# Denoising of EEG Signals in Brain–Computer Interfaces Using Extended Kalman-Filtered Recurrent High-Order Neural Networks

Enrique Martínez-Madrid, Jesús A. Medrano-Hermosillo, Juan. A. Ramírez-Quintana, Iván R. Urbina-Leos, Oscar J. Suarez, *Senior Member, IEEE*, and Francisco Ronay Lopez Estrada

***Abstract*—This paper presents a Recurrent High-Order Neural Network (RHONN) for EEG denoising in brain–computer interface (BCI) applications. Trained online using the Extended Kalman Filter (EKF), the proposed approach effectively suppresses EOG, EMG, and ECG artifacts under non-stationary conditions while preserving EEG temporal structure. Experimental results show that RHONN achieves performance comparable to or better than conventional filters and deep learning models, including MLPs, RNNs, GANs, Transformers and autoencoders. A key advantage of the RHONN–EKF framework is its very low computational cost. By modeling each EEG channel with a single high-order neuron, the method reduces computational load by more than 90% compared to state-of-the-art models, making it suitable for real-time, resource-constrained BCI systems and consistent with Green AI principles.**

**Link to graphical and video abstracts, and to code:**
**https://latamt.ieeer9.org/index.php/transactions/article/view/10557**

***Index Terms*—Electroencephalography, brain–computer interface,Recurrent High Order Neural Network, EKF, real-time signal processing, noise reduction**

## I. Introduction

ELECTROENCEPHALOGRAPHY (EEG) is a noninvasive and widely accessible tool for monitoring brain activity in clinical and brain–computer interface (BCI) applications [1]–[3]. Advances in consumer-grade devices and signal processing have improved data quality, interpretability, and robustness [4]–[6], enabling applications such as emotion recognition and driver distraction detection [7], [8]. EEG is essential in medicine for diagnosing and treating neurological disorders and in BCIs for rehabilitation and communication in patients with severe motor impairments [9]. However, EEG signals are highly susceptible to noise and artifacts, and their non-stationary, subject-dependent nature limits generalization, complicating analysis and requiring extensive preprocessing for deep learning models [10], [11]. Deploying these models in real-time, resource-constrained systems remains challenging due to high computational cost [12]–[15]. To address these issues, in this work we propose a RHONN trained with an EKF for EEG denoising. This nonlinear, dynamic approach reduces computational complexity compared to conventional methods and handles each EEG channel independently, making it suitable for clinical and real-time BCI applications. See Fig. 1.

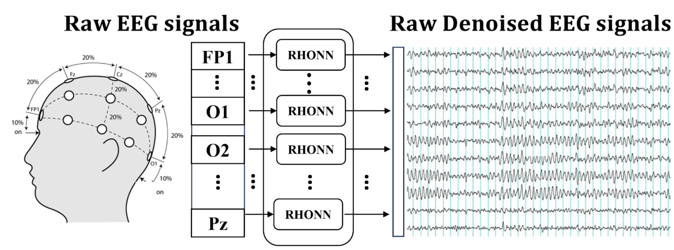


Fig. 1. Proposed block diagram of a multichannel EEG noise reduction system based on RHONN.

The remainder of this article is organized as follows. Section 2 reviews related work, while Section 3 presents the mathematical foundations of high-order neural networks. Section 4 describes the selected dataset, and Section 5 outlines the applied preprocessing steps. The evaluation methodology is detailed in Section 6, followed by the results and comparison with state-of-the-art models in Section 7. Finally, the main conclusions and future research directions are presented.

## II. State of the Art

A wide range of approaches have been proposed for EEG artifact removal, spanning classical signal processing, statistical modeling, and, more recently, deep learning-based techniques [16]. Among the most established methods, time–frequency techniques such as the wavelet transform and short-time Fourier transform (STFT) [17], [18] have been extensively used due to their capability to represent EEG signals in both time and frequency domains. These approaches are particularly effective for nonstationary signals, allowing transient artifacts such as muscle activity or eye blinks to be localized and attenuated. For instance, wavelet-based methods decompose EEG signals into multiresolution components, enabling selective filtering of noise-dominated coefficients, while STFT-based approaches provide a sliding-window spectral representation that facilitates artifact identification in specific time intervals.

The associate editor coordinating the review of this manuscript and approving it for publication was Gladston Moreira (*Corresponding author: Jesús Alfonso Medrano-Hermosillo*).

E. Martínez-Madrid, Jesús Alfonso Medrano-Hermosillo, J. A. Ramírez-Quintana, and I. R. Urbina-Leos are with Tecnológico Nacional de México, Chihuahua, Chihuahua 31310, Mexico (e-mails: m24060173@chihuahua.tecnm.mx, jesus.mh@chihuahua.tecnm.mx, juan.rq@chihuahua.tecnm.mx, and m24061301@chihuahua.tecnm.mx).

O. J. Suarez is with Universidad de Pamplona, Pamplona, Norte de Santander, Colombia (e-mail: oscar.suarez@unipamplona.edu.co).

F. R. Lopez-Estrada is with Tecnológico Nacional de México, Tuxtla Gutiérrez, Chiapas 29050, Mexico (e-mail: frlopez@tuxtla.tecnm.mx).

In parallel, conventional filtering and decomposition techniques have been widely adopted as baseline solutions. Finite impulse response (FIR) filters and notch filters are commonly used to suppress power-line interference and narrowband noise [19], whereas band-pass filters (BPF) isolate physiologically relevant EEG frequency bands such as delta, theta, alpha, and beta rhythms. More advanced decomposition methods, such as variational mode decomposition (VMD) and wavelet packet transform (WPT), have been proposed to further enhance signal separation by adaptively decomposing EEG signals into intrinsic modes or subbands [20]. Hybrid approaches combining wavelet transforms with conventional filters have also demonstrated improved performance in removing artifacts while preserving signal integrity [21]. Despite their effectiveness, these techniques typically require careful manual tuning of parameters such as filter order, cutoff frequencies, and decomposition levels, as well as prior knowledge of the spectral characteristics of both the signal and the noise. Moreover, recent studies have explored improved time–frequency frameworks that integrate adaptive or empirical decomposition strategies to better handle motion artifacts and nonlinear distortions in EEG recordings [18]. While these methods provide enhanced flexibility, they still rely heavily on heuristic parameter selection and may suffer from limited generalization across different subjects or recording conditions. Consequently, although classical signal processing techniques remain computationally efficient and interpretable, their dependency on manual configuration and limited adaptability restrict their applicability in dynamic and real-world EEG environments. On the other hand, statistical approaches, such as Kalman filtering and independent component analysis (ICA) [14], [22]–[24], provide adaptive and data-driven solutions for artifact separation. Kalman-based methods model EEG signals as stochastic dynamic systems, enabling recursive estimation and real-time tracking of noise and underlying neural activity. In contrast, ICA decomposes multichannel EEG recordings into statistically independent components, facilitating the identification and removal of artifact-related sources such as ocular and muscular activity. Despite their effectiveness, these approaches rely on assumptions such as linearity, Gaussianity, or stationarity, which are often violated in real EEG signals. Additionally, ICA-based methods may require manual component selection, while Kalman-based techniques can become computationally demanding in nonlinear settings.

In recent years, deep learning-based methods have gained significant attention due to their ability to learn complex nonlinear representations directly from data [10], [12], [25]. Architectures such as convolutional neural networks (CNNs) have been widely adopted for their capability to extract local temporal and spatial features from EEG signals, while recurrent models, including LSTMs, are effective in capturing temporal dependencies and sequential patterns [26], [27]. Autoencoder-based approaches have been extensively used for unsupervised denoising, learning compact latent representations that preserve relevant neural activity while suppressing artifacts [28]. Similarly, generative adversarial networks (GANs) have shown promising results by learning the distribution of clean EEG signals and generating artifact-free reconstructions [29]. In addition, Transformer-based models and hybrid architectures have been introduced to capture long-range dependencies and global contextual information, significantly improving denoising performance, particularly in complex and highly corrupted signals [30]–[32]. These models are capable of handling the nonlinear and nonstationary nature of EEG signals, making them highly effective for artifact removal. However, their performance often comes at the cost of increased computational complexity, requiring large annotated datasets, extensive training procedures, and significant hardware resources. This limits their applicability in real-time systems or embedded platforms. Furthermore, challenges related to generalization across subjects and varying recording conditions remain an open issue, as highlighted in recent studies [6], [10], [11].

However, despite their accuracy, these approaches often require large datasets and high computational resources, limiting their use in real-time applications. In addition, their performance may degrade under unseen conditions, raising concerns about generalization [6], [10], [11]. To address these limitations, the proposed RHONN model introduces a nonlinear adaptive learning framework that ensures efficient real-time implementation while maintaining robustness and stability, making it suitable for multichannel EEG denoising.

## III. Preliminaries

### A. RHONN Model

The general mathematical formulation of a discrete-time RHONN model, which extends the first-order Hopfield neural network [33], [34], is mathematically defined as

$$\chi_i(k+1) = \omega_i^\top(k)\, z_i\big(x(k), u(k)\big), \quad y_i(k) = \chi_i(k), \tag{1}$$

$(i = 1, \cdots, n)$ where $\chi_i$ is the state of the $i$-th neuron,$y_i(k)$ is the output of the neural network, $u(k)$ is the system input, $n$ is the state dimension, $\omega_i(k)$ is the vector of adaptive weights and $z_i(x(k), u(k))$ is defined as the nonlinear regressor vector

$$z_i(x(k), u(k)) = \begin{bmatrix} z_{i1} \\ z_{i2} \\ \vdots \\ z_{iL} \end{bmatrix} = \begin{bmatrix} \prod_{j\in I_1} \xi_{i_j}^{d_{i_j}(1)} \\ \prod_{j\in I_2} \xi_{i_j}^{d_{i_j}(2)} \\ \vdots \\ \prod_{j\in I_{L_i}} \xi_{i_j}^{d_{i_j}(L)} \end{bmatrix}, \tag{2}$$

where, $L$ indicates the number of higher-order connections, while $\{I_1, I_2, \ldots, I_L\}$ corresponds to a collection of unordered subsets of the set $\{1, 2, \ldots, n\}$. The term $d_{i_j}(k)$ is defined as a non-negative integer, with $\xi_i$ defined by

$$\begin{aligned} \xi_i &= \begin{bmatrix} \xi_{i_1} & \cdots & \xi_{i_n} & \xi_{i_{n+1}} & \cdots & \xi_{i_{n+m}} \end{bmatrix}^T \\ &= \begin{bmatrix} S(x_1) & \cdots & S(x_n) & u_1 & \cdots & u_m \end{bmatrix}^T . \end{aligned} \tag{3}$$

### B. Training Algorithm for RHONN

The extended Kalman filter is a nonlinear extension of the Kalman filter designed to estimate the states of nonlinear systems subject to additive process and measurement noise

[33], [35], [36]. For training neural networks using the EKF, the network weights are treated as the states to be estimated. In this context, the error between the network output and the measured EEG signal is modeled as additive white noise. Accordingly, the objective of training is to determine the optimal weight values that minimize the prediction error. The corresponding equations for the algorithm are

$$\omega_i(k+1) = \omega_i(k) + \eta_i\, K_i(k)\, e_i(k), \tag{4}$$

$$K_i(k) = P_i(k)\, H_i(k)\, \left[R_i(k) + H_i^T(k)\, P_i(k)\, H_i(k)\right]^{-1}, \tag{5}$$

$$P_i(k+1) = P_i(k) - K_i(k)\, H_i(k)\, P_i(k) + Q_i(k), \tag{6}$$

and

$$e_i(k) = gt_i(k) - y_i(k), \tag{7}$$

where $e_i(k)$ denotes the error at time $k$ for the $i$-th output, $gt_i(k)$ represents the desired nonlinear reference signal, and $y_i(k)$ is the output of the neural network at time $k$. $P_i(k) \in \mathbb{R}^{L_i \times L_i}$ denotes the error covariance matrix of the weight estimates, and $w_i \in \mathbb{R}^{L_i}$ is the vector of network weights, where $L_i$ is the total number of weights in the network. The scalar $\eta_i$ is the learning rate, $K_i(k) \in \mathbb{R}^{L_i \times m}$ is the Kalman gain matrix, $Q_i(k) \in \mathbb{R}^{L_i \times L_i}$ is the process noise covariance matrix, and $H_i(k) \in \mathbb{R}^{L_i \times m}$ is the observation matrix. Each row of $H_{ij}(k)$ is defined as [36]

$$H_{ij} = \left[\frac{\partial\, y_i(k)}{\partial w_{ij}}\right]^{\top}. \tag{8}$$

With $i = 1, ..., n$ and $j =, ...L_i$. Typically, $P_i$, $Q_i$, and $R_i$ are initialized as diagonal matrices with entries $P_i(0)$, $Q_i(0)$, and $R_i(0)$, respectively. The schematic diagram used for the training process is shown in Fig. 2.

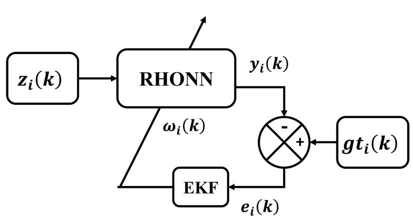


Fig. 2. Training scheme of the RHONN using the EKF.

*1) RHONN Convergence:*

**Theorem 1.** *Consider the $i$-th neuron of the RHONN model presented in* (1)*, where $z_i(\cdot)$ is defined by* (2)–(3)*, and the tracking error is defined in* (7)*. Assume that there exists an ideal constant parameter vector $\omega_i^\star$ such that the desired signal can be expressed as*

$$gt_i(k) = \omega_i^{\star\top} H_i(k) + \nu_i(k), \quad |\nu_i(k)| \le \bar{\nu}_i, \tag{9}$$

*where $H_i(k) := z_i(x(k), u(k))$ denotes the regressor vector, $\tilde{\omega}_i(k) = \omega_i(k) - \omega_i^\star$ is the weight estimation error, and $\nu_i(k)$ represents a bounded approximation error or lumped modeling uncertainty. On the other hand, the EKF-based weight update is given by Eqs.* (4), (5)*, and* (6)*, with $R_i(k) > 0$, $Q_i(k) \ge 0$, $\eta_i > 0$, and bounded regressor $\|H_i(k)\| \le \bar{h}_i$. Based on the previous assumptions, consider the composite Lyapunov candidate function*

$$V_i(k) = \frac{1}{2}e_i^2(k) + \frac{1}{2}\tilde{\omega}_i^\top(k) P_i^{-1}(k)\tilde{\omega}_i(k), \tag{10}$$

*where $P_i(k) > 0$, ensuring that $V_i(k)$ is positive definite. Then, there exist positive constants $\underline{\alpha}_i$, $\underline{\mu}_i$, and $C_i$ such that the one-step difference satisfies*

$$\begin{aligned}\Delta V_i(k) \le &- \underline{\alpha}_i\, e_i^2(k) \\ &- \underline{\mu}_i\, \tilde{\omega}_i^\top(k) H_i(k) R_i^{-1}(k) H_i^\top(k) \tilde{\omega}_i(k) \\ &+ C_i\, \bar{\nu}_i^2.\end{aligned} \tag{11}$$

*Consequently:*

1) *If $\bar{\nu}_i > 0$, the pair $(e_i, \tilde{\omega}_i)$ is* semi-globally uniformly ultimately bounded *(SGUUB).*
2) *If $\bar{\nu}_i = 0$, then $e_i(k) \to 0$ as $k \to \infty$.*
3) *If, in addition, $Q_i(k) \equiv 0$ and $H_i(k)$ is persistently exciting (PE) [37], then $\tilde{\omega}_i(k) \to 0$ as $k \to \infty$.*

*Proof.* Consider the Lyapunov candidate defined in (10), with error $e_i(k)$, parameter error $\tilde{\omega}_i(k)$, covariance $P_i(k) > 0$, regressor $H_i(k)$, noise covariance $R_i(k) > 0$, and Kalman gain [37] defined in (5). From the definition of the desired signal

$$gt_i(k) = \omega_i^{*\top} H_i(k) + \nu_i(k), \tag{12}$$

and the network output $y_i(k) = \omega_i^\top(k) H_i(k)$, the tracking error can be written as

$$e_i(k) = gt_i(k) - y_i(k) = -\tilde{\omega}_i^\top(k) H_i(k) + \nu_i(k), \tag{13}$$

where $|\nu_i(k)| \le \bar{\nu}_i$. Taking the forward difference of the Lyapunov function, we have

$$\Delta V_i(k) = V_i(k+1) - V_i(k) = \Delta V_{e,i}(k) + \Delta V_{\omega,i}(k), \tag{14}$$

and using the EKF parameter update

$$\tilde{\omega}_i(k+1) = \tilde{\omega}_i(k) + \eta_i K_i(k) e_i(k), \tag{15}$$

we obtain

$$\begin{aligned}\Delta V_{\omega,i}(k) = &\frac{1}{2}\tilde{\omega}_i^\top \left[P_i^{-1}(k+1) - P_i^{-1}(k)\right]\tilde{\omega}_i \\ &+ \eta_i e_i\, \tilde{\omega}_i^\top P_i^{-1}(k+1) K_i \\ &+ \frac{1}{2}\eta_i^2 e_i^2\, K_i^\top P_i^{-1}(k+1) K_i.\end{aligned} \tag{16}$$

Using the Joseph-form inequality [38]

$$P_i^{-1}(k+1) \ge P_i^{-1}(k) + H_i(k) R_i^{-1}(k) H_i^\top(k), \tag{17}$$

we obtain

$$\begin{aligned}&\Delta V_{\omega,i}(k) - \frac{1}{2}\tilde{\omega}_i^\top H_i R_i^{-1} H_i^\top \tilde{\omega}_i \\ &\quad\le \eta_i e_i\, \tilde{\omega}_i^\top P_i^{-1}(k+1) K_i + \frac{1}{2}\eta_i^2 e_i^2\, K_i^\top P_i^{-1}(k+1) K_i.\end{aligned} \tag{18}$$

By using the identity $P_i^{-1} K_i = H_i (H_i^\top P_i H_i + R_i)^{-1}$ and the matrix inversion lemma, we have

$$K_i^\top P_i^{-1} K_i \le (H_i^\top P_i H_i + R_i)^{-1}, \tag{19}$$

Now, applying Cauchy–Schwarz and Young's inequalities [39], [40], we obtain

$$\eta_i e_i\,\tilde{\omega}_i^\top P_i^{-1} K_i \le \frac{\mu_i}{2}\,\tilde{\omega}_i^\top H_i R_i^{-1} H_i^\top \tilde{\omega}_i + \frac{\eta_i^2}{2\mu_i}\,e_i^2\,(H_i^\top P_i H_i + R_i)^{-1}. \tag{20}$$

Thus,

$$\Delta V_{\omega,i}(k) \le \left(\frac{1}{2}+\frac{\mu_i}{2}\right)\tilde{\omega}_i^\top H_i R_i^{-1} H_i^\top \tilde{\omega}_i + \eta_i^2 e_i^2\, c_{K,i}. \tag{21}$$

Now, considering the error term

$$\Delta V_{e,i}(k) = \frac{1}{2}\big(e_i^2(k+1) - e_i^2(k)\big), \tag{22}$$

and using the bound

$$|e_i(k)| \le \|\tilde{\omega}_i(k)\|\|H_i(k)\| + |\nu_i(k)|, \tag{23}$$

we obtain, using Young's inequality,

$$\Delta V_{e,i}(k) \le -\frac{a_i}{2} e_i^2 + C_{e,i}\bar{\nu}_i^2, \tag{24}$$

for some $C_{e,i} > 0$. Summing both contributions yields

$$\Delta V_i(k) \le \left(\tfrac{1}{2}+\tfrac{\mu_i}{2}\right)\tilde{\omega}_i^\top H_i R_i^{-1} H_i^\top \tilde{\omega}_i - \underline{\alpha}_i e_i^2 + C_i \bar{\nu}_i^2. \tag{25}$$

By rearranging terms, we obtain

$$\Delta V_i(k) \le -\underline{\mu}_i\,\tilde{\omega}_i^\top H_i R_i^{-1} H_i^\top \tilde{\omega}_i - \underline{\alpha}_i\, e_i^2 + C_i\,\bar{\nu}_i^2. \tag{26}$$

Under the conditions $R_i(k) > 0$, $P_i(k) > 0$, bounded regressor, and bounded adaptation gain $\eta_i$, the Lyapunov function is non-increasing up to a bounded term. In the noise-free case ($\bar{\nu}_i = 0$), the inequality becomes negative definite, ensuring asymptotic convergence of both the tracking error and the parameter estimation error. In the presence of bounded disturbances, the system is semi-globally uniformly ultimately bounded (SGUUB).

*Remark on the assumptions:* In the present EEG denoising application, the regressor $H_i(k) = z_i(x(k), u(k))$ remains bounded, as it is constructed from normalized signals and bounded activation functions. The covariance matrices $P_i(k)$ and $R_i(k)$ are initialized as positive definite and remain so under the EKF update, ensuring well-posed adaptation dynamics.

Regarding the persistent excitation (PE) condition, it is only required to guarantee asymptotic convergence of the parameter estimation error $\tilde{\omega}_i(k)$, and is not necessary for the SGUUB or tracking error convergence results. In practice, the nonstationary nature of EEG signals, together with the variability introduced by different artifact conditions, provides sufficiently rich excitation. However, PE is adopted here as a standard sufficient theoretical assumption and is not explicitly verified from the dataset. This treatment is consistent with common practice in adaptive control and system identification literature, where PE is used as a sufficient condition for parameter convergence but is not always directly verified in experimental data [37], [39].

□

*Remark.* The result can be naturally extended to the complete RHONN by stacking all neuron weight vectors as $W(k) = \mathrm{col}_i\big(\omega_i(k)\big)$, and defining the corresponding estimation error $\tilde{W}(k) = W(k) - W^\star$, where $W^\star = \mathrm{col}_i(\omega_i^\star)$. Similarly, define the global error vector $e(k) = \mathrm{col}_i(e_i(k))$, and construct the block-diagonal matrices $P(k)$, $H(k)$, and $R(k)$ with diagonal blocks $P_i(k)$, $H_i(k)$, and $R_i(k)$, respectively. In this case, the global Lyapunov function is given by

$$V(k) = \tfrac{1}{2}\,\|e(k)\|^2 + \tfrac{1}{2}\,\tilde{W}^\top(k) P^{-1}(k)\tilde{W}(k).$$

Due to the block-diagonal structure of the matrices, the Joseph-form inequality and the bounds derived for each neuron extend directly to the full network. Consequently, the following inequality holds

$$\Delta V(k) \le -\,\underline{\alpha}\,\|e(k)\|^2 - \underline{\mu}\,\tilde{W}^\top(k) H(k) R^{-1}(k) H^\top(k) \tilde{W}(k) + C\,\bar{\nu}^2,$$

where $\underline{\alpha} = \min_i \underline{\alpha}_i > 0$, $\underline{\mu} = \min_i \underline{\mu}_i > 0$, and $C = \sum_i C_i$, with $\bar{\nu} = \max_i \bar{\nu}_i$. Therefore, the same stability properties established for each neuron individually extend to the overall stacked system, guaranteeing semi-global uniform ultimate boundedness of the full RHONN under EKF-based adaptation.

### C. Proposed Model

Based on the general definition given in Eq. (1) and considering the case of a single neuron, the proposed model is defined as

$$y(k) = \omega_i^\top(k)\, z_i(k), \tag{27}$$

where the vector $z_i(k)$ was initially constructed by defining the vector $\boldsymbol{\xi}_i(k)$, for this case, $\boldsymbol{\xi}_i(k)$ consists in

$$\xi_i(k) = \begin{bmatrix}\xi_1\\ \xi_2\\ \xi_3\end{bmatrix} = \begin{bmatrix} S(y(k-1))\\ S(y(k-2))\\ x(k)\end{bmatrix}, \tag{28}$$

where $y(k-1)$ and $y(k-2)$, together with the current system input $x(k)$ defined the vector. The higher-order interactions for the vector $z_i(k)$ are constructed by incorporating cross-product terms of the signals contained in $\boldsymbol{\xi}_i(k)$, such that

$$z_i(k) = \begin{bmatrix}\xi_1 & \xi_1\cdot\xi_2 & \xi_1\cdot\xi_3 & \xi_2\cdot\xi_3\end{bmatrix}^T. \tag{29}$$

Finally, the network output is obtained as

$$\begin{aligned} y(k) = {} & \omega_1\cdot\tanh\big(y(k-1)\big)\\ & + \omega_2\cdot\tanh\big(y(k-1)\big)\cdot\tanh\big(y(k-2)\big)\\ & + \omega_3\cdot\tanh\big(y(k-1)\big)\cdot x(k)\\ & + \omega_4\cdot\tanh\big(y(k-2)\big)\cdot x(k)\\ & + \omega_5\cdot x(k). \end{aligned} \tag{30}$$

The diagram representing the architecture of the proposed RHONN is shown in Fig. 3. For the proposed model, the observation matrix is obtained by substituting Eq. (8), resulting in

$$H_i(k) = \frac{\partial\left[\omega_i^\top(k)\, z_i(k)\right]}{\partial\omega_i} = z_i(x(k), u(k)). \tag{31}$$

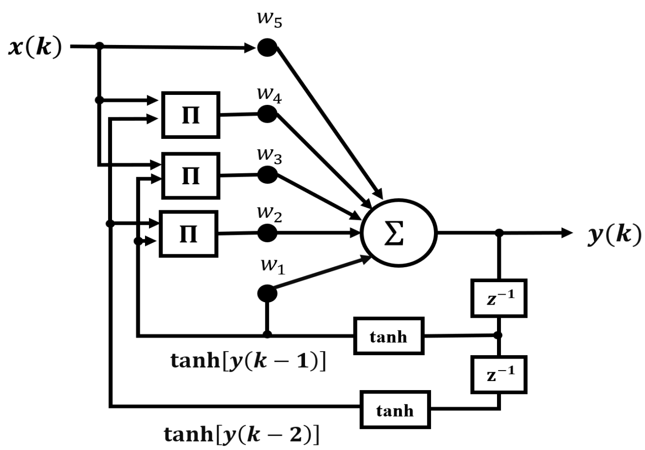


Fig. 3. Architecture of the proposed model.

## IV. Datasets

Two public datasets were used to train and evaluate the proposed network, following standard criteria: EEGDenoiseNet [41] and MIT-BIH Arrhythmia [42], [43]. EEGDenoiseNet [41] is a widely adopted benchmark dataset for EEG denoising in the literature [30], [44], [45]. It comprises EEG, EOG, and EMG recordings from 52 subjects acquired using a 64-channel system at 512 Hz. EEG signals were bandpass filtered (1–80 Hz) and notch filtered, resampled to 256 Hz, and artifact-cleaned using Independent Component Analysis with ICLabel [46], [47]. EOG and EMG signals were filtered (0.3–10 Hz and 1–120 Hz, respectively), resampled, and segmented into 2-second windows. All signals were variance-normalized and visually inspected. The dataset composition is summarized in Table I. The MIT-BIH Arrhythmia Database was considered as a complementary dataset. Although it contains ECG signals, its inclusion is motivated by the fact that cardiac activity can interfere with EEG recordings [1], [23], introducing spectral components that may be misinterpreted as neural activity. This enables the evaluation of the proposed model under more challenging conditions involving physiological artifacts. A summary of the dataset is presented in Table II.

TABLE I
Description of the Data Contained in EEGDenoiseNet

| Signal | Format | Amount |
|---|---|---|
| EEG | Numpy | 4514 |
| EOG | Numpy | 3400 |
| EMG | Numpy | 5598 |

TABLE II
Summary of the MIT-BIH Arrhythmia Database

| Attribute | Value |
|---|---|
| Recordings | 48 |
| Duration | $\sim$30 min |
| Channels | 2 |
| Sampling rate | 256 Hz |

## V. Preprocessing

After dataset selection, several preprocessing steps were applied to ensure consistency across recordings and to properly prepare the data for model training and validation. The following subsections describe the procedures applied to both synthetic EEG signals and real noisy EEG recordings.

### A. Training and Validation Data Preparation

As an initial step for training, validation, and characterization of the proposed model, 4514 clean EEG signals from the EEGDenoiseNet dataset [41] were used as reference data to generate controlled noisy versions.

Noisy signals were generated by adding white Gaussian noise, a common method for simulating EEG contamination [48], resulting in

$$x_{\text{noisy}} = x_{\text{clean}} + \alpha \cdot \mathcal{N}(0, \sigma^2), \tag{32}$$

where $\alpha = 0.20$ (20% noise level). signal was independently normalized using the z-score method [49], computing the mean and standard deviation per signal to avoid data leakage and improve training stability. The dataset was randomly split into 80% for training and 20% for testing at the signal level, preserving the pairing between clean and noisy signals. Although not explicitly subject-independent, this strategy, together with independent noise generation and signal-wise normalization, reduces potential information leakage. The resulting datasets are summarized in Table III, and the preprocessing pipeline is illustrated in Fig. 4.

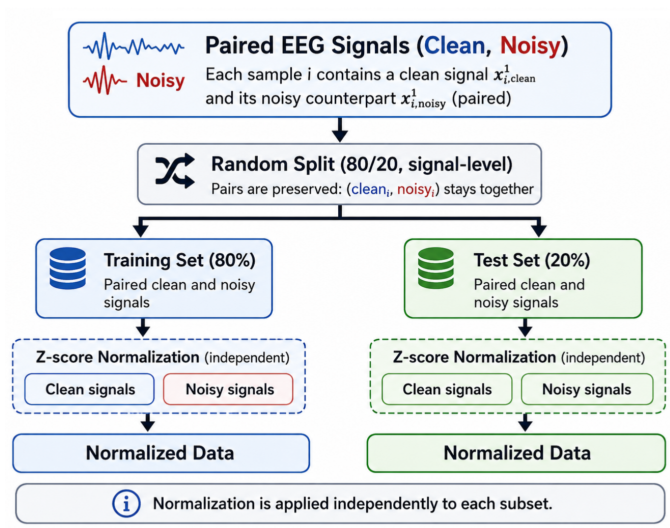


Fig. 4. Diagram of the preprocessing and partitioning of the EEG data.

TABLE III
EEG Dataset Partitions after Preprocessing

| Dataset | Signals |
|---|---|
| EEG_GTtrain_data | 3611 |
| EEG_noisetrain_data | 3611 |
| EEG_GTtest_data | 903 |
| EEG_noisetest_data | 903 |

### B. Artifact-Contaminated EEG Signal Generation

As a second stage of the experimental framework, additional datasets were generated to assess the robustness of the proposed model under controlled contamination by physiological

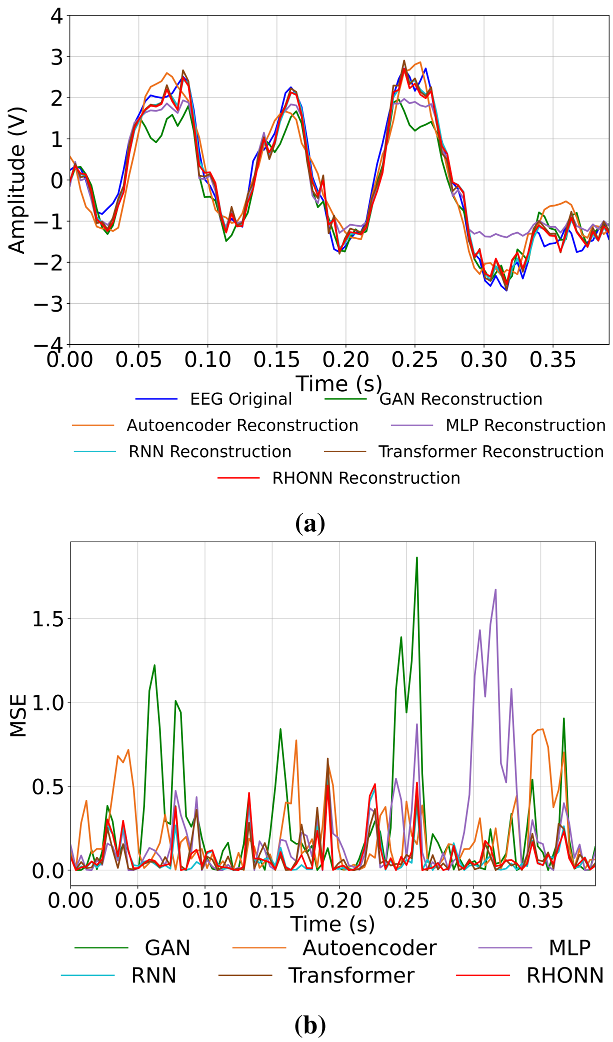


Fig. 5. Comparison of benchmark models: (a) noise reduction performance, and (b) MSE evolution for different denoising techniques.

artifacts, moving beyond the synthetic Gaussian noise considered in the previous subsection toward more realistic interference conditions. To achieve this, contaminated EEG signals were constructed by linearly combining clean EEG segments with physiological artifacts, specifically electrooculographic (EOG), electromyographic (EMG), and electrocardiographic (ECG) signals, extracted from preselected databases, thereby enabling the emulation of biologically plausible distortions commonly observed in real EEG recordings. The contamination process is defined as $x_{\text{noisy}} = x_{\text{clean}} + \lambda \cdot \text{artifact}$, where the parameter $\lambda$ regulates the level of artifact contribution in the resulting signal. To analyze the effect of different contamination levels, two values of $\lambda$ (0.3 and 0.7) were considered. These values were applied to different artifact types, resulting in multiple contamination scenarios. For each condition, 3400 signals were generated, each consisting of 512 samples with a sampling frequency of 256 Hz. A summary of the evaluated scenarios is provided in Table IV.

Prior to evaluation, all signals were independently normalized following the same procedure described in Subsection A and illustrated in Fig. 4.

TABLE IV
ARTIFACT-CONTAMINATED EEG SCENARIOS FOR ROBUSTNESS EVALUATION

| Combination | Number of Signals | Sampling Rate | Length |
|---|---|---|---|
| EEG + EMG | 3400 | 256 Hz | 512 samples |
| EEG + EOG | 3400 | 256 Hz | 512 samples |
| EEG + ECG | 3400 | 256 Hz | 512 samples |
| EEG + ECG + | 3400 | 256 Hz | 512 samples |

## VI. MODEL EVALUATION

To evaluate the denoising capability of the proposed model, its performance was assessed using a set of standard quantitative metrics widely adopted in the EEG signal processing literature [12]–[15], [25], [30], [44], [45], [49]–[52].The overall evaluation procedure is summarized in Algorithm 1.

**Algorithm 1** Evaluation Procedure for EEG Denoising

**Require:** Noisy EEG $x_{\text{noisy}}$, Clean EEG $x_{\text{clean}}$
**Ensure:** Evaluation metrics

```
x_rec ← Model(x_noisy)
for each segment i do
    e^(i) ← x_clean^(i) − x_rec^(i)
    Compute MSE, SNR, PCC, SCC
    Compute RI based on PCC
    Compute SNR Gain
    Compute RRMSE (temporal and spectral)
    Compute SSIM and MI
end for
Aggregate metrics (mean and standard deviation)
return Evaluation results
```

Reconstruction accuracy was quantified using the mean squared error (MSE), which measures the average squared deviation between clean and reconstructed signals. Linear similarity was assessed via the Pearson correlation coefficient (PCC), from which the relative improvement index (RI) was computed with respect to the noisy input. Monotonic relationships were further evaluated using the Spearman rank correlation coefficient (SCC). Signal quality was characterized using the signal-to-noise ratio (SNR) in decibels (dB), comparing the power of the clean signal to the reconstruction error; the SNR gain was defined as the difference between the SNR of the reconstructed and noisy signals. Additionally, the relative root mean square error (RRMSE) was computed in both temporal and spectral domains. Structural similarity was measured using the Structural Similarity Index (SSIM), while Mutual Information (MI) quantified nonlinear dependencies between signals, where higher SSIM and MI values indicate better preservation of EEG content.

## VII. RESULTS

### A. *Results across Temporal Intervals*

Fig. 5(a) shows the first 100 samples of the EEG signal, in which the reconstructed outputs of each model are directly compared with the original signal.

### B. *MSE and RRMSE Analysis*

Fig. 5(b) illustrates the sample-wise mean squared error (MSE) after Gaussian noise suppression for a representative EEG signal from the test set. The proposed RHONN consistently yields the lowest reconstruction error compared with the other evaluated approaches. To further validate this observation, the average MSE was computed across all 903 EEG signals in the test set for each model. The results are summarized in Table V, with the best-performing model highlighted in bold.

TABLE V
MEAN MSE RESULTS UNDER GAUSSIAN NOISE

| Model | MSE ($\mu \pm \sigma$) |
|---|---|
| **RHONN** | $\mathbf{0.0664 \pm 0.0050}$ |
| GAN [13] | $0.1725 \pm 0.0202$ |
| Autoencoder [41] | $0.2171 \pm 0.0311$ |
| MLP [14] | $0.1060 \pm 0.0144$ |
| RNN [26] | $0.0693 \pm 0.0049$ |
| Transformer [30] | $0.0820 \pm 0.0051$ |

The RHONN model achieves an MSE of $0.0664 \pm 0.0050$, outperforming the RNN by approximately 4%. Compared to the Transformer , MLP, GAN and autoencoder , the reductions are approximately 19%, 37%, 61%, and 69%, respectively. These results highlight the superior reconstruction accuracy of the proposed approach, particularly when compared to generative-based models, as well as its improved stability. Similar performance trends can be observed in the temporal and spectral RRMSE metrics presented in Table VI. Since lower RRMSE values correspond to better reconstruction accuracy, the best-performing result for each metric is highlighted in bold. RHONN slightly outperforms the RNN in the temporal domain (approximately 2%) while achieving nearly equivalent performance in the spectral domain (approximately 1% difference). In contrast, the Transformer and MLP exhibit intermediate performance, whereas the GAN and autoencoder present the highest errors across both metrics, indicating lower robustness to Gaussian noise.

TABLE VI
TEMPORAL RRMSE AND SPECTRAL RRMSE OF MODELS UNDER GAUSSIAN NOISE

| Model | Temporal RRMSE | Spectral RRMSE |
|---|---|---|
| **RHONN** | $\mathbf{0.2578 \pm 0.0096}$ | $0.1057 \pm 0.0191$ |
| GAN [13] | $0.4150 \pm 0.0245$ | $0.2814 \pm 0.0591$ |
| Autoencoder [41] | $0.4651 \pm 0.0337$ | $0.1910 \pm 0.0515$ |
| MLP [14] | $0.3252 \pm 0.0215$ | $0.2571 \pm 0.0428$ |
| **RNN** [26] | $0.2633 \pm 0.0093$ | $\mathbf{0.1047 \pm 0.0212}$ |
| Transformer [30] | $0.2865 \pm 0.0090$ | $0.2565 \pm 0.0090$ |

### C. SNR and Gain Results

Regarding SNR performance, Table VII shows that RHONN and RNN are the only methods that achieve an improvement over the initial SNR, whereas all other models result in signal degradation. The best-performing model for each noise level is highlighted in bold. In particular, RHONN outperforms the RNN by approximately 21.5% in SNR gain, obtained by comparing their respective improvements (1.02 dB vs. 0.84 dB). In contrast, the Transformer exhibits a near-zero improvement, the MLP shows a moderate degradation, and the GAN and autoencoder present the largest losses in signal quality, significantly exceeding those of the other approaches. Overall,

TABLE VII
PERFORMANCE EVALUATION OF MODELS BASED ON THE IMPROVEMENT AND DEGRADATION OF THE SIGNAL-TO-NOISE RATIO (SNR) UNDER GAUSSIAN NOISE (DB)

| Model | Initial SNR | Denoised SNR | Gain |
|---|---|---|---|
| **RHONN** | $\mathbf{10.75 \pm 0.27}$ | $\mathbf{11.78 \pm 0.32}$ | $\mathbf{1.02 \pm 0.23}$ |
| GAN [13] | $10.75 \pm 0.27$ | $7.65 \pm 0.51$ | $-3.10 \pm 0.54$ |
| Autoencoder [41] | $10.75 \pm 0.27$ | $6.67 \pm 0.63$ | $-4.08 \pm 0.68$ |
| MLP [14] | $10.75 \pm 0.27$ | $9.77 \pm 0.55$ | $-0.98 \pm 0.52$ |
| RNN [26] | $10.75 \pm 0.27$ | $11.59 \pm 0.30$ | $0.84 \pm 0.20$ |
| Transformer [30] | $10.75 \pm 0.27$ | $10.86 \pm 0.27$ | $0.11 \pm 0.75$ |

RHONN achieves the best global performance, followed by the RNN, while generative-based models demonstrate the poorest behavior under Gaussian noise conditions.

### D. Performance under Various Noise Combinations

This section presents a comparative analysis of EEG signal denoising performance in the presence of physiological artifacts, including EMG, EOG, ECG, and their combinations. The models are evaluated under two contamination levels (30% and 70%), as summarized in Tables VIII and IX, following evaluation protocols commonly adopted in previous EEG artifact removal studies [53]–[56]. The tables report the RI%, SSIM, and MI metrics for each method. The best result for each metric is highlighted in bold.

As shown in Tables VIII and IX, the RHONN model maintains consistent performance across different contamination levels. Under moderate noise conditions ($\lambda = 30\%$), RHONN ranks among the top-performing models across all noise combinations, particularly leading in mutual information (MI) and achieving the best performance in scenarios with ocular and cardiac artifacts, with improvements in noise reduction (RI) on the order of 5–10% over the second-best model. In scenarios involving muscular or combined artifacts, although not always ranked first, it consistently remains within the top two, with differences generally below 5%, indicating stable and competitive performance. Under severe noise conditions ($\lambda = 70\%$), RHONN stands out primarily for its noise reduction capability, achieving the highest RI in most cases, with marginal advantages typically below 2–5% compared to the second-best model. However, in certain configurations, models such as the Autoencoder, MLP, or Transformer exhibit advantages of approximately 10–15% in metrics such as SSIM or MI, suggesting improved structural or information preservation. Overall, these results indicate that RHONN provides the most robust and well-balanced performance, particularly in terms of noise suppression, while other approaches may excel in preserving signal characteristics under specific conditions.

### E. Computational Cost and Energy Efficiency Analysis

To assess the practical feasibility of the proposed approach, a computational cost analysis was conducted using several metrics, including the number of trainable ($\theta$) and non-trainable ($\tilde{\theta}$) parameters, floating-point operations (FLOPs), inference latency ($\tau$), memory footprint (MF), model size (MS), throughput ($R$), and energy consumption ($E$). All evaluated

TABLE VIII
PERFORMANCE COMPARISON UNDER DIFFERENT PHYSIOLOGICAL NOISE COMBINATIONS ($\lambda = 30\%$)

| Model | EEG+EMG | | | EEG+EOG | | | EEG+ECG | | | EEG+EMG+EOG | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | RI | SSIM | MI | RI | SSIM | MI | RI | SSIM | MI | RI | SSIM | MI |
| RHONN | **25.26** | **0.77** | **2.03** | **6.70** | **0.93** | **2.20** | **5.09** | **0.82** | **3.19** | **23.82** | **0.76** | **1.84** |
| GAN [13] | 18.86 | 0.71 | 1.75 | -22.13 | 0.80 | 1.77 | -44.24 | 0.74 | 1.92 | 18.80 | 0.71 | 1.75 |
| Autoencoder [41] | **26.63** | **0.78** | **1.86** | **−5.75** | 0.76 | 1.75 | -4.96 | 0.73 | 1.78 | **26.53** | **0.78** | 1.81 |
| MLP [14] | -2.78 | 0.64 | 1.86 | -7.40 | 0.84 | **1.96** | **−0.56** | **0.90** | **2.43** | -2.81 | 0.64 | **2.03** |
| RNN [26] | 23.88 | 0.76 | 1.84 | -7.81 | **0.93** | 1.93 | -4.09 | 0.90 | 2.32 | 19.21 | 0.75 | 1.83 |
| Transformer [30] | 1.48 | 0.72 | 1.87 | 1.60 | 0.95 | 2.08 | 2.31 | 0.89 | 3.02 | 1.37 | 0.72 | 1.87 |

TABLE IX
PERFORMANCE COMPARISON UNDER DIFFERENT PHYSIOLOGICAL NOISE COMBINATIONS ($\lambda = 70\%$)

| Model | EEG+EMG | | | EEG+EOG | | | EEG+ECG | | | EEG+EMG+EOG | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | RI | SSIM | MI | RI | SSIM | MI | RI | SSIM | MI | RI | SSIM | MI |
| RHONN | **53.09** | **0.65** | **1.70** | **11.41** | 0.68 | 1.66 | **24.50** | 0.83 | **3.19** | **25.39** | 0.62 | 1.65 |
| GAN [13] | 25.44 | 0.62 | 1.65 | -14.73 | 0.73 | 1.69 | -35.32 | 0.63 | 1.89 | 18.57 | 0.63 | 1.68 |
| Autoencoder [41] | **52.41** | **0.72** | **1.93** | -1.53 | **0.88** | **1.87** | 1.49 | 0.64 | 1.69 | **26.32** | **0.72** | 1.72 |
| MLP [14] | 8.09 | 0.57 | 1.72 | **4.89** | 0.84 | **2.09** | **5.00** | **0.84** | **2.40** | 8.06 | 0.57 | **1.93** |
| RNN [26] | 18.61 | 0.63 | 1.68 | -3.68 | **0.88** | 1.85 | 0.81 | **0.84** | 2.33 | 23.04 | **0.65** | **1.70** |
| Transformer [30] | -20.93 | 0.49 | 1.59 | 10.94 | 0.91 | 1.93 | 24.47 | 0.83 | 3.02 | 16.46 | 0.61 | 1.81 |

models (GAN, Autoencoder, MLP, RNN, and Transformer) were tested under identical experimental conditions to ensure a fair comparison. The experiments were performed on a workstation equipped with an AMD Ryzen 7 5700 processor operating at 3.80 GHz, 16 GB of RAM, and an NVIDIA GeForce GTX 1650 graphics card with 4 GB of dedicated VRAM.

All experiments were conducted using EEG segments of 512 samples, consistent with the signal length considered throughout the previous sections, thereby ensuring a fair and realistic comparison among all evaluated models. Energy consumption was indirectly estimated from the computational resources utilized during inference, taking into account both execution time and system workload. As shown in Table X, RHONN achieves substantial reductions in computational complexity across all evaluated metrics. Specifically, it reduces the number of parameters by up to 99.99% compared with GAN, Autoencoder, and Transformer architectures, and by more than 90% relative to MLP and RNN models, while also decreasing FLOPs by over 95%. Bold values denote the best result obtained for each metric among all evaluated models. In terms of inference latency, RHONN achieves reductions of up to 98% compared with GAN and Autoencoder, and between 50% and 90% relative to the remaining models. These improvements translate into significantly higher throughput and enhanced processing capability. The compact nature of RHONN is further reflected in its memory footprint and model size. Compared with Autoencoder and Transformer architectures, RHONN achieves reductions exceeding 99%, while reductions greater than 90% are obtained relative to GAN, MLP, and RNN models. These results highlight the efficiency of the proposed architecture and its suitability for deployment on resource-constrained platforms. From an energy-efficiency perspective, RHONN exhibits the lowest estimated consumption among all evaluated models, achieving reductions of approximately 50%–70% compared with GAN and Autoencoder architectures while remaining competitive with other lightweight alternatives.

TABLE X
COMPARATIVE ANALYSIS OF COMPUTATIONAL AND ENERGY METRICS ACROSS MODELS

| Metric | RHONN | GAN | Autoenc. | MLP | RNN | Transf. |
|---|---|---|---|---|---|---|
| $\theta$ | **5** | 29,154 | 1,570,542 | 61 | 10 | 300,097 |
| $\tilde{\theta}$ | **0** | 0 | 1,422 | 0 | 0 | 0 |
| FLOPs | **422** | 10,285,056 | 3,137,459 | 143,360 | 18,944 | 2,372,858 |
| $\tau$ (s) | **0.0032** | 0.1828 | 0.3877 | 0.0199 | 0.0925 | 0.0460 |
| MF (KB) | **0.00395** | 3.5938 | 3.5938 | 0.00654 | 0.1912 | 0.0039 |
| MS (KB) | **40** | 71,680 | 19,010,944 | 2,048 | 1,024 | 3,945 |
| $R$ (samples/s) | **313** | 5 | 3 | 50 | 11 | 22 |
| $E$ (W) | **1.078** | 2.61 | 2.17 | 2.63 | 1.128 | 2.99 |

## VIII. CONCLUSIONS

This work presented a Recurrent High-Order Neural Network (RHONN) trained using the Extended Kalman Filter (EKF) for artifact removal in EEG signals within brain–computer interface (BCI) applications. The proposed approach combines the nonlinear approximation capability of RHONN with the adaptive estimation properties of the EKF, enabling effective suppression of physiological noise such as EOG, EMG, and ECG under both synthetic conditions and real-world noisy EEG signals. A key advantage of the method is its low computational complexity, as each EEG channel is modeled using a single high-order neuron, significantly reducing the number of trainable parameters compared to deep learning-based alternatives, and making the approach suitable for real-time and resource-constrained environments. Experimental results show that the proposed RHONN achieves competitive or superior performance compared to state-of-the-art methods, including MLP, RNN, GAN, autoencoder, and Transformer-based architectures, obtaining lower MSE, improved RRMSE, and higher SNR values, while maintaining robust performance in the presence of real noise and effectively preserving relevant neural information. This behavior is also reflected in additional metrics such as SSIM, relative improvement (RI), and mutual information. From a theoretical perspective, Lyapunov-based stability analysis guarantees the boundedness of the estimation error under bounded disturbances and asymptotic convergence in the noise-free case. Future work will focus on extending the proposed framework to multichannel coupled RHONN, developing adaptive mechanisms for EKF hyperparameter tuning, implementing the model on embedded hardware for real-time applications, and exploring hybrid or attention-based architectures to further enhance performance in complex real-world EEG scenarios.

## REFERENCES


[1] E. Niedermeyer and F. da Silva, *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, 2005.

[2] G. Cisotto and D. Chicco, "Ten quick tips for clinical electroencephalographic (EEG) data acquisition and signal processing," *PeerJ Computer Science*, vol. 10, p. e2256, sep 3 2024. [Online]. Available: http://dx.doi.org/10.7717/peerj-cs.2256

[3] Y. Li, W. Zeng, W. Dong, D. Han, L. Chen, H. Chen, Z. Kang, S. Gong, H. Yan, W. T. Siok, and N. Wang, "A Tale of Single-Channel Electroencephalography: Devices, Datasets, Signal Processing, Applications, and Future Directions," *IEEE Transactions on Instrumentation and Measurement*, vol. 74, pp. 1–20, 2025. [Online]. Available: http://dx.doi.org/10.1109/TIM.2025.3556900

[4] J. Sabio, N. Williams, G. McArthur, and N. Badcock, "A scoping review on the use of consumer-grade EEG devices for research," *PLOS ONE*, vol. 19, no. 3, p. e0291186, mar 6 2024. [Online]. Available: http://dx.doi.org/10.1371/journal.pone.0291186

[5] Z. Xue, Y. Zhang, H. Li, H. Chen, S. Shen, and H. Du, "Instrumentation, Measurement, and Signal Processing in Electroencephalography-Based Brain-Computer Interfaces: Situations and Prospects," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1–28, 2024. [Online]. Available: http://dx.doi.org/10.1109/TIM.2024.3417598

[6] W. Weng, Y. Gu, S. Guo, Y. Ma, Z. Yang, Y. Liu, and Y. Chen, "Self-supervised Learning for Electroencephalogram: A Systematic Survey," *ACM Computing Surveys*, vol. 57, no. 12, pp. 1–38, jul 12 2025. [Online]. Available: http://dx.doi.org/10.1145/3736574

[7] K. Erat, E. Şahin, F. Doğan, N. Merdanoğlu, A. Akcakaya, and P. Durdu, "Emotion recognition with EEG-based brain-computer interfaces: a systematic literature review," *Multimedia Tools and Applications*, vol. 83, no. 33, pp. 79 647–79 694, mar 1 2024. [Online]. Available: http://dx.doi.org/10.1007/s11042-024-18259-z

[8] G. Li, Y. Yuan, D. Ouyang, L. Zhang, B. Yuan, X. Chang, Z. Guo, and G. Guo, "Driver Distraction From the EEG Perspective: A Review," *IEEE Sensors Journal*, vol. 24, no. 3, pp. 2329–2349, feb 1 2024. [Online]. Available: http://dx.doi.org/10.1109/JSEN.2023.3339727

[9] F. Barreto, L. Carrere, and C. Tabernig, "Cognitask: Bci system based on p300 for cognitive therapies," *IEEE Latin America Transactions*, vol. 20, no. 6, pp. 884–890, Feb. 2022. [Online]. Available: https://latamt.ieeer9.org/index.php/transactions/article/view/6030

[10] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: a systematic review," *Journal of neural engineering*, vol. 16, no. 5, p. 051001, 2019. [Online]. Available: https://doi.org/10.1088/1741-2552/ab260c

[11] S. Shafiezadeh, G. Duma, G. Mento, A. Danieli, L. Antoniazzi, F. Del Popolo, P. Bonanni, and A. Testolin, "Methodological issues in evaluating machine learning models for eeg seizure prediction: Good cross-validation accuracy does not guarantee generalization to new patients," *Applied Sciences*, vol. 13, no. 7, p. 4262, 2023. [Online]. Available: https://doi.org/10.3390/app13074262

[12] W. Xiong, L. Ma, and H. Li, "A general dual-pathway network for eeg denoising," *Frontiers in Neuroscience*, vol. 17, p. 1258024, 2024. [Online]. Available: https://doi.org/10.3389/fnins.2023.1258024

[13] E. Brophy, P. Redmond, A. Fleury, M. De Vos, G. Boylan, and T. Ward, "Denoising eeg signals for real-world bci applications using gans," *Frontiers in Neuroergonomics*, vol. 2, p. 805573, 2022. [Online]. Available: https://doi.org/10.3389/fnrgo.2021.805573

[14] S. Gaamouri, M. Bousbia Salah, and H. Rachid, "Denoising ecg signals by using extended kalman filter to train multi-layer perceptron neural network," *Automatic Control and Computer Sciences*, vol. 52, 11 2018. [Online]. Available: https://doi.org/10.3103/S0146411618060044

[15] W. Lu, "A review of eeg noise reduction technology based on dccrn-lstm," *Theoretical and Natural Science*, vol. 58, pp. 182–187, 2024. [Online]. Available: https://doi.org/10.54254/2753-8818/58/20241381

[16] A. Chaddad, Y. Wu, R. Kateb, and A. Bouridane, "Electroencephalography signal processing: A comprehensive review and analysis of methods and techniques," *Sensors*, vol. 23, no. 14, p. 6434, 2023. [Online]. Available: https://doi.org/10.3390/s23146434

[17] W. Yan and Y. Wu, "A time-frequency denoising method for single-channel event-related eeg," *Frontiers in Neuroscience*, vol. 16, p. 991136, 2022. [Online]. Available: https://doi.org/10.1016/j.bspc.2020.102337

[18] P. Gajbhiye, R. Tripathy, A. Bhattacharyya, and R. Pachori, "Novel approaches for the removal of motion artifact from eeg recordings," *IEEE Sensors Journal*, vol. 19, no. 22, pp. 10 600–10 608, 2019. [Online]. Available: https://doi.org/10.1109/JSEN.2019.2931727

[19] M. Saber, "Removing powerline interference from eeg signal using optimized fir filters," *J. Artif. Intell. Metaheuristics*, vol. 1, no. 1, pp. 8–19, 2022. [Online]. Available: :https://doi.org/10.54216/JAIM.010101

[20] C. Kaur, A. Bisht, P. Singh, and G. Joshi, "Eeg signal denoising using hybrid approach of variational mode decomposition and wavelets for depression," *Biomedical Signal Processing and Control*, vol. 65, p. 102337, 2021. [Online]. Available: https://doi.org/10.1016/j.bspc.2020.102337

[21] M. Felja, A. Bencheqroune, M. Karim, and G. Bennis, "Removing artifacts from eeg signal using wavelet transform and conventional filters," *WSEAS Transactions on Information Science and Applications*, vol. 17, pp. 177–183, 2020. [Online]. Available: https://doi.org/10.37394/23209.2020.17.22

[22] A. Abutaleb, H. Abdelaleem, and K. Hewedy, "Stochastic models for the eeg frequencies," *International Journal of Signal Processing*, vol. 6, 2021. [Online]. Available: https://www.iaras.org/home/caijsp/stochastic-models-for-the-eeg-frequencies

[23] F. Alturki, K. AlSharabi, A. Abdurraqeeb, and M. Aljalal, "Eeg signal analysis for diagnosing neurological disorders using discrete wavelet transform and intelligent techniques," *Sensors*, vol. 20, no. 9, p. 2505, 2020. [Online]. Available: https://doi.org/10.3390/s20092505

[24] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural networks*, vol. 13, no. 4-5, pp. 411–430, 2000. [Online]. Available: https://doi.org/10.1016/S0893-6080(00)00026-5

[25] Y. An, H. Lam, and S. Ling, "Auto-denoising for eeg signals using generative adversarial network," *Sensors*, vol. 22, no. 5, p. 1750, 2022. [Online]. Available: https://doi.org/10.3390/s22051750

[26] R. Williams, "A learning algorithm for continually running fully recurrent neural netwokrs," *Neural Computation*, vol. 1, pp. 256–263, 1989.

[27] A. Lins, E. Torres, N. Nobre, A. Brito, and E. Ribeiro, "Towards automatic eeg signal denoising by quality metric optimization," in *2020 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2020, pp. 1–7. [Online]. Available: https://doi.org/10.1109/IJCNN48605.2020.9207504

[28] X. Zhang, "Deep Learning-Based Techniques for Electroencephalogram (EEG) Signal Denoising," *Transactions on Computer Science and Intelligent Systems Research*, vol. 5, pp. 922–927, aug 12 2024. [Online]. Available: http://dx.doi.org/10.62051/rrve8560

[29] M. Gabardi, A. Saibene, F. Gasparini, D. Rizzo, and F. Stella, "A multi-artifact EEG denoising by frequency-based deep learning," *AIxHMI@AI*IA*, 2023. [Online]. Available: https://arxiv.org/abs/2310.17335

[30] C. Chuang, K. Chang, C. Huang, and A. Bessas, "Art: Artifact Removal Transformer for Reconstructing Noise-Free Multichannel Electroencephalographic Signals," *arXiv.org*, 2024. [Online]. Available: https://arxiv.org/abs/2409.07326

[31] M. Versaci and F. La Foresta, "Eeg Data Analysis Techniques for Precision Removal and Enhanced Alzheimer's Diagnosis: Focusing on Fuzzy and Intuitionistic Fuzzy Logic Techniques," *Signals*, vol. 5, no. 2, pp. 343–381, may 31 2024. [Online]. Available: http://dx.doi.org/10.3390/signals5020018

[32] A. Maitin, J. Romero, and Á. García, "Survey of Machine Learning Techniques in the Analysis of EEG Signals for Parkinson's Disease: A Systematic Review," *Applied Sciences*, vol. 12, no. 14, p. 6967, jul 9 2022. [Online]. Available: http://dx.doi.org/10.3390/app12146967

[33] A. Alanis, "Exploring kalman filtering applications for enhancing artificial neural network learning," *Algorithms*, vol. 18, no. 9, 2025. [Online]. Available: https://www.mdpi.com/1999-4893/18/9/587

[34] E. Sanchez, A. Alanís, and A. Loukianov, *Discrete-Time Block Control*. Springer, 2008.

[35] I. Urbina, J. Medrano, A. Rodríguez, F. López-Estrada, L. Djilali, and A. Badillo-Olvera, "Adaptive real-time neural control of permanent magnet dc motors under uncertainty," in *22nd International Conference*

*on Electrical Engineering, Computing Science and Automatic Control (CCE)*, 2025, pp. 1–6. [Online]. Available: https://doi.org/10.1109/CCE67728.2025.11271987

[36] I. Urbina, J. Medrano-Hermosillo, A. Rodriguez-Mata, F. Lopez-Estrada, O. Suarez, and A. Luna-Gómez, “Super-twisting-based online learning in high-order neural networks for robust backstepping control of dc motors under uncertainty,” *Processes*, vol. 14, p. 1019, 03 2026. [Online]. Available: https://doi.org/10.3390/pr14061019

[37] L. Ljung, *System Identification: Theory for the User*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 1999.

[38] B. Anderson and J. Moore, *Optimal Filtering*. Mineola, NY: Dover Publications, 2012.

[39] H. Khalil, *Nonlinear Systems*, 3rd ed. Upper Saddle River, NJ: Prentice Hall, 2002.

[40] S. Sastry, *Nonlinear Systems: Analysis, Stability, and Control*. New York, NY: Springer, 1999.

[41] H. Zhang, M. Zhao, C. Wei, D. Mantini, Z. Li, and Q. Liu, “Eegdenoisenet: a benchmark dataset for deep learning solutions of eeg denoising,” *Journal of Neural Engineering*, vol. 18, no. 5, p. 056057, 2021. [Online]. Available: https://doi.org/10.1088/1741-2552/ac2bf8

[42] G. Moody and R. Mark, “Mit-bih arrhythmia database,” *(No Title)*, 1992. [Online]. Available: https://doi.org/10.13026/C2F305

[43] A. Goldberger, L. Amaral, L. Glass, J. Hausdorff, P. Ivanov, R. Mark, J. Mietus, G. Moody, C. Peng, and H. Stanley, “Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals,” *circulation*, vol. 101, no. 23, pp. e215–e220, 2000. [Online]. Available: https://doi.org/10.1161/01.cir.101.23.e215

[44] M. Hossain, S. Mahmud, A. Khandakar, N. Al-Emadi, F. Chowdhury, Z. Mahbub, M. Reaz, and M. Chowdhury, “Multiresunet3+: A full-scale connected multi-residual unet model to denoise electrooculogram and electromyogram artifacts from corrupted electroencephalogram signals,” *Bioengineering*, vol. 10, no. 5, p. 579, 2023. [Online]. Available: https://doi.org/10.3390/bioengineering10050579

[45] Y. Cai, Z. Meng, and D. Huang, “Dhct-gan: Improving eeg signal quality with a dual-branch hybrid cnn–transformer network,” *Sensors*, vol. 25, no. 1, p. 231, 2025. [Online]. Available: https://doi.org/10.3390/s25010231

[46] A. Hyvärinen and E. Oja, “Independent component analysis: algorithms and applications,” *Neural Networks*, vol. 13, no. 4, pp. 411–430, 2000.

[47] L. Pion-Tonachini, K. Kreutz-Delgado, and S. Makeig, “Iclabel: An automated electroencephalographic independent component classifier, dataset, and website,” *NeuroImage*, vol. 198, pp. 181–197, 2019. [Online]. Available: https://doi.org/10.1016/j.neuroimage.2019.05.026

[48] Y. Gordienko, N. Gordienko, V. Taran, A. Rojbi, S. Telenyk, and S. Stirenko, “Effect of natural and synthetic noise data augmentation on physical action classification by brain–computer interface and deep learning,” *Frontiers in Neuroinformatics*, vol. 19, p. 1521805, 2025. [Online]. Available: https://doi.org/10.3389/fninf.2025.1521805

[49] I. Dryden and K. Mardia, *Statistical shape analysis: with applications in R*. John Wiley & Sons, 2016.

[50] S. Stalin, V. Roy, P. Shukla, A. Zaguia, M. Khan, P. Shukla, and A. Jain, “A machine learning-based big eeg data artifact detection and wavelet-based removal: An empirical approach,” *Mathematical Problems in Engineering*, vol. 2021, no. 1, p. 2942808, 2021. [Online]. Available: https://doi.org/10.1155/2021/2942808

[51] W. Sun, Y. Su, X. Wu, and X. Wu, “A novel end-to-end 1d-rescnn model to remove artifact from eeg signals,” *Neurocomputing*, vol. 404, pp. 108–121, 2020. [Online]. Available: https://doi.org/10.1016/j.neucom.2020.04.029

[52] B. Wang, F. Deng, and P. Jiang, “Eegdir: Electroencephalogram denoising network for temporal information storage and global modeling through Retentive Network,” *Computers in Biology and Medicine*, vol. 177, p. 108626, 7 2024. [Online]. Available: http://dx.doi.org/10.1016/j.compbiomed.2024.108626

[53] R. Jiang, S. Tong, J. Wu, H. Hu, R. Zhang, H. Wang, Y. Zhao, W. Zhu, S. Li, and X. Zhang, “A novel eeg artifact removal algorithm based on an advanced attention mechanism,” *Scientific Reports*, vol. 15, no. 1, p. 19419, 2025. [Online]. Available: https://doi.org/10.1038/s41598-025-98653-1

[54] B. Kalita, N. Deb, and D. Das, “Aneeg: leveraging deep learning for effective artifact removal in eeg data,” *Scientific Reports*, vol. 14, no. 1, p. 24234, 2024. [Online]. Available: https://doi.org/10.1038/s41598-024-75091-z

[55] Y. Meng, C. Yuan, W. Feng, D. Li, J. Nan, Y. Xia, F. Zhu, and J. Song, “Tf-denoiser: A time-frequency domain joint method for eeg artifact removal,” *Electronics*, vol. 15, no. 1, p. 132, 2025. [Online]. Available: https://doi.org/10.3390/electronics15010132

[56] H. Li, F. Feng, J. Kang, J. Zhang, X. Gong, T. Lu, S. Li, Z. Sun, and J. Solé-Casals, “A 2 dm: Enhancing eeg artifact removal by fusing artifact representation into the time-frequency domain,” *Cognitive Computation*, vol. 17, no. 2, p. 84, 2025. [Online]. Available: https://doi.org/10.1007/s12559-025-10442-0


**Enrique Martinez-Madrid** received his B.Sc. degree in Electronic Engineering from the Tecnológico Nacional de México, Mexico, in 2022, and the M.Sc. degree in Electronic Engineering from the División de Estudios de Posgrado e Investigación of Tecnológico Nacional de México, Chihuahua, Mexico, in 2026. He is currently pursuing the Ph.D. degree in Electronic Engineering at the Tecnológico Nacional de México, Chihuahua, Mexico. He is also a Professor teaching engineering and technology-related courses at the Universidad Tecnológica de Chihuahua, Chihuahua, Mexico. His current research interests include electrical signal processing, artificial intelligence, data science, embedded systems, power electronics, control systems, and adaptive methods for intelligent systems.

**Jesús A. Medrano-Hermosillo** received his Ph.D. degree in Electrical Engineering from Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional. He is currently a Professor-Researcher at Tecnológico Nacional de México, where he contributes to the electrical and electronic signal processing research line within the Master of Science in Electronic Engineering and the Doctorate in Science in Electronic Engineering programs. His research interests include control of electromechanical systems, intelligent systems optimization, high-order neural networks, and advanced control methodologies.

**Juan A. Ramírez-Quintana** received his B.Sc., M.Sc., and Ph.D degrees in Electronic Engineering in 2004, 2007, and 2014, respectively. From 2008 to 2011, he was a researcher and a teaching assistant at different universities. He is currently a Researcher and Professor at the Tecnologico Nacional de Mexico campus I.T. Chihuahua and the Director of the PVR Lab. His current research includes computer vision, signal processing, computational intelligence, and machine learning. He is the author of patents and more than 70 papers published in journals, proceedings, and book chapters. He is a member of the National Research System of Mexico (SNII) and the Academia Mexicana de Computación (AMEXCOMP).

**Iván R. Urbina-Leos** received his B.Sc. degree in Electronic Engineering from the Tecnológico Nacional de México, Mexico, in 2024, where he is currently pursuing the M.Sc. degree in Electronic Engineering Sciences. Currently, he works professionally at Micron Semiconductor. His current academic and professional interests are span automatic control systems, power electronics, digital systems, and embedded systems. Furthermore, he is particularly interested in implementing recurrent high-order neural networks (RHONN) and artificial intelligence for electric motor control.

**Oscar J. Suarez** received his B.S.E.E. degree from Universidad Pontificia Bolivariana, Bucaramanga, Colombia, in 2011, and the M.S.E.E. degree in industrial control from the Universidad de Pamplona, Pamplona, in 2015. He received the D.Sc. degree in Electrical Engineering from the Center for Research and Advanced Studies of the National Polytechnic Institute (CINVESTAV-IPN), Guadalajara, Mexico in 2019. Since 2024, he has been a Professor with the Mechatronics Engineering Department, Universidad de Pamplona, Pamplona, Colombia. His current research interests include nonlinear control, neural networks, artificial intelligence, and complex networks.

**Francisco-Ronay López-Estrada** received his Ph.D. degree in Automatic Control from the University of Lorraine, France, in 2014. He has been with Tecnológico Nacional de México, IT Tuxtla Gutiérrez, since 2008. He is on the Editorial Board of the Mathematical and Computational Applications and Int. J. of Applied Mathematics and Computer Science journals. He is the author/co-author of more than 70 JCR-indexed research journal papers and several international conferences. His research interests are fault diagnosis and fault-tolerant control based on convex LPV and Takagi-Sugeno models and their applications.