \exp(-\frac{\|\vec{x}-\vec{y}\|^2}{2\sigma^2}) \quad (9)$$

The RBF kernel is maximum when $x = y$ and has a width of σ, which defines the smoothness of the decision boundary. The decision function in (8) then takes its maximum value when the input vectors $\vec{x}$ and support vectors $\vec{x_i}$ are identical.

Multilayer Network

Support vector machines were used in a multilayer network to classify each patient as AMI/non-AMI based on computed features and changes in features. Fig. 4 shows a block diagram, in which preprocessing and the 1st and 2nd layer SVM are shown.

A radial basis function (RBF) kernel was chosen for all SVM with σ = 15 and C = 1. The 1st layer SVM consisted of multiple SVM modules that simultaneously analyzed changes in 3D ECG markers from serial ECGs as well as the marker values from the patient's first ECG. Each SVM in this layer was trained on a subset of the patient data. SVM 1.1 was trained on serial ECG changes from subset A (30 ED pts, 50% AMI). SVM 1.2 was trained on 3D ECG marker values



SVM	Set
1.1	A
1.2	A
1.3	B
1.4	C
2.1	D

SVM	Marker Type
1.1	SECG
1.2	3D ECG
1.3	3D ECG
1.4	SECG + 3D ECG
2.1	Binary from Level 1

Figure 5. Description of training sets for each SVM in MLSVM. The Venn diagram shows the overlap of the 4 training sets. The tables show the training set and marker types.

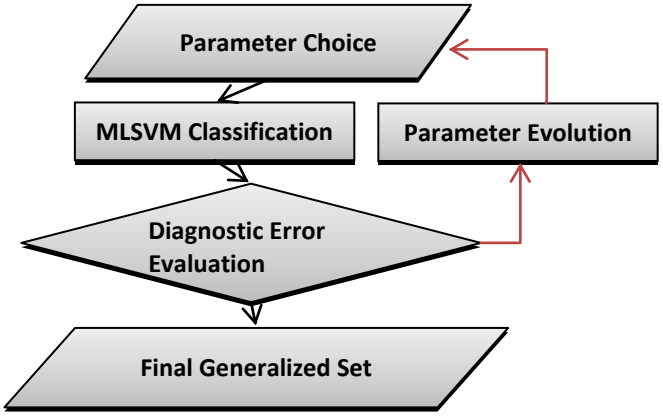


Figure 6: Block Diagram of optimization of parameters via GA. The red line shows where evolutionary selection from the GA search occurs.

from the subset A. SVM 1.3 was trained on 3D ECG marker values from subset B (30 catheterization lab pts, 50% AMI). SVM 1.4 was trained on SECG changes and 3D ECG marker values from subset C (30 combined pts, 50% NSTEMI, 50% non-NSTEMI) from all 201 pts. The training sets are summarized in Figure 5.

The binary outputs of the 1st layer became features for the 2nd layer. The 2nd layer consisted of a single SVM that integrated 1st layer outputs with higher order characterizations of the patients to give a final classification of AMI or non-AMI. SVM 2.1 was trained on subset D (24 combined pts, 50% AMI) based on the aforementioned features. In total, 70 patients were used for training due to the overlap between subsets A, B, C, and D.

C. Genetic Algorithm Optimization

Genetic algorithms are a set of evolutionary algorithms that operate on the principles of natural selection: mutation, selection, crossover, and reproduction [6]. A number of potential solutions to minimization problems are evaluated using a user defined fitness function. These solutions undergo the aforementioned principles and reproduce for new, fitter generations. The process repeats until the change in an error function ceases to exceed a specified value.

The selection of features and training data were optimized so to minimize the error rate of specificity and sensitivity for the network. Features were reduced from 227 to 60 as their fitness was determined from classification error using the generalized multi-layer SVM on all patients. A 227 length bit string was used for genetic algorithm, in which a 1 represented the inclusion of a feature, and a 0 represented exclusion. The initial string included all features, and the algorithm iteratively modified the string until classification error was minimized. Following feature reduction, training patients were chosen by using the same fitness function with an additional constraint on the number of training patients to be less than 100. These patients constituted a training set for which the SVM could be most generalized. Along with patients, parameters such as σ and C

TABLE I. DIAGNOSTIC ACCURACY

Metric	Mean +/- St. Dev			Min	Max
Sensitivity	86.82%	+/-	4.23%	75.00%	100.00%
STEMI	90.47%	+/-	5.08%	75.00%	100.00%
NSTEMI	83.18%	+/-	7.01%	60.00%	100.00%
Specificity	91.05%	+/-	2.10%	86.67%	98.33%
PPV	86.67%	+/-	2.79%	80.00%	97.30%
NPV	91.27%	+/-	2.58%	84.13%	100.00%

were optimized to ensure the best decision boundary. The optimization process is depicted in Fig. 6.

D. Testing

The proposed algorithm was tested on all 201 pts, all non-train pts (131 pts), and 1000 random subsets of all 201 pts consisting of the following: 20 STEMI, 20 NSTEMI, and 60 Non-AMI pts. Additionally, blind testing was performed on a set of 12 pseudo-ischemia pts. These pts had been previously diagnosed with Benign Early Repolarization, a condition that displays ST segment elevation but no AMI.

III. RESULTS

Table I presents the mean, min, and max values for metrics computed on the randomized subsets as described in section II.D. Additionally, on all 201 pts, the proposed algorithm attained a sensitivity of 86.61%, a specificity of 91.01%, a positive predictive value (PPV) of 92.38%, and a negative predictive value (NPV) of 84.38%.On the 131 non-train pts, it attained a sensitivity of 85.71%, a specificity of 88.33%, a PPV of 89.55%, and a NPV of 84.13%.

On this data set, 2 expert cardiologists averaged 55.29% sensitivity and 83.72% specificity in AMI detection based on first collected ECG. The cardiologists interpreted the ECGs according to their expert medical training. The mean performance of the proposed algorithm improved sensitivity by 31.53% and specificity by 7.33% [4]. Based on a McNemar's test, the sensitivity improvement was statistically significant ($p < 0.05$) [8]. Assuming 8% AMI prevalence in the general population presenting to Emergency Departments with symptoms of chest pain, our diagnostic numbers on the non-train pts would produce a PPV of 38.97% and NPV of 98.61% compared to 22.80% and 95.56% for expert cardiologists, respectively [9]. Our PPV improves that of cardiologists by 16.17%. Finally, 11 out of 12 pseudo-ischemia pts were correctly classified as non AMI, for a specificity of 91.67%.

IV. CONCLUSIONS

The proposed algorithm performed strongly, as exhibited by the highly improved sensitivity and PPV as compared to cardiologists' average. The stellar performance on various metrics demonstrates two points: the viability of using serial ECGs as classification features and the algorithm robustness as a diagnostic tool for AMI detection. The high performance on the blinded pseudo-ischemia set indicates that the algorithm is not fooled by ST segment instability in non-AMI patients.

The combination of genetic algorithm optimized multi-layer SVM and serial ECG analysis shows promise in improving diagnostic accuracy of AMI/non-AMI patients in Emergency Departments in a cost-effective manner. This observation is important because approximately 6 million patients present to U.S. EDs each year with chest pain [10]. Approximately 5.5 million of those patients do not have Acute Myocardial Infarction, but rather some other clinical condition, such as heartburn, gall stones, etc. [10]. Of them, about 26% have AMI ruled out by a first diagnostic triage in the ED, typically comprising at least a 12-lead ECG study and blood troponin levels (a biomarker for cardiac injury). The remaining 74%, or approximately 4 million, are kept around in the hospitals for cardiac additional testing, until it is subsequently discovered that most of these patients do not suffer from AMI. Based on the 8% AMI prevalence rate [9], out of the 4 million patients kept around for further testing, 320,000 actually have AMI. It has been shown that each 30 min of therapy delay can drastically reduce their probability of surviving AMI. Our 31.53% improvement in AMI detection sensitivity may results in timely delivery of life-saving therapeutic interventions to about 100,000 of additional patients annually.

We conclude that these preliminary results show sufficient promise for our proposed algorithm. Future randomized clinical studies will expand the training of the multilayer support vector machines and demonstrate performance in diverse test settings.

V. REFERENCES

[1] Mirvis, D.M. and A.L. Goldberger. Electrocardiography, Chapter 12 in Braunwald's *Heart Disease: A Textbook of Cardiovascular Medicine,* Eighth Edition 2008, Libby P Ed.

[2] S.M. Salerno, P.C. Alguire, and H.S. Waxman, "Competency in interpretation of 12 lead electrocardiograms: a summary and appraisal of the published evidence," *Annals of Internal Medicine* (2003) 138:751-759.

[3] A.E. Zadeh, A. Khazaee, and V. Ranaee. "Classification of the electrocardiogram signals using supervised classifiers and efficient features", Computer Methods and Programs in Biomedicine (2010) 99:179-194.

[4] B.Bojovic,L. Hadzievski, D. Panescu, I. Gussak, S. George, and B. Vajdic. "Visual 3Dx: Algorithms for quantitative 3-dimensional analysis of ECG signals," *EMBC 2009. Annual International Conference of the IEEE* (2009) 6751-6754.

[5] B.J. Wenzel, S. George, D. Lakkireddy *et al*. "Algorithm for quantitative 3 dimensional analysis of ECG signals improves myocardial diagnosis over cardiologists in diabetic patients", *EMBC 2011. Annual International Conference of the IEEE* (2011) 965-968.

[6] Y.H. Said, "On Genetic Algorithms and their Applications", Handbook of Statistics (2005) 24: 359-390.

[7] C. Cortes and V. Vapnik. "Support-Vector Networks," *Machine Learning*, vol. 20, pp. 273-297, 1995.

[8] J. Rice, John, Mathematical Statistics and Data Analysis, 2nd Ed. Belmont, California: Duxbury Press, 1995.

[9] D.R. Zucker, J.L. Griffith, J.R. Beshansky, and H.P. Selker, "Presentations of Acute Myocardial Infarction in Men and Women "Journal of General Internal Medicine(1997)12:79-87.

[10] L.F. McCaig and C.W. Burt, "National Hospital Ambulatory Medical Care Survey 2003. Emergency Department Summary: Advance Data from Vital and Health Statistics," No. 358. Hyattsville, MD: National Center for Health Statistics; 2005.