ECG Decomposition using Cascaded Spline Projection Residual Auto Encoders

Kaveh Samiee[1], Péter Kovács[2]

[1] GE Healthcare, Helsinki, Finland [2] Department of Numerical Analysis, Eötvös Loránd University, Budapest, Hungary

Abstract

Recent advancements in remote or handheld patient monitoring devices have led to the development of novel domain-specific AI architectures that enable more accurate and faster real-time ECG diagnosis. We present a data-driven framework for decomposition of ECG signals based on B-Spline Variable Projection Neural Networks (VPNN) and cascaded residual auto encoders (AE). We use VPNN with B-spline bases in regressor mode. Hence, the output of each VPNN layer is an estimation of the input. ECG segment is passed through a set of cascaded VPNN regressors, where the input of each VPNN layer is the residual of the ECG segment and the output of its preceding VPNN regressor. In such a topology, the output of each VPNN can be interpreted as a component of the input representing specific frequency and morphological characteristics. The effectiveness of the decomposition framework was demonstrated in detection of ventricular tachycardia (VT) and ventricular flutter (VFL) ECG segments with a sensitivity rate and a false positive rate of 92% *and* 9.5%*, respectively.*

1. Introduction

Conventional heart rhythm analysis techniques of ECG signals encompass a wide range of human-assisted feature extraction techniques that are used to derive significant and illustrative insights about cardiac health. For most of the morphological and time domain human-assisted feature descriptors, delineation of ECG waves is a prerequisite step. For instance, to calculate RR-intervals and QT-intervals, QRS waves must be first detected and then the locations of P and T waves are estimated accordingly. Thus, the effectiveness of such feature descriptors are highly correlated with the beat detection algorithm performance. Alternatively, using spectral analysis it is feasible to perform both QRS detection and delineation of ECG waves steps in frequency domain. In both time and frequency domains, common signal decomposition methods such as Empirical mode decomposition [1], Wavelet decomposition [2] and Sparse rational decomposition [3] can be uses as a basis to enhance discriminatory power of human-assisted feature descriptors in ECG processing. All of these methods require a priori knowledge in order to be adapted effectively to each specific signal and application. To overcome this limitation, we propose a new signal decomposition technique which can be trained in unsupervised manner to adjust itself into input signals. As a case study, we evaluate its robustness by applying it to single-channel ECG segment in order to distinguish ventricular tachycardia (VT) and ventricular flutter (VFL) arrhythmia from other types of heart rhythms.

2. Related works

The theory of variable projection (VP) [4] provides a very effective framework for solving separable nonlinear least squares (SNLLS) problems of the form

$$\min_{\theta} r_2(\theta) := \min_{\theta} \left\| y - \Phi(\theta)\Phi^{+}(\theta) y \right\|_2^2, \qquad (1)$$

where y represents the signal with N samples, and $\Phi^{+}(\theta)$ stands for the Moore–Penrose pseudoinverse of the matrix $\Phi(\theta) \in \mathbb{R}^{N\times m}$. For a proper choice of θ, the input signal y can be well approximated by its orthogonal projection $\tilde{y}(\theta)$ onto the columnspace of $\Phi(\theta)$, i.e.,

$$y \approx \tilde{y}(\theta) := \Phi(\theta)(\Phi^{+}(\theta)y) = \Phi(\theta)c(\theta), \qquad (2)$$

where $c(\theta)$ denotes the coefficients of the underlying projection. Golub and Pereyra referred to r_2 as the VP functional, which has since been widely applied in signal processing and related fields [5]. Particularly, this approach proved to be very efficient in cases when $q \ll m$, i.e., a larger set of linear parameters $c(\theta) \in \mathbb{R}^m$ is controlled by a small number of nonlinear variables $\theta \in \mathbb{R}^q$.

The projection in Eq. (1) can be interpreted as a layer with trainable weights θ, allowing the embedding of the VP operator into neural network architectures [6]. The resulting VP layer can operate in two modes based on its output. In regression mode, the projected signal $\tilde{y}(\theta)$ is forwarded to subsequent layers, i.e., the VP layer preforms filtering on the input signal. In classification mode, the coefficient vector $c(\theta)$ is forwarded to the next layer, facilitating knowledge-driven feature learning. This means that the VP layer learns how to represent the input signal in the subspace spanned by the columns of $\Phi(\theta)$. In addition to that, expert knowledge can be injected into the network by selecting the structure and parametrization of the projector in a way that aligns with the requirements of the target application.

Note that in Eq. (2), the gradients of both the coefficient vector $c(\theta)$ and the projected signal $\tilde{y}(\theta)$ can be explicitly calculated provided that the partial derivatives of the jth column of $\Phi(\theta)$ exist with respect to all coordinates of θ (see e.g., Lemma 4.1 and Theorem 4.3 in Ref. [4]). In general, if the computation of $\Phi_j(\theta)$ and $\Phi_j(\theta)/\partial\theta_k$ is available for all possible j and k, then the forward and backward passes of the corresponding VP layer can be easily implemented. In the next section, we define these terms for free-knot splines.

3. Spline based VP layers

To date, the basic structure of VPNet has been generalized to artificial and spiking neural networks (SNNs) [6,7], and applied to various real-world problems, such as road abnormality detection [8], classification of cardiac arrhythmias [6] and visually evoked potentials [9]. In these applications, the VPNet concept was realised by using a parametrized variation of the classical Hermite functions. Utilizing our former work [10], we propose a new implementation of VPNet that is based on free-knot splines.

Splines are smooth piecewise polynomial curves that include a wide class of functions, such as Hermite, Bézier, and Catmull-Rom splines. In this study, we will consider the B-spline basis for representing piecewise polynomial functions $s : [a, b] \to \mathbb{R}$ of class $C^\ell[a, b]$ whose shape is controlled by two parameters: the degree ℓ of smoothness, and the so-called knot vector τ_n. B-splines can be defined recursively as follows:

$$B_{0,k}(\tau_n; x) = \begin{cases} 1 & \text{if } x \in [t_k, t_{k+1}), \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

$$B_{\ell,k}(\tau_n; x) = (t_{k+\ell+1} - t_k) \cdot [t_k, \ldots, t_{k+\ell+1}]\,(x - t)_+^\ell\,,$$

where $t_k \in [a, b]$ are the knots, and $(x - t)_+^\ell = (\max\{x - t, 0\})^\ell$ denote the so-called truncated power function (TPF). Note that variable x of each TPF is fixed, and the $(\ell+1)$th divided differences $[t_k, \ldots, t_{k+\ell+1}]\,(x - t)_+^\ell$ are computed for the second variable t. This formulation allows to compute the partial derivatives of the B-spline functions $B_{\ell,k}$ $(\ell \geq 1)$ with respect to the free knots as follows:

$$\frac{\partial B_{\ell,j}(\tau_n; x)}{\partial t_k} = [t_j, \ldots t_k, t_k \ldots, t_{j+\ell+1}]\,(x - t)_+^\ell, \tag{4}$$

provided that $j \leq k \leq j + \ell + 1$, and it is zero otherwise.

In order to integrate free-knot splines into the VPNet framework, we consider a sequence of knots $a = t_0 < t_1 < \ldots < t_{n-1} < t_n = b$ with the following boundary conditions:

$$a = t_0 = t_{-1} =, \ldots = t_{-\ell}, \tag{5}$$

$$b = t_n = t_{n+1} = \ldots = t_{n+\ell}\,. \tag{6}$$

We fix the outer knots, i.e., t_0 and t_n, whereas the inner knots $t_1, \ldots, t_{n-1}$ are assumed to be variable. Now, we connect Eq. (1) and Eqs. (3)-(4) by defining the parametrized knot vector:

$$\tau_n(\theta) := (t_{-\ell}, \ldots, t_0, \theta_1, \ldots, \theta_{n-1}, t_n, \ldots, t_{n+\ell})^T, \tag{7}$$

where $t_k = \theta_k$ $(k = 1, \ldots, n-1)$ hold for the inner knots. Then, the forward and backward passes of the corresponding spline based VP layer can be implemented by taking

$$\Phi_j(\theta) = B_{\ell,j}(\tau_n(\theta)), \quad \frac{\partial \Phi_j(\theta)}{\partial \theta_k} = \frac{\partial B_{\ell,j}(\tau_n(\theta))}{\partial \theta_k}, \tag{8}$$

where the B-splines $B_{\ell,j}$ are sampled uniformly on the interval $[a, b]$ with N samples, thus we have omitted x from the notation. The number of free knots n, and the degree of smoothness ℓ are treated as hyperparameters, which also determine the dimension of the coefficient vector $c(\theta)$ that is $m = n + \ell$ in Eq. (2). These hyperparameters must be set before training such VP layers.

Due to their flexibility, splines have a wide range of applications including computer aided design [11], regression analysis [12], signal and image processing [13]. The proposed spline based VP layer facilitates the incorporation of expert knowledge from these fields into trainable frameworks.

4. Cascaded VP residual auto encoders

4.1. VP as auto encoder

Auto encoders (AEs) are neural network topologies that consist of two main components: encoder part which performs non-linear mapping of input data y into a compact representation called latent space; decoder part which reconstructs the original input by projecting the latent space back to the input space. The objective is then to minimize the reconstruction error in an unsupervised manner. AEs are non-linear systems due to semi-linear or non-linear behavior of neurons' activation functions.

As described in section 2, a VP layer in regression mode estimates the input signal y by first mapping it to coefficient space $c(\theta)$ and then back projection of coefficients using $\Phi(\theta)$ to obtain $\tilde{y}(\theta)$. The two subsequent projection steps of a VP regressor are analogous to the encoder and decoder components of a conventional AE architecture. In fact, $c(\theta)$ represents the latent space of the input data obtained by $\Phi^{+}(\theta)$. When using spline functions in a VP layer, the projection consists of linear combination of piecewise polynomial functions. Moreover, higher-degree B-spline functions increase the curvature of the VP regressor. Thus, similar to AEs, a VP layer in regression mode is also a non-linear system.

4.2. Cascaded spline projection residual auto encoders

When dealing with a high dimensional data, shallow AEs have a limited ability to represent the input data into a low-dimensional and meaningful latent space. Moreover, their construction accuracy drops by increasing the compression ratio between input and latent spaces. One strategy to address such limitations is to increase the depth of AE by introducing new hidden-layers into the architecture. Alternatively, several shallow AEs can be stacked together to enhance the performance of a shallow AE. Cascaded auto encoders (CAEs) are constructed by layering multiple AEs at top of one another. In such an architecture, each AE aims at reconstruction of the previous AE output. Compared to the deeper counterpart, cascaded AEs are computationally less complex and more suitable for applications where computational resources are constrained. By adding residual connections between the input nodes in CAE, each AE layer can learn to reconstruct the residual of the previous AE layer. In residual CAEs, each AE learns a partial representation of input data associated to certain features and characteristics.

The cascaded spline projection residual AEs model is constructed by deploying B-spline VP layers in regression mode as AE layers in the residual CAE framework. Figure 1 illustrates the topology of the proposed architecture.

4.3. ECG decomposition

The objective of signal decomposition is to disentangle each signal into its constituent components. Each component represents a specific set of intrinsic features of the signal such as temporal dynamics and underlying patterns. Thus, in the presence of complex and non-stationary signals, it is easier to characterize each signal by analysing its components. Based on the proposed cascaded spline projection residual AEs architecture, we develop a new signal decomposition technique capable of breaking down signal segments into set of empirically derived components. Let us assume a residual CAE model with $(L \geq 1)$ B-spline VP regressors, the transfer function of the network can be written as:

$$f(y) \approx \tilde{y} := \sum_{l=1}^{L} \tilde{y}(\theta_l), \tag{9}$$

where $\tilde{y} \in \mathbb{R}^n$ is the reconstruction of the input signal $y \in \mathbb{R}^n$, and $\tilde{y}(\theta_l)$ is the output of the l_{th} B-spline VP regressor calculated according to Eq. (2). For set of P signal segments with length of N samples, $\{\theta_l\}_{l=1}^{L}$ parameters of the model are optimized to minimize the average reconstruction error:

$$\underset{\{\theta_l \in \mathbb{R}^q\}_{l=1}^{L}}{\arg\min} \frac{1}{P} \sum_{i=1}^{P} \left\| y^{(i)} - f(y^{(i)}) \right\|_2^2. \tag{10}$$

To ensure that each B-spline VP layer distinctly captures the time and frequency information of the input signal compared to the other layers, we intentionally configure them with splines functions having different degrees of smoothness denoted as ℓ in Eq. (3).

5. Experiments

The publicly accessible MIT-BIH Malignant Ventricular Ectopy Database (VFDB) [14] is used in our experiments. This database includes 22 two-channel ECG recordings from subjects who experienced episodes of sustained VT, VFL, and ventricular fibrillation (VFIB). Each ECG record is with duration of half an hour and was digitized with 12-bit resolution and at sampling frequency of 250 Hz. VFDB contain 15 different rhythms annotations, including VFL, VFIB, normal sinus rhythm, atrial fibrillation, and other rhythms. In total, VFDB contains 60 episodes of VFL and 89 episodes of VT. In this study, we only use lead II of each ECG record.

A cascaded spline projection residual AE is constructed using five B-spline VP regressors that are initialized with only a small number of coefficients to permit localization of basis functions. Additionally, we arrange the B-splines VP regressors with varying degrees of smoothness, i.e. order, in the ascending range of $\ell = 6, \ldots, 2$. Thus, each ECG segment is first passed through the highest order B-spline VP regressor. We note that higher order B-spline VP regressors have a higher chance to fit high frequency components of input signal. Single-channel ECG signals from VFDB dataset are first split into two-second segments, and then rhythm annotations are propagated accordingly. All segments marked as noise are then discarded. In one-class novelty detection manner, the topology is trained on all segments except VT and VFL ones. The average normalized percent root mean squared difference (PRD) of 3.68% and 4.21% were obtained on training and test sets, respectively. By training a linear one-class Support Vector Machine (SVM) classifier on the resulting PRD values, VT/VFL segments are identified with an average accuracy of 82.6%. Additionally, by training the SVM classifier on coefficients of the B-spline VP regressors, a sensitivity rate of 92% and a false positive rate of 9.5% were obtained. Figure 2 show a decomposition example of a two-second

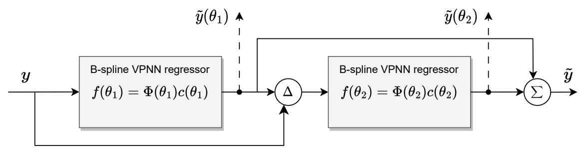

Figure 1. A cascaded spline projection residual AEs model with two B-pline VPNN regressors.

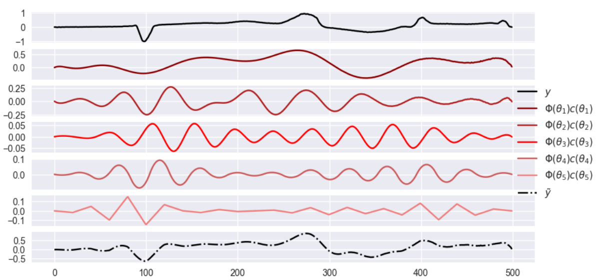

Figure 2. Decomposition and reconstruction of a 2 second ECG segment with ventricular ectopic activity, (MSE is roughly 0.008).

long ECG signal segment from VFDB dataset obtained using the cascaded spline projection residual AE model.

6. Conclusions

In this paper we presented a novel signal decomposition framework which is trainable and can adapt itself according to underlying patterns of the input. The framework can be used for feature extraction or compression of signals. Due to the compact topology of VPNN regressors, the framework is composed of a small set of parameters, making it a suitable choice for edge computing and wearable devices. The principal used for decomposing signals into set of non-orthogonal components is generic and the framework is applicable to any composite signal. In future works, orthogonality of components should be assured in order to obtain a more compact and discriminatory representation of signals.

7. Code Availability

The data and code of this study are openly available at: `https://github.com/KavehSam/CBSPRAE`

Acknowledgements

This project was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

References

[1] Pal S, Mitra M. Empirical mode decomposition based ecg enhancement and qrs detection. Computers in biology and medicine 2012;42(1):83–92.

[2] Martínez JP, Almeida R, Olmos S, Rocha AP, Laguna P. A wavelet-based ecg delineator: evaluation on standard databases. IEEE Transactions on biomedical engineering 2004;51(4):570–581.

[3] Samiee K, Kovács P, Kiranyaz S, Gabbouj M, Saramaki T. Sleep stage classification using sparse rational decomposition of single channel eeg records. In 2015 23rd European Signal Processing Conference (EUSIPCO). IEEE, 2015; 1860–1864.

[4] Golub GH, Pereyra V. The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate. SIAM Journal on Numerical Analysis SINUM 1973;10:413–432.

[5] Golub GH, Pereyra V. Separable nonlinear least squares: The variable projection method and its applications. Inverse problems 2003;19(2):R1–R26.

[6] Kovács P, G. B, Huber C, Huemer M. VPNet: variable projection networks. International Journal of Neural Systems IJNS 2021;32:2150054:1–19.

[7] Kovács P, Samiee K. Arrhythmia detection using spiking variable projection neural networks. In Proc. 49th Computing in Cardiology (CinC), *Tampere, Finland, Sept. 4–7*. 2022; 1–4.

[8] Dózsa T, Radó J, Volk J, Kisari A, Soumelidis A, Kovács P. Road abnormality detection using piezoresistive force sensors and adaptive signal models. IEEE Transactions on Instrumentation and Measurement 2022;71:1–11.

[9] Dózsa T, Böck C, Bognár G, Meier J, Kovács P. Color classification of visually evoked potentials by means of Hermite-functions. In Proc. of the 55st Annual Asilomar Conf. on Signals, Systems, and Computers. 2021; 251–255.

[10] Kovács P, Fekete AM. Nonlinear least-squares spline fitting with variable knots. Applied Mathematics and Computation 2019;354:490–501.

[11] Piegl L. On NURBS: a survey. IEEE Computer Graphics and Applications 1991;11(1):55–71.

[12] Perperoglou A, Sauerbrei W, Abrahamowicz M, Schmid M. A review of spline function procedures in R. BMC Medical Research Methodology 2019;19(46):1–16.

[13] Unser M. Splines: a perfect fit for signal and image processing. IEEE Signal Processing Magazine 1999;16(6):22–38.

[14] Greenwald SD. The development and analysis of a ventricular fibrillation detector. Ph.D. thesis, Massachusetts Institute of Technology, 1986.

Address for correspondence:

Kaveh Samiee
Kuortaneenkatu 2, 00510 Helsinki, Finland.
kaveh.samiee@ge.com