ISSN: 2664-9640

Applications of Deep Learning in the Hearing Aid Industry

Ye Cheng[1,*], Sheng Chen[2]

[1] School of Shanghai for Science and Technology University, Shanghai, 200000, China

[2] School of Shanghai for Science and Technology University, Shanghai 200000, China

[*]Corresponding Author

Abstract

Hearing loss represents a severe global public health challenge, making the technological innovation of its primary intervention-hearing aids-critically important. Although modern digital hearing aids have achieved substantial progress in sound amplification, effectively improving speech intelligibility in noisy environments remains a major technical bottleneck. In recent years, the rapid advancement of artificial intelligence, particularly deep learning (DL), has presented disruptive and transformative opportunities for otology and the hearing aid industry. This paper systematically reviews and synthesizes the current status and developmental trends of DL across multiple dimensions, including otological medical image analysis, hearing assessment and prediction, core performance enhancement of hearing aids (such as speech enhancement, noise reduction, and smart environmental adaptation), and personalized device manufacturing. Research indicates that DL is profoundly transforming every stage from disease diagnosis and auditory rehabilitation to the device itself with unprecedented breadth and depth, propelling the otology and hearing aid industry toward a more precise, intelligent, and highly personalized future.

Keywords

Deep Learning; Hearing Aids; Hearing Loss; Neural Networks.

1. Introduction

According to the World Health Organization (WHO), over 1.5 billion people worldwide suffer from varying degrees of hearing loss. As the core intervention in auditory rehabilitation, the technological development of hearing aids has transitioned from analog to digital signal processing (DSP)[1]. However, current mainstream digital hearing aids still face severe challenges when processing complex acoustic environments. The traditional technologies upon which they rely-including rule-based linear/nonlinear signal amplification[2], fixed or limited directional microphone arrays[3], and classical noise reduction algorithms like Wiener filtering and spectral subtraction[4]-often struggle to effectively separate target speech from interfering noise in complex scenarios, such as non-stationary noise environments or multi-talker conditions (Speech-in-Noise, SiN)[5]. This limitation directly leads to significant listening fatigue for users and reduces overall satisfaction with the device.

Deep learning an artificial intelligence technology inspired by the neural network structure of the human brain, demonstrates a remarkable capability to automatically extract complex features and patterns from massive, high-dimensional datasets[6]. In recent years, driven by significant advancements in computational power and the rapid development of edge computing[7], DL technologies have successfully transitioned from research laboratories to achieve breakthrough applications in real-time signal processing for medical devices. This paper aims to systematically review the developmental trajectory of DL in the field of otology, with a particular focus on its key applications in the R&D and manufacturing processes of modern smart hearing aids, thereby providing a valuable reference for the future innovation and development of auditory rehabilitation technologies.

2. Development of Deep Learning in Otology

2.1. Medical Image Analysis and Disease Diagnosis

In the field of medical image analysis, the advent of Fully Convolutional Networks (FCNs)[8], particularly the U-Net architecture and its derivative models (e.g., 3D U-Net[9], V-Net[10]), has brought about a paradigm shift in computer-aided diagnosis (CAD). Recently, the introduction of Vision Transformers[11] (ViTs) has overcome the limitations of local receptive fields, enabling algorithms to capture global dependencies within images. Driven by these fundamental algorithmic breakthroughs, an increasing number of image-assisted diagnostic models have bypassed the bottlenecks of traditional image processing techniques (e.g., thresholding and edge detection), successfully transitioning into clinical practice. Initially, the application of deep learning in otology focused primarily on otoendoscopic image analysis. Researchers utilized classical convolutional neural networks (CNNs), such as ResNet[12], to classify 2D color images of the tympanic membrane[13], thereby assisting general practitioners in the rapid screening of common superficial lesions like acute otitis media[14], middle ear effusion[15], and tympanic membrane perforation[16]. Furthermore, with the advancement of 3D convolutional networks and multimodal fusion technologies, deep learning has achieved significant progress in otological imaging (e.g., temporal bone CT[17] and inner ear MRI[18]). Utilizing models like CNNs, algorithms can now automatically segment inner ear structures[19] (such as the cochlea and semicircular canals) and identify lesions including acoustic neuromas[20], otitis media, and cholesteatomas[21]. This automated imaging evaluation not only enhances the diagnostic efficiency of physicians but also provides precise 3D anatomical references for subsequent auditory implantation surgeries, such as cochlear implants.

2.2. Audiogram Analysis and Prediction

Traditional audiological tests and the interpretation of their results have long been characterized by significant subjectivity and a heavy reliance on clinical experience. Clinicians and audiologists must routinely synthesize multidimensional datasets-such as pure-tone audiometry (PTA), acoustic immittance, and speech audiometry-to formulate diagnostic strategies. This process is not only resource-intensive but also highly susceptible to significant inter-rater variability, particularly when evaluating atypical cases. With the continuous maturation of underlying architectures like Multilayer Perceptrons (MLPs), Recurrent Neural Networks (RNNs), and multimodal fusion networks, the paradigm of audiological evaluation is undergoing a profound transition from a manual, experience-driven approach to data-driven, intelligent computer-aided diagnostic systems.

In the context of cross-sectional static diagnosis, modern deep learning architectures can map high-dimensional input vectors from a patient's PTA threshold matrices, speech discrimination scores (SDS), objective auditory electrophysiological features (e.g., otoacoustic emissions [OAEs][22] and auditory brainstem responses [ABRs][23]), and multisource medical history data extracted from structured electronic medical records (EMRs). Leveraging the non-linear feature mapping capabilities of deep networks, these algorithms can automatically and accurately classify the type (conductive, sensorineural, or mixed) and severity of hearing loss with an accuracy comparable to that of experienced specialists. Furthermore, they are highly sensitive in detecting subtle subclinical electrophysiological and abnormal acoustic features that are often indiscernible in traditional audiograms, thereby providing a reliable adjunctive basis for the early screening of hidden hearing loss.

Regarding longitudinal dynamic prognosis and disease progression tracking, deep learning exhibits exceptional capabilities in time-series modeling. By utilizing deep architectures tailored for time-series analysis, such as Long Short-Term Memory[5] (LSTM) networks, researchers can extensively mine massive datasets from longitudinal audiological cohorts. These models can precisely fit the future deterioration trajectories of patients' hearing thresholds, demonstrating remarkable statistical superiority, particularly in assessing the natural progression of presbycusis and predicting the probability of hearing recovery in cases of idiopathic sudden sensorineural hearing loss. This disease progression prediction mechanism, based on temporal deep feature representation, effectively fills a methodological gap in the longitudinal dynamic assessment of clinical audiology. Its core clinical value lies in driving the paradigm shift of auditory interventions from reactive, passive acoustic compensation to early, proactive intervention. It assists specialists in precisely determining the optimal time windows for early hearing aid fitting or cochlear implantation, thereby establishing a lifelong, personalized, closed-loop auditory rehabilitation system for hearing-impaired patients.

3. Application of Deep Learning in Hearing Aid Manufacturing

3.1. Speech Enhancement and Noise Reduction

Speech enhancement stands out as one of the most transformative and central applications of deep learning in hearing aid technology, significantly overcoming the inherent limitations of traditional noise reduction algorithms, particularly in low signal-to-noise ratio (SNR) environments. Conventional methods, such as spectral subtraction[24] or Wiener filtering[25], frequently introduce undesirable artifacts like "musical noise" when processing complex background noise, while inevitably degrading and distorting the target speech signal.

The widespread adoption of U-Net and its variants has enabled the development of more sophisticated time-frequency masking techniques[26]. Among these, Ideal Binary Masking[27] (IBM) and Ideal Ratio Masking[28] (IRM) exemplify typical deep learning-based approaches.Through end-to-end learning on large-scale acoustic datasets, these advanced models accurately capture the nonlinear mapping between speech signals and various types of noise (including, but not limited to, ambient wind noise, traffic noise, and multi-talker babble) in the time-frequency domain. Consequently, they generate highly precise masking maps, which effectively suppress non-target sound sources during signal reconstruction, thereby maximizing the preservation and extraction of desired speech components.

Extensive clinical trials and objective acoustic evaluations have demonstrated that hearing aids integrating advanced DNN-based noise reduction algorithms can significantly enhance the perceived SNR in complex acoustic scenarios[29]. This improvement effectively reduces users' cognitive load and listening effort, leading to substantial gains in subjective sound quality and overall quality of life. These technological advancements represent a qualitative leap in the performance of hearing aids within complex, real-world acoustic environments.

3.2. Edge Computing and Ultra-Low Latency Implementation

The physical constraints inherent in hearing aid design pose significant challenges: devices must be compact, operate with extremely limited battery capacity (requiring power consumption in the milliwatt range), and maintain processing latency below 10 milliseconds to avoid perceptual artifacts such as auditory desynchronization and the occlusion effect. To deploy large-scale deep learning models onto the digital signal processing (DSP) chips within hearing aids, engineers employ techniques like model pruning, knowledge distillation, and quantization. Furthermore, leading manufacturers, such as Oticon with its Intent series, have begun integrating specialized hardware accelerators (coprocessors) designed for running deep neural networks (DNNs) directly into their devices[30, 31]. This integration enables real-time deep learning inference locally, fundamentally altering the traditional hardware architecture of hearing aids .

3.3. Real-time Acoustic Environment Classification

Utilizing model architectures such as Convolutional Neural Networks (CNNs) or more advanced Attention Mechanisms, the system can effectively map extracted acoustic features to predefined acoustic scene categories[32, 33]. These categories encompass a wide range of typical daily scenarios, including extremely quiet environments, noisy indoor and outdoor settings, traffic noise, music appreciation, and outdoor activities affected by wind noise.

In contrast to traditional statistical, probability-based classification methods like Hidden Markov Models, deep learning-based models exhibit superior performance when processing highly non-stationary and dynamic noise environments. Empirical data indicate that their environmental classification accuracy exceeds 97%[33]. This capability for high-precision, low-latency, real-time environment recognition allows hearing aids to accurately identify the current acoustic background within milliseconds. Consequently, the device can instantly and seamlessly switch to the optimal preset processing strategy-such as adjusting the noise reduction level, directivity, or gain compensation curve-thereby delivering a continuously optimized auditory experience to the user.

3.4. Application in Hearing Aid Manufacturing Process

In the precision manufacturing workflow of hearing aids, 3D printing technology is progressively supplanting traditional methods that rely on labor-intensive manual operations[34, 35]. Historically, dispensers were required to meticulously adjust digital ear canal models, derived from patient scans, to differentiate between left and right ear canal geometries. This was followed by intricate 3D shell modeling, a process that was not only time-consuming but also susceptible to variations in model precision due to differences in operator experience[36].

With the maturation of advanced algorithms, including deep learning-based 3D Convolutional Neural Networks (3D CNNs) and point cloud processing techniques[37], systems can now directly process high-resolution ear canal scan data (e.g., digital models in STL format). These algorithms automatically identify and differentiate anatomical features of the left and right ear canals, generating accurate, personalized 3D shell models without the need for manual intervention[38].

This AI-driven automated modeling workflow significantly reduces the production lead time, from ear canal scanning to final enclosure fabrication. Moreover, it substantially enhances both the shell-to-ear canal fit and overall wearing comfort. The resulting 3D models are then directly transferable to 3D printing equipment for high-precision fabrication using biocompatible materials. Consequently, the customized hearing aid shells produced offer markedly higher production efficiency compared to traditional manufacturing methods.

4. Conclusion

This paper systematically examines the innovative applications and profound impact of deep learning technology within the domains of otology and hearing aids. By providing a comprehensive overview of its use in medical image analysis, audiogram interpretation and prediction, speech enhancement and noise reduction, real-time acoustic environment classification, edge computing for hearing aids, and hearing aid manufacturing processes, this work demonstrates how deep learning effectively addresses the limitations of traditional techniques. It highlights the technology's role in driving these fields toward enhanced intelligence, precision, and personalization. In conclusion, deep learning is transforming the

ISSN: 2664-9640

fields of otology and the hearing aid industry with unprecedented depth and scope. It not only fundamentally enhances diagnostic accuracy and therapeutic efficacy but is also crucial for improving user experience and fostering technological innovation. This paper aims to establish a robust foundation and offer forward-looking insights for the future development of auditory rehabilitation technologies, thereby facilitating improved intervention and care for hearing-impaired individuals globally.

References

[1] KIM S-W, KIM M-J, KIM J-S. High-performance DSP platform for digital hearing aid SoC with flexible noise estimation [J]. IET Circuits, Devices and Systems, 2019, 13(5): 717-22.

[2] GHOSH S, MEENA T, GEORGE N V. Low-rank enhanced Hammerstein-spline adaptive filter for sparsity-aware nonlinear feedback cancellation in hearing aids [J]. Signal Processing, 2026, 244.

[3] RAMANI G, KAVITHA D, LATHA G C P, et al. Adaptive Dual-Channel Pulse-Coupled Neural Network with S-Transforms for Speech Source Localization Using Binaural and Monaural Microphone Arrays in Hearing Aids; proceedings of the 2025 International Conference on Multi-Agent Systems for Collaborative Intelligence, ICMSCI 2025, January 20, 2025 - January 22, 2025, Erode, India, F, 2025 [C]. Institute of Electrical and Electronics Engineers Inc.

[4] IQBAL Y, ZHANG T, GUNAWAN T S, et al. A Hybrid Speech Enhancement Technique Based on Discrete Wavelet Transform and Spectral Subtraction [J]. IEEE Access, 2025, 13: 39765-81.

[5] HONG C, XIAOHU P, QIANG D, et al. Classification and Recognition of Noise-Induced Hearing Loss Based on P300 Event-Related Potential and LSTM-Attention Network; proceedings of the 16th International Conference on Bioinformatics and Biomedical Technology, ICBBT 2024, May 24, 2024 - May 26, 2024, Chongqing, China, F, 2024 [C]. Association for Computing Machinery.

[6] GOWRI K, SHIVAKUMAR B L. Machine, Deep, and Reinforcement Learning Algorithms and Applications [M]. Apple Academic Press. 2026: 55-83.

[7] RATHEE G, GARG S, KADDOUM G, et al. Accessibility and ensured quality of life for disabled people using trusted edge computing [J]. Alexandria Engineering Journal, 2023, 82: 35-42.

[8] CHEN Q, XU J, KOLTUN V. Fast Image Processing with Fully-Convolutional Networks [Z]. arXiv. 2017

[9] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation; proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention, MICCAI 2015, October 5, 2015 - October 9, 2015, Munich, Germany, F, 2015 [C]. Springer Verlag.

[10] MILLETARI F, NAVAB N, AHMADI S-A. V-Net: Fully convolutional neural networks for volumetric medical image segmentation; proceedings of the 4th International Conference on 3D Vision, 3DV 2016, October 25, 2016 - October 28, 2016, Stanford, CA, United states, F, 2016 [C]. Institute of Electrical and Electronics Engineers Inc.

[11] LIU Z, LIN Y, CAO Y, et al. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows; proceedings of the 18th IEEE/CVF International Conference on Computer Vision, ICCV 2021, October 11, 2021 - October 17, 2021, Virtual, Online, Canada, F, 2021 [C]. Institute of Electrical and Electronics Engineers Inc.

[12] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception-v4, inception-ResNet and the impact of residual connections on learning; proceedings of the 31st AAAI Conference on Artificial Intelligence, AAAI 2017, February 4, 2017 - February 10, 2017, San Francisco, CA, United states, F, 2017 [C]. AAAI press.

[13] BAARAN E, COMERT Z, CELIK Y. Convolutional neural network approach for automatic tympanic membrane detection and classification [J]. Biomedical Signal Processing and Control, 2020, 56.

[14] KOVVALI O, MOTATI L S. Accessible Otitis Media Screening with a Deep Learning-Powered Mobile Otoscope; proceedings of the 2nd International Workshop on Applications of Medical Artificial Intelligence, AMAI 2023, October 8, 2023 - October 8, 2023, Vancouver, BC, Canada, F, 2024 [C]. Springer Science and Business Media Deutschland GmbH.

[15] GRAIS E M, NIE L, ZOU B, et al. An advanced machine learning approach for high accuracy automated diagnosis of otitis media with effusion in different age groups using 3D wideband acoustic immittance [J]. Biomedical Signal Processing and Control, 2024, 87.

[16] GUENNOUNI K, BERRAJAA A, BERRAJAA I, et al. Tympanic Perforations Diagnosis: A Revolutionary Approach by Artificial Intelligence for Increased Accuracy and Swift Management; proceedings of the 7th International Conference on Big Data and Internet of Things, BDIoT 2024, April 17, 2024 - April 19, 2024, Fez, Morocco, F, 2024 [C]. Springer Science and Business Media Deutschland GmbH.

[17] NIKAN S, OSCH K V, BARTLING M, et al. Pwd-3dnet: A deep learning-based fully-automated segmentation of multiple structures on temporal bone ct scans [J]. IEEE Transactions on Image Processing, 2021, 30: 739-53.

[18] ZHANG S, GENG X, MA Y, et al. Segmentation and Measurement of Inner Ear Auditory Nerve Based on Deep Learning; proceedings of the 9th International Conference on Intelligent Computing and Signal Processing, ICSP 2024, April 19, 2024 - April 21, 2024, Hybrid, Xi'an, China, F, 2024 [C]. Institute of Electrical and Electronics Engineers Inc.

[19] BUSWINKA C J. Applications of Computer Vision on the Biology of the Inner Ear [D]; ProQuest LLC, 2025.

[20] ZHU Q, LI H, CASS N D, et al. Acoustic neuroma segmentation using ensembled convolutional neural networks; proceedings of the Medical Imaging 2022: Biomedical Applications in Molecular, Structural, and Functional Imaging, March 21, 2022 - March 27, 2022, Virtual, Online, F, 2022 [C]. SPIE.

[21] RAJAN D, SUMATHY R. Advancing Otolaryngology Through Artificial Intelligence for Ear Disease Detection; proceedings of the 3rd International Conference on Sustainable Computing and Data Communication Systems, ICSCDS 2025, August 6, 2025 - August 8, 2025, Erode, India, F, 2025 [C]. Institute of Electrical and Electronics Engineers Inc.

[22] VARUGEESE A, SHAHINA A, NAWAS K, et al. EarNet: Biometric Embeddings for End to End Person Authentication System Using Transient Evoked Otoacoustic Emission Signals [J]. Neural Processing Letters, 2022, 54(1): 21-41.

[23] KASSJANSKI M, KULAWIAK M, PRZEWOZNY T. Development of an AI-based audiogram classification method for patient referral; proceedings of the 17th Conference on Computer Science and Intelligence Systems, FedCSIS 2022, September 4, 2022 - September 7, 2022, Sofia, Bulgaria, F, 2022 [C]. Institute of Electrical and Electronics Engineers Inc.

[24] NEUMAN A C, WEISS M, LEVITT H. NOISE REDUCTION FOR THE HEARING IMPAIRED; proceedings of the Second International Conference on Rehabilitation Engineering, Combined with the RESNA 7th Annual Conference, Ottawa, Ont, Can, F, 1984 [C]. Rehabilitation Engineering Soc of North America.

[25] KATES J M. Feedback cancellation in hearing aids; proceedings of the 1990 International Conference on Acoustics, Speech, and Signal Processing, ICASSP 1990, April 3, 1990 - April 6, 1990, Albuquerque, New Mexico, USA, F, 1990 [C]. Publ by IEEE.

[26] NAIR A A, KOISHIDA K. Cascaded Time + Time-Frequency Unet for Speech Enhancement: Jointly Addressing Clipping, Codec Distortions, and Gaps; proceedings of the 2021 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2021, June 6, 2021 - June 11, 2021, Virtual, Toronto, ON, Canada, F, 2021 [C]. Institute of Electrical and Electronics Engineers Inc.

[27] CHEN P, NGUYEN B T, IWAI K, et al. Threshold-Based Combination of Ideal Binary Mask and Ideal Ratio Mask for Single-Channel Speech Separation [J]. Information (Switzerland), 2024, 15(10).

[28] NARAYANAN A, WANG D. Ideal ratio mask estimation using deep neural networks for robust speech recognition; proceedings of the 2013 38th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2013, May 26, 2013 - May 31, 2013, Vancouver, BC, Canada, F, 2013 [C]. Institute of Electrical and Electronics Engineers Inc.

[29] ZHAO Y, WANG D, MERKS I, et al. DNN-based enhancement of noisy and reverberant speech; proceedings of the 41st IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2016, March 20, 2016 - March 25, 2016, Shanghai, China, F, 2016 [C]. Institute of Electrical and Electronics Engineers Inc.

[30] WANG D. Deep learning reinvents the hearing aid [J]. IEEE Spectrum, 2017, 54(3): 32-7.

[31] NOSSIER S A, RIZK M R M, MOUSSA N D, et al. Enhanced smart hearing aid using deep neural networks [J]. Alexandria Engineering Journal, 2019, 58(2): 539-50.

[32] PRESANNAKUMAR K, MOHAMED A. Deep learning based source identification of environmental audio signals using optimized convolutional neural networks [J]. Applied Soft Computing, 2023, 143.

[33] DING B, ZHANG T, WANG C, et al. Acoustic scene classification: A comprehensive survey [J]. Expert Systems with Applications, 2024, 238.

[34] SICKEL K, BALOCH S, MELKISETOGLU R, et al. Toward automation in hearing aid design [J]. CAD Computer Aided Design, 2011, 43(12): 1793-802.

[35] SLABAUGH G, TONG F, MCBAGONLURI F, et al. 3-D shape modeling for hearing aid design [Applications Corner] [J]. IEEE Signal Processing Magazine, 2008, 25(5): 98-102.

[36] UNAL G. Nonparametric joint shape learning for customized shape modeling [J]. Comput Med Imaging Graph, 2010, 34(4): 298-307.

[37] SULLIVAN E O, ZAFEIRIOU S. 3D Landmark Localization in Point Clouds for the Human Ear; proceedings of the 15th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2020, November 16, 2020 - November 20, 2020, Buenos Aires, Argentina, F, 2020 [C]. Institute of Electrical and Electronics Engineers Inc.

[38] DRITSAS S, CHUA K W, GOH Z H, et al. Classification, registration and segmentation of ear canal impressions using convolutional neural networks [J]. Med Image Anal, 2024, 94: 103152.