Interspeech 2025
17-21 August 2025, Rotterdam, The Netherlands

Individualized speech enhancement for hearing-impaired listeners

Chuan Wen, Sarah Verhulst

Hearing Technology @ WAVES, Dept. of Information Technology, Ghent University, Belgium

chuan.wen@ugent.be, sarah.verhulst@ugent.be

Abstract

Despite significant progress in speech enhancement (SE), improving speech intelligibility and perceptual quality in noisy environments for hearing-impaired individuals remains challenging. This paper presents an individualized speech-enhancement (ISE) framework that integrates noise reduction (NR) and sound amplification for hearing loss compensation within a unified system. Our fully differentiable, closed-loop design incorporates biophysically realistic auditory models. The framework features two pathways: one simulating the auditory response of a normal-hearing (NH) system to denoised speech and the other modeling a hearing-impaired (HI) system's response to noisy speech. The ISE model is trained by minimizing the difference between NH and HI auditory responses. Experimental results show that the ISE model enhances speech intelligibility and perceptual quality in noisy conditions for HI listeners, offering a promising foundation for advancing personalized noise reduction strategies.

Index Terms: Individualized speech enhancement, hearing-aid processing, closed loop, deep neural network

1. Introduction

Speech enhancement (SE) plays a crucial role in speech communication by improving the perceptual quality and intelligibility of speech signals degraded by noise, reverberation, and distortion. Enhancing speech quality benefits a wide range of applications, including automatic speech recognition (ASR), speaker identification, and hearing-aid design.

Recent advancements in deep neural networks (DNNs) have significantly improved SE performance. However, most SE models are developed under the assumption that listeners have normal hearing [1, 2, 3, 4]. Research has shown that DNN-based SE models provide only limited speech intelligibility improvements for individuals with hearing impairment [5]. Hearing loss substantially degrades auditory perception, making it difficult for affected individuals to understand speech in noisy environments. Conventional SE models alone are insufficient for improving speech intelligibility for this population, as simple noise reduction does not address the underlying auditory deficits. Amplification is often necessary to compensate for the reduced sensitivity and impaired speech perception in hearing-impaired listeners.

Despite progress in DNN-based hearing-aid technologies [6, 7, 8, 9], two key challenges remain. First, current hearing aids primarily compensate for outer-hair-cell (OHC) loss and elevated hearing thresholds but do not adequately address other aspects of sensorineural hearing loss (SNHL), such as cochlear synaptopathy (CS), which involves damage to auditory nerve synapses. Recent studies have introduced differentiable frameworks that aim to compensate for both OHC and CS-related hearing deficits [10, 11], but improvements in speech intelligibility in noisy environments remain marginal [10]. Although a closed-loop hearing-aid model was proposed recently [12] that integrates noise reduction and amplification, it did not account for CS-related deficits. Second, existing approaches typically implement noise reduction and amplification as separate modules, leading to increased computational complexity and processing latency. The requirement for multiple deep learning models can make real-time processing in hearing devices impractical. Addressing these challenges requires an integrated approach that simultaneously performs noise reduction and hearing loss compensation within a unified framework.

To address these challenges, we propose a framework for designing individualized speech enhancement (ISE) models that integrate noise reduction and hearing loss compensation within a unified system. The framework is built upon dCoNNear [11], a DNN-based biophysical model of human auditory processing. A key feature of this framework is its adaptability to an individual's specific hearing impairment profile. Since ambient sound plays a crucial role in environmental awareness, we introduce a coefficient parameter to regulate the level of denoising thereby ensuring an optimal balance between noise suppression and situational awareness. The ISE models are optimized by leveraging biophysical differences in auditory processing between normal-hearing and hearing-impaired pathways. The main contributions of this paper are as follows: 1) developing a biophysically-inspired framework for individualized speech enhancement that integrates noise reduction and sound amplification for hearing-impaired listeners in a single DNN-based model, 2) evaluating the effectiveness of the proposed approach in restoring speech intelligibility and quality in noise for individuals with OHC loss and CS, and 3) examining the impact of varying noise reduction levels on perceptual quality and intelligibility.

2. Framework

The closed-loop framework for designing individualized speech enhancement (ISE) models is illustrated in Figure 1. It consists of two pathways: one representing the auditory nerve (AN) response r_f of a normal-hearing (NH) auditory system to a denoised speech signal, and the other representing the response $\hat{r_f}$ of a hearing-impaired (HI) system to a noisy speech signal. The framework is based on dCoNNear [11], which simulates auditory processing stages for both pathways, including cochlear, inner hair cell (IHC), and auditory nerve fiber (ANF) responses. The ISE model takes two inputs: a noisy speech signal x and a coefficient parameter, called the "target SNR," which controls the level of noise reduction. The enhanced speech output $\hat{y}$ from

 10.21437/Interspeech.2025-1009

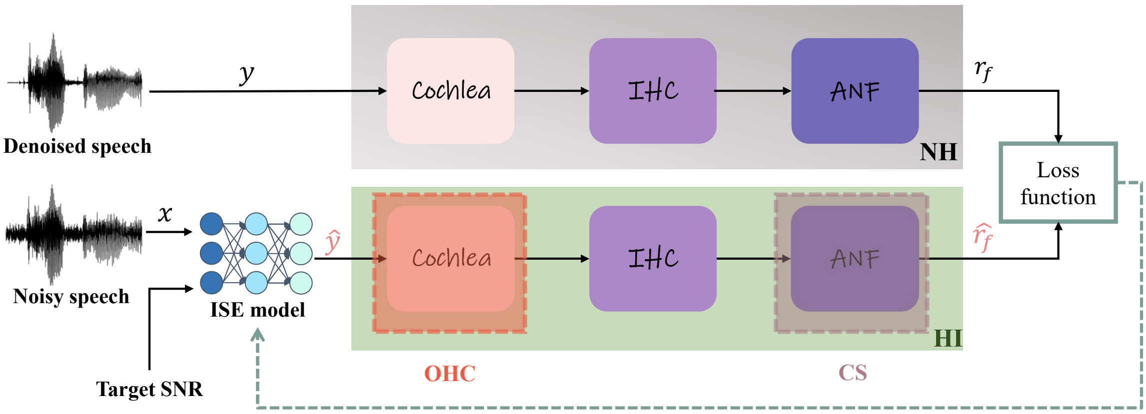

Figure 1: *Diagram of the closed-loop framework for designing individualized speech-enhancement (ISE) models.*

the ISE model is then processed through the HI pathway, while the target enhanced speech samples y are sent through the NH pathway. The ISE model is trained by minimizing a predefined loss function that quantifies the difference between the simulated NH and HI responses, ensuring that the enhancement process aligns with the auditory characteristics of normal hearing.

2.1. Individualized speech enhancement model

The ISE model is built upon the dCoNNear architecture [11], which consists of a sequence of stacked FIR-like memory blocks with exponentially increasing dilation factors [13]. This design enables the network to capture long-term dependencies in the speech signal. Specifically, M memory blocks with dilation factors 2^{m-1} are repeated R times, resulting in a total of $N = M \times R$ blocks. To facilitate learning and mitigate the vanishing gradient problem, residual connections are incorporated between the memory blocks. Additionally, skip connections are introduced to aggregate outputs from each memory block through a weighted summation. The aggregated values are then passed through a non-linear activation function and a convolutional layer to produce the final model outputs. The hyperparameters of the ISE model are summarized in Table 1.

Table 1: *Hyperparameters of the ISE model. K represents the kernel length of the 1D dilated convolutional layer in each memory block, while H denotes the number of channels in the hidden layers.*

M	6
R	2
K	32
H	200
Activation	PreLu
Param. (million)	1.5
GMACs	9.72

2.2. Individualisation of auditory elements

Both normal-hearing (NH) and hearing-impaired (HI) auditory peripheries utilize biophysically inspired CNN-based models to accurately simulate human cochlear, inner hair cell (IHC), and auditory nerve fiber (ANF) processing. These models include $\text{dCoNNear}_{\text{cochlear}}$, $\text{dCoNNear}_{\text{IHC}}$, and $\text{dCoNNear}_{\text{ANF}}$ [11]. To individualize the hearing-impaired auditory elements, the NH auditory modules were modified to simulate varying degrees of outer hair cell (OHC) loss and cochlear synaptopathy (CS) in the HI pathway. OHC loss was modeled by retraining the NH cochlear model using transfer learning, incorporating a specific gain-loss profile or an individual audiogram [14]. The NH auditory nerve (AN) response (r_f) was computed as a weighted sum of the three ANF responses, with parameters H_{NH} = 13, M_{NH} = 3, and L_{NH} = 3 representing the distribution of ANF types. To simulate CS, the model was adjusted by modifying the weights H_{HI}, M_{HI}, and L_{HI} accordingly [10].

3. Training

The training procedure involved two main stages: first, individualizing the auditory modules to match specific hearing-impaired profiles, and second, training the ISE model within the closed-loop framework to minimize the difference between NH and HI auditory pathways. In this study, two types of hearing loss were simulated. The first, “flat30” represents a uniform 30 dB loss across all frequency bands. The second, “CS7,0,0” , simulates cochlear synaptopathy (CS), characterized by a complete loss of low spontaneous rate (LSR) and medium spontaneous rate (MSR) auditory nerve fibers (ANFs), with H_{HI} = 0 and M_{HI}= 0, along with an approximate 46% loss of high spontaneous rate (HSR) ANFs, where L_{HI} = 7. Once the auditory elements were individualized, the ISE models were optimized within the closed-loop system, while the auditory module parameters remained fixed throughout training.

3.1. Dataset

Step 1: To individualize auditory models, a subset of 2,860 randomly selected utterances from the TIMIT speech corpus [15] was employed to train cochlear models corresponding to specific hearing profiles. The dataset was partitioned into 2,310 samples for training and 550 for validation. The training was supervised using target outputs derived from biophysical models of cochlear processing [16], which simulates basilar membrane (BM) vibrations, inner hair cell (IHC), and auditory nerve fiber (ANF) at 201 center frequencies spanning 100 Hz to 12 kHz.

Step 2: To train the ISE models based on the closed-loop framework shown in Figure 1, we used the INTERSPEECH 2021 DNS Challenge dataset [17, 18], which consists of speech recordings from 11,350 speakers and over 600 diverse noise scenarios. For noise dataset, Audioset [19] was used for training, while Freesound [20] was utilized for testing. Noisy speech samples were synthesized by randomly selecting noise segments and mixing them with clean utterances at signal-to-noise ratios (SNRs) ranging from -5 dB to 5 dB, with 1 dB increments. The target dataset, representing denoised speech, was generated by applying a randomly chosen target SNR within the range of 10 dB to 30 dB. The dataset used for model training comprised 100 hours of speech, with an additional 5 hours reserved for validation. For performance evaluation, a test set of 1 hour was generated for each SNR condition (-5 dB, 0 dB, and 5 dB). The corpus was upsampled to 20 kHz to align with the frequency range of the human hearing model in the closed-loop system, and normalized to 70 dB sound pressure level (SPL) with reference $p_0 = 2 \times 10^{-5}$.

3.2. Training configurations

To optimize cochlear model individualization, we adopted the training methodology from [11], using mean absolute error (MAE) as the optimization criterion. The trained dCoNNear-based modules simulate cochlear responses across 201 center frequencies (100 Hz–12 kHz). For ISE model training, we selected 21 equally spaced frequency channels to improve efficiency and optimized the model by minimizing the mean squared error (MSE) between normal-hearing (NH) and hearing-impaired (HI) auditory nerve responses. The framework was implemented in PyTorch using a fixed learning rate of 0.0001 and the Adam optimizer [21].

4. Evaluation

To evaluate the effectiveness of the proposed ISE models, we conducted a comparative analysis against a generic noise reduction (NR) model and a closed-loop hearing-aid (HA) model. The models are defined as follows:

1) NR: A generic noise reduction model based on the proposed ISE model but trained independently, without incorporating the closed-loop framework. This model was optimized using the widely adopted scale-invariant signal-to-distortion ratio (SI-SDR) loss function [22].
2) HA: A closed-loop hearing-aid model derived from [11], designed to compensate for hearing loss profiles “flat30” ($\text{HA}_{\text{flat30}}$) and “CS7,0,0” ($\text{HA}_{\text{CS700}}$).
3) ISE: The proposed ISE model, which conducts noise reduction and compensates for “flat30” ($\text{ISE}_{\text{flat30}}$) and “CS7,0,0” ($\text{ISE}_{\text{CS700}}$) simultaneously.

Evaluation was conducted using a 1-hour test set generated from the DNS Challenge dataset, with an SNR of 0 dB at 70 dB SPL. For the “flat30” OHC loss condition, speech quality was assessed using the Hearing-Aid Speech Quality Index (HASQI) version 2, as proposed by Kates and Arehart [23]. Speech intelligibility was evaluated using the Hearing-Aid Speech Perception Index (HASPI) version 2 [24].

To evaluate the performance of ISE models for “CS7,0,0”, the effectiveness of the simulated restoration was quantified using three objective metrics: speech-to-reverberation modulation energy ratio (SRMR) [25] and normalized root-mean-square error (NRMSE) [10]. The NRMSE is defined as: Let y_i and $\hat{y}_i$ denote the population responses of the NH and HI auditory responses, where

$$y_i = \sum_{j=1}^{K} r_{i,j}, \quad \hat{y}_i = \sum_{j=1}^{K} \hat{r_f}_{i,j} \tag{1}$$

We then compute the root-mean-square error

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\big(y_i - \hat{y}_i\big)^2}, \tag{2}$$

and normalize it by the maximum true value:

$$\text{NRMSE} = \frac{\text{RMSD}}{\max\limits_{1 \le i \le N} y_i}. \tag{3}$$

Here, i indexes the time frames, j indexes the center frequencies, N is the total number of frames, and K is the number of frequency channels.

5. Results

The performance of the ISE models was evaluated for outer hair cell (OHC) loss “flat30” and cochlear synaptopathy (CS) “CS7,0,0”, focusing on sound quality and speech intelligibility. The evaluation metrics were computed for each test sample and subsequently averaged across the test dataset derived from the DNS Challenge.

5.1. Our-hair-cell loss

Table 2 presents HASPI scores for different methods under varying denoising levels, evaluated for simulated normal-hearing and hearing-impaired listeners with “flat30” OHC loss. For simulated normal-hearing listeners, the generic NR model exhibited increased HASPI scores across all denoising levels, with scores increasing as the target SNR decreased. This trend suggests that reducing noise suppression levels can enhance speech intelligibility. For simulated hearing-impaired listeners with “flat30” OHC loss, while the generic NR model also demonstrated improved HASPI scores compared to unprocessed samples when target SNRs were below 20 dB, the HA model outperformed the NR model for all denoising levels. The ISE model, particularly at a target SNR of 10 dB, achieved the highest performance, yielding approximately a 10% improvement in intelligibility compared to the unprocessed system. The performance of the ISE model declined as noise reduction increased, with intelligibility scores decreasing at higher target SNRs. However, when the target SNR was below 20 dB, the ISE model outperformed the HA model. These findings indicate that selecting an optimal denoising level is crucial for the ISE model, as excessive noise suppression may inadvertently degrade speech components, leading to reduced intelligibility.

Table 3 presents HASQI scores for different models, evaluated for both simulated normal-hearing and hearing-impaired listeners with “flat30” OHC loss. For simulated normal-hearing listeners, the generic NR model consistently increased HASQI scores across all target SNR levels, with higher denoising levels yielding higher scores. This trend suggests that increasing noise reduction improves perceived sound quality. For simulated hearing-impaired listeners with “flat30”, the NR model demonstrated an improvement in HASQI scores compared to unprocessed samples, with scores increasing as the denoising level increased. However, the HA model showed lower (worse) HASQI scores than unprocessed speech. This outcome can be

Table 2: *Comparing the HASPI scores (%) of different methods for simulated normal-hearing and hearing-impaired listeners with target SNRs from 10 to 30 dB. Higher scores correspond to better performance.*

Loss types	Models	10 dB	15 dB	20 dB	25 dB	30 dB
NH	Noisy			82.9		
	NR	88.84	86.10	85.58	84.92	84.23
HI flat30	Noisy			73.19		
	NR	76.34	75.87	73.69	72.16	71.17
	HA			77.35		
	ISE	82.46	79.16	76.73	74.52	73.22

attributed to the compressive amplification characteristics of the HA model, which adjust gain dynamically by amplifying quiet sounds while compressing louder ones. As a result, background noise becomes overamplified, degrading overall speech quality. The ISE model exhibited improved HASQI scores relative to the unprocessed system, with scores rising as the target SNR increased. Notably, the ISE model outperformed the NR model when target SNRs exceeded 20 dB. These findings indicate that an optimal target SNR is critical for maximizing speech quality for hearing-impaired listeners.

Table 3: *Comparison of the HASQI scores (%) of different methods for simulated normal-hearing and hearing-impaired listeners with target SNRs from 10 to 30 dB. Higher scores correspond to better performance.*

Loss types	Models	10 dB	15 dB	20 dB	25 dB	30 dB
NH	Noisy			18.82		
	NR	24.02	25.73	27.86	28.82	30.61
HI flat30	Noisy			21.65		
	NR	30.59	31.38	31.45	32.95	33.78
	HA			18.1		
	ISE	26.56	27.87	31.49	33.88	34.65

5.2. Cochlear synaptopathy

Table 4 presents the NRMSE and SRMR scores for cochlear synaptopathy “CS7,0,0”, a condition for which most conventional hearing aids lack targeted treatments. Across both metrics, the ISE model (target SNR = 10 dB) achieved the best performance among all compared systems. Additionally, the HA model exhibited lower NRMSE scores than the unprocessed samples, consistent with findings from [10], which reported that the HA model performed worse in terms of NRMSE for speech-in-noise. This result may be attributed to the HA model’s tendency to overamplify background noise, which disrupts the restoration of the auditory nerve (AN) population. A similar pattern was observed in the NR model, likely because it focuses solely on noise reduction without addressing AN population restoration, thereby limiting its effectiveness. These findings highlight the ISE model’s potential to effectively compensate for “CS7,0,0” in speech-in-noise scenarios.

5.3. Disscussion

In this study, the ISE model was trained using the MSE between the NH and HI outputs. However, given the complex interaction between denoising and amplification in the ISE model, as well as the highly nonlinear nature of auditory processing [10], exploring more effective loss functions could lead to more precise hearing loss compensation and improved sound quality. Future work will investigate loss functions that better capture the perceptual characteristics of impaired hearing.

A key feature of the proposed framework is the introduction of the target SNR parameter, which regulates the level of noise reduction. The results indicated that increasing the denoising level led to a decline in intelligibility scores while improving speech quality. This trade-off suggests that excessive noise suppression may compromise essential speech components, thereby reducing intelligibility. Additionally, ambient sounds play a vital role in helping listeners maintain awareness of their surroundings. To further investigate this balance, a listening test will be conducted to examine how users select the optimal target SNR, considering intelligibility, sound quality, and environmental perception.

Table 4: *Comparing the performance of different methods for CS using the objective measures NRMSE and SRMR. For SRMR, higher scores correspond to better performance and lower scores correspond to better performance for NRMSE.*

	Noisy	HA	NR	ISE (10 dB)
NRMSE (%)	22.35	20.88	25.63	19.26
SRMR	4.91	4.52	8.82	6.94

6. Conclusion

This study proposed a framework for designing individualized speech enhancement (ISE) models that integrate both noise reduction and sound amplification. The framework was evaluated for two types of hearing loss: outer hair cell (OHC) loss and cochlear synaptopathy (CS). The results demonstrated that the ISE models effectively enhanced speech intelligibility and quality in noisy environments for hearing-impaired listeners across both loss types. A coefficient parameter was introduced to regulate noise reduction levels, enabling a balance between speech intelligibility, sound quality, and environmental awareness. Additionally, by integrating denoising and amplification within a unified system, the ISE model reduces computational complexity compared to conventional multi-module approaches. This efficiency makes it well-suited for real-time applications in low-resource hearing devices. Overall, the proposed ISE framework offers a promising direction for precise, DNN-based individualized speech enhancement, paving the way for future advancements in hearing aid technology.

7. Acknowledgements

This work was supported by FWO Machine Hearing 2.0 (216318G) and EIC-Transition EarDiTech (101058278).

8. References

[1] L. Liu, H. Guan, J. Ma, W. Dai, G. Wang, and S. Ding, “A mask free neural network for monaural speech enhancement,” *arXiv preprint arXiv:2306.04286*, 2023.

[2] Y. Song and N. Madhu, “Phase reconstruction in single channel speech enhancement based on phase gradients and estimated clean-speech amplitudes,” in *ICASSP 2024-2024 IEEE Interna-*

tional Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2024, pp. 1461–1465.

[3] X. Xu, Y. Zhang, W. Tu, and Y. Yang, "An efficient and interpre table speech enhancement network via deep dictionary learning," in *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).* IEEE, 2024, pp. 10 481–10 485.

[4] J. Yu, Y. Luo, H. Chen, R. Gu, and C. Weng, "High fidelity speech enhancement with band-split rnn," *arXiv preprint arXiv:2212.00406*, 2022.

[5] C. Zheng, H. Zhang, W. Liu, X. Luo, A. Li, X. Li, and B. C. Moore, "Sixty years of frequency-domain monaural speech enhancement: From traditional to deep learning methods," *Trends in Hearing*, vol. 27, p. 23312165231209913, 2023.

[6] Z. Tu, J. Zhang, N. Ma, J. Barker *et al.*, "A two-stage end-to-end system for speech-in-noise hearing aid processing," *Proc. Clarity*, pp. 3–5, 2021.

[7] X. Chen, Y. Shi, W. Xiao, M. Wang, T. Wu, S. Shang, Q. Meng, and N. Zheng, "A cascaded speech enhancement for hearing aids in noisy-reverberant conditions," in *Proc. Clarity Workshop on Machine Learning Challenges for Hearing Aids*, 2021.

[8] C.-C. Lee, H.-W. Chen, R. Chao, T.-T. Liu, and Y. Tsao, "Citear: A two-stage end-to-end system for noisy-reverberant hearing-aid processing," in *Proc. Clarity-CEC2-2022*, 2022.

[9] S. Cornell, Z.-Q. Wang, Y. Masuyama, S. Watanabe, M. Pariente, and N. Ono, "Multi-channel target speaker extraction with refinement: The wavlab submission to the second clarity enhancement challenge," *arXiv preprint arXiv:2302.07928*, 2023.

[10] F. Drakopoulos and S. Verhulst, "A neural-network framework for the design of individualised hearing-loss compensation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 31, pp. 2395–2409, 2023.

[11] C. Wen, G. Torfs, and S. Verhulst, "Artifact-free sound quality in dnn-based closed-loop systems for audio processing," *arXiv preprint arXiv:2501.04116*, 2025.

[12] P. Leer, J. Jensen, L. H. Carney, Z.-H. Tan, J. Østergaard, and L. Bramsløw, "Hearing-loss compensation using deep neural networks: A framework and results from a listening test," *IEEE Transactions on Audio, Speech and Language Processing*, 2025.

[13] S. Zhang, M. Lei, Z. Yan, and L. Dai, "Deep-fsmn for large vocabulary continuous speech recognition," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).* IEEE, 2018, pp. 5869–5873.

[14] A. Van Den Broucke, D. Baby, and S. Verhulst, "Hearing-impaired bio-inspired cochlear models for real-time auditory applications," in *21st Annual Conference of the International Speech Communication Association (INTERSPEECH 2020).* International Speech Communication Association (ISCA), 2020, pp. 2842–2846.

[15] J. S. Garofolo, L. F. Lamel, W. M. Fisher, J. G. Fiscus, and D. S. Pallett, "Darpa timit acoustic-phonetic continous speech corpus cd-rom. nist speech disc 1-1.1," *NASA STI/Recon technical report n*, vol. 93, p. 27403, 1993.

[16] S. Verhulst, A. Altoe, and V. Vasilkov, "Computational modeling of the human auditory periphery: Auditory-nerve responses, evoked potentials and hearing loss," *Hearing research*, vol. 360, pp. 55–75, 2018.

[17] C. K. Reddy, H. Dubey, K. Koishida, A. Nair, V. Gopal, R. Cutler, S. Braun, H. Gamper, R. Aichner, and S. Srinivasan, "Interspeech 2021 deep noise suppression challenge," in *INTERSPEECH*, 2021.

[18] C. K. Reddy, H. Dubey, V. Gopal, R. Cutler, S. Braun, H. Gamper, R. Aichner, and S. Srinivasan, "Icassp 2021 deep noise suppression challenge," in *ICASSP*, 2021.

[19] J. F. Gemmeke, D. P. Ellis, D. Freedman, A. Jansen, W. Lawrence, R. C. Moore, M. Plakal, and M. Ritter, "Audio set: An ontology and human-labeled dataset for audio events," in *2017 IEEE international conference on acoustics, speech and signal processing (ICASSP).* IEEE, 2017, pp. 776–780.

[20] E. Fonseca, J. Pons Puig, X. Favory, F. Font Corbera, D. Bogdanov, A. Ferraro, S. Oramas, A. Porter, and X. Serra, "Freesound datasets: a platform for the creation of open audio datasets," in *Hu X, Cunningham SJ, Turnbull D, Duan Z, editors. Proceedings of the 18th ISMIR Conference; 2017 oct 23-27; Suzhou, China.[Canada]: International Society for Music Information Retrieval; 2017. p. 486-93.* International Society for Music Information Retrieval (ISMIR), 2017.

[21] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[22] J. L. Roux, S. Wisdom, H. Erdogan, and J. R. Hershey, "Sdr – half-baked or well done?" in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 626–630.

[23] J. M. Kates and K. H. Arehart, "The hearing-aid speech quality index (hasqi) version 2," *Journal of the Audio Engineering Society*, vol. 62, no. 3, pp. 99–117, 2014.

[24] ——, "The hearing-aid speech perception index (haspi) version 2," *Speech Communication*, vol. 131, pp. 35–46, 2021.

[25] T. H. Falk, C. Zheng, and W.-Y. Chan, "A non-intrusive quality and intelligibility measure of reverberant and dereverberated speech," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 7, pp. 1766–1774, 2010.