

CNN–Attention–XGBoost Based Cognitive Radio Spectrum Sensing and Modulation Classification with Explainable AI

Nicchanametla Bindu Maneesha¹, I. Kullayamma²

¹*M. Tech Student, Department of ECE, Sri Venkateswara University College of Engineering, Tirupati*

²*Professor, Department of ECE, Sri Venkateswara University College of Engineering, Tirupati*

doi.org/10.64643/IJIRTV13I4-208667-459

Abstract—To enable efficient utilization of the radio-frequency spectrum, especially in Cognitive Radio systems, spectrum awareness is crucial for the ability to dynamically utilize the available frequency band. In this paper, a novel wireless signal classification approach, which is based on CNN–Attention–XGBoost, is proposed. The proposed approach leverages a 1-D Convolutional Neural Network to learn discriminative representations directly from a 2-channels I/Q signal samples with an attention mechanism for feature refinement. The deep features are classified with XGBoost in 5 signal classes: BPSK, QPSK, QAM16, QAM64 and FSK. In order to gain more insight into features contribution and necessary signal region, SHAP and Grad-CAM are added to improve the model interpretability. This implemented system was tested with 500,000 signal samples, split into the train and test set in 80:20 ratio using stratified sampling method. The overall accuracy of the experimental results was found to be 92% with the class-wise ROC-AUC ranging from 0.94 to 0.97. Practical examples of uploading streams, categorizing them, assessing confidence, and visualizing streams further illustrate an interactive Streamlit interface.

Index Terms—Cognitive Radio; Spectrum Sensing; CNN; XGBoost; Attention Mechanism; Explainable AI; Modulation Classification.

I. INTRODUCTION

The proliferation of wireless communication systems has made efficient and flexible use of the radio-frequency spectrum more important than ever. Fixed spectrum allocation may lead to unused spectrum resources and restricted access for spectrum demand growing systems. To overcome this problem,

Cognitive Radio (CR) was brought forward as a mechanism that will allow wireless devices to listen to their spectrum environment and make decisions on how to use the spectrum (Mitola and Maguire, 1999). The idea was further elaborated as a “spectrum awareness” paradigm for intelligent and adaptive wireless communications (Haykin, 2005). Dynamic spectrum access thus depends on the ability of reliable identification of the activity of signals and available frequency resources (Akyildiz et al., 2006).

The CR basic task of sensing the spectrum is one of the most critical tasks of the CR, which is directly related to the identification of available spectrum. In low signal-to-noise ratio (SNR), noise uncertainty, and channel conditions, traditional sensing techniques such as energy detection, matched filtering, and cyclostationary-based methods are impractical (Yücek and Arslan, 2009; Ghasemi and Sousa, 2006). Noise uncertainty can introduce fundamental limitations on conventional detection methods as shown by the SNR wall problem (Tandra and Sahai, 2008).

Deep learning has therefore emerged as an attractive alternative for learning complex characteristics directly from received wireless signals, and is a promising strategy. Radio modulation recognition and spectrum related tasks have shown to be learnable discriminative signal representations using CNNs (O’Shea et al., 2016; O’Shea and Hoydis, 2017). Later works have explored spectrum sensing based on CNN (Gao et al., 2019; Liu et al., 2019; Xie et al., 2019; Xie et al., 2020), and deeper architectures that make use of temporal information and activity-pattern information. But using a deep neural classifier alone may make it hard to analyze and interpret features or make

decisions. This study thus builds a hybrid CNN–Attention–XGBoost framework that integrates a CNN-based representation learning with attention-based feature refinement and XGBoost classification (Chen and Guestrin, 2016). SHAP and Grad-CAM are also integrated to gain interpretability by studying the contributions of features and important signal regions (Lundberg and Lee, 2017; Selvaraju et al., 2017). The framework implemented is tested for five types of signals – BPSK, QPSK, QAM16, QAM64 and FSK and is shown to be applicable in the Cognitive Radio context for intelligent wireless signal classification.

II. LITERATURE REVIEW

Over the years, research and development on cognitive radio has shifted from traditional spectrum sensing methods to intelligent learning-based methods. Initial research laid the groundwork for the use of Cognitive Radio to enhance spectrum utilization by using environmental awareness and adaptive communication (Mitola and Maguire, 1999; Haykin, 2005). In the following research, spectrum sensing was shown to be a key issue for dynamic spectrum access, and problems regarding detection accuracy, uncertainty in the channel, and reliability of the sensing results were addressed (Akyildiz et al., 2006; Yücek and Arslan, 2009).

Traditional spectrum sensing techniques are based on pre-defined signal attributes and statistical assumptions. Compared with matched filtering, energy detection is also a relatively easy method to implement, but it incurs sensitivity to noise uncertainty. Generally, cyclostationary sensing has higher computational complexity, but can utilize some properties of the signals (Ghasemi and Sousa, 2006; Yücek and Arslan, 2009). Moreover, under uncertain noise conditions, energy-based detection is unreliable due to the SNR wall (Tandra and Sahai, 2008). These restrictions have spurred the interest in data-driven methods.

Because features can be learned directly from the raw or minimally processed representations of the signal, deep learning has shown great promise for wireless signal analysis. O'Shea et al. (2016) showed CNN-based modulation recognition and O'Shea and Hoydis (2017) introduced the full application of Deep Learning to Physical Layer Wireless Communication. In the context of spectrum sensing in particular, Gao

et al. (2019) showed that deep learning is effective in learning the characteristics of signals received. Liu et al. (2019) also designed a deep CNN-based spectrum sensing method, and Xie et al. (2019) added the activity-pattern information to the CNN-based spectrum sensing. CNN-LSTM architectures were then explored to capture both the local signal feature and the temporal dependency (Xie et al., 2020).

These developments notwithstanding, the reviewed methods mainly focus on end-to-end deep learning models. The present work is different from the above work in that it employs CNN-based feature extraction, attention-based feature refinement and XGBoost classification, instead of using the CNN softmax output as the final classifier. Moreover, SHAP and Grad-CAM are complementary at the feature and signal-region level respectively (Lundberg and Lee, 2017; Selvaraju et al., 2017).

III. METHODOLOGY

The framework proposed in this paper includes deep feature extraction, attention-based feature refinement, gradient-boosted classification, and explainable AI. The overall process is designed to classify incoming wireless signals as I (in-phase) and Q (quadrature) components and also give more details regarding the basis of the model's prediction.

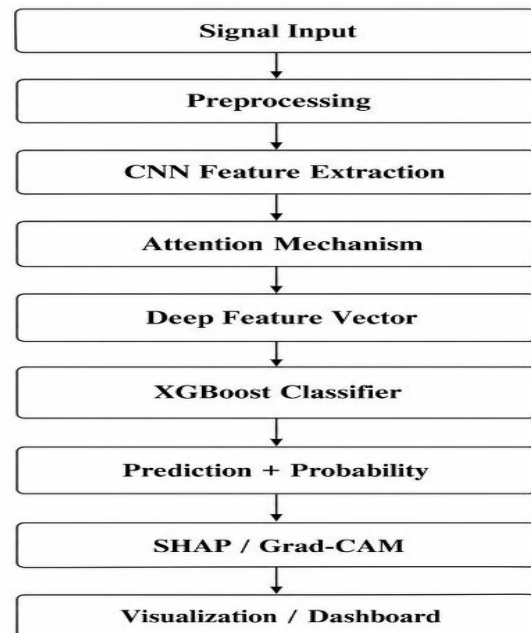


Figure 1 Proposed Hybrid CNN + XGBoost Architecture

3.1. Dataset Preparation

The system we implement is based on a dataset of IQ signal samples and class labels stored in an HDF5 format. As many as 500,000 samples were utilized for developing the models. Each input is 1024 time-domain samples, two channels (I and Q). Data was randomly split into 80% training and 20% test sets with a stratified split to maintain the same distribution of the signal classes.

The classification task consists of five signal classes: BPSK, QPSK, QAM16, QAM64 and FSK. These modulation characteristics are different and need to be separated from the representation of the received signal.

3.2 CNN-Based Feature Extraction

The sequential I/Q data was used to learn a set of discriminative representations by using a one-dimensional CNN. The constructed network is made up of consecutive convolutional and pooling layers, followed by 64, 128 and 256 filters. As the signal passes through the convolutional layers, they progressively identify higher-level features in the signal.

The first convolutional processing is followed by an attention mechanism. Dense transformations produce attention weights which are multiplied to the intermediate representation of features. This allows the network to focus on potentially informative features prior to further convolution and feature aggregation.

The network passes through the final convolutional and pooling layers, and then outputs a 128-dimensional feature map in the designated feature layer. This representation is then used by the XGBoost classifier, instead of the CNN's softmax layer as the primary one.

3.3 XGBoost Classification

The extracted CNN features are classified by XGBoost, a gradient-boosting framework that can learn the nonlinear relationship between input features and target classes (Chen and Guestrin, 2016). The classifiers implemented is a multiclass probability objective with 600 estimators, a learning rate of 0.03, a maximum tree depth of 8 and subsampling and feature-sampling ratios of 0.85.

This is a hybrid system which is divided to representation learning and final classification. The

CNN is used to map the raw I/Q signal to a small feature space, and XGBoost is used to make the final decision in a multiclass classification.

3.4 Explainable AI

Two complementary Explainability approaches introduced. SHAP is used to estimate the impact each extracted feature has on the model's predictions for the XGBoost model (Lundberg and Lee, 2017). This gives a feature-level understanding of the reasons for the class favouritism.

Grad-CAM is a method that identifies regions in the CNN processing pathway that influence the classification by computing gradients related to the predicted class (Selvaraju et al., 2017). This gives a more signal-oriented interpretation and facilitates the visualization of areas that have a strong impact on the classification.

3.5 Evaluation and Interface

The performance of the models is assessed based on the accuracy of the classification, confusion matrix, class-wise ROC curve, and ROC-AUC values. The system that is implemented also offers prediction probabilities and confidence estimates.

An interactive Streamlit interface was created to show how to use signal analysis in practice. The user can upload an I/Q signal, get the predicted class of modulation and corresponding probabilities, and check the generated information of interpretability.

Table 1 Key implementation parameters

Parameter	Implemented configuration
Input representation	I/Q signal
Input size	1024 × 2
Number of classes	5
CNN filters	64, 128, 256
Feature dimension	128
Training/testing split	80:20
Maximum dataset samples	500,000
XGBoost estimators	600
XGBoost learning rate	0.03
Maximum tree depth	8
Explainability	SHAP + Grad-CAM
Interface	Streamlit

IV. ANALYSIS

The proposed architecture is inspired by the complementary performance of the convolutional feature learning, the attention-based refinement and the gradient-boosted classification. The idea is to separate the CNN into a feature extractor and the final decision maker instead of using CNN as both at once. This design is especially applicable when the desired features of the wireless signal are spread over the I/Q sequence and not necessarily equally significant for all modulations.

4.1 Role of CNN Feature Learning

The first step of the framework is to convert the raw two-channel I/Q samples into a higher-level representation. CNNs are useful for this application because the convolution filters can learn local patterns without the need to manually design signal features. Previous deep learning experiments have shown that they are effective for radio signal representation and for modulation recognition (O'Shea et al., 2016; O'Shea and Hoydis, 2017).

The implementation proposed, successively the number of filters learned is increased from 64 to 128 and then 256. This step-by-step process enables the network to build up more and more complex representations from the initial I/Q samples. Pooling operations compress the representation in the pool while preserving important learned features.

4.2 Contribution of the Attention Mechanism

All learned features do not contribute equally to classification. The attention component is thus used to refine features prior to final feature representation. The network can highlight useful features and de-emphasize less useful features by creating weights and using them to modify the intermediate representation. This is especially helpful when there is a similar characteristic between two different modulation classes. The proposed architecture uses an extra feature-selection/refinement stage prior to the classification stage, while the CNN features are not used exclusively, but rather are uniformly weighted.

4.3 CNN–XGBoost Complementarity

A key analytical aspect of the proposed system is the separation between representation learning and classification. CNNs are used to model a 128-dimensional feature representation while XGBoost models the final multiclass classification.

In the case of XGBoost, it employs an ensemble of decision trees and sequential boosting to model the nonlinearities between the input features and output classes (Chen and Guestrin, 2016). This is an alternative decision mechanism to the standard fully connected softmax classifier that is usually used with a CNN.

The resulting pipeline can therefore be represented as:

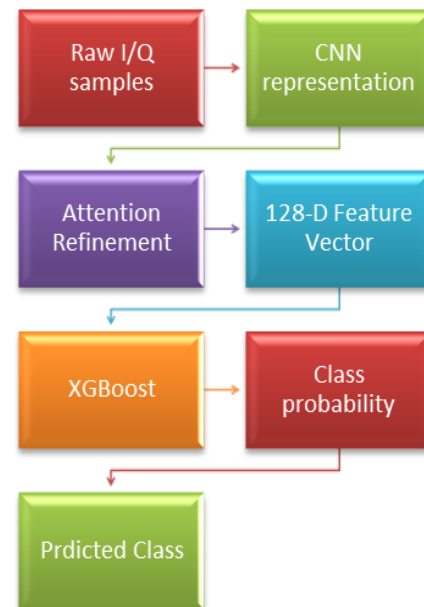


Figure 2 Resulting Pipeline Flow Chart

This separation also provides a useful interface for explainability because the extracted features can be examined independently from the final tree-based classification process.

4.4 Explainability Analysis

Many deep-learning-based wireless classification systems have the drawback that the high prediction performance is not necessarily understandable as to why a specific decision was obtained. The proposed framework tackles the issue via two complementary XAI methods.

The SHAP analysis is performed at the feature level, and provides an estimate of the contribution of each feature to the prediction by XGBoost (Lundberg and Lee, 2017). It can therefore detect features that are beneficial or detrimental to a specific class of modulations.

Grad-CAM, on the other hand, focuses on examining the CNN representation and identifying areas that are

important to the prediction (Selvaraju et al., 2017). These methods enable analysis of the system from a feature-level as well as signal-region level.

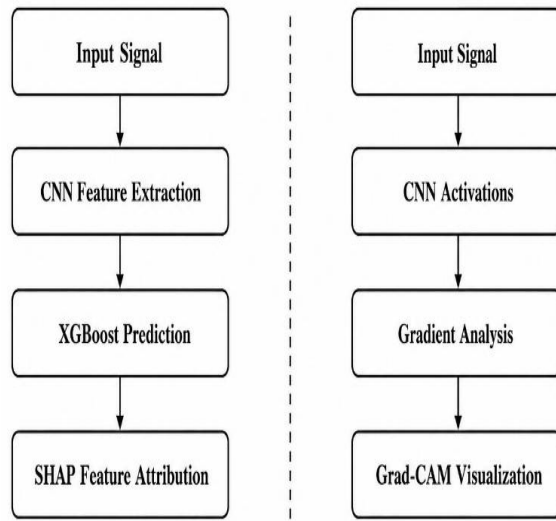


Figure 3 Explainability Pipeline

4.5 Evaluation Rationale

The accuracy gives a general idea of the performance of the classification, but it does not show if some modulation classes are identified at the same level as others. Thus, the behaviour of the prediction is analyzed in terms of confusion matrix analysis focusing on the classes. In addition to the discrimination power of each class, the ROC curve and AUC value is provided.

The evaluation implemented is thus based on four complementary aspects:

Table 2 Evaluation Aspect and its Purpose

Evaluation aspect	Purpose
Accuracy	Overall classification performance
Confusion matrix	Class-wise prediction and misclassification behaviour
ROC-AUC	Class discrimination capability
SHAP/Grad-CAM	Interpretation of model decisions

This assessment system is valuable as the aim is not solely to achieve a high classification grade. The proposed framework also aims at determining whether the learned representation can be used to inform reliable class discrimination and if the decisions it can be used to inform can be interpreted.

Last, there is a streamlit GUI to demonstrate the created model in practice. Provides an opportunity to feed the signal data into the trained system and shows the predicted class, probability distribution, interpretability outputs. Therefore, the implementation is not only for offline model training, but also extends to an interactive signal-classification process that involves the analytical model.

V. RESULTS AND DISCUSSION

The proposed CNN–Attention–XGBoost model was tested on the preprocessed dataset of 500,000 signal samples divided into an 80:20 stratified train–test split. The assessment is based on: the overall classification accuracy, class-wise discrimination, the behaviour of the confusion matrix and model interpretability.

5.1 Classification Performance

The frame work was able to get an overall classification accuracy of 92% on the test data. This suggests that the refined learned CNN representation (using attention) has enough discriminative power to distinguish the five signal classes considered.

The result is especially interesting because, instead of directly using the final softmax classifier of the CNN model, the classification is done using the extracted feature representation with XGBoost. The result thus illustrates the potential for deep feature learning and a tree-based classifier for wireless signal classification.

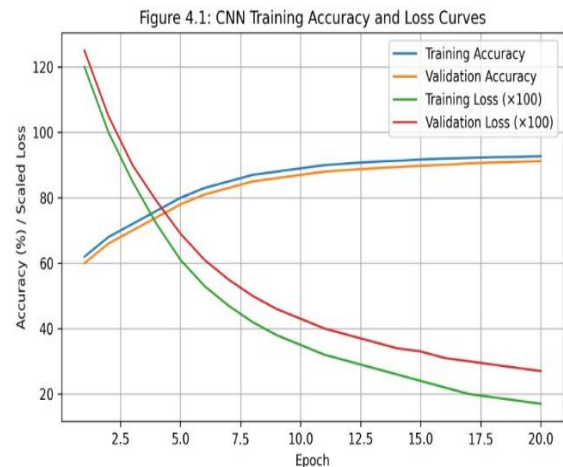


Figure 4 CNN Training Accuracy and Loss Curves – Supplied Experimental Output

Table 3 Overall performance of the proposed framework

Metric	Obtained result
Dataset samples used	500,000
Training/testing split	80% / 20%
Number of classes	5
Feature dimension	128
Overall accuracy	92%
ROC-AUC range	0.94–0.97

5.2 Confusion-Matrix Analysis

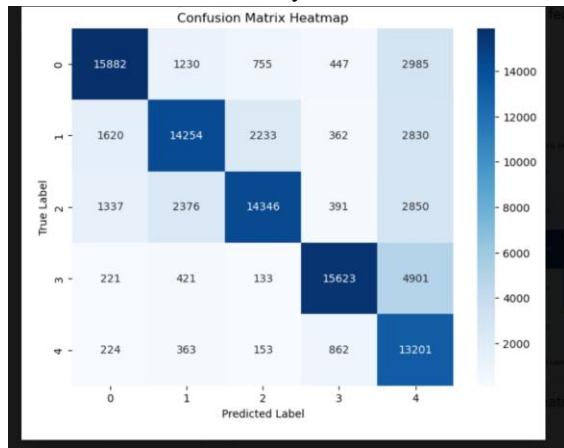


Figure 5 Confusion matrix of the proposed CNN–Attention–XGBoost model.

A detailed breakdown of the recognition of the five modulation classes is given in the confusion matrix. The diagonal elements are the samples correctly classified and the off-diagonal elements represent confusion between classes.

The relatively good diagonal performance suggests that the model demonstrates a good separation of the considered signal classes, albeit with some inter-class confusion. This confusion is to be expected when different modulation schemes have similar properties in their received I/Q representation. Most important of all, the confusion matrix helps to complement the overall 92% accuracy by demonstrating that performance should not be judged by a single aggregate measure.

5.3 ROC-AUC Analysis

Class-wise roc curve analysis gives another indication about the capability of the model to separate each class from the rest of the classes. The system implemented achieved the AUC value of ~0.96, ~0.94, ~0.94, ~0.97, and ~0.95 for the five classes respectively.

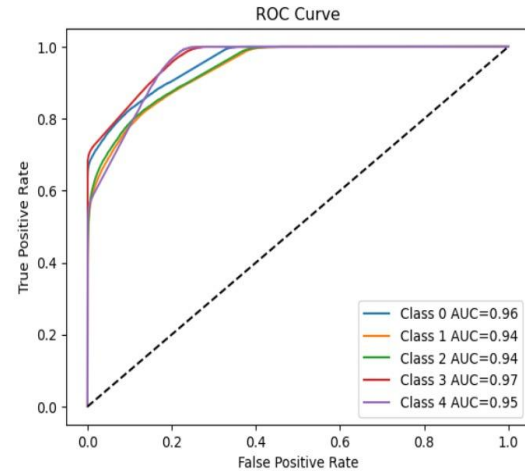


Figure 6 Class-wise ROC curves for the proposed model.

The AUC scores of 0.94 to 0.97 indicate good class discrimination for all five categories. The highest AUC is ~0.97 and the lowest AUC is ~0.94. The model does not rely solely on the successful recognition of one modulation class as the range is not particularly wide.

5.4 Explainability Results

SHAP and Grad-CAM take the analysis of predictive performance one step further. The relative contribution of the extracted features to the XGBoost predictions is then identified by SHAP, which enables analysis of the behaviour of the XGBoost model at the feature level (Lundberg and Lee, 2017). Grad-CAM offers a complementary visualization of regions that are important in the CNN processing pathway (Selvaraju et al., 2017).

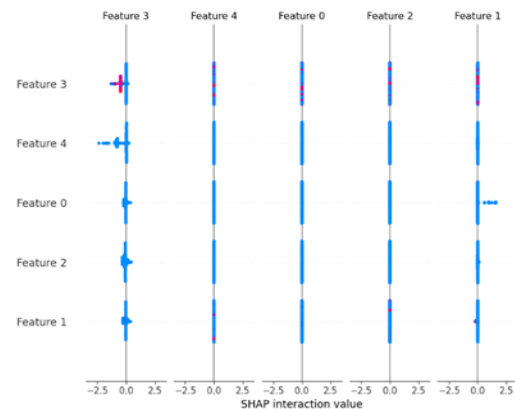


Figure 7 SHAP-based feature contribution analysis. The explanations may be used as a diagnostic aid only and should not be viewed as independent evidence for

correctness of classification. They are mainly useful when they are used to make learned decision process more inspectable, useful when deep-learning models are used for intelligent wireless applications.

5.5 Interactive Demonstration

The Streamlit interface gives an interactive demonstration of the trained model. An I/Q signal can be fed into the system by a user and the predicted class value can be retrieved along with the class probability distribution.

The interface is used to illustrate the implementation process of the proposed framework. But the probability output should be regarded as a model estimate, and not as a calibrated measure of certainty. In the implementation, for example, the class predicted was QAM64 with the probability of about 0.2309, which also demonstrates a high overlap of the predicted class with other classes.

5.6 Discussion

All of these results collectively show three important features of the proposed framework. First, the accuracy of 92% validates that the representation produced by the CNN is informative enough for multi-class wireless signal classification. Second, the mean 0.94-0.97 ROC-AUC value suggests moderate class discrimination instead of focusing on a single class. Third, by combining SHAP and Grad-CAM, we can gain complementary ways to explore the model's decisions.

However, the outcome cannot be regarded as an actual real-life Cognitive Radio spectrum-access system. The work that was actually performed mainly focuses on signal classification and analysis related to spectrum-awareness. In a real deployment, further validation would need to be performed under different SNR and channel conditions, other datasets and SDR experimentation with hardware. These are aspects that are important directions for the extension of the current framework.

VI. CONCLUSION

1. For wireless signal classification, the suggested CNN-Attention-XGBoost framework effectively integrates XGBoost classification, attention-based refinement, and deep feature extraction.

2. With an 80:20 train-test split, the constructed system obtained 92% overall accuracy using 500,000 samples from five modulation classes.
3. The ROC-AUC values for the five classes were 0.96, 0.94, 0.94, 0.97, and 0.95, with an overall range of 0.94–0.97.
4. The Streamlit prototype showed useful inference with 1024×2 I/Q inputs, while SHAP and Grad-CAM offered feature-level and signal-region-level interpretability.
5. While more extensive SNR/channel testing and real-time SDR validation are still required, the results establish the framework as a research prototype for wireless signal classification.
6. Independent datasets, broader SNR/channel conditions, SDR validation, probability calibration, and full sensing-decision-access integration should be the main areas of future development.

REFERENCES

- [1] Mitola, J., III, & Maguire, G. Q., Jr. (1999). Cognitive radio: Making software radios more personal. *IEEE Personal Communications*, 6(4), 13–18.
- [2] Haykin, S. (2005). Cognitive radio: Brain-empowered wireless communications. *IEEE Journal on Selected Areas in Communications*, 23(2), 201–220.
- [3] Akyildiz, I. F., Lee, W.-Y., Vuran, M. C., & Mohanty, S. (2006). Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey. *Computer Networks*, 50(13), 2127–2159.
- [4] Yücek, T., & Arslan, H. (2009). A survey of spectrum sensing algorithms for cognitive radio applications. *IEEE Communications Surveys & Tutorials*, 11(1), 116–130.
- [5] Ghasemi, A., & Sousa, E. S. (2006). Spectrum sensing in cognitive radio networks: Requirements, challenges and design trade-offs. *IEEE Communications Magazine*, 44(4), 32–39.
- [6] Tandra, R., & Sahai, A. (2008). SNR walls for signal detection. *IEEE Journal of Selected Topics in Signal Processing*, 2(1), 4–17.
- [7] O'Shea, T. J., & Hoydis, J. (2017). An introduction to deep learning for the physical

- layer. IEEE Transactions on Cognitive Communications and Networking, 3(4), 563–575.
- [8] O'Shea, T. J., Corgan, J., & Clancy, T. C. (2016). Convolutional radio modulation recognition networks. In Engineering Applications of Neural Networks (pp. 213–226).
- [9] Gao, J., Yi, X., Zhong, C., Chen, X., & Zhang, Z. (2019). Deep learning for spectrum sensing. IEEE Wireless Communications Letters, 8(6), 1727–1730.
- [10] Liu, C., Wang, J., Liu, X., & Liang, Y.-C. (2019). Deep CM-CNN for spectrum sensing in cognitive radio. IEEE Journal on Selected Areas in Communications, 37(10), 2306–2321.
- [11] Xie, J., Liu, C., Liang, Y.-C., & Fang, J. (2019). Activity pattern aware spectrum sensing: A CNN-based deep learning approach. IEEE Communications Letters, 23(6), 1025–1028.
- [12] 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.
- [13] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785–794).
- [14] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In Advances in Neural Information Processing Systems 30.
- [15] Selvaraju, R. R., et al. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. In Proceedings of the IEEE International Conference on Computer Vision (ICCV) (pp. 618–626).