# Performance Analysis of Convolutional Neural Networks with Different Window Functions for Automatic Modulation Classification

Ha-Khanh Le[1], Van-Phuc Hoang[1], Van-Sang Doan[2*], Minh-Thien Hoang[1], and Ngoc Phong Dao[3]

[1]Institute of System Integration, Le Quy Don Technical University, Hanoi, Vietnam
[2]Faculty of Communication and Radar, Vietnam Naval Academy, Nha Trang, Vietnam
[3]Posts and Telecommunications Institute of Technology, Vietnam
[*]Correspondence: doansang.g1@gmail.com

*Abstract*—**Automatic modulation classification is one of the main tasks for developing information security and radio signal surveillance systems. It is also becoming increasingly significant for spectrum monitoring, management, and secure communications. Recently, deep learning models have been widely applied in many fields due to their outstanding feature extraction and classification accuracy. In this paper, the automatic modulation classification performance of several deep convolutional neural networks (CNN) is analyzed on the Fast Fourier Transform (FFT) based signal spectrum and the Short Time Fourier Transform (STFT) based signal spectrogram. By using ResNet18 and MobileNet cross-combined with FFT and STFT input data, the simulation shows that the STFT data provides a higher AMC accuracy than that of FFT data. On the same STFT data, the ResNet18 model outperforms three other models (SqueezeNet, GoogleNet, and MobileNet) in classifying 26 modulation types under the influence of five levels of fading noise with SNRs from -20 dB to +18 dB. Besides, the impact of different window functions in STFT is also investigated. Numerical results indicate that the considered window functions cause an insignificant difference in the AMC accuracy.**

*Index Terms*—**Automatic modulation classification, Deep learning, Convolutional neural network, spectrum.**

## I. INTRODUCTION

Today, along with the development of science and technology, wireless communication is indispensable to people's lives and work. One of the most important technical parameters of wireless communication systems is the data transfer rate, which should be improved. As a solution, modulating the signal using different modulation schemes is expected to achieve efficient data transmission in a complex radio environment. Therefore, automatic modulation classification (AMC) is essential in a communication receiver [1]. The modulation classifier needs to provide precisely the modulation type to help the demodulator demodulate the receiving signal correctly, which is then decoded to get accurate information [2].

Traditional AMC methods can be divided into Likelihood-based (LB) and feature-based (FB). The LB methods consider the probability ratio of the received signal in modulation groups. However, the LB methods require prior knowledge of the channel parameters leading to a computational burden [3]. In contrast, the FB methods can gain less computational complexity and not depend on prior information. Nevertheless, input feature sets heavily affect the performance of FB methods. These features must be manually extracted to distinguish modulation types and may not be feasible under different disturbance conditions. Furthermore, the search for representative features requires significant consideration in terms of data.

Over the years, we have seen the rapid development of deep learning network architectures, which are widely applied in computer vision, natural language processing, and other domains. They have also gradually evolved to classify signal modulation types [4]. Scientists and researchers in wireless communications have recently used a similar approach to improving state-of-the-art signal recognition models in wireless networks. For example, O'Shea et al. [5], [6] investigated the effect of a convolution neural network (CNN) to classify radio frequency (RF) signals with in-phase and quadratic-phase (IQ) data in the time domain of the RadioML2018 dataset. Similarly, other deep learning models, such as ResNet [7], DenseNet [8] and CLDNN [6], are also demonstrated for their efficiency in signal processing. In 2021, Doan *et al.* [9], [10], [11] also studied the signal classification models, but the models are designed for the raw IQ data only.

In the studies mentioned above, most modulation classifiers use I/Q signal data to train and evaluate neural network models. The signal features are not stationary for signals that vary in time, degrading the prediction accuracy of trained networks. Therefore, it becomes more challenging for signal analysis if the signal contains many frequency components. Fortunately, the frequency components can be more easily separated if the signal is presented in the frequency domain. In addition, extracting the signal features provides more information about the signal amplitude and phase under the noise conditions [12], [13]. The AMC accuracy can be improved by converting IQ signals into frequency spectra and time-frequency spectrograms using the Fast Fourier Transform (FFT) [14] and the Short Time Fourier Transform (STFT) [15]. This paper focuses on signal pre-processing using the Fast Fourier transform and

the Short Time Fourier transform to investigate the dependency of frequency spectra on modulation classifiers.

In addition, the paper refers to the effect of using different windows and window parameters for the modulation classification performance.

Regarding the dataset issue, previous studies often use the RadioML2018 dataset [16] for AMC and achieve high signal classification efficiency. In real applications, due to the influence of the environment and weather, many phenomena further complicate the characteristics of the signal, especially multipath fading channels. Therefore, the paper uses the HisarMod2019.1 dataset published in [17] to generalize the assessment of signal classification ability. The HisarMod2019.1 dataset consists of 26 modulated signals belonging to five different modulation groups. It provides wireless signals under ideal, static, Rayleigh, Rician ($k = 3$), and Nakagami-m ($m = 2$) channel conditions.

The rest of this paper is organized as follows. Background knowledge on the system model, signal analysis, and CNN is presented in Section II. Then, Section III discusses the comparison results of classification performance between FFT-based and STFT-based input models. In this section, the performance analysis of different window functions is also presented. Finally, Section IV will conclude the research results and orientate future works.

## II. PROBLEM STATEMENT

Assume that the transmitter signal $s(n)$ of a wireless communication system travels through the channel $h(n)$ and is received by a receiver. The received signal $r(n)$ can be given as follows:

$$r(n) = s(n) * h(n) + c(n) \tag{1}$$

where $c(n)$ is the additive noise.

### A. Fast Fourier Transform

In the receiver, the signal $r(n)$ can be converted from the time domain to the frequency domain using the Fourier transform, which is a mathematical function defined as follows:

$$R(\omega) = \sum_{n=-\infty}^{+\infty} r(n)e^{-j\omega n} \tag{2}$$

where $\omega$ is the angular frequency. In the digital receiver, the frequency $\omega$ is discrete, so the discrete Fourier transform (DFT) of $r(n)$ with the length $N$ is applied as follows:

$$R(k) = \sum_{n=0}^{N-1} r(n)e^{-j\frac{2\pi}{N}kn} \tag{3}$$

### B. Short Time Fourier Transform

The Fourier transform of the signal only provides power density in the frequency domain. In reality, the frequency of a modulated signal may change over time. Accordingly, the Fourier transform seems to be unsuitable for signals whose frequency varies with time. For this reason, the Short-Time Fourier Transform is proposed to determine the signal power density in both time and frequency domains. Specifically, the principle of STFT is described as follows. A longer time signal is divided into shorter segments of equal length and then compute the Fourier transform on each shorter segment. The combination of Fourier spectra on such shorter segments generates a changing spectrum over time, known as a spectrogram or waterfall. As a result, the output of STFT is a time-frequency spectrogram. The continuous time STFT transform of the signal $r(t)$ is expressed as follows:

$$R(\tau, \omega) = \int_{-\infty}^{+\infty} r(t)\mathrm{w}(t - \tau)e^{-j\omega t}dt \tag{4}$$

where $\mathrm{w}(\tau)$ is the window function. The function of STFT is a function of Fourier transform multiple by the window function, so STFT is also called a windowed or time-dependent Fourier transform. The STFT at a time is considered the local spectrum of $r(t)$ around time $\tau$ because the relatively short window suppresses the signal out of the neighborhood. The discrete STFT transform of the signal $r(n)$ is defined as follows:

$$R(m, f) = \sum_{-\infty}^{+\infty} r(n)\mathrm{w}(n - m)e^{-j2\pi fn}, \tag{5}$$

where $w(n)$ is the discrete window function.

### C. Window functions

In this subsection, we analyze six window functions (rectangular, Hamming, Hanning, Kaiser, Blackman and Chebyshev) with different parameters to assess the modulation classification accuracy. These window functions are briefly described as follows:

*1) Rectangular window:* The Rectangular window is ideal for looking for the exact frequency of a signal and has the length required to avoid discontinuities [18]. This window is given by the following equation:

$$w_{\mathrm{R}}(n) = \begin{cases} 1 & 0 \leq n \leq N - 1 \\ 0 & \text{elsewhere} \end{cases} \tag{6}$$

*2) Hamming window:* The Hamming window is a taper formed by a raised cosine with non-zero endpoints, optimized to minimize the nearest side lobe [18]. Hamming window is defined as follows:

$$w_{Hm}(n) = \begin{cases} 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right) & 0 \leq n \leq N - 1 \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

The approximate main-lobe width equals to $8\pi/(N-1)$ and the exact main-lobe width equals to $6.27\pi/(N-1)$.

*3) Hanning window:* The Hanning window method has the advantage of reducing side lobes [18]. It outperforms other window methods, such as the rectangular window in frequency resolution. This window is given by the equation as follows:

$$w_{\mathrm{Hn}}(n) = \begin{cases} 0.5 - 0.5\cos\frac{2\pi n}{N-1} & 0 \leq n \leq N - 1 \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

The approximate main-lobe width is $8\pi/(N-1)$ while the exact main-lobe width is $5.01\pi/(N-1)$.

| Model | Depth | Trainning Parameter (Millions) | Size (MBs) |
|-------|-------|-------------------------------|------------|
| ResNet18 | 18 | 11.7 | 44 |
| GoogleNet | 18 | 1.24 | 5.2 |
| MobileNet | 53 | 3.5 | 13 |
| SqueezeNet | 19 | 1.24 | 4.6 |

*4) Kaiser window:* This window function is proposed by J. F. Kaiser [19]. The Kaiser Window has maximum attenuation based on the width of the main lobe. The coefficients of a Kaiser window are computed from the following equation:

$$
w_{\mathrm{K}}(n) = \begin{cases} \dfrac{I_0\left\{\beta\sqrt{1-[1-2n/(N-1)]^2}\right\}}{I_0(\beta)} & 0 \leq n \leq N-1 \\ 0 & \text{otherwise} \end{cases}
\tag{9}
$$

*5) Blackman window:* The Blackman window is applied to reduce estimator variance [20]. The Blackman window is defined as follows:

$$
w_B(n) = 0.42 + 0.5\cos\left(\frac{2\pi n}{L-1}\right) + 0.08\cos\left(\frac{4\pi n}{L-1}\right) \tag{10}
$$

*6) Chebyshev window:* The Chebyshev window can reduce the main-lobe width with a specific side-lobe height. The window design reduces the relative peak sidelobe height by having all the sidelobes be the same height. This is the window function generated by Harris (1978) [21]. with the following formula:

$$
\begin{aligned}
W_0(k) &= \frac{T_N\left(\beta\cos\left(\frac{\pi k}{N+1}\right)\right)}{T_N(\beta)} \\
&= \frac{T_N\left(\beta\cos\left(\frac{\pi k}{N+1}\right)\right)}{10^\alpha}, 0 \leq k \leq N
\end{aligned}
\tag{11}
$$

where $T_n(x)$ is the $n-th$ Chebyshev polynomial of the first kind evaluated in $x$, which can be computed using

$$
T_n(x) = \begin{cases} \cos\left(n\cos^{-1}(x)\right) & \text{if } -1 \leq x \leq 1 \\ \cosh\left(n\cosh^{-1}(x)\right) & \text{if } x \geq 1 \\ (-1)^n\cosh\left(n\cosh^{-1}(-x)\right) & \text{if } x \leq -1 \end{cases}
\tag{12}
$$

and $\beta = \cosh\left(\frac{1}{N}\cosh^{-1}\left(10^\alpha\right)\right)$.

### D. Model Training and Testing

Recent research on deep learning networks has focused on leveraging state-of-the-art models like ResNet18, SqueezeNet, GoogleNet, and MobileNet to improve signal classification accuracy. These networks have achieved remarkable results in many application areas. Table I reports the network parameters, including the model's size and depth. The models are applied to achieve the highest performance depending on the different research objects. The paper focuses on signal preprocessing with different models to evaluate the efficiency of signal modulation classification.

## III. RESULTS

### A. Dataset

In this study, the dataset HisarMod2019.1 [17] is used. The dataset contains 26 different modulation signals belonging to five other modulation groups and is affected by five types of fading noise with different types of noise, ideal channel circumstances, static, Rayleigh, Rician, and Nakagami-m. The dataset consists of five major modulation groups, as follows:

- Analog modulation: AM-DSB, AM-SC, AM-USB, AM-LSB, FM, PM.
- FSK modulation: 2FSK, 4FSK, 8FSK, 16FSK.
- PAM modulation: 4PAM, 8PAM, 16PAM.
- PSK modulation: BPSK, QPSK, 8PSK, 16PSK, 32PSK, 64PSK.
- QAM modulation: 4QAM, 8QAM, 16QAM, 32QAM, 64QAM, 128QAM, 256QAM.

Each modulation type comprises 1500 signals with a length of 1024 I/Q samples. For five different types of noise, the models require 780,000 signals, including 520,000 signals used for training and the remainder for testing. On the test dataset, the AMC accuracy is tested. Twenty epochs, a mini-batch size of 64, and an initial learning rate of 0.001 are used to train the models. The device used for the simulation is a computer with a 3.70 GHz CPU, 2x16GB RAM, and an NVIDIA GeForce RTX 3060ti GPU.

### B. Discussion

- *Comparison between FFT and STFT:*

In the first comparison, two networks (ResNet18 and MobileNet) are cross-combined with two preprocessing techniques (FFT and STFT) to produce four variants, including ResNet18-FFT, ResNet18-STFT, MobileNet-FFT, and MobileNet-STFT. Through these four networks, we can evaluate which kind of data (FFT or STFT) delivers higher AMC accuracy and which network (ResNet18 or MobileNet) yields better performance. Results in Fig. 1 show that ResNet18 is better than MobileNet and STFT data provides higher accuracy than FFT data. In particular, the ResNet18-STFT gains the highest accuracy with 95.5% for SNR = +8 dB. Thanks to residual blocks, ResNet18 can extract more representative information of modulations that can improve the AMC accuracy, even with FFT data.

In the second comparison, different models using STFT (including GoogleNet-STFT, ResNet18-STFT, SquezeNet-STFT and MobileNet-STFT) are taken into account to evaluate the AMC ability. Fig. 2 reveals that the ResNet18-STFT model achieves the best results in terms of accuracy. We observe that SqueezeNet-STFT and GoogleNet-STFT are lightweight models giving very low accuracies ($< 10\%$ for SqueezeNet-STFT and $< 50\%$ for GoogleNet-STFT). ResNet18-STFT and MobileNet-STFT have greater numbers of trainable parameters, so they achieve significantly higher accuracies ($> 90\%$ for SNR $\geq 5$ dB) compared to SqueezeNet-STFT and GoogleNet-STFT.

The experiment results in Fig. 3 also show that the classification efficiency for all modulation types yields a noticeable improvement when the signal-to-noise ratio range reaches 10 dB or more. Specifically, modulation types with high classification accuracy, 95% or more with signal-to-noise ratios from
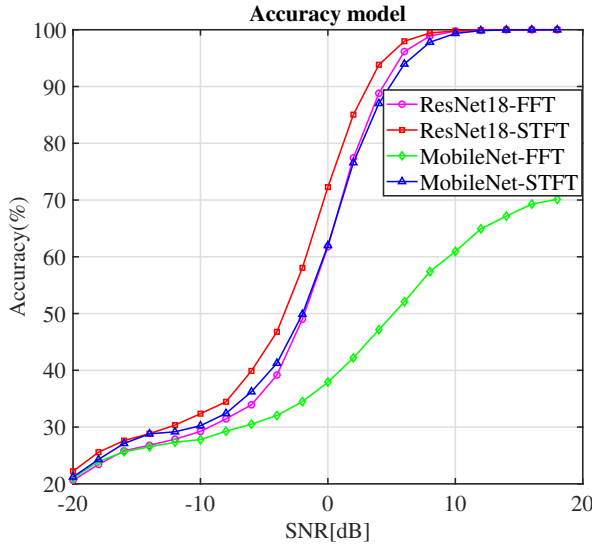
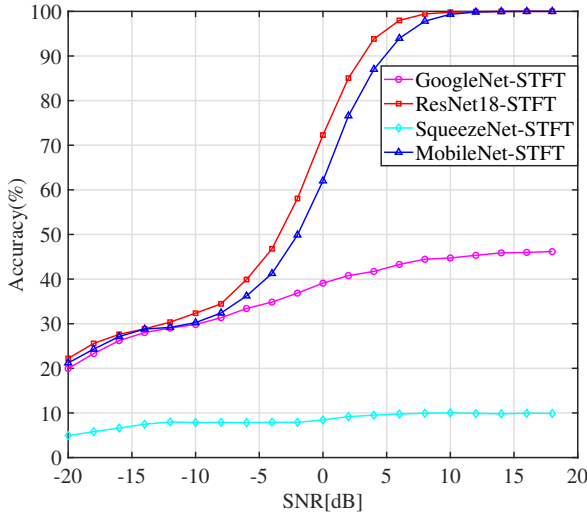Fig. 1. Comparison of the classification accuracy of different CNNs using STFT and FFT.



Fig. 2. Classification accuracy results of models using STFT.

0dB to +18 dB include AM-DSB, AM-DSB-SC, AM-USB, AM-LSB, FM, PM, 2FSK, 4FSK, 8QAM, BPSK. Modulation types 4QAM, 8FSK, 16FSK, QPSK, 4PAM give relatively high classification accuracy of 80% to 95% at SNRs from +10 dB to +18 dB. As for high-order modulations, especially at low SNRs, ResNet18-STFT has not yet achieved good classification performance.

- *Comparison between different windows:*

Next, six window functions (including rectangular, Hamming, Hanning, Kaiser, Blackman, and Chebyshev) are employed for STFT transform. ResNet18-STFT is chosen to evaluate the influence of window functions on the ACM performance. Numerical results in Table II report that changing the window functions also makes the change of AMC accu-
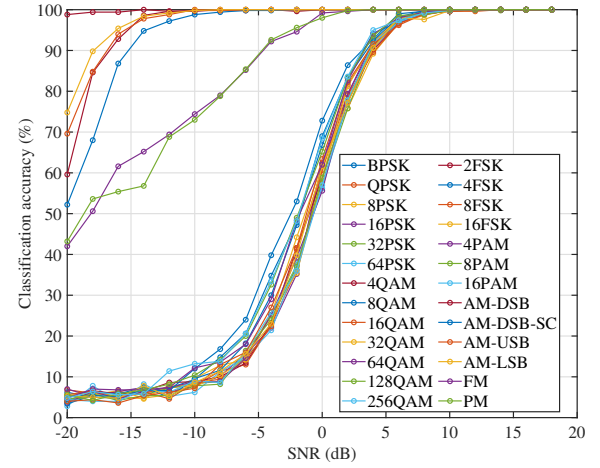


Fig. 3. Classification accuracy of ResNet18-STFT for modulation types.

TABLE II
ACCURACY OF RESNET18 WITH DIFFERENT WINDOW FUNTIONS

| Window | Accuracy | | | | |
|---|---|---|---|---|---|
| | -20dB | -4dB | 6dB | 18dB | Average Accuracy |
| Rectangular | 20.44 | 46.03 | 98.13 | 100 | **62.21** |
| Hamming | 22.22 | 46.78 | 97.97 | 100 | 64.72 |
| Hanning | 21.07 | 46.32 | 97.02 | 100 | 63.8 |
| Kaiser | 21.16 | 45.26 | 98.09 | 99.99 | 64.04 |
| Blackman | 21.08 | 48.06 | 98.22 | 99.98 | 64.86 |
| Chebyshev | 22.23 | 48.13 | 98.34 | 100 | **65.18** |

racy but is not significant. In detail, the Chebyshev window provides the highest average accuracy of 65.18%. Meanwhile, the rectangular window gives the lowest accuracy of 62.2%, and the remaining windows have roughly equal accuracy, with an average accuracy of about 64%.

Moreover, this paper compares and evaluates the AMC accuracy of ResNet18-STFT by varying the window length of the Chebyshev window and the number of overlapped samples. It can be seen in Fig. 4 that increasing the overlapped samples results in improving the AMC accuracy. In this work, ResNet18-STFT obtains the highest accuracy for 75% overlap of the window length compared to 50% and 25%.

The window length facilitates the high resolution in the frequency domain but degrades the resolution in the time domain. Therefore, choosing a suitable window length is also important to improve the AMC accuracy. Indeed, experimental results in Fig. 5 show that the classification accuracy is the lowest for the window length of 512, and the highest for the window length of 128.

## IV. CONCLUSION

The performance of automatic classification modulation using deep learning depends on many factors such as model, data, signal, and signal preprocessing. This paper has demonstrated that the signal preprocessing method and the appropriate selection of window parameters can improve the
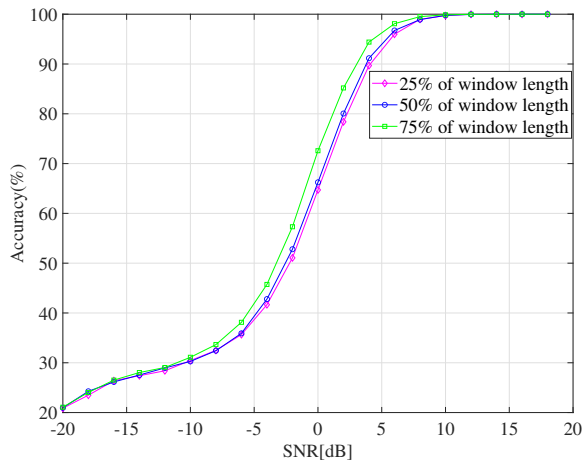
Fig. 4. Classification accuracy with different values of the overlap percentage.
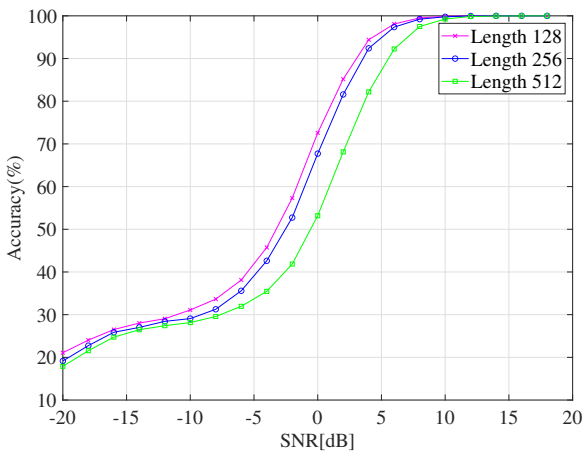


Fig. 5. Classification accuracy with different values of the window length.

accuracy of modulation classification. Specifically, simulation results have shown that the ResNet18 model combining signal preprocessing using STFT with the Chebyshev window can provide the best AMC accuracy compared to other setups considered in this paper. In the future, we will develop signal preprocessing methods with lightweight models, which can guarantee high speed and high AMC accuracy. Furthermore, the experimental measurements will be performed to verify the proposed method.

### ACKNOWLEDGMENT

### REFERENCES

[1] T. Huynh-The, C. -H. Hua, Q. -V. Pham and D. -S. Kim, "MCNet: An Efficient CNN Architecture for Robust Automatic Modulation Classification," in IEEE Communications Letters, vol. 24, no. 4, pp. 811-815, April 2020, doi: 10.1109/LCOMM.2020.2968030.

[2] J. Ma, S.-C. Lin, H. Gao, and T. Qiu, "Automatic modulation classification under non-Gaussian noise: A deep residual learning approach," *Proc. IEEE Int. Conf. Commun. (ICC)*, pp. 1–6, May. 2019.

[3] J. L. Xu, W. Su, and M. Zhou, "Likelihood-ratio approaches to automatic modulation classification," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 41, no. 4, pp. 455–469, Jul. 2011.

[4] T. Huynh-The et al., "Automatic Modulation Classification: A Deep Architecture Survey," *IEEE Access*, vol. 9, pp. 142950-142971, 2021, doi: 10.1109/ACCESS.2021.3120419.

[5] T.J. O'Shea, J. Corgan, T. C. Clancy, "Convolutional Radio Modulation Recognition Networks," *Communications in Computer and Information Science*, vol. 629, pp 213–226, Jun. 2016.

[6] T.J. O'Shea, J. Corgan, T. C. Clancy, "Over-the-air deep learning based radio signal classification," *IEEE J. Sel. Topics Signal Process*, vol. 12, no. 1, pp. 168–179, 2018.

[7] J. Ma, S.-C. Lin, H. Gao, and T. Qiu, "Automatic modulation classification under non-Gaussian noise: A deep residual learning approach," *Proc. IEEE Int. Conf. Commun. (ICC)*, pp. 1–6, May. 2019.

[8] G. Huang, Z. Liu, and K. Q. Weinberger, "Densely connected convolutional networks," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2261-2269, Nov. 2017.

[9] T. Huynh-The, C. H. Hua, V. S. Doan, and D. S. Kim, "Accurate modulation classification with reusable-feature convolutional neural network," *Proc. 2020 IEEE Eighth International Conference on Communications and Electronics (ICCE)*, Phu Quoc Island, Vietnam, pp. 12–17, Jun. 2021.

[10] V. -S. Doan, T. Huynh-The, V. -P. Hoang and D. -T. Nguyen, "MoDANet: Multi-Task Deep Network for Joint Automatic Modulation Classification and Direction of Arrival Estimation," in IEEE Communications Letters, vol. 26, no. 2, pp. 335-339, Feb. 2022, doi: 10.1109/LCOMM.2021.3132018.

[11] H.-K. Le, V.-S. Doan, và V.-P. Hoang, "Ensemble of Convolution Neural Networks for Improving Automatic Modulation Classification Performance", UD-JST, vol 20, pp. 25–32, Jun 2022.

[12] S. Rajendran, W. Meert, D. Giustiniano, V. Lenders, and S. Pollin, "Deep learning models for wireless signal classification with distributed lowcost spectrum sensors," *IEEE Trans. Cognit. Commun. Netw.*, vol. 4, no. 3, pp. 433–445, Sep. 2018.

[13] S. Haykin and P. Setoodeh, "Cognitive radio networks: The spectrum supply chain paradigm," IEEE Trans. Cogn. Commun. Netw., vol. 1, no. 1, pp. 3–28, 2015.

[14] E. O. Brigham and R. E. Morrow, "The fast Fourier transform," in IEEE Spectrum, vol. 4, no. 12, pp. 63-70, Dec. 1967, doi: 10.1109/MSPEC.1967.5217220.

[15] J. Allen, "Short term spectral analysis, synthesis, and modification by discrete Fourier transform," in IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 25, no. 3, pp. 235-238, June 1977, doi: 10.1109/TASSP.1977.1162950.

[16] T. J. O'Shea and N. West, "Radio machine learning dataset generation with gnu radio," in Proceedings of the GNU Radio Conference, vol. 1, 2016.

[17] K. Tekbıyık, A. R. Ekti, A. Görçin, G. K. Kurt, C. Keçeci, "Robust and Fast Automatic Modulation Classification with CNN under Multipath Fading Channels," *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, pp. 1-6, May. 2020.

[18] Oppenheim, Alan V., Ronald W. Schafer, and John R. Buck. Discrete-Time Signal Processing. Upper Saddle River, NJ: Prentice Hall, 1999.

[19] J. Kaiser and R. Schafer, "On the use of the I0-sinh window for spectrum analysis," in IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 28, no. 1, pp. 105-107, February 1980, doi:10.1109/TASSP.1980.1163349.

[20] Sulistyaningsih et al., "Performance Comparison of Blackman, Bartlett, Hanning, and Kaiser Window for Radar Digital Signal Processing," 2019 4th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), 2019, pp. 391-394, doi: 10.1109/ICITISEE48480.2019.9003791.

[21] F. J. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," in Proceedings of the IEEE, vol. 66, no. 1, pp. 51-83, Jan. 1978, doi: 10.1109/PROC.1978.10837.