IET Signal Processing

Review Article

IET Journals
The Institution of Engineering and Technology

Review of noise removal techniques in ECG signals

ISSN 1751-9675
Received on 14th March 2020
Revised 9th July 2020
Accepted on 27th July 2020
E-First on 15th September 2020
doi: 10.1049/iet-spr.2020.0104
www.ietdl.org

Shubhojeet Chatterjee[1] ✉, Rini Smita Thakur[1], Ram Narayan Yadav[1], Lalita Gupta[1], Deepak Kumar Raghuvanshi[1]

[1]*Department of Electronics and Communication Engineering, Maulana Azad National Institute of Technology, Bhopal, India*
✉ *E-mail: shubhojeet28@gmail.com*

Abstract: An electrocardiogram (ECG) records the electrical signal from the heart to check for different heart conditions, but it is susceptible to noises. ECG signal denoising is a major pre-processing step which attenuates the noises and accentuates the typical waves in ECG signals. Researchers over time have proposed numerous methods to correctly detect morphological anomalies. This study discusses the workflow, and design principles followed by these methods, and classify the state-of-the-art methods into different categories for mutual comparison, and development of modern methods to denoise ECG. The performance of these methods is analysed on some benchmark metrics, viz., root-mean-square error, percentage-root-mean-square difference, and signal-to-noise ratio improvement, thus comparing various ECG denoising techniques on MIT-BIH databases, PTB, QT, and other databases. It is observed that Wavelet-VBE, EMD-MAF, GAN2, GSSSA, new MP-EKF, DLSR, and AKF are most suitable for additive white Gaussian noise removal. For muscle artefacts removal, GAN1, new MP-EKF, DLSR, and AKF perform comparatively well. For base-line wander, and electrode motion artefacts removal, GAN1 is the best denoising option. For power-line interference removal, DLSR and EWT perform well. Finally, FCN-based DAE, DWT (Sym6) soft, MABWT (soft), CPSD sparsity, and UWT are promising ECG denoising methods for composite noise removal.

1 Introduction

Electrocardiogram (ECG) is a non-linear non-stationary quasi-periodic time series [1]. It is a wide-spread tool to examine the electrical and muscular functions of the heart. It is a time-varying bio-signal reflecting the ionic current flow, which causes contractions and subsequent relaxations in the cardiac fibres and provide indirect insight into the blood flow to the heart muscle [2]. It gives information about heart rate, rhythm, and electrical activity. The information contained within ECG is both physiological and pathological, which are integral to the diagnosis of heart diseases. ECG monitoring and subsequent analyses find a lot of applications in the medical domain. However, ECG is susceptible to different types of noises, which might distort the morphological features and the interval aspects of the ECG leading to a false diagnosis and improper treatment of patients. ECG is recorded by measuring the potential difference between two electrodes placed on the patient's skin.

Fig. 1*a* shows various components of ECG signal which are of two types, viz., morphological features: P-wave, QRS-complex, T-wave, and U-wave and interval features: PR-segment, ST-segment, PR interval, ST interval, RR interval, and so on [3–5]. ECG signals have a wide variety of applications in the medical domain such as cardiorespiratory monitoring, seizure detection and monitoring, ECG-based biometrics authentication, real-time analysis of electrocardiographic rhythm, heart-rate variability analysis using smart electrocardiography patch, and study of cardiac ischemia [6–11]. These applications require a proper determination of the morphological and interval aspects of the recorded ECG signal, which are susceptible to various kinds of predominant noises such as base-line wander (BW), muscle artefacts (MA) or electromyogram (EMG) noise, channel noise (additive white Gaussian noise, AWGN), power-line interference (PLI), and miscellaneous noises such as composite noise (CN), random noise, electrode motion artefacts (EM), and instrumentation noise, making it challenging to determine disease-specific morphological anomalies in the ECG signals.

ECG denoising is the process of estimating latent clean ECG signal from its noisy observation. Consider an N-sample observed data $\tilde{x}(n)$, where $n = 0, 1, 2, \ldots, N-1$, such that

$$\tilde{x}(n) = x(n) + \in_n \tag{1}$$

where $\{ \in_n \}$ is some noise process with variance σ^2. The denoiser aims at estimating the signal $x(n)$ with the help of a denoising method (Fig. 1*b*). The denoiser converts $\tilde{x}(n)$ into $\hat{x}(n)$. Lesser the difference between $x(n)$ and $\hat{x}(n)$, better is the performance of the denoising model.

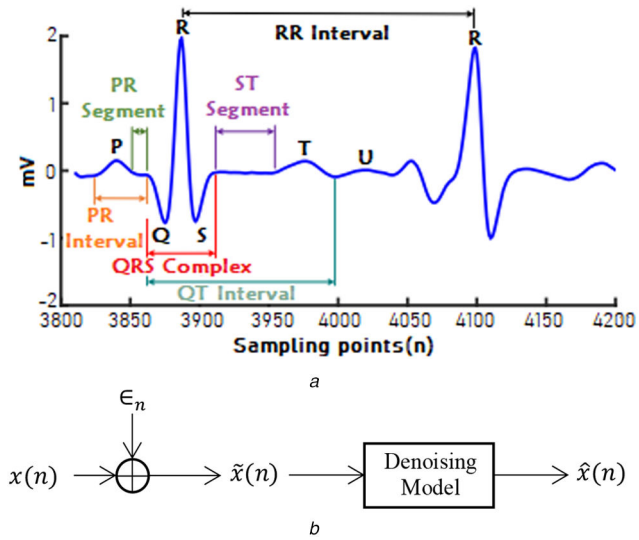

Fig. 1 *Denoising mechanism of a typical cardiac cycle*
(a) Typical cardiac cycle (heartbeat) from MIT-BIH Arrhythmia database (ECG 100 recording) [3], ***(b)*** Addition of noise $\in_n$ to clean ECG $x(n)$ to generate noisy observation $\tilde{x}(n)$ and estimation of clean ECG using the denoising model to get $\hat{x}(n)$

Table 1 Predominant noises in ECG signals

Type	Causes	Peak amplitude and duration	Spectrum range	Effects
baseline wander [12, 13]	respiration, body movements, poor electrode contact, and skin-electrode impedance	depend on electrode properties, electrolyte properties, skin impedance, and the subject movement	ranges between 0.05 and 1 Hz	distorts ST-segment and other LF components of the ECG signal
power-line interference [15, 16]	inductive and capacitive couplings of ubiquitous power lines in the ECG signal acquisition circuitry	amplitude: 50% of the peak-to-peak ECG signal amplitude	narrowband noise centred at 50/60 Hz with a bandwidth <1 Hz	distorts amplitude, duration, and shape of low-amplitude local waves of the ECG signal
muscle artefacts [17]	electrical activity of muscles during periods of contraction or due to a sudden body movement	amplitude: 10% of the peak-to-peak ECG signal amplitude	bandwidth ranges between 20 and 1000 Hz	alter the shapes of local waves of the ECG signal

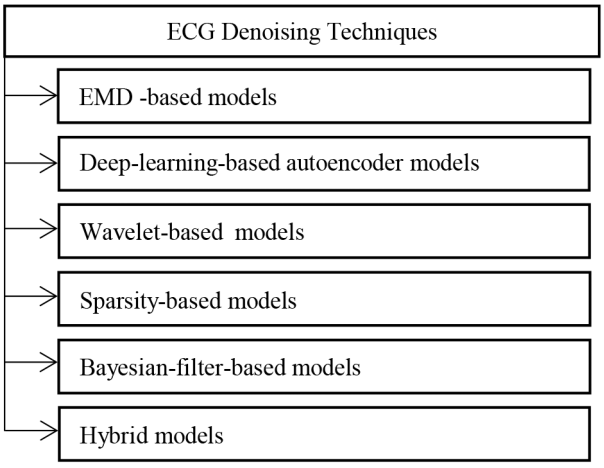

Fig. 2 *Classification of ECG signal denoising techniques*

1.1 Predominant noises in ECG

These noises and artefacts lie within the spectral range of interest and manifest themselves pre-dominantly as morphological features similar to the inherent aspects of the ECG or similar to any disease-specific aspects. A brief description of predominant noises in ECG is given below.

1.1.1 Base-line wander: BW is a low-frequency (LF) artefact caused mainly by respiration, body movements, bad electrode contact, and skin-electrode impedance [12, 13]. The amplitude and duration of the wander depend on electrode properties, electrolyte properties, skin impedance, and body movements [13]. This drift in the baseline is of magnitudes as high as around 15% of full-scale deflection (FSD), the peak-to-peak ECG amplitude over a frequency range of 0.15–0.3 Hz. Abnormal breathing rate and electrode movement alter the ECG by increasing the wandering frequency and causing motion artefacts, respectively [12]. BWs distort the ST-segment and other LF components of the ECG signal causing the wrong diagnosis of myocardial infarction, Brugada syndrome, and other ST-segment related abnormalities [14].

1.1.2 Power-line interference: PLI noises are caused by inductive and capacitive couplings of $50/60 \pm 0.2$ Hz power lines during ECG signal acquisition. It is narrowband with a bandwidth of <1 Hz and with an amplitude of up to 50% FSD [15, 16]. The intermixing of the PLI contents with the ECG distorts morphology of the signals. This leads to P-wave distortions, leading to the wrong diagnosis of atrial arrhythmias like atrial enlargement and fibrillation [14].

1.1.3 Muscle artefacts: MA or EMG noise is caused by electrical activities in muscles, which arise from eye and muscle movements and heartbeat. Typical sources of MA are muscle movements near the head region, like neck movements, swallowing, and so on [17]. The electrical activities due to muscle contractions last for a duration of around 50 ms between DC and 10,000 Hz, the amplitude being around 10% FSD [13]. EMG leads to distortion of local waves of the ECG signals due to a frequency match in the range of 0.01–100 Hz. This makes it challenging to denoise the signals for proper recognition of various ECG arrhythmias [14].

1.1.4 Channel noise: Channel noise is induced in ECG signals when they are transmitted through a channel with poor channel conditions, e.g. AWGN [18].

1.2 Miscellaneous noises

CN refers to a mixture of various noises, for example, a mixture of EM, BW, and MA [19], or a mixture of EM, BW, PLI, and MA [13]. In [20], the noises present in the atmosphere/surroundings are referred to as CN. Random noise is added to clean ECG signals to consider the worst-case scenario. EM are transient base-line changes (rapid drifts) caused by changes in the electrode-skin impedance with electrode motion. The amplitude and duration of the motion artefacts are 500% of peak–peak ECG amplitude and 300–500 ms, respectively [14]. All predominant noises with their causes, peak amplitude, duration, spectrum range, and effects are provided in Table 1.

1.3 Classification of the ECG denoising techniques

The ECG denoising methods have been classified into different categories, as mentioned in Fig. 2. The first category belongs to ECG denoising using EMD, which is a local and adaptive method in the frequency–time analysis. Empirical mode decomposition (EMD) is a data-driven mechanism which is proposed by Huang *et al.* [21], suited for non-linear and non-stationary signals [22]. The techniques included in the second and third categories are statistical and are used to extract a statistical-based model of the noisy signal [23]. The second category includes deep-learning-based autoencoder models (DAEs), which aim at regenerating a clean ECG signal from a corrupted version of the same by optimising the objective function. The wavelet-based methods fall into the third category and use the wavelet transform (WT) as the base for denoising ECG by decomposing the signal, deciding the type of thresholding and reconstructing the signal. The fourth category utilises the sparsity property of ECG for sparse optimisation to denoise ECG signals. An important denoising approach is based on adaptive filtering, which deals with model-based Bayesian filters such as the extended Kalman filter (EKF), Extended Kalman Smoother (EKS), and unscented Kalman Filter (UKF). The fifth category uses Bayesian filters to introduce changes in the conventional dynamic ECG model of Kalman filter to denoise ECG signals. The last category is hybrid that combines different methods available in the literature. Some filtering techniques like conventional filtering [24] and adaptive filtering [25] help to denoise ECG signals. Non-local means (NLM) has been explored for denoising ECG signals [26]. Also, various optimisation techniques like the total variation regularised least-squares problem or the related fused lasso problem (1DTVD) [27],

1. Input: noisy ECG signal $\tilde{x}(n)$
2. Sifting process
3. Repetitive application of the sifting process on proto IMF $h_k(n)$ until the stopping criterion (SD) gives the first IMF $c_1(n)$
4. Residue: $r_1(n) = \tilde{x}(n) - c_1(n)$
5. $r_1(n)$ might contain some useful signal information, hence the algorithm is run with $r_1(n)$ in place of $\tilde{x}(n)$
6. Hence, $\tilde{x}(n) = \sum_{l=1}^{L} c_l(n) + r_L(n)$
   $\tilde{x}(n)$ has L IMFs and the algorithm terminates at the L^{th} iteration, where the residue is either a constant, a monotonic slope, or a function with only one extremum
7. Thresholding or filtering or adaptive selection of IMFs ($\{c_l(n)\}$) for the purpose of signal reconstruction to get the denoised signal $\hat{x}(n)$
$$\hat{x}(n) = \sum_{l=1}^{L} \hat{c}_l(n) + r_L(n)$$
   where $\hat{c}_l(n)$ is filtered IMF
8. Output: $\hat{x}(n)$: denoised ECG signal

Fig. 3 *Algorithm 1: denoising algorithm of EMD-based denoiser [21]*

MM technique [28], genetic algorithm minimisation of a new noise variation estimate (GAMNVE) [29], and so on, help denoise ECG. Some conventional statistical techniques available in literature are principal component analysis and independent component analysis. ECG denoising based on these methods is well demonstrated in [30–33].

2 Techniques for ECG noise removal

2.1 EMD-based models for ECG signal denoising

EMD is an adaptive iterative algorithm through which a signal is decomposed into a series of its oscillatory segments, known as intrinsic mode functions (IMFs). With this iterative decomposition of signals, EMD separates the full signal into ordered elements with frequencies ranged from higher to lower frequencies in each IMF level [34]. The decomposition of the EMD procedure is based on the local time characteristics of the signal, thus it applies to non-linear and non-stationary processes [35]. EMD relies on an entirely data-driven mechanism that does not require any a priori known basis, as opposed to data analysis methods like Fourier transform [36].

2.1.1 Architecture of EMD-based ECG denoiser: Noisy ECG $\tilde{x}(n)$ is given to the denoiser as an input. First, all the maxima of the noisy input are joined using cubic spline interpolation, giving $e_u(n)$ as the upper envelope. Similarly, the minima are connected to get the lower envelope, i.e. $e_l(n)$. The next step is to obtain the mean of the envelopes, as shown below:

$$m_1(n) = [e_u(n) - e_l(n)]/2 \tag{2}$$

This mean is subtracted from the signal to obtain the first proto-IMF

$$h_1(n) = \tilde{x}(n) - m_1(n) \tag{3}$$

This procedure is called the sifting process. An IMF holds two characteristics: (i) the number of zero-crossings and extrema differ by at most one, (ii) the IMF waveform is symmetric concerning the local mean [35]. Therefore, if $h_1(n)$ does not satisfy the characteristics, all the steps from interpolation to subtracting the mean from the signal are repeated until $h_k(n)$ becomes an IMF, where k is the iteration variable. To stop the iteration, a certain terminating criterion is required. More than seven different stopping criteria have been proposed after the development of the EMD algorithm [37]. A rather important criterion is as follows [38]:

$$\text{SD} = \sum_{n=0}^{N-1} |h_{k-1}(n) - h_k(n)|^2 / h_k^2(n) \tag{4}$$

If SD (standard difference) < ξ, where ξ is a set threshold typically ranging between 0.2 and 0.3 [39], the iteration stops. Hence, repetitive application of the sifting process on proto IMF $h_k(n)$ until the stopping criterion gives the first IMF $c_1(n)$. Next, a residue is obtained by subtracting the IMF from the noisy input

$$r_1(n) = \tilde{x}(n) - c_1(n) \tag{5}$$

This residue might contain some important signal information. Hence, the algorithm is now repeated on the residue again and again until the residue is either a constant or a monotonic slope, or a function with only one extremum. Hence, the signal is broken down as follows:

$$\tilde{x}(n) = \sum_{l=1}^{L} c_l(n) + r_L(n) \tag{6}$$

Thus, $\tilde{x}(n)$ has L IMFs as the algorithm terminates at the Lth iteration. The above algorithm of EMD is proposed by Huang *et al.* [21]. After the extraction of IMFs using EMD, the final stage is thresholding or filtering of the obtained IMFs, followed by reconstruction of the signal to obtain the denoised signal $\hat{x}(n)$. Algorithm for EMD-based ECG signal denoising is given in Fig. 3.

2.1.2 Methodologies: Blanco-Velasco *et al.* [36] have concluded after an analysis of EMD on clean and noisy ECG that it is possible to decontaminate the noisy signal and at the same time preserve the QRS complex by applying temporal processing in the EMD domain, which otherwise is difficult due to the similar spectral presence of both noise and ORS complex in the high-frequency (HF) band. Removal of lower-order IMFs would have led to severe QRS complex distortion. Hence, the denoising process devised to tackle noisy ECG is explained in Algorithm 2 (see Fig. 4).

Sonali *et al.* [40] also used an approach similar to Algorithm 2 (Fig. 4) to denoise ECG. The difference lies in the last step where smoothing of the QRS complexes takes place to get a noise-free ECG signal by maintaining the biphasic or triphasic shape of the QRS complex. Pan *et al.* [41] have used EMD to remove the BW from noisy ECG. EMD is used to decompose $\tilde{x}(n)$ into 14 IMFs and one residue. As BW is usually a LF phenomenon, it is involved mainly in the last components of IMFs. Therefore, the sum of the last three IMFs can be regarded as BW. Hence, $\hat{x}(n)$ is obtained by combining the remaining IMFs and the residue. Chacko and Ari [42] have used spectral flatness (SF) measure to identify the noisy IMFs, which are further denoised using Butterworth filters. The EMD and cross-validation (EMDCV) method [43] joins cross-validation and thresholding approaches to smooth the IMF. Narsimha *et al.* [44] have used EMD to free noisy ECG of PLI, and BW. The power-line contaminated signal is broken into ten IMFs. The power-line information, being in 50–60 Hz frequency range, lies in the first IMF [45]. By reducing its contribution, the signal is regenerated to get a PLI-suppressed version of the noisy input. Here, the first IMF is reduced to one-fourth of its value. Hence, the reconstructed signal is given as follows:

$$\hat{x}(n) = \frac{1}{4} c_1(n) + \sum_{l=2}^{L} c_l(n) \tag{7}$$

BW can be removed by decomposing the signal into 15 IMFs and reconstructing the signal with suppressing the last IMF, which has the BW noise (0.2–0.5 Hz).

Samadi and Shamsollahi [46] have used a combination of soft-thresholding and instantaneous half period (IHP) thresholding to

1. Input: noisy ECG signal $\tilde{x}(n)$
2. Delineation of the QRS complex:
   (a).Identification of R-peak in the noisy input $\tilde{x}(n)$
   (b).Use of (1) through (6) of algorithm 1 to decompose the noisy signal $\tilde{x}(n)$
   (c).$d(n) = c_1(n) + c_2(n) + c_3(n)$, the summation of first three IMFs
   (d).Two local minima in $d(n)$ on either side of the R-peak are identified
   (e).Boundary determination of the QRS complex by locating closest zero-crossings on the left-hand side of the left minimum and on the right-hand side of the right minimum
3. Windowing to preserve QRS complex:
   a typical window is tapered cosine window
4. Determination of the number of noisy IMFs by statistical test: a hypothesis testing is conducted to determine the noise order O, which is actually the number of noisy IMFs. To avoid loss of information, a limit is set on the noise order as follows:
   $O_{actual} = \min(O, \beta)$, where β is typically 5
5. Partial reconstruction of the signal:
   (a).A window function $\lambda_l(n)$ is constructed by concatenating the window functions $w_{lj}(n)$, where $w_{lj}(n)$ denotes the variable size window for the j^{th} QRS complex in the l^{th} IMF
   $$\lambda_l(n) = \sum_{j=1}^{N_q} w_{lj}(n)$$
   Where N_q is the number of QRS complexes in the l^{th} IMF
   (b).$\lambda_l(n)$ eliminates noise and retains QRS-complex information. In order to avoid abrupt changes in the QRS complex, complement of $\lambda_l(n)$ is used. The complementary function ($\overline{\lambda}_l(n)$) allows a negligible noise in lower IMFs to reduce distortion in the reconstructed signal
   $\overline{\lambda}_l(n) = 1 - \lambda_l(n) \ \forall n$
6. Output: the reconstructed signal:
   $$\hat{x}(n) = \sum_{l=1}^{O_{actual}} \lambda_l(n)\, c_l(n) + \sum_{l=1}^{O_{actual}} \overline{\lambda}_l(n)\, c_l(n) + \sum_{l=O_{actual}+1}^{L} c_l(n) + r_L(n)$$

Fig. 4 *Algorithm 2: denoising algorithm by Blanco-Velasco et al. [36]*

estimate the IMFs of the clean signal $x(n)$ from those of the noisy signal $\tilde{x}(n)$. The use of soft-thresholding leads to the overhead requirement of precise noise detection, as the thresholding function is applicable to the noisy parts.

The soft-thresholding operation can be shown as follows [35]:

$$\hat{c}_l(n) = T(c_l(n), t) \tag{8}$$

where T is a pre-processing function to estimate $\hat{c}_l(n)$ from $c_l(n)$, and t is the value of the threshold. There are several different ways of thresholding. One method uses the global threshold for noise removal [35], where the threshold $t = (\text{MAD}_l/0.6745)\sqrt{2\log(L)}$, where $\text{MAD}_l = \text{Median}(c_l(n) - \text{Median}(c_l(n)))$ and L is the length of the signal. The pre-processing function for soft-thresholding is as follows:

$$\hat{c}_l(n) = \begin{cases} c_l(n) - t, & c_l(n) \geq t \\ 0, & |c_l(n)| < t \\ c_l(n) + t, & c_l(n) \leq -t \end{cases} \tag{9}$$

In the IHP thresholding, the pre-processing function is as follows:

$$\hat{c}_l(n) = \begin{cases} c_l(n), T_l^a \geq \chi, & t_l^a < t \leq t_l^{a+1} \\ 0, & \text{else} \end{cases} \tag{10}$$

where T_l^a is called the instantaneous half period and is defined for each oscillation as

$$T_l^a = t_l^{a+1} - t_l^a \tag{11}$$

and χ is the threshold of the process. The ath zero crossing in the lth IMF occurs at t_l^a. Tang and Qin [47] have proposed the use of 3σ rule (proposed in [48]) to determine the threshold and have used hard-thresholding to estimate $\{\hat{c}_l(n)\}$. Anapagamini and Rajavel [49] have proposed an algorithm to remove PLI from $\tilde{x}(n)$. Initially, the first IMF is filtered out and an intermediate version of the signal is obtained. Next, the first EMF is passed through a low-pass FIR filter, with a cut-off frequency of 31 Hz. To the intermediate version, this filtered low-pass output is added. The first IMF is removed because PLI is a HF phenomenon. The purpose of adding the filtered output is to retain the signal information lost during the process of IMF removal. The reconstructed signal is shown as follows:

$$\hat{x}(n) = \sum_{l=2}^{L} c_l(n) + \text{LPFIR}(c_1(n)) \tag{12}$$

where LPFIR denotes the low-pass FIR filter.

An algorithm for estimating BW is given in [49], where BW is predicted as the cumulative sum of the residue and the IMFs with frequency content <0.5 Hz, as BW is a LF artefact. The reconstructed signal, in this case, is as follows:

$$\hat{x}(n) = \sum_{l=1}^{\eta} c_l(n) \tag{13}$$

where η is the number of IMFs with frequency content ≥0.5 Hz. EMD though successful in ECG denoising leads to the problem of mode-mixing, which is a phenomenon where a certain IMF contains more than one oscillation component in each mode with dramatically disparate time scales [50]. Ensemble EMD (EEMD) solves the problem of mode-mixing. The algorithm of EEMD is given in Fig. 5.

Chang [34] has used EEMD and EMD to decompose to $\tilde{x}(n)$ into L IMFs and a residue. The filtering of IMFs, in this case, is based on designing an optimum band-pass filter to allow a certain range of IMFs. $\hat{x}(n)$ is obtained as follows:

$$\hat{x}(n) = \begin{cases} \sum_{l=\text{lower}}^{\text{upper}} c_l(n), & \text{EMD process} \\ \sum_{l=\text{lower}}^{\text{upper}} c_l(n), & \text{EEMD process} \end{cases} \tag{14}$$

Here, lower > 1 and upper < L for the band-pass filter design. The (lower, upper) pair with minimum MSE is determined for each filter by a sequential search approach. Wu and Huang [51] have proposed a new EEMD approach, in which an ensemble of white noise-added signal is sifted to obtain the mean, which is the true result. Nguyen and Kim [50] have used EEMD to obtain the IMFs of $\tilde{x}(n)$. For denoising purpose, Kullback–Leibler divergence is used to discriminate signal-dominant IMFs from the noise-dominant IMFs by measuring the probabilistic similarity between the PDFs of $\tilde{x}(n)$ and those of each mode. A genetic-algorithm-

1. Input : $x(n)$: clean ECG input
2. Generation of s different versions of $\tilde{x}(n)$, each having a different noise series $\in_n$ of the same power embedded in it
3. The noisy versions are denoted as $\tilde{x}_i(n)$, such that $\tilde{x}_i(n) = x(n) + \in_i (n)$
4. Each $\tilde{x}_i(n)$ is decomposed using EMD to get IMFs, and a residue
5. The IMFs are denoted as $c_{il}(n)$ and the residue is denoted as $r_i(n)$, for i^{th} noisy version and l^{th} IMF level. $r_i(n)$ follows the same termination criterion as given in the EMD algorithm
6. Now, mean (ensemble) for each IMF scale is obtained as: $EEMD\big(c_l(n)\big) = \frac{1}{s}\sum_{i=1}^{s} c_{il}(n)$
7. Mean residue: $EEMD\big(r_i(n)\big) = \frac{1}{s}\sum_{i=1}^{s} r_i(n)$
8. Thresholding or filtering or adaptive selection of ensemble IMFs
9. Output: $\hat{x}(n) = \sum_{l=1}^{L} EEMD'\big(c_l(n)\big) + EEMD\big(r_i(n)\big)$ where $EEMD'\big(c_l(n)\big)$ is filtered ensemble IMF

Fig. 5 *Algorithm 3: algorithm for EEMD [34]*

based IMF thresholding technique is used to reduce the noise in the noisy IMFs without losing much signal information by optimisation of threshold parameters of each IMF. Thus, $\hat{x}(n)$ is obtained by combining the signal-dominant and denoised IMFs. In [22], an automatic time-efficient noise adaptive IMF-selection procedure has been suggested. In this procedure, EEMD is performed to obtain IMFs. Next, to select the appropriate IMFs, correlation coefficients of different combinations of IMFs, and the ground truth signal are calculated. The IMFs which on removal, increase the correlation coefficient, are considered as the artefacts. Lastly, a matrix of correlation coefficients is constructed to select the combination of IMFs. Jenitta and Rajeswari [52] have combined EMD/EEMD with BLMS (block least means square algorithm, detailed in [53]) to denoise ECG of BW and PLI via an adaptive filter. First, $\tilde{x}(n)$ is decomposed using either EMD or EEMD. The obtained IMFs and a reference signal are input to an FIR filter, whose coefficients are updated using BLMS based on the error output of the FIR filter. The reference input contains noisy IMF components, which are closely related to the noisy components in ECG signals due to which adaptive filter converges faster. BW corrupts the lower frequency components of the ECG signal, hence two lower-order IMF components and the residue forms the reference. For PLI removal, the reference input contains the first three higher-order IMFs. Kaergaard *et al.* [54] have presented an analysis of EEMD-BLMS algorithm to denoise ECG signals. An LMS-based ECG denoising technique has also been discussed in [54, 55]. The technique of EEMD is further extended by Colominas *et al.* [56] to provide complete EEMD with adaptive noise to achieve negligible reconstruction error.

2.2 DAEs for ECG denoising

In this category of ECG denoising, deep-learning-based models are constructed based on the functionality of a denoising autoencoder (DAE). It is an initial unsupervised learning step that maps inputs to intermediate representations [57]. An autoencoder is a machine-learning model that has the function of regenerating an input signal with as much accuracy as possible. It works as a combination of two non-linear subparts, viz., encoder and decoder. The autoencoder-based denoising approach is discussed by Vincent *et al.* [57].

2.2.1 Architecture of DAE: First, a vector $\boldsymbol{x} \in [0,1]^d$ is given as an input to the system, d is the dimension of the input vector. Then, $\boldsymbol{x}$ is corrupted through a stochastic mapping $\tilde{\boldsymbol{x}} \sim q(\tilde{\boldsymbol{x}}|\boldsymbol{x})$ according to a certain destruction rate from x_j to zero, where $j = 1, 2, 3, \ldots, d$ [57].

In stochastic corruption process, a certain fraction of components of $\boldsymbol{x}$ are chosen randomly, and their value is set to 0, while the others are left unaltered. All information about the chosen components is thus removed from that particular input pattern, and the autoencoder is trained to fill the artificially introduced information gap, i.e. this corrupted signal $\tilde{\boldsymbol{x}}$ is used to train a deep network to regenerate the full signal from the partial information. It is imperative to note that the corruption process q is only used during training, but not for propagating representations from the raw input to higher-level representations [57, 58].

The encoder converts $\tilde{\boldsymbol{x}}$ to $\boldsymbol{z}$, which is the partial information. This non-linear transformation is as follows:

$$\boldsymbol{z} = f(\boldsymbol{W}\tilde{\boldsymbol{x}} + \boldsymbol{b}) \tag{15}$$

The decoder converts $\boldsymbol{z}$ back to reconstructed data $\hat{\boldsymbol{x}}$. This non-linear transformation is as follows:

$$\hat{\boldsymbol{x}} = g(\boldsymbol{W}'\boldsymbol{z} + \boldsymbol{b}') \tag{16}$$

f is the feature-extraction function used to represent the input $\tilde{\boldsymbol{x}}$ in the feature space [59]. The function is in the parameterised form with weight and bias parameters $\boldsymbol{W}$ and $\boldsymbol{b}$. The parameters of the decoder are $\boldsymbol{W}'$ and $\boldsymbol{b}'$. f and g are the non- linear activation functions used in the deep-learning models [19]. $\boldsymbol{\Phi} = \{\boldsymbol{W}, \boldsymbol{b}, \boldsymbol{W}', \boldsymbol{b}'\}$ forms the overall parameter set used to define the encoder and the decoder non-linear parameterised equations, where $\boldsymbol{\theta} = \{\boldsymbol{W}, \boldsymbol{b}\}$ are the weight and bias matrices of the encoder, respectively, while $\boldsymbol{\theta}' = \{\boldsymbol{W}', \boldsymbol{b}'\}$ are the weight and bias matrices of the decoder, respectively. These parameters are optimised by minimising the reconstruction error, which is the error between $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$. The average reconstruction error of the DAE is as follows:

$$\boldsymbol{\theta}^*, \boldsymbol{\theta}'^* = \arg\min_{\boldsymbol{\theta},\boldsymbol{\theta}'} \boldsymbol{E}_{q^o(X,\tilde{X})}\big[L_H\big(X, g_{\boldsymbol{\theta}'}\big(f_{\boldsymbol{\theta}}\big(\tilde{X}\big)\big)\big)\big] \tag{17}$$

where $\boldsymbol{E}_{p(X)}[f(X)]$ stands for expectation of $f(X)$ when X has a probability density $p(X)$, X is the random variable assigned to $\boldsymbol{x}$. The average reconstruction function of the basic autoencoder undergoes a set of changes to attain the above form, as detailed in [57]. The average reconstruction function of the basic autoencoder is

$$\boldsymbol{\theta}^*, \boldsymbol{\theta}'^* = \arg\min_{\boldsymbol{\theta},\boldsymbol{\theta}'} \frac{1}{n}\sum_{i=1}^{n} L\big(x^{(i)}, g_{\boldsymbol{\theta}'}\big(f_{\boldsymbol{\theta}}\big(x^{(i)}\big)\big)\big) \tag{18}$$

where L is the loss function defined as $L(\boldsymbol{x}, \hat{\boldsymbol{x}}) = \| \boldsymbol{x} - \hat{\boldsymbol{x}} \|^2$ and n is the number of training samples. In an alternate approach, the traditional loss function is replaced by a loss function with Bernoulli distance, and it can be measured with the following reconstruction cross-entropy:

$$L_H(\boldsymbol{x}, \hat{\boldsymbol{x}}) = (\beta_{\boldsymbol{x}} || \beta_{\hat{\boldsymbol{x}}}) \tag{19}$$

where β indicates that the negative log-likelihood function $L_H(\boldsymbol{x}, \hat{\boldsymbol{x}})$ follows Bernoulli distribution. When $L(\boldsymbol{x}, \hat{\boldsymbol{x}})$ is replaced by $L_H(\boldsymbol{x}, \hat{\boldsymbol{x}})$, the reconstruction function becomes

$$\boldsymbol{\theta}^*, \boldsymbol{\theta}'^* = \arg\min_{\boldsymbol{\theta},\boldsymbol{\theta}'} \boldsymbol{E}_{q^o(X)}[L_H(X, g_{\boldsymbol{\theta}'}(f_{\boldsymbol{\theta}}(X))] \tag{20}$$

where $q^o(X)$ is the empirical distribution associated with the input. $q^o(X)$ is replaced by $q^o\big(X, \tilde{X}\big)$, a joint distribution, due to stochastic mapping [57]. The basic difference between a basic autoencoder and a DAE is that $\boldsymbol{z}$ is a deterministic function of $\tilde{\boldsymbol{x}}$ rather than $\boldsymbol{x}$ and thus the result of a stochastic mapping of $\boldsymbol{x}$. Hence, we get (17) as an objective function minimised by stochastic gradient descent. The pictorial representation of the autoencoder architecture is shown in Fig. 6. Algorithm for the DAE-based ECG denoising is given in Fig. 7.

2.2.2 Methodologies: Chiang *et al.* [19] proposed fully convolutional network (FCN)-based DAE for ECG signal denoising. A FCN is a modification of CNN, in which the last fully connected layer is replaced by a convolutional layer, but the remaining layers, viz., convolutional layers, activation functions, max-pooling layers are not modified [60]. The encoder contains a series of six layers, with each layer consisting of a convolutional layer, a batch normalisation layer [61], and an activation layer. The encoder encodes the noisy ECG data ($\tilde{\boldsymbol{x}}$) into a low dimensional feature map ($\boldsymbol{z}$). The activation layer contains exponential linear units [62]. The convolutional framework is shown in Table 2. The filter size is fixed at 16 × 1. For a noisy ECG input of size 1024 × 1 (size of $\tilde{\boldsymbol{x}}$), a down-sampling with stride 2 is achieved as the feature map turns out to be of size 32 × 1 (size of $\boldsymbol{z}$). The decoder shows inverse symmetry to contain a series of six layers, in which the convolutional layers are replaced by deconvolutional layers. Deconvolutional layers project a single activation input to multiple outputs. The deconvolutional framework is inversely symmetric to the convolutional framework. All the layers are trained with a dropout probability of 0.5 to determine the parameter set $\boldsymbol{\Phi}$, which is determined by optimising the following objective function:

$$L(\boldsymbol{\theta}) = \sum_{i=1}^{d} \| x_i - \hat{x}_i \|^2 \tag{21}$$

The output layer is a deconvolutional layer with 1 filter of size 16 × 1 and stride of 1. The architecture of the FCN-based DAE is given in Fig. 8.

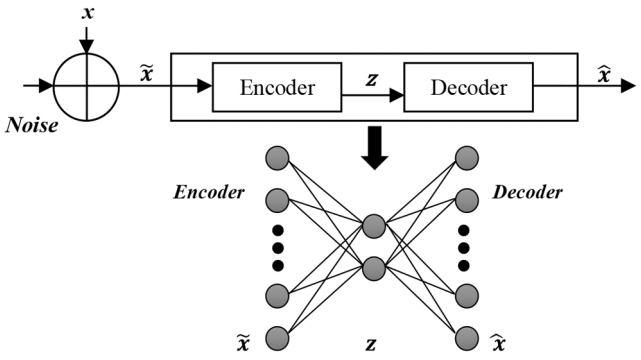

Fig. 6 *Schematic diagram of DAE [19], where $\boldsymbol{x}$: normalised input, $\tilde{\boldsymbol{x}}$: corrupted input, $\boldsymbol{z}$: partial information of input, and $\hat{\boldsymbol{x}}$: reconstructed output*

1. Input: $\boldsymbol{x} \in [0,1]^d$, normalized input vector, where d is the dimension of $\boldsymbol{x}$
2. Corrupt $\boldsymbol{x}$ through a stochastic mapping $\tilde{\boldsymbol{x}} \sim q(\tilde{\boldsymbol{x}}|\boldsymbol{x})$ according to a destruction rate from x_j to 0, where $j = 1,2,3,\ldots, d$
3. a. Encode $\tilde{\boldsymbol{x}}$ to hidden representation $\boldsymbol{z}$: $\boldsymbol{z} = f(\boldsymbol{W}\tilde{\boldsymbol{x}} + \boldsymbol{b})$
   b. Decode $\boldsymbol{z}$: $\hat{\boldsymbol{x}} = g(\boldsymbol{W}'\boldsymbol{z} + \boldsymbol{b}')$
   c. Encoding and decoding take place such that the parameters are optimized by minimizing the reconstruction error as given in (17)
4. Output: $\hat{\boldsymbol{x}}$

Fig. 7 *Algorithm 4: denoising algorithm of DAE [57]*

Table 2 Convolutional framework design

Layer	Number of filters	Stride
1	40	2
2	20	2
3	20	2
4	20	2
5	40	2
6	1	1

Xiong *et al.* [59] proposed a deep neural network with stacked contractive DAE (CDAE) to denoise the ECG signal. A deep-fully-connected neural network is built with the architecture of a CDAE as the underlying basis. The DNN has two hidden layers, viz., h_1 and h_2. $\boldsymbol{\Phi}_1$ and $\boldsymbol{\Phi}_2$ are the parameter sets for the two hidden layers. Each hidden layer has an encoder and a decoder. First, the DNN is pre-trained. During this, the parameter set for the encoder $\boldsymbol{\Phi}_{\text{en}} = \{\boldsymbol{W}, \boldsymbol{b}\}$ is initialised for each of the hidden layers. The input ($\boldsymbol{x}$) is corrupted to $\tilde{\boldsymbol{x}}$ and further converted to $\boldsymbol{z}_1$ by the encoder, which forms the feature vector. This is given as input to h_2 and the same process is repeated. The next step is fine-tuning of the DNN. It is done by updating all DNN parameters. The DNN updates all parameters through stochastic gradient descent to minimise the error between its output ($\hat{\boldsymbol{x}}$) and the target output ($\boldsymbol{x}$). After the training is over, noisy ECG signals are given to the stacked-CDAE DNN architecture for the purpose of denoising. The loss function with the distance of Bernoulli is as follows:

$$L\left(\boldsymbol{x}^{(i)}, \hat{\boldsymbol{x}}^{(i)}\right) = -\sum_{k=1}^{d} \left(x_k^{(i)} \log \hat{x}_k^{(i)} + \left(1 - x_k^{(i)}\right) \log\left(1 - \hat{x}_k^{(i)}\right)\right) \tag{22}$$

Each layer in the DNN is trained individually to minimise the reconstruction error. It is given as

$$\boldsymbol{\theta}^*, \boldsymbol{\theta}'^* = \arg\min_{\boldsymbol{\theta}, \boldsymbol{\theta}'} \sum_{i=1}^{n} \frac{1}{n} \left[L\left(\boldsymbol{x}^{(i)}, \hat{\boldsymbol{x}}^{(i)}\right) + \lambda \| J_{f_{\boldsymbol{\theta}}}\left(\tilde{\boldsymbol{x}}^{(i)}\right) \|_{\text{F}}^2 \right] \tag{23}$$

$\lambda \| J_{f_{\boldsymbol{\theta}}}\left(\tilde{\boldsymbol{x}}^{(i)}\right) \|_{\text{F}}^2$ is the penalty term with the penalty weight λ to penalise the sensitivity of the training input $\tilde{\boldsymbol{x}}$ and n is the number of training samples.

The corruption process q is only used during training. After training, each layer receives the uncorrupted output of the previous layer. The architecture of the DNN-based CDAE is given in Fig. 9. Chiang *et al.* [19] compared the performance of the FCN-based DAE with CNN-based DAE and DNN-based DAE. The dropout probability for the fully-connected layers is 0.5. The DNN has a structure similar to the FCN with 13 layers in the network and compresses the input 32 times to get the feature map. In [63], Xiong *et al.* have proposed an improved DAE for signal enhancement. To improve DAE-based ECG denoising, a generative adversarial network (GAN), which is a generator-discriminator model, has been proposed, in which the generator generates fake samples close to real data, and the discriminator discriminates between the real and the fake, as given in [64, 65]. In [64, 65], the authors have proposed novel GAN-based ECG denoising techniques. In [66], a deep recurrent denoising neural network (DRDNN) is used, which is a specific hybrid of a deep recurrent neural network and a DAE.

2.3 Wavelet-based models for ECG signal denoising

A WT expands the signal in terms of a localised wavelet function in both time and frequency. WT can provide good time resolution at HF and good frequency resolution at LF. Hence, a WT is quite suitable for ECG signal analysis. It is the decomposition of a signal into a set of basis functions consisting of contractions, expansions, and translations of a mother function $\psi(x)$, called the mother wavelet [67]. Dyadic WT (DWT) is useful in analysing ECG signals because of its fast computation and its multiresolution property [68]. DWT and its multiresolution property are discussed below.

For any function $f(x)$, its square-integrability can be defined as follows:

$$f(x) \in L^2\{\mathbb{R}\} \Leftrightarrow \int_{-\infty}^{+\infty} |f(x)|^2 \, \mathrm{d}x < \infty \tag{24}$$

where $L^2\{\mathbb{R}\}$ is the space of square-integrable functions in $\mathbb{R}$.

It follows from [69] if a function is square-integrable, then it can be expressed as follows:

$$f(x) = \sum_{k \in Z} c_k \phi_{ok}(x) + \sum_{j < J, k \in Z} d_{jk} \psi_{jk}(x) \tag{25}$$

where $c_k = \int f(x)\phi_{ok}(x)\,\mathrm{d}x$, integrable over $\mathbb{R}$, $d_{jk} = \int f(x)\psi_{jk}(x)\,\mathrm{d}x$, integrable over $\mathbb{R}$, parameter J controls the maximum resolution, j denotes the decomposition stage, k denotes the translation in time domain, and Z is the set of integers. $\psi_{jk}(x)$ are called wavelets, derived from a single mother wavelet ψ according to a relation

$$\psi_{jk}(x) = 2^{j/2}\psi\left(2^j x - k\right) \tag{26}$$

The mother wavelet is chosen properly to make sure the family $\{\psi_{jk}(x)\}$ forms an orthogonal basis for $L^2\{\mathbb{R}\}$. Similarly, the functions $\phi_{ok}(x)$ are derived from a function ϕ called the father wavelet or scaling function, by dilation and translation

$$\phi_{ok}(x) = 2^{o/2}\phi(2^o x - k) \tag{27}$$

$\psi_{jk}(x)$ and $\phi_{ok}(x)$ are the time-frequency atoms of the DWT. Typically, wavelets are well localised in time and frequency. To achieve this, wavelets are usually compactly supported in either time or frequency, but never both to follow the uncertainty principle. From the filter bank point of view using multiresolution analysis, c_k refers to the approximate coefficients at level o and d_{jk} refers to the detail coefficients at level j [70]. Let $h(n)$ and $g(n)$ are the discrete finite impulse response functions of $\phi(x)$ and $\psi(x)$, then two-scale relationships exist for the time-frequency atoms [68]

$$\begin{aligned} \phi(x) &= \sum_{n \epsilon C} h(n)\phi(2x - n) \\ \psi(x) &= \sum_{n \epsilon C} g(n)\psi(2x - n) \end{aligned} \tag{28}$$

The filterbank implementation of DWT is shown in Fig. 10, where LPF ($H(w)$) and HPF ($G(w)$) represent Fourier counterparts of finite support filters.

Filtering followed by sub-sampling constitutes one level of decomposition. The detail coefficients $d(k)$ and approximate coefficients $a(x)$ are obtained as follows [71]:

$$\begin{aligned} d(k) &= \sum_n x(n) \cdot g(2k - n) \\ a(k) &= \sum_n x(n) \cdot h(2k - n) \end{aligned} \tag{29}$$

For inverse DWT, an inversely symmetrical approach is used. The above theory of forward and reverse WT is combined with the concept of thresholding to denoise the noisy ECG signal $\tilde{x}(n)$.

2.3.1 Architecture of wavelet-based ECG denoiser: The wavelet-based ECG denoiser aims at estimating the clean signal

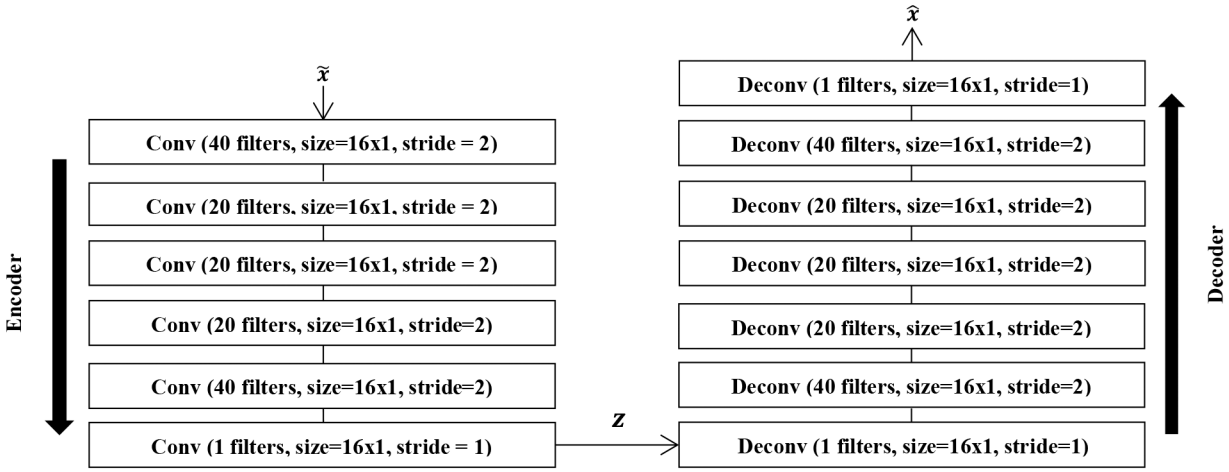

Fig. 8 *Architecture of FCN-based DAE [19]*

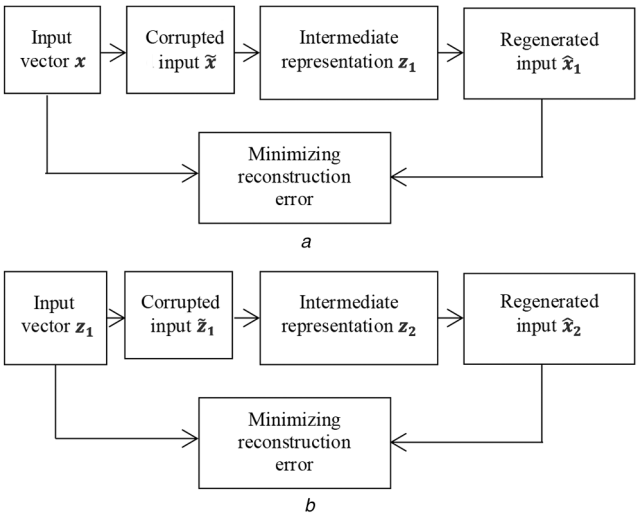

Fig. 9 *Architecture of DNN-based CDAE [59]*
(a) CDAE for hidden layer 1, ***(b)*** CDAE for hidden layer 2

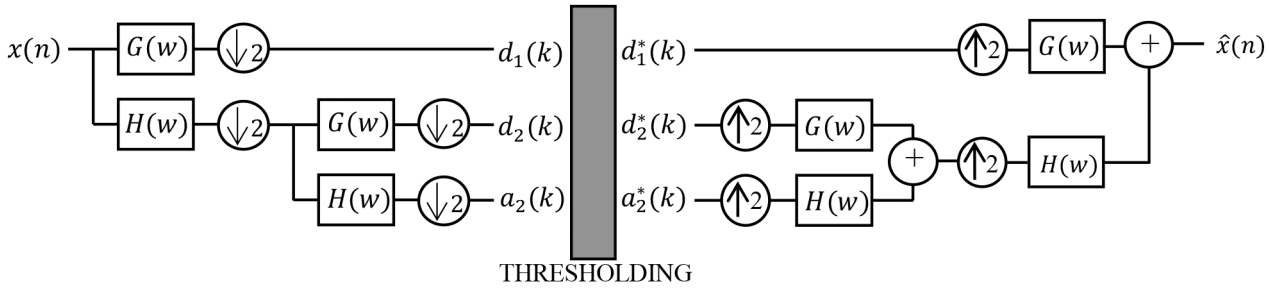

Fig. 10 *Scale-2 filter bank structure for DWT[69, 70]*

Table 3 Mother wavelets mostly used in ECG signal denoising

Author	Mother wavelet	Noise acted upon
Lin *et al.* [72]	Symlets5 (Sym5)	EMG artefact
Sawant and Patti [73]	Biorthogonal 2.4	EMG artefact
	Db9	PLI
Seljuq *et al.* [74]	Daubechies9 (Db9)	ECG noise
Li *et al.* [75]	Sym6	ECG noise
El-Dahshan [76]	Db7 or Db8	low-level ECG noise
	Sym5 or Sym6	high-level ECG noise
Biswas *et al.* [77]	Bior 5.5	PLI

$x(n)$ in the observed signal $\tilde{x}(n)$ with the help of WT, and inverse WT. Conventionally, forward WT is expressed as follows:

$$\boldsymbol{w} = \boldsymbol{W}\tilde{\boldsymbol{x}} \tag{30}$$

where $\boldsymbol{W}$ is the forward transform matrix and $\boldsymbol{w}$ is the set of wavelet coefficients (detail and approximation), but (30) in the filter bank approach is carried out by iterative decomposition of the LF bands. In the decomposition process, down-sampling is used as shown in Fig. 10. Effective estimation of the signal $\boldsymbol{x}$ requires proper selection of the mother wavelet. Different mother wavelets help in proper representation of the signal in wavelet domain, some of which have been listed in Table 3.

The next step is to shrink the coefficients obtained. Shrinking of wavelet coefficients is advised to produce estimates with desirable properties [78]. The wavelet estimators are minimax for a wide range of loss functions and the general class of functions. The estimators are simple, practical, fast, adaptable to spatial and spectral inhomogeneities, and readily extendable to high dimensions [78]. Shrinking is performed by thresholding the wavelet coefficients. Thresholding converts $\boldsymbol{w}$ to $\boldsymbol{w}^*$. It can be of two types.

Hard-thresholding: Given a wavelet coefficient w and a threshold $t > 0$, the hard-thresholded value is given by

$$T_{\text{hard}}(w;t) = \begin{cases} w, & |w| \geq t \\ 0, & |w| < t \end{cases} \tag{31}$$

Soft-thresholding: The soft-thresholded value is given by

$$T_{\text{soft}}(w;t) = \begin{cases} w - t, & w \geq t \\ w + t, & w \leq -t \\ 0, & |w| < t \end{cases} \tag{32}$$

Hard-thresholding retains the sharpness of the original signal due to its discontinuous characteristic, but it is susceptible to HF noises. On the other hand, soft-thresholding tends to distort the reconstructed signal, due to its smooth nature [2]. Different thresholding techniques have been discussed in Table 4.

Finally, in the conventional approach, the inverse transform is performed to get the following:

$$\hat{\boldsymbol{x}} = \boldsymbol{W}^{\mathrm{T}}\boldsymbol{w}^* \tag{33}$$

where $\hat{\boldsymbol{x}}$ is the estimate of $\boldsymbol{x}$ at the points $\{n\}$, but (33) is carried out by iterative reconstruction in the filter bank approach, where upsampling is used as shown in Fig. 10. Algorithm for wavelet-based ECG denoising is given in Fig. 11.

2.3.2 Methodologies: Based on the common thresholding schemes mentioned in Table 4, researchers have developed different thresholding methods to denoise ECG signals. Smith *et al.* [90] used polynomial threshold operators to obtain the polynomial coefficients, optimised utilising the least-square minimisation. Tulsani and Gupta [91] used the bee colony algorithm to upgrade the polynomial coefficients. Lin *et al.* [72] used soft-thresholding for feature extraction followed by sqtwolog rule to remove MA. Sawant and Patti [73] used soft-thresholding followed by heusure rule for EMG artefact removal and hard-thresholding followed by sureshrink rule for PLI removal. Zhang *et al.* [92] used the Mallat-algorithm-based wavelet decomposition followed by composite thresholding, i.e. hard- and soft-thresholding at different DWT levels to deal with ECG noise. In [93], a novel stationary WT-based filtering strategy has been proposed. Alfaouri and Daqrouq [94] have proposed a mechanism to denoise ECG signals, in which the error between the noisy and the original detail coefficients helps to determine the threshold for a five-level WT using Daubechies as the mother wavelet. Zhang *et al.* [95] used a blend of soft- and hard-thresholding to denoise ECG signal, the combining ratio is an element of the greatest and the base values of the obtained wavelet coefficients. Mithun *et al.* [96] denoised ECG signals using DWT with Discrete Meyer (Dmey) as the wavelet base in an adjustable fashion, adjustable utilising a parameter to switch hard- and soft-thresholding modes. This parametrised thresholding function takes the advantages of both hard- and soft-thresholding functions into consideration. Singh and Tiwari [97] have proposed a procedure for optimal selection of mother wavelet to denoise ECG while retaining the peaks of the signal. A newly improved wavelet thresholding (NIWT) method [98] uses a thresholding function based on the sigmoid function, and the universal threshold to denoise ECG. The non-decimated lifting transform (NLT) [99] is based on a lifting algorithm that removes one non-decimated second generation wavelet coefficient at a time. The universal hard-thresholding wavelet shrinkage (UHTWS) method [100] applies a hard-thresholding function to the wavelet coefficients when considering the universal threshold. Mallaparapu *et al.* [101] created a thresholding function by combining the hard and non-negative garrote functions to denoise the ECG signal. Wang *et al.* [102] made a paradigm shift from single-wavelet non-negative-garrote-based thresholding to a double-wavelet scheme to tackle ECG signal noise. Patil and Holambe [103] used an adaptive method that utilised the level-dependent minimax rule to denoise the ECG signal. Yi and Song [104] demonstrated the use of an innovative thresholding function based on level-dependent sqtwolog rule to tackle ECG noise with the aid of actual signal situations to switch hard- and soft-thresholding modes. Poornachandra [105] has formulated the subband level-dependent median (S-median) threshold based on DWT for the recuperation of the ECG signals tainted by the noises. Awal *et al.* [70] have advanced the modified S-median threshold with an extra level-dependent adjustment factor based on the optimal wavelet function and the optimal decomposition level to adapt to the CN in the ECG signals. Kumar *et al.* [106] have proposed a joint approach for denoising, detection, compression, and wireless transmission of

Table 4 Thresholding techniques mostly used in ECG signal denoising

Author	Method	Method description
Donoho and Johnstone [79]	exact minimax	this method uses pre-determined thresholds to minimise a constant term in the upper bound for the minimax risk of estimating a function using a thresholding estimator
Donoho and Johnstone [79]	universal	$T_{\mathrm{UV}} = \sqrt{(2\log n)}\hat{\sigma}$, where n is the number of data points and $\hat{\sigma}$ is an estimate of the noise level σ
Donoho and Johnstone [79] and Stein [80]	sure	the SureShrink chooser species a threshold value t_j for each resolution level j in a WT
Nason [81, 82], Neumann and Spokoiny [83], Weyrich and Warhola [84], Wang [85] Donoho and Johnstone [86]	cross-validation	it is a technique that attempts to minimise the prediction error generated by comparing a prediction based on a subset of the data and comparing it to the remainder of the data
Abramovich and Benjamini [87]	false-discovery rate	instead of choosing a threshold a wavelet coefficient is either kept or discarded in the decomposition of the noisy data if a hypothesis test decides that the coefficient is non-zero or zero
Vidakovic [88]	Bayesian methods	a Bayesian approach is adopted to estimate the wavelet coefficients θ_{jk} in $w_{jk} \sim N(\theta_{jk}, \sigma^2)$. Prior distributions for σ^2 and θ are obtained with the following assumed conditional distribution: $w\|\theta - f(w\|\theta)\|$ after observing w (the noisy wavelet coefficient), hypothesis H_0: $\theta = 0$ is tested versus H_1: $\theta \neq 0$. If this hypothesis is rejected, then θ is estimated by w
Ogden [89]	Ogden's methods	selection thresholding is based on level-by-level hypothesis testing of coefficients. Data-analytic thresholding is based on observation of plots of cumulative sums of the squares of the coefficients at a particular level

1. Input: $\tilde{x}(n)\ s.t. \tilde{x}(n) = x(n) + \in_n$, where $\{\in_n\}$ is some noise process with variance σ^2, where n is the index of $\tilde{x}(n)$
2. $d_l(k)$ and $a_L(k)$, detail coefficients and approximate coefficients, are obtained. The coefficients are collectively denoted as $w(k)$
3. Shrinking is performed by thresholding to convert $w(k)$ to $w^*(k)$
4. Finally, the inverse wavelet transform is taken on $w^*(k)$ to estimate the signal $x(n)$ in the noisy environment, using the filter bank architecture

5 Output: $\hat{x}(n)$

Fig. 11 *Algorithm 5: denoising algorithm of wavelet denoiser [68, 69]*

ECG signals using a modified Bior 3.1 WT. Kumar *et al.* [107] have proposed a new power and hardware efficient filter bank architecture that uses a cascade of three LPFs to denoise ECG signals. In [108], undecimated WT (UWT) is used for ECG denoising. The main principal of the method is signal decomposition using stationary WT. In [109], empirical WT (EWT) is used to denoise ECG signals. In [110], multi-adaptive bionic WT has been used to denoise ECG signals. In [111], the wavelet-based variational Bayesian estimation theory for signal denoising is used. DWT with an adaptive dual threshold filter (ADTF) is an algorithm based on three steps of denoising, viz., the DWT decomposition, the ADTF step, and the highest peaks correction step [112].

2.4 Sparsity-based models for ECG signal denoising

In this category of denoising, the ECG signal is denoised based on the sparse decomposition of the signal. The signal is fragmented into segments, and every segment is broken into sparse parts and residues. Then, these sparse parts are used to estimate clean signals, when the useful information in the signals is considered to be sparse [113]. The segments are decomposed using a non-linear optimisation method to find the sparsest representation.

2.4.1 Architecture of sparse ECG denoiser: In this model, the sparsity property of the ECG signal is utilised. Here, $\boldsymbol{x}$ is sparse and the denoiser aims to estimate $\boldsymbol{x}$ in $\tilde{\boldsymbol{x}}$ by drawing $\hat{\boldsymbol{x}}$ as close to $\boldsymbol{x}$ as possible. The task is to find a sparse signal $\hat{\boldsymbol{x}} = \boldsymbol{D}\hat{\boldsymbol{\alpha}}$ from a sparse dictionary that approximates the observed signal as closely as possible, where the sparsity constraint $\hat{\boldsymbol{\alpha}}$ is given in (35). This sparse signal is the perfect estimation of the input ECG signal, making the loss L minimum, where

$$L = \frac{1}{2} \| \tilde{\boldsymbol{x}} - \boldsymbol{D}\boldsymbol{\alpha} \|_2^2 + \mu \| \boldsymbol{\alpha} \|_0 \tag{34}$$

Here, μ is a sparseness-balancing regularisation parameter, and $\| \boldsymbol{\alpha}_0 \|$ is $\boldsymbol{\alpha}'s$ zero norm. It represents the number of non-zero values in the vector $\boldsymbol{\alpha}$. The first step in sparsity-based denoising methods is to construct a dictionary $\boldsymbol{D}$. The selection of a proper sparse dictionary is detrimental to the denoising process as it directly affects it. Every segment of the signal uses different linear combinations of atoms from the dictionary. Zhu and Li [113] have investigated denoising on several standard dictionaries in terms of symmetry, anti-symmetry, and similarities amongst ECG signals. In [114], Fourier and wavelet coefficients were presented as transform bases for smooth signals. Also, Peyre [115] has put forward several dictionaries consisting of orthogonal bases, which can help find an optimal orthogonal base for the signal. Once the dictionary is known, (34) helps to set $\boldsymbol{\alpha}$ as an unknown parameter. Thus, the estimate of $\boldsymbol{x}$ is computed as follows:

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \| \boldsymbol{\alpha} \|_0 \quad \text{s.t.} \quad \| \tilde{\boldsymbol{x}} - \boldsymbol{D}\boldsymbol{\alpha} \|_2 < \Lambda, \quad \boldsymbol{x} = \boldsymbol{D}\hat{\boldsymbol{\alpha}} \tag{35}$$

Finding the sparsest representation of a signal in the dictionary is a non-deterministic polynomial problem. There are two ways to solve the above problem, viz., greedy pursuit algorithms, and convex relaxation algorithms. The first type of algorithms attempt to reduce the error between the signal and its approximation iteratively, which are discussed in [116, 117]. The second converts the non-deterministic polynomial problem into a linear programming problem, such as the basic pursuit algorithm, as explained in [118]. Algorithm for sparsity-based ECG denoising is given in Fig. 12.

2.4.2 Methodologies: Wang *et al.* [28] have proposed two algorithms of sparse optimisation in the process of ECG denoising using MM technique. The first one is based on symmetric penalty sparse difference model, and the other is a mix of both symmetric and asymmetric penalty sparse difference models. In the symmetric penalty sparse difference model, the estimate for the original ECG with a sparse K-order derivative, $\hat{\boldsymbol{x}}$ is given by

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \left\{ F(\boldsymbol{x}) = \frac{1}{2} \| \boldsymbol{H}(\tilde{\boldsymbol{x}} - \boldsymbol{x}) \|_2^2 + \sum_{i=1}^{K} \lambda_i \sum_{n=0}^{N_i - 1} \phi([\boldsymbol{D}_i \boldsymbol{x}]_n) \right\} \tag{36}$$

where $\tilde{\boldsymbol{x}} = \boldsymbol{x} + \boldsymbol{f} + \boldsymbol{n}$, $\boldsymbol{x}$ is the ECG signal, $\boldsymbol{f}$ is the low-pass BW signal, $\boldsymbol{n}$ is the additive stationary white Gaussian noise with variance σ^2, λ_i is the ith difference order (regulation parameter), $\boldsymbol{D}_i$

1. Input: $\tilde{\boldsymbol{x}}$, where $\tilde{\boldsymbol{x}}$ is the signal observed in the noisy environment
2. Find $\hat{\boldsymbol{x}} = \boldsymbol{D}\hat{\boldsymbol{\alpha}}$, where $\boldsymbol{D}$ is a sparse dictionary constructed to suit $\boldsymbol{x}$, $\hat{\boldsymbol{\alpha}}$ is the sparsity constraint, and $\boldsymbol{\alpha}$ is an unknown parameter set according to:
$$L = \frac{1}{2}\|\tilde{\boldsymbol{x}} - \boldsymbol{D}\boldsymbol{\alpha}\|_2^2 + \mu\|\boldsymbol{\alpha}\|_0$$
Such that L is as low as possible
3. Find the estimate of $\boldsymbol{x}$ according to the following:
$\hat{\boldsymbol{\alpha}} = arg\ \min_{\boldsymbol{\alpha}}\|\boldsymbol{\alpha}\|_0\ s.t.\|\tilde{\boldsymbol{x}} - \boldsymbol{D}\boldsymbol{\alpha}\|_2 < \Lambda, \quad \boldsymbol{x} = \boldsymbol{D}\hat{\boldsymbol{\alpha}}$, where Λ is a small-value limiting parameter
4. Solve the above non-deterministic polynomial problem by using greedy pursuit or convex relaxation
5. Output: $\hat{\boldsymbol{x}}$

Fig. 12 *Algorithm 6: denoising algorithm of sparsity-based denoiser [113]*

1. Input: $\tilde{\boldsymbol{x}}, \boldsymbol{A}, \boldsymbol{B}, regularization\ parameter\ (\lambda_i), i = 0, \ldots, K$
2. $\boldsymbol{b} = \boldsymbol{B}^T\boldsymbol{B}\boldsymbol{A}^{-1}\tilde{\boldsymbol{x}}$
3. $\boldsymbol{x} = \tilde{\boldsymbol{x}}$ (initialization)
Repeat
4. $|\Lambda_i|_{n,n} = \frac{\Phi'(|\boldsymbol{D}_i\boldsymbol{x}|_n)}{|\boldsymbol{D}_i\boldsymbol{x}|_n}, i = 0, \ldots, K$
5. $\boldsymbol{I} = \sum_{i=0}^{K}\lambda_i\boldsymbol{D}_i\Lambda_i\boldsymbol{D}_i$
6. $\boldsymbol{E} = \boldsymbol{B}^T\boldsymbol{B} + \boldsymbol{A}^T\boldsymbol{I}\boldsymbol{A}$
7. $\boldsymbol{x}^{(k+1)} = \boldsymbol{A}\left[\boldsymbol{E}^{(k)}\right]^{-1}\boldsymbol{b}$
If the convergence conditions are met, output $\boldsymbol{x}$, otherwise, $k = k + 1$, and go through steps 4-7
8. $\boldsymbol{H} = \boldsymbol{B}\boldsymbol{A}^{-1}$
$\boldsymbol{f} = \tilde{\boldsymbol{x}} - \boldsymbol{x} - \boldsymbol{H}(\tilde{\boldsymbol{x}} - \boldsymbol{x})$
Output: $\boldsymbol{x}, \boldsymbol{f}$

Fig. 13 *Algorithm 7: symmetric penalty sparse difference algorithm (SPSD) [28]*

1. Input: $\tilde{\boldsymbol{x}}, r \geq 1, \boldsymbol{A}, \boldsymbol{B}, regularization\ parameter\ (\lambda_i), i = 0, \ldots, K$
2. $[\boldsymbol{b}]_n = \frac{1-r}{2}$
3. $\boldsymbol{z} = \boldsymbol{B}^T\boldsymbol{B}\boldsymbol{A}^{-1}\tilde{\boldsymbol{x}} - \lambda_0\boldsymbol{A}^T\boldsymbol{b}$
4. $\boldsymbol{x} = \tilde{\boldsymbol{x}}\ (initialization)$
Repeat
5. $[\boldsymbol{\Gamma}]_{n,n} = \begin{cases} \frac{1+r}{4|x_n|}\ ; |x_n| \geq \rho \\ \frac{1+r}{4\rho}\ ; |x_n| \leq \rho \end{cases}, [\Lambda_i]_{n,n} = \frac{\Phi'(|\boldsymbol{D}_i\boldsymbol{x}|_n)}{|\boldsymbol{D}_i\boldsymbol{x}|_n}, i = 0, \ldots, K$
6. $\boldsymbol{P} = \boldsymbol{B}^T\boldsymbol{B} + \boldsymbol{A}^T\boldsymbol{K}\boldsymbol{A}$
7. $\boldsymbol{K} = 2\lambda_0\boldsymbol{\Gamma} + \sum_{i=1}^{K}\lambda_i(\boldsymbol{D}_i)^T\Lambda_i\boldsymbol{D}_i$
8. $\boldsymbol{x} = \boldsymbol{A}\boldsymbol{P}^{-1}\boldsymbol{z}$
Until convergence
9. $\boldsymbol{H} = \boldsymbol{B}\boldsymbol{A}^{-1}, \boldsymbol{f} = \tilde{\boldsymbol{x}} - \boldsymbol{x} - \boldsymbol{H}(\tilde{\boldsymbol{x}} - \boldsymbol{x})$
$Output: \boldsymbol{x}, \boldsymbol{f}$

Fig. 14 *Algorithm 8: compound penalty sparse difference algorithm (CPSD) [28]*

is the ith order difference matrix, ϕ is the penalty function and the number of constraints of punishment depends on K, and $\boldsymbol{H}$ is the high-pass filter matrix. The above equation could promote sparsity of both $\boldsymbol{x}$ and its K-order derivative (difference). The high-pass filter is the same as the equation $\boldsymbol{H} = \boldsymbol{B}^{-1}\boldsymbol{A}$, where $\boldsymbol{A}$ and $\boldsymbol{B}$ are banded convolution matrices representing LTI systems. $\boldsymbol{H}$ represents a cascade of LTI systems. The function $\phi_A = |x|$ leads to the problem of l_1-norm regularisation, i.e. the function is not differentiable at zero. To overcome this issue, a differentiable approximation of the l_1 penalty function is used. It is given as follows:

$$\phi_B(x) = |x| - \rho\log(|x| + \rho) \tag{37}$$

where the constant $\rho > 0$ should be small enough to ensure that the penalty function is smooth and the original non-differentiable penalty function is sparse ($\phi = \phi_B$). Hence, the symmetric penalty sparse difference algorithm is given in Fig. 13.

When the signal $\boldsymbol{x}$ is asymmetrically sparse, asymmetric penalty sparse difference model is used, hence asymmetric penalty is used for punishment, and θ is defined as

$$\theta(x_n; r) = \begin{cases} x, & x \geq 0 \\ -rx, & x < 0 \end{cases} \tag{38}$$

where $r \geq 1$ is a positive quantity. However, the above equation is not differentiable at $x = 0$. Hence, a different form is opted for, given as follows:

$$\theta_p(x_n; r) = \begin{cases} x, & x > \rho \\ f(x), & |x| \leq \rho \\ -rx, & x < -\rho \end{cases} \tag{39}$$

where $f(x)$ is found to be

$$f(x) = \frac{1+r}{4\rho}x^2 + \frac{1-r}{2}x + \frac{(1+r)\rho}{4}, \quad x \leq \rho \tag{40}$$

r and \rho mean the same as defined before.

When positive ECG signal peaks are taken into consideration, asymmetric penalty is used for the original signal, and symmetric penalty is used for the difference signal. This algorithm is called compound penalty sparse difference algorithm. This method ensures enhanced sparsity and enhanced ECG details. In this case, the estimate $\hat{\boldsymbol{x}}$ is given as follows:

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} F(\boldsymbol{x}) = \frac{1}{2}\parallel \boldsymbol{H}(\tilde{\boldsymbol{x}} - \boldsymbol{x}) \parallel_2^2 + \lambda_0\sum_{n=0}^{N-1}\theta_p(x_n; r) + \sum_{i=1}^{K}\lambda_i\sum_{n=0}^{N_i-1}\phi([\boldsymbol{D}_i\boldsymbol{x}]_n)\Bigg\} \tag{41}$$

where $\phi = \phi_B$ for the optimal solution $\boldsymbol{x}$ and λ_0 is the 0-order difference and $\theta_p(x_n; r)$ is described using (39) and (40). The above scheme is discussed in Algorithm 8 (see Fig. 14).

Mourad [119] with the help of Wang *et al.* [120] and Harris and Yuan [121] has developed a new algorithm to extract QRS complex from noisy ECG that works on the principle of group-sparsity maximisation (MM technique) and singular spectrum analysis (SSA). In [119], another version of the SSA procedure (GSLPF) is also considered for denoising. In [122], a dictionary learning based sparse representation (DLSR) framework is presented, where a frequency-localised sparse representation is used to remove BW and PLI. For AWGN and MA, a time-localised sparse representation is used for denoising ECG signals. A conventional algorithm known as sparsity-assisted signal smoothing (SASS) is used to denoise ECG signals [123]. Devi *et al.* [124] have compared two sparsity-based adaptive filtering techniques, viz., least mean squares (LMS), and recursive least squares (RLS). LMS predicts the optimum filter coefficients by reducing the error signal in a LMS manner, whereas RLS does it in a recursive manner by reducing the weighted linear least-squares error function. Jin *et al.* [125] have proposed the l_1-norm based sparsity recovery algorithm. Here, the classical sparsity recovery methods, like the l_1-norm penalised least squares method, have been applied directly to boost sparsity recovery. A sparse signal $x(n)$ is observed in a noisy environment, where $\boldsymbol{x} = \boldsymbol{A}\hat{\boldsymbol{x}}$, $\boldsymbol{A}\ \epsilon\ \mathbb{R}^{r\times c}$ is the basis vector of the signal space, and $\hat{\boldsymbol{x}}\ \epsilon\ \mathbb{R}^c$ is the representation coefficients, where

$c = N$, and $r = N$. The problem can be formulated as an l_1-norm optimisation problem given below:

$$\min_{\hat{\boldsymbol{x}}} \| \hat{\boldsymbol{x}} \|_1 \quad \text{s.t.} \quad \boldsymbol{x} = \boldsymbol{A}\hat{\boldsymbol{x}} \tag{42}$$

An approximate solution to the above equation can be found by using the basic pursuit denoising problem [118]

$$\arg\min_{\hat{\boldsymbol{x}}} \left\{ \frac{1}{2} \| \tilde{\boldsymbol{x}} - \boldsymbol{A}\hat{\boldsymbol{x}} \|_2^2 + \lambda \| \hat{\boldsymbol{x}} \|_1 \right\} \tag{43}$$

where λ is the regularisation parameter. The detailed steps are given in Algorithm 9 (see Fig. 15).

Since the classical sparsity recovery method based on l_1-norm penalty leads to underestimated peaks, a different method has been proposed by Jin *et al.* [125] that uses a novel penalty for sparsity recovery, which keeps the convexity of the least squares cost function minimised and avoids underestimation of the values of the waveform. By replacing the l_1-norm penalty with the generalised minimax concave (GMC) penalty, the sparsity problem takes the following form:

$$\arg\min_{\hat{\boldsymbol{x}}} \left\{ \frac{1}{2} \| \tilde{\boldsymbol{x}} - \boldsymbol{A}\hat{\boldsymbol{x}} \|_2^2 + \lambda \psi_{\boldsymbol{B}}(\hat{\boldsymbol{x}}) \right\} \tag{44}$$

where $\psi_{\boldsymbol{B}}(x)$ is the GMC penalty function defined as follows:

$$\psi_{\boldsymbol{B}}(\hat{\boldsymbol{x}}) = \| \hat{\boldsymbol{x}} \|_1 - S_{\boldsymbol{B}}(\hat{\boldsymbol{x}}) \tag{45}$$

where $\boldsymbol{B} \in \mathbb{R}^{r \times c}$ and $S_{\boldsymbol{B}} \triangleq \inf_{\boldsymbol{v} \in \mathbb{R}^c} \{ \| \boldsymbol{v} \|_1 + \frac{1}{2} \| \boldsymbol{B}(\hat{\boldsymbol{x}} - \boldsymbol{v}) \|_2^2 \}$ is the generalised Huber function [123], $\inf\{ . \}$ denotes the infimum of a function. Here, $c = N$ and $r = N$. Hence, the reformed optimisation problem is as follows:

$$\left(\hat{\boldsymbol{x}}^{\text{opt}}, \boldsymbol{v}^{\text{opt}}\right) = \arg\min_{\hat{\boldsymbol{x}}} \max_{\boldsymbol{v}} F(\hat{\boldsymbol{x}}, \boldsymbol{v}) \tag{46}$$

where

$$F(\hat{\boldsymbol{x}}, \boldsymbol{v}) = \frac{1}{2} \| \tilde{\boldsymbol{x}} - \boldsymbol{A}\hat{\boldsymbol{x}} \|_2^2 + \lambda \| \hat{\boldsymbol{x}} \|_1 - \lambda \| \boldsymbol{v} \|_1 - \frac{\gamma}{2} \| \boldsymbol{A}(\hat{\boldsymbol{x}} - \boldsymbol{v}) \|_2^2$$

with $0 \leq \gamma \leq 1$, where γ is the convexity-guaranteeing factor. The above is a saddle-point optimisation problem that can be solved according to Algorithm 10 (see Fig. 16). It can be inferred from [125] that SASS converges faster than Algorithms 9 and 10 (Figs. 15 and 16), and that it performs better than Algorithm 9 (Fig. 15) in terms of R-pecks, but underestimates Q-peaks.

The sparse representation-based denoising (SRD) scheme proposed by Satija *et al.* [126] uses elementary sine and cosine waveforms for generation of dictionaries corresponding to noises and ECG components, hence allowing for an automatic dictionary learning based on the noise type.

2.5 Architecture of ECG Bayesian-filter-based denoiser

This denoising category investigates different model-based methods used to denoise ECG signals. Model-based methods work on the principle of estimation of hidden states of an underlying model with the help of the estimation theory. These hidden states are observed through a set of measurements. One such method is the Kalman filter (KF). The basic KF assumes a linear model for the system dynamics and observation equations, but most systems are non-linear. Several versions of the original KF have been developed, viz., EKF, EKS, and UKF. Sameni *et al.* [23] give an overview of the extended versions of the conventional KF.

2.5.1 Architecture of ECG Bayesian-filter-based denoiser: This section discusses the non-linear Bayesian framework for ECG signal denoising. The first step is the modification of the dynamic ECG model. McSharry *et al.* [127] have proposed a non-linear dynamic model with the following dynamic state equations in the Cartesian coordinates:

$$\begin{cases} \dot{x} = \rho x - wy \\ \dot{y} = \rho y + wx \\ \dot{z} = \sum_{i \in P, Q, R, S, T} a_i \Delta\theta_i e^{-(\Delta\theta_i^2 / 2b_i^2)} - (z - z_0) \end{cases} \tag{47}$$

where $x, y,$ and z are the state variables, $\rho = 1 - \sqrt{x^2 + y^2}$, $\Delta\theta_i = (\theta - \theta_i) \bmod (2\pi)$, and $\theta = a\tan2(y, x)$ is the four quadrant arctangent of the elements of x and y, with $-\pi \leq a\tan2(y, x) \leq \pi$. $a_i, b_i,$ and θ_i correspond to the amplitude, width, and centre parameters of the above equation's Gaussian terms. In this model, the BW of the ECG signal is modelled with the parameter z_0. $P, Q, R, S,$ and T are the components of the ECG waveforms. Each of the P, Q, R, S, and T waves of the ECG waveform is modeled with a Gaussian function and is located at a specific angular position θ_i. The new form of the dynamic equations in polar coordinates is as follows [128]:

$$\begin{cases} \dot{r} = r(1 - r) \\ \dot{\theta} = \omega \\ \dot{z} = \sum_{i \in P, Q, R, S, T} \frac{\alpha_i \omega}{b_i^2} \Delta\theta_i e^{-(\Delta\theta_i^2 / 2b_i^2)} - (z - z_0) \end{cases} \tag{48}$$

where $\{\alpha_i\}$ are the peak amplitudes of the Gaussian functions used for modelling each of the ECG components, and r and θ are the radial and angular state variables in polar coordinates. The first differential equation is redundant as the other two equations in (48)

1. Input : $\hat{\boldsymbol{x}}^{(0)}$, $\boldsymbol{A}$, $\boldsymbol{A}^{\boldsymbol{T}}$, λ
2. Set ρ = maxeig ($\|\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}\|$), where maxeig (·) denotes the maximum eigenvalue of a matrix
3. Set $\mu : 0 < \mu < {}^{1}/_{\rho}$
4. Number of iterations: N_{iter}
   For $i = 0$ *to* N_{iter} do
   $\boldsymbol{w}^{(i)} = \hat{\boldsymbol{x}}^{(i)} - \mu\boldsymbol{A}^{\boldsymbol{T}}\left(\boldsymbol{A}\hat{\boldsymbol{x}}^{(i)} - \tilde{\boldsymbol{x}}\right)$
   $\hat{\boldsymbol{x}}^{(i+1)} = soft(\boldsymbol{w}^{(i)}, \mu\lambda)$
   End for
5. Return $\hat{\boldsymbol{x}}^{(i+1)}$
6. Output :$\hat{\boldsymbol{x}}^{(i+1)}$

Fig. 15 *Algorithm 9: algorithm for l_1-norm penalised least squares problem [125]*

1. Input: $\hat{\boldsymbol{x}}^{(0)}, \boldsymbol{v}^{(0)}, \boldsymbol{A}, \boldsymbol{A}^T, \lambda$
2. Set $\rho = \max\{1, \gamma/(1-\gamma).maxeig(\boldsymbol{A}^T\boldsymbol{A})$
3. Set $\mu : 0 < \mu < {}^{2}/_{\rho}$
4. Number of iterations: N_{iter}
   For $i = 0$ *to* N_{iter} do
   $\boldsymbol{w}^{(i)} = \hat{\boldsymbol{x}}^{(i)} - \mu\boldsymbol{A}^T\left(\boldsymbol{A}\hat{\boldsymbol{x}}^{(i)} - \tilde{\boldsymbol{x}}\right) + \mu\gamma\boldsymbol{A}^T\boldsymbol{A}(\boldsymbol{v}^{(i)} - \hat{\boldsymbol{x}}^{(i)})$
   $\boldsymbol{u}^{(i)} = \boldsymbol{v}^{(i)} - \mu\gamma\boldsymbol{A}^T\boldsymbol{A}(\boldsymbol{v}^{(i)} - \hat{\boldsymbol{x}}^{(i)})$
   $\hat{\boldsymbol{x}}^{(i+1)} = soft(\boldsymbol{w}^{(i)}, \mu\lambda)$
   $\boldsymbol{v}^{(i+1)} = soft(\boldsymbol{u}^{(i)}, \mu\lambda)$
   end for
5. Return $\hat{\boldsymbol{x}}^{(i+1)}$
6. Output: $\hat{\boldsymbol{x}}^{(i+1)}$

Fig. 16 *Algorithm 10: algorithm for GMC penalised least squares problem [125]*

1. Input: x, y, z: the state variables
2. Conversion of the non-linear dynamic ECG model from Cartesian to polar
3. Linearization of the modified model
4. Observation of s_k and ϕ_k, where $\{s_k\}$ are the noisy ECG observations and $\{\phi_k\}$ are the phase observations
5. Prediction of model parameters prior to the implementation of the filter

Fig. 17 *Algorithm 11: denoising algorithm of Bayesian-filter-based denoiser [23]*

do not depend on r. Hence, it might be removed. This modification has the advantages of making the new model simpler and helping to locate the components of the waveform from the angular viewpoint due to the introduction of the phase parameter θ. The next step is to linearise the modified model. After linearising the model, it is necessary to make observations, viz., noisy ECG observations s_k, and phase observations ϕ_k.

The above observations are related to the state vector according to the following equation:

$$\begin{pmatrix}\phi_k\\ s_k\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\cdot\begin{pmatrix}\theta_k\\ z_k\end{pmatrix} + \begin{pmatrix}u_k\\ v_k\end{pmatrix} \tag{49}$$

where $R_k = E\left\{\begin{pmatrix}u_k\\ v_k\end{pmatrix}^{\mathrm{T}}(u_k v_k)\right\}$ is the observation noise co-variance matrix. Finally, it is needed to predict the model parameters prior to the implementation of the filter. For automatic parameter prediction, the parameters should be estimated from the signal itself. For this, any noisy ECG is transformed to a three-dimensional representation by plotting the noisy ECG versus the periodic phases that are assigned to each sample in polar coordinates on the unit circle. It is now possible to estimate the dynamic model parameters for the given ECG signal. For this, the mean and variance of the phase-wrapped ECG are calculated for all phases between 0 to 2π. This gives the average of the ECG waveform. Algorithm for ECG Bayesian-filter-based denoising is given in Fig. 17.

2.5.2 Methodologies: Hesar and Mohebbi [129] have proposed a marginalised particle EKF that overcomes the demerits of EKF and particle filter. This uses a novel combination of marginalised particle filter and EKF. At first, the generation of three-state polar EDM from two-state EDM is carried out by introducing angular velocity to the state model as an AR state giving rise to 'marginalised particle extended Kalman filter' (MP-EKF) algorithm. ECG parameter extraction is carried out using approaches similar to [23, 130]. First, the R-peaks of ECG cycles in the signal are detected and assumed to be positioned at $\theta = 0$. The ECG samples between two sequential R-peaks have a phase between 0 and 2π (or $-\pi$ and π). By estimating the mean and variance of each sample in the overlaid ECG cycles, a phase-wrapped ECG mean waveform (ECG(θ)) and a phase-wrapped ECG variance waveform (σECG(θ)) are built. A particle weighting strategy is considered that evaluates and weights the particles at each time step based on their distance to the noisy measurements and ECGsynth. ECGsynth is a synthetic ECG signal constructed using the feature parameters extracted from ECG(θ). A statistical distance metric called the 'Mahalanobis distance' is used as the closeness evaluation measure. This weighting strategy is used in each time step for each particle. Another method has been proposed by Haser and Mohebbi [131] to denoise ECG that proves to be better than the previous one. The denoising performance of 'Old MP-EKF' in the previous work benefited from a synthetic signal called 'ECGsynth' which had the same length as the original ECG signal and was built using the EDM characteristic parameters of that signal. To build this signal, linear phase observations were used but here in this algorithm, non-linear phase observations are used. The synthetic ECG signals made with the linear phase wrapping approach have temporal delays, especially when the angular frequency changes dramatically from one beat to another one. These delays would impair the performance of 'MP-EKF' especially in processing signals containing abnormal beats. On the other hand, the synthetic ECG signals built based on the non-linear phase wrapping approach have acceptable harmonic correspondences with their associated ECG signals. This fact helps 'New MP-EKF' to have better and acceptable results. Some linear Bayesian model-based benchmark methods for ECG denoising are EKF/EKS frameworks proposed in [23] which use linear phase observations ('EKF3 linear' and 'EKS3 linear') and EKF/EKS frameworks with non-linear phase observations [132] ('EKF3 non-linear' and 'EKS3 non-linear'), with the number of EDM states as 3. Akhbari *et al.* [132] discuss two EKF-based approaches, viz., EKF25 with linear phase observation, and EKF25 with non-linear phase observation. In the first approach, the first observation is the linearly approximated phase of the ECG signal, and the other observation is the overall summation of the estimated states providing an enhanced estimation of the overall cardiac beat. However, individual estimated ECG states (P-wave, QRS complex and T-wave) may have rising and falling trends resulting from inaccuracies in modelling dynamic baseline changes. The second approach is EKF25 with non-linear phase observations. This approach uses non-linear phase observations as linear phase observations are inaccurate, especially for signals with major RR-interval deviations or signals with frequent abnormal beats that appear irregularly in an unsteady fashion in the ECG. For constructing a better approximation of the ECG phase observation, Niknazar *et al.* [133] have used the dynamic time warping (DTW) method. The DTW method measures the similarity between two sequences which may vary in time or speed, to obtain an optimal match between the two with certain restrictions [134]. In [135], an adaptive Kalman filter (AKF) bank architecture is used in denoising ECG signals.

2.6 Hybrid methods for ECG signal denoising

For years, researchers have tried combining different denoising methodologies to denoise ECG signals to improve the performance of the basic denoising techniques, i.e. researchers have tried to combine techniques from different domains to denoise ECG signals to get better results in terms of $\mathrm{SNR_{imp}}$, RMSE, and PRD. This section of the paper discusses the successful hybrid approaches towards ECG signal denoising.

2.6.1 Methodologies: Rakshit and Das [136] have attempted to combine the known method of EMD, and the concept of adaptive switching mean filter (ASMF). In this algorithm, the advantages of both methods are exploited to reduce ECG signal noise. Conventional EMD rejects the initial IMFs or attempts at using a window-based approach for reducing HF noises, but in this scheme, a wavelet-based soft-thresholding method is used for the reduction of HF noises, and this is followed by an ASMF operation to enhance the quality. Singh and Pradhan [137] have proposed a novel ECG denoising approach based on a combination of variational mode decomposition (VMD), the NLM estimation, and the discrete WT (DWT) filtering technique. To achieve the effective ECG denoising goal, the noisy ECG signal is decomposed into narrow-band variational mode functions (VMFs) using VMD method, with the number of modes (k) being pre-defined. This parameter is empirically determined. The aim is to filter out noise from the narrow-band VMFs generated. To achieve that, the centre frequency information associated with each VMF is used to exclusively divide them into lower-frequency and higher-frequency signal groups. The higher frequency VMFs were filtered out using DWT-thresholding technique. The lower frequency VMFs are denoised through the NLM estimation technique. The non-recursive nature of VMD enables the parallel processing of NLM estimation and DWT filtering. The traditional DWT-based approaches need large decomposition levels to filter LF noises and at the same time, NLM technique suffers from the rare-patch effect in HF region. On the contrary, in that framework, both NLM and DWT approaches complement each other to overcome their

individual ill-effects. The signal reconstruction is performed using the denoised HF and LF VMFs.

Kumar *et al.* [20] have combined FFT, and an adaptive R-peak detection algorithm to denoise and detect ECG signals. Jain *et al.* [138] have proposed a new method for ECG denoising, which incorporates an EMD algorithm with Riegmann Liouvelle (RL) fractional integral filtering and Savitzky–Golay (SG) filtering. In this method, noisy ECG signal is decomposed into its IMFs, from which noisy IMFs, corrupted with HF and LF noises, are identified by noisy-IMF identification methodologies. To denoise the signal, RL fractional integral filtering and SG filtering are applied on noisy IMFs corrupted with HF and LF noises, respectively; ECG signal is reconstructed with denoised IMFs and remaining signal dominant IMFs to obtain noise-free ECG signal. In the work proposed by Rajankar and Talbar [139], ECG denoising is achieved using wavelet neural network by approximating signal to a maximum possible accuracy. The feed-forward back propagation neural network with ten neurons and two hidden layers is designed with conjugate gradient optimisation. The library mother wavelets such as Daubachies, Symlet, and so on, are used as the activation function of one hidden layer for ECG signal estimation. Kaergaard *et al.* [54] have also presented an analysis of DWT-NN to denoise ECG signals. Kabir and Shahnaz [140] have proposed a hybrid technique, in which the EMD-enhanced output is further given to the DWT-based denoiser. The DWT-based denoiser uses adaptive soft-thresholding with universal threshold, where the threshold value is adaptable to signal intensities. Kumar *et al.*[141] have combined NLM and EMD for noise cancelation. The proposed algorithm has four steps, viz., detection of the fiducial point (R-peak detection), differential standard deviation calculation, EMD framework, and finally NLM framework. Li and Li [142] use a combination of EMD and WT. The EMD-denoised signal is further improved using adaptive wavelet filtering, where the threshold is obtained to keep the error between $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$ as small as possible. The error is given as follows:

$$E(\hat{\boldsymbol{x}}, \boldsymbol{x}) = \frac{1}{N}\sum_{i=0}^{N-1} \| \hat{x}_i - x_i \|_2^2 \tag{50}$$

Generally, the first IMF in the EMD process corresponds to HF noise, which can be filtered out. However, the second IMF contains information pertaining to the QRS complex and is also noisy. This does not allow direct removal of the second IMF. A solution to this problem is given by Lu *et al.* [143]. The ECG dynamic model is an excellent tool to capture ECG morphological features. This model can be used to filter a noisy ECG by fitting (51) to an ECG beat, and the filtered signal can preserve much clinical information. The vertical displacement of ECG is determined by an ordinary differential equation given below:

$$z(\theta) = -\sum_{i \epsilon P,Q,R,S,T} a_i \frac{\Delta\theta_i}{b_i^2} e^{-(\Delta\theta_i^2/2b_i^2)} \tag{51}$$

The advantage of this model-based filtering lies in the physiological meaning behind the dynamic model. In practice, the ECG model $z(\theta)$ is fitted to a noisy signal $s(\theta)$ by minimising the squared error between them. That is, we need to find

$$\min_{a_i, b_i, \theta_i} \| z(\theta) - s(\theta) \|^2 \tag{52}$$

where $i \in \{P, Q, R, S, T\}$ and $\theta \in [0, 2\pi]$ and solve (52) by a gradient descent method in the parameter space. DWT has also been explored with VMD, and NLM separately for denoising ECG signals, giving rise to VMD-DWT [144], and NLM-DWT [145]. In [146], DWT is combined with a technique called HMM that identifies the generation process that is behind the observed data. This gives rise to a new method termed signal denoising by clustering and soft thresholding (SDCST). HMM identifies the detail wavelet coefficients coming from a noisy source. In [147], the pre-processing stage of the proposed ECG detector system uses band-pass filtering with the help of fractional order low-pass filters, high-pass filters, and fractional order derivative to denoise ECG. This puts forward a non-linear and non-differentiable problem, which is solved by cuckoo search algorithm. Yadav *et al.* [148] have proposed an algorithm to denoise ECG signals, which is a combination of NLM and WT. The proposed algorithm (NLWT) involves three sequential steps: the SDM extraction, the shrinkage of transform coefficients, and the aggregation.

3 Performance analysis of ECG signal denoising techniques

In this paper, various ECG denoising techniques have been compared for a better understanding of their effectiveness in the removal of the predominant noises that get added to pure ECG signals. Almost all the popular methods have been tested on two main ECG databases, viz., the MIT-BIH Arrhythmia Database (MIT-BIHA), and the MIT-BIH Normal Sinus Rhythm Database (MIT-BIH NSR). In some cases, the noise signals are taken from the MIT-BIH Noise Stress Test Database (MIT-BIH NST).

3.1 Test databases

The MIT-BIH Arrhythmia database has 48 records of 30 min each. These records were digitised with a sampling frequency of 360 Hz at a resolution of 11 bits per sample over a 10 mV range. The noises found in the records are BW, MA, PLI, and EM [149]. The MIT-BIH Normal Sinus Rhythm database includes 18 long-term ECG recordings of subjects not having any significant arrhythmias [150]. These two databases are used to analyse different denoising methods adopted by different researchers, but in some cases, noise is added to the records of the above two databases. Thus, noise signals are taken from the MIT-BIH Noise Stress Test database and added to the parent records of the above databases. This database includes 12 30 min ECG recordings and 3 30 min recordings of noise typical in ambulatory ECG recordings. The later three noise records are composed of the former 12 records by selecting intervals that contain BW, MA, and EM. These ECG signal recordings are created using two clean recordings (118 and 119) from the MIT-BIH Arrhythmia (MITBIHA) database, to which calibrated amounts of noise from a record containing EM are added [151]. ECG denoising techniques can also be tested on databases other than MIT-BIH like PTB [152], BIDMC [153], St. database, Apnea ECG, Fantasia, ADB [154], QT [155], CUDB [156], SCDB [157], and so on.

3.2 Performance parameters

There are three benchmark metrics for the analysis of different denoising methods. These are root-mean-square error (RMSE), percentage-root-mean-square difference (PRD), and improvement in signal-to-noise ratio (SNR_{imp}). RMSE is the root of the squared error difference between the denoised and original ECG signals. It is used for determining the variance between the output predicted by the denoising model and the actual signal. A smaller value of RMSE implies the better performance of the model. PRD computes the total distortion present in the denoised signal. A lower PRD represents a better quality of the denoised signal. SNR_{imp} is the improvement in the SNR levels between the input and the output. For the denoising model shown in Fig. 1*b*, the performance evaluation criteria can be defined as follows:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{n=0}^{N-1} \left[x(n) - \hat{x}(n)\right]^2} \tag{53}$$

$$\text{SNR}_{\text{imp}} = \text{SNR}_{\text{out}} - \text{SNR}_{\text{in}} \tag{54}$$

where SNR_{in} and SNR_{out} are as follows:

$$\text{SNR}_{\text{in}} = 10 \times \log_{10}\left(\frac{\sum_{n=0}^{N-1} [x(n)]^2}{\sum_{n=0}^{N-1} [\tilde{x}(n) - x(n)]^2}\right) \tag{55}$$

$$\mathrm{SNR}_{\mathrm{out}} = 10 \times \log_{10}\left(\frac{\sum_{n=0}^{N-1}[x(n)]^2}{\sum_{n=0}^{N-1}[\hat{x}(n) - x(n)]^2}\right) \quad (56)$$

$$\mathrm{PRD} = \sqrt{\frac{\sum_{n=0}^{N-1}[x(n) - \hat{x}(n)]^2}{\sum_{n=0}^{N-1}[x(n)]^2}} \times 100 \quad (57)$$

where N is the number of data points.

3.3 Performance analysis

A qualitative analysis is shown in Figs. 18–23. These figures show a pictorial comparison between six promising methods of ECG denoising, each falling into one of the six major denoising categories as discussed in Fig. 2. A quantitative analysis has been done on the methods with the noise type, $\mathrm{SNR}_{\mathrm{in}}$, the database type, and the record number of the considered database as the classifying parameters. Tables 5–7 discuss the effect of ECG denoising in the presence of AWGN on the MIT-BIH databases.

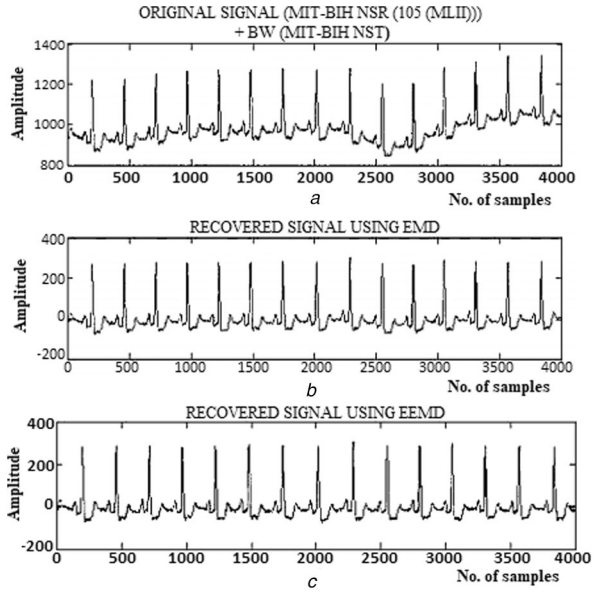

Fig. 18 *ECG denoising using EEMD*

(a) Original ECG (MIT-BIH NSR(105(MLII))) signal with BW, ***(b)*** Denoised ECG signal using EMD, ***(c)*** Denoised ECG signal using EEMD [22]

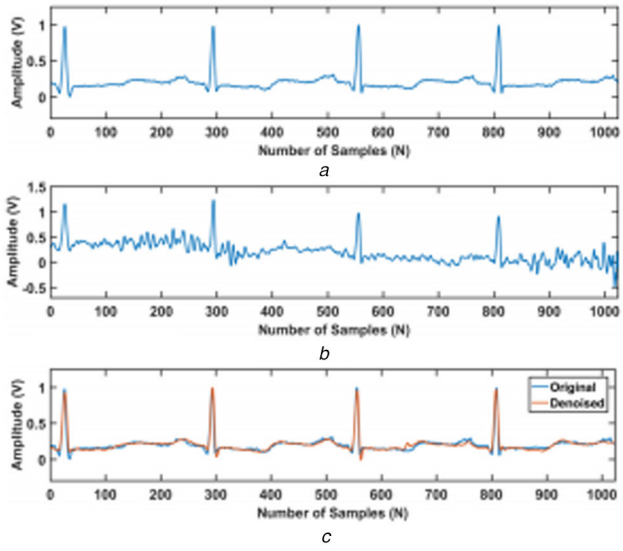

Fig. 19 *ECG denoising using FCN-based DAE*

(a) Original ECG signal (MIT-BIHA(100)), ***(b)*** Noisy ECG signal (BW + MA), ***(c)*** Denoised ECG signal in the deep-learning-based ECG denoising category using FCN-based DAE [19]

In Table 5, four equidistant $\mathrm{SNR}_{\mathrm{in}}$ values from −5 to 10 dB are used for the analysis with records 100 and 103 for record-specific analysis. Performance analysis has also been done on an average basis. For 100, wavelet-VBE performs the best amongst the methods available in the literature, with the $\mathrm{SNR}_{\mathrm{imp}}$ being 28.75 dB for both −5 and 0 dB $\mathrm{SNR}_{\mathrm{in}}$ values, 25 dB for 5 dB $\mathrm{SNR}_{\mathrm{in}}$, and 21.25 dB with the $\mathrm{SNR}_{\mathrm{in}}$ as 10 dB. EEMD-BLMS and EEMD-LMS also perform well.

Also, the WNN (db6) method shows a promising denoising performance at 0 dB $\mathrm{SNR}_{\mathrm{in}}$. For the record 103, IHP-ST EMD works comparatively well. An average analysis on the MIT-BIHA database shows that new MP-EKF gives an $\mathrm{SNR}_{\mathrm{imp}}$ of 15 dB for a highly noisy ECG signal at an $\mathrm{SNR}_{\mathrm{in}}$ of −5 dB, but gets surpassed by the hybrid technique, viz., VMD-NLM-WT, which gives an $\mathrm{SNR}_{\mathrm{imp}}$ of 11.50 dB for an $\mathrm{SNR}_{\mathrm{in}}$ of 0 dB. For 5 and 10 dB $\mathrm{SNR}_{\mathrm{in}}$ values, GAN2 performs better than other denoising methods considered for comparison. An average analysis on the MIT-BIH NSR database shows that new MP-EKF dominates other ECG denoisers for −5, 0, and 5 dB $\mathrm{SNR}_{\mathrm{in}}$ values, but VMD-DWT outperforms New MP-EKF at 10 dB $\mathrm{SNR}_{\mathrm{in}}$. Table 6 discusses the denoising performance for the records 100 and 103, considering MIT-BIHA, and for the records 16265_ECG1 and 16265_ECG2, considering MIT-BIH NSR. For the record 100, 1DTVD performs the best with an $\mathrm{SNR}_{\mathrm{imp}}$ of 8.34 and a PRD of 21.59% at an $\mathrm{SNR}_{\mathrm{in}}$ of 5 dB, and SDCST performs the best with an $\mathrm{SNR}_{\mathrm{imp}}$ of 7.62 and a PRD of 13.17% at an $\mathrm{SNR}_{\mathrm{in}}$ of 10 dB. For the record 103, GSSSA performs the best with an $\mathrm{SNR}_{\mathrm{imp}}$ of 10.2 and a PRD of 17.38% at an $\mathrm{SNR}_{\mathrm{in}}$ of 5 dB, and an $\mathrm{SNR}_{\mathrm{imp}}$ of 9.20 and a PRD of 10.96% at an $\mathrm{SNR}_{\mathrm{in}}$ of 10 dB. It is observed for the record 16265_ECG1 that 1DTVD outperforms other methods with an $\mathrm{SNR}_{\mathrm{imp}}$ of 4.14 and a PRD of 34.95% at an $\mathrm{SNR}_{\mathrm{in}}$ of 5 dB, and SDCST performs the best with an $\mathrm{SNR}_{\mathrm{imp}}$ of 2.99 and a PRD of 22.43% at an $\mathrm{SNR}_{\mathrm{in}}$ of 10 dB. A similar trend is observed in the case of 100. For 16265_ECG2, SDCST performs the best for both the $\mathrm{SNR}_{\mathrm{in}}$ values, viz., 5 and 10 dB. Table 7 discusses the performance of the denoising methods on the records 100, 103, and 105. The performance is analysed on 10 dB $\mathrm{SNR}_{\mathrm{in}}$. EMD-MAF outperforms other methods for all the three records. For all of the above inferences, the same trend can be observed for PRD. Tables 8 and 9 discuss the effect of ECG denoising in the presence of MA on the MIT-BIH databases. In Table 8, four equidistant $\mathrm{SNR}_{\mathrm{in}}$ values from −5 to 10 dB are used for the analysis. An average analysis on the MIT-BIHA database shows that DLSR performs well for all the $\mathrm{SNR}_{\mathrm{in}}$ values, except for 10 dB, where EMD-ASMF outperforms DLSR. For an average analysis on the MIT-BIH NSR database, new MP-EKF works comparatively well for all the $\mathrm{SNR}_{\mathrm{in}}$ values, with the $\mathrm{SNR}_{\mathrm{imp}}$ values being 16, 12.40,

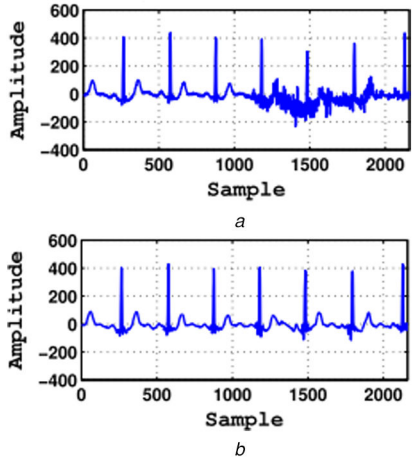

Fig. 20 *ECG denoising using DWT (Sym5)*

(a) Original ECG signal (MIT-BIHA(103)) + noise (MA), ***(b)*** Denoised ECG signal in the wavelet-based ECG denoising category using DWT (Sym5) with soft-thresholding (sqtwolog rule) [72]

8.25, and 3.60 dB, respectively. Table 9 discusses the denoising performance on the basis of SNR_{imp}, RMSE, PRD for the records

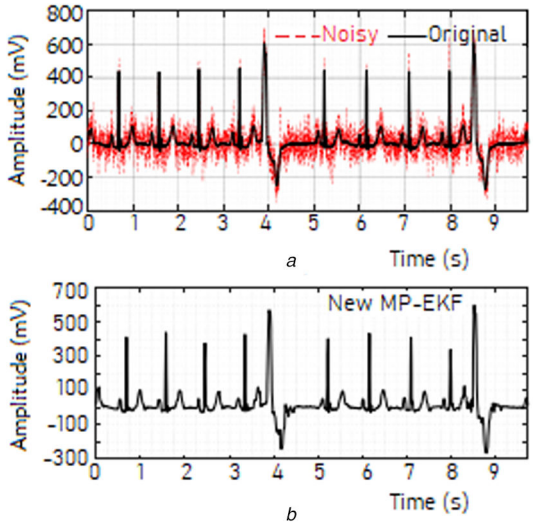

Fig. 21 *ECG denoising using new MP-EKF*
(a) Original ECG signal (MIT-BIHA(119)) and noisy ECG signal (AWGN), ***(b)*** Denoised ECG signal in the Bayesian-filter-based ECG denoising category using new MP-EKF [131]

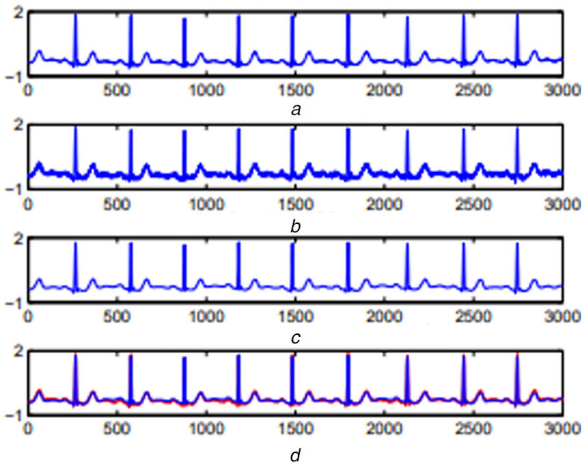

Fig. 22 *ECG denoising using sparsity*
(a) Original ECG signal (MIT-BIHA(103)), ***(b)*** Noisy ECG signal (BW), ***(c)*** Denoised ECG signal in the sparsity-based ECG denoising category using BW correction based on sparsity [28], ***(d)*** Comparison between original ECG signal and denoised ECG signal

103 and 105 of MIT-BIHA. For the record 103, and the record 105, GAN1 is the best method for ECG denoising when 5 and 10 dB SNR_{in} values are considered for each record. Tables 10 and 11 discuss the denoising performance for the records 103, and 105, considering MIT-BIHA. Table 10 compares the methods that denoise BW, and Table 11 compares the methods that denoise EM. For both the noises, GAN1 works the best in terms of all the three metrics. Table 12 discusses the denoising performance on PLI. For 0 dB SNR_{in}, EWT performs comparatively well with an SNR_{imp} of 22.15 dB and a PRD 7.81%. For 5 and 10 dB SNR_{in} values, DLSR works the best for both SNR_{imp} and PRD. Table 13 discusses CN removal, in which DWT (Sym6) (Soft) works well for 100, and MABWT (soft) for 103, for ECG noise. CPSD sparsity works well for 103 for both SNR_{imp} and PRD, for a mix of BW and other noises. Table 14 discusses DAE-based ECG denoising, in which it is evident that FCN-based DAE works the best among the three DAE-based denoisers, viz., DNN-based denoiser, CNN-based denoiser, and FCN-based denoiser for SNR_{imp}, RMSE, PRD at −1 dB as well as 7 dB for CN (MA + EM + BW) removal. Finally, Table 15 discusses the denoising performance for −5, 0, 5, and 10 dB SNR_{in} values on databases other than MIT-BIH. For a mixture of MITDB and QT, it is evident that DLSR works comparatively well on BW, PLI, and MA. For AWGN, DLSR performs well, but ADTF also works well in this case, especially for the 5 dB SNR_{in} value. For a mixture of MIT-BIHA, MIT-BIH NSR, AUDB, and SCDB, in the case of MA, AKF works well for −5 dB SNR_{in}, and EEMD works well for 0, 5, and 10 dB. In the case of AWGN, AKF provides better results at −5 and 0 dB SNR_{in} values. For 5 and 10 dB SNR_{in} values, WT works well. For PTB, in the case of CN, UWT works well for −5 and 0 dB SNR_{in} values.

The superiority of some ECG denoising methods over the others can be attributed to the unique intelligent design of the denoisers. The noted advantage of the GAN2 framework is the generalisation to several noise conditions using a single generative model. New MP-EKF has an adaptive fuzzy based particle weighting strategy along with a non-linear framework to trace ECG signals efficiently. The high performance of GSSSA in correcting the noise is associated with preserving all features of the ECG data. The other methods, on the contrary, introduce distortions to the main features of the corrected ECG data. EMD-MAF benefits from its hybrid nature to improve the QRS quality. DLSR incorporates the advantages of both time and frequency localisation capability sparse representation scheme for efficient enhancement of ECG signal quality. The adversarial de-noising method of GAN1 has a high generalisation capacity that not only helps it to preserve the morphological characteristics of the ECG signals, but also enhances these characteristics to some extent. The CPSP sparsity algorithm benefits from the use of banded matrices. It ensures that the iterative optimisation algorithm is computationally efficient, uses minimal memory, and applies to long data series. FCN-based

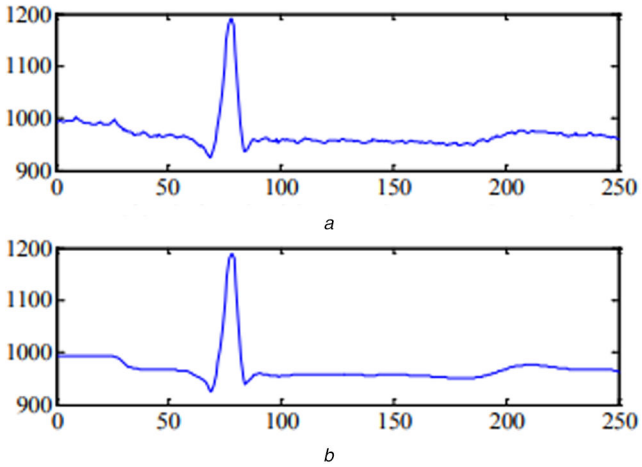

Fig. 23 *ECG denoising using WNN*
(a) Original ECG signal (MIT-BIHA(119)) + AWGN, ***(b)*** Denoised ECG signal in the hybrid ECG denoising category using an optimum ECG denoising with wavelet neural network [139]

Table 5 Performance of various methodologies in the presence of AWGN on MIT-BIH databases

SNR_{in}	−5 dB		0 dB		5 dB		10 dB	
Method	SNR_{imp}	PRD, %	SNR_{imp}	PRD, %	SNR_{imp}	PRD, %	SNR_{imp}	*PRD*, %
record information				*Record 100*				
EMD-NLM [141]	10.50	53.00	9.40	34.00	9.20	20.00	8.20	12.00
Wavelet-VBE [111]	**28.75**	**6.50**	**28.75**	**3.70**	**25.00**	**3.16**	**21.25**	**2.70**
EEMD-BLMS [54]	26.12	8.80	22.00	8.00	17.00	8.00	11.50	8.40
EEMD-LMS [55]	26.00	8.91	21.90	8.04	16.90	8.04	11.30	8.61
C-EEMD [51]	21.20	15.50	18.75	11.50	11.25	16.00	7.50	13.00
C-EMD [21]	6.80	81.00	3.75	65.50	2.50	42.00	−1.25	37.00
WNN [54]	21.25	16.00	18.75	11.50	13.72	12.00	8.75	11.50
CF [24]	2.60	132.00	2.20	78.00	1.15	49.00	−1.88	39.00
AF [25]	6.25	87.00	1.25	87.00	−3.50	84.00	−5.60	60.00
wavelet [97]	10.00	56.00	9.40	34.00	7.50	24.00	6.80	15.00
NLM [26]	6.24	87.00	8.00	40.00	8.50	21.00	7.70	13.00
EMD [36]	7.00	79.00	6.20	49.00	4.96	32.00	2.20	25.00
WNN (db6) [139]	—	—	28.16	3.91	—	—	—	—
record information				*Record 103*				
IHP-ST EMD [46]	**7.84**	**72.11**	**6.14**	**49.32**	**5.89**	**28.54**	—	—
SF method [42]	4.97	100.35	4.98	56.36	4.92	31.92	—	—
record information				*Average (MIT-BIHA)*				
DLSR [122]	9.25	61.00	9.23	35.00	8.62	21.00	8.30	12.16
CE-EMD-AN [56]	7.60	74.00	6.90	45.00	6.25	27.00	5.62	17.00
NLM [26]	9.80	58.00	8.85	36.00	7.90	23.00	6.30	15.00
SASS [123]	8.50	67.00	8.20	40.00	7.90	23.00	7.50	13.34
GMC [125]	10.00	56.00	9.50	34.00	9.00	20.00	7.70	13.03
EMD-NLM [141]	10.50	53.00	9.90	32.00	9.60	19.00	8.20	12.30
EMD [36]	7.00	79.00	6.80	46.00	6.50	27.00	4.60	19.00
LMS-based sparsity [124]	8.39	48.92	—	—	−0.06	35.18	—	—
RLS-based sparsity [124]	8.89	46.36	—	—	−0.66	43.22	—	—
New MP-EKF [131]	**15.00**	**32.00**	11.10	28.00	8.00	22.00	3.50	21.00
VMD-DWT [144]	9.50	60.00	8.62	37.00	7.50	24.00	5.50	17.00
VMD-NLM [137]	—	—	—	—	9.77	12.94	8.27	8.64
GAN2 [65]	—	—	—	—	**14.52**	**6.68**	**14.64**	**6.49**
DWT [97]	—	—	5.50	53.09	5.00	31.62	4.00	19.95
EMD-DWT [142]	—	—	10.40	30.20	8.50	21.13	7.00	14.12
NLM-DWT [145]	—	—	11.00	28.18	9.60	18.62	8.70	11.61
VMD-NLM-WT [137]	—	—	**11.50**	**26.61**	10.60	16.60	9.00	11.22
linear EKF25(4 obs) [132]	11.25	48.70	10.25	30.73	7.90	22.65	—	—
non-linear EKF25 (4 obs) [132]	11.75	45.97	9.50	33.50	8.00	22.39	—	—
linear EKF25(2 obs) [132]	10.70	51.88	10.70	39.17	8.00	22.39	—	—
non-linear EKF25 (2 obs) [132]	10.90	50.70	9.20	34.67	8.50	21.14	—	—
record information				*Average (MIT-BIH NSR)*				
EKF3 linear [23]	9.60	59.00	8.25	39.00	6.75	25.85	4.60	18.62
old MP-EKF [129]	14.00	35.00	10.75	29.00	6.70	26.00	2.25	24.41
EKS3 linear [23]	12.00	45.00	10.25	31.00	8.20	22.00	5.90	16.03
EKF3 non-linear [132]	9.50	60.00	8.25	39.00	7.10	25.00	5.75	16.31
new MP-EKF [131]	**18.00**	**22.00**	**14.50**	**19.00**	**10.40**	**17.00**	5.80	16.22
EKS3 non-linear [132]	12.00	45.00	10.50	30.00	8.90	20.00	7.00	14.12
VMD-DWT [144]	10.50	53.00	6.20	35.00	8.90	20.00	**7.20**	**13.80**

The bold values represent the best performance.

DAE preserves more clinically relevant information than other DAEs. It extracts the original ECG signal from its noisy version by learning to capture the main morphological features of the ECG signal rather than following the noisy signal, which is done by other DAEs.

4 Future scope

This work can be extended by including upcoming denoising techniques for comparative analysis with the existing work and as an aid to a variety of research work, viz., feature extraction, wave detection, heartbeat classification, and so on, where ECG denoising is a major pre-processing step and help develop new methods based on some of these features. This paper can also serve as a valuable source of information, with the help of which modern techniques can be developed with better SNR_{imp}, lower values of RMSE, and PRD. Kumar *et al.* [158] point out that high-sensitivity online ECG detectors do not show satisfactory results in detecting pathological ECG signals due to the presence of noises like BW and PLI. In this case, this paper might help to choose a satisfactory ECG denoising method. Also, following the latest trends, convolutional neural networks, like DnCNN and PDNN, can be improvised for ECG signal denoising [159].

Table 6 Performance of various methodologies in the presence of AWGN on MIT-BIH databases

SNR_{in}	5 dB		10 dB		5 dB		10 dB	
Method	SNR_{imp}	PRD, %	SNR_{imp}	PRD, %	SNR_{imp}	PRD, %	SNR_{imp}	PRD, %
record information	*Record 100*				*Record 103*			
EMD-ASMF [136]	3.85	36.13	2.85	22.78	4.28	34.40	3.31	21.61
SDCST [146]	1.66	49.15	**7.62**	**13.17**	9.00	20.03	8.54	11.86
NIWT [98]	0.08	55.75	0.31	30.58	0.36	54.01	0.27	30.67
EMDCV [43]	4.96	31.84	4.40	19.11	5.18	31.02	4.87	18.09
NLT [99]	4.06	35.28	1.78	25.91	1.57	46.95	−1.69	38.58
GSSSA [119]	—	—	—	—	**10.20**	**17.38**	**9.20**	**10.96**
GSLPF [119]	—	—	—	—	9.40	19.06	8.40	12.02
1DTVD [27]	**8.34**	**21.59**	3.93	20.11	8.48	21.21	4.41	19.05
GAMNVE [29]	5.88	28.65	5.09	17.65	5.71	29.24	5.47	16.94
PLI-SWT [93]	−0.01	56.30	−0.01	31.67	−0.01	56.30	−0.01	31.66
UHTWS [100]	4.90	32.03	4.21	19.52	4.92	31.98	5.56	16.72
record information	*Record 16265_ECG1*				*Record 16265_ECG2*			
EMD-ASMF [136]	1.76	45.95	0.45	30.02	1.96	44.91	1.17	27.68
SDCST [146]	3.86	36.10	**2.99**	**22.43**	**2.27**	**43.34**	**1.66**	**26.13**
NIWT [98]	0.19	55.06	0.15	31.11	0.16	55.28	0.05	31.47
EMDCV [43]	2.49	42.32	2.02	25.36	1.20	49.34	−2.53	44.41
NLT [99]	1.34	50.31	0.80	28.66	−1.22	64.75	−6.01	63.17
1DTVD [27]	**4.14**	**34.95**	−0.37	33.02	−3.52	84.30	−8.52	84.29
GAMNVE [29]	1.98	44.82	1.12	27.82	1.01	50.16	−0.62	34.01
PLI-SWT [93]	−0.11	56.31	−0.02	31.69	−0.01	56.30	−0.01	31.68
UHTWS [100]	−1.54	67.20	−2.19	40.75	−1.34	65.63	−4.04	50.38

The bold values represent the best performance.

Table 7 Performance of various methodologies in the presence of AWGN on MIT-BIH databases

SNR_{in}	10 dB		10 dB		10 dB	
Method	SNR_{imp}	PRD, %	SNR_{imp}	PRD, %	SNR_{imp}	PRD, %
record information	*Record 100*		*Record 103*		*Record 105*	
GMC [125]	7.66	13.10	9.08	11.11	8.58	11.78
EMD-MAF [40]	**19.94**	**3.18**	**18.21**	**3.89**	**13.48**	**6.70**
EMD [45]	9.50	10.59	6.80	14.45	4.50	18.84
VMD-NLM-WT [137]	8.92	11.32	8.56	11.80	8.35	12.09
DWT-LDMR [103]	—	—	0.61	30.67	—	—

The bold values represent the best performance.

Table 8 Performance of various methodologies in the presence of MA on MIT-BIH databases

SNR_{in}	−5 dB		0 dB		5 dB		10 dB	
Method	SNR_{imp}	PRD, %	SNR_{imp}	PRD, %	SNR_{imp}	PRD, %	SNR_{imp}	PRD, %
record information	*Average (MIT-BIHA)*							
DLSR [122]	**12.00**	**44.67**	**10.50**	**29.85**	**9.90**	**17.99**	7.40	13.49
EMD-ASMF [136]	—	—	9.30	34.32	9.14	18.95	**8.79**	**11.52**
DWT (Sym5) [72]	11.00	50.12	9.80	32.36	—	—	—	—
DWT (Dmey) [96]	8.80	64.60	5.10	55.60	—	—	—	—
record information	*Average (MIT-BIH NSR)*							
EKF3 linear [23]	1.20	154.88	1.00	89.12	0.70	51.88	0.20	30.90
old MP-EKF [129]	11.25	48.70	9.20	35.00	6.18	27.61	2.30	24.27
EKS3 linear [23]	1.45	150.49	1.30	86.10	0.90	50.70	0.50	29.85
EKF3 non-linear [132]	1.20	154.88	1.00	89.12	0.80	51.29	0.60	29.51
new MP-EKF [131]	**16.00**	**28.18**	**12.40**	**24.00**	**8.25**	**21.75**	**3.60**	**20.89**
EKS3 non-linear [132]	1.45	150.49	1.30	86.10	1.00	50.12	0.90	28.51
VMD-DWT [144]	0.10	175.79	0.10	98.86	0.10	55.59	0.10	31.26

The bold values represent the best performance.

5 Conclusion

ECG denoising is a pre-processing step in a wide variety of research work. It finds numerous applications in the field of medical sciences. Hence, it is important to know the techniques which help in the noise-free acquisition of ECG and also help in the removal of predominant noises that happen to corrupt clean ECG. In this paper, six major domains of ECG denoising have been explored, so that an analysis of the performance of the methods can help in better diagnosis and proper treatment of patients. The paper analyses conventional domains like wavelet-based ECG denoising and also takes into consideration the newer

Table 9 Performance of various methodologies in the presence of MA on MIT-BIH databases

SNR_{in}	0 dB			5 dB			0 dB			5 dB		
Method	SNR_{imp}	RMSE	PRD, %	SNR_{imp}	RMSE	PRD, %	SNR_{imp}	RMSE	PRDD, %	SNR_{imp}	RMSE	PRDD, %
record information			*Record 103*						*Record 105*			
stacked DAE [57]	18.92	0.046	11.32	14.31	0.044	10.83	22.97	0.036	7.10	19.12	0.032	6.22
improved DAE [63]	21.38	0.034	8.53	18.33	0.027	6.82	24.72	0.030	5.81	20.13	0.028	5.54
GAN1 [64]	**41.36**	**0.004**	**0.86**	**38.24**	**0.003**	**0.69**	**36.49**	**0.007**	**1.50**	**34.55**	**0.005**	**1.05**
WT [97]	19.66	0.044	10.40	10.79	0.067	16.24	22.09	0.040	7.86	19.11	0.032	6.23
GAN2 [65]	—	—	—	20.34	0.018	5.41	—	—	—	24.01	0.021	3.54

The bold values represent the best performance.

Table 10 Performance of various methodologies in the presence of BW on MIT-BIH databases

SNR_{in}	0 dB			5 dB			0 dB			5 dB		
Method	SNR_{imp}	RMSE	PRD, %	SNR_{imp}	RMSE	PRD, %	SNR_{imp}	RMSE	PRD, %	SNR_{imp}	RMSE	PRD, %
record information			*Record 103*						*Record 105*			
CPSD sparsity [28]	12.53	0.095	23.63	12.87	0.055	12.78	—	—	—	—	—	—
stacked DAE [57]	20.38	0.038	9.57	15.77	0.037	9.15	24.90	0.029	5.69	20.47	0.027	5.33
improved DAE [63]	23.78	0.026	6.47	18.89	0.025	6.39	25.40	0.028	5.37	20.45	0.027	5.34
GAN1 [64]	**40.26**	**0.003**	**0.97**	**36.60**	**0.003**	**0.83**	**39.49**	**0.004**	**1.06**	**35.56**	**0.003**	**0.94**
WT [97]	14.87	0.074	18.05	9.90	0.074	17.99	31.53	0.014	2.65	27.71	0.012	2.31
GAN2 [65]	—	—	—	25.70	0.019	2.92	—	—	—	22.01	0.022	4.46

The bold values represent the best performance.

Table 11 Performance of various methodologies in the presence of EM on MIT-BIH databases

SNR_{in}	0 dB			5 dB			0 dB			5 dB		
Method	SNR_{imp}	RMSE	PRD, %	SNR_{imp}	RMSE	PRD, %	SNR_{imp}	RMSE	PRD, %	SNR_{imp}	RMSE	PRD, %
record information			*Record 103*						*Record 105*			
stacked DAE [57]	18.94	0.047	11.30	14.30	0.041	10.84	23.45	0.035	6.72	19.56	0.030	5.92
improved DAE [63]	22.75	0.029	7.29	18.45	0.027	6.72	23.70	0.033	6.53	19.66	0.030	5.85
GAN1 [64]	**38.09**	**0.005**	**1.25**	**34.39**	**0.004**	**1.07**	**34.27**	**0.008**	**1.93**	**30.67**	**0.007**	**1.65**
WT [97]	9.51	0.138	33.46	8.07	0.091	22.21	21.47	0.044	8.44	23.37	0.020	3.82
GAN2 [65]	—	—	—	20.40	0.022	5.37	—	—	—	21.20	0.024	4.90

The bold values represent the best performance.

Table 12 Performance of various methodologies in the presence of PLI on MIT-BIH databases

SNR_{in}	0 dB		5 dB		10 dB	
Method	SNR_{imp}	PRD, %	SNR_{imp}	PRD, %	SNR_{imp}	PRD, %
record information			*Average (MIT-BIHA)*			
EMD-ASMF [136]	14.50	18.84	13.50	11.88	10.50	9.44
DLSR [122]	21.90	8.04	**19.00**	**6.31**	**16.90**	**4.52**
EWT [109]	**22.15**	**7.81**	18.00	7.08	14.10	6.24
EMD [21]	15.70	16.40	11.12	15.62	4.65	18.51

The bold values represent the best performance.

Table 13 Performance of various methodologies in the presence of CN on MIT-BIH databases

SNR_{in}		5 dB		5 dB	
Method		SNR_{imp}	PRD, %	SNR_{imp}	PRD, %
record information		*Record 100*		*Record 103*	
DWT (Sym6) (Soft) [70]	CN (ECG noise)	**9.41**	**19.00**	7.14	25.00
DWT (Sym6) (Hard) [70]		4.92	32.00	4.88	32.00
S-median (Soft) [105]		1.49	47.00	1.46	48.00
S-median (Hard) [105]		0.01	56.00	0.01	56.00
MABWT (Soft) [110]		7.80	23.00	**7.70**	**23.00**
MABWT (Hard) [110]		6.40	27.00	5.80	29.00
CPSD sparsity [28]	CN (BW + other noises)	—	—	**10.94**	**15.96**
wavelet [28]		—	—	−0.15	57.21
median filter [28]		—	—	−1.15	64.19

The bold values represent the best performance.

Table 14 Performance of various methodologies in the presence of CN (EM + BW + MA) on MIT-BIH databases

SNR_{in}		−1 dB			7 dB	
Method	SNR_{imp}	RMSE	PRD, %	SNR_{imp}	RMSE	PRD, %
record information			*Average (MIT-BIHA)*			
DNN-based DAE [19]	11.75	0.094	30.89	5.56	0.077	25.79
CNN-based DAE [19]	14.14	0.072	23.09	8.37	0.056	18.40
FCN-based DAE [19]	**15.49**	**0.063**	**19.68**	**10.54**	**0.044**	**14.05**

The bold values represent the best performance.

Table 15 Performance of various methodologies in the presence of different noises on other databases

Record information		Average											
SNR_{in}		−5 dB			0 dB			5 dB			10 dB		
Method		SNR_{imp}	RMSE	PRD, %	SNR_{imp}	RMSE	PRD, %	SNR_{imp}	RMSE	PRD, %	SNR_{imp}	RMSE	PRD, %
noise type							*BW*						
DLSR [122]	MITDB/QT	**19.89**	**0.05**	**18.02**	**19.70**	**0.03**	**10.38**	**18.61**	**0.02**	**6.66**	**16.13**	**0.01**	**4.99**
DWT [97]		17.23	0.08	25.57	12.78	0.07	24.38	8.20	0.07	23.56	3.32	0.07	23.39
DLSR [122]	QT	19.75	—	18.30	19.50	—	11.00	18.5	—	7.00	16.00	—	5.00
noise type							*PLI*						
DLSR [122]	MITDB/QT	**23.31**	**0.04**	**12.38**	**21.54**	**0.03**	**8.62**	**18.46**	**0.02**	**6.914**	**16.14**	**0.02**	**5.06**
EMD-ASMF [136]		14.19	0.11	34.78	13.52	0.06	21.15	11.86	0.04	14.45	8.96	0.03	11.50
DWT [97]		10.56	0.16	53.81	7.60	0.13	42.10	5.519	0.09	30.02	3.91	0.06	20.27
DLSR [122]	QT	22.50	—	13.00	21.50	—	8.00	18.00	—	7.00	16.00	—	5.00
noise type							*MA*						
DLSR [122]	MITDB/QT	**11.69**	**0.14**	**47.19**	**10.55**	**0.09**	**29.68**	**9.53**	**0.06**	**18.77**	**7.23**	**0.05**	**14.02**
SRD [126]		10.84	0.15	51.29	7.33	0.13	43.65	2.77	0.12	41.80	2.16	0.12	41.61
EMD-ASMF [136]		0.50	0.50	167.83	0.46	0.28	94.78	0.33	0.16	54.16	0.11	0.09	31.10
DLSR [122]	QT	11.25	—	49.00	10.50	—	29.85	9.50	—	18.84	7.50	—	13.00
AKF [135]	MITDB/NSRDB/CUDB/SCDB	**9.90**	—	**56.88**	7.20	—	43.65	3.50	—	37.58	−1.00	—	35.48
EEMD [34]		8.80	—	64.56	**8.00**	—	**89.81**	**6.40**	—	**26.92**	**4.00**	—	**19.95**
WT [97]		6.40	—	85.11	5.50	—	53.09	3.50	—	37.58	0.50	—	29.85
noise type							*AWGN*						
DLSR [122]	MITDB/QT	**9.20**	**0.18**	**62.14**	**9.11**	0.11	**34.95**	**8.44**	**0.06**	**21.28**	**7.85**	**0.04**	**12.98**
SRD [126]		7.80	0.22	72.49	6.47	0.14	47.71	4.19	0.11	35.94	0.66	0.09	31.64
ADTF [112]		9.10	**0.18**	62.27	9.03	**0.10**	35.13	**8.44**	**0.06**	**21.28**	7.35	**0.04**	13.78
DLSR [122]	QT	9.50	—	60.00	9.00	—	36.00	8.75	—	21.00	7.75	—	13.00
AKF [135]	MITDB/NSRDB/CUDB/SCDB	**11.90**	—	**45.19**	**9.90**	—	**31.99**	6.50	—	26.61	0.01	—	31.62
EEMD [34]		10.80	—	51.29	10.30	—	30.55	8.70	—	20.65	6.00	—	15.85
WT [97]		10.50	—	53.09	9.20	—	34.67	**8.90**	—	**20.18**	**7.80**	—	**12.88**
noise type							*CN*						
DRDNN [66]	PTB	14.00	—	35.48	9.00	—	35.48	—	—	—	—	—	—
UWT [108]		**17.00**	—	**25.12**	**30.00**	—	**3.16**	—	—	—	—	—	—

The bold values represent the best performance.

techniques like deep-learning-based DAEs and sparsity-based denoising models. It also throws light on the hybrid techniques, which involve a mix of techniques from different research areas. This paper compares the existing methods and helps us make a few deductions. On the basis of a thorough analysis of the performances of the denoising methods tested on MIT-BIH databases, we can conclude that wavelet-VBE, GAN2, new MP-EKF, SDCST, GSSSA, and EMD-MAF perform comparatively better in case of AWGN. For MA, DLSR, new MP-EKF, and GAN1 outperform other denoising methods. For BW and EM, GAN1 is a good choice for ECG denoising. EWT and DLSR are best suited for PLI noise removal while DWT (Sym6) soft, MABWT (Soft), CPSD sparsity, and FCN-based DAE show promising results for CN removal. To mention, FCN-based DAE is a comparatively preferable denoiser for the noise mixture of EM, BW, and MA among DAE-based denoisers. The analysis also shows that other databases like PTB, QT, CUDB, SCDB, and so on can also be used to test the performance of ECG denoising methods effectively in case of AWGN, MA, BW, PLI, and CN. It can also be seen that methods like DLSR, AKF, EEMD, WT, and UWT show promising results for databases other than MIT-BIH. Thus, the paper can serve as a compiled source of valuable information in the context of ECG signal denoising. The work can be extended by considering newer denoising mechanisms for future analysis.

6 References

[1] Han, G., Lin, B., Xu, Z.: 'Electrocardiogram signal denoising based on empirical mode decomposition technique: an overview', *J. Instrum.*, 2017, **12**, (3), pp. P03010–P03010

[2] Li, W.: 'Wavelets for electrocardiogram: overview and taxonomy', *IEEE Access*, 2019, **7**, pp. 25627–25649

[3] He, R., Wang, K., Li, Q., *et al.*: 'A novel method for the detection of R-peaks in ECG based on K-nearest neighbours and particle swarm optimization', *EURASIP J. Adv. Signal Process.*, 2017, **2017**, pp. 1–14

[4] Lastre-Dominguez, C., Shmaliy, Y.S., Ibarra-Manzano, O., *et al.*: 'Fiducial features extraction for ECG signals using state-space unbiased FIR smoothing'. IEEE Int. Autumn Meeting Power, Electronics Computing (ROPEC), Ixtapa, Mexico, 2018, pp. 1–6

[5] Ye, C., Vijaya Kumar, B.V.K., Coimbra, M.T.: 'Heartbeat classification using morphological and dynamic features of ECG signals', *IEEE Trans. Biomed. Eng.*, 2012, **59**, (10), pp. 2930–2941

[6] Brüser, C., Antink, C.H., Wartzek, T., *et al.*: 'Ambient and unobtrusive cardiorespiratory monitoring techniques', *IEEE Rev. Biomed. Eng.*, 2015, **8**, pp. 30–43

[7] Mporas, I., Tsirka, V., Zacharaki, E.I., *et al.*: 'Seizure detection using EEG and ECG signals for computer-based monitoring, analysis and management of epileptic patients', *Expert Syst. Appl.*, 2015, **42**, (6), pp. 3227–3233

[8] Kang, S.J., Lee, S.Y., Cho, H.I., *et al.*: 'ECG authentication system design based on signal analysis in mobile and wearable devices', *IEEE Signal Process. Lett.*, 2016, **23**, (6), pp. 805–808

[9] Cox, J., Nolle, F.M., Fozzard, H.A., *et al.*: 'AZTEC, a preprocessing program for real-time ECG rhythm analysis', *IEEE Trans. Biomed. Eng.*, 1968, **15**, (2), pp. 128–129

[10] Lee, W.K., Yoon, H., Park, K.S.: 'Smart ECG monitoring patch with built-in R-peak detection for long-term HRV analysis', *Ann. Biomed. Eng.*, 2016, **44**, (7), pp. 2292–2301

[11] Pope, J.H., Aufderheide, T., Ruthazer, R., *et al.*: 'Missed diagnoses of acute cardiac ischemia in the emergency department', *N. Engl. J. Med.*, 2000, **342**, (16), pp. 1163–1170

[12] Clifford, G.D.: 'ECG statistics, noise, artifacts, and missing data', *Adv. Meth. Tools ECG Anal.*, 2006, **6**, pp. 55–99

[13] Friesen, G.M., Jannett, T.C., Jadallah, M.A., *et al.*: 'A comparison of the noise sensitivity of nine QRS detection algorithms', *IEEE Trans. Biomed. Eng.*, 1990, **37**, (1), pp. 85–98

[14] Satija, U., Ramkumar, B., Sabarimalai Manikandan, M.: 'A review of signal processing techniques for electrocardiogram signal quality assessment', *IEEE Rev. Biomed. Eng.*, 2018, **11**, pp. 36–52

[15] Van Alsté, J., Schilder, T.: 'Removal of base-line wander and power-line interference from the ECG by an efficient FIR filter with a reduced number of taps', *IEEE Trans. Biomed. Eng.*, 1985, **32**, (12), pp. 1052–1060

[16] Levkov, C., Mihov, G., Ivanov, R., *et al.*: 'Removal of power-line interference from the ECG: a review of the subtraction procedure', *Biomed. Eng. Online*, 2005, **50**, (4), pp. 1–18

[17] Frølich, L., Dowding, I.: 'Removal of muscular artifacts in EEG signals: a comparison of linear decomposition methods', *Brain. Inform.*, 2018, **5**, (1), pp. 13–22

[18] Velayudhan, A., Peter, S.: 'Noise analysis and different denoising techniques of ECG signal-a survey', *IOSR J. Electron. Commun. Eng.*, 2016, **1**, pp. 40–44

[19] Chiang, H.T., Hsieh, Y.Y., Fu, S.W., *et al.*: 'Noise reduction in ECG signals using fully convolutional denoising autoencoders', *IEEE Access*, 2019, **7**, pp. 60806–60813

[20] Kumar, A., Ranganatham, R., Komaragiri, R., *et al.*: 'Efficient QRS complex detection algorithm based on fast Fourier transform', *Biomed. Eng. Lett.*, 2019, **9**, (1), pp. 145–151

[21] Huang, N.E., Shen, Z., Long, S.R., *et al.*: 'The empirical mode decomposition and the Hubert spectrum for nonlinear and non-stationary time series analysis', *Proc. R. Soc. A, Math. Phys. Eng. Sci.*, 1998, **454**, pp. 903–995

[22] Singh, G., Kaur, G., Kumar, V.: 'ECG denoising using adaptive selection of IMFs through EMD and EEMD'. 2014 Int. Conf. on Data Science & Engineering (ICDSE), Kochi, India, 2014, pp. 228–231

[23] Sameni, R., Shamsollahi, M.B., Jutten, C., *et al.*: 'A nonlinear Bayesian filtering framework for ECG denoising', *IEEE Trans. Biomed. Eng.*, 2007, **54**, (12), pp. 2172–2185

[24] Joshi, S.L., Vatti, R.A., Tornekar, R.V.: 'A survey on ECG signal denoisingtechniques'. 2013 Int. Conf. on Communication Systems and Network Technologies, Gwalior, 2013, pp. 60–64

[25] Marque, C., Bisch, C., Dantas, R., *et al.*: 'Adaptive filtering for ECG rejection from surface EMG recordings', *J. Electromyogr. Kinesiol.*, 2005, **15**, (3), pp. 310–315

[26] Tracey, B.H., Miller, E.L.: 'Nonlocal means denoising of ECG signals', *IEEE Trans. Biomed. Eng.*, 2012, **59**, (9), pp. 2383–2386

[27] Condat, L.: 'A direct algorithm for 1-D total variation denoising', *IEEE Signal Process. Lett.*, 2013, **20**, (11), pp. 1054–1057

[28] Wang, X., Zhou, Y., Shu, M., *et al.*: 'ECG baseline wander correction and denoising based on sparsity', *IEEE Access*, 2019, **7**, pp. 31573–31585

[29] Vargas, R.N., Veiga, A.C.P.: 'Electrocardiogram signal denoising by a new noise variation estimate', *Res. Biomed. Eng.*, 2020, **36**, (1), pp. 13–20

[30] Chawla, M.P.S.: 'PCA and ICA processing methods for removal of artifacts and noise in electrocardiograms: a survey and comparison', *Appl. Soft Comput. J.*, 2011, **11**, (2), pp. 2216–2226

[31] Moody, G.B., Mark, R.G.: 'QRS morphology representation and noise estimation using the Karhunen-Loeve transform',Computer in Cardiology, Jerusalam, Israel, 1989, pp. 269–272

[32] He, T., Clifford, G., Tarassenko, L.: 'Application of independent component analysis in removing artefacts from the electrocardiogram', *Neural Comput. Appl.*, 2006, **15**, (2), pp. 105–116

[33] Barros, A.K., Mansour, A., Ohnishi, N.: 'Removing artifacts from electrocardiographic signals using independent components analysis', *Neurocomputing*, 1998, **22**, (1–3), pp. 173–186

[34] Chang, K.M.: 'Arrhythmia ECG noise reduction by ensemble empirical mode decomposition', *Sensors*, 2010, **10**, (6), pp. 6063–6080

[35] Boudraa, A.O., Cexus, J.C.: 'Denoising via empirical mode decomposition'. Proc. IEEE ISCCSP, Athens, Greece, 2006

[36] Blanco-Velasco, M., Weng, B., Barner, K.E.: 'ECG signal denoising and baseline wander correction based on the empirical mode decomposition', *Comput. Biol. Med.*, 2008, **38**, (1), pp. 1–13

[37] Wu, Z., Huang, N.E.: 'On the filtering properties of the empirical mode decomposition', *Adv. Adapt. Data Anal.*, 2010, **2**, (4), pp. 397–414

[38] Huang, N.E., Attoh-Okine, N.O.: '*The Hilbert–Huang transform in engineering*' (CRC Press, Boca Raton, 2005)

[39] Gurumurthy, S.: 'System design for baseline wander removal of ECG signals with empirical mode decomposition using Matlab', *Int. J. Soft Comput. Eng.*, 2013, **3**, (3), pp. 85–92

[40] Sonali, S., Singh, O., Sunkaria, R.K.: 'ECG signal denoising based on Empirical Mode Decomposition and moving average filter'. 2013 IEEE Int. Conf. Signal Processing Computing Control (ISPCC), Solan, 2013, pp. 1–6

[41] Pan, N., Mang, V., Un, M.P., *et al.*: 'Accurate removal of baseline wander in ECG using empirical mode decomposition'. Proc. 2007 Joint Meeting 6th Int. Symp. Noninvasive Functional Source Imaging Brain Heart Int. Conf. Functional Biomedical Imaging, NFSI ICFBI 2007, Hangzhou, China, 2007, pp. 177–180

[42] Chacko, A., Ari, S.: 'Denoising of ECG signals using Empirical Mode Decomposition based technique'. IEEE-Int. Conf. on Advances in Engineering, Science and Management, ICAESM-2012', Nagapattinam, Tamil Nadu, India, 2012

[43] Hee-Seok, K.D.-H.: 'Hierarchical smoothing technique by empirical mode decomposition', *Korean J. Appl. Stat.*, 2006, **19**, (2), pp. 319–330

[44] Narsimha, B., Suresh, E., Punnamchandar, K., *et al.*: 'Denoising and QRS detection of ECG signals using Empirical Mode Decomposition'. ICCSP 2011 - 2011 Int. Conf. Communications Signal Processing, Calicut, India, 2011, no. 1, pp. 439–442

[45] Weng, B., Blanco-Velasco, M., Barner, K.E.: 'ECG denoising based on the empirical mode decomposition'. 2006 Int. Conf. of the IEEE Engineering in Medicine and Biology Society, New York City, USA, 30 August–3 September 2006, pp. 1–4

[46] Samadi, S., Shamsollahi, M.B.: 'ECG noise reduction using empirical mode decomposition based on combination of instantaneous half period and soft-thresholding'. 2nd Middle East Conf. on Biomedical Engineering, Doha, 2014, pp. 244–248

[47] Tang, G., Qin, A.: 'ECG de-noising based on empirical mode decomposition'. 2008 The 9th Int. Conf. for Young Computer Scientists, Hunan, 2008, pp. 903–906

[48] Zhang, J.X., Zhong, Q.H., Dai, Y.P.: 'The determination of the threshold and the decomposition order in threshold de-noising method based on wavelet transform', *Proc. CSEE*, 2004, **24**, (2), pp. 118–122

[49] Anapagamini, S.A., Rajavel, R.: 'Removal of artifacts in ECG using Empirical mode decomposition'. 2013 Int. Conf. on Communication and Signal Processing, Melmaruvathur, 2013, pp. 288–292

[50] Nguyen, P., Kim, J.M.: 'Adaptive ECG denoising using genetic algorithm-based thresholding and ensemble empirical mode decomposition', *Inf. Sci.*, 2016, **373**, pp. 499–511

[51] Wu, Z., Huang, N.E.: 'Ensemble empirical mode decomposition: a noise-assisted data analysis method', *Adv. Adapt. Data Anal.*, 2009, **1**, (1), pp. 1–41

[52] Jenitta, J., Rajeswari, A.: 'Denoising of ECG signal based on improved adaptive filter with EMD and EEMD'. 2013 IEEE Conf. on Information & Communication Technologies, Thuckalay, Tamil Nadu, India, 2013, pp. 957–962

[53] Olmos, S., Sörnmo, L., Laguna, P.: 'Block adaptive filters with deterministic reference inputs for event-related signals: BLMS and BRLS', *IEEE Trans. Signal Process.*, 2002, **50**, (5), pp. 1102–1112

[54] Kaergaard, K., Jensen, S.H., Puthusserypady, S.: 'A comprehensive performance analysis of EEMD-BLMS and DWT-NN hybrid algorithms for ECG denoising', *Biomed. Signal Process. Control*, 2016, **25**, pp. 178–187

[55] Diniz, P.S.R., Diniz, P.S.R.: 'The least-mean-square (LMS) algorithm', in '*Adaptive filtering*' (Springer, USA, 1997), pp. 71–131

[56] Colominas, M.A., Schlotthauer, G., Torres, M.E.: 'Improved complete ensemble EMD: a suitable tool for biomedical signal processing', *Biomed. Signal Process. Control*, 2014, **14**, (1), pp. 19–29

[57] Vincent, P., Larochelle, H., Bengio, Y., *et al.*: 'Extracting and composing robust features with denoising autoencoders'. Proc. of the 25th Int. Conf. on Machine Learning, Helsinki, Finland, 2008, pp. 1096–1103

[58] Vincent, P., Larochelle, H., Lajoie, I., *et al.*: 'Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion', *J. Mach. Learn. Res.*, 2010, **11**, pp. 3371–3408

[59] Xiong, P., Wang, H., Liu, M., *et al.*: 'A stacked contractive denoising auto-encoder for ECG signal denoising', *Physiol. Meas.*, 2016, **37**, (12), pp. 2214–2230

[60] Long, J., Shelhamer, E., Darrell, T.: 'Fully convolutional networks for semantic segmentation', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2017, **39**, (4), pp. 640–651

[61] Ioffe, S., Szegedy, C.: 'Batch normalization: accelerating deep network training by reducing internal covariate shift'. 32nd Int. Conf. Machine Learning. ICML 2015, Lille, France, 2015, vol. 1, pp. 448–456

[62] Clevert, D.A., Unterthiner, T., Hochreiter, S.: 'Fast and accurate deep network learning by exponential linear units (ELUs)'. 4th Int. Conf. Learning Represent. ICLR 2016 - Conf. Track Proc., San Juan, Puerto Rico, 2016, pp. 1–14

[63] Xiong, P., Wang, H., Liu, M., *et al.*: 'ECG signal enhancement based on improved denoising auto-encoder', *Eng. Appl. Artif. Intell.*, 2016, **52**, pp. 194–202

[64] Wang, J., Li, R., Li, R., *et al.*: 'Adversarial de-noising of electrocardiogram', *Neurocomputing*, 2019, **349**, pp. 212–224

[65] Singh, P., Pradhan, G.: 'A new ECG denoising framework using generative adversarial network', *IEEE/ACM Trans. Comput. Biol. Bioinf.*, 2020, doi: 10.1109/TCBB.2020.2976981, (Early Access)

[66] Antczak, K.: 'Deep recurrent neural networks for ECG signal denoising', 2018, pp. 1–8. Available at http://arxiv.org/abs/1807.11551

[67] Akansu, A.N., Haddad, P.A., Haddad, R.A., *et al.*: '*Multiresolution signal decomposition: transforms, subbands, and wavelets*' (Academic Press, USA, 2001, 2nd edn.)

[68] Chen, B., Li, Y., Zeng, N.: 'Centralized wavelet multiresolution for exact translation invariant processing of ECG signals', *IEEE Access*, 2019, **7**, pp. 42322–42330

[69] Nason, G.P.: 'Choice of the threshold parameter in wavelet function estimation', *Wavelets Stat.*, 1995, **2**, pp. 261–280

[70] Awal, M.A., Mostafa, S.S., Ahmad, M., *et al.*: 'An adaptive level dependent wavelet thresholding for ECG denoising', *Biocybern. Biomed. Eng.*, 2014, **34**, (4), pp. 238–249

[71] Donoho, D.L., Johnstone, I.M., Kerkyacharian, G., *et al.*: 'Asymptopia? Shrinkage: wavelet', *J. R. Stat. Soc. B*, 2002, **57**, (2), pp. 301–369

[72] Lin, H.Y., Liang, S.Y., Ho, Y.L., *et al.*: 'Discrete-wavelet-transform-based noise removal and feature extraction for ECG signals', *IRBM*, 2014, **35**, (6), pp. 351–361

[73] Sawant, C., Patii, H.T.: 'Wavelet based ECG signal de-noising'. IEEE First Int. Conf. on Networks & Soft Computing (ICNSC2014), Guntur, India, 2014, pp. 20–24

[74] Seljuq, U., Himayun, F., Rasheed, H.: 'Selection of an optimal mother wavelet basis function for ECG signal denoising'. 17th IEEE Int. Multi Top. Conf. Collab. Sustain. Dev. Technol. IEEE INMIC 2014 - Proc., Karachi, Pakistan, 2014, pp. 26–30

[75] Li, D., Zhang, H., Zhang, M.: 'Wavelet de-noising and genetic algorithm-based least squares twin SVM for classification of arrhythmias', *Circuits Syst. Signal Process.*, 2017, **36**, (7), pp. 2828–2846

[76] El-Dahshan, E.S.A.: 'Genetic algorithm and wavelet hybrid scheme for ECG signal denoising', *Telecommun. Syst.*, 2011, **46**, (3), pp. 209–215

[77] Biswas, U., Hasan, K.R., Sana, B., *et al.*: 'Denoising ECG signal using different wavelet families and comparison with other techniques'. 2nd Int. Conf. Electrical Engineering Information Communication Technology. iCEEiCT 2015, Dhaka, 2015, pp. 1–6

[78] Banerjee, S., Gupta, R., Mitra, M.: 'Delineation of ECG characteristic features using multiresolution wavelet analysis method', *Meas. J. Int. Meas. Confed.*, 2012, **45**, (3), pp. 474–487

[79] Donoho, D.L., Johnstone, I.M.: 'Ideal spatial adaptation by wavelet shrinkage', *Biometrika*, 1994, **81**, (3), pp. 425–455

[80] Stein, C.M.: 'Estimation of the mean of a multivariate normal distribution', *Ann. Stat.*, 1981, **9**, (6), pp. 1135–1151

[81] Nason, G.P.: 'Wavelet function estimation using cross-validation', *Lect. Notes Stat.*, 1995, **1**, (103), pp. 261–261

[82] Nason, G.P.: 'The WaveThresh package: wavelet transform and thresholding software for S', *Journal of Computational and Graphical Statistics*, 1994, **3**, (2), pp. 163–191

[83] Neumann, M.H., Spokoiny, V.G.: 'On the efficiency of wavelet estimators under arbitrary error distributions', *Math. Meth. Stat.*, 1995, **4**, (2), pp. 137–166

[84] Weyrich, N., Warhola, G.T.: 'De-noising using wavelets and cross validation', '*Approximation theory, wavelets applications*' (Kluwer Academic Publishers, Dordrecht, Netherlands, 1995), pp. 523–532

[85] Wang, Y.: 'Function estimation via wavelets for data with long-range dependence'. Proc. of 1994 Workshop on Information Theory and Statistics, Alexandria, USA, 1994, p. 100

[86] Donoho, D.L., Johnstone, L.M.: 'Adapting to unknown smoothness via wavelet shrinkage', *J. Am. Stat. Assoc.*, 1995, **90**, (432), pp. 1200–1224

[87] Abramovich, F., Benjamini, Y.: 'Adaptive thresholding of wavelet coefficients', *Comput. Stat. Data Anal.*, 1996, **22**, (4), pp. 351–361

[88] Vidakovic, B.: 'Nonlinear wavelet shrinkage with Bayes rules and Bayes factors', *J. Am. Stat. Assoc.*, 1998, **93**, (441), pp. 173–179

[89] Ogden, T., Parzen, E.: 'Data dependent wavelet thresholding in nonparametric regression with change-point applications', *Comput. Stat. Data Anal.*, 1996, **22**, (1), pp. 53–70

[90] Smith, C.B., Agaian, S., Akopian, D.: 'A wavelet-denoising approach using polynomial threshold operators', *IEEE Signal Process. Lett.*, 2008, **15**, pp. 906–909

[91] Tulsani, H., Gupta, R.: 'Wavelet based optimized polynomial threshold function for ECG signal denoising'. 2015 IEEE 2nd Int. Conf. on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 2015, pp. 1563–1566

[92] Zhang, S., Gao, J., Yang, J., *et al.*: 'A Mallat based wavelet ECG de-noising algorithm', *Appl. Mech. Mater.*, 2013, **263**, pp. 2267–2270

[93] Marco, G., Alberto, B., Taian, V.: 'A novel wavelet-based filtering strategy to remove powerline interference from electrocardiograms with atrial fibrillation', *Physiol. Meas.*, 2017, **38**, (5), p. R27

[94] Alfaouri, M., Daqrouq, K.: 'ECG signal denoising by wavelet transform thresholding', *Am. J. Appl. Sci.*, 2008, **5**, (3), pp. 276–281

[95] Zhang, J., Lin, J.L., Li, X.L., *et al.*: 'ECG signals denoising method based on improved wavelet threshold algorithm'. Proc. 2016 IEEE Advanced Information Management Communicates Electronic Automation Control Conf. (IMCEC), Xi'an, 2016, pp. 1779–1784

[96] Mithun, P., Pandey, P.C., Sebastian, T., *et al.*: 'A wavelet based technique for suppression of EMG noise and motion artifact in ambulatory ECG'. Proc. Annual Int. Conf. IEEE Engineering Medicine Biology Society (EMBS), Boston, MA, 2011, pp. 7087–7090

[97] Singh, B.N., Tiwari, A.K.: 'Optimal selection of wavelet basis function applied to ECG signal denoising', *Digit. Signal Process. A Rev. J.*, 2006, **16**, (3), pp. 275–287

[98] Han, G., Xu, Z.: 'Electrocardiogram signal denoising based on a new improved wavelet thresholding', *Rev. Sci. Instrum.*, 2016, **87**, (8), p. 084303

[99] Knight, M.I., Nason, G.P.: 'A 'nondecimated' lifting transform', *Stat. Comput.*, 2009, **19**, (1), pp. 1–16

[100] Percival, D., Andrew, W.: '*Wavelet methods for time series analysis*' (Cambridge University Press, United Kingdom, 2006, 1st edn.)

[101] Mallaparapu, K., Krishna, B.A., Masthan, S., *et al.*: 'Analysis of denoising on different signals using new thresholding function'. IEEE Conf. Signal Processing Communication Engineering Systems (SPACES 2018), Vijaywada, 2018, pp. 154–162

[102] Wang, C.Y., Huang, D.J., Xie, X.Q., *et al.*: 'A lifting double-wavelet algorithm for ECG signal denoising'. 2016 IEEE Int. Conf. Consumer Electronics (ICCE-TW 2016), Taiwan, 2016, pp. 1–2

[103] Patil, H.T., Holambe, R.S.: 'New approach of threshold estimation for denoising ECG signal using wavelet transform'. Annual IEEE India Conf. (INDICON 2013), Mumbai, India, 2013, pp. 1–4

[104] Yi, L., Song, Y.Q.: 'A method of wavelet-based dual thresholding de-noising for ECG signal'. 7th IEEE Int. Congress Image Signal Processing (CISP 2014), Dalian, 2014, pp. 1085–1089

[105] Poornachandra, S.: 'Wavelet-based denoising using subband dependent threshold for ECG signals', *Digit. Signal Process. A Rev. J.*, 2008, **18**, (1), pp. 49–55

[106] Kumar, A., Komaragiri, R., Kumar, M.: 'Design of wavelet transform based electrocardiogram monitoring system', *ISA Trans.*, 2018, **80**, pp. 381–398

[107] Kumar, A., Komaragiri, R., Kumar, M.: 'Heart rate monitoring and therapeutic devices: a wavelet transform based approach for the modeling and classification of congestive heart failure', *ISA Trans.*, 2018, **79**, pp. 239–250

[108] Hernández, O., Olvera, E.: 'Noise cancellation on ECG and heart rate signals using the undecimated wavelet transform'. Int. Conf. on eHealth, Telemedicine, and Social Medicine, Cancun, Mexico, 2009, pp. 145–150

[109] Singh, O., Sunkaria, R.K.: 'ECG signal denoising via empirical wavelet transform', *Australas. Phys. Eng. Sci. Med.*, 2017, **40**, (1), pp. 219–229

[110] Sayadi, O., Shamsollahi, M.B.: 'Multiadaptive bionic wavelet transform: application to ECG denoising and baseline wandering reduction', *EURASIP J. Adv. Signal Process.*, 2007, **2007**, p. 11, Article ID 41274, doi:10.1155/2007/41274

[111] Amindavar, H., Naraghi, F.: 'Wavelet-based variational Bayesian ECG denoising'. 3rd Int. Conf. on Pattern Analysis and Image Analysis, IPRIA 2017, Shahrekord, Iran, 2017, pp. 7–12

[112] Jenkal, W., Latif, R., Toumanari, A., *et al.*: 'An efficient algorithm of ECG signal denoising using the adaptive dual threshold filter and the discrete wavelet transform', *Biocybern. Biomed. Eng.*, 2016, **36**, (3), pp. 499–508

[113] Zhu, J., Li, X.: 'Electrocardiograph signal denoising based on sparse decomposition', *Healthc. Technol. Lett.*, 2017, **4**, (4), pp. 134–137

[114] Candes, E.J., Tao, T.: 'Near optimal signal recovery from random projections: universal encoding strategies?', *IEEE Trans. Inf. Theory*, 2006, **52**, (12), pp. 5406–5425

[115] Peyré, G.: 'Best basis compressed sensing', *IEEE Trans. Signal Process.*, 2010, **58**, (5), pp. 2613–2622

[116] Tropp, J.A.: 'Greed is good: algorithmic results for sparse approximation', *IEEE Trans. Inf. Theory*, 2004, **50**, (10), pp. 2231–2242

[117] Tropp, J.A., Gilbert, A.C.: 'Signal recovery from random measurements via orthogonal matching pursuit', *IEEE Trans. Inf. Theory*, 2007, **53**, (12), pp. 4655–4666

[118] Chen, S.S., Donoho, D.L., Saunders, M.A.: 'Atomic decomposition by basis pursuit', *SIAM Rev.*, 2001, **43**, (1), pp. 129–159

[119] Mourad, N.: 'ECG denoising algorithm based on group sparsity and singular spectrum analysis', *Biomed. Signal Process. Control*, 2019, **50**, pp. 62–71

[120] Wang, Z., Wan, F., Wong, C.M., *et al.*: 'Adaptive Fourier decomposition based ECG denoising', *Comput. Biol. Med.*, 2016, **77**, pp. 195–205

[121] Harris, T.J., Yuan, H.: 'Filtering and frequency interpretations of singular spectrum analysis', *Phys. D, Nonlinear Phenom.*, 2010, **239**, (20–22), pp. 1958–1967

[122] Rakshit, M., Das, S.: 'Hybrid approach for ECG signal enhancement using dictionary learning-based sparse representation', *IET Sci. Meas. Technol.*, 2019, **13**, (3), pp. 381–391

[123] Selesnick, I.: 'Sparse regularization via convex analysis', *IEEE Trans. Signal Process.*, 2017, **65**, (17), pp. 4481–4494

[124] Devi, R., Tyagi, H.K., Kumar, D.: 'Performance comparison and applications of sparsity based techniques for denoising of ECG signal'. 2019 6th Int. Conf. on Signal Processing and Integrated Networks, SPIN 2019, Noida, India, 2019, pp. 346–351

[125] Jin, Z., Dong, A., Shu, M., *et al.*: 'Sparse ECG denoising with generalized minimax concave penalty', *Sensors (Switzerland)*, 2019, **19**, (7), pp. 1–21

[126] Satija, U., Ramkumar, B., Manikandan, M.S.: 'Noise-aware dictionary-learning-based sparse representation framework for detection and removal of single and combined noises from ECG signal', *Healthc. Technol. Lett.*, 2017, **4**, (1), pp. 2–12

[127] McSharry, P.E., Clifford, G.D., Tarassenko, L., *et al.*: 'A dynamical model for generating synthetic electrocardiogram signals', *IEEE Trans. Biomed. Eng.*, 2003, **50**, (3), pp. 289–294

[128] Sameni, R., Shamsollahi, M.B., Jutten, C., *et al.*: 'Filtering noisy ECG signals using the extended Kalman filter based on a modified dynamic ECG model', *Comput. Cardiol.*, 2005, **32**, pp. 1017–1020

[129] Hesar, H.D., Mohebbi, M.: 'ECG denoising using marginalized particle extended Kalman filter with an automatic particle weighting strategy', *IEEE J. Biomed. Health Inf.*, 2017, **21**, (3), pp. 635–644

[130] Sayadi, O., Shamsollahi, M.B.: 'ECG denoising and compression using a modified extended Kalman filter structure', *IEEE Trans. Biomed. Eng.*, 2008, **55**, (9), pp. 2240–2248

[131] Hesar, H.D., Mohebbi, M.: 'An adaptive particle weighting strategy for ECG denoising using marginalized particle extended kalman filter: an evaluation in arrhythmia contexts', *IEEE J. Biomed. Health Inf.*, 2017, **21**, (6), pp. 1581–1592

[132] Akhbari, M., Shamsollahi, M.B., Jutten, C., *et al.*: 'ECG denoising and fiducial point extraction using an extended Kalman filtering framework with linear and nonlinear phase observations', *Physiol. Meas.*, 2016, **37**, (2), pp. 203–226

[133] Niknazar, M., Rivet, B., Jutten, C.: 'Application of dynamic time warping on Kalman filtering framework for abnormal ECG filtering'. ESANN 2012 Proc., 20th European Symp. Artificial Neural Networks, Computational Intelligence Machine Learning, Bruges, Belgium, April 2012, pp. 139–144

[134] Raghavendra, B.S., Bera, D., Bopardikar, A.S., *et al.*: 'Cardiac arrhythmia detection using dynamic time warping of ECG beats in e-healthcare systems'. 2011 IEEE Int. Symp. a World Wireless, Mobile Multimedia Networks, WoWMoM 2011 - Digit. Proc., Lucca, Italy, 2011, pp. 1–6

[135] Danandeh Hesar, H., Mohebbi, M.: 'An adaptive Kalman filter bank for ECG denoising', *IEEE J. Biomed. Health Inf.*, 2020, doi: 10.1109/JBHI.2020.2982935 (Early Access)
[136] Rakshit, M., Das, S.: 'An efficient ECG denoising methodology using empirical mode decomposition and adaptive switching mean filter', *Biomed. Signal Process. Control*, 2018, **40**, pp. 140–148
[137] Singh, P., Pradhan, G.: 'Variational mode decomposition based ECG denoising using non-local means and wavelet domain filtering', *Australas. Phys. Eng. Sci. Med.*, 2018, **41**, (4), pp. 891–904
[138] Jain, S., Bajaj, V., Kumar, A.: 'Riemann liouvelle fractional integral based empirical mode decomposition for ECG denoising', *IEEE J. Biomed. Health Inf.*, 2018, **22**, (4), pp. 1133–1139
[139] Rajankar, S.O., Talbar, S.N.: 'An optimum ECG denoising with wavelet neural network'. 2015 Int. Conf. on Pervasive Computing (ICPC), Pune, 2015, pp. 1–4
[140] Kabir, M.A., Shahnaz, C.: 'Denoising of ECG signals based on noise reduction algorithms in EMD and wavelet domains', *Biomed. Signal Process. Control*, 2012, **7**, (5), pp. 481–489
[141] Kumar, S., Panigrahy, D., Sahu, P.K.: 'Denoising of electrocardiogram (ECG) signal by using empirical mode decomposition (EMD) with non-local mean (NLM) technique', *Biocybern. Biomed. Eng.*, 2018, **38**, (2), pp. 297–312
[142] Li, N., Li, P.: 'An improved algorithm based on EMD-wavelet for ECG signal de-noising'. 2009 Int. Joint Conf. on Computational Sciences and Optimization, Sanya, Hainan, 2009, pp. 825–827
[143] Lu, Y., Yan, J., Yam, Y.: 'Model-based ECG denoising using empirical mode decomposition'. 2009 IEEE Int. Conf. on Bioinformatics and Biomedicine, Washington, DC, 2009, pp. 191–196
[144] Lahmiri, S.: 'Comparative study of ECG signal denoising by wavelet thresholding in empirical and variational mode decomposition domains', *Healthc. Technol. Lett.*, 2014, **1**, (3), pp. 104–109
[145] Singh, P., Pradhan, G., Shahnawazuddin, S.: 'Denoising of ECG signal by non-local estimation of approximation coefficients in DWT', *Biocybern. Biomed. Eng.*, 2017, **37**, (3), pp. 599–610
[146] Vargas, R.N., Veiga, A.C.P.: 'Electrocardiogram signal denoising by clustering and soft thresholding', *IET Signal Process.*, 2018, **12**, (9), pp. 1165–1171
[147] Kumar, A., Komaragiri, R., Kumar, M.: 'Design of efficient fractional operator for ECG signal detection in implantable cardiac pacemaker systems', *Int. J. Circuit Theory Appl.*, 2019, **47**, (9), pp. 1459–1476
[148] Yadav, S.K., Sinha, R., Bora, P.K.: 'Electrocardiogram signal denoising using non-local wavelet transform domain filtering', *IET Signal Process.*, 2015, **9**, (1), pp. 88–96
[149] Moody, G.B., Mark, R.G.: 'The impact of the MIT-BIH arrhythmia database', *IEEE Eng. Med. Biol. Mag.*, 2001, **20**, (3), pp. 45–50
[150] Goldberger, A., Maral, L., Glass, L., *et al.*: 'The MIT-BIH normal sinus rhythm database – PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals', *Circulation*, 2000, **101**, (23), pp. 215–220
[151] Moody, G.B., Muldrow, W.E, Mark, R.G.: 'A noise stress test for arrhythmia detectors', *Comput. Cardiol.*, 1984, **11**, pp. 381–384
[152] Bousseljot, R., Kreiseler, D., Schnabel, A.: 'Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das internet', *Biomed. Tech., Biomed. Eng.*, 1995, **40**, (s1), pp. 317–318
[153] Baim, D.S., Colucci, W.S., Monrad, E.S., *et al.*: 'Survival of patients with severe congestive heart failure treated with oral milrinone', *J. Am. Coll. Cardiol.*, 1986, **7**, (3), pp. 661–670
[154] Elgendi, M.: 'Fast QRS detection with an optimized knowledge-based method: evaluation on 11 standard ECG databases', *PLOS One*, 2013, **8**, (9), p. e73557
[155] Goldberger, A.L., Amaral, L.A., Glass, L., *et al.*: 'Physiobank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals', *Circulation*, 2000, **101**, (23), pp. e215–e220
[156] Creighton University Ventricular Tachyarrhythmia Database. Available at https://physionet.org/physiobank/database/cudb/
[157] Sudden Cardiac Death Holter Database. Available at https://physionet.org/content/sddb/1.0.0/
[158] Kumar, A., Komaragiri, R., Kumar, M.: 'From pacemaker to wearable: techniques for ECG detection systems', *J Med Syst.*, 2018, **42**, (34), https://doi.org/10.1007/s10916-017-0886-1
[159] Thakur, R.S., Yadav, R.N., Gupta, L.: 'State-of-art analysis of image denoising methods using convolutional neural networks', *IET Image Process.*, 2019, **13**, (13), pp. 2367–2380