


## Article

# Speech Enhancement for Hearing Aids with Deep Learning on Environmental Noises

Gyuseok Park <sup>1,2,†</sup> , Woohyeong Cho <sup>3,†</sup>, Kyu-Sung Kim <sup>2,4</sup> and Sangmin Lee <sup>1,3,\*</sup><sup>1</sup> Department of Electronic and Computer Engineering, Inha University, Incheon 22212, Korea; gyuseok.park@gmail.com<sup>2</sup> Research Institute for Aerospace Medicine, Inha University, Incheon 22212, Korea; stedman@inha.ac.kr<sup>3</sup> Department of Smart Engineering, Program in Biomedical Science & Engineering, Inha University, Incheon 22212, Korea; wakeiy@inha.edu<sup>4</sup> Department of Otolaryngology Head & Neck Surg., Inha University Hospital, Incheon 22332, Korea

\* Correspondence: sanglee@inha.ac.kr

† These authors contributed equally to this work.

Received: 20 August 2020; Accepted: 31 August 2020; Published: 2 September 2020



**Abstract:** Hearing aids are small electronic devices designed to improve hearing for persons with impaired hearing, using sophisticated audio signal processing algorithms and technologies. In general, the speech enhancement algorithms in hearing aids remove the environmental noise and enhance speech while still giving consideration to hearing characteristics and the environmental surroundings. In this study, a speech enhancement algorithm was proposed to improve speech quality in a hearing aid environment by applying noise reduction algorithms with deep neural network learning based on noise classification. In order to evaluate the speech enhancement in an actual hearing aid environment, ten types of noise were self-recorded and classified using convolutional neural networks. In addition, noise reduction for speech enhancement in the hearing aid were applied by deep neural networks based on the noise classification. As a result, the speech quality based on the speech enhancements removed using the deep neural networks—and associated environmental noise classification—exhibited a significant improvement over that of the conventional hearing aid algorithm. The improved speech quality was also evaluated by objective measure through the perceptual evaluation of speech quality score, the short-time objective intelligibility score, the overall quality composite measure, and the log likelihood ratio score.

**Keywords:** hearing aids; speech environment; environmental noise; deep neural networks; convolutional neural networks; noise classification; noise reduction

## 1. Introduction

The people with impaired hearing have difficulty hearing because they have a higher hearing threshold and narrower dynamic range than a person with normal hearing [1,2]. Consequently, hearing aids are worn as a way to reduce the difficulties caused by hearing loss, with hearing aids compensating for that loss [3]. Hearing aids are small electronic devices that amplify speech in noisy environments and improve the quality of communication for the hearing impaired.

Daily life is full of sound, and hearing aid technology is being constantly developed to reduce noise such as car horns, restaurants noise, the buzzing from electrical equipment, and random voices in the surroundings. Consequently, hearing aid technologies need to reduce environmental noise and amplify voices in order to improve speech intelligibility. However, one of the major complaints of hearing aid users is the inability of hearing aids to completely reduce environmental noise, along with the associated amplification of unexpected noise together with speech [4]. This is because the actual hearing aid operates in scenarios of various, irregular, and random environmental noise.

Digital hearing aids operated various audio signal processing algorithms, among which the speech enhancement algorithms are extremely important. Furthermore, speech enhancement algorithms for hearing aids, which are dissimilar to general signal processing algorithms, need to be realistic and specific due to the expectations of people with impaired hearing.

A typical speech enhancement algorithm for hearing aids classifies environmental noise near the hearing aid itself, estimating the noise power from the input signal, in order to reduce the estimated noise. For example, popular hearing aid noise classification algorithms include the support vector machine (SVM) [5] and gaussian mixture model (GMM) [6], popular noise estimation algorithms include the minimum statistics (MS) [7], minima controlled recursive averaging (MCRA) [8], and improved minima controlled recursive averaging (IMCRA) [9] algorithms, and popular noise reduction algorithms include the spectral subtraction (SS) [10], noise reduction using the wiener filter [11], the minimum mean square error short-time spectral amplitude (MMSE-STSA) [12], and the log minimum mean square error (logMMSE) algorithms [13].

Yu et al. [14] used the minimum mean square error-log scale amplitude (MMSE-LSA) estimator and modified for improving hearing aids speech perception. Nayan et al. [15] proposed a noise reduction and speech enhancement algorithm for hearing aid using the matrix wiener filter and compared the performance with multichannel wiener filter. Chandan et al. [16] used a power spectrum estimation and calculated the gain using spectral subtraction for speech enhancement of hearing aids, and Yuyong and Sangmin [17] proposed a speech enhancement algorithm for hearing aids using the IMCRA for noise estimation and removing the noise using logMMSE. After that, the noises were estimated using adaptive IMCRA parameters based on the noise classification using GMM, the parameters of which were used for noise reduction [18]. Though various conventional algorithms for hearing aids were proposed to improve the speech quality, they did not apply deep learning techniques for speech enhancement of hearing aids, which was known to be effective for noise classification and speech recognition technology.

The speech quality evaluation components of most typical speech enhancement algorithms for hearing aids use popular noise databases which are unnatural and homogenized, such as the NOISEX-92 database [19], the NOIZEUS database [20], the CHiME background noise data [21], and the AURORA-2 corpus [22]. However, the actual hearing aids operates in various irregular, and random noisy environments. The actual environmental noise may comprise two or more types of noise—for example, the random babble of subway sounds—and unexpected sounds may suddenly occur.

In this study, a speech enhancement algorithm was proposed that considers the environmental noise in which hearing aids operate and improve the speech quality using deep learning techniques. In order to accurately assess the type of environmental noise, the noise classification results with convolutional neural networks (CNNs) were used to accurately recognize the type of environmental noise [23], and a noise reduction algorithm using deep neural networks (DNNs) was applied to improve the speech enhancement.

## 2. Materials and Methods

The DNNs used for the proposed speech enhancement and the results of previous studies on noise classification are described in Sections 2.1 and 2.2, respectively. Section 2.3 includes an explanation of the process and the application of the proposed algorithm to efficient speech enhancement for hearing aids. Section 2.4 introduces the conditions pertaining to self-recorded noises and determination of the DNNs architecture and hyperparameters. In order to evaluate the performance of speech quality, the objective measurements are described in Section 2.5. Overall, this section deals with the process of speech enhancement and the operation of the network presented in this paper based on the environmental noise for hearing aids.

### 2.1. Deep Neural Networks

Artificial neuron networks (ANNs) are a computational model based on the structure and functions of biological neural networks [24]. DNNs are ANNs with multiple layers between the input and output layers, and deep learning is a machine learning technique that constructs artificial neural networks [25–27]. As the number of hidden layers increases, the characteristics of DNNs become more detailed and complex, and derive high-level function from input information. Based on this network model, the work of extracting features and determining the form of unstructured data is called deep learning.

DNNs consist of three types of neuron layers: the input layer, the hidden layers, and the output layer [28]. The input layer receives input data, and the output layer returns output data. The hidden layers perform mathematical computations on the inputs. Depending on the weight and transfer function of the input values, an activation function is applied to the next layer, and activation is implemented in networks until the neurons reach the output layer [29]. The difference between the value of output layer and the expected target value are compared using a cost function. Reducing it by changing the weight determines the performance of the network model.

### 2.2. Classification of Environmental Noise

To properly reduce the noise in a noisy environment, different noise reduction algorithms have to be applied to each noise. For that, the current noise characteristics need to be extracted according to the hearing aid's environment and distinctly classified into various noise categories.

The noise classification result using CNNs was applied for speech enhancement. CNNs are a kind of deep learning techniques and are widely used for image classification while maintaining the spatial information of the image. Convolutional and pooling layers are added between the input and output layers of CNNs, for excellent performance in processing data composed of multi-dimensional arrays, such as color images [23]. The feature map of the input data is produced by moving a convolution filter in the convolution layer, and the values obtained from the final feature maps are then extracted to reduce the computational complexity and improve the accuracy of the pooling layer [30].

In order to improve the classification rate of the environmental noise, a spectrogram image of the noise was used to transform sound signals in the time frequency-domain into image signals. In addition, a sharpening mask and median filter were applied to improve the noise classification rate [31,32]. The environmental noise used the same dataset as the experiment outlined in this paper. The classification results using CNNs were better than the conventional noise classification algorithms of hearing aids. This improved noise classification can contribute to the enhancement of speech for hearing aids [23].

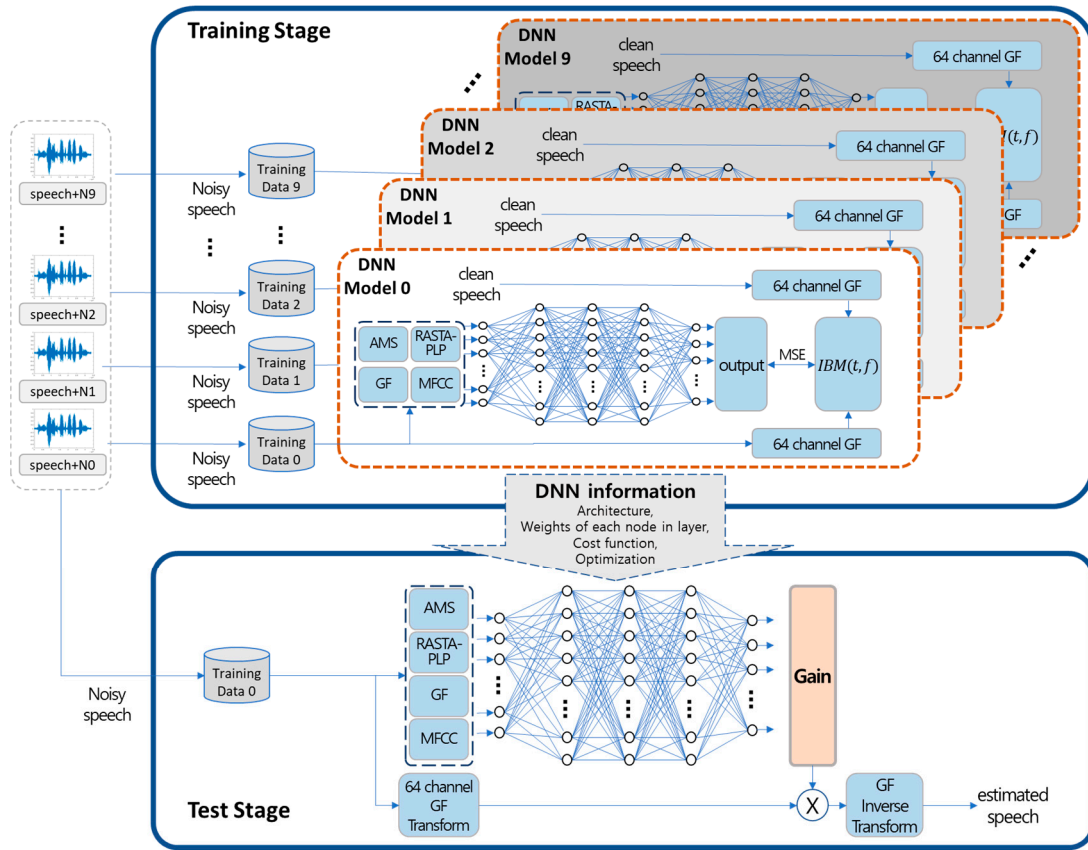
### 2.3. Proposed Algorithm

The noise reduction algorithm was operated using the stored learning model according to the environmental noise which had been trained using ten kinds of noise conditions for an effective noise reduction performance. For hearing aids operation, the DNNs were pre-trained and stored in each type of environmental noise.

The storage elements of the trained model were the training configuration, such as the architecture of the model, the weight of each node in the layers, the cost function, and the optimization method. In particular, since the weight of each node was different based on the type of environmental noise, the weight value varied depending on the noise classification. As a result, the weight set increased in proportion to the type of environmental noise learnt using the DNNs, and it was possible to reduce this according to the noise.

The process of the training stage and test stage of the DNNs for noise reduction are shown in Figure 1. First, the characteristic feature sets of noisy speech were extracted to generate input data for the DNNs [33]. The number of input data were 123, which consisted of mel-frequency

cepstral coefficients (MFCCs), amplitude modulation spectrogram (AMS), relative spectral transformed perceptual linear prediction coefficients (RASTA-PLP), and 64 gamma-tone filterbanks (GFs) [34]. The DNNs had three hidden layers and an output layer, and there were 1024 nodes in the hidden layer.



**Figure 1.** Proposed speech enhancement algorithm using the deep neural networks (DNNs) based on noise classification.

In the training stage, the cost function of the output layer was calculated as the mean square error between 64 IBM derived from the GFs and 64 output data applied to the sigmoid activation function. The ideal binary mask (IBM) is defined as [35]:

$$IBM(t, f) = \sqrt{\frac{S^2(t, f)}{N^2(t, f)}} \quad (1)$$

where  $S^2(t, f)$  and  $N^2(t, f)$  are the power of the speech signal and noisy speech signal, the range of values being from 0 to 1.

In the test stage, 64 output data ( $t_1, t_2, \dots, t_{64}$ ), which were the gain values of the power spectrum with the gamma-tone filterbank, were generated from 123 characteristic features of noisy speech from the input data [36]. The gain values were multiplied by the power spectrum that had been applied to the GF of the noisy speech, as follows:

$$|\hat{S}(k)|^2 = \sum_{k=1}^{64} t_k |Y_G(k)|^2 \quad (2)$$

The gain values were inversely transformed from the GF to the enhanced speech. Subsequently, the estimated enhanced speech was generated and evaluated using objective measures.

## 2.4. Experimental Setting

Since the noise databases, for example the NOISEX-92, the NOIZEUS, the CHiME background noise and the AURORA-2, are processed and machined to be applicable to most audio signal processing, it is easy to compare the results with those of other studies. However, the databases do not provide the variety of mixed sound found in real life.

Ten types of noises were recorded in the actual environments in which hearing aids are used: white noise (white, N0), café noise around Inha University, Korea (café, N1), interior noise in a moving car (car interior, N2), fan noise in a laboratory (fan, N3), laundry noise in a laundry room (laundry, N4), noise in the library at Inha University (library, N5), normal noise in a university laboratory (office, N6), various noises in a restaurant (restaurant, N7), noise in subway car (subway, N8), and traffic noise around an intersection (traffic, N9). Each noise was recorded three times at different times on different days, and the noise data for each noise type were generated for 30 min. The noises were recorded at 44.1 kHz, which is the highest sampling frequency of microphone recording to minimize loss of audio data as possible, and the noise data was need to down-sampled to 16 kHz, because 44.1 kHz was too high to use in hearing aids signal processing. The process of down-sampling was that low-pass filter with 8 kHz of cut-off frequency was used to prevent the effect of aliasing. Unlike the noise database, the speech database used TIMIT because speech needs to be provided with a constant speed and a uniform tone to mix with the noise. Noisy speech used a mixture of TIMIT sentences with recorded noises at 0, 5, 10, and 15 SNR.

The DNNs used three hidden layers, and each layer had 1024 nodes and rectified linear unit (ReLU). The standard backpropagation algorithm with dropout regularization was used to train the DNNs, the dropout rate of which was 0.2. The stochastic gradient descent algorithm with Adagrad was used for training the weights in the DNNs. A momentum rate of 0.5 was used in the first five epochs, after which the rate was increased to 0.9. The cost function in the last layer was the mean square error (MSE). The batch was set to 1024, the learning rate was 0.001, and the network was trained for 20 epochs.

The noise dataset was generated and the DNNs architecture designed with MATLAB R2019a (MathWorks Inc, Natick, MA, USA) and Parallel Computing Toolbox in MATLAB was use for deep learning.

## 2.5. Performance Evaluation

For objective speech quality evaluation, the results of speech enhancement were compared using the perceptual evaluation of speech quality (PESQ), the short-time objective intelligibility (STOI), the log likelihood ratio (LLR), and the overall quality composite measure (OQCM). The PESQ scored from −0.5 to 4.5, and the STOI scored from 0.0 to 1.0, with a higher score being an indication of better quality. The LLR scored from 0.0 to 2.0, in which a lower score is an indication of better speech quality. the OQCM was calculated using a combination of the existing objective evaluation measures to form a new measure [37], as follows:

$$OQCM = 1.594 + 0.805 \cdot PESQ - 0.512 \cdot LLR - 0.007 \cdot WSS \quad (3)$$

The higher values of the OQCM represent a lower signal distortion and higher quality of speech. PESQ, LLR, and the weighted spectral slope (WSS) represent the perceptual evaluation of speech quality, the log likelihood ratio, and weighted slope spectral distance in each frequency band, respectively.

## 3. Results

The results of speech enhancement for hearing aids with various objective measurements and the detailed performance of the proposal algorithm, compared to a conventional hearing aid algorithm, are outlined below.

### 3.1. Speech Enhancement

This section presents experimental results of the proposed speech enhancement for hearing aids using DNNs in each noise type. Table 1 shows the results of applying the speech enhancement algorithm (After) according to each objective evaluation measure. The quality of speech enhancement in noisy signals with large SNRs produced better results than those with low SNRs, in all noise types. In particular, the subway noise, fan noise, and laundry noise exhibited higher PESQ, STOI, and OQCM, and lower LLR. The reason is that the three noises had regular patterns and the babble noise interfered less when speech enhancement was applied using the DNNs.

**Table 1.** Evaluation of speech quality by proposed noise reduction algorithm based on noise classification.

Noise	SNR (dB)	PESQ Score		STOI Score		OQCM Score		LLR Score	
		Before	After	Before	After	Before	After	Before	After
Café	0	1.086	1.315	0.642	0.759	1.415	1.850	0.956	0.624
	5	1.195	1.578	0.750	0.832	1.731	2.255	0.729	0.450
	10	1.419	1.966	0.837	0.889	2.117	2.727	0.530	0.317
	15	1.802	2.460	0.907	0.934	2.608	3.251	0.359	0.214
Car interior	0	1.071	1.472	0.712	0.820	1.533	2.257	0.690	0.225
	5	1.165	1.788	0.780	0.862	1.812	2.629	0.507	0.164
	10	1.369	2.162	0.845	0.901	2.162	3.025	0.357	0.122
	15	1.710	2.586	0.901	0.936	2.591	3.442	0.241	0.091
Fan	0	1.075	1.456	0.685	0.808	1.400	2.107	1.002	0.469
	5	1.179	1.758	0.768	0.856	1.705	2.503	0.765	0.338
	10	1.396	2.128	0.839	0.898	2.094	2.930	0.547	0.235
	15	1.758	2.624	0.901	0.934	2.571	3.435	0.364	0.163
Laundry	0	1.066	1.455	0.644	0.808	1.336	2.034	1.011	0.577
	5	1.149	1.756	0.753	0.869	1.631	2.448	0.787	0.421
	10	1.352	2.159	0.840	0.915	2.013	2.914	0.578	0.303
	15	1.711	2.635	0.903	0.947	2.491	3.411	0.396	0.210
Library	0	1.080	1.460	0.671	0.793	1.484	2.054	1.056	0.686
	5	1.167	1.694	0.758	0.849	1.725	2.377	0.875	0.555
	10	1.458	2.134	0.854	0.905	2.193	2.895	0.614	0.391
	15	1.780	2.514	0.908	0.938	2.596	3.299	0.450	0.294
Office	0	1.059	1.349	0.629	0.770	1.355	1.893	1.083	0.678
	5	1.136	1.609	0.729	0.834	1.633	2.278	0.863	0.520
	10	1.316	1.962	0.822	0.888	1.992	2.714	0.642	0.382
	15	1.649	2.404	0.896	0.931	2.448	3.194	0.447	0.270
Restaurant	0	1.070	1.322	0.614	0.755	1.245	1.814	1.347	0.784
	5	1.143	1.610	0.739	0.842	1.528	2.237	1.126	0.599
	10	1.310	1.987	0.847	0.905	1.888	2.697	0.878	0.448
	15	1.632	2.474	0.922	0.948	2.354	3.221	0.635	0.319
Subway	0	1.094	1.486	0.715	0.814	1.554	2.215	0.774	0.331
	5	1.215	1.811	0.780	0.858	1.847	2.611	0.578	0.241
	10	1.453	2.189	0.843	0.898	2.228	3.029	0.406	0.168
	15	1.825	2.637	0.900	0.933	2.690	3.479	0.266	0.117



Table 1. Cont.

Noise	SNR (dB)	PESQ Score		STOI Score		OQCM Score		LLR Score	
		Before	After	Before	After	Before	After	Before	After
Traffic	0	1.076	1.367	0.668	0.783	1.417	1.953	1.075	0.623
	5	1.171	1.666	0.775	0.854	1.721	2.371	0.834	0.454
	10	1.380	2.074	0.859	0.907	2.090	2.845	0.618	0.320
	15	1.754	2.559	0.920	0.944	2.566	3.348	0.428	0.225
White	0	1.034	1.526	0.701	0.827	1.224	2.062	1.577	0.854
	5	1.058	1.880	0.811	0.894	1.378	2.479	1.491	0.729
	10	1.143	2.307	0.898	0.942	1.598	2.946	1.343	0.598
	15	1.351	2.769	0.953	0.971	1.933	3.430	1.133	0.472
Average <sup>1</sup>	0	1.071	1.421	0.668	0.794	1.396	2.024	1.057	0.585
	5	1.158	1.715	0.764	0.855	1.671	2.419	0.856	0.447
	10	1.360	2.107	0.848	0.905	2.038	2.872	0.651	0.328
	15	1.697	2.566	0.911	0.938	2.485	3.351	0.472	0.238

Average <sup>1</sup>: The average score of ten noises in each SNR dB.

As shown by the average of ten noises outlined in Table 1, all objective measurements of the proposed enhancement algorithm's application (After) demonstrated improved speech quality over the noisy speech which did not apply the algorithm (Before). In particular, the improvement of speech quality using OQCM, which was the closest subjective measure, increased 45% in the 0 dB SNR, 45% in the 5 dB SNR, 41% in the 10 dB SNR, and 35% in the 15 dB SNR tests, respectively.

### 3.2. Comparison Algorithms

In this session, the quality of speech enhancement using DNNs based on environmental noise classification was compared to the quality of the conventional hearing aid algorithm. In order to compare the quality of speech enhancement with the proposed algorithm, the results of the speech enhancement algorithm that were not based on the noise classification results and the conventional speech enhancement algorithm for hearing aids are presented in Table 2. The conventional speech enhancement algorithm for hearing aids classified the environmental noise using the Gaussian Mixture Models with covariance matrix, estimated the noise using MCRA by applying the optimized five parameters for each noise type, and reduced the noise using logMMSE [16].

**Table 2.** Evaluation of speech quality by conventional and proposed speech enhancement based on noise classification.

Evaluation Measure	SNR (dB)	IMCRA + logMMSE with Classification	Proposed DNNs	
			without Classification	with Classification
PESQ score	0	1.320	1.324	1.421
	5	1.593	1.554	1.715
	10	1.978	1.921	2.107
	15	2.451	2.440	2.566
STOI score	0	0.719	0.765	0.794
	5	0.807	0.839	0.855
	10	0.879	0.897	0.905
	15	0.931	0.939	0.942
OQCM score	0	1.382	1.872	2.024
	5	1.734	2.230	2.419
	10	2.145	2.679	2.872
	15	2.720	3.223	3.351
LLR score	0	0.726	0.684	0.585
	5	0.510	0.537	0.447
	10	0.359	0.395	0.328
	15	0.226	0.281	0.238

The PESQ and STOI scores of speech enhancement using the proposed DNNs without the classification result were similar to or slightly higher than the results of the speech enhancement using the conventional algorithm for hearing aids. However, the results of the speech enhancement using the proposed DNNs with the classification result produced higher PESQ, STOI, and OQCM and lower LLR scores.

In particular, in the case of noisy speech with a low SNR, the improvement of speech quality was seen to be greater. Consequently, the speech enhancement for hearing aids can be expected to improve by applying different noise reduction algorithms depending on the type of noise.

#### 4. Conclusions

In this study, the speech enhancement for hearing aids was investigated in actual self-recorded noisy environments. In order to improve speech quality, the environmental noises were classified using convolutional networks, and noise reduction using DNNs was applied based on the classified noise. The environmental noise was recorded in ten places most related to the environment in which hearing aids are used, and the objective evaluation of speech quality improvement was made using the PESQ, STOI, OQCM, and LLR scores.

With the proposed algorithm, comprising noise reduction based on noise classification, the PESQ score increased 2.17% in the 0 dB SNR, 3.50% in the 5 dB SNR, 3.69% in the 10 dB SNR, and 2.62% in the 15 dB SNR tests, respectively, than when the classification results were not applied. The STOI score increased 3.23% in the 0 dB SNR, 2.71% in the 5 dB SNR, 1.89% in the 10 dB SNR, and 1.30% in the 15 dB SNR tests, respectively. The OQCM, which was calculated using a combination of the existing objective evaluation measures, increased 0.203 in the 0 dB SNR, 0.243 in the 5 dB SNR, 0.225 in the 10 dB SNR, and 0.161 in the 15 dB SNR tests, respectively. The LLR score was 7.57% lower in the 0 dB SNR, 7.79% lower in the 5 dB SNR, 5.86% lower in the 10 dB SNR, and 4.20% lower in the 15 dB SNR tests, respectively, than the performance of the noise reduction without applying the classification results.

The proposed speech enhancement of hearing aids provided the best quality of speech in various and irregular noisy environments of the typical hearing aid user. Because recorded actual noises were closer to environments of hearing aid use, noise datasets were effectively used to improve the speech quality for hearing aid users. The speech enhancement using deep learning in hearing aid resulted in the improved speech quality compared with conventional speech enhancement algorithms in hearing aids. In addition, the proposed speech enhancement algorithm, which was applied different DNNs models according to the classified noise, had better performance of speech quality than the case without applying classified noise results. In summary, the increased noise classification rate using CNNs (with two types of image filters) improved the noise reduction performance using DNNs, and through the proposed algorithm, the quality of speech enhancement of the hearing aid improved, resulting in increased listening satisfaction for the people with impaired hearing. As hearing aid chips are advanced and speech signal processing for hearing aids is developed continuously, more complex algorithms could be applied to the hearing aids and more detailed hearing compensation would be provided to the diverse characteristics of hearing.

**Author Contributions:** Conceptualization, G.P., W.C., K.-S.K. and S.L.; data curation, G.P., W.C., K.-S.K. and S.L.; formal analysis, G.P., W.C., K.-S.K. and S.L.; funding acquisition, S.L.; investigation, G.P., W.C., K.-S.K. and S.L.; methodology, G.P., W.C., K.-S.K. and S.L.; project administration, S.L.; resources, S.L.; software, G.P. and W.C.; supervision, S.L.; validation, G.P., W.C., K.-S.K. and S.L.; visualization, G.P. and W.C.; writing—original draft, G.P.; and writing—review and editing, G.P., W.C., K.-S.K. and S.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (NRF-2020R1A2C2004624) and supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. 2018R1A6A1A03025523).

**Conflicts of Interest:** The authors declare no conflict of interest.



## References

1. Festen, J.M.; Plomp, R. Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *J. Acoust. Soc. Am.* **1990**, *88*, 1725–1736. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Hygge, S.; Ronnberg, J.; Larsby, B.; Arlinger, S. Normal-hearing and hearing-impaired subjects' ability to just follow conversation in competing speech, reversed speech, and noise backgrounds. *J. Speech Lang. Hear. Res.* **1992**, *35*, 208–215. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Plomp, R. Auditory handicap of hearing impairment and the limited benefit of hearing aids. *J. Acoust. Soc. Am.* **1978**, *63*, 533–549. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Duquesnoy, A. Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons. *J. Acoust. Soc. Am.* **1983**, *74*, 739–743. [\[CrossRef\]](#)
5. Suykens, J.A.; Vandewalle, J. Least squares support vector machine classifiers. *Neural Process. Lett.* **1999**, *9*, 293–300. [\[CrossRef\]](#)
6. Kim, G.; Lu, Y.; Hu, Y.; Loizou, P.C. An algorithm that improves speech intelligibility in noise for normal-hearing listeners. *J. Acoust. Soc. Am.* **2009**, *126*, 1486–1494. [\[CrossRef\]](#)
7. Martin, R. Noise power spectral density estimation based on optimal smoothing and minimum statistics. *IEEE Trans. Speech Audio Process.* **2001**, *9*, 504–512. [\[CrossRef\]](#)
8. Cohen, I.; Berdugo, B. Noise estimation by minima controlled recursive averaging for robust speech enhancement. *IEEE Signal Process. Lett.* **2002**, *9*, 12–15. [\[CrossRef\]](#)
9. Cohen, I. Noise spectrum estimation in adverse environments: Improved minima controlled recursive averaging. *IEEE Trans. Speech Audio Process.* **2003**, *11*, 466–475. [\[CrossRef\]](#)
10. Loizou, P.C. *Speech Enhancement: Theory and Practice*; CRC Press: Boca Raton, FL, USA, 2013.
11. Van den Bogaert, T.; Doclo, S.; Wouters, J.; Moonen, M. Speech enhancement with multichannel Wiener filter techniques in multimicrophone binaural hearing aids. *J. Acoust. Soc. Am.* **2009**, *125*, 360–371. [\[CrossRef\]](#)
12. Eddins, D.A. *Sandlin's Textbook of Hearing Aid Amplification*; Taylor & Francis: Abingdon, UK, 2014.
13. Ephraim, Y.; Malah, D. Speech enhancement using a minimum mean-square error log-spectral amplitude estimator. *IEEE Trans. Acoust. Speech Signal Process.* **1985**, *33*, 443–445. [\[CrossRef\]](#)
14. Rao, Y.; Hao, Y.; Panahi, I.M.; Kehtarnavaz, N. Smartphone-based real-time speech enhancement for improving hearing aids speech perception. In Proceedings of the 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 16–20 August 2016; pp. 5885–5888.
15. Modhave, N.; Karuna, Y.; Tonde, S. Design of multichannel wiener filter for speech enhancement in hearing aids and noise reduction technique. In Proceedings of the 2016 Online International Conference on Green Engineering and Technologies (IC-GET), Coimbatore, India, 19 November 2016; pp. 1–4.
16. Reddy, C.K.; Hao, Y.; Panahi, I. Two microphones spectral-coherence based speech enhancement for hearing aids using smartphone as an assistive device. In Proceedings of the 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 16–20 August 2016; pp. 3670–3673.
17. Jeon, Y.; Lee, S. Low-Complexity Speech Enhancement Algorithm Based on IMCRA Algorithm for Hearing Aids. *J. Rehabil. Welf. Eng. Assist. Technol.* **2017**, *11*, 363–370.
18. Jeon, Y. A Study on Low-Complexity Speech Enhancement with Optimal Parameters Based on Noise Classification. Ph.D. Thesis, Inha-University, Incheon, Korea, February 2018.
19. Varga, A.; Steeneken, H.J. Assessment for automatic speech recognition: II. NOISEX-92: A database and an experiment to study the effect of additive noise on speech recognition systems. *Speech Commun.* **1993**, *12*, 247–251. [\[CrossRef\]](#)
20. Hu, Y.; Loizou, P.C. Subjective comparison and evaluation of speech enhancement algorithms. *Speech Commun.* **2007**, *49*, 588–601. [\[CrossRef\]](#)
21. Barker, J.; Vincent, E.; Ma, N.; Christensen, H.; Green, P. The PASCAL CHiME speech separation and recognition challenge. *Comput. Speech Lang.* **2013**, *27*, 621–633. [\[CrossRef\]](#)
22. Hirsch, H.-G.; Pearce, D. The Aurora experimental framework for the performance evaluation of speech recognition systems under noisy conditions. In *ASR2000-Automatic Speech Recognition: Challenges for the New Millenium*; ISCA tutorial and research workshop (ITRW): Paris, France, 2000.

23. Park, G.; Lee, S. Environmental Noise Classification Using Convolutional Neural Networks with Input Transform for Hearing Aids. *Int. J. Environ. Res. Public Health* **2020**, *17*, 2270. [[CrossRef](#)]
24. Rosenblatt, F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol. Rev.* **1958**, *65*, 386. [[CrossRef](#)]
25. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536. [[CrossRef](#)]
26. McClelland, J.L.; Rumelhart, D.E.; Group, P.R. Parallel distributed processing. *Explor. Microstruct. Cognit.* **1986**, *2*, 216–271.
27. Fukushima, K.; Miyake, S. Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition. In *Competition and Cooperation in Neural Nets*; Springer: Berlin/Heidelberg, Germany, 1982; pp. 267–285.
28. Liu, W.; Wang, Z.; Liu, X.; Zeng, N.; Liu, Y.; Alsaadi, F.E. A survey of deep neural network architectures and their applications. *Neurocomputing* **2017**, *234*, 11–26. [[CrossRef](#)]
29. Hopfield, J.J. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 2554–2558. [[CrossRef](#)] [[PubMed](#)]
30. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems, Proceedings of the Neural Information Processing Systems Conference, Lake Tahoe, NV, USA, 3–6 December 2012*; Pereira, F., Burges, C.J.C., Bottou, L., Weinberger, K.Q., Eds.; NIPS: San Diego, CA, USA, 2012; pp. 1097–1105.
31. Yang, C.-C. Improving the overshooting of a sharpened image by employing nonlinear transfer functions in the mask-filtering approach. *Optik* **2013**, *124*, 2784–2786. [[CrossRef](#)]
32. Hong, S.-W.; Kim, N.-H. A study on median filter using directional mask in salt & pepper noise environments. *J. Korea Inst. Inf. Commun. Eng.* **2015**, *19*, 230–236.
33. Wang, Y.; Narayanan, A.; Wang, D. On training targets for supervised speech separation. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2014**, *22*, 1849–1858. [[CrossRef](#)] [[PubMed](#)]
34. Wang, Y.; Han, K.; Wang, D. Exploring monaural features for classification-based speech segregation. *IEEE Trans. Audio Speech Lang. Process.* **2012**, *21*, 270–279. [[CrossRef](#)]
35. Wang, D. On ideal binary mask as the computational goal of auditory scene analysis. In *Speech Separation by Humans and Machines*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 181–197.
36. Han, K.; Wang, D. A classification based approach to speech segregation. *J. Acoust. Soc. Am.* **2012**, *132*, 3475–3483. [[CrossRef](#)]
37. Coelho, R.F.; Nascimento, V.H.; de Queiroz, R.L.; Romano, J.M.T.; Cavalcante, C.C. *Signals and Images: Advances and Results in Speech, Estimation, Compression, Recognition, Filtering, and Processing*; CRC Press: Boca Raton, FL, USA, 2018.

