OPTIMIZED SIZE-PERFORMANCE MODEL FOR INTERFERENCE REJECTION IN DIGITAL COMMUNICATIONS FOR THE ICASSP 2024 CHALLENGE

*Ashkan Jafari Fesharaki[1]** *Hossein Alizadeh Piralidehi[1]** *Ali Hasanzadeh Karkan[2]* *François Leduc-Primeau[2]*

[1]Department of Electrical and Computer Engineering, University of Tehran
[2]Department of Electrical Engineering, Polytechnique Montreal

ABSTRACT

This study summarizes our submitted model to the ICASSP 2024 interference extraction challenge. In this challenge, the goal is to extract the Signal of Interest (SOI) from the mixture signal of SOI and the interference signal. Our submitted model features the following. First, we changed the baseline's signal estimation concept to symbol estimation. Second, we use the given prior Signal-to-Interference Ratio (SIR) information in the training. Third, we suggest a method to reduce memory requirements during training while maintaining comparable performance to baselines. Fourth, we customize the training loss function regarding baselines' weaknesses. Finally, after applying our proposed ideas to a baseline method, it results in a significantly smaller model with superior performance compared to baselines.

1. INTRODUCTION

In the ICASSP 2024 SP Challenge, the objective is to develop an engine to extract a signal of interest (SOI) that uses QPSK or OFDM QPSK modulations from a composite signal. The challenge involves handling four unknown interference structures using data-driven methods, requiring the development of eight specialized models in total [5]. In [1], it is shown that communication signals with separable sources in time and/or frequency can be effectively separated using masking and classical filtering methods. The primary challenge in this field is separating signals that overlap in both time and frequency with a single antenna receiver, lacking spatial diversity. This is known as single-channel source separation (SCSS). Numerous methods in the literature address SCSS in digital communication signals. A common approach is maximum likelihood sequence estimation, using algorithms like particle filtering [2] and per-surviving processing [3]. However, these methods necessitate prior knowledge of signal models, which may be unavailable in practice. In [4], a CNN-U-NET model is presented for the data-driven interference rejection, using the CNN as a synchronization step for finding the starting point of the added interference. However, while this model achieves great results compared to previous studies in both BER and MSE, it has a large size and parameters leading to high resource requirements and possible slow convergence. In this report, we propose four ideas that can be applied to a baseline UNET or WaveTorch NET, resulting in a smaller and lower complexity model while achieving the same or better results in MSE and BER.

2. METHODOLOGY

In this study, our contribution to propose a size-performance optimized interference rejector model is divided into four sections and has been applied to baseline UNET for evaluation.

2.1 Symbol Estimation

To extract the Signal of Interest (SOI) from the mixture signal, previous studies used two Neural Network models with a 40960-size input and output. We transformed the initial Signal-Estimation model into a Symbol-Estimation model based on the challenge's assumption. Considering two SOI types (QPSK and OFDM QPSK), we came up with two different structures (Figure 1).

2.1.1 QPSK Type (Signal-to-Symbol Estimator): The model would be designed to receive the mixture signal as the input and estimate the symbols of SOI. This results in a model with an input size of 40960 and an output size of 2560. By simplifying the estimation problem, we have the freedom to decrease the number of filters in convolutional layers. The proposed structure applied to baseline UNET will have the same length encoder and less deep decoder with fewer filters compared to the baseline UNET.

2.1.2 OFDM QPSK Type (Symbol-to-Symbol Estimator): The model would be designed to receive the demodulated mixture signal as the input and estimate the symbols of SOI. This results in a model with an input size of $28672 = 56 * 512$ and an output size of 28672 where $56 = \bar{K} = K - 8$ and $512 = \frac{N}{K+T_{cp}}$. It needs to be mentioned that in the demodulating part before input, the cyclic prefix and the zero padding are going to be removed so no information is going to be lost. As for the QPSK-type model, we have the freedom to decrease the number of filters in convolutional layers. The proposed structure applied to baseline UNET results in encoder and decoder parts with reduced depth and fewer filters compared to the baseline UNET. The model goal change offers dual advantages. Firstly, it simplifies UNET or WaveTorch NET, reducing resource requirements (e.g. GPU RAM); Secondly, in the QPSK model, it accelerates convergence by reducing the output size from 40960 to 2560. Additionally, in the OFDM QPSK model, interference power decreases during preprocessing by eliminating non-informative signal parts like cyclic prefix and zero padding.

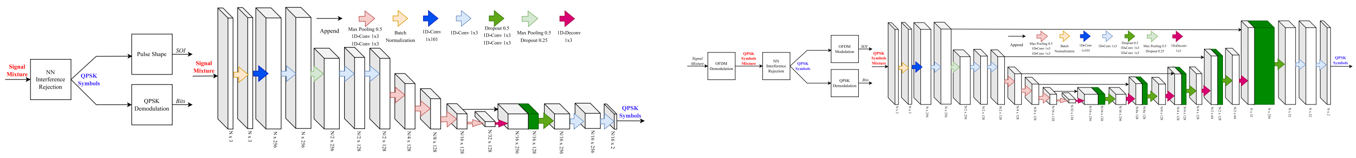
2.2 Prior Information on Interference to Signal Ratio (ISR)

Through experimentation, we found that the challenge in extracting SOI is not mainly due to interference power but a lack of information on its value. To address this, we added a repetitive ISR vector, increasing input channels from 2 to 3. While challenge data included ISR metadata, real-world scenarios lack this, prompting the suggestion of a simple neural network to predict ISR in dB. This modification helps the model distinguish interference powers for more accurate output estimation.

2.3 Training

2.3.1 Suboptimal Method for Training Process: Previous studies suggest optimal model training involves creating a sizable dataset

*These authors contributed equally to this work.



a) Overall model structure for QPSK SOI-type mixture input

b) Overall model structure for OFDM QPSK SOI-type mixture input

Fig. 1: Proposed Engine's Structure after applying the mentioned modifications to baseline UNET for QPSK and OFDM-QPSK type mixtures

with an equal number of samples at each Signal-to-Interference Ratio (SIR) value. In this challenge, the baseline methods used a high number of samples per SIR (e.g. 4000), demanding substantial system memory. Our approach involves implementing a loop where, in each iteration, a dataset with a reduced number of samples per SIR is generated and fed to the model for training. While suboptimal compared to baseline models, this method presents the same dataset volume to the model, aiding training with lower memory requirements.

2.3.2 Activation Functions: In our challenge with vector inputs fluctuating between negative and positive values, ReLU's tendency to discard negatives poses a risk of the 'Dying ReLU' problem. Baseline models, using ReLU in all layers except the last (Linear), risk bias term inflation, affecting subsequent layers and extending training time. To address this, we adopted LeakyReLU activation functions, preserving negatives. This mitigates the 'Dying ReLU' issue, fostering faster convergence and enhancing the model's ability to capture nuanced information in mixed-sign inputs.

2.4 Customized Loss Function

In prior studies, it was surprising to find that different Neural Network models, while one excels in MSE, may perform worse in BER. This is because one aims for overall signal reconstruction, while the other focuses on precise bit detection. To address this, we tailored our model's Loss Function with dual constraints: maximizing signal reconstruction and enhancing bit detection precision. Our proposed loss function comprises two terms to achieve these objectives. Our loss function is given by

$$\boldsymbol{L} = \boldsymbol{MSE}\{Y, \hat{Y}\} + \psi\, \boldsymbol{BCE}\,\{S(Y), 1 - Sigmoid(\hat{Y})\},$$

where S(Y)=0 if Y>=0, S(Y)=1 if Y<0, and ψ is a hyperparameter. Note that $\hat{Y}$ and Y are a 2-channel vector of real and imaginary parts of estimated and ground truth symbols.

2.5 Other Details

We implemented and tested the proposed ideas to the baseline UNET using TensorFlow. In this challenge the mixtures of QPSK and OFDM QPSK types with eight unknown interference frames were investigated, for which we propose one structure for each of the modulations (Figure 1). We trained our model utilizing transfer learning for each type of mixture, in a loop (refer to 2.3.1) iterating 10 times with 10 epochs for training in each iteration using Adam optimizer with a learning rate of 0.0003.

3. EVALUATION AND RESULTS

In this challenge, MSE and BER scores were calculated by a specific formulation mentioned in the contest's manual where the smaller the score, the better the performance in each MSE and BER [5]. According to Table 1, after applying the ideas mentioned in previous sections, while the baseline UNET model has 147% larger model size in QPSK and 115% larger model size in OFDM-QPSK SOI type, the proposed model got 138% better MSE score and equal BER score. Moreover, although WaveTorch model has 38% better MSE and 146% better BER than the proposed UNET model, it has 132% larger size in QPSK and 100% larger size in OFDM-QPSK SOI type. As the proposed ideas do not rely on the which models is used, they can be applied to WaveTorch NET, making it much simpler and smaller in terms of model size, while performing better or at worse equal to baseline WaveTorch NET.

Table 1: Size-Performance comparison in terms of trainable parameters and MSE/BER score achieved in the challenge [5]

Model	Total Number of Parameters		MSE score	BER score
	QPSK	OFDM QPSK		
Baseline UNET	4,222,698	4,222,698	-36.3	-42
Proposed UNET	1,702,920	1,961,320	-86.6	-42
WaveTorch NET	3,964,674	3,964,674	-119	-78

In Fig 2, the average performance over eight types of mixture signal of baseline models and the proposed UNET model is demonstrated. It is shown that after applying the proposed ideas to the baseline UNET, it outperforms and matches the baseline UNET in terms of MSE respectively, while having significant smaller size and requiring significant lower resource requirements for training. Moreover, WaveTorch complex structure shows its strength compared to baseline UNET and modified UNET; however, it is a significant larger model rather than the modified UNET.

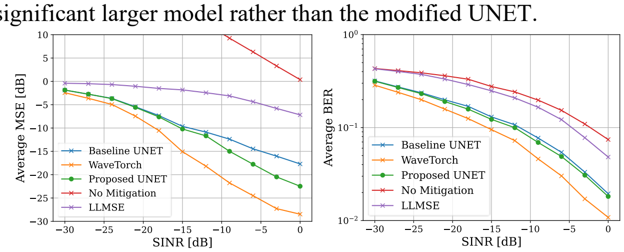


Fig. 2: Performance comparison: Applied proposed ideas to baseline UNET vs. baseline UNET and WaveTorch NET

4. CONCLUSION

We propose resource-efficient modifications for interference rejection models, achieving superior performance with reduced resource requirements and smaller model size compared to the baseline UNET. Future work can explore these ideas in the context of WaveTorch NET as the proposed ideas are conceptual and universal, not relying on the model's structure.

5. REFERENCES

[1] M. G. Amin, "Interference mitigation in spread spectrum communication systems using time-frequency distributions," IEEE Trans. Signal Process., vol. 45, no. 1, pp. 90–101, Jan. 1997.

[2] T. Shilong, C. Shaohe, Z. Hui, and W. Jian, "Particle filtering based single-channel blind separation of co-frequency MPSK

signals," in IEEE Int. Symp. Intell. Signal Process. and Commun. Syst., Feb. 2007, pp. 582–585.
[3] T. Shilong, Z. Hui, and G. Na, "Single-channel blind separation of two QPSK signals using per-survivor processing," in IEEE Asia Pac. Conf. Circuits Syst. (APCCAS), Dec. 2008, pp. 473–476.
[4] A. Lancho et al., "Data-Driven Blind Synchronization and Interference Rejection for Digital Communication Signals," GLOBECOM 2022 - 2022 IEEE Global Communications Conference, Rio de Janeiro, Brazil, 2022, pp. 2296-2302.
[5] T. Jayashankar et al., "The data-driven radio frequency signal separation challenge," ICASSP 2024 -2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP0, Boston, USA, 2024.