2024 39th Conference on Design of Circuits and Integrated Systems (DCIS) | 979-8-3503-6439-2/24/$31.00 ©2024 IEEE | DOI: 10.1109/DCIS62603.2024.10769191

Increasing the Accuracy of Spectrogram-based Spectrum Sensing Trained by a Deep Learning Network Using a Resnet-18 Model

A. Rojas*, G. Jovanovic Dolecek*, J.M. de la Rosa† and G. Liñán-Cembrano†,
*Dept. of Electronics, National Institute of Astrophysics, Optics, and Electronics (INAOE), Puebla, Mexico.
†Instituto de Microelectrónica de Sevilla, IMSE-CNM (CSIC/Universidad de Sevilla), Sevilla, Spain.

***Abstract*—This paper presents the evaluation of an image restoration approach based on a bilateral Gaussian filter (BGF) to address signal identification for spectrum sensing. The proposed methodology uses synthetic Long-Term Evolution (LTE) and 5G New Radio (NR) signals generated in MATLAB to build two spectrogram datasets. A Resnet-18 model has been trained with and without enhanced spectrograms to show the benefits of spectrogram preprocessing. Several case studies are considered using noisy and restored spectrograms. A noise recognition of 84.81% is obtained, showing the benefits of the proposed approach for spectrum sensing[1].**

I. Introduction

Spectrum sensing (SS) is one of the most important tasks for cognitive radio (CR) [1] because of its capability of gathering information from the electromagnetic environment allowing CR systems to detect the best band without interfering with the operation of primary users (PU) [2]. SS can be applied over the broadband spectrum composed of multiple Radio Frequency (RF) signals. Detecting them is important for many applications including RF spectrum management, signal intelligence, electronic warfare, new-generation telecommunication network solutions, etc [3]. RF signals in the time domain can be transformed into a spectral image in the frequency domain, redirecting the problem to the image processing area.

In the literature only a few works use this approach such as [4] which defines bounding boxes to locate frequency-hopping heterogeneous signals over spectrograms. A survey of the current literature regarding spectrum sensing using deep learning (DL) and image processing is presented in [5] showing the necessity of more research that combines image processing and DL techniques for spectrum sensing. To the best of our knowledge, this is the first time that an image restoration approach such as the bilateral Gaussian filter (BGF) is proposed as preprocessing for spectrum sensing systems based on spectrogram images.

In this work, we focus on fifth-generation (5G) New Radio (NR) and Long-Term Evolution (LTE) signals present in the same wideband spectrogram in next-generation wireless networks [6]. Applying the Fast Fourier Transform (FFT) to the combination of these modulated signals facilitates the challenging task of discerning them in the form of a complex envelope. The main idea of using spectrograms is to convert the information into a visually informative image that leverages from DL backbones such as Resnet-18, Resnet-50, MobilenetV2, etc [6]. Spectrogram datasets have been used previously in the literature. From a reconfigurable intelligent surface SS system reported in [7] to more intricated systems such as the improved encoder-decoder architecture for accurate SS in [6]. However, the background noise in spectrograms may hide the foreground pixels (signals) in the noise. This work addresses this problem by denoising spectrograms used to train a semantic segmentation network. Denoising is achieved using a Gaussian bilateral image restoration filter.

Following this introduction, the rest of the paper is organized as follows. Section II describes the proposed methodology. Section III presents the results using a Resnet-18 architecture to address semantic segmentation. Section IV discusses the advantages and disadvantages of using this type of preprocessing for SS. Finally, conclusions are drawn in Section V.

II. Methodology

The presented method denoises spectrogram images as a preprocessing step for training a DL network for SS. We provide the benefits of this denoising method using the Resnet-18 model, which was chosen due to its shallower 18-layer architecture compared to other residual networks [8]. Fig. 1 shows the two processes implemented in this work. The first one, depicted in Fig. 1a, does not use preprocessing, and the second one (Fig. 1b), uses bilateral filtering.

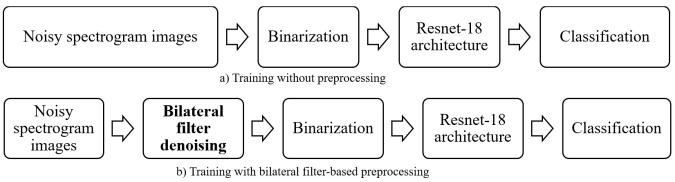

Fig. 1. Proposed method with and without preprocessing.

[1]This work was supported in part by Grants PID2019-103876RB-I00, PID2022-138078OB-I00, and PDC2023-145808-I00, funded by MCIN/AEI/10.13039/501100011033, by the EU ESF Investing in your future and by ERDF A way of making Europe.

A. Noisy spectrogram images

Based on the MATLAB example in [9], a synthetic spectrogram dataset was generated using LTE [10] and 5G [11] toolboxes. The dataset contains ground truth labels for each pixel in the spectrogram indicating if it belongs to one of three classes: LTE, NR, or Noise. Multiple signal-to-noise (SNR) ratios were implemented through an additive white Gaussian noise (AWGN) channel to provide more realistic scenarios. The SNR range goes from -20 dB up to 20 dB. More details regarding the 4G and 5G wireless standard features simulated can be found in [9]. A total of 10000 spectrograms were generated, 1000 corresponding to each SNR value $\{-20 : 20, step = 5\}$ dB. Some examples of pristine and noisy spectrograms are presented in Fig. 2 showing the variety of scenarios simulated in this image dataset.

B. Bilateral Gaussian filter (BGF) denoising

The bilateral filter is commonly used for denoising images accounting for harsh edges [12]. The filter uses a kernel to smooth pixels in a defined neighborhood. In this case, the kernel is Gaussian. The BGF recognizes pronounced edges and avoids smooth pixels with largely distinct intensities. In the proposed methodology, the images were transformed to the *L*a*b** color space [13] before applying the proposed filter [9].

C. Binarization

Binary images consist of logical values (0 or 1). Using a binary image for image processing and DL benefits from a smaller memory size used to store information and produce faster classification results. In this work, every spectrogram was binarized before training a DL architecture. The Otsu thresholding algorithm [14] was used to create two image datasets. The first dataset contains binarized spectrograms without any preprocessing. The second dataset is composed of binary images computed after the BGF smooths its surface. Fig. 3 presents examples of binarized spectrograms, one for each dataset, together with its pristine and noisy versions.

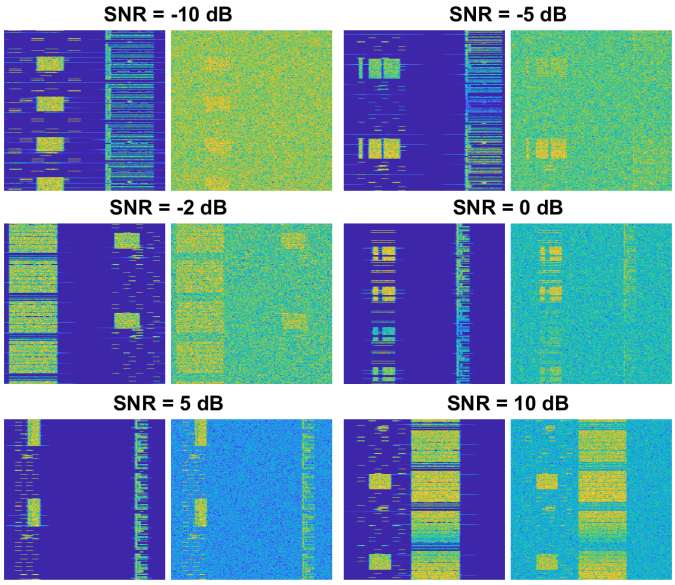

Fig. 2. Some examples of spectrograms with different SNR values. For each SNR, the left image shows the pristine spectrogram while the right image presents the corrupted spectrogram with AWGN.

D. Resnet-18 architecture

In a deep residual learning framework, shortcut connections are applied to connect layers inside the network, providing identity mapping and improving the training performance [8]. The Resnet-18 model was trained for semantic segmentation for this SS system, which classifies pixel images according to specific labels. This system comprises three classes: LTE, NR, and Noise. The last category refers to the background noise in every spectrogram representing pixels that do not include primary user signals (LTE or 5G NR).

III. Results

A. Training and validation

Two training processes were conducted to assess the application of the BGF to denoise spectrograms. The first training was executed using the noisy binarized spectrograms and the second was performed applying the restored binarized spectrograms. Both training processes were applied to the same Resnet-18 model, using the stochastic gradient descent with momentum (SGDM) optimizer and training parameters detailed in Table I.

TABLE I
TRAINING PARAMETERS

Parameter	Value
Batch size	20
Epochs	20
Image input size	256x256
Initial learning rate	0.02
Learning rate drop period	10
Learning rate drop factor	0.1

Each dataset was divided into training (80% of images), validation (15%), and test (5%) sets. The training and validation results are summarized in Table II.

TABLE II
TRAINING AND VALIDATION RESULTS FOR BOTH DATASETS

Dataset	Training iterations (epochs)	Validation accuracy
Noisy images	1750 (5)	85.33 %
Restored images	1300 (4)	80.74 %

The validation accuracy is close to each other for both datasets. However, the Resnet-18 architecture needed fewer training iterations for the restored images (spectrograms) dataset.

B. Test cases

Four test cases arise using both noisy and restored spectrograms as input to the spectrum sensing system. These cases are illustrated in Table III and explained below:

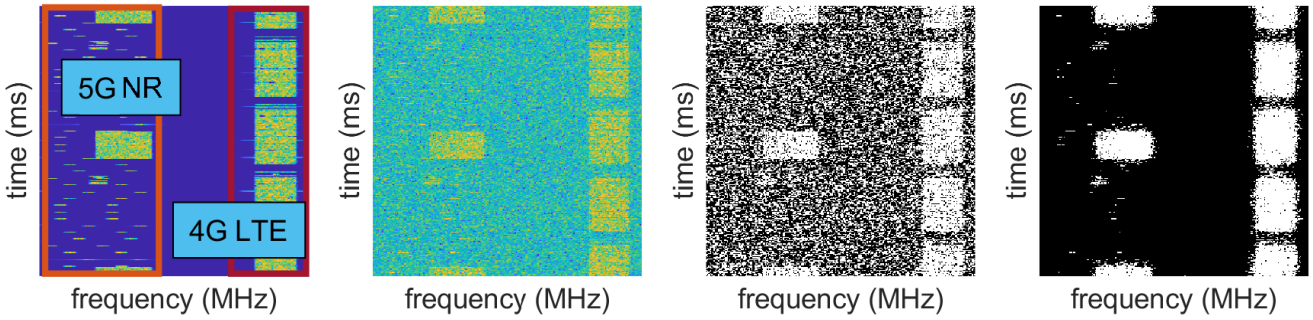

Fig. 3. a) Pristine spectrogram, b) Noisy spectrogram (SNR = 0 dB), c) Binarization of the noisy spectrogram, d) Binarization of the noisy spectrogram after bilateral filtering.

TABLE III
TEST CASES SUMMARY

Input spectrograms	Resnet-18 Trained Models	
	Noisy images	Restored images
Noisy images	Case 4	Case 3
Restored images	Case 1	Case 2

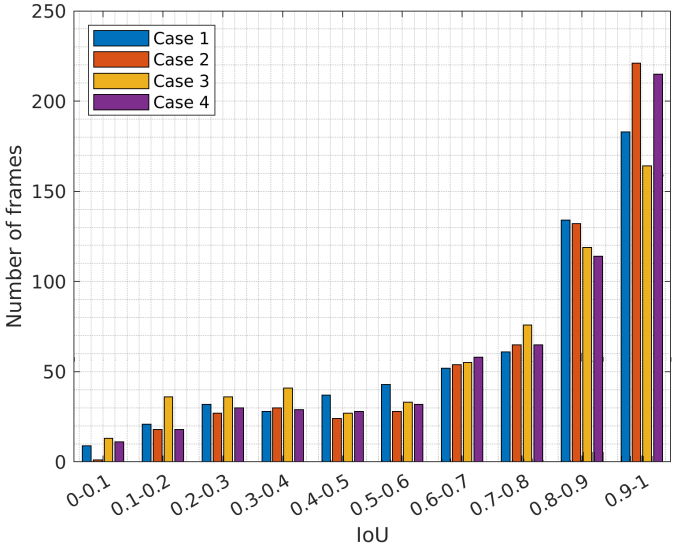

Fig. 4. Mean IoU for the four test cases.

- **Case 1:** Testing the Resnet-18 trained with original (noisy) images using as input restored images.
- **Case 2:** Testing the Resnet-18 trained with restored images using the restored images as input.
- **Case 3:** Testing the Resnet-18 trained with restored images using the original (noisy) images as input.
- **Case 4:** Testing the Resnet-18 trained with original (noisy) images using the original (noisy) images.

C. Evaluation

Fig. 4 presents the histogram of the per-image intersection over union (IoU) for the test cases. For each class (LTE, NR, Noise), IoU is the ratio of correctly classified pixels to the total number of ground truth and predicted pixels in that class.

It is interesting to notice that Case 2 has the largest number of frames correctly classified for the IoU $0.9-1$, followed by Case 4. However, this does not happen for the IoU range of $0.8-0.9$. Case 1 has the largest amount of frames for this range. Case 3 could be considered to have the lowest performance since it has the highest number of frames for lower IoU values such as $0-0.1$, $0.1-0.2$, etc.; in comparison to the other cases. Since these semantic segmentation models have been used for spectrum sensing, Fig. 5 evaluates the probability of detection (P_d) vs. the SNR range evaluated. P_d is the total count of correctly located signals with respect to the total number of signals in the spectrogram. The lowest SNR (-20 dB) is one of the harshest scenarios which is why it was necessary to account for the background noise using a denoising filter. Case 2 shows the highest P_d for the lowest SNRs from -20 dB to 0 dB.

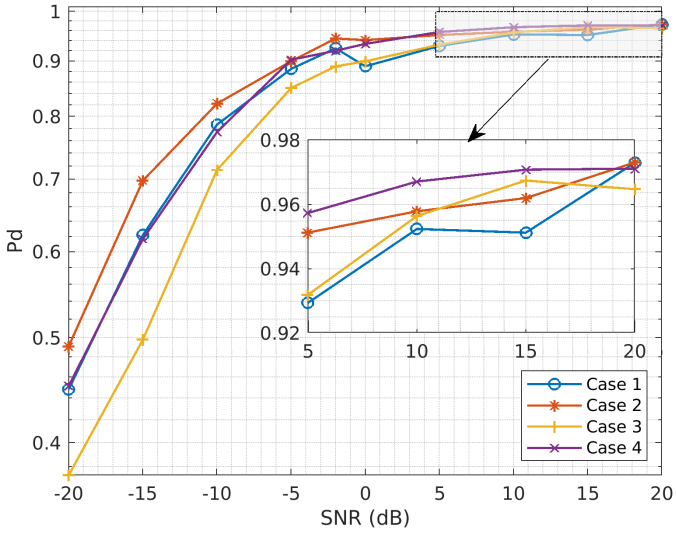

Fig. 5. Probability of detection for each test case.

For the SNR range from 5 dB to 20 dB, P_d has similar values for all the cases which makes sense since the noise level is much smaller than the other SNR range. The two spectrum sensing systems provide similar P_d when dealing with smaller noise environments. However, addressing noisy environments (SNR $<$ 0 dB) is important for spectrum sensing, and the preprocessing based on the bilateral Gaussian filter provides better results as proven in Fig. 5.

Fig. 6. Normalized confusion matrices for the four test cases.

IV. Discussion

Fig. 6 shows the normalized confusion matrices for the four test cases allowing us to assess the model's capability of correctly identifying the true vs. predicted classes. These results were calculated by evaluating the same image test set. For Case 4, the lowest noise recognition value is obtained because the network was trained using noisy spectrograms without any preprocessing. Cases 1 and 2 present similar noise recognition values, constituting the systems evaluated using input-restored spectrograms. Case 2 provides the best results, both for noise recognition and signal (LTE and NR) recognition. This shows that the Resnet-18 model benefits from using restored spectrograms for training and testing. The Noise classification value of 84.81% is the highest accuracy for this class compared to the other three cases. This is beneficial for spectrum sensing since a cognitive radio would need to make more certain decisions about only noise regions (i.e. unoccupied portions of the spectrum) as they define gaps for SU transmissions. We summarize the advantages and disadvantages of the BGF as a preprocessing method for SS in the following subsections.

A. Advantages

- Better noise recognition as seen on the confusion matrix for Case 2 in Fig. 6.
- Less training time since it takes fewer epochs to train the model with restored spectrograms as reported in Table II.
- Higher P_d for low SNR scenarios as seen in Fig.5.

B. Disadvantages

- Extra processing step for the SS system. Denoising the images includes performing more image processing computations.
- Smaller accuracy for recognition of 5G NR signals. However, since the objective of SS is finding spectrum holes this effect can be neglected.
- Similar results without preprocessing for high SNR scenarios. In a real scenario, the SNR would probably have strong noise variance impeding an accurate detection when the background noise is not addressed.

V. Conclusion

This work presents a method to increase the accuracy of spectrogram-based spectrum sensing trained by a deep learning (DL) network using a Resnet-18 model. It was proposed that spectrograms be preprocessed to remove background noise before DL training. A system for removing background noise in spectrograms based on Gaussian bilateral filtering was applied. The proposed methodology allows smoothing of the intensities of the background pixels without significantly affecting the foreground pixels, which in this case consist of LTE (4G) and NR (5G) signals. The Resnet-18 semantic segmentation network was trained with and without preprocessing the input spectrograms in order to validate the proposed methodology's benefits. The paper's results show the highest noise area recognition accuracy of 84.81% for Case 2, thus proving the proposed approach's benefits. In future work, we plan to assess other DL architectures such as Resnet-50, MobilenetV2, etc.

References

[1] J. Mitola and G. Q. Maguire, "Cognitive radio: Making software radios more personal," *IEEE personal communications*, vol. 6, no. 4, pp. 13–18, 1999.

[2] K. M. Captain and M. V. Joshi, *Spectrum sensing for cognitive radio: Fundamentals and applications*. CRC Press, 2021.

[3] A. Olesiński and Z. Piotrowski, "A radio frequency region-of-interest convolutional neural network for wideband spectrum sensing," *Sensors*, vol. 23, no. 14, p. 6480, 2023.

[4] H. Q. Nguyen, H. P. Nguyen, and B. T. Nguyen, "An image processing approach to wideband spectrum sensing of heterogeneous signals," in *International Conference on Cognitive Radio Oriented Wireless Networks*, pp. 211–221, Springer, 2018.

[5] A. Rojas and G. J. Dolecek, "A review of spectrum sensing techniques based on machine learning," *Encyclopedia of Information Science and Technology, Sixth Edition*, pp. 1–21, 2024.

[6] G.-V. Nguyen, C. V. Phan, and T. Huynh-The, "Accurate spectrum sensing with improved DeepLabV3+ for 5G-LTE signals identification," in *Proceedings of the 12th International Symposium on Information and Communication Technology*, pp. 221–227, 2023.

[7] S. Kayraklik, I. Yildirim, E. Basar, I. Hokelek, and A. Gorcin, "Practical implementation of RIS-aided spectrum sensing: A deep-learning-based solution," *IEEE Systems Journal*, 2024.

[8] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[9] Mathworks, "Spectrum sensing with deep learning to identify 5G and LTE signals." https://la.mathworks.com/help/comm/ug/spectrum-sensing-with-deep-learning-to-identify-5g-and-lte-signals.html#d123e4355, 2023.

[10] Mathworks, "LTE toolbox documentation." https://la.mathworks.com/help/lte/, 2024.

[11] Mathworks, "5G toolbox documentation." https://la.mathworks.com/help/5g, 2024.

[12] R. G. Gavaskar and K. N. Chaudhury, "Fast adaptive bilateral filtering," *IEEE Transactions on Image Processing*, vol. 28, no. 2, pp. 779–790, 2018.

[13] S. Murali and V. Govindan, "Shadow detection and removal from a single image using lab color space," *Cybernetics and Information Technologies*, vol. 13, no. 1, pp. 95–103, 2013.

[14] K. Saddami, K. Munadi, Y. Away, and F. Arnia, "Improvement of binarization performance using local otsu thresholding.," *International Journal of Electrical & Computer Engineering (2088-8708)*, vol. 9, no. 1, 2019.