Enhanced Spectrum Sensing Techniques For 5G New Radio (NR) And Long-Term Evolution (LTE) Utilizing Unet-Based Deep Learning Network

[1] Huan Nguyen-Duy, [2] Thien Huynh-The
[1, 2]*Ho Chi Minh City University of Technology and Education (HCMUTE), Vietnam*
[1, 2]*The Intelligent Multimedia and Advanced Computing (IMAC) Laboratory*
[1]huan2931@gmail.com
[2]thienht@hcmute.edu.vn

***Abstract*—Spectrum sensing has a crucial role in the next generation new radio and long-term evolution network, allowing devices to identify and accommodate demand for wireless connectivity. In this paper, we addressed spectrum sensing challenges and introduced a innovative wireless communication network called WiComNet to tackle over-the-air communication sensing tasks. In our proposed network, which employed an architecture based on Unet, we integrated several advanced deep learning methodologies aimed at enhancing sensing performance. These included depthwise separable convolution, recurrent residual block, and atrous spatial pyramid pooling techniques, among others. Relying on simulation results, WiComNet acquired synthetic signal dataset of two wireless communication types in differenct signal-to-noise, our proposed outperformed a impressive result with high performance for spectrum sensing missions. Our implementation and pre-trained model were available at: Spectrum sensing base on deep learning Github project**

***Index Terms*—Wireless communication network (WiComNet), spectrum sensing, 5G new radio (5G NR), long-term evolution (LTE), Unet, semantic image segmentation, signal processing.**

I. Introduction

A. Spectrum sensing

Spectrum sensing is defined as identifying the availability and non-availability of wireless communication network or radio network in particular frequency bands. When frequency allocation scheme cannot provide high data rate demand constantly, it leads to waste spectrum resources that can utilize to allocate for others such 5G new radio (NR) or long-term evolution (LTE) at the same time. Therefore, spectrum sensing for 5G NR and LTE became a main topic that contained many challenges, addressing the research direction to increase effectiveness performance and utilizing limited spectrum resource. The figure 1 shows the example of spectrum sensing for 5G and LTE application.

In the past, numerous sensing techniques were showed in [**?**], using sensing algorithms, Multi-Dimensional spectrum sensing, channel estimation, and cooperative sensing. Nevertheless, spectrum sensing tasks met a big challenge in hardware requirements [**?**], this task required high sampling rates, high resolution analog to digital converters (ADC) with large dynamic ranges, and high performance of processor. On the other hand, cognitive wireless communication can be identified immediately by sensing duration and frequency. However, this method had trades off between performance and sensing algorithms. Since their introduction, researches spent efforts to tackle improved spectrum sensing problems, addressing various challenges and introducing innovative solutions to increase prediction ratio and performance in cognitive wireless communication network. Furthermore, prominent surveys on this topic were introduced in [**?**], [**?**], [**?**], and [**?**].

Fig. 1. Spectrum sensing for 5G and LTE application.

B. Related work

In recent years, many works on cognitive wireless communication and radio network were introduced to enhance sensing performance. In general, the energy detector based sensing was common way of spectrum sensing because it offered low computational and implementation complexities, waveform base sensing increased sensing performance, cyclostationarity based sensing for detecting and matching filtering primary user transmissions [**?**]. On the other hand, several innovative deep learning (DL) methods were presented to tackle spectrum sensing performance. Applying convolution neural network for improved waveform classification [**?**], the work [**?**] introduced innovative DL architecture to improve the spectrum sensing prediction ratio for 5G and LTE signals. Other works utilized

deeplabv3+ [?] and DetectNet [?] to increase accuracy for spectrum sensing task. Nevertheless, while sensing algorithms had a trades-off between hardware performance and computation complexities, DL models offered higher parameters to reach a high level of precision that required considerable hardware resources.

C. Contribution

In this work, we utilized DL methods to tackle sensing tasks and proposed a innovative DL architecture based on Unet called wireless communication network (WiComNet), which reduced significantly complexity and enhanced sensing performance for 5G and LTE signals. Unet was a fully convolutional network [?] using special for semantic segmentation images, it was designed with parallel encoder-decoder and connecting each layer by skip connections. Furthermore, Unetpp and other variant models were showed by the authors [?] that improved dramatically prediction ratio for image segmentation tasks. Nevertheless, the drawback of these models was that they offered significantly complexities and required considerable hardware resources. In term of decreasing complexities, we presented a WiComNet architecture that incorporated depthwise separable convolution [?] and recurrent residual convolution [?] [?] [?], replacing for Unet blocks in each layer of encoder and decoder paths. On the other hand, atrous pyramid spatial pooling modules [?] were adapted between encoder and decoder paths that enhanced features extraction in multiple scales. In summary, we made the following two main contributions:

1. Addressing spectrum sensing for 5G and LTE problems and utilizing DL methods to tackle spectrum sensing tasks.
2. Introduction a innovative WiComNet architecture based on Unet reduces significant complexities and enhances sensing performance for 5G and LTE signals.

II. Methodology

A. Signal processing

Nowadays, over-the-air communication systems have a vital role in diverse scenarios and applications [?]. This work mainly considers about 5G NR and LTE communication systems. In practical terms, received signals (RX signal) from different sources in figure **??** can be defined by formula (**??**), where $y(t)$ denotes for the RX signal, $x(t)$ is the transmitted signal (TX signal), the channel response $h(t)$, and $n(t)$ presents for the addition of noise.

$$y(t) = (x(t) *^{1} h(t)) + n(t) \tag{1}$$

Short-time Fourier transformer (STFT) is a widely used signal processing technique for analyzing the frequency content of a signal over time. The RX signal (**??**) can be presented in frequency domain by STFT function that is defined in (**??**). In detail, $Y(f,t)$ represents for the spectrogram of the RX signal at frequency f and time t, $y(t)$ is the input RX signal, $w(t-\tau)$ denotes for the window function using for segmenting the signal into short frames. Furthermore, the energy spectrogram is illustrated by the equation (**??**), with $E(t)$ denotes for energy at time t.

$$Y(f,t) = \int_{-\infty}^{\infty} y(\tau) \cdot w(t-\tau) \cdot e^{-j2\pi f\tau}\, d\tau \tag{2}$$

$$E(t) = \sum_{f} |Y(f,t)|^2 \tag{3}$$

The signal-to-noise (SNR) in both 5G and LTE systems can be calculated by comparing the power of the signal to the power of the noise. In wireless communication systems, SNR is typically defined as the ratio of the average received signal power to the average noise power over a specified bandwidth. In over-the-air communication systems, the signal power depends on factors such as the transmitted power, path loss, antenna gains, and fading effects. We define P_{signal} as signal power of RX 5G and LTE signals, it is demonstrated by formula (**??**) with P_{tx} is the transmitted power (TX power) and G denotes for the channel gain respectively. On the other hand, the noise power P_{noise} in (**??**) can be effected by thermal noise power P_{thermal}, atmospheric noise $P_{\text{atmospheric}}$, and powers are generated from inter-symbols P_{ISI}, co-channel P_{CCI}, adjacent channel P_{ACI} interference, respectively.

$$P_{\text{signal}} = P_{\text{tx}} \cdot G \tag{4}$$

$$P_{\text{noise}} = P_{\text{thermal}} + P_{\text{ISI}} + P_{\text{CCI}} + P_{\text{ACI}} + P_{\text{atmospheric}} \tag{5}$$

SNR formulas can be defined in (**??**) and (**??**). While SNR denotes the SNR ratio between P_{signal} and P_{noise}, SNR_{dB} expresses for SNR ratio in decibels (dB).

$$SNR = \frac{P_{\text{signal}}}{P_{\text{noise}}} \tag{6}$$

$$SNR_{\text{dB}} = 10 \cdot \log_{10}\left(\frac{P_{\text{signal}}}{P_{\text{noise}}}\right) \tag{7}$$

There are several spectrogram generation samples in figure **??** that presents for 5G NR, LTE, and overlapping 5G and LTE, respectively.

(a) (b) (c)

Fig. 2. The spectrogram of received signal samples: (a) LTE, (b) 5G NR, (c) overlapping 5G NR and LTE

[1]*: representation for convolution operation

B. WiComNet: A robust and cost-efficient convolution network

In this work, we proposed the WiComNet that was defined as a robust and cost-efficient convolution network (convnet) utilizing Unet-based architecture [**?**]. In general, this work utilized depthwise separable convolutions (DSC) [**?**] as convolution operation that replaced for traditional convolutions. DSC uses inoperative depthwise convolution and pointwise convolution to reduce significantly the number of convolution operators. The detail of DSC can be illustrated in formula (**??**). Where $F_{\text{DSC}}(f)$ denotes for the output of DSC convolution operation with the number of features extraction f , $Dwconv(3\times3,1)_{\text{BN}}$ represents for depthwise convolution and batch normalization active function, it is connected directly by $Conv(1\times1,f)_{\text{Relu}}$.

$$F_{\text{DSC}}(f) = Dwconv(3\times3,1)_{\text{BN}} * Conv(1\times1,f)_{\text{Relu}} \quad (8)$$

On the other hand, this work applied atrous spatial pyramid pooling (ASPP) modules [**?**] between encoder and decoder paths of WiComNet to enhance the ability of feature extraction in multiple scales. ASPP modules incorporate parallel four convolution layers, which include a 1×1 convolution layer and three 3×3 convolution layers in difference dilatation rates. The output of ASPP module is synthetic by a concatenation layer. The figure **??** depicts ASPP module with r_1, r_2, r_3 represent for dilatation rates of three 3×3 convolution layers.

Fig. 3. Atrous spatial pyramid pooling module

The output of ASPP module can be calculated by (**??**). Where $F_{\text{ASPP}}(f, r_1, r_2, r_3)$ illustrates for the output of ASPP module, while r_1, r_2, r_3 represents for dilatation rates, f indicates for the number of input features. Furthermore, F_{Concat} denotes for the output of concatenation operation and $Conv$ denotes for convolution layers, respectively.

$$F_{\text{ASPP}}(f, r_1, r_2, r_3) = F_{\text{Concat}}\left(Conv(1\times1, f), Conv(3\times3, f, r_1), Conv(3\times3, f, r_2), Conv(3\times3, f, r_3)\right) \quad (9)$$

Inspired by the architectural innovation of the Recurrent Residual Block (RRB) as proposed in [**?**] and [**?**], we have revolutionized our network design. Specifically, we had overhauled both the encoder and decoder paths of our network by substituting the conventional Unet convolution blocks with RRB blocks. In WiComNet architecture, the output of each RRB block is seamlessly integrated with a cascade of DSC blocks, augmented by a residual path. This augmentation significantly amplifies the richness and complexity of the feature representation captured by our network, it is showed by figure **??**. The formula (**??**) illustrates for the output of RRB block, F_{add} denotes for addition operation between both input $F_{\text{in}}(f)$ and $Conv(1\times1, f)$ from previous layer, in which $F_{\text{in}}(f)$ is the result of previous layer, $Conv(1\times1, f)$ and f are convolution layer and the number of input features respectively.

$$F_{\text{RRB}}(f) = F_{\text{add}}\left(F_{\text{in}}(f), Conv(1\times1, f)\right) \quad (10)$$

Fig. 4. Wireless communication network

TABLE I
NETWORKS COMPLEXITY COMPARISON

Network	No.Layers	No.Params
WiComNet	**125**	**7.8M**
Unet	59	31.3M
Unetp	101	42.5M
UnetE	101	42.3M
Unetpp	101	43.5M
ConvNet	100	20.6M
Deeplabv3	100	20.6M

The figure **??** depicts for WiComNet architecture that was redesigned by Unet-based architecture [**?**]. The WiComNet architecture comprised five layers each for the encoder and decoder paths, ASPP modules were integrated into skip connection paths across multiple layers to enhance feature extraction across various scales. Our proposing design offered lowest complexity comparing to original Unet architecture, and other variants. The table **??** compares complexities among our design with other architectures.

III. DATASET, HARDWARE RESOURCES, RESULT AND DISCUSSIONS

A. Dataset and training resources

In this work, considering realistic challenges (e.g., cost and privacy), we used 5G toolbox in MatLab to generate a synthetic 5G and LTE spectrum with three difference types that include 5G, LTE, overlapping 5G and LTE frames, respectively. In particular, our dataset contained 5000 samples for each type, which 128×128 image size combining with SNR range from 0 to 30 dB. We slipped dataset with 80%, 10%, and 10% for training, validation, and test respectively. In detail, the training process was executed on a state-of-the-art computing system renowned for its high-performance capabilities, featuring a formidable 3.00GHz central processing unit (CPU) combining with cutting-edge graphic processing unit (GPU) RTX 2080. The computing resources include a substantial 16GB of memory capacity. Furthermore, training option sets up with initialize learning rate 0.001, using $piecewise$ learning rate schedule respectively. Training process executed in 40 epochs and validated every 1000 iterations. For more detail information about dataset and training resources, following to table **??**.

TABLE II
HARDWARE RESOURCES AND TRAINING OPTIONS

Hardware Resources		Training Options	
CPU	3.0GHz	No. epochs	40
GPU	RTX 2080	Learning rate	0.001
Memory	16GB	Learning rate schedule	$piecewise$
MatLab version	R2023	Validation frequency	1000
Dataset Information			
Category	**No. Samples**	**Image size**	**SNR (dB)**
LTE	5000	128×128	[0 30]
5G	5000	128×128	[0 30]
5G and LTE	5000	128×128	[0 30]

B. Simulation and evaluation result

We evaluated networks in four metrics, which include global accuracy, average loss, weight intersection-over-union (IoU), and mean boundary F1 (BF) score. These metrics are widely used to evaluate prediction ratio of semantic segmentation. Following to table **??** for the detail information of our results. Weight IoU is defined in formula (**??**) as a metric commonly used in the evaluation of semantic segmentation models, where TP_c, FP_c, and FN_c denotes for the number of true positive, false positive, and false negative pixels for class c, respectively. The formula (**??**) presents for mean BF score with N is the total number of classes or objects.

$$IoU_c = \frac{TP_c}{TP_c + FP_c + FN_c} \tag{11}$$

$$\text{MeanBFScore} = \frac{1}{N}\sum_{i=1}^{N} \frac{2 \cdot \text{Precision}_i \cdot \text{Recall}_i}{\text{Precision}_i + \text{Recall}_i} \tag{12}$$

TABLE III
SIMULATION RESULT COMPARISON

Network	G.Accuracy	Avg.Loss	W.IoU	M.BFScore
WiComNet	**97.233**	**0.0713**	**94.611**	**87.32**
Unet	98.227	0.0766	96.539	90.56
Unetp	98.07	0.0754	96.245	89.206
UnetE	97.844	0.0835	95.821	88.056
Unetpp	97.981	0.0801	96.074	88.892
ConvNet	39.696	1.2742	19.328	15.612
Deeplabv3	46.266	1.1725	33.878	14.397

Fig. 5. Confusion matrix at difference SNR levels: (a) 0dB, (b) 15dB, (c) 30dB, (d) [0 80] dB

In reference to table **??** and **??**, WiComNet outperformed superior performance, achieving an impressive 97.223% global accuracy. In comparison, ConvNet and Deeplabv3 exhibited considerably lower global accuracy of 39.696% and 46.266%, respectively. Our proposal network consistently maintained a consistently high level of accuracy and performance, comparing to that of Unet-based architectures, while significantly reducing complexity. Specially, it achieved the same global

Fig. 6. Accuracy evaluation under difference noise levels

accuracy while reducing the number of parameters by approximately 80%, comparing to Unet, Unetpp, Unetp, and UnetE [**?**] [**?**]. The figure **??**, **??** present the confusion matrix generated during the evaluation of WiComNet model, alongside the training accuracy evaluation process, comparing our model with others. Following the confusion matrix, this study achieved notable prediction outcomes across four distinct classes in difference SNR ranges. It is true that the prediction performance increased corresponding with higher SNR levels, we evaluated our results using test samples spanning various SNR levels, ranging from -5dB to 30dB. Outcome in figure **??** depicted that our model acquired significant predictive accuracy across difference SNR ranges, demonstrating a robust prediction performance, particularly at higher 5dB levels. In summary, these results suggested that WiComNet substantially outperformed strongly semantic segmentation classification performance in all categories (i.e., 5G NR, LTE, noise, and unknown).

Regarding evaluation results in figure **??** using several samples in various SNR levels, this work evaluated corresponding three overlapping 5F and LTE spectrograms at 10dB, 30dB, and 60dB SNR levels. In general, the classification accuracy improved comprehensively SNR levels exceeding 30dB (e.g., the proportion of corrected prediction labels). Respectively, with a 10dB increase in noise level, the prediction of 5G and LTE labels was not relatively accurate, with an incorrect prediction in the 5G label. Nevertheless, it exhibited a significant improvement in classification results for 5G, LTE, and noise at 30dB. Therefore, the results yielded precise label predictions at higher SNR levels surpassing 30dB. Specially, achieving accurate predictions at 60dB. These consequence demonstrated that our network outperformed outstanding performance that is suitable for real-world applications.

(a)

(b)

(c)

Fig. 7. Visualization of received spectrograms in various SNR levels: (a) 10dB, (b) 30dB, (c) 60dB

IV. Conclusion

In this paper, we presented a innovative WiComNet utilizing Unet-based architecture, outperforming impressive performance in all difference SNR levels of 5G and LTE prediction spectrograms. Indeed, our study tackled performance problems when using DL for the spectrum sensing undertaking. Our proposing network reduced significantly complexity comparing to Unet-based architectures and other variants, utilizing cutting-edge DL techniques to create robust and cost efficient network. WiComNet attained classification robustness in simulation on 5G and LTE spectrograms dataset, combining with high proportion of addition of noise, making it suitable for real-world conditions.