Engineered Science
DOI: https://dx.doi.org/10.30919/es1903

Synthetic Generation as Denoiser for P Mitrale Contour Extraction from a Noisy Electrocardiogram Dataset

Krishnadas Bhagwat,[1] M. Supriya,[2] Sreeja Kochuvila[3] and Abhilash Ravikumar[1,*]

Abstract

Akin to software as service, synthetic generation as denoiser addresses explain ability while detecting left atrial enlargement (LAE) in a noisy ECG. In our approach to denoise the ECG, we have taken an unconventional look at synthetic generation. Our objective caters to bring down cost per function by synthetic generator mimicking principles of cardiac atrial depolarization. To this effect, our results demonstrate the effective use of Gaussian function convolution for right and left atrial depolarization vector. Gaussian lobe inherently captures the underlying signal characteristics favoring explainable computation, which served our objective as well. The hump tracer and spread analyzer algorithms showcase simple, effective way to extract parameters from a noisy ECG dataset. This also opens up a portal for patient specific modeling solutions, which have gained much traction. We demonstrate our approach by delineating P contour and detecting the P mitrale (pathology) pattern from the PTB-XL (ECG) dataset. The low compute complexity of our approach O(K), K discrete samples, ensures real time process amenability. Our contribution also brings out trade-offs in generalized denoising solutions such as Empirical mode decomposition (EMD) and neural schemes as compared with tailored solutions such as synthetic generation as denoiser (SGAD).

Keywords: Electrocardiogram; Gaussian fitting; Explain ability; Left atrial enlargement; P mitrale; Empirical mode decomposition.
Received: 22 July 2025; Revised: 03 November 2025; Accepted: 24 November 2025
Article type: Research article.

1. Introduction

Synthetic generation has varied applications akin to software as service (SaaS). Broadly, the applications range from deep fake images, test vector generator, data augmenter, data explorer, fusion imaging[1-4] to mention a few. Our contribution takes an unconventional look at synthetic generation, by formulating it as, synthetic generation as a denoiser (SGAD). We use it to denoise the raw electrocardiogram (ECG) signal and showcase an approach to infer the obscure pathology signature (P Mitrale) amidst static noise. ECG is a time series recording of cardiac electrical activity and the first point of contact in diagnosis in clinical setting.[5,6] ECG serves as a key tool in the diagnosis and evaluation of various cardiovascular malfunctions. As a non-invasive and readily accessible test, ECG is a defacto measurement employed in assessing cardiac health. Clinicians use ECG in diagnosing a wide spectrum of cardiac disorders and as a monitoring tool as well.

The ECG waveform is a representation of the electrical impulses generated during each heartbeat. The ECG provides clues and insights into the heart's rhythm, rate plus overall cardiac functioning. Clinicians often use ECG to identify irregularities such as arrhythmias, atrial fibrillation, heart blockages and structural remodeling such as atrial enlargement.[7] Furthermore, ECG serves as a first line assessment in routine health check-ups and pre-surgical evaluations. Its wide spread use extends beyond hospitals to ambulances and emergency settings, where it aids in rapid diagnosis and immediate intervention during cardiac emergencies.[8,9] Automated ECG processing often calls for pre-processing and filtering (denoising) is one of them.[10,11] The novelty of our contribution extends the application usage model by serving dual purpose of denoise (no explicit filtering) and generate. The scope is further extended by accounting explainable computation into the solution space. Explain ability is a space which has gained significance in

[1]*Nanoelectronics Research Laboratory, Department of Electronics and Communication Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Bangalore, Karnataka, 560035, India*
[2]*Department of Computer Science and Engineering, Amrita School of Computing, Amrita Vishwa Vidyapeetham, Bangalore, Karnataka, 560035, India*

context of AI based approaches (XAI).[12] The prominence in explainable computations is justified in applications dealing with insight gaining, analysis of pathology influence on the signals. Explain ability provides the clinicians with rationale of how a model arrives at a particular diagnosis or prediction.[13] Besides for diagnosis and analysis solutions there are regulatory compliance requirements set up such as General Data Protection Regulation (GDPR), United States Food and Drug Administration (FDA) and Ethics guidelines for trustworthy AI (European Commission) to name a few.

Our contribution specifically has addressed denoising the static noise influenced on the P contour of ECG without any explicit filtering. The rationale behind P contour selection being left atrial enlargement (LAE), a pathology reflects in the P contour of ECG of lead II as a bimodal P wave in most cases. ECG however may not be sufficient for a comprehensive understanding of LAE. While ECG signals can suggest atrial enlargement, echocardiography helps identify the underlying causes of LAE.[14] Common etiologies include hypertension, valvular heart disease, or other conditions that lead to increased atrial pressure or volume. In the current contribution, we have attempted modelling some aspects of explainable computation by involving a delay parameter in convolution of the right left atrial depolarization vectors. LAE is a condition characterized by the enlargement or dilation of the left atrium, one of the four chambers of the heart. This cardiac adaptation is often a response to increased pressure or volume overload, such as in conditions like hypertension or valvular heart disease.[15] LAE can have discernible effects on the morphology of the P wave in an ECG. The P wave represents atrial depolarization, and changes in the left atrium's size can alter the P wave's characteristics.[16] In the presence of LAE, the P wave may exhibit prolonged duration or changes in shape, commonly resulting in a bifid or notched appearance, and referred in medical jargon as "P mitrale". Denoising challenges arise when attempting to extract a clear P contour impacted by LAE from an ECG amidst static noise influence. The enlarged atrium introduces additional complexity to the signal, and distinguishing between true atrial depolarization and noise becomes crucial for accurate interpretation, may call for ensemble methods for confirmation.

In biomedical applications, the real ECG set composed of raw signals is besieged with noise[17,18] and artifacts. The high frequency noise in ECG has varied sources such as ambient EM (electromagnetic) signals via cables and instrumentation amplifiers. Power line interference and its harmonics are also a common observation in real ECG signals. The common artifacts observed are the low frequency tones due to respiratory movements or motion due to the subject. Base line drift a low frequency tone is common in ECG caused by temperature variations and bias in acquisition equipment. Overlap of signal and noise in case of ECG is not uncommon.

Defacto acceptance range for diagnostic ECG is 0.05 to 40 or 100 Hz.[19,20] The low frequency cutoff is dictated by the ST segments of ECG. Baseline wander also overlaps this range and for heart rate variability analysis, addressing this becomes paramount. Notch filtering is a common solution to the power line interference (50/60 Hz). But these filters may also contribute to noise (usually at high frequencies). The real signal is also influenced by static noise whose spectral components are spread out the entire range and is of random nature. The EP study (cardiac electrophysiology) has environment which is dominated by various sources of noise.[20] The EM consists of static and dynamic electric fields usually capacitively coupled to object under consideration.

De-noising raw ECG signals poses unique challenges, particularly when considering the influence of pathological conditions.[14] Pathological changes in the heart can manifest as irregularities in the ECG signal, including abnormal rhythms and fluctuations in amplitude. The challenges in denoising become pronounced when attempting to extract relevant information from ECG recordings affected by pathologies such as arrhythmias or myocardial infarction. The characteristic features introduced by these conditions may overlap with noise, making it difficult to distinguish pathological signals from undesirable artifacts. Conventional denoising techniques face the challenge of preserving diagnostically relevant in formation while effectively eliminating noise.[14] Advanced signal processing approaches, tailored to address the specific characteristics of pathological ECGs, are crucial for enhancing the accuracy of signal interpretation and supporting more precise clinical diagnostics. The integration of machine learning algorithms, capable of learning and adapting to the intricate patterns introduced by cardiac pathologies, represents a promising avenue for overcoming the challenges inherent in denoising raw ECG signals affected by various heart-related disorders. However, these come at a high computation cost and model development turnaround time.

Machine learning solutions in ECG denoising encounter several challenges inherent to the complexity of physiological signals and the variability in noise patterns. In the first place, acquiring labeled datasets for training deep learning models

[3]*Department of Electronics and Communication Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Bangalore, Karnataka, 560035, India*

[*]*Email:* r_abhilash@blr.amrita.edu (Abhilash Ravikumar)

poses a significant challenge. The need for large, diverse, and accurately annotated datasets is crucial to ensuring the robustness and generalization of machine learning models. Access to such datasets, especially those encompassing various cardiac conditions and noise sources, can be resource-intensive plus require extensive collaboration between medical professionals and data scientists. Additionally, ensuring the ethical and privacy-compliant use of sensitive patient data adds another layer of complexity to dataset acquisition. The interpretability of machine learning models is a critical concern in the context of ECG denoising, which is another significant hurdle. Deep neural networks, particularly those with intricate architectures like convolutional neural networks (CNNs) or recurrent neural networks (RNNs), often function as "black boxes", making it challenging to understand how they arrive at specific denoising decisions. Interpretability plays a key role in medical applications where clinicians need to trust and comprehend the underlying reasoning of the algorithm. Addressing these challenges requires ongoing research into developing efficient and interpretable machine learning models, as well as establishing robust frameworks for data sharing and collaboration within the medical community. Added to these conditions are ethical considerations and regulatory frameworks for accounting trust to ensure the responsible deployment of machine learning solutions in ECG denoising and other healthcare applications.

Literature scan reveals several approaches using time domain, time-frequency combination (wavelet based), digital filtering, Kalman filtering, fractional calculus, empirical mode decomposition, and Fourier transform, to denoise ECG signals. When selecting an appropriate ECG signal denoising technique, essential considerations include computational complexity, parameter selection, hardware software partitioning if real time. Wavelet transform-based techniques are favored for ECG denoising.[21] The wavelet schemes typically use filter bank.

Wavelet Transform: The wavelet transform has a longer history and was developed over several decades. The concept of wavelets was introduced in the late 1970s and early 1980s by mathematicians Jean Morlet and Alex Grossmann.[14] The continuous wavelet transform (CWT) and discrete wavelet transform (DWT) were subsequently developed. Wavelet based methods have been widely applied in signal processing, including image processing, audio signal analysis, and biomedical signal processing.[22,23]

Empirical Mode Decomposition (EMD): EMD is a relatively newer technique compared to wavelet transform. It was introduced by Huang *et al*. in the late 1990, specifically designed for analyzing non-linear and non-stationary signals.[24] It decomposes a signal into intrinsic mode functions (IMFs), allowing adaptive and data-driven analysis of the signal's components.[25] EMD gained attention for its ability to handle signals with rapidly changing characteristics and has been applied in various fields, including biomedical signal processing.[25,26] EMD also has percolated into real time processing hardware solutions in ECG.[27] While EMD is quite effective in denoising, it is a generic decomposition which is domain agnostic and one has to account for convergence conditions when it comes to building the IMFs.[28,29,30] The other challenges which EMD poses are that of mode mixing.[31] Much of the literature scan provides insights in addressing these challenges as well specifically the enhanced empirical mode decomposition (EEMD).[32,28]

State of art in denoising mechanisms: The state of the art in ECG denoising solutions reflect a dynamic landscape marked by continuous advancements in signal processing, machine learning, and computational techniques. Multiple variants of machine learning models provide a good treatise as in.[33-36] Wavelet-based methods remain prominent, leveraging multi-resolution analysis for noise reduction while preserving important signal features. Additionally, sophisticated thresholding techniques, including adaptive and data-driven approaches, contribute to improved denoising performance. Machine learning algorithms, particularly deep learning models like CNNs[37,38] and RNNs, have gained traction for their ability to automatically learn complex patterns from ECG data, offering adaptive and robust denoising capabilities. Hybrid approaches that integrate wavelet transform with machine learning strategies aim to harness the strengths of both methodologies. Real-time applications[11,21] such as wearable devices for continuous monitoring, are becoming more feasible with optimized algorithms and hardware capabilities. The growing emphasis on explain ability and interpretability in healthcare algorithms is influencing the development of denoising solutions that not only provide high accuracy but also solve the black box problem.

Gaps in denoising solutions: Most of the denoising solutions known so far (to the best of our knowledge) are not tailored for specific use cases, but the wavelet differs in this aspect. It does provide a framework to design the suitable basis functions accounting the signal characteristics. Specific use case in the context of the current solution is to make the solution pathology aware. In a way, pathology influences the characteristics of the signal. By making the solution pathology aware, the underlying causation can contribute to better comprehension and hence explainable computation. Patient specific models have gained much traction[39,40] and our solution caters seamless portal to such schemes. primary

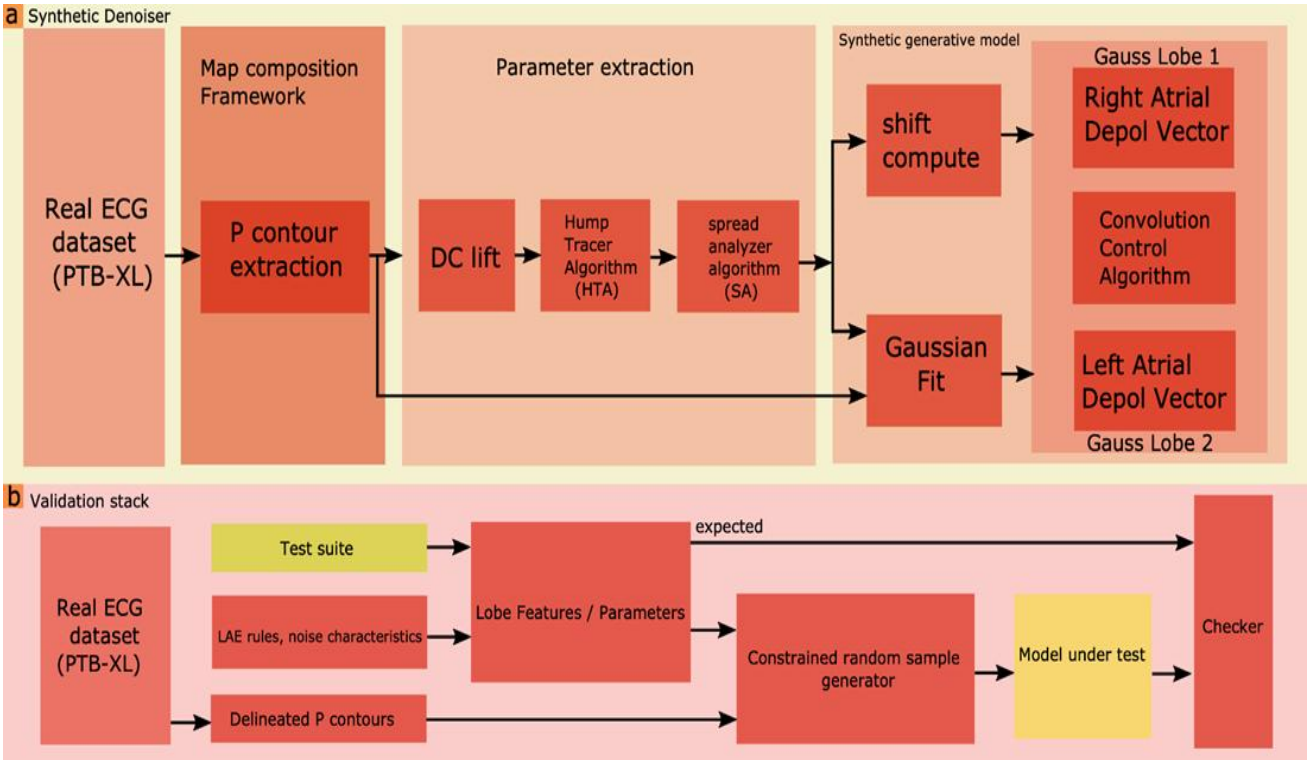

Fig. 1: (a) Functional view synthetic denoiser (b) Validation scheme.

challenge in wavelet, lies in the careful selection of the wavelet and its parameters. The effectiveness of a wavelet-based denoising approach depends on the choice of appropriate wavelet function. The configuration parameters for tuning the wavelet for ECG characteristic under consideration also add to the challenges. Different wavelets possess distinct properties, and selecting an ill-suited one can cause sub optimal denoising results. Additionally, determining the optimal threshold for soft or hard thresholding, a common step in wavelet-based denoising complicates the problem space. The threshold needs to strike a delicate balance between noise removal and preservation of essential signal features, requiring a nuanced understanding of the signal's nature. Another challenge involves the adaptability of wavelet-based methods to the inherent variability of ECG signals. Ensuring that a wavelet-based denoising scheme can effectively adapt to non-stationary property of ECG is crucial for its success. Achieving this adaptability requires a robust choice of wavelet. Balancing these challenges with the need for computational efficiency, especially in real-time applications, adds an additional layer of complexity to the implementation of wavelet-based denoising solutions. Addressing these challenges necessitates understanding of both the mathematical intricacies of wavelet transform and the specific features of ECG signal. This calls for a multidisciplinary approach involving signal processing expertise, clinical insights, and computational proficiency. Traditional filtering approach is a direct approach, it separates the noise from the signal but is prone to change the characteristics of the signal too.[41] And in general, filtering falls in pre-processing of signals. Our objective aimed to leverage the noise in itself, without the need to pre-process. The rationale being low computational burden, a heuristic route was used. To facilitate such computations, we leveraged domain insights of ECG and the underlying causation principles. This is one of the stark contrasts in our approach as compared with most other solutions. The other objective we have met to a significant extent is that of explainable computations. Our approach was augmenting machine learning techniques in ECG processing, by building a synthetic model which is driven by the parameters extracted from the ECG records. The synthetic model is built using the atrial depolarization dynamics by a simple yet effective Gaussian function (Fig. 1). Our framework for synthetic generative model is completely discrete algorithmic framework and validation aspects were explored by use of constrained random sample generation mechanisms.[42,43] A noteworthy mention is with regards to patient specific models[44,45] which has gained traction, will find the approach, of much use when plugged in the processing pipeline. A limitation of the current model however is a focus only on P mitrale case. However, this can be extended as an approach, at the cost of understanding relevant cardiac insights and plugging it in the extensible framework. Methodology wise, we have made use of Python framework, where we

leveraged P contour delineation from our previous work.[46] For SGAD contribution we developed DC lift, hump tracer and spread analyzer algorithms and the Gaussian convolution scheme as a synthetic generator of P contour. The complete solution space is as covered in Fig. 1. We use PTB-XL dataset[47] annotated LAE record to extract and decipher P Mitrale pathology.

Specifically with respect to denoising ECG, several works have brought out architecture trade-offs to address signal processing requirements. Notable mentions include fusion methods, semi-supervised methods using framework approach, combination of filter and neural processing, cycle GAN in combination with attention mechanisms and GAN with convolution mechanisms. In the enhanced heart sound anomaly detection, Peipei Zeng *et al*[48] uses the framework approach named WCOS. This is a semi-supervised framework, which has combined wavelet, convolution autoencoder and support vector machine to enhance detection. A notable attribute is the ability to identify the anomaly amidst the noise, with noise reduction network. The wavelet reconstruction provides significant noise suppression capability, by exposing the detail coefficient that mainly contain high frequency noise components (Peipei Zen *et al*). The subsequent processing stage uses this processed signal which is clean. One of the limitations which the work has brought out is computational complexity during the training process. The study demonstrates that WCOS significantly outperforms traditional methods, reducing area under curve (AUC) standard deviation by up to 54.1% compared to baseline models. The approach is particularly useful for early congenital heart disease detection, ensuring higher sensitivity in identifying abnormal heart sounds. In another notable work, though the paper is non-medical signal pattern, it brings out an interesting question which pursues, to find out a robust pattern recognition method in presence of noise.[49] The study has compared fifteen different methods and found that Ridgelet plus fast Fourier transform descriptor is most robust to additive Gaussian white noise for an aircraft and Chinese character dataset. The study also computes computational time for the different methods. In Myocardial perfusion Single Photon Emission Computed Tomography (SPECT (MPS)) images, Antti Sohlberg *et al*[50] have provided denoising methods to overcome the artefacts caused by low-count statistics. They have compared several denoising methods such as CNN, RES, UNET and cGAN. The comparison was done in two modes, with and without DL based denoising. Inclusive of denoising as a use case, the medical image analysis using deep learning algorithms Mengfang Li *et al* reviews several noteworthy[51] publications. It brings out real time analysis of intricate datasets along with other notable medical image signal challenges. Categorizing of state-of-the art DL techniques, such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Generative Adversarial Networks (GANs), Long Short-term Memory (LSTM) models, and hybrid models are explored. Amar *et al* brings out a study on medical image denoising using filters and neural network,[52] in which they also use a Kaggle CT scan dataset on an autoencoder as denoising. The have observed the metrics peak signal to noise ratio (PSNR), structural similarity index (SSIM) of two different images. Rajesh Patil *et al*[53] has done a comparative study and analysis using wavelets and specifically on gaussian noise influenced multimodal images (MRI). They have used MATLAB 2020B set up to process the noise signal coefficients and denoise the gaussian noise impacted MRI, along with quantifying the noise variance. Quantitative accuracy[54] for deep learning based denoising in oncological PET is studied and investigated by Lu W *et al*. The specific emphasis is on the challenge of noise levels due to low injection dosage, causing low SNR. Amidst this backdrop, to improve the diagnostic accuracy involving deep learning method, by a U net architecture was employed.

Zhang P *et al*[55] provides a novel fusion method of denoising ECG end to end, using (ECA-net) efficient channel attention and cycle GAN. They introduce a new loss function to extract global and local features. Inferring motion artefacts and other noise in a wearable ECG acquisition mode is a feasible way to proceed for processing as per Zhang P *et al*. A novel ECG denoising[56] approach engaging GAN with goal of circumventing traditional thresholding makes use of a denoising frame work comprising of CNN architecture by Singh P *et al*. They have made use of MIT-BIH database for all qualitative and quantitative analysis and performed end to end GAN training showcasing improved noise performance. Holgado *et al*[57] brings out characterization noise in their work on ECG with emphasis on long term ECG monitoring. They bring out an important distinction between qualitative and quantitative score and the notion of clinical noise which is differentiated by noise level influence. Machine learning techniques are showcased to categorize noise severity using a database annotated as per clinical noise taxonomy as gold standard. The models make use of signal quality indexes from time, frequency, and statistical domains to distinguish valid and invalid ECG segments. To prevent overfitting, the methodology ensures balanced classes, patient separation, and rotation in the test set. As per Holgado *et al*, their proposal showcases strong classification performance, with recall, precision, and F1 scores using a single-layer perceptron. Tutuko *et al*,[58] have come up with single lead ECG processing

Algorithm 1: Pseudo code for MAX Update Algorithm (Hump tracer).

```
Require:
1:        ▷ Initialization
2:        MAX = -10 ▷ initialize the MAX as the lower than available in database, the objective is to trace the positive humps
3:        mxbufindx = [] ▷ max buffer index: buffer to track the index of humps of P contour, used to push entries into it
4:        pbufin = delineated_psamples (PTBXL)
5:        mxbuftrace=[] ▷ state encodes: [MAXUPDATE, CHECK]
6:        state = CHECK; smpl_indx = 0 ▷ PWIDTH: number of samples representing P contour width.
7:      ▷ smpl_indx: index of sample in the buffer pbufin
Ensure:
7:        while (smpl_indx < PWIDTH)
8:          smpl = pbufin[smpl_indx]
9:          state MAXUPDATE:
10:             If (smpl >= MAX)
11:                 MAX = smpl
12:                 mxbufindx.push (index of smpl)
13:                 mxbuftrace.push(MAX)
14:                 state = updateState (MAX)
15:             else
16:                 state = updateState(CHECK)
17:         state CHECK:
18:             If (smpl < MAX)
19:                 mxbufindx.push (MAX)
20:                 mxbuftrace.push (MAX)
21:                 state = updateState (CHECK)
22:             else
23:                 state = updateState (MAXUPDATE)
24:         smpl_indx++
25:       endwhile
```

for cardiac pathology detection. The notable mentions are on limitations in identification of suitable wavelet basis function and threshold, a time-consuming process. The denoising autoencoder proposed, is an end-to-end learning using convolutional bidirectional long short term memory for delineation of PQRST wave and isoelectric lines. Unsupervised learning scheme is used in the autoencoder for addressing denoising and a noteworthy mention is of 98.59% accuracy on average.

The organization of the paper here on starts withMethods which depicts algorithm details in the processing pipe. Fig. 1 captures the complete solution space developed and we describe the algorithms in sequence in relation to the figure. We then cover results, discussions inclusive of EMD and Neural scheme of denoising and finally the conclusion addresses the insights gained, application avenues and limitations.

2. P mitrale denoiser method

2.1 Dataset used and scheme overview

In this section we cover the functional components of synthetic generation architecture. Fig. 1 depicts the complete solution space developed for denoising of the P mitrale contour from a noisy raw PTB-XL[47] dataset. The entire framework operates in a discrete algorithmic setup. The sample index resolution chosen is 2 ms to match the PTB-XL record500, 500 Hz sampling frequency. We have used Python framework version 3.11.5. PTB-XL dataset is a discrete raw ECG record of 10 seconds length. The records are annotated and covers broad range of diagnostic classes. Standard 12 lead signals are present in the record at both 500 Hz resolution and 100 Hz resolution. The 500 Hz record consists of signal with 4096 samples, spawning the 10 second duration. We make use of the PTB-XL record (01000/01447_hr) which is annotated as LAE. The dataset is impacted by static noise (Fig. 2). For P contour extraction (Fig. 1)we make use of our developed algorithm library in map composition framework.[46] We briefly revisit the P contour extraction method for reference in the current paragraph. The R peak extraction algorithm extracts the R peaks of the ECG. Firstly, R peak identification is accomplished and later the samples representing the R-R interval are extracted. The entire framework is a discrete

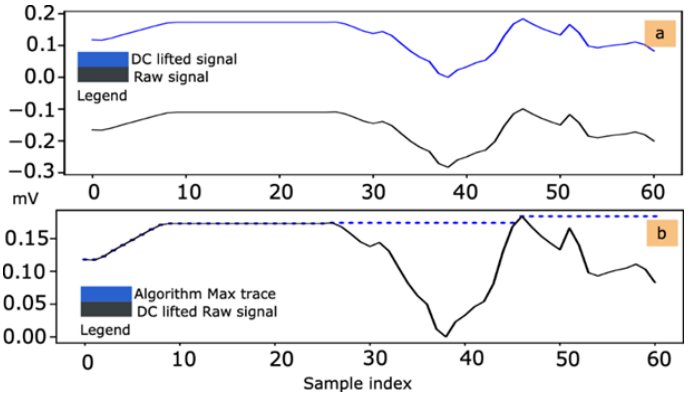

Fig. 2: (a) Raw signal and DC lifted signal (b) Hump max trace.

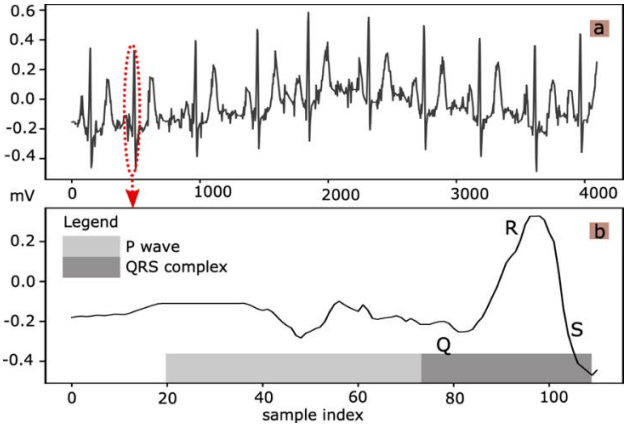

Fig. 3: (a) Raw P contour signal (b) Zoom of second cycle in (a).

Algorithm 2: Spread analyzer: Tracking the hump index buffer, locating Peak1 (lobe1).

Require:

1: ▷ Initialization
2: i = 0 ▷ index variable to trace the index buffer
3: prev= trk_indxbuf [0]
4: ▷ Max peak2 (Mitrale lobe2) resides at this location (Fig. 5)
5: MAXINDEX = length (trk_indxbuf) ▷ number of humps tracked
6: indx_spread = [] ▷ buffer to track the distance between humps
7: indx_value = [] ▷ buffer to store values
8: threshold = 5 ▷ programmable threshold, spread metrics
9: Peaks.push(prev) ▷ Peak2, the maximum Mitrale peak the second one, Fig. 5

Ensure:

13: while (i < MAXINDEX)
14: current = trk_indxbuf[i]
15: spread= abs (current − prev)
16: indx_spread.push (spread)
17: indx_value.push (current)
18: if (spread >= threshold)
19: check_lae_rules (i, spread, threshold)
20: Peaks.push(current)
21: prev= current
22: i = i + 1
23: endwhile
24: Pk1index = Peaks [1]

Algorithm 3: Pseudo code: generation of Gauss atrial depolarization lobes.

Require:

1: ▷ Initialization

2: stddev = 9 ▷ standard deviation

3: peak =0.15 ▷ Gauss peak (mV)

4: lobe width = 40 ▷ width: number of samples

5: lobeshift=15 ▷ samples by which lobe2 is shifted from lobe1

6: pwidth=70 ▷ samples representing width of P, the maximum

7: params = [stddev, peak, lobe width, pwidth]

8: ▷width represents number of samples at 2ms step

9: ▷ 500 Hz sampling frequency

Ensure:

13: lobe1 = gengauss(params) ▷ lobe1: right atrial depol vector

14: lobe1_zapnd = zero_append (lobe1, right append, pwidth)

15: lobe2 = gengauss(params) ▷ lobe1: left atrial depol vector

16: lobe2_zapnd = zero_append (lobe1, right append, pwidth)

17: pwave = lobe1_zapnd + lobe2_zapnd

18: ▷ pwave: composing P waves with right and left depol vectors

Algorithm 4: Pseudo code: an example of discretized EMD algorithm.

Require:

1: residue = signal

2: i = 1

3: nIMFs = 10 ▷ IMF extraction iteration index

▷ limiting the number of IMF to be extracted

4: IMF = []

5: NS = 10

6: UEnlp = [] ▷ buffer to store extracted IMFs

▷ fixed stopping condition, an example

▷ Upper and lower envelope buffer allocation

7: LEnlp = []

8: parameters = [tmax, tmin, xmax, xmin]

9: ▷ sifting process extracted parameters ▷ max time, min time, max amplitude signal, min amplitude signal

10:

Ensure:

11: while (i < nIMFs)

12: IMF[i] = residue

13: sift = 1

14: while(sift < NS)

15: parameters = extractExtrema(signal)

16: UEnlp, LEnlp= genEnvelope(params)

17: ▷ generate upper and lower envelopes

18: IMF[i] = IMF[i] − Mean (UEnlp, LEnlp)

19: sift + +

20: endwhile

21: residue = residue − IMF[I]

22: i + +

23: endwhile

algorithmic one, inclusive of current denoiser method. P extraction is a back trace algorithm which works on the samples fed from the R-R interval. From the Q fiducial point of the QRS complex, the algorithm traces backwards until several conditions for P onset are tested to extract the P contour. To sum up the algorithms are the Algorithm 1, the hump tracer; Algorithm 2, spread analyzer; Algorithm 3, the gaussian lobe generator and the Algorithm 4 which is a discretized EMD setup for comparison of denoise with the current scheme.

2.2 Noise invariant hump tracer algorithm

The hump tracer algorithm (Algorithm 1) traces the peaks of the delineated P samples from the PTBXL dataset[47] as a snippet depicted in Fig. 3). The Fig. 2 is the output of hump tracer algorithm snippet. The tracer algorithm feeds into index spread analyzer algorithm which is the next stage of the piped processing. The static noise influenced raw ECG signal has appearance of humps. The objective is to track the progression of peaks as the signal is traversed from P contour onset to end of P contour. The P mitrale, P pulmonale in a clean P signal will be composed of two peaks. The first peak comprising the right atrial depolarization vector and the second peak comprising the left atrial depolarization vector. In a non-pathology P contour, there will be only one peak. The algorithm refers to Peak1 and Peak2 in this context. The hump tracer is a simple effective tracker of the progression of peaks encountered. In signal influenced by static noise, there are possibilities of multiple peaks.

2.3 Index spread analyzer algorithm using LAE conditional checks

The spread analyzer algorithm (Algorithm 2)checks for the distance between the peaks encountered. For the P Mitrale as well as P pulmonale, the criteria accepted is that of separation between the peaks to be at least 30 mS.[59] The spread analyzer first builds up the spread buffer by computing the difference between the current and previous index in theindex buffer (named trkindxbuf). This buffer is the input to the spread analyzer algorithm. The hump tracer output loadsthe index buffer. The spread analyzer checks for the first encountered spread distance and identifies the Peak1 of theright atrial depolarization vector. The left atrial enlargement rules consist of the distance between peaks, and is extensible for other rules as well.

2.4 Synthetic Gaussian atrial lobe generation

The P wave generation is based on explainable atrial depolarization lobes. There are two lobes, right atrial lobe and left atrial lobe. The lobes are modelled using simplistic Gauss function. The left atrial depolarization is initiated by the right atrial depolarization vector. The final P wave is a composite and is formed by convolution of right and left lobes. This model serves the explainable P wave generation dynamics as depicted in (Fig. 1(a) (synthetic generative model) and results as in Fig. 4.

$$gaussian\_data = gausspeak * e^{-B*x^2} \quad (1)$$

$$B = 1/(stddev)^2 \quad (2)$$

In the Eq. (1,2), stddev represents standard deviation.[60] The Gauss peak and standard deviation are obtained from a normal healthy P wave inference as mentioned in the Table 1.[19,59] The model is first initialized with parameters derived from the knowledge base,[59] Table 2; also the Algorithm 3 makes use of this knowledge base. The left atrial lobe when generated is shifted from the right atrial lobe explained by the vector being initiated by the right depolarization vector. The shift is impacted by various pathology as well and specifically left atrial enlargement. The composite P wave is then obtained as convolution of left and right lobes. The algorithm being discrete, we make use of zero padding extension to facilitate ease of addition to the tail end of lobe 1. The lobe 2 is also zero padded to the start of the lobe, since lobe 2 initiated by right lobe, accounted by the shift parameter.

Table 1: Normal P wave attributes.Reproduced from.[19,59]

Feature	Value	Description
P duration	100 ms	From onset of P till return to baseline.
PQ segment	40 ms	From end of P to the Q fiducial point
P amplitude	0.2 mV	Peak of P contour
LAE P duration	>= 120 ms	Bifid P wave lead II Peak1, Peak2 separated by at least 30 ms

Table 2: Synthetic Gaussian generative parameters.

Attributes	Description
Sample index	Integer index of sample with resolution of 2 ms per step index
Gauss lobe width	40 samples
stddev	Standard deviation, 9 (Eq. 1)
gausspeak	0.15 mV default
padding	Zero padded samples to account for PQ segment (40 ms ([59])
lobeshift	Gauss lobe 2 shift from lobe 1, shift is variable from 10 samples onwards up to a max 20 samples (Fig. 4)

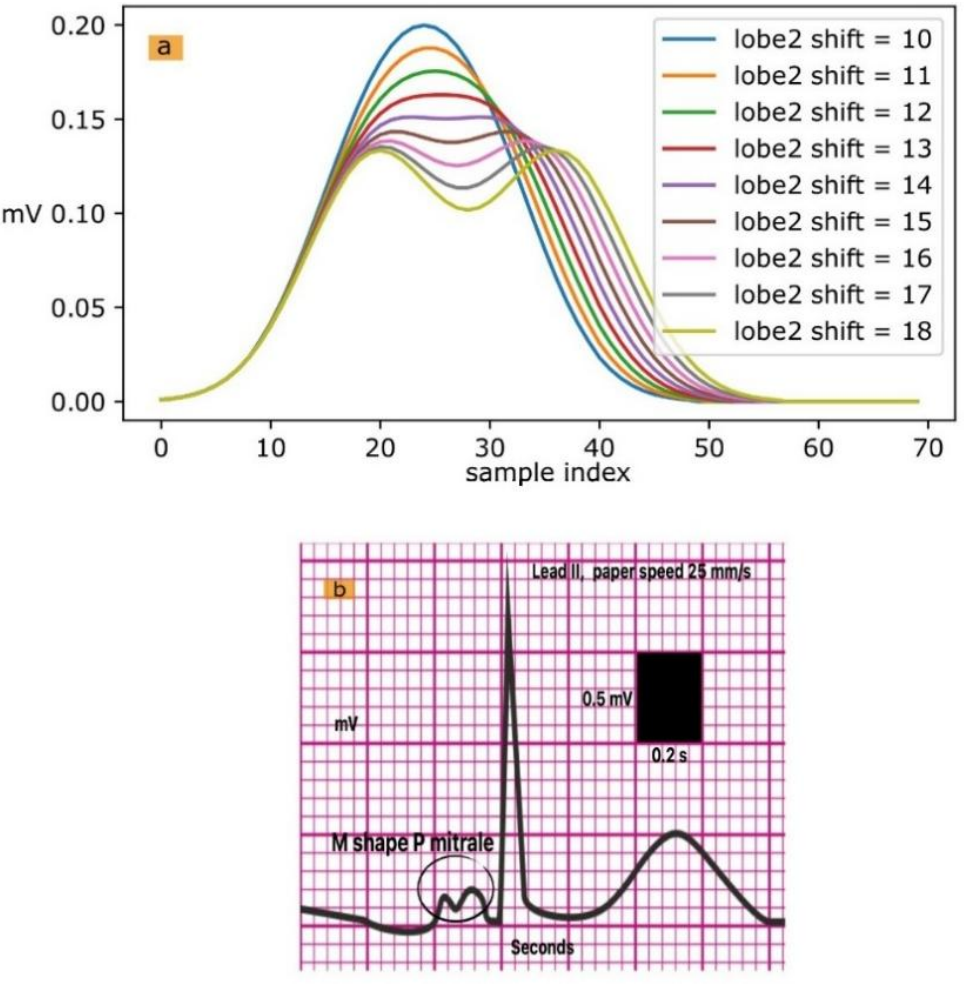

Fig. 4: (a) Explainable P mitrale model, (b) P mitrale pattern from a real ECG paper print (Adapted image, courtesy: Koley TK. Atrial enlargement. Reproduced from.[64]

2.5 Validation aspects

Entire framework being a discrete algorithm based the validation has used random vector generation with constraints as per the specifications for the corresponding algorithms (Fig. 1 b). The hump tracer has to identify the peaks in the input signals and also trace the signal curve while curve has rising trend. To validate this, we generated synthetic input vectors. These synthetic input vectors are to test the algorithm. The basic tests cover the formal signal shape requirements, with input vectors upfront designed to have some peaks and the test run compares the expected peaks with algorithm. The test suite is designed to create peaks and control the lobe separation and then use it as constraint to random sample generation. This covers possible scenarios, in a directed test vector as well. It

feeds the expected vectors to the checker which compares the model output vector with it. The test suite mechanism provides a formal mechanism to capture different scenarios for coverage and validation. The expectation of the synthetic denoiser solution is that it expects a P contour rightly delineated before it is fed to its pipeline. This stems from the P mitrale, pulmonale pathology which states that the P duration is greater than or equal to 120 ms. The validation stack captures the LAE rules and is flexible enough to adapt to other requirements as well.

2.6 Comparing setup for SGAD with EMD and autoencoder

To compare the performance computation wise, separate setups for EMD and autoencoder as denoiser were created. The EMD was used from Python library, PyEMD. The input signal to it was the same signal as used by SGAD, the P delineated sample from the PTB XL. The results of EMD setup are shown in the Fig. 6. Autoencoder was used from Tensor flow library Python and we built the following setup to train the encoder (Fig. 7). We used the Mitrale signal influenced with noise, to generate a clean signal serving as a template reference to the autoencoder. The reference was then subjected to Gaussian noise with randomization to train the autoencoder decoder. Post training, we subjected the trained autoencoder to predict the mitrale signal influenced with noise, basically as a denoiser, the output of which is in the figure. We also run the same pattern to denoise. We then measured the computation time taken by each algorithm. For EMD and autoencoder though, it is the time taken to filter the signal. Whereas for SGAD, we do not explicitly filter, which is a novel approach. For SGAD, the time taken to extract parameters and the final gaussian fitting has been accounted and measured. Our scheme specifically being inference amidst noise without explicit filtering, the mean square error does not justify, as there is no filtering per say, hence we chose the focus on computation time, as a metric.

3. Results and discussion

The results obtained by synthetic denoiser are depicted in the Fig. 2, 4 and 5, whereas the comparison with EMD and autoencoder as denoiser setup are in Fig. 6-8 respectively. The P mitrale (the M shape) is a bifid appearance and two prominent peaks are visible as can be seen in the real ECG example (Fig. 4 b P mitrale pattern from a real ECG paper print (Adapted image, courtesy: Koley TK. Atrial enlargement. In: Rapid Review of ECG. Singapore: Springer; 2024, 138, Figure 11-4[64] which is a paper print out of ECG lead II. These represent outcome of (Fig. 1a) the piped processing computational elements. P contour delineated samples are fed to the denoising scheme. The samples are static noise influenced hence do not have smooth contours. De-noising forms part of pre-processing of ECG signals. Our scheme differs from the traditional filtering methods, in that it is designed as static noise aware. The static noise is random and the spectral content of this noise encompasses a wide range. The raw signal (Fig. 3) from the PTB-XL dataset is a 10 second record and is annotated by cardiologist as left atrial enlargement pathology. Our objective hence was to leverage the underlying principles which can decipher the pathology pattern even amidst the noise. The LAE pathology confirmation has to be quantified with other clinical context too such as lead V1 P termination shape. Besides an ensemble of contours, multiple factors need to be accounted before confirmation of the enlargement. Mitrale shape is one of the strong indicators and our solution provides a minimal compute burden and an approach to inference through the physical principles. The DC lift scheme rationale is a constraint influenced by the subsequent synthetic explainable fitting model primarily for simplifying the computations. The output of this computation is as in Fig. 3, a mere shifting (y axis shift) of the signal by a fixed margin decided by the most negative sample in the P contour buffer, this pushes the entire P contour to have positive samples. The challenge in P contour delineation amidst static noise arises due to baseline impacted by the noise, a possibility. To circumvent this, the delineation accounts that a general normal PQ segment is around 40 ms.[59] We rely for P delineation the Q fiducial point in the ECG lead II which is the one which ends with most negative of the QRS complex. The QRS identification in the map composition relies on first identifying the R peaks, which are the most prominent in any ECG, the R-R samples are then fed to P delineation. Whereas traditional filtering is agnostic to the domain insights, the current scheme is developed to be noise aware and with explain ability in focus. The Fig. 3 is a raw ECG signal and we show the zoom of P contour of it, which is to be decoded as P mitrale.

DC lift and hump tracer are the pre-processing stages for the synthetic explainable gaussian fit model. We have taken a different look at the fitting as well, by modeling the rationale of atrial depolarization. We call this method of fitting as explainable fitting, where we rely on the explain ability aspect, the reasons of P morphology to decipher the meaning amidst the static noise. The hump tracer is akin to the standard sample and hold circuit, tracing the humps encountered in the signal as depicted in Fig. 3(b). The spread analyzer, analyzes the spread of the index of samples as recorded by the hump tracer algorithm. Based on the morphology and static noise, there can be multiple possibilities, however, the analyzer objective is to find the LAE conditions, the Mitrale shape, which is to locate the lobe peaks, separated by at least a margin of 30 ms. The analyzer is a back trace algorithm which starts with peak 2 and proceeds to locate the peak 1. There are cases where the

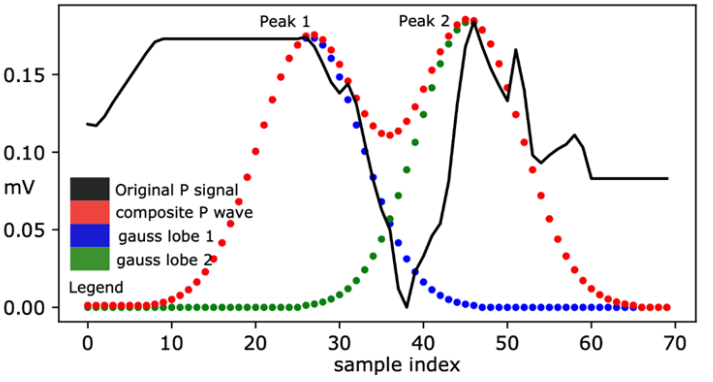

Fig. 5: Gaussian explainable fitting for P Mitrale as extracted from PTB-XL.

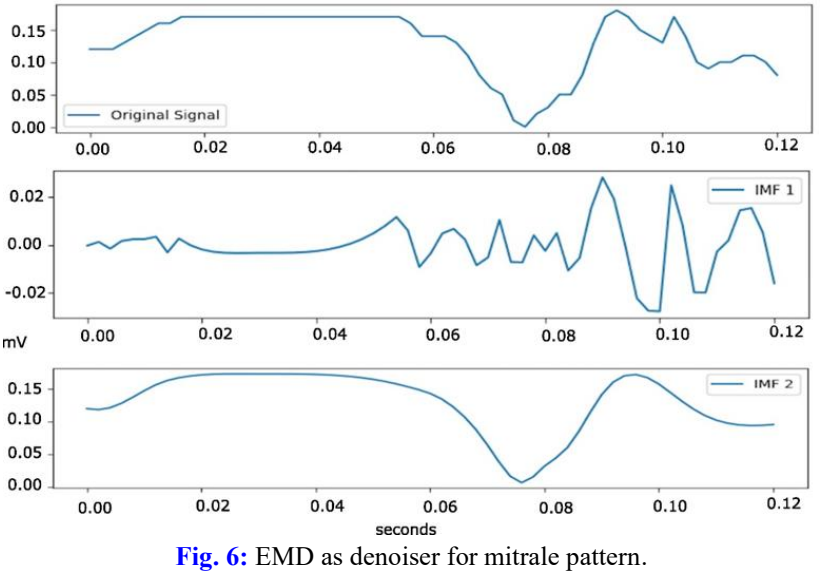

Fig. 6: EMD as denoiser for mitrale pattern.

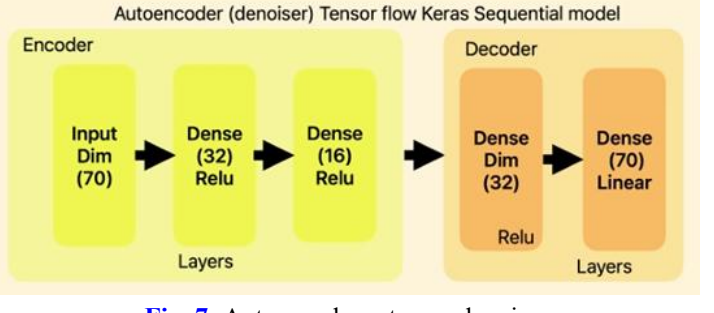

Fig. 7: Autoencoder setup as denoiser.

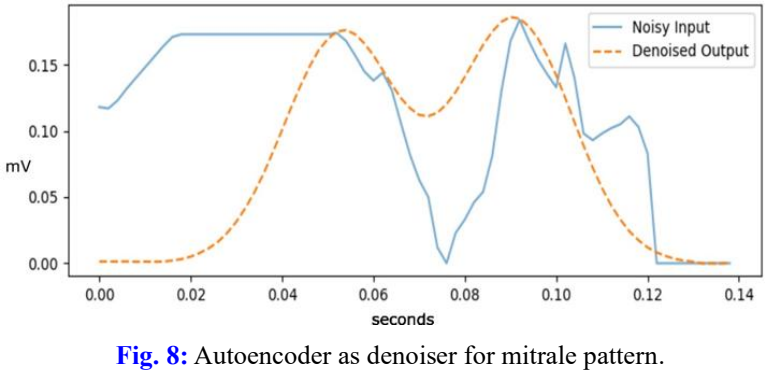

Fig. 8: Autoencoder as denoiser for mitrale pattern.

maximum spread of the indices due to noise can mislead, however it is the first spread which satisfies the condition which is qualified to belong to peak 1. The index buffer only tracks the peaks, so for the lows of the signal, the buffer does not store the indices. This facilitates the spread analyzer to compute only the required spreads. The analyzer computes the spread for all the indices from peak 1 onwards tracing it till the onset of the P contour. One of the advantages of explainable fitting is that it relies on the physical principles behind the P morphology. Thus, it is able to leverage the static noise and there is no need to filter it out first. It analyzes the underlying pattern amidst the noise samples which saves the computational burden of filtering out. Importantly, explainable computation, is a mandate in any automatic interpretation healthcare solutions and thus favors our solution. Our approach to explain ability is a direct approach, where we model the underlying principles which gives rise to specific morphology. We started with P morphology as a Gaussian wave for its strong resemblance with the P contour. The other aspect for selection of Gaussian function was that most underlying physical signals, biomedical signals are a good gaussian resemblance. Gaussian function also has excellent properties which fit the atrial depolarization dynamics. The P mitrale shape is well explained by the two gaussian vectors shifted apart post a shift limit. Computation wise, when we consider discrete structures, the processing calls for padding, shifting, addition which are not just low cost but debug friendly as well. This saves development cost of the solution. The synthetic generative model (Fig. 1(a)), is driven by parameters extracted from real ECG dataset and it facilitates patient specific model development. Parameters from atrial cellular depolarization dynamics[61] can also be coupled to the synthetic model which serves to extend the functionality of the solution space.

The 2020 European society of cardiology (ESC) Guidelines and 2023 American heart association, American college of cardiology (AHA/ACC) for the Diagnosis and Management of Atrial Fibrillation discussed LAE in detail but it does not mention on the P wave duration. As of May 2025, there hasn't been a newer ESC guideline specifically updating the management of LAE. We have referenced latest clinical documentation (and updated the current document) and articles for bringing the context, numerical and the significance from Doudi *et al*,[62] (National Center for Biotechnology Information, US (NCBI)) and Chen *et al*.[63]

The clinical significance of P duration as per clinical literature understanding is that P wave ascribes to Atrial depolarization mechanism. The ECG is comprised of atrial depolarization (P wave) followed by ventricular depolarization. The Right atrial depolarization is followed by Left atrial Page of 9 16 depolarization, where the right depolarization vector propagates to the left atria and initiates the left depolarization. In the context of AHA/ESC guidelines the duration aspect is not explicitly mentioned, to the best of our knowledge and search. However, several other sources and clinical books have brought out these duration and mechanism, significance (esp. Douedi *et al*). The 120 ms cutoff for P-wave duration as a marker of LAE is a longstanding convention in electrocardiography. It is based on the understanding that normal P-wave duration is typically less than 120 ms. Prolongation beyond this threshold suggests delayed atrial conduction, often due to left atrial enlargement. Prolongation of the P wave to a duration exceeding 120 ms classically associates with left atrial abnormality and is referred to as P mitrale,[62] Douedi *et al*. The P duration criterion is referenced in various clinical studies and reviews. As per Chen *et al*,[63] a P-wave duration ≥120 ms is abnormal and is a criterion for partial inter atrial block, attributing it to a conduction delay between the right and left atria.

With regards to validation, we have added multi-case tests and quantified the results from the additional runs. Total of 1200 tests were run, where noise is gaussian distribution, with noise levels randomized between (0.001, 0.02 (mV)) based on record. The diversity of the test includes varying the mitrale peaks, width between the peaks and noise levels. The rationale behind 1200 tests is by analysis that the algorithm is discrete and that Mitrale pattern is deterministic comprised of two peaks, duration between two peaks and the depth of curve between the peaks. An approximation of the sample space of these variables along with the constraint comes to (P1, V, P2, W) to be considered, where P1 is Mitrale first peak, P2 is Mitrale second peak, V is the depth of the contour in the center and W the distance between the peaks. When we consider each of these variables as discrete, the combination comes to a maximum of (N x N x N x N, test cases). To begin with we take N equals to 5 discrete values and get test cases as 625. We then randomize these to increase the number of test cases and check accuracy with increased test cases. Beyond 1000 tests, with the constrained randomization, the accuracy gets stabilized at 95.42%. To a discrete algorithm for screening purpose this can be considered sufficient. The SGAD algorithm primarily is for inference of P mitrale pattern and we have obtained the accuracy of 95.42% for inference with total run of tests 1200, randomized. The multi-case tests are based on randomization of the 01447_hr record, because not all the records available in the PTB-XL are discernible. To the best of our knowledge there are no static noise “only”, influenced dataset besides PTB-XL. PTB-XL has LAE with

all types of noise mixed as well, viz burst mode noise. The scope of our work is restricted to detect Mitrale pattern in signals impacted by only static noise. Even in PTB XL, for LAE influenced there are only 54 records which are static noise only influenced. Within 54 records as per our scan, 01447_hr has noise levels and signal quality which can be discernible and hence.

With regards to quantitative metrics, computation time for EMD is 3 ms, for autoencoder the time required to training is 3.63 seconds, prediction took 28 ms, for SGAD the computation time is 0.143 ms (we account only the hump tracer, spread analyzer, parameter extraction and gaussian convolution fitting the pattern). The very low computation time for SGAD is due to the fact that it does not involve any explicit filtering. The tests were run on Apple M1 Pro (8 core, base clock 3.2 GHz), Sequoia 15.41. Python 3.12.4 version.

The SGAD has emphasis on interpretable computation and the notion we highlight is where we model the P component generation in the cardiac system as explained by the bio mechanism. The P corresponds to the depolarization of the Atria and is an electrical vector. This is considered as a gaussian lobe, for the reason that it inherently represents the P shape. This initiates the Right atrial depolarization, symbolically the right-side P vector. There is delay in the propagation of the left atrial electrical P vector to the right atrial P vector formation. This phenomenon can be represented as convolution of two gaussian lobes, the left lobe and right lobe. The phase shift between two lobes will decide the final P (composite) vector shape. Intuitively, when two gauss lobes are sufficiently close to each other, and we sum them up (convolve), the final P shape is also a gaussian curve. However, as the two lobes move apart, and upon convolution, there will be a dip in the center. This mechanism is explainable (interpretable) by a mere phase shift between two lobes. As the phase shift between the two lobes vary, the composite waveform will capture the delay between the two lobes as a morphology change which is bimodal. This has been depicted in the results figure with shift varied and corresponding morphology change. The bimodal pattern is recognized by the hump tracer and spread analyzer by looking for peaks, P1, P2, separated by at least 30 ms (as per previous clinical literature [59]), as referenced in 2024[64] this is considered to be 40 ms (Figure 4 b, two prominent peaks are visible as can be seen in the real ECG example (P mitrale pattern from a real ECG paper print (Adapted image, courtesy: Koley TK. Atrial enlargement. In: Rapid Review of ECG. Singapore: Springer; 2024 p. 138, Fig. 11-4[64]))), which is a paper print out of ECG lead II.). By identifying the peak values and the separation between them, the parameter is fed to the Gaussian convolution model, which then shows the composite P waveform.

3.1 Example compute complexity EMD solution, neural denoising methods and synthetic denoiser

EMD is a generalized and very effective decomposition solution applicable to varied application scenarios involving non-linearity and non-stationary signals. We discuss here a sample compute complexity of the solution using EMD and in general scenarios of importance. Our inference from varied exploration and literature study is that tailoring the solution for the application usage model provides a control in terms of trade off with respect to complexity. EMD is a data driven method adaptive to signal nature, decomposing it into several zero-mean signal components IMFs by an iterative procedure called sifting process. The sifting objective is to iterate till a stopping condition, convergence is met. The sifting process is a key component of the Empirical Mode Decomposition (EMD) method. Based on EMD[65] we have come up with the pseudo code (Algorithm 4) for a discretized version in our setup for comparison. The Eq. (3) depicts the EMD complexity for the discretized version which we have arrived at for our setup.

It is used to decompose a signal into a finite number of intrinsic mode functions (IMFs) that are simpler and easier to analyze. The sifting process can be described as follows:

1) Identify all the local maxima and minima of the signal.

2) Connect the local maxima and minima with cubic splines to form the upper and lower envelopes of the signal.

3) Calculate the mean of the upper and lower envelopes, and subtract it from the original signal to obtain the first IMF

4) Repeat steps 1-3 on the residual signal to obtain the second IMF, and so on, until the residual signal is a simple function

$$EMD\_complexity = O(nIMFS * NS * K) \quad (3)$$

where K is the number of discrete samples of the signal, nIMFs is the number of intrinsic mode functions extracted, NS is the sifting process iteration. Detailed analysis of compute complexity is available in.[65] The complexity is also a function of signal characteristics and the convergence conditions utilized for sifting process. Neural schemes of denoising have gained popularity because of its full-scale automation and domain agnostic nature of use case, varied application areas. There are multiple variants in neural schemes. One example is the autoencoder variant which is used in ECG denoising. The challenge here in lies with training procedure. The underlying principle of operation is simple and robust, where it learns a compressed representation of the input signal and then

Table 3: Comparison of denoising methods.

Model Attributes	Synthetic denoiser	EMD denoiser	Autoencoder denoiser
Objective	A directed explainable gaussian fitting amidst static noise for P contour of ECG to detect the Mitrale pathology pattern	A generalized decomposition model of signal with primary objective of denoising.	A generalized model with varied applications targets, a neural scheme
Methods	A discrete python framework of algorithms, synthetic model built from underlying explainable atrial depolarization dynamics using Gaussian lobes and extraction of parameters from raw ECG dataset	Use of an iterative sifting process to decompose the signal into intrinsic mode functions and a residue. Use of relevant decomposed signals to generate a smooth signal	It works by learning a compressed representation of the input signal and then reconstructing it back to its original form. Involves a training procedure to learn the underlying patterns.
Strengths	Directed towards explanation, rationale by fusing synthetic generation with explainable fitting	Effective in non-linear, non-stationary signal representation handling using space-time dynamics capture	Scores high in full scale automatic pattern recognition and has wide application target base
Limitations	Method, approach is coupled to underlying pattern and needs application domain insights	Sifting procedure is iterative and has challenges in convergence.	Training procedure is time consuming and has a need of example dataset if supervised scheme is involved. Black box nature, not inherently explanatory
Use case	Decision making in case of pathology patterns where explainable computation gains significance	Generalized denoising scheme with wide application target base	Wide application target base in denoising solution and generalized scheme.
Real time amenability	Being a directed method and low compute complexity is amenable in real time solutions	It is real time amenable, has its challenges though in terms of convergence	For real time amenability the hardware software partitioning trade-offs have to be dealt with. Does not specifically suit for real time unless crafted for it upfront.
Compute complexity	Low to moderate (Eq. (4))	High (Eq. (3))	Very high[46] in relation to directed and EMD methods
Computation time	0.143 ms	3 ms	3.63 seconds training, 28 ms prediction

reconstructs it back to its original form. The compressed representation is a lower dimensional version of the input signal that captures its most important features.

During training, the autoencoder learns to reconstruct the input signal from its compressed representation while minimizing the difference between the reconstructed signal and the original signal. Once trained, the autoencoder can be used to denoise new ECG signals by feeding them through the network and using the reconstructed output as the denoised signal. Based on our exploration and experimental set up of generator adversarial network (GAN), where we used it to generate a P contour, we estimate the compute complexity to be in the similar range, which is in the order of 10 K to 25 K epochs.[46] Our solution, the synthetic generator as denoiser is a context dependent algorithm coupled to the pathology. The approach does heuristic exploits of the domain, where in, the general model is built using underlying physical principles. Post this the explainable fitting uses the synthetic generator to fit the observed pattern by using the pathology pattern P mitrale. The compute complexity of the synthetic generator considering Algorithms 1, 2 is as in Eq. (4).

$$SynthDenoiser\_complexity = O(K) \quad (4)$$

where K is the number of discrete samples of the signal, which in our case is limited to 80 considering the parameter extraction piped computation stage. From the perspective of explanation, we mention that the solution provides an effective approach to building explain-able computation blocks, mimicking the cardiologist way. The limitation which we foresee in our scheme is that it is coupled to specifics of underlying pathology. Thus, much of the underlying pattern insights are domain tied and need to be captured upfront in the solution space. So, per pathology pattern, we might need to investigate the nature and its coupling to building the explanatory computation block. Thus, it is observed that the very strength to which it was targeted at has become weakness in some other view. The limitation of SGAD apart from current focus on P mitrale, is that it does not account for mixed type of noise, such as static and burst mode noise together. The current algorithm works only for static noise, which is a random noise spread across the spectrum. In the current form, SGAD cannot infer right atrial enlargement as it designed for two peaks and left atrial enlargement. In the right atrial enlargement, the P composite has only one peak and its value is considered >= 0.25 mV. Since it is value (threshold) based, inferring it based on threshold amidst noise without explicit filtering is a limitation for SGAD. We have made use of Python library PyEMD[66] in our setup to obtain denoised Mitrale pattern for comparison (Fig. 6). Fig. 7 and Fig. 8 are the results of denoising in autoencoder setup utilizing the Python library, Tensor flow[67] and Keras.[68] We also note the comparison of different denoising methods with synthetic denoiser in Table 3.[46,67,68]

4. Conclusion

Synthetic generator as a denoiser is an approach in stark contrast with existing denoising solutions. The scheme demonstrates building explainable computation blocks as explain ability takes significance in pathology decision making. The fitting using the Gaussian lobes comes out as inherently explainable, of the underlying atrial depolarization dynamics. It is amenable to capture the underlying pathology patterns in terms of rationale. The rationale is represented by atrial depolarization vectors as a Gaussian function convoluted using underlying cardiac function, providing relevant insight. The parameter extraction adapts to the ECG dataset as well be targeted at receiving inputs from the atrial cellular dynamics. Thus, patient specific models will find the current approach, of much relevance when plugged in the processing pipeline. The development cost in any solution is dictated by complexity, methods and the scale. It works in the favor of synthetic generator as denoiser, because it combines varied functionality in a single package, thus bring down the economics.

Real time amenability too is in the favor of directed solutions rather than very generalized solutions, where complexity can turn out to be hindrance. In the current contribution we have demonstrated use of constrained random generation in validation stack, providing an effective debug and validation scheme for the discrete algorithmic framework. We also addressed the perils of generalization, full automation, the black box which hampers explanatory solutions. While we showcased the strength of our solution as explain-able and low cost per function, the same can be viewed as a limitation, when it comes to coupling every pathology insight into the building of explainable block. We find that there is a vast scope in this area of building explainable computational blocks.

Acknowledgments

The authors gratefully acknowledge Sri Mata Amritanandamayi Devi (Amma) Chancellor, Amrita Vishwa Vidyapeetham, for her inspiration and for providing financial support for the Article Processing Charges (APC) of the publication.

Conflict of Interest

There is no conflict of interest.

Supporting Information

Not applicable.

CRediT Statement

Krishnadas Bhagwat: Idea, Conceptualization of frame work, Design and development of algorithms. **Supriya M** and **Sreeja Kochuvila:** Document reviews, Reviews, Discussions and validation. **Abhilash Ravikumar:** Supervision, logistics and discussions.

References

[1] R. R. Nair, T. Singh, Retraction Note: MAMIF: multimodal adaptive medical image fusion based on B-spline registration and non-subsampled shearlet transform, *Multimedia Tools and Applications*, 2024, **83**, 84455, doi: 10.1007/s11042-024-19941-y.

[2] T. Babu, T. Singh, D. Gupta, and H. Shahin, Prediction of normal and grades of cancer on colon biopsy images at different magnifications using minimal robust texture and morphological features, *Multimedia Tools and Applications*, 2020, **11**, 695-701, doi: 10.37506/ijphrd.v11i1.532.

[3] N. Aishwarya, N. G. Praveena, B. Rajalakshmi, R. Reshma, B. Rizwana Begum and D. Sowmya, Detection of Brain Tumor by Image Fusion based on Convolutional Neural Network, *International Journal of Advanced science and Technology*, 2020, **29**, 6500-6509.

[4] N. Aishwarya, C. Bennila Thangammal, An image fusion framework using morphology and sparse representation, *Multimedia Tools and Applications*, 2018, **77**, 9719-9736, doi: 10.1007/s11042-017-5562-4.

[5] A. Bayés de Luna, Basic Basic Electrocardiography: Normal and Abnormal ECG Patterns, Wiley-Blackwell, *General & Introductory Medical Science*, 2007, 352, ISBN - 9781405104838.

[6] F. M. Kusumoto, ECG Interpretation, *From Pathophysiology to Clinical Application*, Springer Cham, 2009, 1–400, doi: 10.1007/978-3-030-40341-6.

[7] A. Gacek and W. Pedrycz, ECG Signal Processing, Classification and Interpretation, *A Comprehensive Framework of Computational Intelligence*, Springer London, 2012, 1–350, ISBN - 9780857298683.

[8] G. S. Wagner, D. G. Strauss, Marriott's Practical Electrocardiography, 12th Edition, *Ovid Technologies*, Lippincott Williams & Wilkins, 2014, 532, ISBN - 1451146256, 9781451146257.

[9] A. Goldberger, Z. Goldberger, and A. Svilkin, Goldberger's Clinical Electrocardiography: A Simplified Approach, Amsterdam, Elsevier, *Science Direct*, 2017, 1–500, ISBN - 9780323390540.

[10] S. Kaplan Berkaya, A. K. Uysal, E. Sora Gunal, S. Ergin, S. Gunal, M. B. Gulmezoglu, A survey on ECG analysis, *Biomedical Signal Processing and Control*, 2018, **43**, 216-235, doi: 10.1016/j.bspc.2018.03.003.

[11] V. Gupta, M. Mittal, V. Mittal, R-peak detection using chaos analysis in standard and real time ECG databases, *Irbm*, 2019, **40**, 341-354, doi: 10.1016/j.irbm.2019.10.001.

[12] C. M. Cutillo, K. R. Sharma, L. Foschini, S. Kundu, M. MacKintosh, K. D. Mandl, T. Beck, E. Collier, C. Colvis, K. Gersing, V. Gordon, R. Jensen, B. Shabestari, N. Southall, Machine intelligence in healthcare: perspectives on trustworthiness, explainability, usability, and transparency, *NPJ Digital Medicine*, 2020, **3**, 47, doi: 10.1038/s41746-020-0254-2.

[13] A. F. Markus, J. A. Kors, P. R. Rijnbeek, The role of explainability in creating trustworthy artificial intelligence for health care: a comprehensive survey of the terminology, design choices, and evaluation strategies, *Journal of Biomedical Informatics*, 2021, **113**, 103655, doi: 10.1016/j.jbi.2020.103655.

[14] Microsoft, Microsoft Copilot, 2024, https://copilot.github.com.

[15] D. Rusinaru, Y. Bohbot, C. Kowalski, A. Ringle, S. Maréchaux, C. Tribouilloy, Left atrial volume and mortality in patients with aortic stenosis, *Journal of the American Heart Association*, 2017, **6**, e006615, doi: 10.1161/jaha.117.006615.

[16] T. A. Gaziano, J. M. Gaziano, Global burden of cardiovascular disease, *Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine*, Amsterdam, Elsevier, 2012, 1-20, doi: 10.1016/b978-1-4377-0398-6.00001-9.

[17] G. M. Friesen, T. C. Jannett, M. A. Jadallah, S. L. Yates, S. R. Quint, H. T. Nagle, A comparison of the noise sensitivity of nine QRS detection algorithms, *IEEE Transactions on Biomedical Engineering*, 1990, **37**, 85-98, doi: 10.1109/10.43620.

[18] V. S. Chouhan and S. S. Mehta, Detection of QRS complexes in 12-lead ECG using adaptive quantized threshold, *International Journal of Computer Science and Network Security*, 2008, **8**, 155–163.

[19] G. D. Clifford, *Advanced Methods and Tools for ECG Data Analysis*, Artech House, 2006, 1–450, ISBN - 9781580539661.

[20] K. L. Venkatachalam, J. E. Herbrandson, S. J. Asirvatham, Signals and signal processing for the electrophysiologist, *Circulation: Arrhythmia and Electrophysiology*, 2011, **4**, 965-973, doi: 10.1161/circep.111.964304.

[21] A. Kumar, R. Komaragiri, M. Kumar, From pacemaker to wearable: techniques for ECG detection systems, *Journal of Medical Systems*, 2018, **42**, 34, doi: 10.1007/s10916-017-0886-1.

[22] L. Gao, Y. Gan, J. Shi, A novel intelligent denoising method of ECG signals based on wavelet adaptive threshold and mathematical morphology, *Applied Intelligence*, 2022, **52**, 10270-10284, doi: 10.1007/s10489-022-03182-3.

[23] P. B. Patil, M. S. Chavan, A wavelet based method for denoising of biomedical signal, *International Conference on Pattern Recognition, Informatics and Medical Engineering*, Salem, India, March 21-23, 2012, 278-283, doi: 10.1109/ICPRIME.2012.6208358.

[24] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, H. H. Liu, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proceedings of the Royal Society of London Series A: Mathematical, Physical and Engineering Sciences*, 1998, **454**, 903-995, doi: 10.1098/rspa.1998.0193.

[25] Z. Wu, N. E. Huang, Ensemble empirical mode decomposition: a noise-assisted data analysis method, *Advances in Adaptive Data Analysis*, 2009, **1**, 1-41, doi: 10.1142/s1793536909000047.

[26] M.-T. Lo, L.-Y. Lin, W.-H. Hsieh, P. C. Ko, Y.-B. Liu, C. Lin, Y.-C. Chang, C.-Y. Wang, V. H. Young, W.-C. Chiang, J.-L. Lin, W.-J. Chen, M. H. Ma, A new method to estimate the amplitude spectrum analysis of ventricular fibrillation during cardiopulmonary resuscitation, *Resuscitation*, 2013, **84**, 1505-1511, doi: 10.1016/j.resuscitation.2013.07.004.

[27] M.-H. Lee, K.-K. Shyu, P.-L. Lee, C.-M. Huang, Y.-J. Chiu, Hardware implementation of EMD using DSP and FPGA for online signal processing, *IEEE Transactions on Industrial Electronics*, 2011, **58**, 2473-2481, doi: 10.1109/TIE.2010.2060454.

[28] X. Ye, Y. Hu, J. Shen, R. Feng, G. Zhai, An improved empirical mode decomposition based on adaptive weighted rational quartic spline for rolling bearing fault diagnosis, *IEEE Access*, 2020, **8**, 123813-123827.

[29] A. Tabrizi, L. Garibaldi, A. Fasana, S. Marchesiello, Influence of stopping criterion for sifting process of empirical mode decomposition (EMD) on roller bearing fault diagnosis, *Advances in Condition Monitoring of Machinery in Non-Stationary Operations*, 2013, 389-398, doi: 10.1007/978-3-642-39348-8_33.

[30] C. van Jaarsveldt, G. W. Peters, M. Ames, M. Chantler, Tutorial on empirical mode decomposition: basis decomposition and frequency adaptive graduation in non-stationary time series, *IEEE Access*, 2023, **11**, 94442-94478.

[31] K. Peng, H. Guo, X. Shang, EEMD and multiscale PCA-based signal denoising method and its application to seismic P-phase arrival picking, *Sensors*, 2021, **21**, 5271, doi: 10.3390/s21165271.

[32] J. Zhang, R. Yan, R. X. Gao, Z. Feng, Performance enhancement of ensemble empirical mode decomposition, *Mechanical Systems and Signal Processing*, 2010, **24**, 2104-2123, doi: 10.1016/j.ymssp.2010.03.003.

[33] S. Chatterjee, R. S. Thakur, R. N. Yadav, L. Gupta, D. K. Raghuvanshi, Review of noise removal techniques in ECG signals, *IET Signal Processing*, 2020, **14**, 569-590, doi: 10.1049/iet-spr.2020.0104.

[34] P. Vincent, H. Larochelle, Y. Bengio, P.-A. Manzagol, Extracting and composing robust features with denoising autoencoders, *Proceedings of the 25th International Conference on Machine Learning – ICML '08*, Helsinki, Finland, ACM, July 5-9, 2008, 1096-1103, doi: 10.1145/1390156.1390294.

[35] P. Xiong, H. Wang, M. Liu, F. Lin, Z. Hou, X. Liu, A stacked contractive denoising auto-encoder for ECG signal denoising, *Physiological Measurement*, 2016, **37**, 2214-2230, doi: 10.1088/0967-3334/37/12/2214.

[36] P. Xiong, H. Wang, M. Liu, S. Zhou, Z. Hou, X. Liu, ECG signal enhancement based on improved denoising auto-encoder, *Engineering Applications of Artificial Intelligence*, 2016, **52**, 194-202, doi: 10.1016/j.engappai.2016.02.015.

[37] S. Li, J. Li, Y. Zhang, X. Zhang, and Y. Li, A novel VMD-SSA-SVD algorithm for ECG signal denoising based on deep learning, *IEEE Access*, 2021, **9**, 107155–107166, doi: 10.1109/ACCESS.2021.3100021.

[38] Y. Li, X. Zhang, Y. Li, Y. Wang, and Y. Zhang, Denoising of ECG signals using deep learning models, *Journal of Healthcare Engineering*, 2021, **2021**, 1–12, doi: 10.1155/2021/6695698.

[39] Y. Zhang and Y. Wang, Accurate ECG classification based on spiking neural network and channel-wise attentional module, *Electronics*, 2022, **11**, 1889, doi: 10.3390/electronics11121889.

[40] K. Luo, J. Li, Z. Wang, and A. Cuschieri, Patient-specific deep architectural model for ECG classification, *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2017, 1466–1470, doi: 10.1109/BIBM.2017.8217860.

[41] G. Lenis, N. Pilia, A. Loewe, W. H. W. Schulze, O. Dössel, Comparison of baseline wander removal techniques considering the preservation of ST changes in the ischemic ECG: a simulation study, *Computational and Mathematical Methods in Medicine*, 2017, **2017**, 9295029, doi: 10.1155/2017/9295029.

[42] J. Pierzchlewski, T. Arildsen, Generation and analysis of constrained random sampling patterns, *Circuits, Systems and Signal Processing*, 2016, **35**, 3619-3643, doi: 10.1007/s00034-015-0216-0.

[43] D. Yllanes, Efficient sampling of constrained high-dimensional parameter spaces with generative models, *The European Physical Journal C*, 2021, **81**, 1–10, doi: 10.1140/epjc/s10052-021-09956-0.

[44] K. Luo, J. Li, Z. Wang, A. Cuschieri, Patient-specific deep architectural model for ECG classification, *Journal of Healthcare Engineering*, 2017, **2017**, 4108720, doi: 10.1155/2017/4108720.

[45] Suran Galappaththige, Richard A. Gray, C. Mendonca Costa, S. Niederer, P. Pathmanathan, Credibility assessment of patient-specific computational modeling using patient-specific cardiac modeling as an exemplar, *PLoS Computational Biology*, 2022, **18**, e1010541, doi: 10.1371/journal.pcbi.1010541.

[46] K. Bhagwat, M. Supriya, A. Ravikumar, Map composition framework for synthetic P morphology, *Biomedical Signal Processing and Control*, 2023, **79**, 104063, doi: 10.1016/j.bspc.2022.104063.

[47] P. Wagner, N. Strodthoff, R.-D. Bousseljot, D. Kreiseler, F. I. Lunze, W. Samek, T. Schaeffter, PTB-XL, a large publicly available electrocardiography dataset, *Scientific Data*, 2020, **7**, 154, doi: 10.1038/s41597-020-0495-6.

[48] P. Zeng, S. Kang, F. Fan, J. Liu, Enhanced heart sound anomaly detection *via* WCOS: a semi-supervised framework integrating wavelet, autoencoder and SVM, *Frontiers in Neuroinformatics*, 2025, **19**, 1530047, doi: 10.3389/fninf.2025.1530047.

[49] G. Y. Chen and A. Krzyzak, An experimental study for the effects of noise on pattern recognition algorithms, *Engineering Letters*, 2025, **33**, 1185-1192.

[50] A. Sohlberg, T. Kangasmaa, C. Constable, A. Tikkakoski, Comparison of deep learning-based denoising methods in cardiac SPECT, *EJNMMI Physics*, 2023, **10**, 9, doi: 10.1186/s40658-023-

00531-0.
[51] M. Li, Y. Jiang, Y. Zhang, H. Zhu, Medical image analysis using deep learning algorithms, *Frontiers in Public Health*, 2023, **11**, 1273253, doi: 10.3389/fpubh.2023.1273253.
[52] A. R. Palwankar, Medical images denoising using filters and neural network: comparison through implementations, *Journal of Information Systems Engineering and Management*, 2025, **10**, 501-511, doi: 10.52783/jisem.v10i9s.1249.
[53] R. Patil, S. Bhosale, Multi-modal medical image denoising using wavelets: a comparative study, *Biomedical and Pharmacology Journal*, 2023, **16**, 2271-2281, doi: 10.13005/bpj/2803.
[54] W. Lu, J. A. Onofrey, Y. Lu, L. Shi, T. Ma, Y. Liu, C. Liu, An investigation of quantitative accuracy for deep learning based denoising in oncological PET, *Physics in Medicine & Biology*, 2019, **64**, 165019, doi: 10.1088/1361-6560/ab3242.
[55] P. Zhang, M. Jiang, Y. Li, L. Xia, Z. Wang, Y. Wu, Y. Wang, H. Zhang, An efficient ECG denoising method by fusing ECA-Net and CycleGAN, *Mathematical Biosciences and Engineering*, 2023, **20**, 13415-13433, doi: 10.3934/mbe.2023598.
[56] P. Singh, G. Pradhan, A new ECG denoising framework using generative adversarial network, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2021, **18**, 759-764, doi: 10.1109/TCBB.2020.2976981.
[57] R. Holgado-Cuadrado, C. Plaza-Seco, L. Lovisolo, M. Blanco-Velasco, Characterization of noise in long-term ECG monitoring with machine learning based on clinical criteria, *Medical & Biological Engineering & Computing*, 2023, **61**, 2227-2240, doi: 10.1007/s11517-023-02802-5.
[58] B. Tutuko, A. Darmawahyuni, S. Nurmaini, A. E. Tondas, M. Naufal Rachmatullah, S. B. P. Teguh, F. Firdaus, A. I. Sapitri, R. Passarella, DAE-ConvBiLSTM: End-to-end learning single-lead electrocardiogram signal for heart abnormalities detection, *PLoS One*, 2022, **17**, e0277932, doi: 10.1371/journal.pone.0277932.
[59] A. Bayés de Luna, M. Fiol-Sala, A. Bayés-Genis, A. Baranchuk, *Clinical Electrocardiography*, Fifth edition, John Wiley & Sons Ltd, 2022, ISBN - 9781119488980.
[60] M. Abramowitz and I. A. Stegun, Eds, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Washington, DC: National Bureau of Standards, 1964, ISBN - 9780486612720.
[61] K. Bhagwat, M. Supriya, S. Kochuvila, A. Ravikumar, A synthetic cardiac episode generator for explainable, pathology based action potential using heuristic polynomial signatures, *Biomedical Signal Processing and Control*, 2024, **87**, 105552, doi: 10.1016/j.bspc.2023.105552.
[62] S. Douedi and H. Douedi, P wave, StatPearls [Internet], Treasure Island, FL: StatPearls Publishing, 2023, https://www.ncbi.nlm.nih.gov/books/NBK551635/.
[63] L. Y. Chen, A. L. P. Ribeiro, P. G. Platonov, I. Cygankiewicz, E. Z. Soliman, B. Gorenek, T. Ikeda, V. P. Vassilikos, J. S. Steinberg, N. Varma, A. Bayés-de-Luna, A. Baranchuk, P wave parameters and indices: a critical appraisal of clinical utility, challenges, and future research: a consensus document endorsed by the international society of electrocardiology and the international society for holter and noninvasive electrocardiology, *Circulation: Arrhythmia and Electrophysiology*, 2022, **15**, 2, doi: 10.1161/circep.121.010435.
[64] T. K. Koley, Atrial enlargement, *Rapid Review of ECG*, Singapore: Springer Nature Singapore, 2024, 135-142, doi: 10.1007/978-981-99-9116-7_11.
[65] Y.-H. Wang, C.-H. Yeh, H. V. Young, K. Hu, M.-T. Lo, On the computational complexity of the empirical mode decomposition algorithm, *Physica A: Statistical Mechanics and its Applications*, 2014, **400**, 159-167, doi: 10.1016/j.physa.2014.01.020.
[66] D. Laszuk, PyEMD: Python implementation of Empirical Mode Decomposition, GitHub repository, 2017, Available online: https://github.com/laszukdawid/pyemd. Documentation: https://pyemd.readthedocs.io.
[67] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, TensorFlow: Large-scale machine learning on heterogeneous systems, 2015, Software available at https://www.tensorflow.org.
[68] F. Chollet, Keras, GitHub repository, 2015. Available: https://keras.io.

Publisher's Note: Engineered Science Publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access
This article is licensed under a Creative Commons Attribution 4.0 International License, which permits the use, sharing, adaptation, distribution and reproduction in any medium or format, as long as appropriate credit to the original author(s) and the source is given by providing a link to the Creative Commons license and changes need to be indicated if there are any. The images or other third-party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or

exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/ licenses/by/4.0/.

©The Author(s) 2025.