DEMUCS FOR DATA-DRIVEN RF SIGNAL DENOISING

*Çağkan Yapar** *Fabian Jaensch** *Jan C. Hauffen** *Francesco Pezone*†
*Peter Jung**¶ *Saeid K. Dehkordi** *Giuseppe Caire**

*Technical University of Berlin, †Sapienza University of Rome, ¶DLR

ABSTRACT

In this paper, we present our radio frequency signal denoising approach, RFDEMUCS,[1] for the 2024 IEEE ICASSP RF Signal Separation Challenge. Our approach is based on the DEMUCS architecture [1], and has a U-Net structure with a bidirectional LSTM bottleneck. For the task of estimating the underlying bit-sequence message, we also propose an extension of the DEMUCS that directly estimates the bits. Evaluations of the presented methods on the challenge test dataset yield MSE and BER scores of -118.71 and -81, respectively, according to the evaluation metrics defined in the challenge.

***Index Terms*—** signal denoising, interference rejection, single-channel source separation, deep learning, supervised learning.

1. INTRODUCTION

Wireless communication links are susceptible to interference from other electromagnetic sources, such as other wireless communication devices, or non-communication devices e.g. a microwave oven. In a wireless communication system, the receiver is interested in estimating the bit-sequence message that is transformed at the transmitter into an appropriate electromagnetic waveform, the signal-of-interest (SOI). As part of a communication system design, a description (e.g., pulse shape, modulation, timing, frequency, error coding) of this transformation is available at the receiver, while information about the structure of the interfering signals is usually not.

Estimating the message bits or SOI with the highest accuracy under such interference is an imperative part of a wireless communication system. Traditional radio frequency (RF) signal separation (denoising) methods typically rely on prior knowledge of the interfering signal model, the assumption of non-overlapping time/frequency bands of the SOI and the interfering signal, or the availability of multiple antennas. Unfortunately, these requirements are not met in many realistic wireless communication scenarios, which this challenge addresses. A reasonable choice for problems with a lack of accurate modeling and an abundance of data is to employ data-driven solutions, specifically deep neural networks (DNN).

Denoising audio signals using DNNs has been a successful and very productive practice in recent years, e.g. [1, 2], achieving state-of-the-art performance. DNNs have also been used in other denoising tasks, such as that of seismic signals [3] and gravitational waves [4]. Recent work, e.g. [5], has considered the application of DNNs originally designed for audio signal denoising also to the RF signal denoising task, and the current paper is another attempt where the application of the DEMUCS architecture [1, 6] is considered.

The **task** of this signal processing grand challenge is to estimate both the SOI waveform and its underlying bit-sequence message from the given signal mixture $y = s + b \in \mathbb{C}^{40960}$, where s is the SOI, and b is an interference signal, which is a time-series segment from one of the frames of the EMISignal1, CommSignal2, CommSignal3, or CommSignal5G1 datasets provided by the organizers. Two different types of SOI are considered. The first SOI is a single-carrier signal with QPSK and Gray coding of uniformly distributed message bits, and modulated by a root-raised cosine pulse shaping function. The second SOI is an orthogonal frequency division multiplexing (OFDM) signal—a multi-carrier signal that is comprised of 64 (while 8 of them being inactive) orthogonal subcarriers, each carrying a QPSK symbol. Note that, for SOI 1 (QPSK) and SOI 2 (OFDMQPSK), the estimated bit-sequence messages $\hat{m}$ involve 5120 bits, and 57344 bits ($40960/80 = 512$ OFDM symbols $\times 56$ active subcarriers $\times 2$ bits per QPSK-symbol), respectively. For further details of the SOIs and their mathematical descriptions we refer to [7].

2. METHODOLOGY

2.1. Dataset

For **training**, we used a script provided by the organizers to generate 240000 sample mixtures with different random target SINRs (between -33 dB and 3 dB) for the four interference types. **Test** mixtures were provided for various target SINRs from -30 dB to 0 dB in 3 dB increments, 100 cases per SINR.

2.2. DEMUCS for RF Signal Denoising and Bit Regression

DEMUCS [1, 6] is an encoder/decoder architecture composed of a convolutional encoder, a unidirectional or bidirectional LSTM (we considered the latter) applied on the encoder's output, and a convolutional decoder, with the encoder and decoder linked with U-Net skip connections. It is characterized by its number of layers L in encoder/decoder, initial number of hidden channels H, layer kernel size K, stride S and resampling factor U.

[1] `https://github.com/CagkanYapar/RFDemucs`

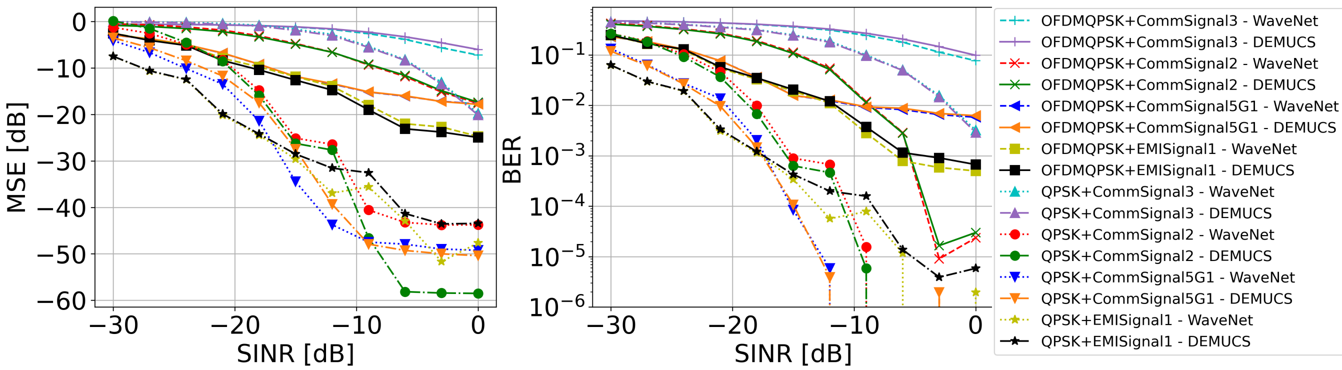

Fig. 1: Comparison of the test MSE (Mean Squared Error) and BER (Bit Error Rate) accuracies of the baseline WaveNet [2,8] provided by the challenge organizers with those of the DEMUCS [1] architecture we adopted. **Left:** MSE, **Right:** BER.

Except for QPSK + CommSignal3, we set $H = 64$, $S = U = 2$, and for the former we set $H = 80$ and $S = U = 4$. For all scenarios, we set $K = 8$ and $L = 5$. Unlike [1], we did not normalize the input by its standard deviation, as it has been our experience that this is detrimental to performance.

We also propose an extension of the DEMUCS for direct **bit regression** by appending the DEMUCS architecture with a fully connected layer that is applied to consecutive disjoint blocks of an appropriately chosen number of output samples (e.g., 64 for QPSK SOI) of DEMUCS and outputs the bits (e.g., 8 bits for QPSK) during the inference phase after applying a "hard decision" on the threshold value of 0.5. This idea of extending DEMUCS was inspired by the "Bit Regression" baseline method from the *Single-Channel RF Challenge.*[2]

2.3. Training

For each combination of SOI and interference, we train a separate DEMUCS model. For SOI estimation and bit regression, we use the MSE loss for training. We also considered training the bit regression DNN with cross entropy loss. For the QPSK+CommSignal2 and QPSK+CommSignal3 cases, we used a learning rate of $3 \cdot 10^{-4}$, otherwise $3 \cdot 10^{-5}$. We used ReduceLROnPlateau scheduler with a patience of 3 and EarlyStopping with a patience of 12 and Adam optimizer [9].

3. RESULTS

According to our results shown in Fig. 1, DEMUCS closely follows WaveNet in most scenarios, yielding an overall MSE score of DEMUCS (WaveNet) -118.71 (-119.35) and a BER score of DEMUCS (WaveNet) -81 (-78) in the challenge test set (cf. [7] for the definitions/formulations of these scores. The lower the better on both scores). Our results show that almost (except for QPSK + CommSignal5G1) in all considered settings, the estimation of the SOI followed by the matched filtering baseline method yields better BER performance than the bit regression DEMUCS. Another observation is that for bit regression, using the MSE loss instead of the cross-entropy loss leads to higher accuracy, supporting the findings in [10].

Remark 1: *Surprisingly, the common practice of early stopping based on validation loss failed in the QPSK + CommSignal3 setting for the* DEMUCS *architecture, as the validation loss continued to decrease along with the training loss. We note that the training/validation dataset generation script provided by the challenge organizers is based on extracting random frames from a "global" dataset,* INTERFERENCESET *[7], and thus the training and validation datasets are not guaranteed to be disjoint, which may have played a role in the overfitting in this exceptional case. To remedy the overfitting, in this setting we used the* TESTSET1EXAMPLE *dataset for validation, which is not part of the* INTERFERENCESET.

[2] https://github.com/RFChallenge/rfchallenge_singlechannel_starter

4. REFERENCES

[1] A. Défossez, G. Synnaeve, and Y. Adi, "Real time speech enhancement in the waveform domain," in *Proc. Interspeech*, 2020.

[2] D. Rethage, J. Pons, and X. Serra, "A wavenet for speech denoising," in *Proc. 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018, pp. 5069–5073.

[3] A. Novoselov, P. Balazs, and G. Bokelmann, "SEDENOSS: SEparating and DENOising Seismic Signals with dual-path recurrent neural network architecture," *Journal of Geophysical Research: Solid Earth*, vol. 127, no. 3, pp. e2021JB023183, 2022.

[4] H. Shen, D. George, E. A. Huerta, and Z. Zhao, "Denoising gravitational waves with enhanced deep recurrent denoising auto-encoders," in *Proc. ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 3237–3241.

[5] G. C. Lee, A. Weiss, A. Lancho, Y. Polyanskiy, and G. W. Wornell, "On neural architectures for deep learning-based source separation of co-channel OFDM signals," in *Proc. ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023, pp. 1–5.

[6] A. Défossez, N. Usunier, L. Bottou, and F. Bach, "Music source separation in the waveform domain," *arXiv preprint arXiv:1911.13254*, 2019.

[7] T. Jayashankar, B. Kurien, A. Lancho, G. C. Lee, Y. Polyanskiy, A. Weiss, and G. Wornell, "The data-driven radio frequency signal separation challenge," in *To appear in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, April 2024.

[8] A. van den Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. Senior, and K. Kavukcuoglu, "WaveNet: A generative model for raw audio," *arXiv preprint:1609.03499*, 2016.

[9] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, San Diego, CA, USA, May 2015.

[10] L. Hui and M. Belkin, "Evaluation of neural architectures trained with square loss vs cross-entropy in classification tasks," *arXiv preprint:2006.07322*, 2020.