KKU ENGINEERING JOURNAL 2016;43(S2):317-322 Research Article

KKU Engineering Journal

https://www.tci-thaijo.org/index.php/kkuenj/index

Deriving the 12-lead ECG from an EASI-lead system via support vector regression

Piroon Kaewfoongrungsi and Daranee Hormdee*

Embedded system research and development group, Department of computer engineering, Faculty of engineering, Khon kaen university, Khon Kaen 40002, Thailand.

Received April 2016
Accepted June 2016

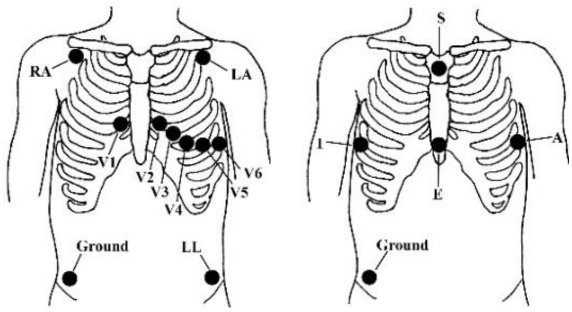
Abstract

The measuring of all 12-lead electrocardiogram (ECG) is often cumbersome and impractical especially on a long term monitoring. In 1988, Gordon Dower has introduced an EASI-lead system, where only 5 electrodes are used. 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 Support Vector Regression (SVR) is 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 SVR against those of Linear regression and those of Dower's method. The experimental results have shown that the best performance amongst those methods with the minimum of RMSE for all signals with the standard 12-lead ECG was obtained by SVR, followed by Linear regression and Dower's equation, respectively.

Keywords: ECG, 12-lead system, EASI-lead system, Linear regression, Dower's method, Support vector regression, PhysioNet database

1. Introduction

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 the standard 12-lead ECG for diagnosis of physician requires 9 electrodes to be placed strategically on the body and one electrode to be connected to ground as shown in Figure 1(a) [1-2].



(a) Standard 12-lead ECG system (b) EASI-lead system

Figure 1 (a) Standard 12-lead ECG system and (b) EASI-lead system

The development of ECG systems with reduced number of electrodes for increases mobility of patients and reduces cost of a device are started in the 1940 [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].

Previously, in 1968, Dower presented a case for the first category. In 1988, Dower, again, and team [6] set an example for the latter category, by deriving the 12-lead ECG from four completely new (EASI) electrodes, as shown in Figure 1(b). 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. In 2012, Oleksy [7] proposed 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. This yielded to less error compared to the previous Dower's method.

Up till recently, the previous works mostly focused on Linear regression as the synthesis approach to derive the 12-lead ECG signals from EASI-lead system. This paper attempts to present nonlinear regression with support vector regression (SVR) as the alternative method as opposed to Dower's or Linear regression.

2. Literature reviews

2.1 Dower's method

The synthesis method implemented in Dower's method [7] used paired signals A-I (primarily X, or horizontal vector

*Corresponding author. Tel.: +6683 360 3232
Email address: darhor@kku.ac.th; darhor@gmail.com
doi: 10.14456/kkuenj.2016.119

component), E-S (primarily Y, or vertical vector component) and A-S (containing X, Y, plus Z, the anteriorposterior component) derive as a weighted linear sum of these 3 base signals as in the Equation (1).

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

Where $\mathrm{L_{Derived}}$ represents any surface ECG lead and a, b, and c represent empirical coefficients. These coefficients, developed by Dower, are positive or negative values with accuracy up to 3 decimal points.

2.2 Linear regression

Linear regression [8] is the oldest and most widely used predictive model. The goal is to minimize the sum of the squared errors to fit a straight line to a set of data points. The Linear regression model fits a linear function for derive the 12-lead ECG signals from EASI-lead system. The function is as follow:

$$Y_n = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 \quad (2)$$

Where Y_n is the transfer function of lead n signal, n is 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, β_0 is the constant and β_1 ,..,β_4 are coefficients of X_1,... X_4 from the fold providing the minimum RMSE of lead n signal. X_1 is lead E, X_2 is lead A, X_3 is lead S, X_4 is lead I.

2.3 Support vector regression method

Support vector regression [9-11] in the past it has been used to solve nonlinear problems. Basic idea behind 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 (3).

$$f(X) = \langle W \cdot \mathrm{K}(X) \rangle + b \quad (3)$$

Where W is the weight vector, X is the input column vector, K is the kernel function for mapping data to higher dimension, b is the bias value. The dataset used to train with SVR is $\{(X_i, Y_i)\}_{i=1}^{l}$, $X \in \mathbb{R}^n$, $Y \in \mathbb{R}$ where X_i is the input data vector, Y_i is desired output vector, X is input space, Y is output space. In the function $f(X)$ has the deviation (ε) called "loss function" and all input data X_i that give value of $f(X)$ within $\pm\varepsilon$ interval are called "support vector" as shown in Figure 2.

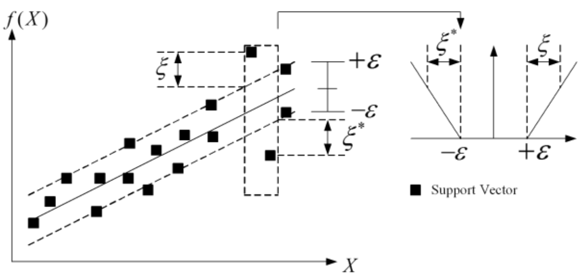


Figure 2 Soft margin $\varepsilon -$ insensitive in linear SVR

From Figure 2, the optimization was used to find weight vector (W) as in Equation (4-6).

$$\min_{W,\xi_i,\xi_i^*} \frac{1}{2}\|W\|^2 + C\sum_{i=1}^{l}(\xi_i, +\xi_i^*) \quad (4)$$

Subject to

$$\left.\begin{array}{c} Y_i - \langle W \cdot K(X, X_i)\rangle - b \le \varepsilon + \xi_i \\ \langle W \cdot K(X, X_i)\rangle + b - Y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0 \end{array}\right\}$$

Where C is a constant variable, W is weight vector obtained by solving with optimization problem as in Equation (5).

$$W = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)K(X_i) \quad (5)$$

Substitute equation (5) into (1), the function $f(X)$ can be written as in Equation (6).

$$f(X) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)K(X, X_i) + b \quad (6)$$

Where $K(X, X_i)$ is a kernel mapping function between X and X_i.

The performance of SVR is majorly dependent on kernel function being used. In this paper used three kernels for mapping function to map input data to a higher dimension as in Equation (7-9). The parameter ε was set to 0.001 and parameter C was set to 5,000.

RBF kernel

$$K_{RBF}(X, X_i) = \exp(-\|X - X_i\|^2/2\sigma^2) \quad (7)$$

ERBF kernel

$$K_{ERBF}(X, X_i) = \exp(-\|X - X_i\|/2\sigma^2) \quad (8)$$

Spherical kernel

$$K_{Spherical}(X, X_i) = 1 - \frac{3}{2}\left(\frac{\|X - X_i\|}{\sigma}\right) + \frac{1}{2}\left(\frac{\|X - X_i\|}{\sigma}\right)^3 \quad (9)$$

Where σ is the bandwidth of the kernel function.

2.4 The experimental methodology

The experiments are conducted to compare various synthesis methodologies for deriving the 12-lead ECG from EASI-lead system. All dataset used in this work are obtained from PhysioNet database [12] 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 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 n[th] round, fold#n is used for testing while the remaining folds are used for training. For instance, in the 2[th] 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 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 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), Linear regression in Equation (2) and SVR in Equation (6).

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

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 signal, the RMSEs of each lead signal can be determined from the calculated signals and the ones from the PhysioNet dataset.

3. Results

The testing results with 5-fold cross-validation to find RMSE of Dower's method, Linear regression and SVR for 12 signals are listed in Table 1-2.

From Table 1-2, the minimum RMSEs of 12-lead are highlighted of each fold. The constant, coefficients, parameter σ of those folds with the minimum of RMSE used for derived ECG of 12 signals. The parameter σ of those folds with the minimum of RMSE used for derived ECG of 12 signals with RBF, ERBF and Spherical kernel function are shown in Table 3.

Then, using the transfer function models for each signal is tested with blind test data (the 10% part) and tested with the whole 4,810 data samples to find RMSE from all methods are shown and compared in Table 4.

In Figure 3 illustrates the relative of average RMSE errors for Dower's method, Linear regression and SVR using RBF, ERBF and Spherical kernel function.

Table 1 Root mean squared error with Dower's method and Linear regression

Signals	Dower's method					Linear regression				
	Fold#					Fold#				
	1	*2*	*3*	*4*	*5*	*1*	*2*	*3*	*4*	*5*
Lead I	33.608	30.026	29.693	30.993	28.152	24.380	23.373	25.086	26.846	23.861
Lead II	34.885	31.966	30.140	35.927	34.266	40.678	37.272	35.344	41.899	40.131
Lead III	54.207	47.657	42.845	54.354	46.284	47.419	42.953	40.908	51.353	44.795
Lead aVR	25.672	24.062	24.508	24.917	25.700	23.714	22.505	22.823	24.059	24.254
Lead aVL	40.243	35.191	32.393	38.959	32.672	31.746	29.126	28.967	35.214	29.755
Lead aVF	44.897	40.078	36.279	46.002	40.899	42.462	38.477	36.111	44.901	40.819
Lead V1	27.421	25.007	25.286	29.801	23.904	20.115	17.880	20.466	27.402	20.187
Lead V2	41.022	37.179	37.895	44.646	41.476	40.981	37.045	37.696	44.856	41.359
Lead V3	50.933	46.322	44.833	52.422	43.699	48.055	44.943	44.525	51.549	44.171
Lead V4	53.287	50.880	56.162	64.026	55.620	54.586	50.523	55.933	63.799	55.354
Lead V5	31.169	30.070	29.124	34.890	31.224	24.043	22.057	23.111	29.470	25.179
Lead V6	23.477	19.720	17.422	19.670	18.782	10.954	10.418	9.750	11.955	9.857

Table 2 Root mean squared error (mV) with SVR using RBF, ERBF and Spherical kernel function

Fold#	RBF kernel function											
	Signals (lead)											
	I	*II*	*III*	*aVR*	*aVL*	*aVF*	*V1*	*V2*	*V3*	*V4*	*V5*	*V6*
1	14.590	17.663	20.153	16.137	13.267	19.396	6.210	16.249	15.131	20.865	8.980	5.265
2	14.142	15.405	18.392	14.074	12.741	17.093	6.096	14.954	15.297	21.457	9.773	5.156
3	14.464	20.409	21.793	18.645	12.897	22.199	6.678	15.030	14.449	25.434	11.118	5.197
4	14.739	15.885	17.767	14.513	12.062	17.116	6.673	15.440	13.224	20.681	8.849	4.859
5	14.603	15.028	20.530	13.729	15.533	17.554	6.392	15.847	13.204	31.391	9.148	5.200
Fold#	**ERBF kernel function**											
	Signals (lead)											
	I	*II*	*III*	*aVR*	*aVL*	*aVF*	*V1*	*V2*	*V3*	*V4*	*V5*	*V6*
1	3.869	6.361	8.842	5.883	5.735	7.931	2.280	4.992	6.521	10.416	4.512	2.708
2	3.416	6.989	7.405	6.457	3.996	7.599	2.398	5.359	7.176	10.031	3.582	4.781
3	4.518	8.051	7.794	7.304	3.911	8.407	4.677	7.330	5.286	9.176	6.608	5.593
4	3.485	4.460	6.065	4.131	4.268	5.377	2.417	5.678	7.180	9.492	4.234	2.801
5	3.271	4.281	5.994	4.002	4.051	5.270	2.726	5.522	5.108	9.981	4.724	1.940
Fold#	**Spherical kernel function**											
	Signals (lead)											
	I	*II*	*III*	*aVR*	*aVL*	*aVF*	*V1*	*V2*	*V3*	*V4*	*V5*	*V6*
1	3.836	6.352	8.827	5.803	5.688	7.924	2.257	4.908	6.482	10.287	4.480	2.656
2	3.351	6.995	7.420	6.391	4.011	7.619	2.384	5.269	7.102	9.880	3.552	4.743
3	4.191	7.666	7.739	7.003	4.049	8.155	4.534	7.252	5.225	9.105	6.412	5.500
4	3.454	4.444	6.043	4.060	4.189	5.357	2.390	5.597	7.098	9.316	4.190	2.763
5	3.259	4.280	5.994	3.910	4.024	5.276	2.707	5.427	5.050	9.802	4.708	1.884

Table 3 The parameter σ of RBF, ERBF and Spherical kernel function

Signals	RBF		ERBF		Spherical	
	Fold#	σ	Fold#	σ	Fold#	σ
Lead I	2	1	5	2	5	10
Lead II	5	0.3	5	2	5	12
Lead III	4	0.3	5	2	5	11
Lead aVR	5	0.3	5	3	5	12
Lead aVL	4	0.3	3	1	2	2
Lead aVF	2	0.3	5	2	5	12
Lead V1	2	0.5	1	3	1	23
Lead V2	2	0.6	1	3	1	15
Lead V3	5	0.3	5	2	5	8
Lead V4	4	0.3	3	3	3	10
Lead V5	4	0.4	2	3	2	21
Lead V6	4	0.8	5	4	5	26

Table 4 Root mean squared error (mV) tested with blind test data versus with 4,810 data samples

Derived lead	Tested with blind test data of 10%.					Tested with 4,810 data samples.				
	Dower	Linear	SVR			Dower	Linear	SVR		
			RBF	ERBF	Spherical			RBF	ERBF	Spherical
Lead I	35.288	25.665	14.704	3.635	3.608	30.529	24.476	14.507	1.868	1.843
Lead II	32.476	37.704	15.613	6.584	6.587	33.216	33.216	16.877	2.856	2.857
Lead III	54.648	46.660	18.638	7.767	7.773	49.501	45.370	19.727	3.639	3.628
Lead aVR	24.088	22.037	11.477	4.655	4.576	24.595	23.152	15.419	2.165	1.992
Lead aVL	41.950	32.607	12.868	4.051	4.043	36.269	30.904	13.299	2.185	2.175
Lead aVF	43.574	40.427	18.565	7.508	7.521	41.757	40.375	18.671	3.362	3.367
Lead V1	27.438	20.460	6.244	2.232	2.212	26.053	20.868	6.409	1.276	1.251
Lead V2	40.801	40.807	16.807	6.408	6.331	40.254	40.205	15.500	3.082	2.960
Lead V3	49.371	47.581	13.624	6.530	6.461	47.409	46.273	14.261	3.568	3.521
Lead V4	54.262	53.723	21.377	12.008	11.901	55.419	55.407	23.96	5.600	5.470
Lead V5	35.083	24.579	9.218	8.591	8.558	31.436	24.670	9.573	2.164	2.124
Lead V6	23.152	12.079	5.224	3.733	3.682	20.001	10.655	5.135	1.531	1.459
Average	**38.511**	**33.694**	**13.697**	**6.142**	**6.104**	**36.370**	**32.964**	**14.445**	**2.775**	**2.721**

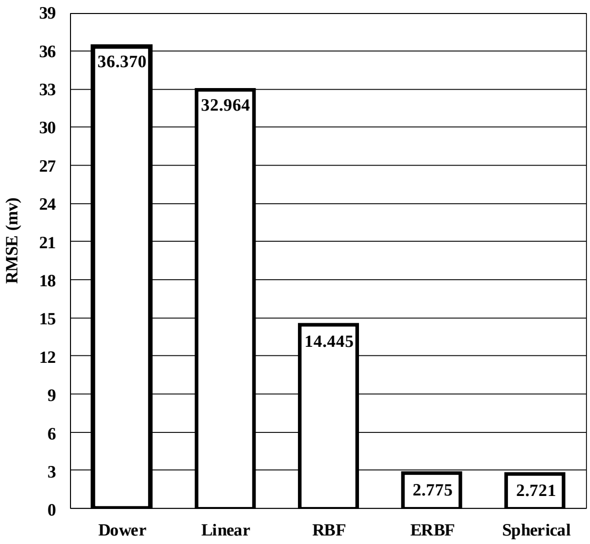
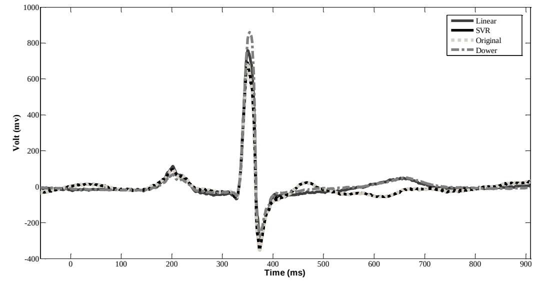
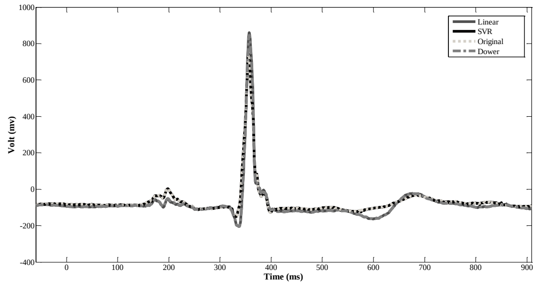


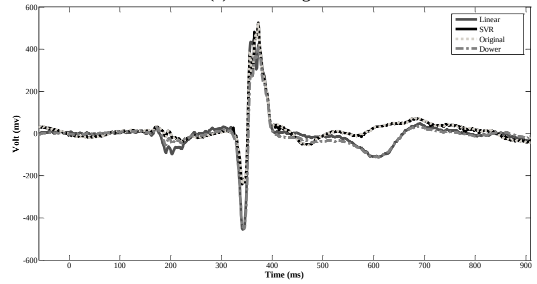
Figure 3 Comparison of average RMSEs errors from Dower’s method, Linear regression and SVR



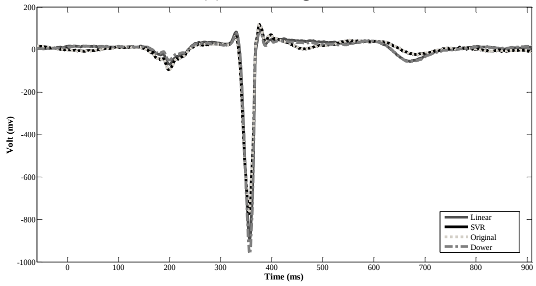
(a) Lead I Signal.



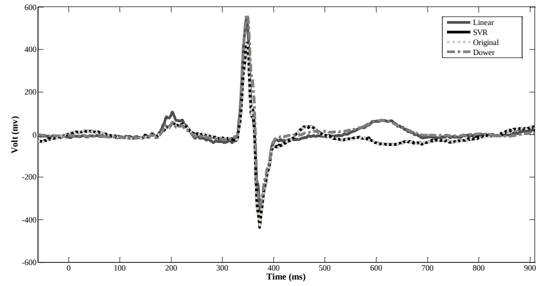
(b) Lead II Signal.



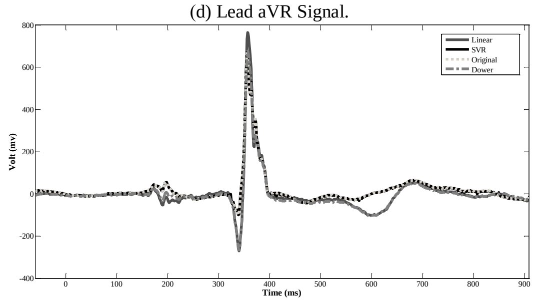
(c) Lead III Signal.



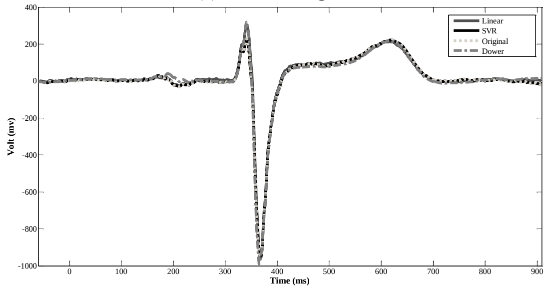
(d) Lead aVR Signal.



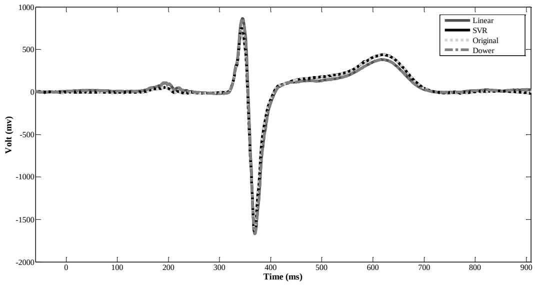
(e) Lead aVL Signal.



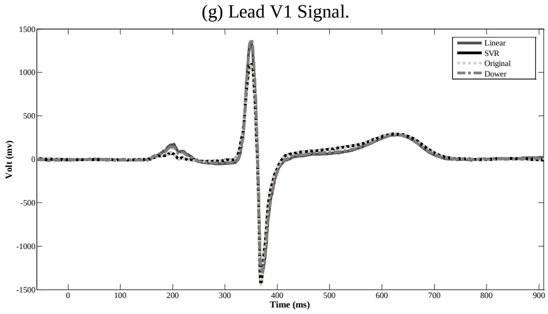
(f) Lead aVF Signal.



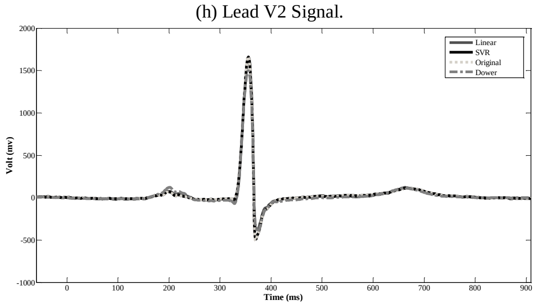
(g) Lead V1 Signal.



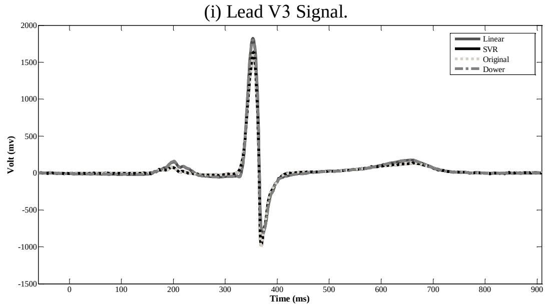
(h) Lead V2 Signal.



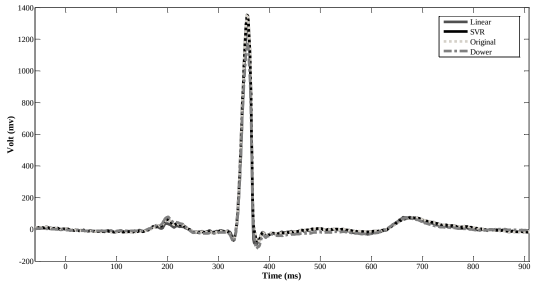
(i) Lead V3 Signal.



(j) Lead V4 Signal.



(k) Lead V5 Signal.



(l) Lead V6 Signal.

Figure 4 Derived vs original signals of 12-lead ECG signals

Lastly, the result graphs of 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 measured using standard 12-lead ECG method, derived using EASI-lead system by Dower's method, Linear regression and SVR using Spherical kernel function are shown in Figure 4(a-l).

4. Discussion and conclusions

This paper has presented SVR for deriving the standard 12-lead ECG from EASI-lead system. The experimental results from Table 4 and Figure 3-4 are showed that the best performance in this work, was obtained from the SVR using Spherical kernel function method, followed by SVR using ERBF, RBF kernel function, Linear regression and Dower's method, respectively. Therefore, it is obvious to conclude that nonlinear regression with support vector regression is worth chosen for deriving the 12-lead ECG from EASI-lead system.

As for future works, other regression and machine learning techniques to improve the performance of deriving the 12-lead ECG signals from EASI-lead system should be investigated further

5. References

[1] Berbari EJ. Principles of Electrocardiography. In: Bronzino JD. The Biomedical Engineering Handbook 2nd ed. Boca Raton: CRC Press LLC; 2000. p. 231-240.

[2] www.bem.fi [Internet]. New York: Bioelectromagnetism - Principles and Applications of Bioelectric and Biomagnetic Fields, [updated 2014; cited 2016 Mar 26]. Available from: http://www.bem.fi/book/.

[3] Tomasic I, Trobec R. Electrocardiographic Systems With Reduced Numbers of Leads—Synthesis of the 12-Lead ECG. IEEE Reviews in Biomedical Engineering 2014;7:126-142.

[4] Dower GE. A lead synthesizer for the Frank system to simulate the standard 12-lead Electrocardiogram. J Electrocardiol 1968;1:101-116.

[5] Frank E. An accurate clinically practical system for spatial vectorcardiography. Circulation 1956;13:737-749.

[6] Dower GE, Yakush A, Nazzal SB, Jutzy RV, Ruiz CE. Deriving the 12-lead electrocardiogram from four (EASI) electrodes. Journal of Electrocardiology 1988;21:182-187.

[7] Oleksy W, Tkacz E, Budzianowski Z. Improving EASI ECG Method Using Various Machine Learning and Regression Techniques to Obtain New EASI ECG Model. International Journal of Computer and Communication Engineering 2012;1:287-289.

[8] DTREG. Linear Regression [Internet]. 2014 [cited 2016 Mar 26]. Available from: https://www.dtreg.com/solution/view/28.

[9] Smola AJ, Scholkopf B. A tutorial on support vector regression. Technical report. ESPRIT working group in neural and computational leaning; 1998 Oct. No.: NC2-TR-1998-030.

[10] Theera-Umpon N, Boonprasert U. Leeway Prediction of Oceanic Disastrous Target via support vector regression. JACIII 2004;8:606-612.

[11] Yodjaiphet A, Theera-Umpon N, Auephanwiriyakul S. Electrocardiogram reconstruction using support vector regression. IEEE International Symposium on Signal Processing and Information Technology (2012); 2012 Dec 12-15; Ho Chi Minh, Vietnam. New Jersey: IEEE; 2013. p. 269-273.

[12] PhysioNet. PhysioNet database [Internet]. 2007 [cited 2016 Mar 26]. Available from: http://www.Physionet.org/challenge/2007/data/.