

## GERAÇÃO DE ECG SINTÉTICO COM AUTOCODIFICADOR VARIACIONAL

Eduardo P.L. Jaqueira, Renato Candido e Magno T.M. Silva

Escola Politécnica, USP, SP

eduardo.jaqueira@usp.br, renatocan@lps.usp.br, magno.silva@usp.br

### Objetivos

O aumento de dados (*data augmentation*) tem por objetivo a expansão e/ou balanceamento de bancos de dados para obter melhores métricas de classificação em aprendizado de máquina. Neste trabalho, sinais de eletrocardiograma (ECG) sintéticos foram gerados utilizando um modelo de rede neural do tipo autocodificador variacional (*variational autoencoder* – VAE). Os sinais sintéticos foram incluídos ao conjunto de treinamento de uma rede perceptron multicamada (*multilayer perceptron* – MLP) para classificação de arritmias cardíacas. A inclusão desses dados proporcionou uma melhora na taxa de acerto da rede.

### Métodos e Procedimentos

A classificação de arritmias cardíacas depende de uma análise precisa da atividade elétrica do coração, registrada no sinal de ECG. Esse sinal contém informações clínicas que permitem que essa tarefa seja realizada por especialistas qualificados, o que pode ser demorado e é inviável quando se necessita de resultados imediatos. Diante disso, soluções baseadas em técnicas de aprendizado de máquina para a classificação de arritmias cardíacas têm sido amplamente estudadas [1]. Entretanto, altas taxas de erro dessas soluções fazem com que elas ainda estejam longe de serem empregadas na prática. Em [2], foram propostas soluções baseadas em aprendizado de máquina, incluindo redes MLP, para classificação automática de arritmias cardíacas. Foram utilizados sinais de ECG do banco de dados MIT-BIH *Arrhythmia Database* (MITDB), levando-se em conta quatro classes: batimentos do nó sinoatrial (N), supraventriculares ectópicos (S), ventriculares ectópicos (V) e fusão de batimentos normais e ventriculares ectópicos (F). Algumas soluções propostas em [2] alcançaram métricas de classificação superiores às da literatura para as classes S e V.

Apesar de terem sido consideradas técnicas usuais para lidar com classes desbalanceadas em [2], acredita-se que métricas de classificação superiores só possam ser alcançadas com um banco de dados balanceado. Na classificação de arritmias cardíacas, a técnica de *data augmentation* com modelos gerativos também tem sido usada para gerar sinais de ECG sintéticos com o objetivo de balancear os bancos de dados [3]. Neste trabalho, é utilizada uma rede neural do tipo VAE [4] para gerar sinais de ECG sintéticos da classe S.

O VAE é formado por duas estruturas: o codificador (C) e o decodificador (D) [4]. Para a aplicação em questão, C e D foram compostos de redes MLP com duas e três camadas, respectivamente. A saída do C foi reparametrizada [5]. Em ambas as redes, consideraram-se uma combinação do erro quadrático médio e da divergência de Kullback-Leibler [6] como função custo, LeakyReLU como função de ativação, normalização de *batch*, 800 épocas, *minibatch* de 25 e passo de adaptação de  $5 \times 10^{-5}$ . No treinamento do VAE, um sinal de ECG real  $x$  com 320 amostras é fornecido ao codificador com 320 neurônios na camada de entrada. Esse sinal é então expandido para a camada seguinte contendo 512 neurônios e em seguida comprimido, produzindo três vetores de tamanho 10:  $z$ ,  $\mu$  e  $\varphi$ . O vetor  $z$  é fornecido ao decodificador que procura “recuperar” o sinal anterior, expandindo esse vetor ao longo das camadas. Primeiramente para uma camada de 512 neurônios e chegando a uma camada de 320 neurônios, obtendo a saída  $y$ . O esquema da arquitetura descrita para essa rede pode ser observada na Figura 1.

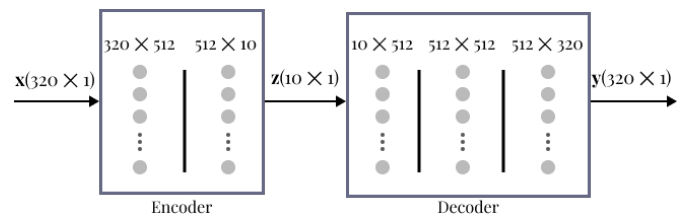


Figura 1: Arquitetura esquemática da rede VAE.

### Resultados

O modelo gerativo foi treinado com batimentos da classe S do MITDB. Como entrada, foram consideradas amostras de batimentos individuais centralizados em seus picos, normalizadas para o intervalo  $[-1, 1]$  e organizadas em vetores de tamanho 320. Após o treinamento, os sinais sintéticos gerados foram comparados com modelos representativos da classe S por meio da métrica *dynamic time warping* (DTW) [7]. Assim, sinais gerados que apresentam DTW acima de um determinado limiar são descartados por serem considerados muito diferentes do “padrão” da classe.

A fim de verificar a variabilidade dos sinais da classe S, um sinal da aleatório dessa classe foi comparado aos

demaís sinais da mesma classe contidos no MITDB. A distribuição das distâncias DTW obtidas a partir desse estudo podem ser observadas na Figura 2, o que indica que um único *template* é insuficiente para definir o “padrão” esperado da classe. Diante disso, definiram-se seis *templates* igualmente espaçados no espectro de distâncias analisado como representantes da classe S. Esses *templates* podem ser observados na Figura 3.

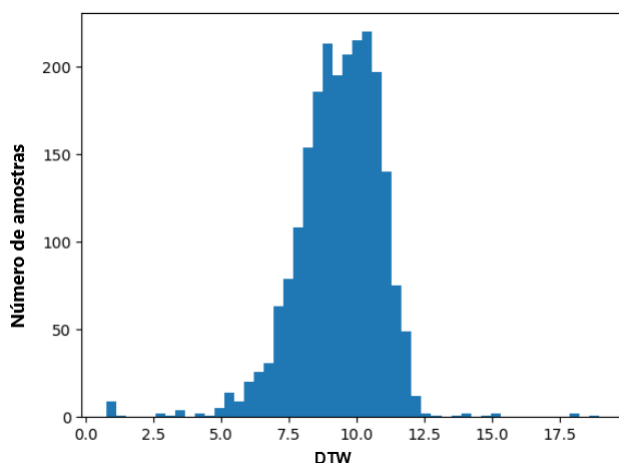


Figura 2: Histograma de DTW ao se comparar um sinal da classe S com os demais sinais dessa classe contidos no MITDB.

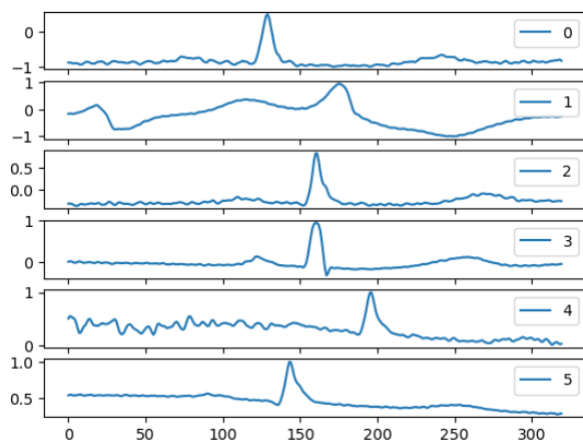


Figura 3: *Templates* selecionados para a classe S.

Em posse do conjunto de *templates* da classe S, composto por 6 elementos, calculou-se a menor distância, denotada por  $s_3$ , entre os batimentos gerados e os elementos do conjunto de *templates*. A partir dessa distância, definiu-se um limiar ( $2s_3$ ). Foram considerados como aceitáveis apenas os sinais com valores de DTW inferiores ao limiar.

Foram gerados aproximadamente  $12 \times 10^3$  sinais contendo três batimentos com 960 amostras, tendo um sinal S centralizado. Esses sinais foram alimentadas à rede MLP de [2] em complemento ao banco de dados MITDB,

e utilizados para retreinar a rede. Os resultados obtidos para as classificações acuratas de cada treinamento, utilizando batimentos reais como conjunto de teste, podem ser observadas na Tabela 1.

Tabela 1: Taxa de acerto na classificação de batimentos.

Banco de dados	N	S	V	F
MITDB	78,0%	11,0%	73,0%	3,2%
MITDB <i>augmented</i>	80,0%	13,0%	73,0%	11,0%

## Conclusões

Como podemos observar, a expansão do banco de dados com sinais sintéticos da classe S causa não só uma melhora na taxa de acerto para a classe S, de 11% para 13%, mas também uma melhora na classificação de outras classes, como na classe N de 78% para 80% e na classe F de 3% para 11%. A melhora considerável na classificação dos sinais referentes a classe F pode ser melhor observada ao se analisar a matriz de confusão da classificação dos sinais. A matriz obtida sem a inclusão dos dados sintéticos indica que o erro de classificação de sinais da classe S como pertencente a classe F se encontrava em torno de 9%. Esse erro passou para apenas 0,5% quando da implementação de *data augmentation*. Desse modo, o aumento do banco de dados por sinais sintéticos apresentou um benefício para a taxa de acerto da rede MLP utilizada. Dando continuidade à pesquisa, pretende-se utilizar um modelo mais robusto para a geração dos sinais e encontrar uma forma de gerar sinais mais variados representativos de uma mesma classe. Além disso, pretende-se gerar sinais sintéticos das classes F e V, a fim de obter-se um banco de dados mais balanceado.

## Referências Bibliográficas

- [1] S. K. Berkaya *et al.*, “A survey on ECG analysis,” *Biomed. Signal Process. and Control*, vol. 43, pp. 216–235, 2018.
- [2] N. Nagata, R. Candido, M. T. M. Silva, “Combinações de redes neurais e discriminantes lineares para classificação de arritmias cardíacas,” in *Anais do Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT)*, Fortaleza, CE, 2021.
- [3] E. Adib, F. Afghah, and J. J. Prevost, “Synthetic ECG signal generation using generative neural networks,” *arXiv*, 2021, disponível em <https://arxiv.org/abs/2112.03268>.
- [4] D. P. Kingma, M. Welling, *An Introduction to Variational Autoencoders*, Now Publishers, Norwell, MA, 2019.
- [5] DINH, Laurent. Reparametrization in deep learning. 2018.
- [6] S. Kullback, R. A. Leibler, “On information and sufficiency,” *Annals of Mathematical Statistics*, vol. 22, p. 79-86, 1951.
- [7] P. Senin, *Dynamic time warping algorithm review*, Technical Report, University of Hawaii, 2008.