Program Digest, 2019 IEEE International Ultrasonics Symposium (IUS)
Glasgow, Scotland, October 6-9, 2019

A convolution neural network based machine learning approach for ultrasonic noise suppression with minimal distortion

1st Mohamed Elamir Mohamed
Comfort and driving assistance
Valeo Schalter und Sensoren
Bietigheim-Bissingen, Germany

2nd Heinrich Gotzig
Comfort and driving assistance
Valeo Schalter und Sensoren
Bietigheim-Bissingen, Germany

3rd Raoul Zoellner
Hochschule Heilbronn
Heilbronn, Germany

4th Patrick Maeder
Technische Universitaet Ilmenau
Ilemanu, Germany

***Abstract*—In this paper we present a novel machine learning approach for noise suppression in the signal generated by automotive industrial grade active ultrasonic sensors. A convolutional neural network (CNN) based machine learning approach is presented. State of art noise suppression methods are also discussed and used as benchmark against the proposed machine learning approach. The results of numerous simulated scenarios as well as actual sensor measurement campaigns are presented and discussed. Several metrics are derived to quantify the quality of the signal and give an indication of the performance of the different approaches of noise suppression. These derived metrics assess the performance of the different approaches in terms of amount of noise suppressed and amount of distortion introduced to the signal of interest.**

***Index Terms*—neural networks, ultrasonic, machine learning, noise suppression, denoising**

I. Introduction

Ultrasonic sensors are commonly used in the automotive industry for obstacle detection and environment perception. Ultrasonic-based systems are usually comprised of several sensors distributed around the vehicle. These sensors fire ultrasonic waves and report the time of flight (TOF) needed by an echo to bounce back off an obstacle in the vicinity of the vehicle in order to calculate the relative distance between the sensor and the obstacle. The TOF information from several sensors is then processed simultaneously using a form of triangulation to determine the exact position of the obstacle in a two dimensional map centered around the vehicle. This map is then used for several higher-end functionalities such as reporting these distances to the end user, automatic parking and braking on obstacles.

The quality of the ultrasonic signal plays an important role in the ability of the whole system to report correct information to the end user and to perform the higher-level driver assistance functionalities. In real life operation, there are several factors affecting the ability of the system to correctly identify the TOF of the ultrasonic reflected echo such as the presence of ultrasonic noise in the environment, the presence of other vehicles equipped with ultrasonic sensors in the vicinity of the vehicle and ground reflection from uneven terrain in the field of view of the sensor (1).

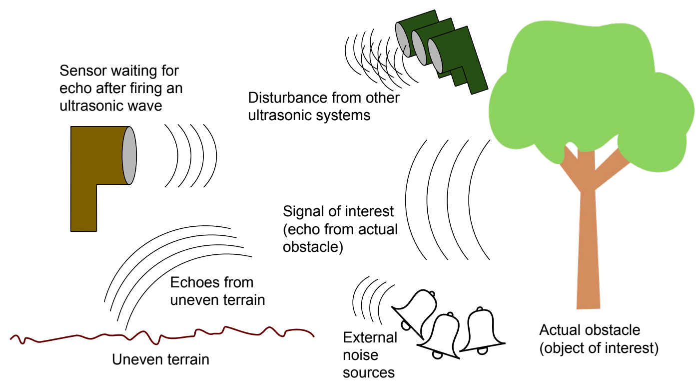

Fig. 1. Ultrasonic based system detecting an echo from an actual obstacle (tree) and being influenced by external disturbance sources such as noise in ultrasonic frequency range, other ultrasonic based systems and echoes from ground reflection

Several algorithms exist to suppress spurious ultrasonic signals and extract the signal of interest (SOI) and then consequently identify correctly the TOF. In this paper we present a novel approach to suppress noise in the measured signal as reported by an industrial grade automotive ultrasonic sensor based on machine learning (ML). Simulation results are presented showing the improvement achieved by this approach in comparison to the state of art algorithms. The results are furthermore consolidated by real life measurement campaigns and emphasized using several metrics of interest.

Initially the existing algorithms will be reviewed followed by a description of the simulation setup used to validate the results. Afterwards we present several classical methods for noise suppression and also a description of the proposed ML based approach and the results pertaining to it under the same simulation conditions. A description of the measurement campaign is then presented and the results for all the denoising algorithms is compared in terms of relevant metrics. We end by a discussion of the implications of these results and conclude.

II. Existing noise suppression algorithms

In this section we discuss the most commonly used denoising methods showing which is most suitable to the nature of ultrasonic signals coming from automotive grade ultrasonic sensors and which is most effective as well. We then discuss the drawbacks of such methodologies. Later on we show how this is addressed and mitigated in the proposed approach.

In recent years several denoising methodologies are found in commercial products such as noise cancellation headsets. This family of denoising algorithms perform active noise cancellation [1] to counteract spurious signals coming from external sources. They rely on the fact that the SOI is known and coming from a definite source such as an MP3 player. This is not applicable for automotive ultrasonic sensors because the signal is already mixed at the source and it is not known apriori which artifacts belong to the SOI and which are spurious and should be suppressed. Therefore we have to rely on other methodologies to identify the spurious noise sources, such as assuming the noise artifacts have only a certain range of frequency components or assuming that the SOI follows a specific model and all other signal components that do not fit this model should be suppressed.

A trivial method for removing noise from a signal is the moving average window technique (eq.1). This method depends mainly on the assumption that the SOI is only found in the low frequency range of the signal and thus the high frequency components are only contributing to noise. Therefore this part of the spectrum with the noise components is what is removed.

$$S_{out}[t] = \frac{1}{N} * \sum_{i=0}^{N} S_{in}[i+t] \tag{1}$$

Where $S_{in}[t]$ is the sampled input signal at a sampling rate higher than the Nyquist rate to avoid any loss of information, N is the size of the averaging window in terms of number of samples and $S_{out}[t]$ is the output from the denoising algorithm.

A more sophisticated method of achieving the same target of suppressing the high frequency noise components given the same assumptions, is using a low pass filter (LPF) (eq.2).

$$S_{out}[n] = \sum_{i=0}^{N} b_i * S_{in}[n-i] \tag{2}$$

Where S_{out} and S_{in} play the same role as in the previous equation and b_i are the weights of the LPF.

This method offers the possibility to tune the filter parameters to achieve minimum distortion of the low frequency components and also to achieve regulated suppression of the high frequency components. This comes with the cost of having longer filters, thus adding to the complexity of the algorithm and number of computations needed to filter the signal and thereby increasing the overall runtime of the algorithm. Given the realtime processing restrictions present in ultrasonic based driver assistance systems (ADAS) the optimum filter design is not always possible.

The assumption that the noise is only in the high frequency part of the signal does not hold in case of ultrasonic sensor signals. Other methods such as adaptive multistage noise suppression filters [2] [3] are more adequate to the problem at hand. Such filters are by nature complex and include a large number of computations in subsequent steps which also lead to increased runtime and failure to satisfy the realtime constraints.

A family of algorithms that require relatively lower processing power and perform denoising over the whole spectrum are discrete wavelet transform (DWT) based methods. The DWT approach described in [6] and [7] achieves good denoising results in terms of suppressing unwanted spurious signals and extracting the SOI. It comes though with the cost that the extracted SOI is distorted, since it is the result of reconstructing the superimposed wavelets passing the designed noise thresholds. This is not suitable for the ultrasonic sensor signal because the SOI is further processed to extract features pertaining to the reflecting obstacle such as height and class of the obstacle which is then fed into the higher layers of the ADAS for end-user functionalities such as braking on obstacles and automatic parking. Thus, the need arises to have a denoising algorithm that achieves similar noise suppressing results while maintaining the integrity of the SOI.

In the following section we present the proposed machine learning based approach. Further, we present a simulation setup that shows initial results of how this proposed method competes with the existing algorithms in terms of noise suppression and outperforms them with regards to maintaining signal integrity and introducing minimal distortion into the signal of interest.

III. The proposed machine learning based approach

In literature there exist several publications discussing the use of machine learning and specifically deep neural networks for signal processing. Here we are interested in the family of algorithms that deal with the separation of signal components. Some approaches focus on the extraction of a specific SOI and discard the rest while other approaches consider all the signal components to be of interest and separates the different components such as [4] and [5]. There, the voice of the singer and the background music are both of interest and the algorithm tries to separate them without compromising signal integrity. The problem with denoising ultrasonic sensor signals is more in the domain of the former family of algorithms, where we only care about the echo from the obstacle and discard all other signal components.

We present here a novel approach to use deep CNNs to perform noise suppression and extract the SOI from an ultrasonic signal mixture containing an echo from an actual obstacle and different types of environment noise in the same spectrum as the echo signal. The main concept of the approach is that the whole measurement is fed into the CNN and the denoised signal is completely regenerated at the output layer of the network. The network is trained on noisy signals with different types and levels of noise using supervised learning. The label at the output layer for each training sample is the same version of the input signal but with no noise component and only the echo features are present.

The structure of the network is the classical hour glass shape with the difference being that it is purely convolutional with no dense layers. The network is compressed using downsampling and decompressed using upsampling as described in (2). The number of layers is optimized to be 7 layers and the number of activation maps is set to a maximum of 128 kernels at the most compressed layer located in the middle of the neural network. The hyperbolic tangent (tanh) is used as activation function to limit the maximum values of the firing of the neurons, especially at the output layer where an extreme value could lead to the occurrence of false positives and unexpected behaviour. The output and the input layers are of the same size as the number of samples present in one measurement from the sensor.

In the following sections we show the results of using this approach on simulation data and actual measurements. we benchmark the results against state of the art algorithms for noise suppression.

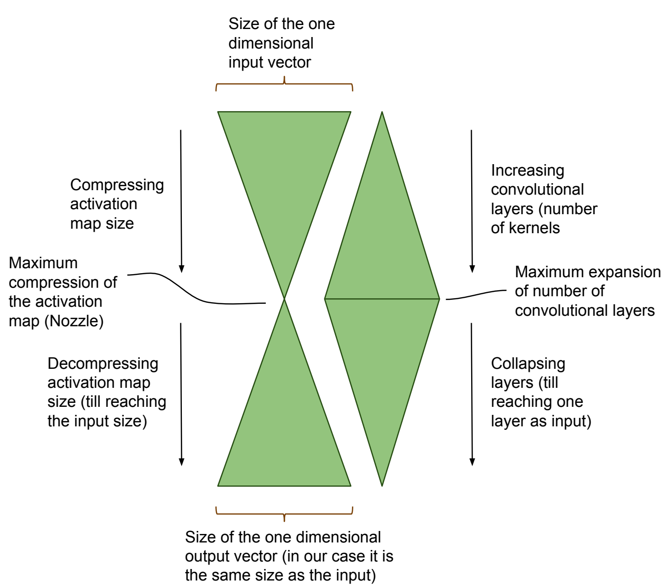

Fig. 2. The convolutional neural network structure of the proposed machine learning approach

IV. Simulation model setup

The natural reverberation frequency of the membrane of the automotive grade sensor used is 51.2 KHz. We generate a carrier wave having this resonance frequency. This carrier wave is then used to modulate an empirical ultrasonic echo shape that was derived based on numerous field examination of the features of the ultrasonic wave reflected off several obstacle classes. We see in eq.(3) how this process is defined.

$$A_m(t) = A_e(t) * A_c sin(2\pi f_c t) \tag{3}$$

Where $A_m(t)$ is the modulated signal, $A_e(t)$ is the base-band empirical echo envelope shape, A_c is the amplitude of the carrier frequency. f_c is the carrier frequency and is set to 51.2 KHz.

On top of this modulated signal we add two components one for noise and one for ground reflections. Both components are superimposed in the modulated state of the simulated signal to account for constructive and destructive interference effects. For the noise components, we generate them using an additive Gaussian white noise (AGWN) model. The power of this added noise signal is then varied to asses the performance of the different suppression mechanisms under different signal to noise ratio (SNR) values. The ground reflection components simulate the weaker reflections coming from uneven terrain. The empirically generated ground reflection signal is done by superimposing a large number of weak echoes having uniformly distributed scatter and a Gaussian distributed range of amplitudes.

The final generated signal is then demodulated using a carrier sinusoidal wave having the same frequency and the resulting envelope curve is then decimated to 200 samples covering a range of 10 ms of ultrasonic wave to resemble the signal output from the actual ultrasonic sensor. From the approximation equation of acoustical wave propagation in air (eq.4), we determine the speed of the ultrasonic wave (eq.5). Accordingly we know that the 10 ms duration is equivalent to a range of 172 cm at room temperature (20°C).

$$S = 331.3 + (0.606 * temp°C) \tag{4}$$

$$S = 343.42\text{m/s} \tag{5}$$

Where S is the speed of the ultrasonic wave propagation in air.

A validation set with a large number of samples is generated using this simulation model covering a range of SNR values. This validation set is then used to assess the performance of the proposed noise suppression algorithm and benchmark it against the DWT method.

In (3) we see an example of applying the DWT based noise suppression algorithm. We see that most of the noise artifacts are completely suppressed but when zooming in on the SOI segment (4) we see that the SOI is deformed, which is the main drawback of this method.

The DWT algorithm used in this example is based on the Daubechies wavelet, as it is suitable for the nature of the ultrasonic sensor signal. We also use a first-level transform to keep the number of computations at a suitable level for realtime constraints. We employ the universal threshold (eq.6) for suppressing the undesired spurious artifacts in the signal. This step includes the soft clipping of the coefficients falling below the threshold.

$$T_u = \frac{\sqrt{2 * \log(length(X)) * median(abs(D))}}{0.6745} \tag{6}$$

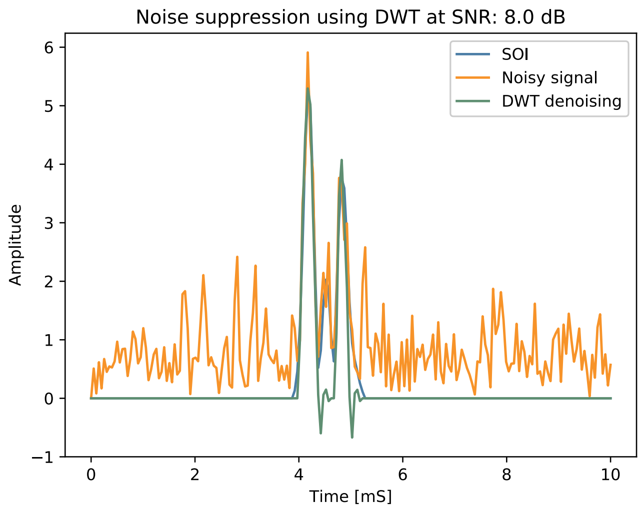

Fig. 3. Applying descrete wavelet transform denoising method on a noisy signal simulated to resemble the signal from an automotive ultrasonic sensor

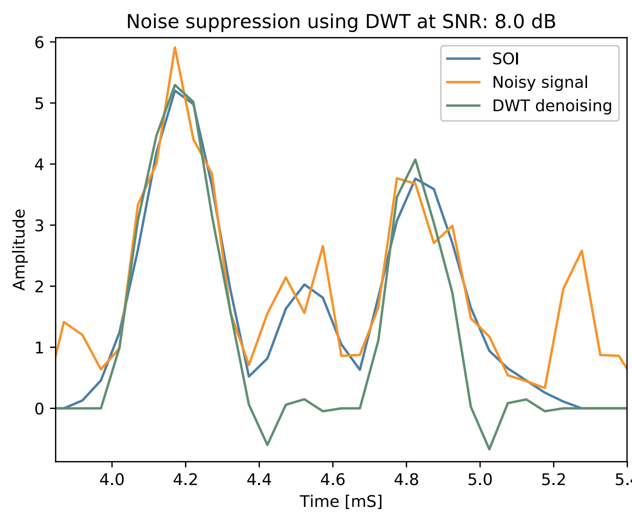

Fig. 4. Applying discrete wavelet transform denoising method on a noisy signal simulated to resemble the signal from an automotive ultrasonic sensor - zooming on the signal of interest which is the simulation of a reflection from an obstacle in the vicinity of the sensor

where T_u is the empirically calculated universal threshold.

In (5) we see an example of noise suppression performed by the proposed machine learning approach over the same signal presented earlier with the DWT approach. The neural network structure is trained over 80% of the simulation data available at all the SNR levels generated. It is then validated on 10% of the samples after each epoch and tested over the remaining 10% of the samples from which this example is chosen. The results shown throughout the paper are only based on test samples that the neural network never had as input during its training phase.

From (5) we see that the noise was suppressed to levels

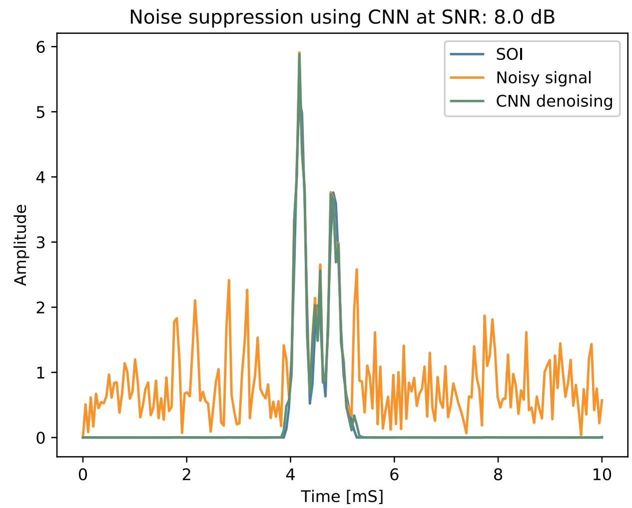

Fig. 5. Applying machine learning based denoising method on a noisy signal simulated to resemble the signal from an automotive ultrasonic sensor

comparable to the DWT method. By further zooming in (6) we see that the SOI retains much more of its original features.

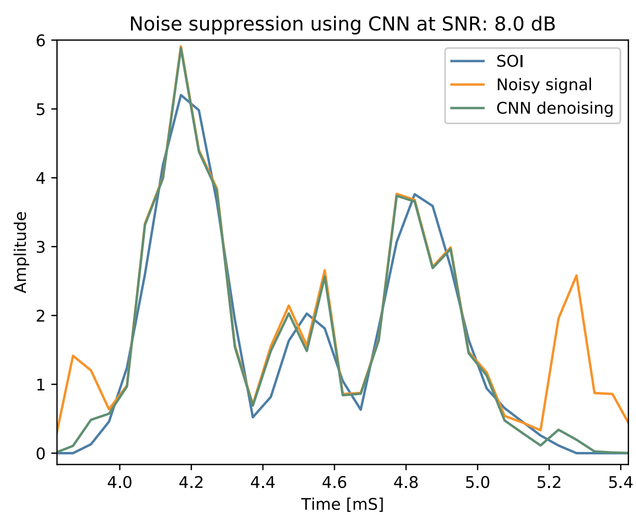

Fig. 6. Applying machine learning based denoising method on a noisy signal simulated to resemble the signal from an automotive ultrasonic sensor - zooming on the signal of interest which is the simulation of a reflection from an obstacle in the vicinity of the sensor

In (7) we see the performance of both algorithms in terms of suppressing noise presented over a range of SNR values [-5, 20] dB. The DWT performs slightly better in this regard. Nevertheless the main advantage obtained from using the machine learning based approach is that it introduces much less distortion compared to DWT as will be demonstrated.

To adequately judge the performance of the algorithm we take the changes in signal energy not originally present in the SOI as a measure of distortion. This will include added

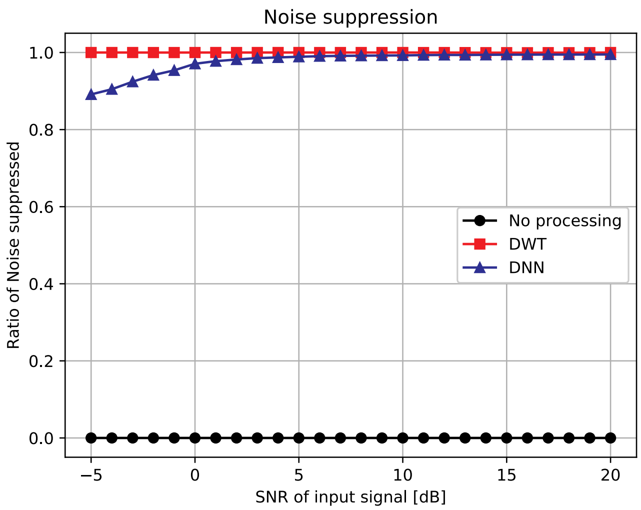

Fig. 7. Plotting the ratio of the noise suppressed to the original noise power present in the signal without any processing and after applying both the discrete wavelet transform denoising method and the machine learning based denoising method

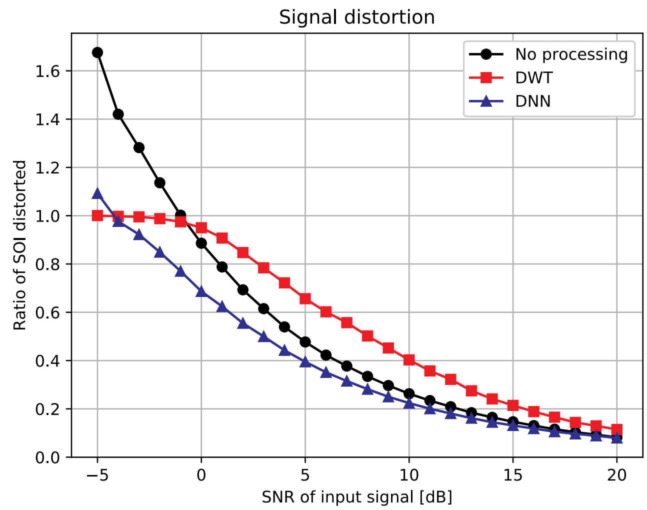

Fig. 8. The amount of distortion introduced to the signal of interest without processing and after applying the discrete wavelet transform denoising method and the machine learning based denoising method

artifacts as well as removed components from the original SOI. This approach will also prevent the trivial solution where the output of the algorithm is an array of zeroes where the noise is completely suppressed but the SOI is also lost. In (8) we show a comparison between the DWT based method and the machine learning based method in terms of the defined distortion metric. We see that even though the 2 algorithms have comparable performance in terms of noise suppression, the machine learning based approach outperforms the DWT method in terms of conserving the SOI integrity and in terms of the introduced distortions to the original shape of the SOI. We also see that at low SNR levels, the DWT distortion saturates at a ratio of 1.0, indicating that the SOI is completely suppressed and the algorithm is not capable of differentiating between the SOI and the noise components of the signal. In contrast, the machine learning approach is still capable of extracting the SOI from the noisy input signal.

It is important to mention that with further optimization of the hyper parameters of the legacy denoising algorithms we could achieve better results over the available trace set. Consequently the performance could improve but only to a certain limited extent. This is manifested in the trade-off between the noise suppression capability and the distortion introduced to the signal.

In the next section we present the measurement campaign we carried out. We show that the results from this measurement campaign are in line with the simulation results, which emphasizes further the superiority of the machine learning based approach over the legacy algorithms in real life scenarios.

V. Measurements and Results

In the simulation setup, it is easy to know the SOI with no added noise on one hand and the amount of noise introduced to the measurement on the other hand as it is synthetically generated. With real life measured data, it is more difficult to extract this information as the echo SOI and the environmental noise are already superimposed at the entry point of the data acquisition setup. Therefore we perform every measurement twice for this measurement campaign. Once with a noise source and once without. The measurement done with no noise source will be considered to be the SOI. This defined SOI is, of course, not really only the SOI as there is an added noise floor coming from the environment and the measuring equipment. However, in a controlled environment, and compared to the magnitude of the noise added by the noise source introduced in the noisy measurement, this noise factor in the SOI can be neglected.

A measurement campaign is carried out to record the echo coming from several obstacles including boxes, poles, car fronts, car sides, bushes and pedestrians set at different distances from the recording sensor. Each measurement is repeated several times, once with no noise source and then with several noise sources including clinging keys, truck brakes, rain, air gun and modified semaphores for the visually challenged. The ensemble of the set of traces with the noise source are used as training, validation and testing sets for the machine learning based approach. The set of scenarios with no noise sources are used as labels. The KPIs in terms of noise suppression and distortion to the SOI are measured using the noisy scenarios and the corresponding non noisy scenarios as SOI.

For the extraction of the KPIs we use the complete set of measurements for the DWT based method and we use 80% of the shuffled measurements for training the machine learning based approach, 10% for validation and the remaining 10% for testing and setting the KPIs. It is important to mention that for the extracted KPIs the samples used were never introduced to the neural network during the training phase, and thus have no influence on the evolution of the kernel weights of the neural network.

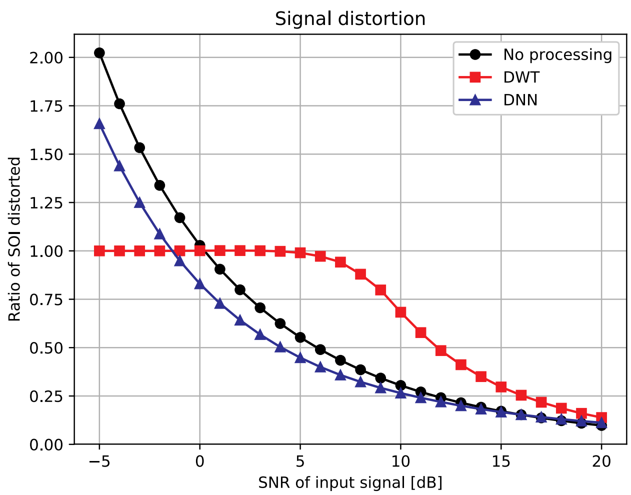

Fig. 9. The amount of distortion introduced to the signal of interest without processing, after applying the discrete wavelet transform denoising method, and after applying the machine learning based denoising method using real measured traces

The DWT approach and the ML based approach are applied on the recorded real measurements. In (9) it is evident that for the SOI distortion, the levels of distortion are higher than the levels reported during the simulation. This is attributed to the fact that the echo shapes are not generated by a simulation model rather the shape of the echo varies in a realistic manner as the data is from real measurements. This proves to be more difficult for the DWT approach to recreate the exact echo from weighted wavelets. It also proves to be more challenging for the machine learning based approach to learn the shape of the echo and generalize in a manner sufficient to recreate the SOI as a congruent denoised signal at the output layer. Another factor is that there is a minor neglected noise component because of the measuring equipment and the surrounding environment that - although minimal - still could not be entirely eliminated.

With these considerations the performance of the ML based approach still proves to be superior over the complete SNR range and follows the same pattern as predicted by the simulation model.

VI. Discussion and Conclusion

It is clear from the results of the simulations and measurements campaigns that the ML based approach is comparable to the DWT approach in terms of noise suppression, and is superior in terms of the levels of distortion introduced to the SOI.

The superior performance is explained by the fact that the ML based approach learns the typical shape of a SOI as well as the different noise patterns from the numerous measurements used to train the CNN, unlike the DWT algorithm which fits the SOI (the echo in this case) to the base wavelet shape.

By employing our algorithm the integrity of the SOI is maintained while simultaneously suppressing the unwanted spurious noise artifacts. This leads to higher quality signal information and better functionality for the driver assistance systems relying upon these ultrasonic sensors.

There are many areas where ML could be employed as supplementary algorithms to provide degrees of confidence and to further assist existing deterministic algorithms. In other domains such as the noise suppression for ultrasonic sensor signals that is presented in this work, the ML approaches could replace existing methods and provide better performance in terms of achieving the main target of the algorithm while introducing less distortion to the SOI and conserving the integrity of the processed signal. In general, ML based signal processing approaches show very promising results with superior KPIs, compared to existing state of art algorithms. They provide potential for improvement in many aspects of signal processing such as signal conditioning, filtering of unwanted artifacts and extraction of useful information and features.

References

[1] Elliott, Stephen J., and Philip A. Nelson. ”Active noise control.” IEEE signal processing magazine 10, no. 4, 1993: 12-35.
[2] Ikeda, Shigeji, and Akihiko Sugiyama. ”An adaptive noise canceller with low signal distortion for speech codecs.” IEEE Transactions on Signal Processing 47, no. 3, 1999: 665-674.
[3] Qian, Wei, and Laurence P. Clarke. ”Adaptive multistage nonlinear filtering and wavelet for medical image enhancement.” In Proceedings of 1st International Conference on Image Processing, vol. 3, pp. 711-715. IEEE, 1994.
[4] Chandna, Pritish, Marius Miron, Jordi Janer, and Emilia Gmez. ”Monoaural audio source separation using deep convolutional neural networks” In International conference on latent variable analysis and signal separation, pp. 258-266. Springer, Cham, 2017.
[5] Uhlich, Stefan, Marcello Porcu, Franck Giron, Michael Enenkl, Thomas Kemp, Naoya Takahashi, and Yuki Mitsufuji. ”Improving music source separation based on deep neural networks through data augmentation and network blending.” In 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 261-265. IEEE, 2017.
[6] Lang, Markus, Haitao Guo, Jan E. Odegard, C. Sidney Burrus, and Raymond O. Wells. ”Noise reduction using an undecimated discrete wavelet transform.” IEEE Signal Processing Letters 3, no. 1 (1996): 10-12.
[7] Lang, Markus, Haitao Guo, Jan Erik Odegard, C. Sidney Burrus, and Raymond O. Wells Jr. ”Nonlinear processing of a shift-invariant discrete wavelet transform (DWT) for noise reduction.” In Wavelet Applications II, vol. 2491, pp. 640-651. International Society for Optics and Photonics, 1995.