DCT–AE for Spectrum Data Augmentation and MLP on Sensing Decision

Satyajit Bhunia[1,2], Seba Maity[2], Santi P. Maity[1]
[1]*Department of Information Technology, Indian Institute of Engineering Science and Technology, Shibpur*, West Bengal, India
[2]*Department of ECE, College of Engineering and Management Kolaghat*, West Bengal, India
[1,2]Email: sjb@cemk.ac.in [2]Email: sebamaity@cemk.ac.in [1]Email: santipmaity@it.iiests.ac.in

***Abstract*—This paper presents a Discrete Cosine Transform (DCT) based Autoencoder (AE) and a light weight Multilayer Perceptron (MLP) framework for Primary User (PU) spectrum data generation/ augmentation followed by sensing decision to develop cognitive radio networks (CRN). The proposed approach employs tailored DCT domain transformation to generate realistic and diverse spectrum snapshots, thereby enhancing the robustness of learning-based spectrum sensing (SS) under low Signal to Noise Ratio (SNR) and generalize Gaussian noise conditions. An AE is utilized to denoise and compress the spectral features, while a downstream MLP performs binary classification to determine the presence or absence of PU activity. We benchmark the proposed method against two recent works on Time–Frequency Cross-Fusion Network (TFCFN) and Short Term Fourier Transform–Convolutional Neural Network (STFT-CNN), which incorporates cross-attention for enhanced spectral feature fusion. Evaluations using probability of detection (P_d), probability of false alarm (P_f), accuracy, and inference latency indicate that the proposed DCT–AE–MLP framework provides strong robustness at very low SNRs while maintaining competitive computational complexity. Specifically, DCT–AE–MLP achieves $P_d \approx 0.98$ at $P_f = 0.05$, maintains $P_d \approx 0.90$ at $P_f = 0.3$, and sustains $P_d \approx 0.72$ at $P_f = 0.7$; it also achieves $P_d \approx 0.99$ at 0 dB, 0.97 at -10 dB, 0.93 at -15 dB, and 0.84 at -20 dB, across the above SNR levels.**

***Index Terms*—Spectrum Sensing, DCT, AE, SNR, TFCFN, STFT–CNN, Data Augmentation, Cognitive Radio, MLP.**

I. Introduction

The ongoing digital transformation driven by the Internet of Things (IoT) has resulted in an exponential rises on wireless devices and data-hungry applications such as smart healthcare, autonomous mobility, immersive media, industrial automation etc [1]. This sharp rise in the density of heterogeneous wireless nodes is exerting unprecedented pressure on the limited radio spectrum. Reports from global regulators consistently highlight that spectrum scarcity is becoming a key bottleneck to guarantee seamless connectivity, low latency, and reliable quality-of-service (QoS) in emerging networks. As modern societies rely more heavily on pervasive digital services, the demand for intelligent, flexible, and energy-efficient spectrum access has intensified, necessitating advanced spectrum-sharing mechanisms [2].

Despite this explosive growth, conventional spectrum regulation still follows a static allocation model, in which fixed frequency bands are licensed, exclusively to specific services over long durations. Although this ensures strong protection of spectrum for primary user (PU), extensive measurement campaigns reveal that many licensed bands remain idle for significant periods of time, resulting in severe underutilization of valuable spectrum resources. To mitigate this inefficiency, the concept of Cognitive Radio (CR) was proposed as an opportunistic spectrum sharing scheme that supports dynamic spectrum access through advanced sensing, learning, and adaptive transmission mechanisms. CR systems allow secondary users (SUs) to opportunistically exploit unused spectrum segments without disrupting primary communications [3], [4].

Spectrum sensing (SS) represents the core functionality of CR systems, serving as the mechanism that determines whether a PU is active or absent over a bandwidth due to data transmission. Several classical approaches have already been reported with relative merits and demerits [5]–[8]. Energy Detection (ED) [5], [6] is one of the simplest and the most widely used technique due to its independence from PU signal characteristics; however, its performance severely deteriorates under noise uncertainty and low SNR scenarios. Cyclostationary Feature Detection (CFD) [7], [8] extracts periodic statistical features inherent in modulated signals, offering robustness to noise but demanding higher computational resources. Matched filtering (MF) [9], while optimal under known PU waveform assumptions, is often impractical given the lack of prior information in real deployments. Advanced statistical approaches, including covariance-based detection [10] and eigenvalue-based sensing [11], offer significant improvements but still face limitations in fading channels such as Rayleigh, Rician, and shadowing. These challenges highlight the need for more resilient and data-adaptive sensing paradigms in challenging wireless environment.

To address the shortcomings of traditional detectors, recent research has shifted toward data-driven SS, leveraging the strengths of machine learning (ML) and deep learning (DL) to address challenges in particular modeling of the wirless channel, intermittent nature of PU data traffic, transmit power variation, hidden node problems etc. Classical ML models, such as Support Vector Machines (SVMs), Naïve Bayes, and Ensemble Clasifiers (EC) have provided improvements by learning discriminative features extracted from hand-engineered signal descriptors. However, present DL architectures, including Convolutional Neural Networks (CNNs) [12], Recurrent Models [13], Autoencoders (AE) [14] and Transformer based frameworks [15], [16] have demonstrated superior ability to learn robust representations directly from raw or transformed spectral data [17]. Such models have achieved enhanced resilience under fading, non-Gaussian noise, interference, extreme low SNR problems etc. AE based feature learning, in particular, has gained attention for its ability to perform unsupervised denoising and latent-space compression [18].

Motivated by these advances, this work proposes a hybrid DCT–AE with Multilayer Perceptron (MLP) framework, designed to generate enriched spectrum data and yields highly accurate PU sensing decisions, under realistic channel conditions. We benchmark the proposed DCT–AE against the recent Time–Frequency Cross Fusion Network (TFCFN) [19] and Short Term Fourier Transform (STFT)–CNN [20] too.

II. Related Works

A. Literature Review

Since the primary focus of this work is to develop data driven SS, a brief literature review is made for SS using ML–DL methods. DL approaches for SS have been widely explored in the recent years. CNNs [21] trained on power spectral density inputs, hybrid CNN–LSTM (Long Short Term Memory) models [22], [23] that capture temporal dependencies, and attention/ transformer based detectors [16] for wideband sensing have all demonstrated significant performance gains over classical energy and cyclostationary detectors [8], [9].

The works, TFCFN [19] uses GRUs to capture time–domain features and CNN layers to learn frequency–domain information, followed by a cross–attention module for feature fusion. STFT–CNN [20] operates as a time–frequency detector by converting the received signal into an STFT spectrogram, treating it as a 2D image for CNN-based pattern extraction. Both models demonstrate strong robustness, achieving a detection probability (P_d) 90% at -16 dB SNR with a fixed false-alarm rate of $P_f = 0.1$ under generalized Gaussian noise.

Autoencoders (AEs) [18], [24] have been used in spectrum sensing (SS) for unsupervised feature learning, denoising, and data generation. Surveys [25], [26] show that they can learn compact latent features and create synthetic samples when data are limited. However, DCT–AE based augmentation for SS is still mostly unexplored. While changing transform–domain coefficients is common in image and audio augmentation [27], its use for RF spectral data has not been well studied.

To the best of authors' knowledge, DCT–AE is explored first time in PU spectrum data augmentation to address data limitation issue for ML/ DL based SS analysis.

B. Scope of the Work

With the emergence of IoT based various applications such as smart health care, vehicular network, industrial automation, intelligent transport, underground mines and tunnels, gas monitoring systems etc., CRN becomes attractive for spectral and energy efficient data transmission. Furthermore, diverse applications face resource constraints in challenging wireless environment, difficult to model accurately through widely used statistical models. This need data extensive studies for wireless channel modeling which need exploration of ML/ DL methods in SS.

In this work, we proposed a DCT–AE module for PU spectrum data generation/augmentation and a light weight MLP classifier SS decision model to the training pipeline. We also compare the proposed approach with recent DL methods such as TFCFN [19] and STFT–CNN [20]. Earlier studies [28]–[30] highlight the use of AEs, ML/DL-based augmentation, and cross-attention techniques for building realistic SS models.

The main PU–SS aspects in challenging wireless environments considered in this work are as follows:

1) Robust Spectrum data Augmentation model by perturbing DCT coefficients to model channel and noise effects;
2) Using an AE to learn compact features and to generate valid synthetic samples; and
3) Applying a light weight classification model, MLP for binary PU detection (transmitted/ not transmitted).

The main contributions of this work are as follows:

- A DCT-domain AE design for PU spectrum augmentation is proposed to preserve spectral structure while adding controlled perturbations. The DCT–AE jointly denoises, compresses, and reconstructs the spectrum, supplying stable features to the MLP detector.
- Simulation results show that the DCT–AE–MLP achieves $P_d \approx 0.93$ at -16 dB (SNR) for $P_f = 0.10$ under generalized Gaussian noise, due to improved denoising and robust latent features. These findings align with [19], which report $P_d \approx 0.90$ at -16 dB for $P_f = 0.10$. In contrast, the STFT–CNN [20] baseline degrades significantly in heavy noise.

III. Problem Formulation and Metrics

A. Spectrum Sensing Model

SS in CR networks can be expressed as a binary hypothesis test observed by a secondary user (SU) receiver. The received discrete-time signal $x[n]$ is modeled as

$$\mathcal{H}_0 : \; x[n] = w[n] \tag{1}$$

representing the case when the PU is absent, i.e. in non–transmitable and the received samples contain only noise. On the other hand, PU data transmission state is shown as hypothesis

$$\mathcal{H}_1 : \; x[n] = hs[n] + w[n] \tag{2}$$

denotes the case when the PU is present and the received samples contain both the PU signal $s[n]$ and additive noise $w[n]$. 'h' represents the channel degradation model due to mobility. The noise term may follow additive white Gaussian noise (AWGN) or generalized Gaussian noise (GGN), depending on the environmental characteristics.

From the received samples, the SU computes a feature vector $\mathbf{f}$, which may represent the power spectrum, spectrogram or raw In-phase and Quadrature (I/Q) samples. A DL classifier then outputs a binary decision

$$\hat{y} \in \{0, 1\}$$

where $\hat{y} = 1$ indicates PU presence and $\hat{y} = 0$ indicates a spectrum hole.

B. Performance Metrics

The following metrics are used to evaluate SS performance

1) Probability of Detection (P_d):

$$P_d = P(\hat{y} = 1 \mid \mathcal{H}_1) \tag{3}$$

This represents the likelihood of correctly detecting the PU presence when it transmits own data. Higher P_d values reduce collision probability between PU and SU data transmission.

2) Probability of False Alarm (P_f):

$$P_f = P(\hat{y} = 1 \mid \mathcal{H}_0) \tag{4}$$

This measures the probability of incorrectly declaring PU activity when the channel is actually free. Lower P_f enables more efficient SU access.

IV. Proposed Methodology

This section presents the proposed DCT–AE based PU spectrum data generation and augmentation framework, integrated with a light weight MLP classifier for robust SS under low SNR conditions. Algorithm 1 nicely presents the work flow of the proposed model, comprising of the following sections–

1) PU spectrum data as image representation,
2) A compact DCT–AE model,
3) PU Spectrum data augmentation and featured data generation module, and
4) A light weight MLP Classifier operating on the learned latent features.

Fig. 1 conceptually illustrates the system where the incoming PU spectrum data is first compressed using the DCT–AE, then expanded by the augmentation module to generate additional data with feature space. Finally, such feature data is fed into the MLP, which uses the latent features to classify whether the PU is transmitting or not.

A. Spectrum as an Image

In the proposed method, the SU receiver checks over a coverage zone, whether PU transmits data over its specific band or not. To accomplish this task, over each sensing period, the SU measures the received signal energy on 2D grid points. These energy values are then mapped as a matrix, where each grid cell shows the PU energy at that point, which can be thought of as a spectrum energy pixel.

This grid forms an image like view of the spectrum. Using this image helps the system learn useful patterns of PU activity more effectively than using only a single energy value. Given a sequence of spectral measurements, we reshape the data into a 32×32 two dimensional patch

$$S \in \mathbb{R}^{32\times 32}. \tag{5}$$

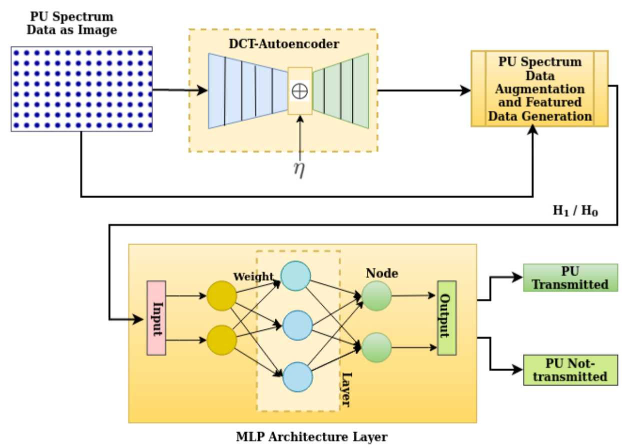

Fig. 1. Architecture of DCT–AE with MLP SS Classifier

This spatial representation enables the model to capture local and global energy patterns related to PU activity. To stabilize subsequent learning, each image is normalized as

$$S \leftarrow \frac{S - \mu}{\sigma}. \tag{6}$$

B. Compact DCT–AE Model

The normalized spectrum image is transformed into the DCT domain

$$C = \mathrm{DCT}(S), \tag{7}$$

and then vectorized as

$$x = \mathrm{vec}(C). \tag{8}$$

A shallow encoder maps the DCT vector into a compact latent representation

$$z = \mathrm{ReLU}(W_e x + b_e), \tag{9}$$

with a fixed latent dimension of 32.

The decoder reconstructs the DCT image as

$$\hat{x} = W_d z + b_d, \tag{10}$$

which is reshaped to $\hat{C}$ and transformed back into the spatial domain as

$$\hat{S} = \mathrm{IDCT}(\hat{C}). \tag{11}$$

The reconstruction loss is computed as

$$\mathcal{L}_{\text{rec}} = \|S - \hat{S}\|_2^2. \tag{12}$$

When performing the joint training with the classifier, the total loss becomes

$$\mathcal{L} = \mathcal{L}_{\text{rec}} + \lambda \mathcal{L}_{\text{clf}}. \tag{13}$$

C. PU Spectrum Data Augmentation and Featured Data Generation

To increase robustness against noise, interference variations, and low SNR conditions, spectrum augmentation is applied directly. Several augmentation operations are performed on C.

1) Multiplicative Jitter:

$$C \leftarrow C \odot (1 + \epsilon_j), \qquad \epsilon_j \sim \mathcal{N}(0, \sigma_j^2). \tag{14}$$

2) Coefficient Dropout:

$$C \leftarrow C \odot M, \qquad M_{ij} \sim \mathrm{Bernoulli}(1 - p_d). \tag{15}$$

3) Laplacian Noise Injection:

$$C \leftarrow C + \epsilon_\ell, \qquad \epsilon_\ell \sim \mathrm{Laplace}(0, b). \tag{16}$$

4) High Frequency Smoothing or Permutation: High-frequency components may be permuted or smoothed to improve invariance while preserving the dominant spectral features. These augmentations expand the training distribution and reduce overfitting to specific noise patterns.

D. Light weight MLP classifier for robust SS

The latent vector z obtained from the encoder is fed into a light weight MLP consisting of a single hidden layer of 64 ReLU neurons

$$h = \mathrm{ReLU}(W_1 z + b_1). \tag{17}$$

The classifier outputs the probability of PU presence/ absence

$$p = \sigma(W_2 h + b_2). \tag{18}$$

A binary decision is then made using a threshold τ

$$\hat{y} = \begin{cases} 1, & p \geq \tau, \\ 0, & p < \tau, \end{cases} \tag{19}$$

where τ is selected to satisfy the target false-alarm rate (e.g., $P_f = 0.10$). The corresponding detection probability P_d is evaluated across a range of SNR levels. Algorithm 1 shows DCT–AE spectrum data augmentation and MLP based decision.

V. Dataset Preparation

To show realistic performance results, we describe here a standard controlled simulation protocol suitable for DL based Spectrum data augmentation and SS decision.

A. Dataset Description

The dataset in this work is fully synthetic and generated inside the script. Each sample is a 32×32 spectrum image created by a custom generator, corresponding to either H_1 (PU present) or H_0 (PU absent). Images contain Gaussian spatial patterns with unit energy and are corrupted by AWGN using a random training SNR from $[-18, -2]$ dB.

a) Training Set: From 3000 samples, 70% are used to train the DCT–AE and the MLP classifier, with balanced H_1/ H_0 classes.

b) Validation Set: 15% of the data is used to track AE convergence and to set the threshold for $P_f \approx 0.1$.

c) Test Set: The final 15% forms the test set, used to compute ROC curves, AUC, and reconstruction metrics.

Algorithm 1: DCT–AE for PU Spectrum Data Generation and Augmentation + MLP Classifier

Require: Training set $\mathcal{D} = \{r^{(i)}, y^{(i)}\}$, val. set $\mathcal{V}$, latent dim d_z, augment params $\{\sigma_j, p_d, b\}$, joint weight λ, epochs E
Ensure: Trained encoder/decoder θ_e, θ_d, classifier θ_c, threshold τ
Preprocess: for each $r^{(i)}$ form $S^{(i)} \in \mathbb{R}^{32\times 32}$ and normalize
$C^{(i)} \leftarrow \mathrm{DCT}(S^{(i)})$, $x^{(i)} \leftarrow \mathrm{vec}(C^{(i)})$
Initialize $\theta_e, \theta_d, \theta_c$
for epoch $= 1$ to E **do**
  Shuffle $\mathcal{D}$, form mini-batches
  for each batch **do**
    for each (x, y) in batch **do**
      Apply DCT augmentations to x (multiplicative jitter, dropout, Laplace noise, optional HF perturbation) $\rightarrow x_{\text{aug}}$
      $z \leftarrow \mathrm{ReLU}(W_e x_{\text{aug}} + b_e)$
      $\hat{x} \leftarrow W_d z + b_d, \quad \hat{S} \leftarrow \mathrm{IDCT}(\mathrm{mat}(\hat{x}))$
      $\ell_{\text{rec}} \leftarrow \|S - \hat{S}\|_2^2$
      $p \leftarrow \sigma(W_2\ \mathrm{ReLU}(W_1 z + b_1) + b_2)$, $\ell_{\text{clf}} \leftarrow \mathrm{CE}(p, y)$
      $\ell \leftarrow \ell_{\text{rec}} + \lambda \ell_{\text{clf}}$; accumulate and update parameters
    end for
  end for
  Optionally validate on $\mathcal{V}$ (early stop / LR schedule)
end for
If classifier trained separately: encode all $x^{(i)}$ to $z^{(i)}$ and train MLP on $\{z^{(i)}, y^{(i)}\}$
Select threshold τ on $\mathcal{V}$ to meet target false-alarm P_f
return $\theta_e, \theta_d, \theta_c, \tau$

d) Evaluation Set (P_d–SNR): For SNR values in $\{-20, -18, \ldots, 0\}$ dB, an extra dataset of 300 samples (150 per class) is generated to estimate P_d and P_f without overlapping training data.

e) Baselines: Two baseline models are used:

- TFCFN [19], combining GRU temporal features and CNN frequency features via cross–attention.
- A deep STFT–CNN [20] classifier, serving as a conventional SS spectrogram baseline.

VI. Results and Discussion

A. Quantitative Comparison

Simulations on SS are performed with prior probability of $P(H_0) = 0.6$. As shown in Fig. 2, the training loss decreases from 0.7047 to 0.0719 and the validation loss from 0.1142 to 0.0255 within the first 20 epochs, after which both curves stabilize, indicating good convergence and generalization.

At $P_f = 0.10$, Table I shows that the proposed DCT-AE achieves higher P_d value than both TFCFN [19] and STFT-

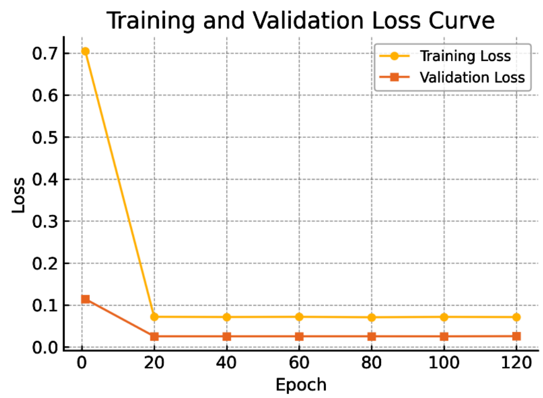

Fig. 2. Training and Validation Loss Graph for Selected Epochs in the proposed DCT–AE–MLP model

TABLE I
DETECTION PERFORMANCE AT $P_f = 0.10$

Method	P_d @ -18 dB	-16 dB	-12 dB	AUC	Acc.	Inf. (ms)
TFCFN [19]	0.78	0.91	0.97	0.95	0.92	18
STFT-CNN [20]	0.70	0.86	0.95	0.92	0.89	12
DCT-AE-MLP (Proposed)	**0.84**	**0.93**	**0.98**	**0.96**	**0.94**	14

CNN [20], especially at very low SNRs such as -18 dB, where the baseline models degrade sharply.

This model achieved $P_d \approx 0.98$ at $P_f = 0.05$, indicating superior performance as shown in Fig. 3. As P_f increases to 0.3, DCT–AE–MLP maintains $P_d \approx 0.90$, outperforming TFCFN (≈ 0.88) and STFT–CNN (≈ 0.86). Near $P_f = 0.7$, DCT–AE still achieves $P_d \approx 0.72$, whereas TFCFN and STFT–CNN drop below 0.70. Across the full P_f range, the proposed DCT–AE consistently yields the highest P_d, indicating superior feature discrimination and improved detection sensitivity relative to existing deep models.

This proposed DCT–AE demonstrates improved performance, especially at low SNR as shown in Fig. 4. The P_d–SNR curve shows degradation under low SNR conditions. At 0 dB, all models achieve $P_d \approx 0.99$. At -10 dB, DCT–AE maintains $P_d \approx 0.97$, outperforming TFCFN (0.96–0.93) and STFT–CNN (0.95–0.92). At -15 dB, DCT–AE achieves $P_d \approx 0.93$, exceeding TFCFN (≈ 0.91) and STFT–CNN (≈ 0.86). Even at -20 dB, DCT–AE sustains $P_d \approx 0.84$, significantly higher than TFCFN (≈ 0.78) and STFT–CNN (≈ 0.70). At an SNR of -12 dB and a prior probability of $P(H_0) = 0.7$, the simulation yields $P_d \approx 0.96$ and $P_f \approx 0.03$ respectively. Therefore, the proposed model exibits superior robustness and noise tolerance on all examined SNR range.

B. Effect of DCT–AE Augmentation

To assess the impact of DCT–AE augmentation, three model variants were tested:

- No augmentation: $P_d \approx 0.74$ at -18 dB
- Time–domain Gaussian augmentation: $P_d \approx 0.78$ at -18 dB
- Proposed DCT–AE augmentation: $P_d \approx 0.84$ at -18 dB

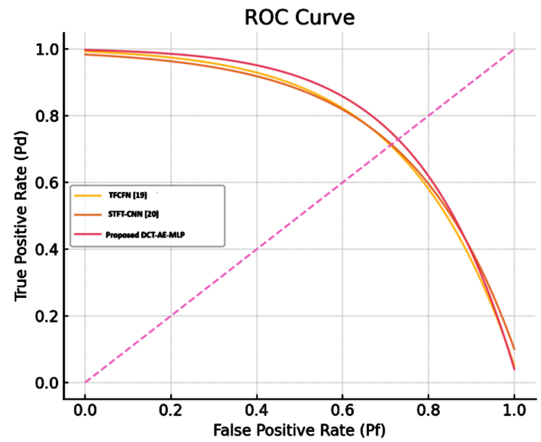

Fig. 3. ROC curves comparing TFCFN [19], STFT–CNN [20] baseline, and the proposed DCT–AE–MLP model

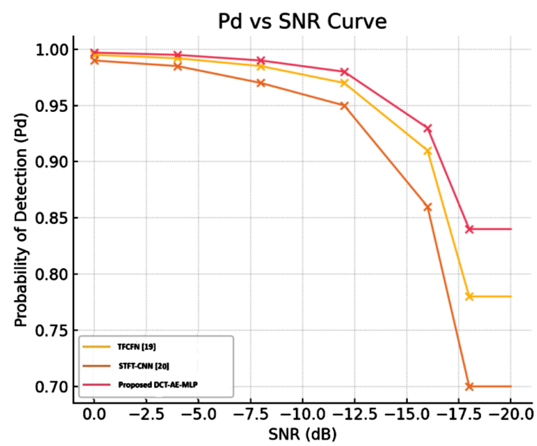

Fig. 4. Probability of Detection (P_d) Vs SNR curves comparing TFCFN [19], STFT–CNN [20] baseline, and the proposed DCT–AE–MLP model

These results indicate that DCT–AE based augmentation offers the highest gain, likely because perturbing transform coefficients simulate realistic spectral variations and channel distortions.

C. Computational Analysis

To support the detection results in Table I for $P_f \approx 0.10$, we compare the computational cost of TFCFN [19] and STFT–CNN [20], with the proposed DCT–AE–MLP architecture. In the STFT–CNN model, each STFT is computed over frames of length W, with a hop size S, for a total of N input samples. This results in approximately $\frac{N}{S}$ frames, and since each frame requires an FFT of size W, the overall STFT computational cost becomes $\mathcal{O}\big(\frac{N}{S} W \log W\big)$, followed by a 2-D CNN whose cost is dominated by convolution operations. TFCFN further increases complexity by using multiple time–frequency branches, causing its Floating-Point Operations (FLOPs) to scale by the number of branches B, and exceeding the cost of STFT–CNN. In contrast, the proposed DCT–AE–MLP a better substitute of the above with a complexity $\mathcal{O}(N \log N)$ and processes a reduced feature vector through lightweight fully connected layers, resulting a smaller computational load. Using STFT-CNN as a baseline (relative FLOPs = 1), TFCFN requires about 1.4–$1.6\times$ more operations, while the proposed method operates at 0.5–$0.7\times$ the cost of STFT–CNN, making it the most efficient of the three.

VII. Conclusion and Future Work

The present work proposes a DCT–AE PU spectrum data generation/ augmentation strategy combined with MLP architecture for robust SS decision. The current evaluation is based on synthetic data and additive white Gaussian noise, and the method as a lightweight baseline rather than a deployment-ready solution. Here, the detection probability (P_d) improves as the SNR increases. At very low SNR, such as $-18\,\text{dB}$, the proposed DCT–AE–MLP maintains high performance with $P_d \approx 0.84$. As the SNR rises to $-10\,\text{dB}$, P_d increases to approximately 0.97, and at $0\,\text{dB}$ the system achieves $P_d \approx 0.99$. These results show that the DCT–AE–MLP remains robust in heavy noise, provides high detection accuracy and computationally efficient across the entire SNR range.

In future, the proposed DCT–AE–MLP model can be combined with a GAN based module for further improvement on data augmentation and the MLP classifier can be replaced with a multi–kernel CNN model to enrich feature sets leading to improved SS performance and extend the framework to realistic interference scenarios, including co-channel interference, and investigate scalability across varying bandwidths and dynamic PU activity.

References

[1] A. Patil, S. Iyer, O. L. A. López, R. J. Pandya, K. Pai, A. Kalla, and R. Kallimani, "A Comprehensive Survey on Spectrum Sharing Techniques for 5G/B5G Intelligent Wireless Networks: Opportunities, Challenges and Future Research Directions," *Computer Networks*, vol. 223, p. 110697, 2024.

[2] M. S. Miah, M. Schukat, and E. Barrett, "A throughput analysis of an energy-efficient spectrum sensing scheme for the cognitive radio–based Internet of Things," *EURASIP Journal on Wireless Communications and Networking*, vol. 2021, no. 1, p. 201, 2021.

[3] A. Paul, S. Das, and S. P. Maity, "Spectrum prediction: Boosting D2D communications in CRNs using POMDP," *Physical Communication*, vol. 71, p. 102704, 2025.

[4] A. Paul and S. P. Maity, "Outage analysis in cognitive radio networks with energy harvesting and Q-Routing," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 6, pp. 6755–6765, 2020.

[5] S. P. Maity, S. Chatterjee, and T. Acharya, "On optimal fuzzy c-means clustering for energy efficient cooperative spectrum sensing in cognitive radio networks," *Digital Signal Processing*, vol. 49, pp. 104–115, 2016.

[6] N. Chaudhary *et al.*, "A Comprehensive Review on Spectrum Sensing Techniques including Energy Detection for Cognitive Radio Networks," *Engineering Review*, vol. 42, no. 3, pp. 1677–1693, 2022.

[7] Y. Liu, Z. Zhong, G. Wang, and D. Hu, "Cyclostationary Detection Based Spectrum Sensing for Cognitive Radio Networks," *Journal of Communications*, vol. 10, no. 1, pp. 1–10, 2015.

[8] A. Kumar, J. Venkatesh, N. Gaur *et al.*, "Cyclostationary and Energy Detection Spectrum Sensing Beyond 5G Waveforms," *Electronic Research Archive*, vol. 31, no. 6, pp. 3400–3416, 2023.

[9] M. U. Muzaffar, A. Rana, A. Yousaf, A. R. Alhammadi, A. A. Khan, and R. Sharqi, "A Review of Spectrum Sensing in Modern Cognitive Radio Networks," *Telecommunication Systems*, vol. 85, no. 2, pp. 347–363, Feb 2024.

[10] Y. Zeng and Y.-C. Liang, "Spectrum-Sensing Algorithms for Cognitive Radio Based on Statistical Covariances," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 4, pp. 1804–1815, 2009.

[11] T. Yucek and H. Arslan, "A Survey of Spectrum Sensing Algorithms for Cognitive Radio Applications," *IEEE Communications Surveys and Tutorials*, vol. 11, no. 1, pp. 116–130, 2009.

[12] A. Mehrabian, M. Sabbaghian, and H. Yanikomeroglu, "CNN-based Detector for Spectrum Sensing with General Noise Models," *IEEE Transactions on Wireless Communications*, vol. 22, no. 2, pp. 1235–1249, 2022.

[13] O. Naparstek and K. Cohen, "Deep Multi-User Reinforcement Learning for Dynamic Spectrum Access in Multichannel Wireless Networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 10, pp. 6650–6663, 2018.

[14] X. Zhao, Y. He, and G. Cao, "Autoencoder-Based Denoising and Feature Learning for Low-SNR Sensing," *IEEE Communications Letters*, vol. 26, no. 3, pp. 339–343, 2022.

[15] J. Yuan, H. Zhang, and Z. Tian, "Transformer-based wideband sensing with channel-aware attention," *IEEE Transactions on Wireless Communications*, 2023.

[16] W. Zhang, Y. Wang, X. Chen, and Z. Tian, "Spectrum Transformer: Wideband Spectrum Sensing Using Multi-Head Self-Attention," in *Proc. IEEE Int. Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2023, pp. 101–105.

[17] S. E. Abdelbaset, M. Y. Hassan, A. A. El-Mowafy, M. A. Elhosseini, and A. R. Emara, "Deep learning-based spectrum sensing for cognitive radio applications," *Sensors*, vol. 24, no. 24, p. 7907, 2024.

[18] K. Berahmand, F. Daneshfar, E. S. Salehi, A. Hassani, and M. P. Fotouhi, "Autoencoders and their applications in machine learning: A survey," *Artificial Intelligence Review*, vol. 57, p. 28, 2024.

[19] H. Xi, W. Guo, Y. Yang, R. Yuan, and H. Ma, "Cross-attention mechanism-based spectrum sensing in generalized Gaussian noise," *Scientific Reports*, vol. 14, p. 23261, 2024.

[20] Z. Chen, Y.-Q. Xu, H. Wang, and D. Guo, "Deep STFT-CNN for Spectrum Sensing in Cognitive Radio," *IEEE Communications Letters*, vol. 25, no. 3, pp. 864–868, Mar 2021. [Online]. Available: http://dx.doi.org/10.1109/lcomm.2020.3037273

[21] A. Kumar, "Spectrum Sensing Beyond 5G System: Deep Learning and Conventional Techniques Analysis," *Multimedia Tools and Applications*, 2025. [Online]. Available: https://doi.org/10.1007/s11042-025-20638-z

[22] S. Peng, L. Zhang, and J. Zhou, "CNN-based spectrum sensing with noisy and corrupted RF environments," *IEEE Access*, vol. 11, pp. 12 540–12 552, 2023.

[23] Y. Liu, Q. Li, and X. Zhang, "LSTM-enhanced temporal feature extraction for dynamic spectrum access," *IEEE Internet of Things Journal*, vol. 10, no. 5, pp. 4521–4535, 2023.

[24] X. Tian, Y. Liu, and H. Wang, "A Deep Convolutional Autoencoder–Enabled Channel Estimation Method in Intelligent Wireless Communication Systems," *Wireless Communications and Mobile Computing*, vol. 2024, pp. 1–12, 2024.

[25] I. Eddahmani, C.-H. Pham, T. Napoléon, I. Badoc, J.-R. Fouefack, and M. El-Bouz, "Unsupervised Learning of Disentangled Representation via Auto-Encoding: A Survey," *Sensors*, vol. 23, no. 4, p. 2362, 2023.

[26] A. Mumuni, F. Mumuni, and N. K. Gerrar, "A Survey of Synthetic Data Augmentation Methods in Machine Vision," *Machine Intelligence Research*, vol. 21, no. 5, pp. 831–869, 2024.

[27] W. H. I. Clark, S. C. Hauser, W. C. Headley, and A. J. Michaels, "Training Data Augmentation for Deep Learning RF Systems," *Journal of Defense Modeling and Simulation: Applications, Methodology, Technology*, vol. 18, no. 3, pp. 253–262, 2021.

[28] S. T. Nguyen and R. Murugan, "Autoencoder-based RF feature learning for spectrum sensing: Methods and applications," *IEEE Communications Letters*, vol. 28, no. 2, pp. 250–254, Feb. 2024.

[29] X. Li, Z. Zhao, Y. Zhang, S. Zheng, and S. Dai, "Spectrum Sensing Algorithm Based on Self-Supervised Contrast Learning," *Electronics*, vol. 12, no. 6, p. 1317, 2023.

[30] L. Li, Z. Zhang, and L. Yang, "Influence of Autoencoder-Based Data Augmentation on Deep Learning-Based Wireless Communication," *IEEE Wireless Communications Letters*, vol. 10, no. 9, pp. 2090–2093, 2021.