

DYNAMIC THRESHOLDING MECHANISM FOR CYCLOSTATIONARY SPECTRUM SENSING USING ARTIFICIAL NEURAL NETWORK UNDER TIME-VARIANT ENVIRONMENTAL CONDITIONS

^A M. SUBA, ^B DR. D. SUSAN

^A ASSISTANT PROFESSOR, DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING, SRINIVASA RAMANUJAN CENTRE, KUMBAKONAM, 612001, INDIA. EMAIL: suba@src.sastra.edu

^B ASSOCIATE PROFESSOR, SCHOOL OF ELECTRICAL AND ELECTRONICS ENGINEERING, SASTRA DEEMED TO BE UNIVERSITY, THANJAVUR, 613401, INDIA. EMAIL: d_susan@ece.sastra.edu

Abstract

Reliable spectrum sensing is critical for cognitive radio networks to ensure efficient utilization of wireless resources and to protect primary users, particularly in dynamic maritime environments where channel conditions vary due to mobility, fading, and interference. Existing spectrum sensing techniques, including TFCFN, STFT-RADN, CNN-LSTM, and DBN-FOA, often suffer from limited adaptability to non-stationary channels, reduced detection accuracy, and suboptimal threshold selection, which can result in high false alarms or missed detections. To overcome these limitations, this work introduces a dynamic thresholding mechanism for cyclostationary spectrum sensing using artificial neural networks under time-variant environmental conditions. The proposed framework integrates Temporal Spectrum Variation Modeling (TSVM) via TSVANet, which combines ANN-based local spectral feature extraction with a Multi-Head Temporal Self-Attention (MHTSA) module to capture both local and long-term temporal dependencies, and Adaptive Learning Rate Optimization (ALRO) via Social Spider–Krill Hybrid (SKH), which dynamically tunes learning parameters for robust convergence under non-stationary data streams. Experimental evaluations demonstrate that the proposed method achieves 99.39% accuracy, 98.32% precision, and 99.09% sensitivity, outperforming existing approaches. The framework enhances spectrum sensing reliability, adaptability, and robustness, enabling efficient cognitive radio operation in complex maritime scenarios.

Keywords: Cyclostationary Spectrum Sensing, Temporal Spectrum Variation Modeling (TSVM), TSVANet, ALRO, SKH

1. INTRODUCTION

Spectrum sensing is a fundamental operation in cognitive radio networks, enabling the identification of available frequency bands without interfering with licensed or primary users [1][2]. With the rapid growth of wireless communication systems and the increasing demand for bandwidth, efficient spectrum utilization has become critical [3]. Cognitive radio networks rely on accurate detection of spectral occupancy to dynamically access underutilized channels, improving communication efficiency, reducing congestion, and supporting diverse societal applications, including maritime, urban, and industrial wireless communications [4][5]. Effective spectrum sensing ensures reliable communication, minimizes service disruption, and safeguards critical transmissions in dynamic and time-varying environments [6][7].

Conventional spectrum sensing techniques, such as energy detection, matched filtering, and cyclostationary feature detection, have been widely explored [8][9]. More recent approaches, including TFCFN, STFT-RADN, CNN-LSTM, and DBN-FOA, have attempted to enhance detection performance under challenging conditions. Despite these advancements, existing methods exhibit significant limitations [10]. Many rely on fixed thresholds or single-feature representations, making them highly sensitive to noise, multipath fading, and varying channel conditions. Several approaches struggle to capture long-term temporal dependencies or adapt to dynamic environments, resulting in reduced detection accuracy, higher false alarm rates, and unreliable primary user identification [11][12].

These challenges underscore the importance of developing more robust and adaptive spectrum sensing mechanisms that can maintain high reliability even under non-stationary channel conditions [13][14]. Addressing these issues is essential for efficient spectrum management, enhanced communication quality, and improved utilization of scarce wireless resources [15][16][17].

The current study focuses on designing and evaluating a framework for cyclostationary spectrum sensing that accounts for dynamic environmental conditions and time-varying channels. It emphasizes accurate feature extraction, preprocessing, and adaptive processing to ensure reliable detection of spectrum occupancy in practical communication scenarios.

The major contribution of the work includes:

- To design a robust dynamic thresholding mechanism for cyclostationary spectrum sensing capable of operating under time-variant and non-stationary environmental conditions.
- To develop a Temporal Spectrum Variation Modeling (TSVM) framework via TSVANet, combining ANN-based local spectral feature extraction with Multi-Head Temporal Self-Attention (MHTSA) to capture both local and long-term temporal dependencies in spectral data.
- To implement an Adaptive Learning Rate Optimization (ALRO) strategy using a Social Spider–Krill Hybrid (SKH), ensuring stable convergence and robust learning under varying channel conditions.

The organization of this work: The literature review using the current methodology is found in Section 2. The system model, comprising the framework and key elements, is presented in Section 3. The result and comparison analysis are shown in Section 4. Finally, the work comes to a close in the section 5 with the conclusion.

2. LITERATURE REVIEW

In 2024, Xi et al. [18] proposed a Time-Frequency Cross Fusion Network (TFCFN) that integrated GRU-based temporal modeling, FFT-based frequency extraction, and CNN-based local feature learning with cross-attention fusion. Evaluations under Gaussian and non-Gaussian noise using GGD showed improved detection ability, robustness, and reduced complexity compared with baseline methods. In 2024, Wang et al. [19] developed a spectrum sensing method based on Short-Time Fourier Transform (STFT) and Residual Attention Dense Network (RADN). By combining residual and dense connections with attention modules, the model enhanced feature extraction and classification of time–frequency spectrograms, achieving superior detection performance under low SNR conditions across multiple modulation schemes. In 2025, Wang et al. [20] introduced a hybrid CNN-LSTM model with multi-head self-attention for cooperative spectrum sensing in multi-user cognitive radio systems. CNN extracted spatial features, LSTM captured temporal dependencies, and attention improved adaptability. Results demonstrated reduced sensing errors and higher accuracy across different user configurations compared with conventional deep learning models. In 2025, Hameed et al. [21] presented a hybrid spectrum sensing technique combining Kernel Least Mean Square (KLMS) filtering and PyramidNet, optimized with Harmonic Elk Herd Optimization (HEHO). Using cyclic spectrum features and average fusion, the method achieved high detection probability, energy efficiency, and throughput, while minimizing false alarms and detection time in CR networks. In 2021, Reddy and Siva [22] proposed a Deep Belief Network (DBN) optimized by the Fruit Fly Optimization Algorithm (FOA) to overcome ANN limitations in cooperative spectrum sensing. FOA tuned DBN parameters such as learning rate, weight decay, penalty factor, and hidden units. The framework demonstrated improved accuracy, lower false alarms, and reduced detection loss.

3. SYSTEM MODEL

The proposed system operates within a Maritime Cognitive Radio Network, where spectrum sensing is performed under time-variant channel conditions. Initially, signals are captured and processed using cyclostationary feature analysis to generate representative datasets. The preprocessing stage extracts relevant features through ACS, transforms them into time-frequency maps via STFT, and enhances signal quality using Z-score normalization and wavelet denoising.

Subsequently, Temporal Spectrum Variation Modeling (TSVM) via TSVANet captures both local spectral patterns and long-term temporal dependencies using ANN-based feature extraction combined with a Multi-Head Temporal Self-Attention (MHTSA) module. Environment-Aware Neural Thresholding (EANT) then refines features through CNNs, models temporal dependencies via GRUs, and generates dynamic thresholds for adaptive decision making. To ensure robust learning under non-stationary conditions, Adaptive Learning Rate Optimization (ALRO) via a Social Spider–Krill Hybrid (SKH) dynamically adjusts network parameters. Finally, the system performs spectrum sensing by comparing processed features against adaptive thresholds to determine spectrum occupancy, enabling reliable detection of primary user activity in varying maritime environments. The general architecture of the system model is represented in Figure 1.

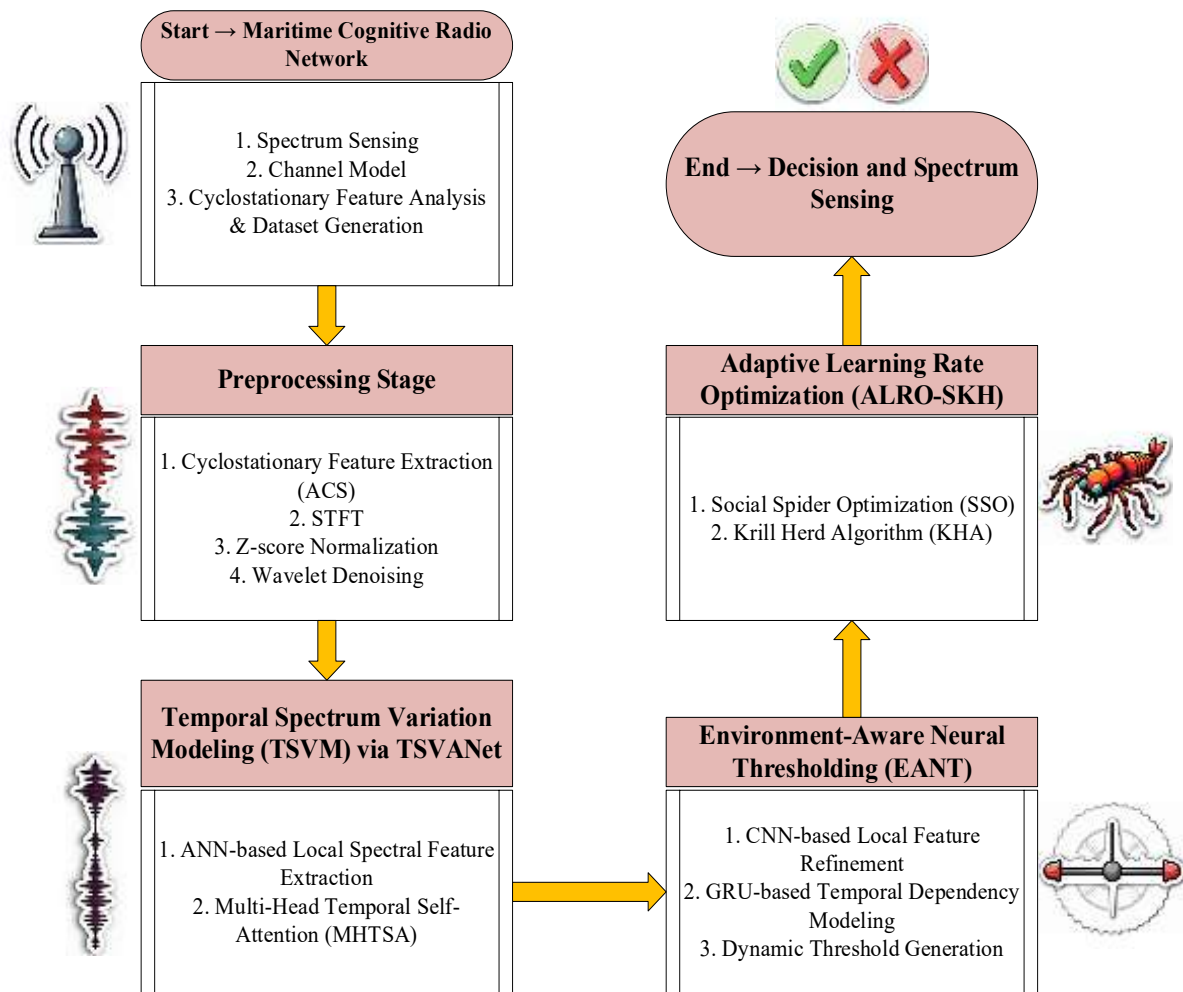


Figure 1: Architecture of the System Model

3.1 Maritime Cognitive Radio Network

The marine two-way radio system that can send and receive messages is used in this model to depict a PU. The fusion center serves as the network's primary node for signal processing and decision-making, and it might be situated on an island or the shore. Unlicensed ships are examples of movable maritime assets that are commonly used as cognitive nodes, or SUs. Every vessel is outfitted with cognitive radio operations devices that are designed to scan the radio environment on a regular basis and discover spectrum resources that PUs isn't using. The fusion center makes a worldwide judgement on spectrum availability after combining the cognitive node's sensing findings. The SUs is then informed of these decision outcomes. Additionally, satellite links offer an alternate route for communication to the fusion center in cases where boats are located far from shore and cannot reach it directly.

3.1.1 Spectrum Sensing

Cognitive radio nodes with sensing capabilities actively scan the radio frequency spectrum using antennas. Through statistical analysis of the signals received, these nodes are able to determine the occupancy status of the frequency bands assigned to PUs. The following is how Eq. (1) represents the SS process mathematically as a binary detection problem:

$$a(t) = \begin{cases} m(t)H_0 \\ S(t) + m(t), H_1 \end{cases} \quad (1)$$

In this model, $a(t)$ represents the received signal. Within the signal that has been received, the signal component is specified as $S(t) = h \times x(t)$, where $x(t)$ is the transmitted signal and h is the channel gain. $m(t)$ represents the white Gaussian additive noise (AWGN). The PU is either present or absent, according to the theories H_0 and H_1 , respectively. The detection strategy is to compare the detection statistic ξ with a threshold ζ . The probabilities of detection (P_A) and false alarm (P_{FB}) are defined as follows in Eq. (2) and Eq. (3):

$$P_A = P_r[\xi \geq \zeta | H_1] = P_r[H_1 | H_2] \quad (2)$$

$$P_{FB} = P_r[\xi \geq \zeta | GH_0] = P_r[H_1 | H_0] \quad (3)$$

3.1.2 Channel Model

At a distance T from the transmitting antenna, the received signal power P_r in free space equals are shown in Eq. (4)

$$P_r = \left(\frac{\lambda}{4\pi T} \right)^2 G_r P_t G_t \quad (4)$$

Where, P_t is the transmitting power. The G_t and G_r are the relative gains of the receiving and transmitting antennas. For the wavelength, the λ .

The multi-path effects brought on by sea surface reflection have an impact on the received signal strength in marine contexts. Direct, specular, and diffuse reflections are all included in the reflection. Diffuse reflection signals are typically regarded as random signals because of their limited correlation with transmitting signals. Consequently, direct and specular reflection signals make up the majority of the received signal, and their power written as in Eq. (5)

$$P_r = \left(\frac{\lambda}{4\pi r}\right)^2 P_t G_r G_t \left|1 + \sum_{i=1}^N \rho_i e^{j\phi_i}\right|^2 \quad (5)$$

where the i th reflection path's specular reflection coefficient is denoted by ρ_i . The phase difference between the direct and reflected paths is represented by the symbol ϕ_i . The number of reflection spots that are effective is denoted by N . The multi-path channel gain h (dB) of any path from transmitter to receiver, ignoring the transmitting and receiving antenna strengths, can be written as in Eq. (6)

$$h = 10\log_{10} \frac{P_t}{P_r} = 101g\left(\frac{\frac{4\pi R_f^2}{3 \times 10^8}}{|1 + \sum_{i=1}^N \rho_i e^{j\phi_i}|}\right) \quad (6)$$

Here, f stands for radio frequency.

3.1.3 Cyclostationary Feature Analysis and Dataset Generation

The PU of a maritime cognitive wireless communication system is usually a marine voice intercom system that uses frequency modulation (FM) as its primary method. Because of this, the signals show second-order cyclostationarity, which makes cyclostationary analysis a useful detection method. With second-order cyclostationarity, the transmitting signal $s(t)$ is the following Fourier series expansion of its periodic Autocorrelation Function (ACF) as Eq. (7):

$$R_s(t, \tau) = \sum_{\alpha=-\infty}^{\infty} R_s^\alpha(\tau) e^{2\pi j \alpha t} \quad (7)$$

in which the autocorrelation interval $R_s^\alpha(\tau)$ is defined as Eq. (8):

$$R_s^\alpha(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T s\left(t + \frac{\tau}{2}\right) s^*\left(t - \frac{\tau}{2}\right) e^{-2\pi j \alpha t} dt \quad (8)$$

where τ is the time delay related to ACF. The T and α denotes period and cycle frequency, respectively. Using the Wiener relation, the CPS can be defined as Eq. (9):

$$S_s^\alpha(f) = \int_{-\infty}^{\infty} R_s^\alpha(\tau) e^{-j2\pi f \tau} d\tau \quad (9)$$

According to the formula given in Eq. (1), the received signal CPS is represented by Eq. (10):

$$S_r^\alpha(f) = \begin{cases} S_n^\alpha(f) & H_0 \\ S_s^\alpha(f) + S_n^\alpha(f)H_1 \end{cases} \quad (10)$$

Where, $S_r^\alpha(f)$ and $S_n^\alpha(f)$ represent the CPS of the received signal and AWGN, respectively. The $S_s^\alpha(f)$ indicates the CPS of the PU signal component. As $n(t)$ is not a cyclisation process, its CPS equals zero at $\alpha = 0$. For the received signal $r(t)$, its periodic ACF is computed by Eq. (11)

$$R_r^\alpha(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T u\left(t + \frac{\tau}{2}\right) u^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi \alpha t} dt \quad (11)$$

Where, $u(t) = r(t)e^{-j\pi \alpha t}$. The results of cross-spectral analysis are defined as Eq. (12)

$$S_r^\alpha(f) = \lim_{\Delta t \rightarrow \infty} \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{\Delta t}{2}}^{\frac{\Delta t}{2}} R_T\left(t, f + \frac{\alpha}{2}\right) R_T^*\left(t, f - \frac{\alpha}{2}\right) dt \quad (12)$$

Where, $R_T(t, f) = \int_{t-\frac{T_0}{2}}^{t+\frac{T_0}{2}} r(u) e^{-j2\pi f u} du$. However, $S_r^\alpha(f)$ cannot be used directly as an estimate of the cyclic spectral density due to its large variance. A smoothing approach is required to provide a more precise approximation. The FFT Accumulation Method (FAM) is employed in this work to smooth the CPS, so attaining a satisfactory balance between cycle leakage, computational efficiency, and cycle aliasing. The FAM algorithm performs better than other methods (including the periodogram technique, the indirect method, and the SSCA algorithm) in terms of computing efficiency, real-time capabilities, and resilience when the signal-to-noise ratio is low. Derive the CPS estimate by discretizing the incoming signal $r(t)$ to be $r(n)$ as Eq. (13).

$$S_r^{\alpha 0 + q\Delta\alpha}(nL, f) = \frac{1}{P} \sum_{r=0}^{P-1} \frac{1}{N_p} R_T\left(rL, f - \frac{\alpha}{2}\right) e^{-j2\pi r \frac{q}{P}} \quad (13)$$

where the P signifies the number of blocks that $r(n)$ is divided into, and NP denotes the total number of points in each block. The value of L is $\frac{N_p}{4}$, and $q = -\frac{P}{2}, -\frac{P}{2} + 1, \dots, \frac{P}{2} - 1$, and $R_T(n, f) = \sum_{k=-\frac{N_p}{2}}^{\frac{N_p}{2}-1} w(k) r(n-k) e^{-j2\pi f(n-k)T_s}$. Where, T_s denotes the sampling period, and $w(k)$ is a kaiser window with a width of $T_\omega = N_p T_s$. The N_p and p are calculated by Eq. (14) and (15).

$$N_p = 2 \left\lceil \log_2 \left(\frac{f_s}{dv} - 1 \right) + 1 \right\rceil \quad (14)$$

$$P = 2 \left\lceil \log_2 \left(\frac{f_s}{Ld\alpha} - 1 \right) + 1 \right\rceil \quad (15)$$

Where, f_s is the sampling frequency, and dv is the desired frequency resolution, and $d\alpha = \frac{1}{dt}$ is the desired cycle frequency resolution.

3.2 Preprocessing Stage

The preprocessing stage plays a crucial role in preparing the raw received signal for reliable spectrum sensing under time-varying environmental conditions. Wireless signals are often corrupted by noise, multipath fading, and mobility-induced distortions, which can severely degrade detection accuracy if left unaddressed. To mitigate these challenges, preprocessing transforms the raw RF input into a more structured and noise-resilient representation. This is achieved by extracting cyclostationary features, converting the signal into a time–frequency domain, and applying normalization and denoising techniques. The resulting feature set provides a stable and discriminative input representation for subsequent analysis.

3.2.1 Cyclostationary Feature Extraction (ACS)

Cyclostationary analysis exploits the periodic statistical properties inherent in modulated communication signals, which remain distinct even in the presence of additive white Gaussian noise (AWGN). Unlike noise, which exhibits stationary characteristics, communication signals often demonstrate periodicity in their autocorrelation function due to carrier frequency, symbol rate, or cyclic prefixes in modulation schemes. This property allows cyclostationary feature extraction to serve as a robust preprocessing step for spectrum sensing.

Let $r(t)$ denote the received signal in the time domain. The autocorrelation function of $r(t)$ is expressed as Eq. (16):

$$R_r(t, \tau) = \mathbb{E}[r(t)r^*(t + \tau)] \quad (16)$$

Where, τ represents the time lag and $(\cdot)^*$ denotes the complex conjugate. For a cyclostationary process, $R_r(t, \tau)$ is periodic in t with period T_0 . This periodicity enables the representation of $R_r(t, \tau)$ in terms of Fourier series coefficients, known as cyclic autocorrelation functions (CAF). The CAF at a cyclic frequency α is given by Eq. (17).

$$R_r^\alpha(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} R_r(t, \tau) e^{-j2\pi\alpha t} dt \quad (17)$$

The CAF highlights spectral correlation at cyclic frequencies α , which are absent in stationary noise, thereby enabling discrimination between noise and legitimate communication signals. The spectral correlation function (SCF), obtained as the Fourier transform of the CAF with respect to the lag variable τ , is expressed as Eq. (18).

$$S_r^\alpha(f) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\infty}^{\infty} R_r^\alpha(\tau) e^{-j2\pi f \tau} d\tau \quad (18)$$

Where, f denotes the spectral frequency. Peaks in $S_r^\alpha(f)$ correspond to cyclic features induced by modulation, providing a robust set of discriminative features for subsequent processing stages. The output is denoted as Eq. (19):

$$\mathcal{F}_{ACS} = \{S_r^\alpha(f) | \alpha \in A, f \in \mathbb{R}\} \quad (19)$$

Where, A denotes the set of considered cyclic frequencies. The feature set $\mathcal{F}_{ACS}$ serves as the input to the time–frequency transformation stage.

3.2.2 Short-Time Fourier Transform (STFT)

The cyclostationary feature set $\mathcal{F}_{ACS}$ provides a frequency-dependent representation of the signal's periodic statistical properties. However, under time-varying channel conditions, such as those caused by user mobility or multipath fading, the statistical structure of the received signal evolve over time. To capture this temporal variability, a joint time–frequency analysis is required. The Short-Time Fourier Transform (STFT) is employed to obtain a localized spectral representation, thereby extending the cyclostationary analysis into the time–frequency domain.

Let $x(t)$ denote the input signal to this stage, where $x(t) \in \mathcal{F}_{ACS}$. The STFT of $x(t)$ with respect to a window function $w(t)$ is defined as Eq. (20):

$$X(t, f) = \int_{-\infty}^{\infty} x(\tau) w(\tau - t) e^{-j2\pi f \tau} d\tau \quad (20)$$

Where, $w(\tau - t)$ is a sliding analysis window centered at time t , and f denotes the frequency variable. The choice of $w(t)$ governs the trade-off between temporal and spectral resolution, with narrower windows providing finer time resolution and wider windows enhancing frequency resolution. The squared magnitude of the STFT yields the spectrogram, expressed as Eq. (21).

$$\mathcal{P}(t, f) = |X(t, f)|^2 \quad (21)$$

Which provides a two-dimensional energy distribution over the joint time–frequency plane. This spectrogram captures the evolution of the signal's cyclostationary characteristics across time, thereby enabling resilience to dynamic environmental variations.

For subsequent processing, the time–frequency representation obtained from the STFT is denoted as Eq. (22):

$$\mathcal{F}_{STFT} = \{\mathcal{P}(t, f) | t \in \mathbb{R}, f \in \mathbb{R}\} \quad (22)$$

This representation not only preserves the cyclic features extracted in the previous stage but also enriches them with temporal dynamics, ensuring robustness against mobility-induced spectral fluctuations. The feature set $\mathcal{F}_{STFT}$ is next subjected to normalization and denoising to further enhance its reliability.

3.2.3 Normalization (Z-score Normalization)

The time–frequency representation $\mathcal{F}_{STFT}$ obtained from the previous stage encapsulates both the spectral and temporal dynamics of the received signal. However, the raw spectrogram values often span a wide range due to variations in signal power, environmental interference, and channel fluctuations. Such variability introduce bias during subsequent learning and decision-making, where higher-energy components could dominate feature representation. To mitigate this issue, Z-score normalization is applied to scale the features into a standardized range with zero mean and unit variance.

Let $\mathcal{P}(t, f)$ denote the energy at time t and frequency f . The Z-score normalized feature value is expressed as Eq. (23):

$$\hat{\mathcal{P}}(t, f) = \frac{\mathcal{P}(t, f) - \mu_{\mathcal{P}}}{\sigma_{\mathcal{P}}} \quad (23)$$

Where, $\mu_{\mathcal{P}}$ represents the mean of all spectrogram values and $\sigma_{\mathcal{P}}$ represents the standard deviation. This transformation ensures that each feature contributes equally to subsequent processing by eliminating scale disparities.

The normalization process is particularly crucial under time-variant environmental conditions, where received signal strengths fluctuate due to fading and mobility. By centering the features around zero and rescaling according to the distribution spread, Z-score normalization enhances numerical stability and prevents skewed learning dynamics in downstream models. The output is denoted as Eq. (24):

$$\mathcal{F}_{\text{Norm}} = \{\hat{\mathcal{P}}(t, f) | (t, f) \in \mathcal{F}_{\text{STFT}}\} \quad (24)$$

which represents the standardized time–frequency feature set. This normalized representation serves as the input to the denoising stage, where residual channel-induced distortions are further suppressed.

3.2.4 Denoising (Wavelet Denoising)

Although normalization alleviates variations in feature magnitudes, the time–frequency representation $\mathcal{F}_{\text{Norm}}$ remains susceptible to distortions introduced by multipath fading, Doppler shifts, and background interference. These artifacts manifest as irregular fluctuations or spurious energy concentrations in the spectrogram, potentially masking the true cyclostationary patterns of the underlying signal. To address this issue, wavelet-based denoising is employed as the final stage of preprocessing.

Wavelet denoising exploits the ability of wavelet transforms to provide localized time–frequency analysis with adaptive resolution. The normalized spectrogram $\hat{\mathcal{P}}(t, f)$ is first decomposed into approximation and detail coefficients through a discrete wavelet transform (DWT). Let $\hat{\mathcal{P}}(t, f) \xrightarrow{\text{DWT}} \{A_j, D_j\}_{j=1}^J$, where A_j and D_j denote the approximation and detail coefficients at decomposition level j , and J signifies the maximum number of levels. Noise components typically dominate the high-frequency detail coefficients, whereas the essential spectral features of the signal are concentrated within the approximation coefficients.

A thresholding operation is then applied to the detail coefficients to suppress noise. This process can be expressed as Eq. (25):

$$\tilde{D}_j(k) = \begin{cases} 0, & |D_j(k)| < \lambda_j, \\ D_j(k), & |D_j(k)| \geq \lambda_j. \end{cases} \quad (25)$$

Where, λ_j denotes the threshold at level j , and k indexes the coefficients. After thresholding, the denoised spectrogram is reconstructed by performing the inverse discrete wavelet transform (IDWT), which is defined as Eq. (26):

$$\tilde{\mathcal{P}}(t, f) = \text{IDWT}(\{A_j, \tilde{D}_j\}_{j=1}^J) \quad (26)$$

The reconstructed spectrogram $\tilde{\mathcal{P}}(t, f)$ retains the essential cyclostationary and temporal-spectral features while significantly reducing interference and fading artifacts. The final denoised output of the preprocessing stage is denoted as Eq. (27):

$$\mathcal{F}_{\text{Denoised}} = \{\tilde{\mathcal{P}}(t, f) | (t, f) \in \mathcal{F}_{\text{Norm}}\} \quad (27)$$

This refined feature set provides a robust, noise-suppressed, and well-structured representation of the received signal, serving as the foundation for subsequent learning and dynamic thresholding mechanisms.

3.3. Temporal Spectrum Variation Modeling (TSVM) via TSVANet (Temporal Spectrum Variation Attention Network)

While the preprocessing stage refines the received signal into a noise-suppressed and normalized time–frequency representation, effective spectrum sensing under time-varying environments further requires modeling the temporal dynamics introduced by Doppler shifts, user mobility, and channel fluctuations. Static spectral analysis alone is insufficient, as the spectrum occupancy patterns evolve over time and demand a framework capable of capturing both local spectral dependencies and long-term temporal variations.

To address this challenge, a novel architecture termed Temporal Spectrum Variation Attention Network (TSVANet) is introduced. TSVANet integrates a feedforward artificial neural network (ANN) with a temporal attention mechanism, thereby combining the strengths of local feature extraction and global dependency modeling. The ANN component captures fine-grained spectral correlations in the preprocessed features, while the attention layer selectively emphasizes relevant temporal contexts, enabling the system to remain robust under mobility-induced spectrum variability. The joint representation produced by TSVANet forms the basis for dynamic threshold adaptation in subsequent decision-making. Figure 2 depicts the architecture of the proposed TSVANet.

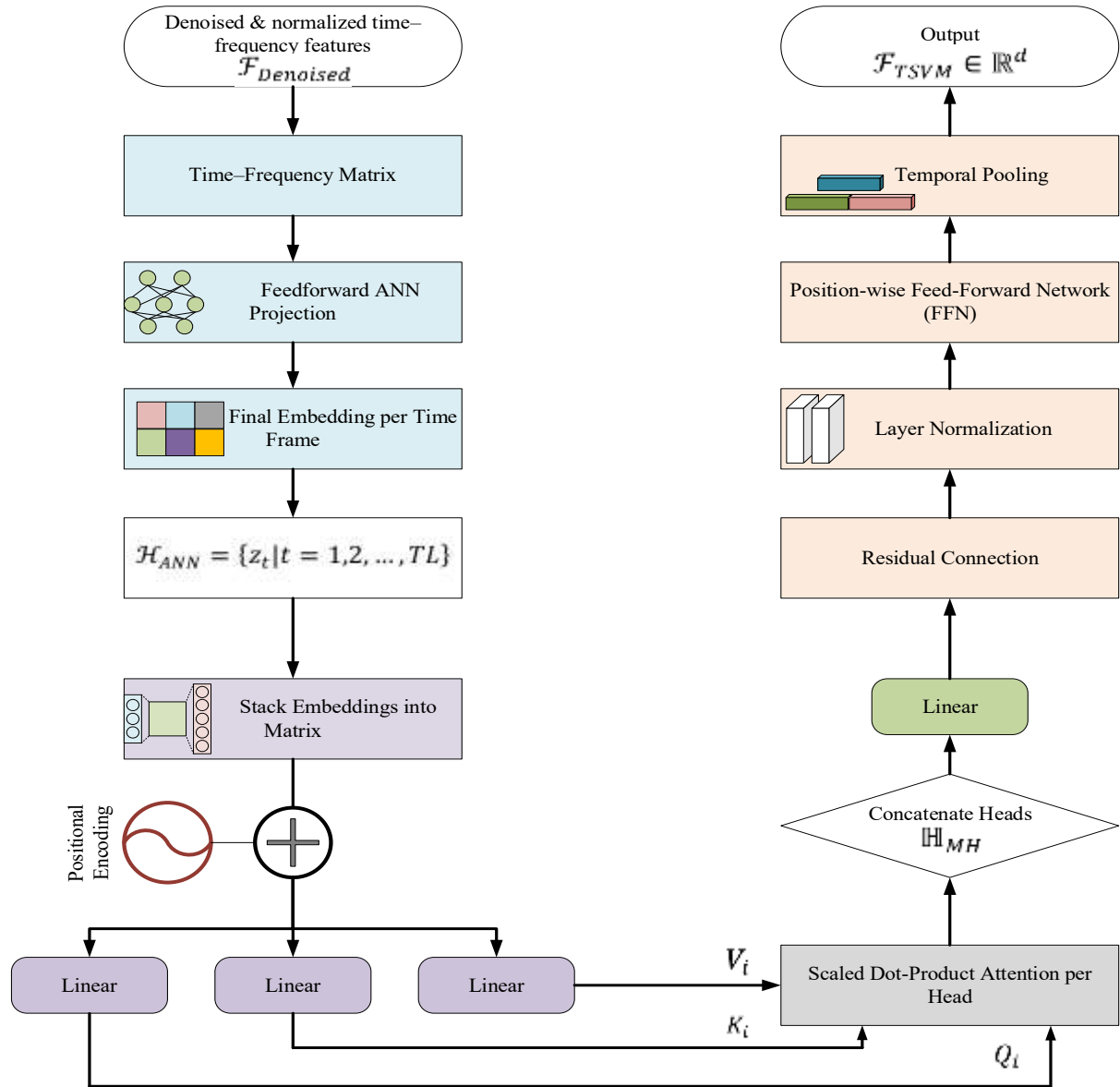


Figure 2: Architecture of the Proposed TSVANet

3.3.1 ANN-based Local Spectral Feature Extraction

The input to the modeling stage is the denoised and normalized time-frequency feature set, denoted as $\mathcal{F}_{\text{Denoised}}$. Although this representation is well-structured and noise-suppressed, it remains high-dimensional and exhibits complex local correlations across frequency bins. To reduce dimensionality and extract discriminative local spectral patterns, a feedforward artificial neural network (ANN) is employed as the first component of TSVANet. Let $X \in \mathbb{R}^{\text{TF} \times \text{FB}}$ represent the preprocessed time-frequency matrix, where TF denotes the number of time frames and FB denotes the number of frequency bins. Each row vector $x_t \in \mathbb{R}^{\text{FB}}$ corresponds to the spectral distribution at time index t . The ANN projects these spectral vectors into a lower-dimensional embedding space through a sequence of fully connected transformations followed by nonlinear activations. Formally, the hidden representation at layer l is expressed as Eq. (28).

$$h_t^{(l)} = \sigma_{\text{na}}(W^{(l)}h_t^{(l-1)} + b^{(l)}), \quad l = 1, 2, \dots, L \quad (28)$$

Where, $W^{(l)}$ and $b^{(l)}$ denote the weight matrix and bias vector at layer l , $\sigma_{\text{na}}(\cdot)$ represents a nonlinear activation function such as ReLU, and $h_t^{(0)} = x_t$. The final output embedding at time index t after L layers is given by $z_t = h_t^{(L)}$.

This transformation produces a compact feature vector $z_t \in \mathbb{R}^d$, where $d \ll \text{FB}$, encapsulating the most informative local spectral dependencies present at time t . The set of embeddings over all time frames forms the local spectral representation, which is defined as Eq. (29):

$$\mathcal{H}_{\text{ANN}} = \{z_t | t = 1, 2, \dots, TL\} \quad (29)$$

The representation $\mathcal{H}_{\text{ANN}}$ serves as the intermediate feature space in TSVANet, preserving local spectral structures while significantly reducing redundancy. This output forms the input to the subsequent attention mechanism, which models long-term temporal dependencies across the sequence of embeddings.

3.3.2 Attention Module: Multi-Head Temporal Self-Attention (MHTSA)

The input to the attention module is the sequence of local spectral embeddings produced by the ANN, denoted by $\mathcal{H}_{\text{ANN}} = \{z_t | t = 1, 2, \dots, TL\}$ with $z_t \in \mathbb{R}^d$. For compactness, stack these embeddings into a matrix $Z \in \mathbb{R}^{TL \times d}$, where each row corresponds to a time-indexed embedding. To enable the attention mechanism to exploit temporal ordering, a positional encoding $\mathbb{P}_t \in \mathbb{R}^d$ is added to each embedding. Using a sinusoidal scheme, the positional encoding is defined component-wise as Eq. (30).

$$\mathbb{P}_{t,2i} = \sin\left(t/10000^{\frac{2i}{d}}\right), \quad \mathbb{P}_{t,2i+1} = \cos\left(t/10000^{\frac{2i}{d}}\right) \quad (30)$$

Here, $i = 0, \dots, \frac{d}{2} - 1$, and the position-augmented input becomes $\tilde{Z} = Z + \mathbb{P}$, where $\mathbb{P}$ collects $\{\mathbb{P}_t\}$.

Multi-Head Temporal Self-Attention (MHTSA) projects $\tilde{Z}$ into query, key and value space and computes multiple parallel attention heads to model diverse temporal relationships. Let ah denote the number of attention heads and d_k the dimensionality of each head such that $d = ah \cdot d_k$. The per-head linear projections are written as Eq. (31).

$$Q_i = \tilde{Z}W_Q^{(i)}, \quad K_i = \tilde{Z}W_K^{(i)}, \quad V_i = \tilde{Z}W_V^{(i)}, \quad i = 1, \dots, ah \quad (31)$$

Where, $W_Q^{(i)}, W_K^{(i)}, W_V^{(i)} \in \mathbb{R}^{d \times d_k}$. Each head computes scaled dot-product attention across all time frames as Eq. (32):

$$\text{head}_i = \text{softmax}\left(\frac{Q_i K_i^T}{\sqrt{d_k}}\right) V_i \quad (32)$$

Yielding $\text{head}_i \in \mathbb{R}^{TL \times d_k}$. The outputs of the ah heads are concatenated and linearly projected to produce the multi-head output as $\mathbb{H}_{\text{MH}} = \text{Concat}(\text{head}_1, \dots, \text{head}_{ah})W_0$, with $W_0 \in \mathbb{R}^{(ahd_k) \times d}$. To promote stable optimization and preserve the original local embeddings, a residual connection and layer normalization are applied as $\mathbb{H} = \text{LayerNorm}(Z + \mathbb{H}_{\text{MH}})$.

To further refine temporal representations and introduce position-wise nonlinearity, a position-wise feed-forward network (FFN) is applied to each time row of $\mathbb{H}$, which is defined as Eq. (33):

$$\text{FFN}(x) = \text{ReLU}(xW_1 + b_1)W_2 + b_2 \quad (33)$$

The FFN output is combined with a second residual connection and layer normalization to produce the final temporally refined matrix, which is mathematically represented as Eq. (34):

$$\mathbb{H}_{\text{attn}} = \text{LayerNorm}\left(\mathbb{H} + \text{FFN}(\mathbb{H})\right) \quad (34)$$

Each row $\mathbb{H}_t^{\text{attn}}$ of $\mathbb{H}_{\text{attn}}$ represents the embedding at time t that has been adaptively informed by the entire temporal context. In set notation, the attention-refined representation is denoted as Eq. (35).

$$\mathcal{H}_{\text{Attn}} = \{\mathbb{H}_t^{\text{attn}} | t = 1, \dots, TL\}, \quad \mathbb{H}_{\text{attn}} \in \mathbb{R}^{TL \times d} \quad (35)$$

The MHTSA formulation allows each time-indexed embedding to attend selectively to all other time frames, thereby capturing long-term dependencies induced by Doppler shifts and user mobility. Multiple heads enable the module to specialize across different temporal scales (for example, slowly varying frequency drift and rapid fading events). Dropout on attention weights, weight decay and layer normalization mitigate overfitting and promote numerical stability during training. Optionally, relative positional encodings replace absolute encodings when shift-invariant temporal relations (time-lag patterns caused by motion) are to be emphasized. The output $\mathcal{H}_{\text{Attn}}$ therefore provides a temporally refined spectro-temporal representation suitable for downstream dynamic-thresholding and decision layers.

The integration of local spectral embeddings from the ANN and long-range temporal dependencies captured through Multi-Head Temporal Self-Attention results in a unified high-level spectro-temporal representation. The refined embedding sequence $\mathcal{H}_{\text{Attn}} = \{\mathbb{H}_t^{\text{attn}}\}_{t=1}^{TL}$ encapsulates both short-term spectral characteristics and mobility-induced temporal variations. To obtain a compact and discriminative representation, temporal pooling is applied across the sequence dimension. Specifically, global average pooling aggregates information over time as Eq. (36).

$$\mathcal{F}_{\text{TSVM}} = \frac{1}{TF} \sum_{t=1}^{TL} \mathcal{H}_{\text{Attn}} \quad (36)$$

Where, $\mathcal{F}_{\text{TSVM}} \in \mathbb{R}^d$ denotes the final feature vector summarizing the temporal spectrum variation. For improved robustness, alternative pooling operators such as max-pooling or attention-weighted pooling can be employed, but average pooling is adopted for its balance between stability and representational compactness.

The resulting vector $\mathcal{F}_{\text{TSVM}}$ is denoted as the Temporal Spectrum Variation Modeling representation (TSVM representation), which serves as the high-level spectro-temporal descriptor of the received signal under time-varying environments. This representation forms the input to the subsequent dynamic thresholding mechanism, ensuring that the detection process is guided not only by instantaneous spectral signatures but also by temporally informed contextual dependencies.

3.4 Environment-Aware Neural Thresholding (EANT)

Accurate spectrum sensing under time-varying wireless environments requires thresholds that adapt dynamically to changing propagation conditions such as mobility-induced Doppler shifts, multipath fading, and unpredictable interference. Traditional fixed-threshold approaches fail to maintain reliable detection performance under such variability, leading to increased false alarms or missed detections. To overcome this limitation, the Environment-Aware Neural Thresholding (EANT) module is introduced. In this stage, the high-level spectro-temporal features $\mathcal{F}_{\text{TSVM}}$, obtained from TSVANet, are mapped into adaptive detection thresholds through a hybrid CNN-GRU

framework. The CNN component captures localized feature variations associated with frequency-selective fading and bursty interference, while the GRU models temporal dependencies, ensuring that thresholding decisions remain consistent across sequential time instants. The output of this stage is a set of dynamic thresholds $\{\theta_t\}_{t=1}^{TL}$, which serve as environment-aware decision boundaries for reliable primary user detection in non-stationary spectrum conditions.

3.4.1 CNN-based Local Feature Refinement

The spectro-temporal representation $\mathcal{F}_{TSVM}$ obtained from TSVANet is first passed through a convolutional neural network (CNN) to refine localized structural features. While $\mathcal{F}_{TSVM}$ encodes both spectral and temporal dynamics, it remains sensitive to short-term distortions introduced by fading, interference spikes, and frequency-selective effects. CNNs are particularly effective in capturing such localized variations due to their ability to apply shared convolutional kernels across the input representation, thereby enhancing robustness to small-scale spectral distortions.

Formally, given the input $\mathcal{F}_{TSVM} \in \mathbb{R}^{TL \times d}$, where TL denotes the temporal length and d signifies the spectral feature dimension, the CNN applies a set of convolutional filters $W^{(c)} \in \mathbb{R}^{k \times d}$, with kernel size k , to extract localized patterns. The convolutional operation is expressed as Eq. (37).

$$z_t^{(c)} = \sigma na(\sum_{i=0}^{k-1} W^{(c)} \cdot \mathcal{F}_{TSVM}[t + i] + b^{(c)}) \quad (37)$$

Where, $z_t^{(c)}$ denotes the convolutional feature at time t, $b^{(c)}$ represents the bias term, and $\sigma na(\cdot)$ represents a nonlinear activation function such as ReLU. The convolutional responses across all filters are aggregated into a feature map as $Z_{CNN} = f_{CNN}(\mathcal{F}_{TSVM})$, where $Z_{CNN} \in \mathbb{R}^{TL \times d'}$ encodes enhanced local representations with d' being the new feature dimension determined by the number of filters.

Through this process, the CNN acts as a localized feature extractor that emphasizes short-term frequency-dependent characteristics while suppressing irrelevant noise artifacts. The refined representation Z_{CNN} forms the input to the subsequent recurrent modeling stage, where long-term temporal dependencies are addressed.

3.4.2 GRU-based Temporal Dependency Modeling

While the CNN module enhances local spectral patterns, reliable spectrum sensing in dynamic wireless environments requires capturing long-term dependencies across sequential time instants. Multipath fading, Doppler shifts, and intermittent interference often introduce temporal correlations that cannot be adequately modeled using convolutional operations alone. To address this challenge, the refined features Z_{CNN} are passed into a Gated Recurrent Unit (GRU) network, which is specifically designed to model sequential dependencies while mitigating the vanishing gradient problem associated with traditional recurrent architectures.

Formally, let $Z_{CNN} = \{z_1, z_2, \dots, z_{TL}\}$, where $z_{ts} \in d'$ denotes the CNN-enhanced feature at time step ts. The GRU computes a hidden state sequence $\{h_1, h_2, \dots, h_{TL}\}$ by iteratively updating its memory as Eq. (38):

$$h_{ts} = (1 - up_{ts}) \odot h_{ts-1} + up_{ts} \odot \widetilde{h_{ts}} \quad (38)$$

Where,

$$re_{ts} = \sigma(W_{re} z_{ts} + U_{re} h_{ts-1} + b_{re}),$$

$$up_{ts} = \sigma(W_{up} z_{ts} + U_{up} h_{ts-1} + b_{up}),$$

$$\widetilde{h_{ts}} = \tanh(W_h z_{ts} + U_h (re_{ts} \odot h_{ts-1}) + b_h)$$

Where, re_{ts} and up_{ts} represent the reset and update gates respectively, $\odot$ denotes element-wise multiplication, and h_{ts} signifies the hidden state capturing the temporal context at step ts. The GRU output is then represented as Eq. (39).

$$\mathbb{H}_{GRU} = f_{GRU}(Z_{CNN}) \quad (39)$$

Where, $\mathbb{H}_{GRU} \in \mathbb{R}^{TL \times d''}$ captures temporally-aware spectral representations with dimension d'' .

By integrating memory dynamics through gated updates, the GRU ensures that spectrum sensing thresholds adapt smoothly to gradual variations in wireless conditions while remaining responsive to sudden temporal changes. This temporally contextualized representation serves as the basis for learning dynamic thresholds in the subsequent environment-aware mapping stage.

3.4.3 Dynamic Threshold Generation

The temporally contextualized representation $\mathbb{H}_{GRU}$ serves as the input for dynamic threshold generation, enabling environment-aware spectrum sensing decisions. Unlike static thresholding approaches that apply a fixed decision boundary regardless of propagation variations, the proposed mechanism learns adaptive thresholds that continuously evolve with the underlying wireless conditions. This ensures robustness against fading, Doppler-induced spectral shifts, and interference fluctuations. To map $\mathbb{H}_{GRU}$ into dynamic decision thresholds, a fully connected projection layer is applied as Eq. (40):

$$\theta_t = f_{FC}(h_{ts}) = W_\theta h_{ts} + b_\theta \quad (40)$$

Where, $\theta_t \in \mathbb{R}$ denotes the adaptive threshold at time step ts, W_θ and b_θ signifies trainable parameters, and h_{ts} is the GRU hidden state at that time step. The complete set of thresholds over a time horizon T_h is expressed as $\Theta = \{\theta_1, \theta_2, \dots, \theta_{T_h}\}$.

The generated thresholds Θ dynamically adjust to both slow and fast temporal variations, thereby improving the reliability of primary user (PU) detection under non-stationary conditions. The CNN-GRU synergy ensures that localized spectral distortions and long-term temporal dependencies are jointly captured before thresholding,

resulting in thresholds that are sensitive to environmental dynamics while maintaining stability across fluctuating spectrum states.

The final output of this stage, denoted as Θ_{EANT} , represents the environment-aware adaptive threshold sequence that forms the decision boundary for subsequent spectrum sensing as $\Theta_{\text{EANT}} = \mathcal{G}_{\text{EANT}}(\mathcal{F}_{\text{TSM}})$.

Here, $\mathcal{G}_{\text{EANT}}(\cdot)$ denotes the mapping function implemented by the CNN–GRU hybrid network, translating spectro-temporal features into dynamic thresholds. These thresholds directly feed into the decision-making stage, where they determine spectrum occupancy based on the presence or absence of primary user signals.

3.5 Adaptive Learning Rate Optimization (ALRO) via Social Spider–Krill Hybrid (SKH)

Neural network training for spectrum sensing is challenged by non-stationary wireless conditions, where fading, Doppler shifts, and interference continuously alter input distributions. Fixed learning rate schedules used in standard optimizers (Adam, RMSprop) often result in instability or premature convergence. To overcome this, an adaptive scheme termed Adaptive Learning Rate Optimization using Social Spider–Krill Hybrid (ALRO-SKH) is introduced.

ALRO-SKH integrates Social Spider Optimization (SSO) for global exploration with the Krill Herd Algorithm (KHA) for local exploitation, enabling dynamic adjustment of learning rates and hyperparameters. This hybrid outer-loop guides gradient-based inner-loop updates, ensuring both adaptability and stability under time-varying environments. By incorporating synthetic and real-world signals during training, ALRO-SKH maintains robust convergence and enhances generalization across diverse channel conditions.

The resulting optimization mechanism ensures reliable learning dynamics and prepares the framework for final decision-making based on adaptive thresholds.

Dynamics of ALRO-SKH

The input to the adaptive optimization stage is denoted as Θ_{EANT} , representing the parameter set generated from the Environment-Aware Neural Thresholding module. The objective of ALRO-SKH is to iteratively refine these parameters through a hybrid optimization strategy that balances global exploration with local exploitation while adapting the learning rate to the dynamic spectrum environment.

In the first stage, Social Spider Optimization (SSO) models the cooperative behavior of spiders in a communal web, where individuals share vibrations to update their positions in the solution space. Each candidate solution θ_i^t at iteration t is updated according to Eq. (41):

$$\theta_i^{t+1} = \theta_i^t + E \cdot \mathcal{V}_i^t + A \cdot (\theta_{\text{best}}^t - \theta_i^t) \quad (41)$$

Where, $\mathcal{V}_i^t$ represents the vibration-induced movement influenced by neighboring solutions, θ_{best}^t denotes the best-performing parameter set at iteration t , and E, A are weighting coefficients controlling exploration and attraction. This process enhances global search capability, preventing premature convergence to suboptimal solutions.

In the second stage, Krill Herd Algorithm (KHA) refines the exploration by modeling the herding dynamics of krill. Each solution is updated based on induced motion, foraging activity, and random diffusion as Eq. (42):

$$\theta_i^{t+1} = \theta_i^t + \mathcal{M}_i^t + \mathcal{F}_i^t + \mathcal{D}_i^t \quad (42)$$

Where, $\mathcal{M}_i^t$ corresponds to the motion induced by the local density of solutions, $\mathcal{F}_i^t$ represents foraging adjustments guided by the global best position, and $\mathcal{D}_i^t$ captures stochastic perturbations that prevent stagnation. KHA thus enhances exploitation by locally optimizing around promising regions discovered by SSO.

The hybridization in ALRO-SKH is achieved by adaptively weighting both strategies during each iteration, which is defined as Eq. (43):

$$\theta_{\text{ALRO}}^{t+1} = \lambda \cdot \theta_{\text{SSO}}^{t+1} + (1 - \lambda) \cdot \theta_{\text{KHA}}^{t+1} \quad (43)$$

Where, $\lambda \in [0, 1]$ dynamically balances exploration and exploitation according to the stability of the training process. When high variability is detected in the spectral environment, λ is increased to emphasize global search, whereas in stable conditions, it decreases to promote local refinement.

The updated parameters $\theta_{\text{ALRO}}^{t+1}$ are used to adapt the learning rate η^t within the gradient-based update rule as Eq. (44):

$$\theta^{t+1} = \theta^t - \eta^t \nabla \mathcal{L}(\theta^t), \eta^t = \mathcal{f}_m(\theta_{\text{ALRO}}^t) \quad (44)$$

Where, $\mathcal{L}(\theta^t)$ denotes the loss function associated with spectrum sensing accuracy, and $\mathcal{f}_m(\cdot)$ represents the adaptive mapping from the hybrid optimizer to the learning rate.

Through this two-level optimization scheme, the network achieves stable convergence under non-stationary data streams while preserving adaptability to unseen channel conditions. The final output of this stage, denoted as Θ_{ALRO} , represents the adaptively optimized parameter set that feeds into the final decision-making stage.

3.6 Decision and Spectrum Sensing

The final stage of the proposed framework translates the adaptively optimized parameters Θ_{ALRO} into actionable spectrum availability decisions. At this stage, the artificial neural network, trained and optimized through the preceding modules, produces dynamic thresholds that are sensitive to time-varying propagation conditions such as mobility, fading, and interference. These thresholds, denoted as τ_{dyn} , evolve adaptively according to the spectral and temporal context provided by the processed feature set.

The decision mechanism operates by comparing the incoming feature representation $\mathcal{F}_{\text{TSM}}$, which encodes high-level spectro-temporal dynamics, against the learned adaptive threshold τ_{dyn} . The decision rule is expressed as Eq. (45):

$$\mathcal{D} = \begin{cases} 1, & \text{if } \|\mathcal{F}_{\text{TSVM}}\| \geq \tau_{\text{dyn}} \quad (\text{Signal Present, Occupied}) \\ 0, & \text{if } \|\mathcal{F}_{\text{TSVM}}\| < \tau_{\text{dyn}} \quad (\text{Signal Absent, Free}) \end{cases} \quad (45)$$

Where, $\mathcal{D} \in \{0,1\}$ denotes the final binary outcome of the spectrum sensing process. By leveraging adaptive thresholds instead of fixed ones, the system avoids the common pitfalls of conventional energy detection, where noise uncertainty or channel variability often leads to false alarms or missed detections.

This decision-making mechanism ensures that primary user activity is detected with high reliability, even under hostile and non-stationary channel conditions. The adaptive nature of the framework allows it to balance detection probability and false-alarm rate dynamically, thereby supporting efficient spectrum utilization while safeguarding primary user transmissions.

With this final stage, the proposed architecture completes its pipeline, offering a robust and intelligent spectrum sensing paradigm capable of sustained operation in real-world cognitive radio environments.

7. RESULT AND DISCUSSION

This section presents the experimental validation of the proposed framework, highlighting its effectiveness under dynamic spectrum environments. The evaluation includes implementation details and parameter settings, followed by a comparative performance analysis with existing spectrum sensing techniques. Results demonstrate improvements in adaptability, detection accuracy, and robustness against non-stationary channel conditions.

7.1 Experimental Setup

The proposed framework has been implemented in Python, using standard deep learning and signal processing libraries. For comparison, several existing techniques have been considered, including TFCFN [18], STFT-RADN [19], CNN-LSTM [20], and DBN-FOA [22]. Both synthetic datasets and real-world spectrum traces have been employed to evaluate performance under varying conditions.

The assessment has been conducted using key metrics such as Accuracy, Precision, Sensitivity, Specificity, F1 Score, Negative Predictive Value (NPV), Matthews Correlation Coefficient (MCC), False Positive Rate (FPR), and False Negative Rate (FNR), ensuring a comprehensive evaluation of detection effectiveness and robustness.

7.2 Performance Analysis Between the Proposed and Existing Techniques

The proposed approach has demonstrated superior performance across all evaluation metrics when compared with state-of-the-art methods, as presented in Table 1 and graphically shown in Figure 3. In terms of Accuracy, the proposed framework has achieved 99.39%, outperforming TFCFN [18] (97.64%), STFT-RADN [19] (95.34%), CNN-LSTM [20] (94.94%), and DBN-FOA [22] (96.24%). This notable improvement in accuracy can be attributed to the incorporation of the Temporal Spectrum Variation Modeling (TSVM) via TSVANet, which effectively captures both local spectral patterns and long-term dependencies through the combination of ANN-based feature extraction and the Multi-Head Temporal Self-Attention (MHTSA) mechanism. By modeling Doppler-induced variations more comprehensively, the system has maintained stability even under dynamic channel conditions, thus reducing misclassification errors.

When evaluated on F1 Score, the proposed model has reached 98.32%, surpassing TFCFN [18] (97.24%), STFT-RADN [19] (94.94%), CNN-LSTM [20] (94.64%), and DBN-FOA [22] (96.04%). The higher F1 Score reflects the balance achieved between precision and sensitivity, which is further strengthened by the Adaptive Learning Rate Optimization (ALRO-SKH) mechanism. The hybrid integration of Social Spider Optimization and Krill Herd Algorithm dynamically adjusts the learning process, ensuring stable convergence under non-stationary data streams and enhancing the generalization ability of the network.

Table 1: Performance comparison of the proposed method with existing spectrum sensing techniques

Method	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)	F1 Score (%)	NPV (%)	MCC (%)	FPR (%)	FNR (%)
Proposed	99.39	98.32	99.09	99.3	98.32	98.39	99.26	1.84	1.75
TFCFN [18]	97.64	96.54	97.54	97.94	97.24	97.69	97.39	2.2	2.1
STFT-RADN [19]	95.34	94.44	94.89	95.24	94.94	95.24	95.74	4.9	5.1
CNN-LSTM [20]	94.94	94.04	94.74	95.14	94.64	95.09	95.39	5	5.2
DBN-FOA [22]	96.24	95.34	95.94	96.44	96.04	96.29	96.24	3.7	3.4

Other metrics also confirm the robustness of the proposed method. For example, Sensitivity (99.09%) and Specificity (99.3%) indicate that the framework has reliably detected primary user activity while minimizing false alarms. Likewise, the MCC value of 99.26% highlights strong correlation across all predicted outcomes, exceeding the best performance among existing methods. Furthermore, the proposed method has consistently produced lower FPR (1.84%) and FNR (1.75%), establishing its reliability under adverse spectrum environments.

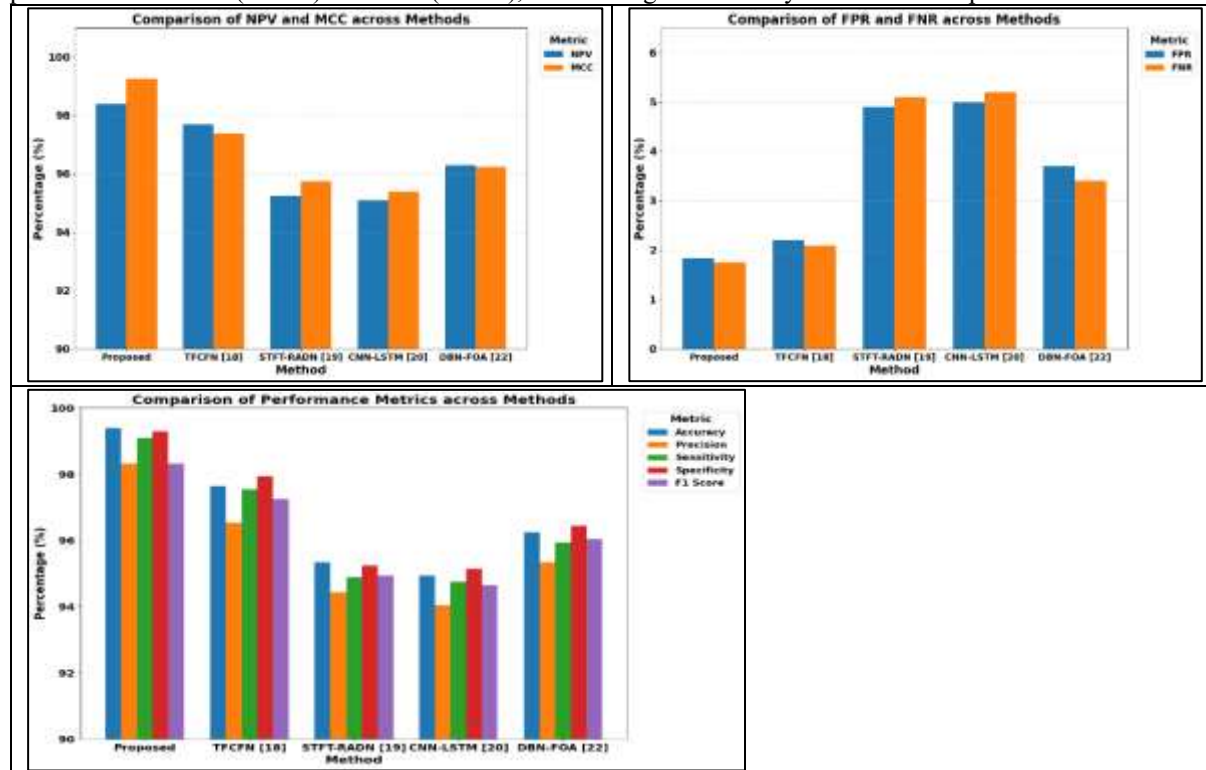


Figure 3: Comparison Analysis of the Proposed Method with Existing Techniques

Overall, the improvements across multiple metrics demonstrate the effectiveness of combining TSVM-based feature modeling with ALRO-SKH optimization, which collectively enhance adaptability, robustness, and spectrum detection accuracy under varying wireless conditions.

5. CONCLUSION

This work has presented a dynamic thresholding framework for cyclostationary spectrum sensing using artificial neural networks under time-variant maritime environments. The system incorporated preprocessing, Temporal Spectrum Variation Modeling via TSVANet, Environment-Aware Neural Thresholding, and Adaptive Learning Rate Optimization through a Social Spider–Krill Hybrid. The framework has been implemented in Python, and experimental results have demonstrated high performance, achieving Accuracy of 99.39%, F1-Score of 98.32%, and MCC of 99.26%. The proposed approach has significantly enhanced spectrum sensing reliability, adaptability, and robustness against non-stationary channel conditions. Future work could explore real-time deployment on embedded platforms, integration with cognitive radio networks at larger scales, and extension to multi-band or multi-user spectrum scenarios.

REFERENCES

- [1]. Nasser, A., Al Haj Hassan, H., Abou Chaaya, J., Mansour, A., & Yao, K. C. (2021). Spectrum sensing for cognitive radio: Recent advances and future challenge. *Sensors*, 21(7), 2408.
- [2]. Nair, R. G., & Narayanan, K. (2023). Cooperative spectrum sensing in cognitive radio networks using machine learning techniques. *Applied Nanoscience*, 13(3), 2353-2363.
- [3]. Xing, H., Qin, H., Luo, S., Dai, P., Xu, L., & Cheng, X. (2022). Spectrum sensing in cognitive radio: A deep learning based model. *Transactions on Emerging Telecommunications Technologies*, 33(1), e4388.
- [4]. Mustafa, W., Ghafoor, H., & Koo, I. (2025). Coverage Enhancement using UAVs for Cognitive Marine Network. *IEEE Access*.
- [5]. Joshi, N., Arora, N., Yadav, H., & Sharma, S. C. (2024). AI-Driven Cognitive Radio Networks for Transforming Industries and Sectors Towards a Smart World. In *Recent Trends in Artificial Intelligence Towards a Smart World: Applications in Industries and Sectors* (pp. 1-35). Singapore: Springer Nature Singapore.
- [6]. Tekbiyik, K., Kurt, G. K., & Lesage-Landry, A. (2024). Federated learning for UAV-based spectrum sensing: Enhancing accuracy through SNR-weighted model aggregation. *arXiv preprint arXiv:2411.11159*.

- [7]. Kim, T. K., Jeon, Y. S., & Min, M. (2021). Training length adaptation for reinforcement learning-based detection in time-varying massive MIMO systems with one-bit ADCs. *IEEE Transactions on Vehicular Technology*, 70(7), 6999-7011.
- [8]. Haldorai, A., Sivaraj, J., Nagabushanam, M., & Kingston Roberts, M. (2022). Cognitive wireless networks based spectrum sensing strategies: A comparative analysis. *Applied Computational Intelligence and Soft Computing*, 2022(1), 6988847.
- [9]. Nawaz, T., & Alzahrani, A. (2023). Machine-learning-assisted cyclostationary spectral analysis for joint signal classification and jammer detection at the physical layer of cognitive radio. *Sensors*, 23(16), 7144.
- [10]. Mei, R., & Wang, Z. (2023). Deep learning-based wideband spectrum sensing: A low computational complexity approach. *IEEE Communications Letters*, 27(10), 2633-2637.
- [11]. Xie, J., Fang, J., Liu, C., & Li, X. (2020). Deep learning-based spectrum sensing in cognitive radio: A CNN-LSTM approach. *IEEE Communications Letters*, 24(10), 2196-2200.
- [12]. Kumar, A., Gaur, N., Chakravarty, S., Alsharif, M. H., Uthansakul, P., & Uthansakul, M. (2024). Analysis of spectrum sensing using deep learning algorithms: CNNs and RNNs. *Ain Shams Engineering Journal*, 15(3), 102505.
- [13]. Zhang, L., Liu, J., Jin, B., & Wei, X. (2025). Meta-Learning Task Relations for Ensemble-Based Temporal Domain Generalization in Sensor Data Forecasting. *Sensors*, 25(14), 4434.
- [14]. Zhang, H., Zhou, F., & Yuen, Q. W. C. (2025). Spectrum Cognition: Semantic Situation for Next-Generation Spectrum Management. *arXiv preprint arXiv:2509.00851*.
- [15]. Srivastava, V., & Prasad, B. (2025). Enhancing Adaptive Spectrum Access: An Intelligent Reflecting Surface Assisted CRN for Future Wireless Communication. *IEEE Access*.
- [16]. Sivagurunathan, P. T., Ramakrishnan, P., & Sathishkumar, N. (2021). Recent paradigms for efficient spectrum sensing in cognitive radio networks: Issues and challenges. In *Journal of Physics: Conference Series* (Vol. 1717, No. 1, p. 012057). IOP Publishing.
- [17]. Sharma, H., Yadav, S., & Kumar, A. (2025). Enhanced spectrum sensing in optical-NOMA for 256-QAM: a hybrid energy and matched filter detection approach. *Journal of Optical Communications*, (0).
- [18]. Xi, H., Guo, W., Yang, Y., Yuan, R., & Ma, H. (2024). Cross-attention mechanism-based spectrum sensing in generalized Gaussian noise. *Scientific Reports*, 14(1), 23261.
- [19]. Wang, A., Zhu, T., & Meng, Q. (2024). Spectrum sensing method based on STFT-RADN in cognitive radio networks. *Sensors*, 24(17), 5792.
- [20]. Wang, K., Chen, Y., Bo, D., & Wang, S. (2025). A novel multi-user collaborative cognitive radio spectrum sensing model: Based on a CNN-LSTM model. *PloS one*, 20(1), e0316291.
- [21]. Ansari, A. H., & Gulhane, S. M. (2025). Deep Learning With Harmonic Elk Herd Optimization for Spectrum Sensing With Cyclostationary in Cognitive Radio Network. *Transactions on Emerging Telecommunications Technologies*, 36(8), e70215.
- [22]. Sonti, S. R., & Prasad, M. S. G. (2021). Deep belief network with FOA-based cooperative spectrum sensing in cognitive radio network. *International Journal of Communication Networks and Distributed Systems*, 27(3), 323-347.