Phonak Insight.

Artificial intelligence in hearing aid technology

Artificial intelligence (AI) enables the creation of more personalized and adaptive solutions, tailoring amplification and noise reduction algorithms to the unique needs of the wearer in any given environment. Innovating in the AI field for 24 years, AutoSense OS™ 5.0 is the latest example of AI-based machine learning from Phonak. AutoSense OS 5.0 accurately identifies and adjusts to a client's unique sound environment, providing improved speech understanding in many different hearing situations.

April 2024: Henning Hasemann, Alena Krylova

Key highlights

- The development of automatic classification systems, like Phonak's AutoSense OS, have been shown to significantly improve speech understanding by adjusting the hearing aid according to the user's environment (Appleton, 2020).
- While deep neural networks (DNN) with enough parameters may perform better than classical signal processing approaches, currently no DNN has the computational power to outperform a classical or machine learning (ML) algorithm.
- Many hearing aid manufacturers have adopted ML approaches for scene classification tasks. This "best of both worlds" approach utilizes the power of "learning" approaches to cover a wide range of situations with high accuracy, without the high energy consumption required due to complex processing.

Considerations for practice

- Understanding the foundations of AI and ML is crucial for recognizing the benefits they bring to hearing aids.
- AI enables the creation of more personalized and adaptive solutions, tailoring amplification, and noise reduction algorithms to the unique needs of each hearing aid user in any given auditory environment.
- Dependent on a variety of factors, modern hearing aids can improve the signal-to-noise ratio (SNR) of a moderately challenging environment by approximately 4-8 dB SNR. In comparison, DNNs have the potential to improve by upwards of 12dB SNR. DNN-based denoising technology offers a potential that modern hearing aids do not have access to yet.

Introduction

Artificial intelligence (AI), a field of study in computer science aiming to replicate human intelligence in machines, is rapidly growing and has revolutionized entire industries. AI-based technologies are successfully applied across various sectors, bringing about a powerful, positive transformation. The European Commission defines AI as "systems that display intelligent behavior by analyzing their environment and taking actions with some degree of autonomy — to achieve specific goals", elaborating that "AI-based systems can be purely software based, acting in the virtual world (e.g. voice assistants and image analysis software) or can be embedded in hardware devices (e.g. advanced robots and autonomous cars)" (European Commision, 2018). From healthcare to finance, manufacturing to education, AI is reshaping traditional processes and opening up new opportunities.

AI applications emerge in a wide array of forms and sizes, featuring a remarkably broad spectrum of capabilities from emulating human-like perception, like vision and hearing, to excelling in communication and complex tasks. Hearing aids have benefitted greatly from advancements in AI, with the primary focus of improving outcomes like speech understanding and reduced listening effort. Despite significant progress in these features, the noise reduction aspect in hearing aids has been addressed primarily using traditional signal processing techniques, such as beamformers, rather than AI. Denoising, defined as real-time separation of speech from noise, involves eliminating undesired background noise to enhance the clarity of the sounds a user wants to hear. Recent evidence suggests that Deep Learning, a specific branch of AI, can further reduce background noise and offers dramatic improvements in speech intelligibility for hearing aid users (Diehl et al., 2023).

Decoding artifical intelligence

Understanding the foundations of AI and ML is crucial for recognizing the significant benefits they bring to hearing aids.

Today, AI is used as a rather broad term that encompasses a wide range of technologies and techniques that vary significantly in their complexity. On the one hand, it is employed to address modern training-based ML approaches that have gained attention for achievements like mastering complex games, generating realistic images, or engaging in human-like conversations. On the other hand, this term also encompasses more traditional programming techniques such as rule-based and expert systems and basic ML algorithms. Classical (non-AI) algorithms are often used to tackle problems with well defined objectives, whereas AI shines best in solving problems for which the objective can be better described with examples rather than rules. The latter is usually the case when the objective itself has a "human element" to it.

Machine learning

A subset of AI that currently sees great momentum is that of ML which optimizes some mathematical formula (Model) containing adjustable values (Parameters) to perform a particular task. This adjustment is generally a statistical procedure that builds on a set of example data points (Training Data). The goal of this procedure is to have the model correctly described by the training data while also providing a correct output for previously unseen or unlearned data. This process is commonly referred to as "generalization". The behavior of the model after this procedure is largely defined by the processed examples, thus this procedure is referred to as "training" or "learning". It is analogous to the human cognitive abilities which allow us to derive knowledge from examples.

After the model is completely trained, it can be used for inference — it will process real input data to produce output data according to the parameters that have been trained (Figure 1).

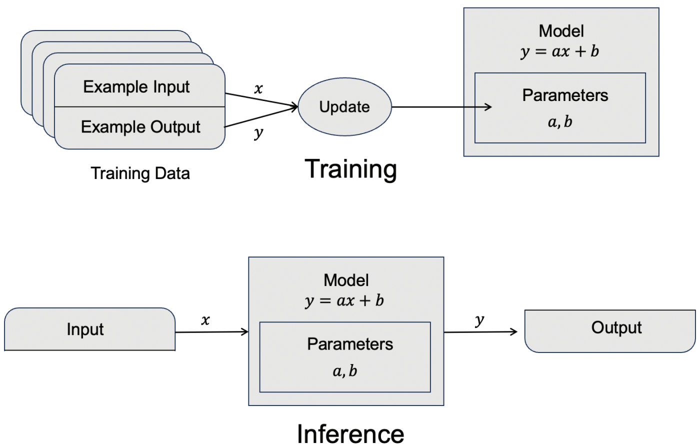

Figure 1: Training and Inference for ML models on the example of a linear model. During training, an update procedure adjusts the model parameters according to some training data. In inference, the model uses the previously trained parameters to compute an output from the given input.

Deep neural networks

Perhaps the most sophisticated and most discussed model type in ML to date are DNNs. Given the right conditions, very complex data with numerous inputs and outputs with complex relationships can be modeled very precisely with this approach. Just as with many other ML methods, the model represents some relatively simple calculations of its input data, that depend on its parameters. The most striking difference to other approaches are that DNN models generally include many more of these calculations and, thus, more parameters (millions or billions).

With a higher number of parameters, given enough data points and a suitable optimization procedure, these kinds of models can adapt to a large variety of problems with unprecedented generalization capabilities. Prominent examples include analyzing the contents of images or generation of believable text which is used for chatbots like ChatGPT. In both examples, the input data varies widely (it is hard to foresee what kind of images or text will be seen in a future input), however the DNNs still perform well. The demands and limitations in terms of computational resources depend a lot on the target domain (Figure 2).

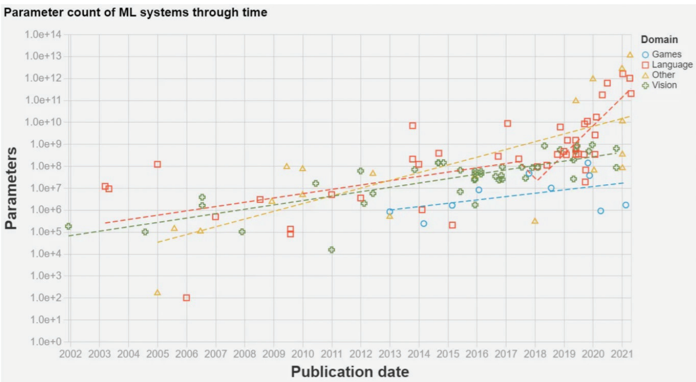

Figure 2: DNN Model size is steadily increasing with the availability of more potent hardware. At the same time, DNN sizes "spread out" across the different domains according to their respective demands and limitations, so that we see 7 orders of magnitude (factor 10 million) difference in model sizes these days (Sevilla 2021).

A variety of factors can impact the success of the training process for a large DNN. A thorough understanding of the specific problem and target function better supports the optimization procedure.

It has been shown that more, high-quality training data will improve the final model performance. This ensures better generalization to a larger variety of different inputs. To allow adaptation to the complexity of the training data, successful models also need to be scaled up in terms of number of computations and parameters (Koutsoukas, 2017; Sun, 2017; Sutton, 2019; Hestness, 2017; Bahri, 2021).

However, with increased parameters and computations, more powerful hardware is required to support the training process, as well as during interference. This is a limitation within the current landscape of hearing aid technology, as no hearing instrument is effectively equipped with the technological capabilities to meet these demands

Minimum viable deep neural network

Depending on the concrete problem at hand, DNNs require a certain "minimum size" (in terms of parameters & computation) to provide a benefit over comparable classical or other kinds of ML algorithms.

We can illustrate this with a simple experiment: We trained DNNs with various sizes and complexities for the task of denoising and compared it to one of the current state-of-the-art denoising algorithms in terms of how well it performs in the task according to some performance metric ("score") and how many parameters and how much computation it requires ("parameter count").

While the DNNs, given enough parameters eventually perform better than the classical approach, no DNN can outperform the classical algorithm with the same number of parameters (Figure 3).

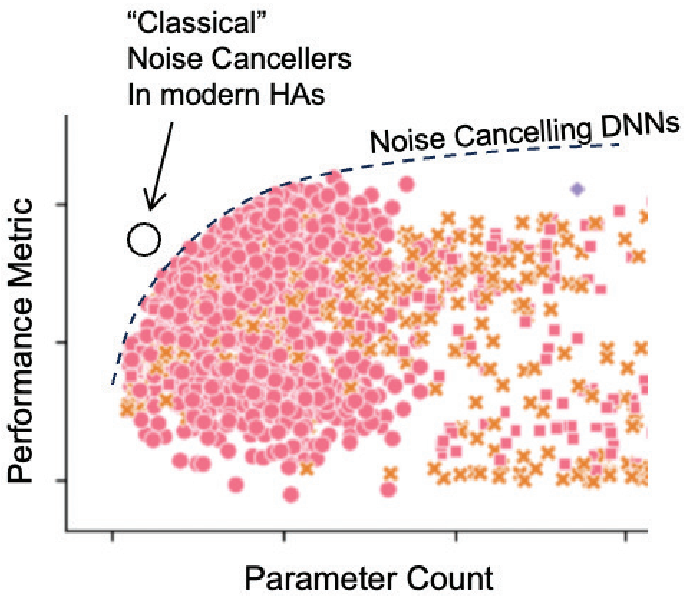

Figure 3: A population of DNNs trained to fulfill the role of noise cancellation compared to a noise cancellation algorithm used in a modern hearing aid. For the parameter counts supported by existing hardware, classical algorithms are still at an advantage."Performance Metric" measures denoising performance.

Various techniques have been developed to reduce and optimize the size of DNNs without impacting performance (Hongrong, 2023). While today's DNNs make use of these optimization techniques, they still generally require significantly more computational resources than classical approaches.

Artifical intelligence in hearing aids

AI has found diverse applications in hearing aids, with extensive ongoing research efforts and the continual emergence of innovative technologies aimed at addressing the challenges encountered by users at several stages of their hearing aid journey. Modern AI-powered hearing aids optimize the fitting process and offer various functionalities such as classification and health and fitness tracking, expanding their utility beyond conventional hearing assistance. This transformation of hearing aids into smart wearables not only enhances their functionality but could also aid in further reducing the stigma associated with hearing aids that persists today.

Still, we believe a more committed application of AI would have the potential to address a long-standing and frequent complaint among hearing aid users: hearing in noisy environments. Background noise has long been recognized as one of the biggest hurdles for hearing aid wearers (e.g., Abrams et al., 2015; Ng & Ronnberg, 2020; Bottalico et al., 2022). Early generations of hearing aids would amplify all sounds indiscriminately. In noisy settings this created a suboptimal listening experience and lead to users' frustration and withdrawal from certain social situations, negatively impacting their quality of life and well-being. The amplification of background noise has been a commonly cited cause for the reluctance to adopt hearing aids (Hougaard & Ruf, 2011; Hartley et al., 2010).

Most modern hearing aids use spatial selectivity techniques such as beamforming to reduce noise, which has been shown to improve speech intelligibility and reduce listening effort (Latzel, M. et al., 2022; Adler, M. & Seitz-Paquette, K., 2023). However, conversation in noise remains one of the situations in which hearing aid owners have the least satisfaction (Appleton-Huber, 2022). Hearing aid wearers encounter a large variety of acoustic environments, each presenting its own unique set of noise challenges in our inherently noisy world. AI has promising potential to play a transformative role in reshaping hearing aids to meet the demands of contemporary life. AI enables the creation of more personalized and adaptive solutions, tailoring amplification and noise reduction algorithms to the unique needs of each hearing aid user in any given auditory environment. Moreover, with their continuous learning capabilities, AI-based hearing aids could improve their performance over time, gradually perfecting the listening experience.

Artificial intelligence and machine learning-based sound processing in hearing aids today

Several AI-based technologies have already been incorporated into hearing aid devices and are commercially available. For example, AutoSense OS has been developed with ML to precisely recognize and adjust to various sound environments, significantly improving speech understanding (Appleton, 2020). By continually analyzing the user's surroundings, it dynamically fine-tunes the hearing aids to suit the noise levels and acoustic conditions. It combines a variety of features of the input audio signal to create customized settings, providing tailored auditory experiences. This eliminates the need for manual adjustments, allowing users to enjoy enhanced hearing performance in any situation.

Sound scene classification

Sound scene classification refers to the task of identifying the appropriate sound scene and how it should be treated. Is the wearer on the street, in a car, in a windy environment or in a restaurant? Each of these situations requires different approaches to sound processing and thus a correct decision is crucial for a good user experience.

This decision is easier for ML approaches to handle while harder to address with classical algorithms due to various factors influencing a good decision (as that has a 'human element' to it). A DNN is often not required (but can be beneficial) for these kinds of classification problems in which there are only a limited number of possible outputs, in contrast to problems like denoising in which the output is a continuous signal. This is like the situation in Figure 3: Only with sufficient computing power will a DNN provide a benefit over other solutions.

Accordingly, many hearing aid manufacturers today have adopted ML (but generally not deep learning) approaches for scene classification tasks, Phonak's AutoSense OS being no exception. This allows a "best of both worlds" approach, we can utilize the power of "learning" approaches to cover a wide range of situations with high accuracy but can avoid the increased computational resources required by the DNN to perform the processing on the hearing aid.

Speech and noise separation

The most challenging situation a hearing aid has to perform in, is one where the target sound, speech, is competing with a lot of noise. One of the interesting metrics in this scenario is that of speech intelligibility, that is, how much of the speech can be understood by the user.

Considerable development has gone into the effective application of multiple microphone technology to spatially improve the signal-to-noise ratio (SNR) in these situations. For this technology to be effective, first the hearing aid has to correctly classify the current sound scene, considering the users preference. From this, it can derive what the target sound is, and whether it is concentrated in a small area (such as a person speaking) or more spread out around the user. This job is done by modern sound scene analysis software such as AutoSense OS.

Second, combining the signals from the microphones to select the direction or directions of interest. This step is commonly referred to as beamforming.

This approach has shown impressive results and is today considered state of the art in denoising technology. This success has led to increased demand: Users explore more challenging sound scenes and expect improved hearing in more and more challenging situations.

In this environment, features like SpeechSensor, part of Phonak's SmartSpeech Technology, accurately detects the direction of the dominant speaker and sends this information to AutoSense OS 5.0 to adjust the directional microphone mode accordingly. This results in an average of 15% better speech understanding in noise when speech is coming from the side and behind (Woodward, J. et al., 2022).

In some of these situations, even beamforming approach may have difficulty fully removing the noise. Beamformers inherently assume that the audio source of interest and the distracting noise can be separated by direction. In certain sound scenes we might however wish to clearly understand speech from all directions but at the same time remove unwanted background noise from the surroundings. Social situations with multiple speakers, such as having a conversation with group of friends in a restaurant, is one example where beamformers may be of limited utility to the user.

The future of speech and noise separation

This very problem is well addressed with DNNs on more powerful devices and today, solutions are found commonplace in video conferencing software, mobile phones and a variety of other applications. Recently, remarkable effects on intelligibility with these methods have been shown that could theoretically be applicable to hearing aids (Diehl, 2023, Andersen, 2021) (Figure 4).

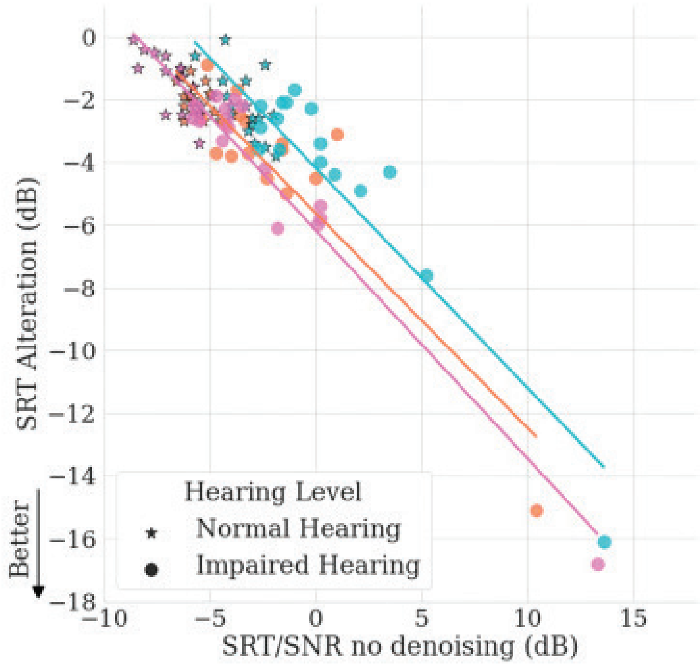

Figure 4: DNN-based denoising can significantly improve intelligibility for HI and NH users in a variety of situations (left is better hearing, lower is more improvement). Here a model with ~4M parameters is used (Diehl, 2023).

Hearing aid manufacturers have started to investigate the use of DNNs for denoising indirectly by using DNNs not only for sound scene classification but also for "steering" the denoising algorithm to improve its performance. While this is a crucial step towards better utilization of the available DNN technology, it is not yet powerful enough to deliver results close to what is seen in Figure 4.

In a moderately noisy situation of about 2–3dB Signal-to-Noise Ratio (SNR), modern hearing aids provide an SNR improvement of around 4–8dB, depending on a variety of factors (Husstedt, 2021). In comparable situations, DNNs have the potential to hit the 12dB SNR mark (Gao, 2023). This shows that the DNN-based denoising technology offers a potential that current hearing aids do not have access to yet.

Several reasons may play into this, but we believe a large factor for this is the challenge in providing capable-enough hardware, with a sufficiently low power consumption, enabling very small form factors.

Conclusion

Artificial intelligence has demonstrated tremendous potential to influence the hearing healthcare industry, from fitting to denoising in difficult environments, but is sometimes advertised in confusing ways. In this Phonak Insight we demystified the inner workings of machine learning and deep neural networks and pointed out the technical limitations that currently hinder their full adoption in the field. Finally, we showed the great potential of a hypothetical application of deep neural networks for advanced denoising in hearing aids that has yet to reach the market.

Bibliography

Adler, M., & Seitz-Paquette, K. (2023). Advancements in beamformer technology demonstrate an advantage of a steered beamformer vs. static for off-axis speech in noisy environments. *Phonak Field Study News* retrieved March 18th, 2024, from www.phonak.com/evidence

Andersen, A. H., Santurette, S., Pedersen, M. S., Alickovic, E., Fiedler, L., Jensen, J., Behrens, T. (2021). Creating Clarity in Noisy Environments by Using Deep Learning in Hearing Aids. *Seminars in Hearing*, 42, 260-281.

Appleton, J. (2020). AutoSense OS 4.0 - significantly less listening effort and preferred for speech intelligibility. *Phonak Field Study News* retrieved March 18th, 2024, from www.phonak.com/evidence

Appleton-Huber, J. (2022). What Is Important to Your Hearing Aid Clients… and Are They Satisfied? Retrieved March 18th, 2024, from https://hearingreview.com/hearing-loss/patient-care/counseling-education/what-important-to-your-hearing-aid-clients-are-they-satisfied

As'ad, H., Bouchard, M., & Kamkar-Parsi, H. (2019). Beamforming designs robust to propagation model estimation errors for binaural hearing aids. IEEE Access, 7, 114837-114850.

Bahri, Y., Dyer, E., Kaplan, J., Lee, J., & Sharma, U. (2021). Explaining Neural Scaling Laws. https://arxiv.org/pdf/2102.06701.pdf.

Balling, L. W., Mølgaard, L. L., Townend, O., & Nielsen, J. B. B. (2023). The collaboration between hearing aid users and artificial intelligence to optimize sound. *Seminars in Hearing*, 42(3), 282-294.

Bottalico, P., Piper, R. N., & Legner, B. (2022). Lombard effect, intelligibility, ambient noise, and willingness to spend time and money in a restaurant amongst older adults. *Sci Rep 12*, 6549. https://doi.org/10.1038/s41598-022-10414-6.

Cheng, H., Zhang, M., Shi, J. Q. (2023). A Survey on Deep Neural Network Pruning: Taxonomy, Comparison, Analysis, and Recommendations. https://doi.org/10.48550/arXiv.2308.06767

Chong, F. Y., & Jenstad, L.M. (2017). A critical review of hearing-aid single-microphone noise-reduction studies in adults and children. *Disability and Rehabilitation: Assistive Technology*, 13(6), 600-608, DOI: 10.1080/17483107.2017.1392619.

Crowson, M. G. (2020). Artificial Intelligence to Support Hearing Loss Diagnostics. *The Hearing Journal*, 73(11), 8-9.

Diehl, P. U., Singer, Y., Zilly, H., Schönfeld, U., Meyer-Rachner, P., Berry, M., Sprekeler, H., Sprengel, E., Pudszuhn, A., Hofmann, V. M. (2023). Restoring speech intelligibility for hearing aid users with deep learning. *Sci Rep.13(1)*, 2719. doi: 10.1038/s41598-023-29871-8. PMID: 36792797; PMCID: PMC9932078.

European Commission (2018). Coordinated Plan on Artificial Intelligence. Available at https://digital-strategy.ec.europa.eu/en/policies/plan-ai

Hartley, D., Rochtchina, E., Newall, P., Golding, M., & Mitchell, P. (2010). Use of hearing aids and assistive listening devices in an older Australian population. *Journal of the American Academy of Audiology*, 21, 642–653.

Hestness, J., Narang, S., Ardalani, N., Diamos, G. F., Jun, H., Kianinejad, H., Patwary, M. A., Yang, T., Zhou, Y. (2017). Deep Learning Scaling is Predictable, Empirically. CoRR abs/1712.00409 https://arxiv.org/pdf/1712.00409.pdf.

Hougaard, S., & Ruf, S. (2011). EuroTrak I: A consumer survey about hearing aids in Germany, France, and the UK. *Hearing Review*, 18(2), 12-28.

Husstedt, H., Kreyenhagen, J., Langhof, L., Kreikemeier, S., Denk, F., Wollermann, S. & Frenz, M. (2021). Using the phase inversion method and loudness comparisons for the evaluation of noise reduction algorithms in hearing aids. *Acta Acust.*, 5(5), 41. https://doi.org/10.1051/aacus/2021036.

Koutsoukas, A., Monaghan, K. J., Li, X., & Huan, J. (2017). Deep-learning: Investigating deep neural networks hyper-parameters and comparison of performance to shallow methods for modeling bioactivity data. *Journal of Cheminformatics 9(1)*, 42. https://doi.org/10.1186/s13321-017-0226-y.

Latzel, M., Lesimple, C., & Woodward, J. (2022). New implementation of directional beamforming configurations shows improved speech understanding and reduced listening effort. *Phonak Field Study News* retrieved March 18th, 2024, from www.phonak.com/evidence

Ng, E. H. N. & Rönnberg, J. (2020). Hearing aid experience and background noise affect the robust relationship between working memory and speech recognition in noise. *International Journal of Audiology*, 59(3), 208-218.

Peng, L., Fang, S., Fan, Y., Wang, M., Ma, Z. (2023). A Method of Noise Reduction for Radio Communication Signal Based on RaGAN. *Sensors 23(1)*, 475. doi: 10.3390/s23010475. PMID: 36617068; PMCID: PMC9823341.

Sevilla, J. (Jul 2, 2021). Parameter counts in Machine Learning. https://towardsdatascience.com/parameter-counts-in-machine-learning-a312dc4753d0.

Sun, C., Shrivastava, A., Singh, S., & Gupta, A. (2017). Revisiting Unreasonable Effectiveness of Data in Deep Learning Era. https://doi.org/10.48550/arXiv.1707.02968

Sutton, R. (2019). The Bitter Lesson. http://incompleteideas.net/IncIdeas/BitterLesson.html

Woodward, J., & Latzel, M. (2022). New implementation of directional beamforming configurations show improved speech understanding and reduced listening effort. *Phonak Field Study News* retrieved March 18th, 2024, from www.phonak.com/evidence