DOI 10.5162/Ultrasonic2025/A16-a4

Enhancing Air-coupled Ultrasonic Inspection: Resolving Signal Overlapping in Long-duration Chirp Signals

Muhammad Tayyib[1], *Musyyab Yousufi*[2], *Paulius Tervydis*[1], *and Linas Svilainis*[1]
[1] *Department of Electronics Engineering, Kaunas University of Technology, Lithuania*
[2] *Centre of Real Time Computer Systems, Kaunas University of Technology, Lithuania*
muhammad.tayyib@ktu.edu

Abstract: Air-coupled ultrasound investigations are crucial for non-destructive testing of thin and delicate materials. Conventional pulse excitation yields suboptimal signal-to-noise and narrow bandwidth. Chirp signals improve these parameters, however, requires long duration to do so, which results in overlapping of reflections. To solve the overlapping this paper proposed the use of residual deep learning approach using convolutional UNet1D architecture and composite loss function. Results show successful separation of overlapping reflections.

Keywords: Chirp signals, 1D convolutional neural network, overlapping reflections, air-coupled ultrasound, residual UNet.

Introduction

The use of air-coupled ultrasound is gaining traction in non-destructive testing (NDT), due to its adaptability across various mediums and non-contact investigations. Short-duration pulse excitation signals are commonly used because of their high temporal resolution and simple interpretation [1]. The main problem with pulse signals is limited bandwidth and low signal-to-noise ratio (SNR), especially in air-coupled ultrasound, where a high impedance mismatch between air and sample leads to losses in energy transmission [1]. An alternative approach involves the use of long-duration chirp signals (linear frequency modulation (LFM)), which offers the ability to increase the bandwidth and duration of signals (Fig. 1), as in [2] it was demonstrated that the longer the signal, the better the ability to compensate for spectral losses.

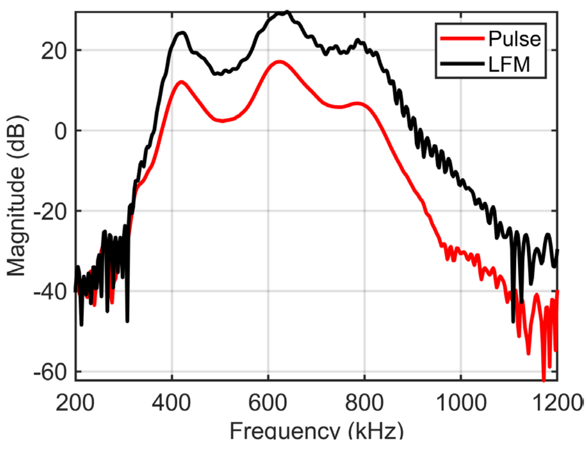

Fig. 1: Spectral response of pulse compared with LFM chirp signal

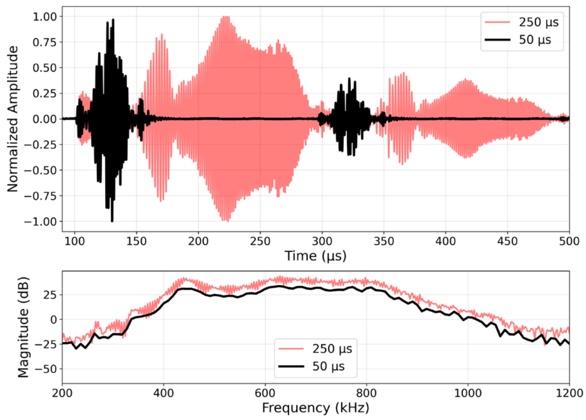

Fig. 2: Temporal response (top), spectral response (bottom) of $50\,\mu s$ *chirp compared with* $200\,\mu s$ *chirp*

The typical air-coupled ultrasound system is composed of transmitting and receiving transducers that cannot be too far from each other; thus, for short durations $10\,\mu s$ to $100\,\mu s$, the investigation could be carried out without acquisition problems; however, in longer signals, i.e. $150\,\mu s$ onward reflections begin to overlap in the received signals (Fig. 2 Top), making it extremely difficult to isolate and interpret individual reflections. Moreover, overlapping signals in the spectral domain (Fig. 2 bottom) results in a serrated response that makes it difficult to use the signal for parameter estimation.

Signal overlapping is traditionally solved by manual gating [3], matched filtering [4], or through frequency domain processing. Manual gating requires selection of time windows to isolate the reflections; however, it becomes ineffective if the signal overlaps and there is

DOI 10.5162/Ultrasonic2025/A16-a4

no distinction between reflections. Matched filtering or inverse filtering provides improved temporal resolution based on the assumption that the system is linear and the impulse response is known, which is almost impractical in air-coupled systems.

Moreover, in the case of time-localized reflection, they suffer from leakage and are not reliable. Deep neural networks (DNN) have been applied for signal separation in recent studies [5, 6], in [5], a deep learning approach classification task was achieved for modulation recognition, while the methodology was evaluated using only simulated data. Similarly, in [6], the results were also calculated with simulated data. Compared to these approaches, real-world experimental methods are needed that not only have the ability to train on experimentally obtained data but, instead of only classification, predict for the data which was not even part of the training.

In the air-coupled ultrasound received signals, if the reflections overlap and there is no gating, the spectral response contains too much interference making it difficult to interpret, as shown in Fig. 2(bottom). The overlapping signal scenario was experimented with using $650\,\mathrm{kHz}$ center frequency air-coupled transducers of different duration of LFM ($10\,\mu\mathrm{s}$ to $350\,\mu\mathrm{s}$). From visible interpretation of even the time domain signal, it was clear that signals up to $100\,\mu\mathrm{s}$ exhibit separate reflections, while longer signals, that is, $150\,\mu\mathrm{s}$, were affected by reflection overlapping. Thus, a novel approach is required to solve the overlapping in reflections to extract meaningful information from long-duration excitation signals.

In this work, a residual deep learning-based approach is proposed that focuses on the solution of overlapping in long-duration excitation signals. The model was trained on short-duration $10\,\mu\mathrm{s}$, $20\,\mu\mathrm{s}$, $50\,\mu\mathrm{s}$ gated signal (only the first reflection was preserved), for overlapped long-duration signals $150\,\mu\mathrm{s}$, $200\,\mu\mathrm{s}$, $300\,\mu\mathrm{s}$, and $350\,\mu\mathrm{s}$. The excitation voltage on the transducers was measured and considered as the input to the model, the output of the deep learning model was the calibration signal with only first reflection, the target function was L_2 norm in both time and frequency domains between the predicted signal and the actual signal. Thus, ensuring both temporal and spectral information for the modeling, the aim was to train the model in such a way that it can predict the non-overlapping reflection from the long-duration overlapping reflections using the knowledge learned during short-duration signals training.

To the best of our knowledge, this is the first application of residual UNet with 1D convolutions (ResUNet1D) to separate overlapping reflections in experimentally obtained time-domain air-coupled ultrasound signals.

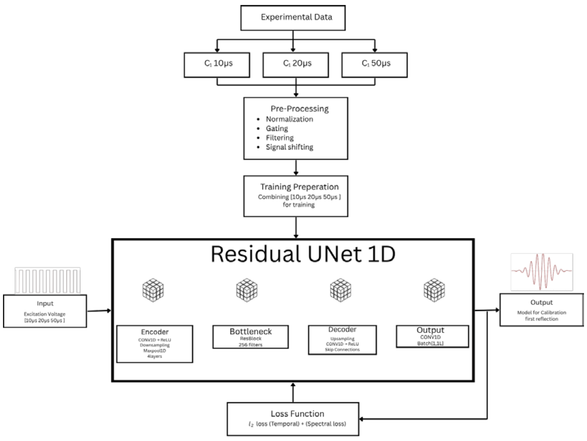

Fig. 3: Proposed Res-UNet1d based training flow diagram

2. Materials and Methods

Data was partitioned in such a way, as for training $10\,\mu\mathrm{s}$, $20\,\mu\mathrm{s}$ and $50\,\mu\mathrm{s}$ signals were used while long durations $150\,\mu\mathrm{s}$, $200\,\mu\mathrm{s}$, $250\,\mu\mathrm{s}$, $300\,\mu\mathrm{s}$, $350\,\mu\mathrm{s}$, were used for the testing where reflections are already overlapping. Data collected and stored in mat format where 100 repetitions were measured for each duration for a better statistical analysis of a diverse training set.

The methodology is shown in Fig. 3. As the excitation voltage for obtaining these signals was based on rectangular chirp, both the excitation voltage and the calibration signals were normalized to ensure a smooth training process; moreover, band-pass filtering was done between $0.15 - 1.5\,\mathrm{MHz}$ to remove any unwanted noise.

The UNet was developed from scratch and the ResUNet1D block was specifically designed for residual tasks involving air-coupled ultrasound modeling as 1D time-domain signals. The residual block consists of two 1D convolutional layers, followed by batch normalization and activated by leaky rectified linear unit (LReLU) functions. Encoding was achieved in four stages; each stage consists of dual residual blocks with progressively increasing dilation $(1\,,\ 2\,,\ 4\,,\ 8\,)$. These dilation rates ensure that the models effectively capture complex temporal features on varying scales.

The bottleneck segment deepens the feature extraction with two residual blocks characterized by high dilation rates, that is 16, capturing long-range dependencies in the signal. The decoder mirrors the encoder path with slight up-gradation of using ConvTranspose1D for up-sampling. The final output was generated by a sequence of a 1x1 convolution reducing feature channels from 64 to 32, a ReLU activation, and a subsequent 1x1 convolution projecting to a

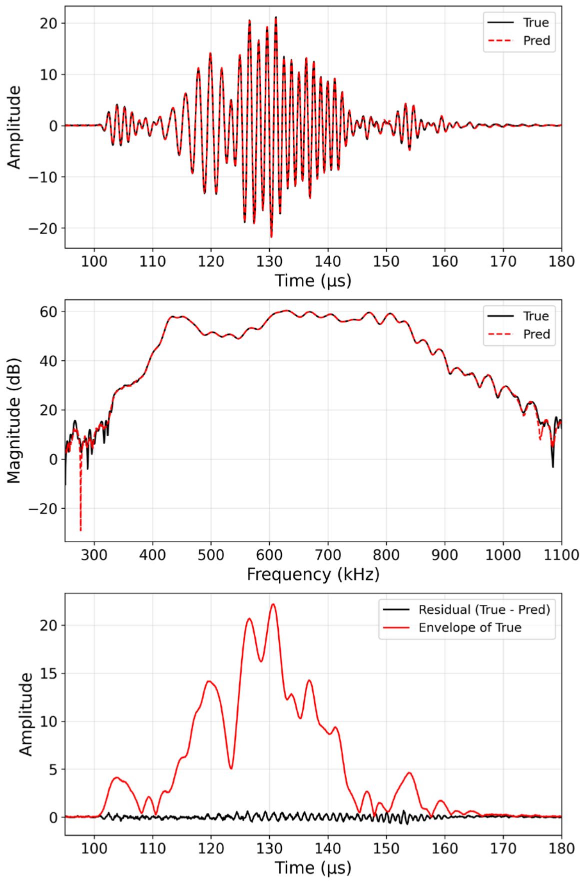

Fig. 4: Temporal (top), spectral (middle) and residual (bottom) response of $50\,\mu\mathrm{s}$ *that was part of the training*

single-channel output, corresponding to the target residual signal

As training was achieved using short duration signals, the fitting results of modeling in comparison to the experiment are shown in Fig. 4, where both temporal (Fig. 4 (top)) and spectral (Fig. 4 (middle)) fitting results are in perfect match. Furthermore, the residual (Fig. 4 (bottom)) shows nearly perfect fitting results since the only remaining part contains noise content. The total loss function ($\mathcal{L}_{\text{total}}$) is described as fitting in both time ($\mathcal{L}_{\text{time}}$) and frequency ($\mathcal{L}_{\text{freq}}$) domains, as:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{time}} + \alpha\,\mathcal{L}_{\text{freq}} \tag{1}$$

$$\mathcal{L}_{\text{time}} = \frac{1}{N}\sum_{n=1}^{N}(y_n - \hat{y}_n)^2 \tag{2}$$

$$\mathcal{L}_{\text{freq}} = \frac{1}{K}\sum_{k=1}^{K} w_k\left(|\mathcal{F}\{y\}_k| - |\mathcal{F}\{\hat{y}\}_k|\right)^2 \tag{3}$$

where $\alpha = 0.1$ is the trade-off weight (frequency loss scaling factor) for the total loss function that was set at such a value to obtain optimal or stable performance across our dataset after several validation experiments. y_n and $\hat{y}_n$ are the true experimental signal and predicted signals in sample n respectively, N is the total number of samples, $\mathcal{F}y_k$ is discrete Fourier transform (DFT) of y at frequency k and w_k is the weighting coefficient for the frequency bin k. The weight w_k is set to 1 for frequency bins within the range $[200, 1200]$kHz, and 0.5 outside this range, which results in a penalty for mismatch outside this passband region.

Results

The Res-UNet1D modeling was successfully achieved while training only for short-duration signals and testing for both short and long-duration signals (long-duration signals were not part of the model development, only part of the validation, i.e. prediction). Fig. 5 plots the time, freq and residual plots of $250\,\mu\mathrm{s}$

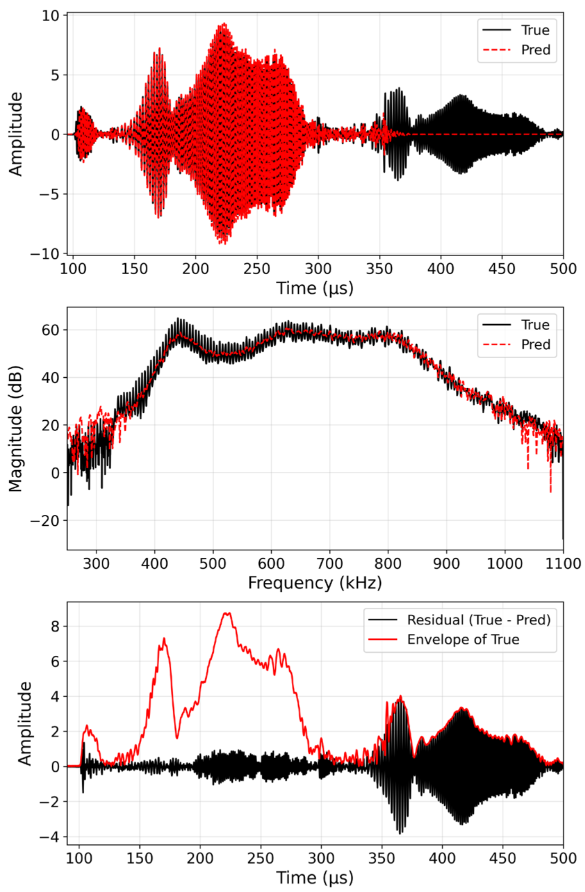

Fig. 5: Temporal (top), spectral (middle) and residual (bottom) response of $250\,\mu\mathrm{s}$ *chirp signal, compared with the model (first reflection) of complete signal*

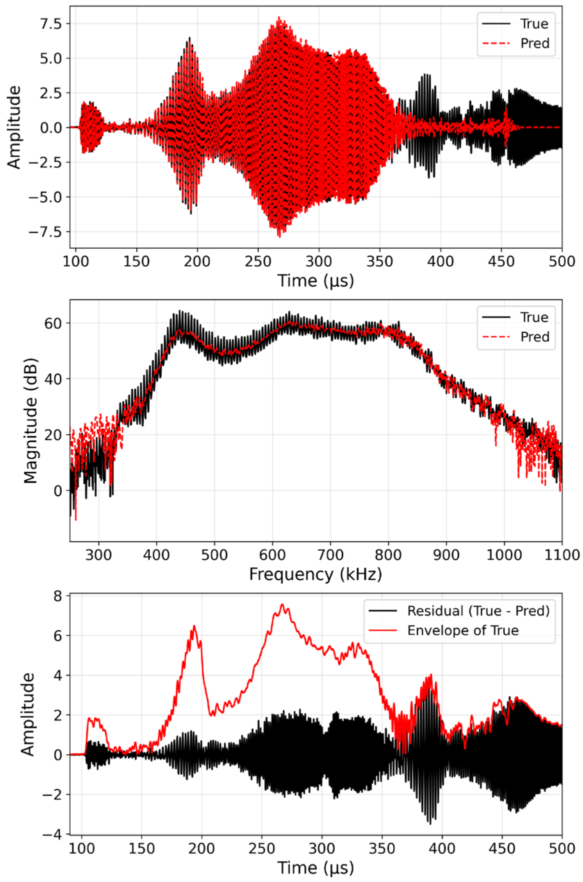

Fig. 6: Temporal (top), spectral (middle) and residual (bottom) response of $350\,\mu s$ *chirp signal, compared with the model (first reflection) of complete signal*

chirp signals, compared with $50\,\mu s$ (Fig. 4) the residual is a bit more, however it is important to note that $250\,\mu s$ was not in the training and the results are purely based on model prediction. Similarly, for the longest $350\,\mu s$ (Fig. 6), the visual results are also acceptable with minor losses of the modeling results.The residual part indicates that learning was successful for the first reflection region (Figs. 4–6 (bottom), under the envelope), while comparison with actual signal validates the proposed approach. Although for the long-duration signals there are some residuals, however, this is little sacrifice in order to achieve modeling.

Conclusion

Res-UNet1D based deep learning approach was proposed for the long-duration reflection overlapping in chirp signals used for air-coupled ultrasound. The results emphasize the importance of a deep learning architecture for modeling and predicting signals with significant overlap. The pre-processing, gating, scaling, normalization, and signal alignment ensured robust modeling of complex ultrasound reflections for experimental data. The close match between predicted and true signals in both the time and frequency domains highlights the model's capability to generalize across different excitation conditions and signal complexities.

Acknowledgment

This research was funded by a Grant No. S-MIP-23-133 from the Research Council of Lithuania.

References

[1] L. Fariñas et al. "Origin, development and applications of air-coupled broadband ultrasounds for the study of tissues and water relations in plant leaves: a review". In: *IEEE Open Journal of Ultrasonics, Ferroelectrics, and Frequency Control* 4 (2024), pp. 77–88. DOI: 10.1109/OJUFFC.2024.3433316.

[2] M. Tayyib and L. Svilainis. "Compensation function for enhanced bandwidth and improved SNR through programmable spectral shape signals". In: *2024 IEEE Ultrasonics, Ferroelectrics, and Frequency Control Joint Symposium (UFFC-JS)*. IEEE, 2024, pp. 1–5. DOI: 10.1109/UFFC-JS60046.2024.10793903.

[3] C. Bruckmann, S. Müller, and C. Zu Siederdissen. "Automatic, fast, hierarchical, and non-overlapping gating of flow cytometric data with flowEMMi v2". In: *Computational and Structural Biotechnology Journal* 20 (2022), pp. 6473–6489. DOI: 10.1016/j.csbj.2022.11.033.

[4] K. Dakic et al. "HybNet: a hybrid deep learning-matched filter approach for IoT signal detection". In: *IEEE Transactions on Machine Learning in Communications and Networking* 1 (2023), pp. 18–30. DOI: 10.1109/TMLCN.2023.3270131.

[5] G. Jajoo and P. Singh. "Modulation classification for overlapped signals using deep learning". In: *IEEE Open Journal of the Communications Society* (2024). DOI: 10.1109/OJCOMS.2024.3416750.

[6] R. Zhang et al. "A reference signal-aided deep learning approach for overlapped signals automatic modulation classification". In: *IEEE Communications Letters* 27.4 (2023), pp. 1135–1139. DOI: 10.1109/LCOMM.2023.3242690.