Deriving ECG from EASI Electrodes via Machine Learning

Piroon Kaewfoongrungsi [1] and **Daranee Hormdee** [2,*]

[1,2] Department of Computer Engineering, Faculty of Engineering, Khon Kaen University 40002, Thailand.
E-mail: [1] kpirun@hotmail.com, [2] darhor@gmail.com

Abstract: The 12-lead Electrocardiogram (ECG) is the standard clinical method of heart disease diagnose. Measuring all 12 leads is often impractical. In 1988, Gordon Dower has introduced an EASI-lead ECG System. In order to gain all 12-lead ECG back from this EASI-lead system, Dower's equation was proposed then. Ever since various attempts have been explored to improve the synthesis accuracy, mostly via linear regression. This paper presents how machine learning was used to find a set of transfer function for deriving the 12-lead ECG from EASI-lead system. The experiments were conducted to compare the results those of Support Vector Regression (SVR), Artificial Neural Networks (ANNs) against those of Dower's method. The results have shown that the best performance amongst those methods with the less RMSE error values for all signals with the standard 12-lead ECG was obtained by SVR, followed ANNs and Dower's equation, respectively.

Keywords-- **12-lead System, EASI Electrodes, Dower's Method, ANNs, SVR, Machine Learning, PhysioNet DataBase.**

1. Introduction

The 12-lead Electrocardiogram (ECG) is the standard clinical method in cardiology and lies at the center of the decision pathway for the evaluation and management of patients and evaluating complicated cardiac arrhythmias, for diagnosing other cardiac disorders. The standard 12-lead ECG signals are Lead I, Lead II, Lead III, Lead aVR, Lead aVL, Lead aVF, Lead V1, Lead V2, Lead V3, Lead V4, Lead V5 and Lead V6 signals. Typically for measuring 12-lead ECG requires 9 electrodes to be placed strategically on the body and one electrode to be connected to ground [1,2] as shown in Figure 1(left).

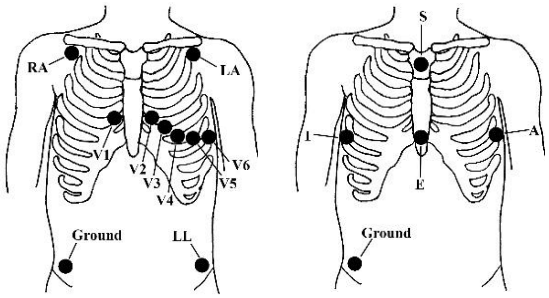


Figure 1. Standard 12-lead ECG system (left) VS EASI-lead system (right)

Measuring all 12 leads is often cumbersome and impractical especially on a long term monitoring. Reducing the number of leads from the standard 12-lead ECG yielding the smaller number of measurement electrodes and consequently fewer wires, is possible by deriving the missing signals from the actual measured electrodes. The development of ECG systems with reduced number of electrodes started in the 1940s [3], but the first notable work on derived 12-lead ECG system came in 1968 [4] with the introduction of a derived 12-lead ECG synthesized from the spatial vectorcardiography previously introduced by Frank [5].

In 1988, Dower, again, and team [6] set an example for deriving the 12-lead ECG from four completely new (EASI) electrodes, as shown in Figure 1(right). After the derived 12-lead ECG system via EASI electrodes has been presented, various improvements on coefficients in Dower's equation have been investigated ever since.

This paper attempts to present machine learning techniques as the alternative methods as opposed to the original Dower's method.

2. Related Works

Considering the derivation of 12-lead ECG from EASI-lead system, despite Dower's method, there have been a number of research on finding a better derivation approach.

In 2012, Oleksy [7] introduced the linear regression method as opposed to Dower's equation, in order to synthesize the standard ECG signals from EASI lead system using E, A, S and I signals as input data.

Recently, the nonlinear regression methodology [8] has been proposed as the synthesis approach to derive the 12-lead ECG signals from EASI leads. This yielded to less error compared to the previous Dower's and linear methods.

Two machine learning techiques explored this paper are Artificial Neural Networks (ANNs) and Support Vector Regression (SVR) with ERBF kernel function. The following subsections briefly revise the basic concepts of Dower's, SVR and ANNs.

2.1 Dower's Method

The synthesis method implemented in Dower's method used paired signals A-I, E-S and A-S derive as a weighted linear sum of these 3 base signals as in the Equation (1).

$$L_{Derived} = a(A - I) + b(E - S) + c(A - S) \quad (1)$$

Where $L_{Derived}$ represents any surface ECG lead and a, b, and c represent empirical coefficients. These coefficients,

* Corresponding Author

developed by Dower, are positive or negative values with accuracy up to 3 decimal points.

2.2 Support Vector Regression (SVR)

SVR in the past it has been used to solve nonlinear problems [9]. Basic idea of SVR is to map input data into higher dimensional space to map nonlinearity in original data as to perform linear in higher dimensional space using a kernel function and construct the separated hyper plane. The SVR function is shown in Equation (2).

$$f(X) = \langle W \cdot \mathrm{K}(X) \rangle + b \quad (2)$$

The performance of SVR is dependent on kernel function being used. In this paper used ERBF kernels for mapping function to map input data to a higher dimension as in Equation (3). The parameter ε was set to 0.001 and parameter C was set to 5,000.

$$K_{ERBF}(X, X_i) = \exp(-\|X - X_i\|/2\sigma^2) \quad (3)$$

Where σ is the bandwidth of the kernel function.

2. Artificial Neural Networks (ANNs)

Artificial Neural Network has been used for synthesis 5 signals (V1, V3, V4, V5 and V6) from 3 leads (Leads I, II and V2) of the standard 12-lead ECG signals [10].

However, in this paper, an ensemble of N multilayer feedforward ANNs trained by means of a supervised back-propagation algorithm was utilized. Each individual ANNs consists of one input layer with 4 input neurons as lead E, A, S and I, one output layer with 12 output neurons (one for each derived signal), 4 hidden layer and N =10 - 60 neurons per hidden layer. The activation function type use a linear activation function for the output neurons and chosen sigmoid transfer function for the hidden layer is shown in Equation (4).

$$f(n) = \frac{2}{1+exp^{(-2n)}} - 1 \quad (4)$$

2. Experimental Methodology

The experiments have been conducted to compare synthesis methodologies for deriving the 12-lead ECG from EASI-lead system. All dataset, used in this work, are obtained from PhysioNet database [11] consisting of 4,810 samples for each signal to shuffle data sets in order to prevent over fitting and using five-fold cross-validation, to find the best parameter.

The following steps present how to derive the transfer function;

1) The total dataset from PhysioNet has been into two parts (90:10). The first '90%' part was used to find kernel parameters for SVR and nodes for ANNs while the last '10%' part was used for blind test.

2) As five-fold cross-validation was utilized in this work, the first 90% dataset was then divided into 5 equal parts/folds. Each round a single fold is used for testing, leaving the other 4 folds for training. In the nth round, fold#n is used for testing while the remaining folds are used for training. For instance, in the 2th round, fold#2 is used for testing while folds#1 and folds#3-5 are used for training. In total 5 rounds are processed. To find the average errors in the regression of each fold, the Root Mean Squared Error (RMSE) in the Equation (10) is used.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(A_t - F_t)^2} \quad (10)$$

Where A_t is the actual value in time t, F_t is the forecast value in time t and n is sample of testing set in each fold.

3) From all 5 folds, the RMSE value of the lead I, lead II, lead III, lead aVR, lead aVL, lead aVF, lead V1, lead V2, lead V3, lead V4, lead V5 and lead V6 signals are considered. In order to find the transfer function of each signal, the fold that provides the minimum RMSE value of that signal must be identified. Then the constant, coefficients, parameter σ from that fold will be substituted into the equation of Dower's method in Equation (1), SVR in Equation (2) and ANNs in Equation (7).

4) After obtaining the transfer function models for each signal is tested with blind test data of 10% to find RMSE value.

5) Finally the big test in order to evaluate these transfer functions can then be started. By feeding the data set from those 4,810 data samples into these 12 transfer functions to get the calculated lead n signals, the RMSE values of each lead signal can be determined from the calculated signals and the ones from the PhysioNet dataset.

3. Results

The test results with 5-fold cross-validation to find RMSE value of Dower's method, SVR and ANNs for 12-lead signals are listed in Table 1-3.

Table 1. RMSE (mV) with Dower's method.

Root Mean Squared Error (*RMSE*)					
Signals	**Fold#**				
	1	***2***	***3***	***4***	***5***
Lead I	33.608	30.026	29.693	30.993	28.152
Lead II	34.885	31.966	30.140	35.927	34.266
Lead III	54.207	47.657	42.845	54.354	46.284
Lead aVR	25.672	24.062	24.508	24.917	25.700
Lead aVL	40.243	35.191	32.393	38.959	32.672
Lead aVF	44.897	40.078	36.279	46.002	40.899
Lead V1	27.421	25.007	25.286	29.801	23.904
Lead V2	41.022	37.179	37.895	44.646	41.476
Lead V3	50.933	46.322	44.833	52.422	43.699
Lead V4	53.287	50.880	56.162	64.026	55.620
Lead V5	31.169	30.070	29.124	34.890	31.224
Lead V6	23.477	19.720	17.422	19.670	18.782

Table 2. RMSE (mV) with SVR using ERBF kernel.

Root Mean Squared Error (*RMSE*)					
Signals	**Fold#**				
	1	***2***	***3***	***4***	***5***
Lead I	3.869	3.416	4.518	3.485	3.271
Lead II	6.361	6.989	8.051	4.460	4.281
Lead III	8.842	7.405	7.794	6.065	5.994
Lead aVR	5.883	6.457	7.304	4.131	4.002
Lead aVL	5.735	3.996	3.911	4.268	4.051
Lead aVF	7.931	7.599	8.407	5.377	5.270
Lead V1	2.280	2.398	4.677	2.417	2.726
Lead V2	4.992	5.359	7.330	5.678	5.522
Lead V3	6.521	7.176	5.286	7.180	5.108
Lead V4	10.416	10.031	9.176	9.492	9.981
Lead V5	4.512	3.582	6.608	4.234	4.724
Lead V6	2.708	4.781	5.593	2.801	1.940

Table 3. RMSE (mV) with ANNs method.

Root Mean Squared Error (*RMSE)*					
Signals	**Fold#**				
	1	***2***	***3***	***4***	***5***
Lead I	15.531	9.701	8.750	11.049	12.214
Lead II	15.057	11.920	10.991	13.882	16.181
Lead III	21.757	15.570	18.367	18.204	12.575
Lead aVR	10.120	10.029	10.836	9.781	9.596
Lead aVL	14.087	10.589	10.880	14.257	11.826
Lead aVF	14.302	16.003	15.173	14.232	13.011
Lead V1	5.862	5.144	5.667	5.705	4.872
Lead V2	12.278	12.235	15.370	12.922	12.814
Lead V3	10.987	14.442	14.943	16.975	14.379
Lead V4	25.985	21.524	25.205	23.915	28.019
Lead V5	9.917	7.056	9.527	9.061	7.904
Lead V6	3.964	3.974	4.515	5.809	5.198

The highlighted showed the minimum RMSE values amongst 5 folds for each of 12 leads. The parameter σ and the number of hidden layers from those folds with the minimum of RMSE value was then used for deriving 12 signal ECG. Then, using the transfer function models for each signal is tested with blind test data of 10% and tested with 4,810 data samples to find RMSE value from all three methods are shown and compared in Table 4-5.

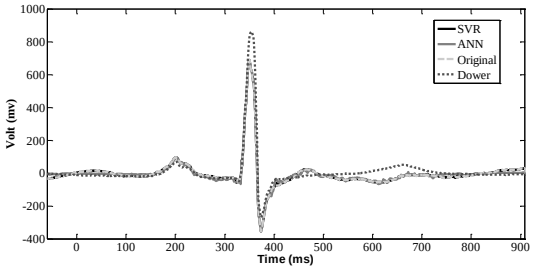
Table 4. RMSE (mV) tested with blind test data amount 10%.

Signals	**Dower**	**SVR**	**ANNs**
Lead I	35.288	3.635	10.628
Lead II	32.476	6.584	11.930
Lead III	54.648	7.767	13.126
Lead aVR	24.088	4.655	12.103
Lead aVL	41.950	4.051	13.446
Lead aVF	43.574	7.508	15.329
Lead V1	27.438	2.232	5.763
Lead V2	40.801	6.408	14.672
Lead V3	49.371	6.53	16.942
Lead V4	54.262	12.008	25.997
Lead V5	35.083	8.591	11.025
Lead V6	23.152	3.733	5.003
Average	**38.511**	**6.142**	**12.997**

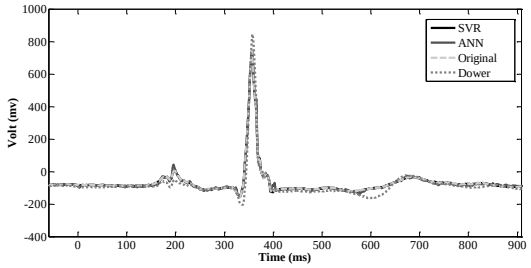
Table 5. RMSE (mV) tested 4,810 data samples.

Signals	**Dower**	**SVR**	**ANNs**
Lead I	30.570	1.868	10.321
Lead II	33.241	2.856	9.252
Lead III	49.559	3.639	12.468
Lead aVR	24.603	2.165	9.379
Lead aVL	36.357	2.185	11.616
Lead aVF	41.781	3.362	13.483
Lead V1	26.093	1.276	4.757
Lead V2	40.280	3.082	11.617
Lead V3	47.440	3.568	13.844
Lead V4	55.584	5.600	19.220
Lead V5	31.483	2.164	7.496
Average	**36.370**	**2.775**	**10.693**

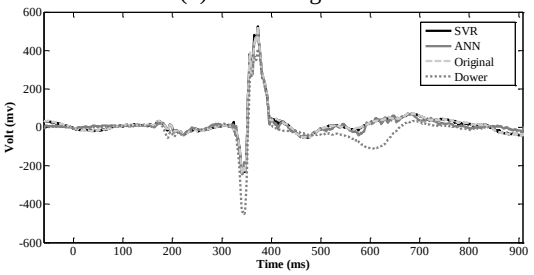
Lastly, plots of 12 signals measured using standard 12-lead ECG method, derived using EASI-lead system by Dower's method, SVR using ERBF kernel function and ANNs are shown in Figure 2(a-l).



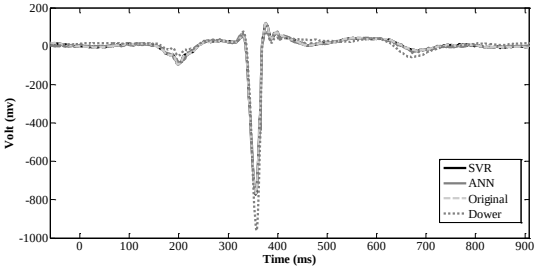
(a) Lead I Signal.



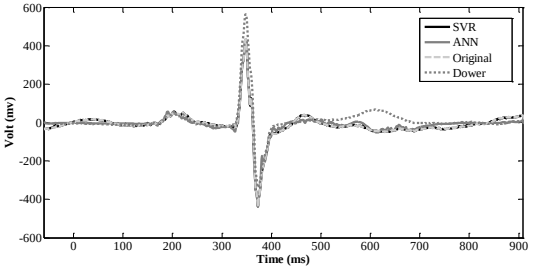
(b) Lead II Signal.



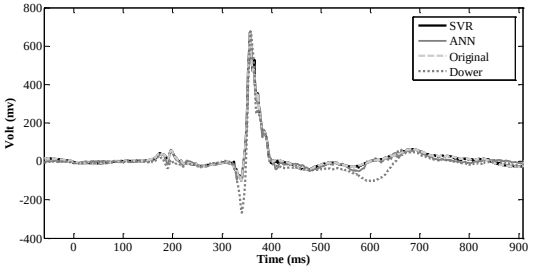
(c) Lead III Signal.



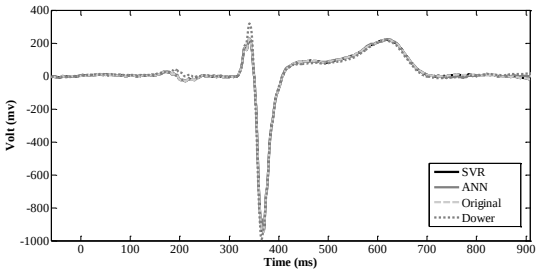
(d) Lead aVR Signal.



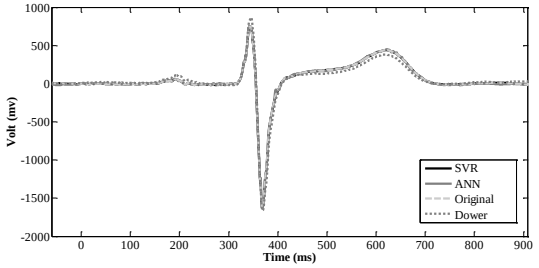
(e) Lead aVL Signal.



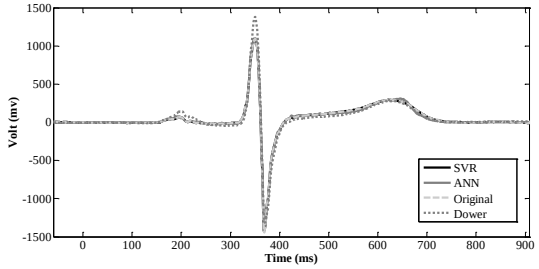
(f) Lead aVF Signal.



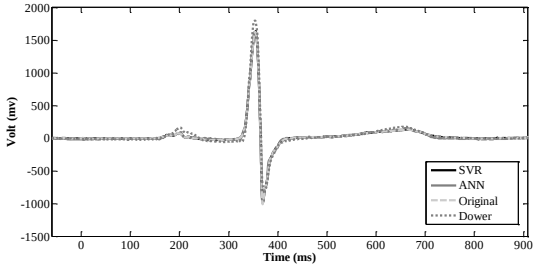
(g) Lead V1 Signal.



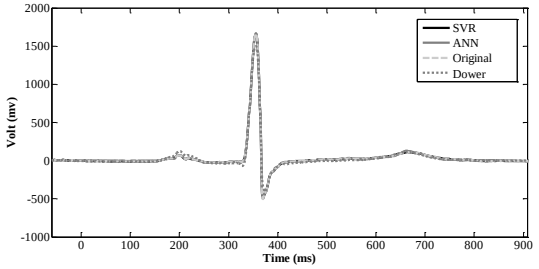
(h) Lead V2 Signal.



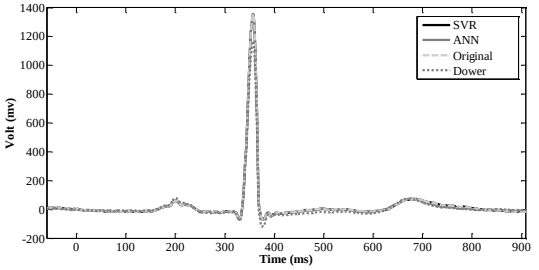
(i) Lead V3 Signal.



(j) Lead V4 Signal.



(k) Lead V5 Signal.



(l) Lead V6 Signal.

Figure 2. Derived VS original signals of 12-lead ECG.

4. Conclusions

This paper has presented machine learning with SVR and ANNs for deriving the standard 12-lead ECG from EASI-lead system. The experimental results showed that the best performance in this work, was obtained from the SVR using ERBF kernel function is less RMSE values and gave average RMSE value of all signals just over 3 mV followed by ANNs and Dower's method, respectively. Therefore, it is obvious to conclude that machine learning with SVR is worth chosen for deriving the 12-lead ECG from EASI-lead system. As for future works, other regression or machine learning techniques to improve the performance for deriving the 12-lead ECG signals from EASI-lead system should be investigated further.

References

[1] E. Joseph and D. Bronzino. "Principles of Electrocardiography," *The Biomedical Engineering Handbook: Second Edition*. 2000.

[2] Bioelectromagnetism [Web page]. http://www.bem.fi/book/.

[3] I. Tomasic and R. Trobec, "Electrocardiographic Systems With Reduced Numbers of Leads—Synthesis of the 12-lead ECG," *Biomedical Engineering, IEEE Reviews in Biomedical Engineering*, vol.7, pp.126- 142, 2014.

[4] G. E. Dower, "A Lead Synthesizer for the Frank System to Simulate the Standard 12-lead Electrocardiogram," *Journal of Electrocardiology*, vol. 1, no. 1, pp. 101–116, 1968.

[5] E. Frank, "An Accurate, Clinically Practical System for Spatial Vectorcardiography," *Circulation*, vol. 13, no. 5, pp. 737–749, May 1956.

[6] G. E. Dower, et al, "Deriving the 12-lead Electrocardiogram from four (EASI) Electrodes," *Journal of Electrocardiology*, Vol.21, Supplement, pp.S182-S187, 1988.

[7] W. Oleksy, E. Tkacz and Z. Budzianowski, "Improving EASI ECG Method Using Various Machine Learning and Regression Techniques to Obtain New EASI ECG Model," *Intl. Journal of Computer and Communication Engineering*, vol. 1, no. 3, pp. 287-289, 2012.

[8] P. Kaewfoongrungsi and D. Hormdee, "Deriving the 12-lead ECG from EASI Electrodes via Nonlinear Regression," *Intl. Journal of Advances in Electronics and Computer Science*, vol.2, pp.106-110, 2015.

[9] A. Yodjaiphet, N. Theera-Umpon and S. Auephan wiriyakul, "Electrocardiogram Reconstruction using Support Vector Regression," *Signal Processing and Information Technology (ISSPIT)*, IEEE International Symposium, pp.000269-000273, 2012.

[10] H. Atoui, et al. "A Novel Neural-Network Model for Deriving Standard 12-Lead ECGs from Serial Three-lead ECGs: Application to Self-Care," *IEEE Transactions on Information Technology in Biomedicine*, vol.14, No.3, pp.883-890, 2010.

[11] PhysioNet Database. [Web page] http://www.Physionet.org/challenge/2007/data/