Supplementary Material: Water Leak Detection and Classification with Complex-valued Neural Networks and Sensor Fusion

D.U. Leonzio, S. Mandelli, P. Bestagini, M. Marcon, S. Tubaro

Introduction

This document serves as supplementary material for the paper titled "Water Leak Detection and Classification with Complex-Valued Neural Networks and Sensor Fusion." Due to space constraints in the original manuscript, we provide additional experiments, results, and details in this document to further support our findings.

The supplementary material is structured into three sections. The first section presents an analysis of the computational complexity of our proposed method compared to baseline approaches, with a focus on inference time measurements across multiple cross-validation folds. The second section provides detailed visualizations of the loss and accuracy functions, offering insights into the training dynamics and convergence behavior of the models. Finally, the third section includes a comparative evaluation of our approach against a Recurrent Neural Network (RNN), highlighting the advantages and potential trade-offs in terms of detection and classification performance.

1 Computational complexity

In this section, we present a computational complexity comparison between the baseline methods and the proposed approach. To conduct this evaluation, we measured the inference time required for each method across five different cross-validation folds. Specifically, we computed both the mean and standard deviation of the inference times, as reported in Table 1 and Table 2. The results indicate that although our method exhibits a slightly higher inference time compared to the baselines, it remains highly efficient, with execution times consistently within the millisecond range. This level of efficiency ensures that our approach is well-suited for real-time or near-real-time applications, where rapid decision-making and low-latency processing are critical requirements.

Table 1: Inference times in milliseconds for Hydrophone and Accelerometer in detection task

Sensor	**Baseline I**	**Baseline II**	**Baseline III**	**Proposed**
Hydrophone	4.8 ± 0.6	69.9 ± 0.1	1.37 ± 0.01	197.3 ± 0.02
Accelerometer	3.5 ± 0.4	10.4 ± 0.1	1.73 ± 0.01	132.1 ± 0.11

Table 2: Inference times in milliseconds for Hydrophone and Accelerometer in classification task

Sensor	**Baseline I**	**Baseline II**	**Baseline III**	**Proposed**
Hydrophone	8.1 ± 0.1	88.7 ± 0.2	8.35 ± 0.10	193.2 ± 0.01
Accelerometer	5.6 ± 0.02	50.3 ± 0.3	2.28 ± 0.01	127.1 ± 0.05

2 Training Curve

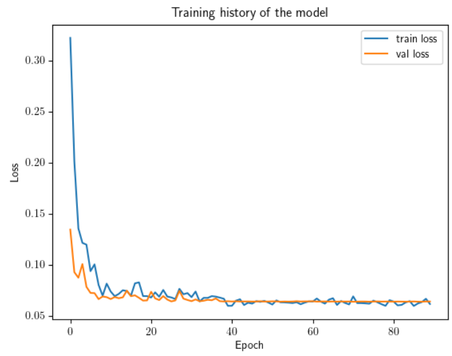
In this section, we present the training curves that, for the sake of readability and clarity, were not included in the main paper. These curves provide valuable insights into the learning process of our model, offering a deeper understanding of its optimization dynamics.

For each sensor type, namely the hydrophone and the accelerometer, we report the loss function and accuracy curves computed on both the training and validation partitions. We provide these plots separately for both the detection and classification tasks to illustrate how the model converges during training and how well it generalizes to unseen data. As described in the main paper, we utilized the cross-entropy loss function for training, with a learning rate set to 0.001. The optimization was carried out using the Adamax optimizer, incorporating a weight decay of 0.01 to promote regularization and prevent overfitting.

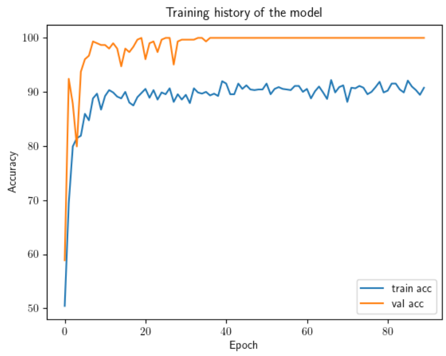
(a) Loss function curve.

(b) Accuracy curve.

Figure 1: Loss function and accuracy curves for hydrophone sensor in detection task.



(a) Loss function curve.



(b) Accuracy curve.

Figure 2: Loss function and accuracy curves for accelerometer sensor in detection task.

(a) Loss function curve.

(b) Accuracy curve.

Figure 3: Loss function and accuracy curves for hydrophone sensor in classification task.

(a) Loss function curve.

(b) Accuracy curve.

Figure 4: Loss function and accuracy curves for accelerometer sensor in classification task.

3 RNN comparison

In this section, we present a comparative analysis of our proposed method against a RNN architecture. The RNN model was specifically designed to process input windows of the raw signal with the same length as the window used to compute the Short Time Fourier Transform (STFT), which serves as the input for the Convolutional Neural Network (CNN) proposed in the main paper. This ensures a fair comparison by maintaining consistency in temporal resolution across both approaches.

We evaluated the RNN model on both detection and classification tasks for each sensor type, namely the hydrophone and the accelerometer. The RNN processes waveform samples using a window length of 2 seconds, allowing it to capture temporal dependencies in the raw signal. As in the main paper, we report results separately for each sensor and task. For the detection task, performance was assessed on both datasets, while for classification, evaluations were conducted exclusively on the laboratory-scale dataset. We report the result in Table 3 for detection and in Table 4 for classification. To compare the methods we decided to adopt the balanced accuracy at fixed threshold in both the tasks.

The results indicate that the proposed method consistently outperforms the RNN-based approach for both detection and classification tasks, demonstrating superior performance using data from both hydrophone and accelerometer sensors. These findings highlight the advantages of our approach, particularly in leveraging spectral representations and sensor fusion to improve accuracy.

As part of future work, we plan to further investigate the potential of alternative architectures, such as advanced RNNs variants or Transformer-based models, to enhance performance and generalization capabilities. This exploration could provide deeper insights into the benefits of different modeling paradigms for water leak detection and classification.

Table 3: Performance Comparison of RNN and Proposed method for Detection Task in terms of balanced accuracy. In bold the best results.

Sensor	**Dataset**	**RNN**	**Proposed**
Hydrophone	D1	0.87	**0.96**
	D2	0.85	**0.90**
Accelerometer	D1	0.85	**0.95**
	D2	0.86	**0.96**

Table 4: Performance Comparison of RNN and Proposed method for Classification Task in terms of balanced accuracy. In bold the best results.

Sensor	**Dataset**	**RNN**	**Proposed**
Hydrophone	D1	0.73	**0.90**
Accelerometer	D1	0.72	**0.90**