2023 IEEE 21st World Symposium on Applied Machine Intelligence and Informatics (SAMI) | 979-8-3503-1986-6/23/$31.00 ©2023 IEEE | DOI: 10.1109/SAMI58000.2023.10044486

A Comprehensive Investigation into The Noise Reduction Techniques for Speech

1st Suryakant Tyagi
Doctoral School of Applied Informatics and Applied Mathematics
Óbuda University, Budapest, Hungary
suryakant.tyagi@phd.uni-obuda.hu

2nd Annamária R. Várkonyi-Kóczy
Inst. of Software Design and Software Development,
Óbuda University, Budapest, Hungary
varkonyi-koczy@uni-obuda.hu

3rd Sándor Szénási
John von Neumann Faculty of Informatics, Óbuda University, Budapest, Hungary
Faculty of Economics and Informatics, J. Selye University, Komárno, Slovakia
szenasi.sandor@nik.uni-obuda.hu

***Abstract* – There has been a ton of study on the use of machine learning for speech processing applications, particularly voice recognition, over the past few decades. However, in recent years, research has concentrated on using deep learning for applications that relate to voice. In several applications, including speech, this new branch of machine learning has outperformed others, making it an extremely appealing topic for research. This paper presents a comprehensive review of the studies on deep learning for speech applications that have been carried out since 2017 when deep learning emerged as a new field of machine learning. This evaluation provides a rigorous statistical analysis that was completed by removing certain data from 184 papers published between 2017 and 2022. The findings presented in this paper shed light on the patterns of research in this field and concentrate attention on fresh research areas.**

Keywords* – *Noise Reduction, Machine Learning, Deep Learning, Speech Processing

I. INTRODUCTION

The background indoor and outdoor noise is altering the speech signal during the recording. The background noise is the major challenge for a lot of sample recording speeches from professional speakers, celebrities, and leaders [1]. One of the disgusting problems is speech mixtures with unwanted noise and the problem is faced by people hearing disturbing noise [2]. The statics term used for describing the randomness of the noise in order of first and second statistics resolves is very complicated without knowing the background noise verification moments [3]. One of the effective algorithms is subtractive noise reduction for improving the SNR on LPC-based spectral parameters related to the performance of speech processors operating with input SNR of 15 dB and below [4]. The improvement of speech devices and spectral estimation is removing the less than or full amount of noise [5]. The deep neural network used in the training stage for optimizing the ideal ratio mask to the gaining function of classic speech enhancement and testing stage implemented by the GF-DNN-IRM model without using the conventional speech enhancement process. The challenging task on CHiME-4 results showed that given algorithm relative world error reduction accuracy result of 6.57% on the Real Data test set without acoustic model retraining in causal mode. The ASR performance cannot be improved by the traditional DNN-IRM method [6]. Deep learning techniques are better to overcome the limitations of conventional methods. Multiplying noisy input spectrogram features. The result is optimized based on long short-term memory (LSTM) units through the recurrent neural network. The relative word-error rate reduction of 47.73% and PESQ improvement of 0.76 based on (LSTM-MT) architecture [7]. This work uses characteristics including short-term energy, periodicity, and spectrum flatness to build a speech activity identification algorithm. The optimum filter for noise reduction performance has been discovered to be the speech distortion-weighted Wiener filter [8]. Discrete wavelets transform (DWT) is employed as a gain function for post-processing, and SNR values are made up utilizing coherent function [9].

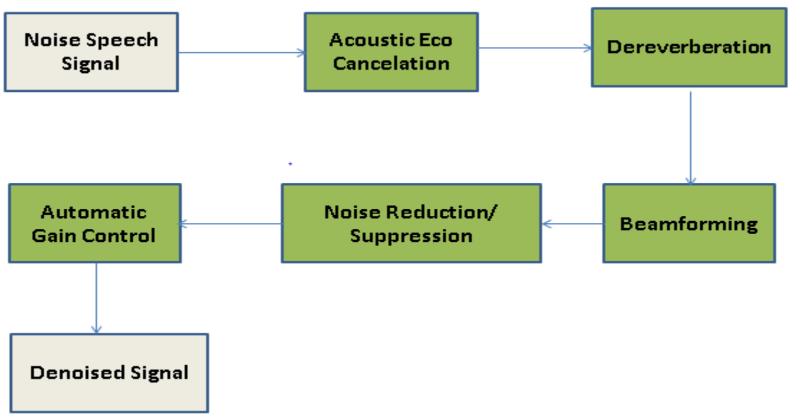

Fig. 1. Architecture of Noise reduction in Speech

After that, utilizing extracted audio features from the denoised audio, the speech recognition decoder detects the speech and turns it into a Unicode document using acoustic, language, and pronunciation models. Additionally, noise reduction models are exposed as a secondary solution for multilingualism noise reduction, supporting English, and detached from the preservation approach [10]. A normal speech signal from the NOIZEUS database and an impaired speech signal from individuals with Parkinson's disease from the MDVR-KCL dataset is used to assess the proposed filter model's ability to reduce noise [11]. The NPAS filtered the crowd noises, demonstrating the simultaneous identification and digitization of individual speakers' speeches [12]. The simulations demonstrate that, in comparison to current ICA methods and other traditional methods, the proposed adaptive ICA method offers greater SNR. Since all real-time communication devices use adaptive ICA, noise cancellation is effectively performed [13]. In order to properly recover Lombard speech created at SNRs as low as 10 dB, both suppression approaches must provide appropriate noise reduction, according to assessments of pitch extraction, energy level, and objective speech intelligibility and quality [14]. The simulation findings support the value of the wavelet denoising technique when applied as a pre-processing step for noise reduction, as well as the superiority of transform domain separation over time domain separation [15]. The experimental results demonstrate that our system outperforms other state-of-the-art methods in terms of noise reduction, reduced voice distortion, and greater speech intelligibility [16].

II. Related work

However, the speech augmentation algorithm's deep learning technique reduces the performance of speech, particularly for undetectable noises and invisible speakers. Deep learning models are therefore constrained to a restricted number of speakers [17]. Therefore, a dense Non-Negative Matrix Factorization method is used to denoise the noisy speech signal before the SER achieves noise robustness in order to overcome this shortcoming and improve the SER performance in noisy settings [18]. Figure 2 shows the year-by-year progress of development activities from 2017 to 2022. The graph shows that there has been a significant advancement in research work over the last three years. This suggests that while this topic is receiving a lot of attention, there is still a lot of research to be done in this field. When real-time data is used, the system becomes complicated, and noise reduction becomes difficult. This depicts the noise reduction techniques for speech progress.

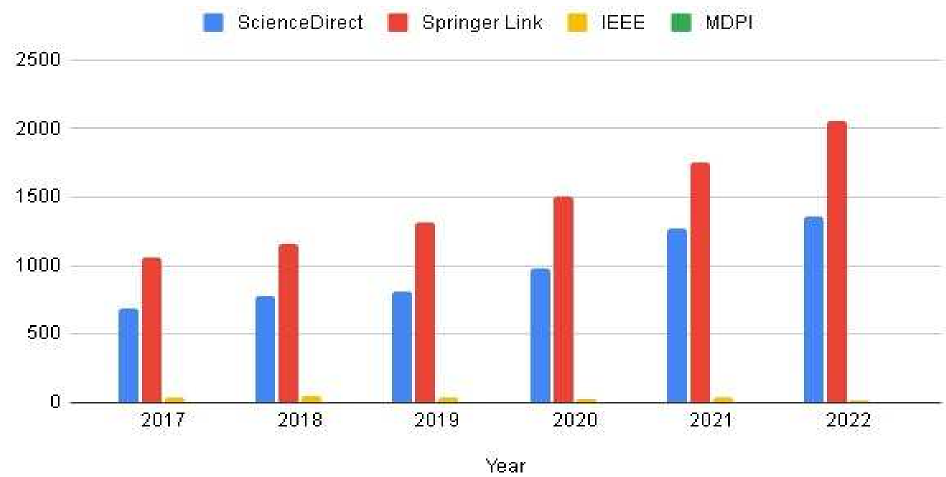

Fig. 2. Recent development activities in SER from 2017 to 2022.

The dual voice activity detection (VAD), convolutive blind source separation (CBSS), and noise cancellation blocks used in the speech-protected noise cancellation system are presented. Different feature extraction techniques are used in the VAD block, as well as a deep neural network (DNN), which extracts features from raw data, to determine the feature that best identifies speech in a noisy environment [19]. The results of the listening test show that, in terms of immediate noise attenuation, the signals processed by the blocking-based algorithms are much favored over the noisy signal. The results of the listening test data analysis also support the judgments made based on the unbiased assessment [20]. The two parts of this research project are the voice enhancement and modeling phases. In order to improve the noisy Tamil voice signal, the modified Modulation Magnitude Estimation based Spectral Subtraction with Chi-Square Distribution based Noise Estimation (SS-NE) approach is proposed [21].

III. OVERVIEW OF MACHINE LEARNING AND DEEP LEARNING

DL has gained more interest recently and has become a new field of study in machine learning. Deep learning has gained more interest recently and has become a new field of study in machine learning. It also emphasizes the numerous voice databases and evaluation measures that DL algorithms use to assess their performance [22]. Finding a deep neural network-based mapping function between noisy and clean speech data is a supervised strategy for improving speech [23]. The main intention of the augmentation system is to develop the perceptual quality of communication or speech. In order to improve the odds of reducing noise and restoring the original signal, artificial intelligence and machine learning algorithms were included in every sector. Deep transfer learning was used in this work to remove noise from the data and restore the original signals. This proposed approach includes a filtration scheme instead of using a convolution layer [24]. They investigated the clinical effectiveness of a novel deep learning-based noise reduction approach for Mandarin-speaking cochlear implant recipients under noisy conditions with challenging noise types at low signal-to-noise ratio (SNR) levels. This study's deep learning-based NR approach is made up of two modules: noise classifier and deep denoising autoencoder [25]. They investigate the clinical effectiveness of a novel deep learning-based noise reduction approach for Mandarin-speaking cochlear implant recipients under noisy conditions with challenging noise types at low SNR levels. This study's deep learning-based NR approach is made up of two modules: noise classifier and deep denoising autoencoder [26]. Multichannel acoustic echo cancellation, microphone array processing and dereverberation techniques for signal enhancement, reliable wake-up word, and end-of-interaction detection, high-quality speech synthesis, and sophisticated statistical models for speech and language learned from large amounts of heterogeneous training data are examples of these technologies. Deep learning has played an important role in all of these fields [27]. In this study, a speech enhancement algorithm based on noise classification was proposed to improve speech quality in a hearing aid environment by using noise reduction algorithms with deep neural network learning. Ten types of noise were self-recorded and classified using convolutional neural networks to evaluate speech enhancement in an actual hearing aid environment. Furthermore, deep neural networks were used to apply noise reduction for speech enhancement in hearing aids based on a noise classification. As a result, the speech quality based on the deep neural network-removed speech enhancements and associated environmental noise classification showed a significant improvement over the conventional hearing aid algorithm [28]. DNN is used to estimate ideal ratio masks at individual time-frequency bins, which are then used to design three potential speech enhancement systems: single-channel ego-noise reduction, multi-channel beamforming, and multi-channel time-frequency spatial filtering [29]. In recent years, noise reduction techniques in speech have gained much attention, especially after the popularity of social media. The peer-reviewed journals have shown a significant increase in the past few years. Science Direct shows 7,184 results on the topic out of which 955 are review articles, 4,935 are research articles, 99 are encyclopedias, 869 are book chapters, 29 are conference abstracts, 22 are editorials, and the rest are other types of articles. On the other hand, MDPI shows 6 search results on this topic. Springer shows 18,920 search results. Out of this, 5,979 are chapters, 6,628 are articles, 6,048 are books, 2,275 are conference papers and proceedings, and the rest are other documents. Finally, IEEEXplore shows 1,165 search results. This has 936 conferences, and 206 journals, the rest are other documents. Figure 3 shows the year-wise development works from 2017 to 2022. The graph shows that in the past six years, there is a huge development in the research work. This suggests that this topic is gaining huge attention but there is still too much research work that needs to be done in this field.

Table 1. Recent novel work in noise reduction techniques.

References	Year	Source	Application	Module/Algorithm
30	2018	Computation and Language	Detecting Hate Speech and Offensive Language	N-gram and TFIDF
31	2018	Innovations in Electronics and Communication Engineering	Noise Destruction and Improvement	GSO optimization algorithm
32	2019	ICAEM	Hate Speech Detection, Semantic word embedding	CNN
33	2019	Computation and Language	Hate Speech Detection	CBOW and RoBERTa, Natural language processing
34	2020	ACM Transactions on Internet Technology	Automatic hate speech detection	Bidirectional Long Short Term, Gated Recurrent Unit
35	2020	Speech and Computer	End-to-End Speech Recognition	MP3 compression
36	2020	12th ACM Conference on Web Science	Hate Speech Detection via Multi-Faceted	DeepHate
37	2020	Computing-Springer	Automatic hate speech detection	KNLPEDNN
38	2021	SAGE Journal	Intelligent detection of hate speech	RF, DT, SVM, NB
39	2021	Journal of Intelligent Systems	Bangla hate speech detection	NLP, Attention-based encoder–decoder model
40	2021	SEPLN 2021	Sexism detection, data augmentation	RF
41	2021	Speech Communication	Sexism Identification	Modulation frequency

IV. RESULT AND DISCUSSION

The 186 papers included in the study can be divided into four primary categories: conference papers, journal papers, workshop papers, and publications from research institutes. As can be seen above, most of the papers used in the study, 67% were classified as conference papers. The remaining 33% were split among journal articles, workshop papers, and research institute articles at rates of 14%, 10%, and 09%, respectively. A thorough statistical analysis of the various conferences and publications in which these papers were published was developed in order to provide even more information about the articles that were used. As can be seen, ICASSP, the "IEEE International Conference on Acoustics, Speech and Signal Processing," published most conference papers or roughly 54% of them. 31% of those were then published in Interspace. The remaining 15% was split between the remaining conferences, with 3% going to Springer, the "IEEE Conference on Signal and Information Processing," 2% going to nature, and then 1% going to each of the other 15 conferences. Different speech-related topics, such as speaker identification, speech emotion recognition, voice enhancement, speech recognition, and speech transcription, were found among the 174 papers. A deeper analysis was done on these papers in order to uncover even more information about the various voice recognition regions each article is focused on because a significant portion of papers falls under the category of speech recognition. There were many different subfields found, including automatic speech recognition, large vocabulary speech recognition, low resource speech recognition, multilingual speech recognition, noise-robust speech recognition, phone recognition, sequence classification, speaker adaptation, and speech separation, among others. In the study publications, many datasets were used to develop and evaluate algorithms. While some databases were private, the majority 89% were public and accessible the in the study publications, many datasets were used to develop and evaluate algorithms.

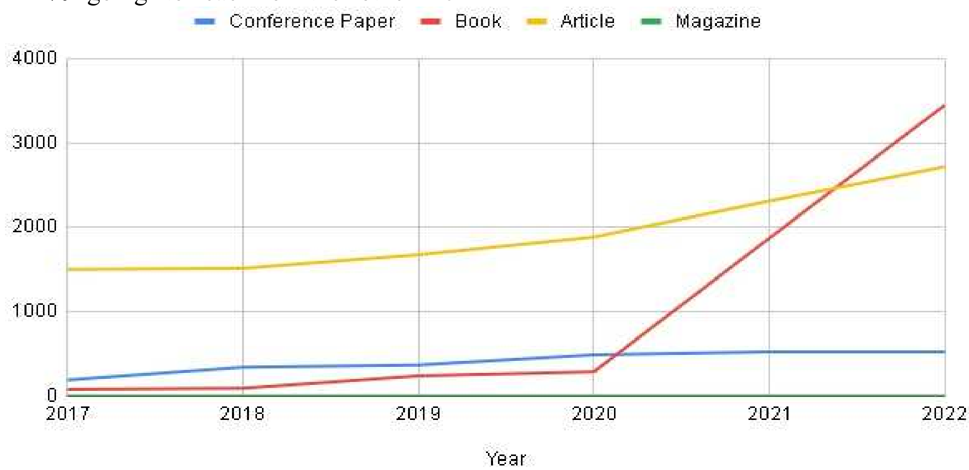

Fig. 4. Development of the number of scientific documents in noise reduction using ML techniques**.**

Some of the databases were secret, but 83% of them were public and accessible online. Algorithms were trained and tested in a variety of environments, including noisy, neutral, and emotional ones. In 68% of the papers, the use of a neutral environment was either specified, or nothing was indicated at all, leaving neutrality to be presumed. Building a noisy robust system was cited in 28% of the studies regarding noisy environments. The use of emotional speech to test and train the algorithms was only addressed in 4% of the articles. Out of 184 studies on the different deep neural network models used in voice recognition, 132 papers employed independent DNN models, and 42 papers used hybrid models of DNN and LSTM.

V. CONCLUSION

This research extracted specific data from 186 papers published between 2017 and 2022 to give a complete statistical analysis of the use of deep learning in speech-related applications. Over 67%, 19%, 14% of the papers that were found were published in ICASSP, majority of conference papers and the journal papers were printed in the IEEE. The speech intelligibility model described here could serve as a useful starting point for investigating and better comprehending the interactions between the effects of auditory signal processing, acoustic preprocessing, and speech intelligibility. A greater understanding of the connections between background noise signal exposure history and speech-in-noise test performance as well as the most sensitive and relevant procedures to assess functional hearing ability in diverse sectors are urgently needed. Additional research into damage-risk interactions is urgently needed because the hazards of supra-threshold function are unclear. The Win and the HINT are two examples of speech-in-noise tests that have been devised in the research lab. It is interesting to observe that the majority of researchers still extract features from voice signals using MFCCs in deep learning models. When utilizing deep learning models, it is worthwhile to give various feature extraction techniques, like linear predictive coding. Another finding is that there isn't much research being done on speech recognition using recurrent neural networks. Since CNN, recurrent neural network models, particularly Long Short-Time Memory are exceptionally effective at voice recognition, authors are strongly encouraged to perform research using them in the future.

REFERENCES

[1] Rahman, M., Pranto, S., Ema, R. R., Anfal, F., & Islam, T. (2022). Application of Noise Reduction Techniques to Improve Speaker Verification to Multi-Speaker Text-to-Speech Input. In International Conference on Innovative Computing and Communications (pp. 43-56). Springer, Singapore.

[2] Hamsa, S., Iraqi, Y., Shahin, I., & Werghi, N. (2021). Dominant voiced speech segregation and noise reduction pre-processing module for hearing aids and speech processing applications. In International Conference on Soft Computing and Pattern Recognition (pp. 395-403). Springer, Cham.

[3] Yeldener, S., & Rieser, J. H. (2000, June). A background noise reduction technique based on sinusoidal speech coding systems. In 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No. 00CH37100) (Vol. 3, pp. 1391-1394).

[4] Nuha, H. H., & Absa, A. A. (2022, July). Noise Reduction and Speech Enhancement Using Wiener Filter. In 2022 International Conference on Data Science and Its Applications (ICoDSA) (pp. 177-180). IEEE.

[5] Kushner, W. M., Goncharoff, V., Wu, C., Nguyen, V., & Damoulakis, J. N. (1989, May). The effects of subtractive-type speech enhancement/noise reduction algorithms on parameter estimation for improved recognition and coding in high noise environments. In International Conference on Acoustics, Speech, and Signal Processing, (pp. 211-214). IEEE.M.

[6] Chen, Y. Hao, Task offloading for mobile edge computing in software defined ultra-dense network, IEEE J. Sel. Areas Commun. 36 (3) (2018) 587–597.

[7] Thakallapelli, A., Ghosh, S., & Kamalasadan, S. (2016, July). Real-time frequency based reduced order modeling of large power grid. In 2016 IEEE Power and Energy Society General Meeting (PESGM) (pp. 1-5). IEEE.

[8] Pejić, Aleksandar, and Piroska Stanić Molcer. "Predictive machine learning approach for complex problem-solving process data mining." *Acta Polytechnica Hungarica* 18, no. 1 (2021): 45-63.

[9] Bagekar, S., & Tank, V. (2018, June). Dual channel coherence based speech enhancement with wavelet denoising. In 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 1826-1830). IEEE.

[10] Aravinda, K. P., Sandeepa, K. G. H., Sedara, V. V., Chamodya, A. K. Y. L., Dharmasena, T., & Abeygunawardhana, P. K. (2021, December). Digital Preservation and Noise Reduction using Machine Learning. In 2021 3rd International Conference on Advancements in Computing (ICAC) (pp. 181-186). IEEE.

[11] Pauline, S. H., Dhanalakshmi, S., Kumar, R., & Narayanamoorthi, R. (2022). Noise reduction in speech signal of Parkinson's Disease (PD) patients using optimal variable stage cascaded adaptive filter configuration. Biomedical Signal Processing and Control, 77, 103802.

[12] Jung, Y. H., Pham, T. X., Issa, D., Wang, H. S., Lee, J. H., Chung, M., ... & Lee, K. J. (2022). Deep learning-based noise robust flexible piezoelectric acoustic sensors for speech processing. Nano Energy, 101, 107610.Wang, H., et al., 2022. Comprehensive review of load forecasting with emphasis on intelligent computing approaches. Energy Reports, 8, pp.13189-13198.

[13] Mohanaprasad, K., Singh, A., Sinha, K., & Ketkar, T. (2019). Noise reduction in speech signals using adaptive independent component analysis (ICA) for hands free communication devices. International Journal of Speech Technology, 22(1), 169-177.Nejad, H.D., et al., 2022. Fuzzy State-Dependent Riccati Equation (FSDRE) Control of the Reverse Osmosis Desalination System With Photovoltaic Power Supply. IEEE Access, 10, pp.95585-95603.

[14] Vaziri, G., Giguère, C., & Dajani, H. R. (2019). Evaluating noise suppression methods for recovering the Lombard speech from vocal output in an external noise field. International Journal of Speech Technology, 22(1), 31-46.

[15] Hammam, H., El-Shafai, W., Hassan, E., Abu El-Azm, A. E., Dessouky, M. I., Elhalawany, M. E., ... & Fathi, E. (2021). Blind signal separation with Noise Reduction for efficient speaker identification. International Journal of Speech Technology, 24(1), 235-250.Gaussian TLBO Method. Applied Sciences, 12(16), p.7959.

[16] Martín-Doñas, J. M., Peinado, A. M., López-Espejo, I., & Gomez, A. (2019). Dual-channel speech enhancement based on extended Kalman filter relative transfer function estimation. Applied Sciences, 9(12), 2520.Tavoosi, J., et al., 2022. A machine learning approach for active/reactive power control of grid-connected doubly-fed induction generators. Ain Shams Engineering Journal, 13(2), p.101564.

[17] Bandela, S. R., & Kumar, T. K. (2021). Unsupervised feature selection and NMF de-noising for robust Speech Emotion Recognition. Applied Acoustics, 172, 107645.Akbari, E., et al., 2022. A Fault-Tolerant Cascaded Switched-Capacitor Multilevel Inverter for Domestic Applications in Smart Grids. IEEE Access.

[18] Develasco, M., Justo, R., Zorrilla, A., & Inés Torres, M. (2022). Automatic analysis of emotions from the voices/speech in Spanish tv debates. Acta Polytech. Hung, 19(5), 149-171.

[19] Usta, Ö. C., & Doğan, H. (2022). Speech protected noise cancellation system in noise dominated environments. Applied Acoustics, 189, 108577.Shakibjoo, et al., A.H. and Vandevelde, L., 2021. Optimized Type-2 Fuzzy Frequency Control for Multi-Area Power Systems. IEEE access, 10, pp.6989-7002.

[20] Azarpour, M., & Enzner, G. (2017). Binaural noise reduction via cue-preserving MMSE filter and adaptive-blocking-based noise PSD estimation. EURASIP Journal on Advances in Signal Processing, 2017(1), 1-17.Cao, Y., et al., 2021. Deep learned recurrent type-3 fuzzy system: Application for renewable energy modeling/prediction. Energy Reports, 7, pp.8115-8127.

[21] Kalamani, M., Krishnamoorthi, M., & Valarmathi, R. S. (2019). Continuous Tamil Speech Recognition technique under non stationary noisy environments. International Journal of Speech Technology, 22(1), 47-58.Tavoosi, J., et al., 2021. Modeling renewable energy systems by a self-evolving nonlinear consequent part recurrent type-2 fuzzy system for power prediction. Sustainability, 13(6), p.3301.

[22] Bhangale, K. B., & Kothandaraman, M. (2022). Survey of Deep Learning Paradigms for Speech Processing. Wireless Personal Communications, 1-37.Band, S.S., et al., 2022. When Smart Cities Get Smarter via Machine Learning: An In-depth Literature Review. IEEE Access.

[23] Xu, Y., Du, J., Dai, L. R., & Lee, C. H. (2014). A regression approach to speech enhancement based on deep neural networks. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 23(1), 7-19.Mosavi, A.H., et al., 2022. Deep learning fuzzy immersion and invariance control for type-I diabetes. Computers in Biology and Medicine, 149, p.105975.

[24] B Wu, B., Li, K., Ge, F., Huang, Z., Yang, M., Siniscalchi, S. M., & Lee, C. H. (2017). An end-to-end deep learning approach to simultaneous speech dereverberation and acoustic modeling for robust speech recognition. IEEE Journal of Selected Topics in Signal Processing, 11(8), 1289-1300.

[25] Chen, Y., Dong, B., Zhang, X., Gao, P., & Li, S. (2021). A Hybrid Deep-Learning Approach for Single Channel HF-SSB Speech Enhancement. IEEE Wireless Communications Letters, 10(10), 2165-2169.Tavoosi, J., et al., 2022. A machine learning approach for active/reactive power control of grid-connected doubly-fed induction generators. Ain Shams Engineering Journal, 13(2), p.101564.

[26] Lai, Y. H., Tsao, Y., Lu, X., Chen, F., Su, Y. T., Chen, K. C., ... & Lee, C. H. (2018). Deep learning–based noise reduction approach to improve speech intelligibility for cochlear implant recipients. Ear and hearing, 39(4), 795-809.Ardabili, S., et al., 2019, September. Deep learning and machine learning in hydrological processes climate change and earth systems a systematic review. In International conference on global research and education (pp. 52-62). Springer, Cham.

[27] Haeb-Umbach, R., Watanabe, S., Nakatani, T., Bacchiani, M., Hoffmeister, B., Seltzer, M. L., ... & Souden, M. (2019). Speech processing for digital home assistants: Combining signal processing with deep-learning techniques. IEEE Signal processing magazine, 36(6), 111-124.

[28] Chen, L. C., Chen, P. H., Tsai, R. T. H., & Tsao, Y. (2022). Epg2s: Speech generation and speech enhancement based on electropalatography and audio signals using multimodal learning. IEEE Signal Processing Letters.

[29] Wang, L., & Cavallaro, A. (2020). Deep learning assisted time-frequency processing for speech enhancement on drones. IEEE Transactions on Emerging Topics in Computational Intelligence, 5(6), 871-881.Samadianfard, S., et al., 2019. Support vector regression integrated with fruit fly optimization algorithm for river flow forecasting in Lake Urmia Basin. Water, 11(9), p.1934.

[30] Prasadh, S. K., Natrajan, S. S., & Kalaivani, S. (2017, November). Efficiency analysis of noise reduction algorithms: Analysis of the best algorithm of noise reduction from a set of algorithms. In 2017 International Conference on Inventive Computing and Informatics (ICICI) (pp. 1137-1140). IEEE.Mohammadzadeh S, D., et al., 2019. Prediction of compression index of fine-grained soils using a gene expression programming model. Infrastructures, 4(2), p.26.

[31] Dalga, D., & Doclo, S. (2013, October). Influence of secondary path estimation errors on the performance of ANC-motivated noise reduction algorithms for hearing aids. In 2013 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (pp. 1-4). IEEE.

[32] Mae, N., Ishimura, M., Makino, S., Kitamura, D., Ono, N., Yamada, T., & Saruwatari, H. (2017, February). Ego noise reduction for hose-shaped rescue robot combining independent low-rank matrix analysis and multichannel noise cancellation. In International Conference on Latent Variable Analysis and Signal Separation (pp. 141-151). Springer, Cham.Gaydhani, Aditya, Vikrant Doma, Shrikant Kendre, and Laxmi Bhagwat. "Detecting hate speech and offensive language on twitter using machine learning: An n-gram and tfidf based approach." arXiv preprint arXiv:1809.08651 (2018).

[33] Anoop, V., & Rao, P. V. (2018). Noise Destruction and Improvement of Speech Signal Quality Using Group Search Optimization (GSO) Algorithm. In Innovations in Electronics and Communication Engineering (pp. 335-342). Springer, Singapore.Sajjad, Muhammad, Fatima Zulifqar, Muhammad Usman Ghani Khan, and Muhammad Azeem. "Hate Speech Detection using Fusion Approach." In 2019 International Conference on Applied and Engineering Mathematics (ICAEM), pp. 251-255. IEEE, 2019.

[34] Nguyen, Thai Binh, Quang Minh Nguyen, Thu Hien Nguyen, Ngoc Phuong Pham, The Loc Nguyen, and Quoc Truong Do. "VAIS hate speech detection system: A deep learning based approach for system combination." arXiv preprint arXiv:1910.05608 (2019).

[35] Corazza, Michele, Stefano Menini, Elena Cabrio, Sara Tonelli, and Serena Villata. "A multilingual evaluation for online hate speech detection." ACM Transactions on Internet Technology (TOIT) 20, no. 2 (2020): 1-22.

[36] Andronic, I., Kürzinger, L., Chavez Rosas, E. R., Rigoll, G., & Seeber, B. U. (2020, October). MP3 compression to diminish adversarial noise in end-to-end speech recognition. In International Conference on Speech and Computer (pp. 22-34). Springer, Cham.

[37] Corazza, Michele, Stefano Menini, Elena Cabrio, Sara Tonelli, and Serena Villata. "A multilingual evaluation for online hate speech detection." ACM Transactions on Internet Technology (TOIT) 20, no. 2 (2020): 1-22.

[38] Al-Makhadmeh, Zafer, and Amr Tolba. "Automatic hate speech detection using killer natural language processing optimizing ensemble deep learning approach." Computing 102, no. 2 (2020): 501-522.

[39] Aljarah, Ibrahim, Maria Habib, Neveen Hijazi, Hossam Faris, Raneem Qaddoura, Bassam Hammo, Mohammad Abushariah, and Mohammad Alfawareh. "Intelligent detection of hate speech in Arabic social network: A machine learning approach." Journal of Information Science 47, no. 4 (2021): 483-501.

[40] Das, Amit Kumar, Abdullah Al Asif, Anik Paul, and Md Nur Hossain. "Bangla hate speech detection on social media using attention-based recurrent neural network." Journal of Intelligent Systems 30, no. 1 (2021): 578-591.

[41] Butt, Sabur, Noman Ashraf, Grigori Sidorov, and Alexander Gelbukh. "Sexism identification using BERT and data augmentationEXIST2021." In International Conference of the Spanish Society for Natural Language Processing SEPLN 2021, IberLEF 2021. 2021.