

# Deep Generative Model and Analysis of Cardiac Transmembrane Potential

Sandesh Ghimire<sup>1</sup>, Linwei Wang<sup>1</sup>

<sup>1</sup>Computational Biomedicine Lab, Rochester Institute of Technology, NY, USA

## Abstract

*It has been shown recently that inverse electrophysiological imaging can be improved by using a deep generative model learned in an unsupervised way so that cardiac transmembrane potential and underlying generative models could be simultaneously inferred from the ECG. The prior and conditional distributions learned in such a way are, however, directly affected by the architecture of neural network used in unsupervised learning. In this paper, we investigate the effect of architecture in learning representation and generalizing to new test cases. By comparing reconstruction of three types of sequence autoencoder, we show that different sequence autoencoders might be focusing on different aspects of TMP and might perform differently according to the metric used to measure reconstruction. We also analyze the latent space in different architectures and discuss important questions raised by these observations.*

## 1. Introduction

Noninvasive electrophysiological imaging refers to the estimation of cardiac electrical signals from the body surface electrocardiograms (ECG). It requires solving an ill-posed inverse problem where the difficulty of estimation increases as we move from estimation of epicardial potential to that of myocardial and endocardial potential. To compensate the lack of information in the ECG signal, prior knowledge about the cardiac potential is used to constrain the inverse solution [1, 2]. An effective way to constrain inverse solution is to use physiological knowledge about cardiac electric propagation. This is realized by constraining the cardiac electrical activity according to a physiological generative model [3, 4] during inference. The physiological model is typically parameterized by model parameters. During inference, it is highly desirable to simultaneously adapt high dimensional inverse signal as well as model parameters given the ECG data. However, reliable inference is very difficult and computationally expensive because of high dimensionality of model parameters and their complex relation with electric signal, which, often, cannot be expressed in closed form. As a workaround, researchers focus on estimation of model pa-

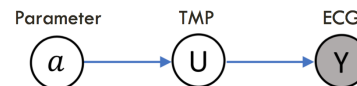


Figure 1. Probabilistic graphical model of ECG generation

rameters only [5], inverse signal only [4] assuming fixed model parameters or inverse signal as well as model error introduced due to error in parameters [6].

Ghimire et al [7] addressed the challenges in simultaneous estimation of parameters as well as cardiac transmembrane potential (TMP) by replacing complex physiological model with a deep generative model learnt by using autoencoder from samples of TMP. In this approach, the quality of relation between latent factor and the TMP as well as prior distribution of latent factor are all directly affected by the architecture of autoencoder used in learning. Hence, in this paper, we try to understand the role of autoencoder architecture in learning and representation of the TMP by comparing reconstruction ability of three different architectures of sequence autoencoder.

## 2. Background

### 2.1. Probabilistic Formulation of Inverse Problem

The biophysical relationship between cardiac transmembrane potential (TMP),  $U$  and body-surface ECG,  $Y$  can be described by a linear measurement model [4]:  $Y = HU$ , where  $Y$  and  $U$  are both matrices with each column denoting one time instant,  $H$  is called the forward matrix and is specific to the heart-torso model of an individual. The inverse problem is to  $U$  from  $Y$ .

We utilize the probabilistic graphical model framework to represent the generation of TMP and ECG as shown in Fig. 1. The forward relation is modeled as a likelihood distribution  $p(Y|U) = \mathcal{N}(HU, \sigma^2\mathbf{I})$ . The prior generation model of TMP given the value of model parameters can be represented with another conditional distribution as  $p(U|a)$ . The ECG matrix  $Y$  is shaded because it is observed random variable. The objective here is to obtain joint posterior distribution  $p(U, a|Y)$ . Typically, there is no closed form relation for  $p(U|a)$  and is given by some simulation model, because of which the posterior distribu-

tion is analytically intractable. Sampling based approach could be used in such cases, but it would require running simulation for a large number of times, and is, thus, computationally quite expensive if not infeasible.

## 2.2. Deep Generative Solution

Ghimire et. al. [7] addresses the problem of joint estimation by replacing original TMP generation model with a deep generative model. To do so, a sequence variational autoencoder is used to learn a low dimensional latent representation. This latent variable becomes low dimensional generative factor,  $Z$  and the decoder of autoencoder becomes the deep generative model. Thus, authors present an equivalent probabilistic graphical model replacing complex relation between physiological parameters and TMP with deep generative model, which is simple enough to allow joint estimation of generative factor and posterior distribution of TMP given the ECG. Simultaneously adapting generative factor with TMP results in its better estimate.

Obviously, the autoencoder used to learn representation in an unsupervised manner plays a crucial role in this task. In this paper, we try to analyze how different architecture choice for autoencoder affects its representation and generalization ability.

## 3. Preliminaries

### 3.1. Autoencoder

Autoencoder is a deep neural network that reconstructs the same signal at the output as is fed in the input by going through multiple layers of neural network. A common architecture of neural network has bottleneck which compresses input signal to a low dimensional latent code and then reconstructs the original signal from latent code. The portion of neural network from input to the latent code is called an encoder and that reconstructs signal from latent code is called decoder. Depending on the training procedure, there are many versions like denoising, contractive or sparse autoencoder. A slightly different version is called variational autoencoder [8] which derives loss function from variational lower bound of data log likelihood and introduces stochasticity to the latent code.

### 3.2. Sequence Encoder Decoder

Autoencoders are usually trained in a setting with vector input and output; sequence autoencoders have been explored only recently in the machine learning literature. To deal with sequences, recurrent neural networks (RNNs) and long short-term memory (LSTM)s are used in both the encoder and decoder inspired by sequence to sequence language translation [9] in natural language processing

(NLP). Typically, to obtain a compressed latent representation from the whole sequence, last hidden code from latent sequence is considered and full sequence is reconstructed from it using multiple layer LSTM in the decoder.

## 4. Architectures

In this paper, we experiment with three architectures: Language, svS and sSS architecture. Language architecture is of the same form as used in Language translation [9], and has deterministic latent space. Other two architectures have stochastic latent vector as described in [10]. Fig. 2 shows a general architecture for a stochastic model at the bottom half. As shown, both encoder and decoder has two layer LSTMs. Both, svS and sSS are in the stochastic setting where there are two networks for mean( $M$ ) and variance( $S$ ) while the Language architecture does not have variance network. The major difference in three architectures is explained in the top half of the Fig. 2. In the language model, the output from last hidden unit of LSTM is directly fed to the decoder and then subsequent predictions are computed recurrently. The svS architecture uses additional fully connected layers to map sequence of latent codes into a vector – hence the name sequence to vector to sequence (svS). In the sSS architecture, however, the hidden codes from all units are represented as a matrix latent code from which input TMP signal is reconstructed through a mirrored architecture.

## 5. Experiments

### 5.1. Implementation details

Training and test sets of transmembrane potential (TMP) were generated by using Aliev Panfilov model [11] on a human-torso geometry model. By varying two parameters: origin of excitation and tissue properties representing myocardial scar, we generated about 600 simulation data with the combination of 17 different tissue property configurations and 35 different origins of excitation. To test generalization ability, test data were selected with different origin of excitation than those used in training.

We used ReLU activation function in both encoder and decoder, ADAM optimizer and a flat learning rate of  $10^{-3}$  in all three architectures.

### 5.2. Results and Discussion

We compare three architectures in their ability to generalize in new test cases. We measure the reconstruction accuracy with four metrics: 1) mean square error (MSE) of TMP, 2) correlation of TMP, 3) dice coefficient of the scar region, 4) correlation of activation time. Fig. 3 compares TMP propagation reconstructed by using three architectures with the ground truth. The Language model

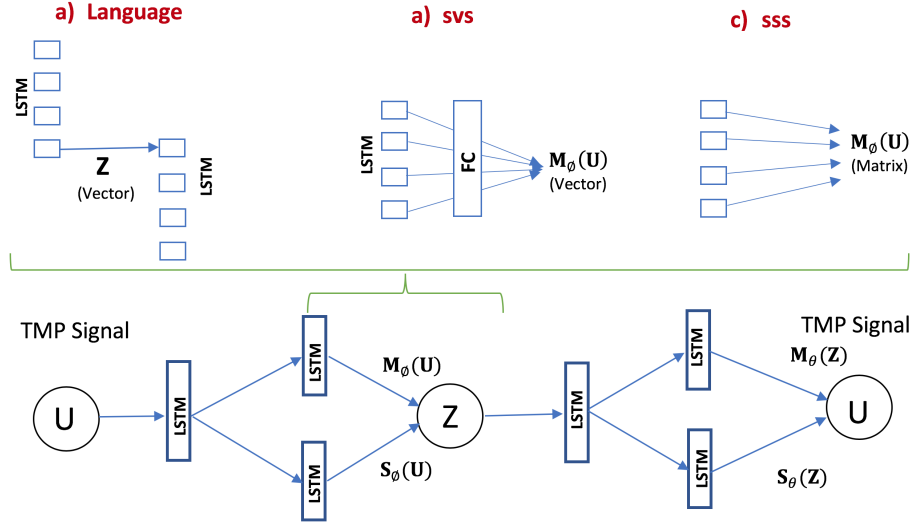


Figure 2. Bottom: Common skeleton for three architectures, Top: Three architectures of differing in their ways of converting output from last layers of LSTM to latent representation

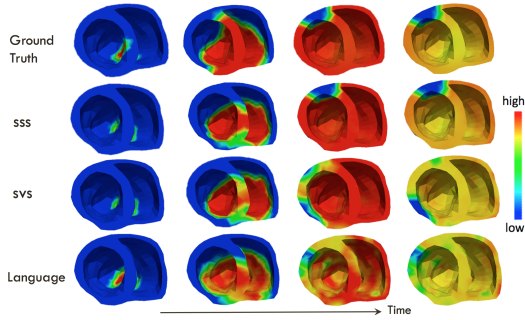


Figure 3. Comparison of transmembrane potential propagation

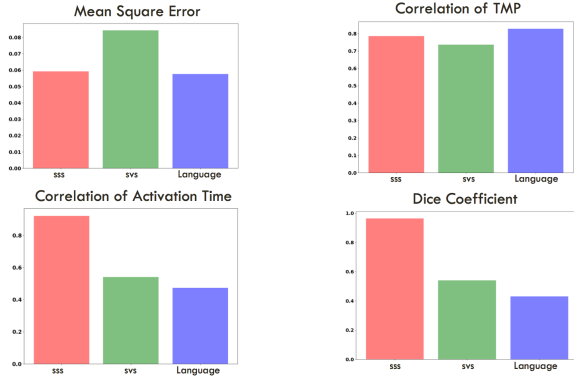


Figure 4. Comparison of reconstruction using three architectures

matches the ground truth better than other architectures at the beginning of the propagation sequence. However, later on, other two methods are qualitatively better. The scar region, however, seems to be better identified by sss architecture compared to other two.

The graphs on Fig. 4 shows average of 20 tests, each performed by randomly drawing 200 samples from the test set. It is interesting that the Language model performs quite good when measured with mean square error and correlation of TMP. But, when we measure dice coefficient of scar and correlation of activation time derived from the reconstructed TMP, the Language architecture performs the worst. It suggests that Language model might be good at preserving temporal consistency but not so much at learning underlying factors. On the other hand, svs and sss architectures seem to be better learning underlying factors, which might be because of the stochastic latent space in these two architectures.

We also visualized the latent representation of the whole dataset- training and test set- of three different architectures. In Fig. 5, top row shows latent point cloud colored according to the segment where origin of excitation lies. Similarly, bottom row shows latent point cloud colored according to the segment where scar lies. The heart is divided into 17 segments according to American heart association (AHA) standard and each color denotes one segment of the heart where scar/origin lies. We observe that the latent representation is clustered by the location of origin of excitation in all three architectures, but not by the location of scar region. It might be because there were not many examples of scar regions from the same segment for the network to generalize. We need further analysis.

The results are thought-provoking. However, we caution that our work is preliminary in that we tried it on a single geometry and relatively small dataset for a deep network. We leave some open questions triggered by these observations: What does it mean that a method performs better with respect to RMSE error but not so well with respect to

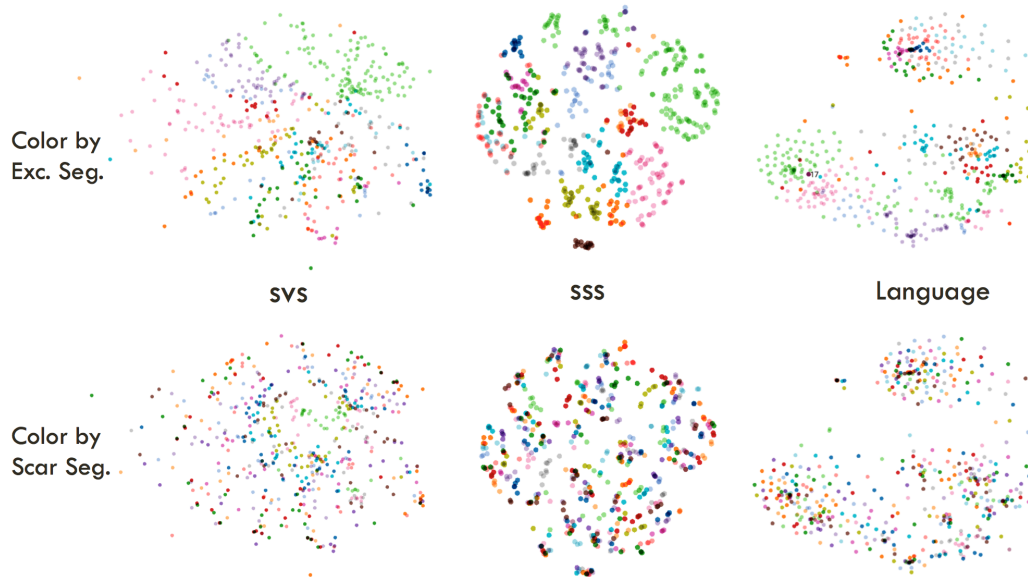


Figure 5. Visualization of point cloud in the latent space corresponding training and test data

error of origin of excitation? Does stochasticity play a role in better representation? Why did the neural network better represented the origin of excitation than the scar region?

## 6. Conclusion

We have shown that different choices of autoencoder architecture affects generalization ability as well as reconstruction of underlying aspects of TMP. The pattern in latent representation was consistent among architectures but differed according to origin or excitation and region of scar. In the future, we will confirm these findings in a larger dataset and investigate deeper to understand these differences.

## Acknowledgements

This work is supported by the National Science Foundation under CAREER Award ACI-1350374 and the National Institute of Heart, Lung, and Blood of the National Institutes of Health under Award R21HL125998.

## References

- [1] Greensite F, Huiskamp G. An improved method for estimating epicardial potentials from the body surface. *IEEE TBME* 1998;45(1):98–104.
- [2] Brooks DH, Ahmad GF, MacLeod RS, Maratos GM. Inverse electrocardiography by simultaneous imposition of multiple constraints. *IEEE TBME* 1999;46(1):3–18.
- [3] Pullan A, Cheng L, Nash M, Bradley C, Paterson D. Non-invasive electrical imaging of the heart: theory and model development. *Annals of Biomedical Engineering* 2001; 29(10):817–836.
- [4] Wang L, Zhang H, Wong KC, Liu H, Shi P. Physiological-model-constrained noninvasive reconstruction of volumetric myocardial transmembrane potentials. *IEEE Transactions on Biomedical Engineering* 2010;57(2):296–315.
- [5] Dhamala J, Ghimire S, Sapp JL, Horáček BM, Wang L. High-dimensional bayesian optimization of personalized cardiac model parameters via an embedded generative model. In *International Conference on MICCAI*. Springer, 2018; 499–507.
- [6] Ghimire S, Sapp JL, Horacek M, Wang L. A variational approach to sparse model error estimation in cardiac electrophysiological imaging. In *International Conference on MICCAI*. Springer, 2017; 745–753.
- [7] Ghimire S, Dhamala J, Gyawali PK, Sapp JL, Horacek M, Wang L. Generative modeling and inverse imaging of cardiac transmembrane potential. In *International Conference on MICCAI*. Springer, 2018; 508–516.
- [8] Kingma DP, Welling M. Auto-encoding variational bayes. *arXiv preprint arXiv13126114* 2013;.
- [9] Sutskever I, Vinyals O, Le QV. Sequence to sequence learning with neural networks. In *Advances in neural information processing systems*. 2014; 3104–3112.
- [10] Ghimire S, Gyawali PK, Sapp JL, Horacek M, Wang L. Improving generalization of sequence encoder-decoder networks for inverse imaging of cardiac transmembrane potential. *arXiv preprint arXiv181005713v1* 2018;.
- [11] Aliev RR, Panfilov AV. A simple two-variable model of cardiac excitation. *Chaos Solitons Fractals* 1996;7(3):293–301.

### Address for correspondence:

Sandesh Ghimire - sg9872@rit.edu; Web : www.sandeshgh.com  
 Computational Biomedicine Lab  
 Rochester Institute of Technology  
 1 Lomb Memorial Drive, Rochester, NY-14623, USA