Jurnal Informatika Universitas Pamulang ISSN: 2541-1004
Penerbit: Fakultas Ilmu Komputer Universitas Pamulang e-ISSN: 2622-4615
Vol. 10, No. 1, Maret 2025(41-49) https://doi.org/10.32493/informatika.v10i1. 48093

Mapping The Landscape of Speech Processing Research: Trends, Insights, and Emerging Directions

Ardi Mardiana[1], Ade Bastian[2*], Muhammad Rifki[3] and Eka Tresna Irawan[4]

[123]Informatika, Universitas Majalengka, Jl. Raya K H Abdul Halim No.103, Majalengka Kulon, Kec. Majalengka, Kabupaten Majalengka, Jawa Barat, Indonesia, 45418
[4]Mayar International Pte. Ltd. 160 Robinson Road #14-04 Singapore Business Federation Center Singapore 068914
e-mail: [1]aim@unma.ac.id, [2]adebastian@unma.ac.id, [3]m.rifki726@gmail.com [4]i@mayar.id
*Corresponding author

Submitted Date: February 12th, 2025 Reviewed Date: February 27th, 2025
Revised Date: March 3th, 2025 Accepted Date: March 31th, 2025

Abstract

Speech processing has become a significant study domain within signal processing, artificial intelligence, and human-computer interaction. This work does a bibliometric analysis to ascertain research trends, notable problems, and prospective directions in voice processing. We assess significant research outputs, including publication growth, influential authors, renowned journals, and collaboration networks during the last two decades, using data sourced from credible scientific sources such as Scopus and Web of Science. The results underscore notable progress in automated voice recognition, speaker identification, and speech synthesis, while simultaneously confronting ongoing issues associated with multilingual datasets, noise resilience, and resource efficiency. Moreover, new technologies, such as deep learning and neural architecture search, are recognized as catalysts for future developments. This bibliometric study aims to provide scholars and practitioners with a comprehensive overview of the existing environment and strategic insights for advancing the voice processing domain.

Keywords: Automatic Speech; Recognition; Bibliometric Analysis; Research Trend; Speech Processing;

1. Introduction

Research in the domain of speech processing has seen substantial advancement in recent decades. The field's roots can be traced to foundational work on enhancing auditory capabilities, such as the initial research on basic speech recognition systems by M. Vilchur (1973), which emphasized the use of signal processing for persons with hearing impairments. This foundational work laid the groundwork for a deeper understanding of how we process sound, as seen in later research by C. Brodbeck and J. Z. Simon (2020), which provided profound insights into the continuous monitoring of sound waves.

Building upon this historical foundation, the field has been revolutionized by the advent of deep learning. Recent studies now emphasize the use of deep learning to enhance the precision and resilience of voice recognition systems, as articulated by S. Latif et al. (2020). This shift has allowed for more robust and accurate models that can handle the complexities of human speech.

Complementing this deep learning approach, self-supervised learning has also emerged as a powerful methodology. These model-based techniques, as significantly contributed by M. Chen et al. (2022), have further pushed the boundaries of what is possible, enabling systems to learn from vast amounts of unlabeled data, thereby reducing the need for extensive manual annotation.

Beyond these technological advancements, the applications of speech processing have also broadened. The field has moved from its audiological origins to play a crucial role in human-computer interaction and various audiological requirements. For example, studies by C. A. Kamm et al. (1997) and S. M. Pichora-Fuller and P. E. Souza (2003) demonstrated how speech processing is now integral to creating more natural interfaces and addressing the specific needs of individuals with hearing difficulties.

Jurnal Informatika Universitas Pamulang ISSN: 2541-1004
Penerbit: Fakultas Ilmu Komputer Universitas Pamulang e-ISSN: 2622-4615
Vol. 10, No. 1, Maret 2025(41-49) https://doi.org/10.32493/informatika.v10i1. 48093

Despite the substantial contributions of speech processing technology across diverse applications, several intricate hurdles persist. A primary problem is the system's resilience to acoustic interference and noise, which remains a barrier to its deployment in real-world settings (Chen, M., Wang, D., & Liu, P., 2022). Research continues to focus on enhancing performance in fluctuating acoustic environments, a challenge highlighted by Blamey et al. (1987) and Van Wassenhove et al. (2005).

Beyond these technical challenges related to the acoustic environment, other obstacles are linguistic and practical in nature. A considerable problem is adapting these technologies to languages with limited resources, as noted by Etard and Reichenbach (2019). This issue is often compounded by the variations among speakers—such as accent, intonation, and speech rate—which complicate the development of systems capable of adapting to diverse user settings (Gerkmann, T., Krawczyk, M., & Doclo, S., 2015). These linguistic and speaker-specific variations present a different kind of complexity, requiring not just noise reduction but a deeper understanding of the speech itself.

Conversely, the primary uses of speech processing technology include the creation of virtual assistants, voice-activated educational systems, and healthcare support tools. This technology is foundational to enhancing human-computer interaction, as demonstrated by Kamm, Biagioni, & Walker (1997), enabling more intuitive and natural ways for people to engage with machines. The benefits of this technology are particularly evident in two key areas. In educational settings, voice processing facilitates language acquisition, as shown by Latif et al. (2020), by providing interactive and responsive learning tools. Simultaneously, the technology's ability to track speech signals, as elucidated by Brodbeck and Simon (2020), was used to enhance communication intelligibility, especially in noisy environments, which was critical for applications like healthcare support tools and other real-world scenarios where clear communication is vital.

Future research must concentrate on creating resource-efficient deep learning models and enhancing the integration of multilingual data. Moreover, using models grounded in self-supervised learning may provide substantial improvements in system performance across diverse auditory and linguistic contexts (Etard, J., & Reichenbach, T., 2019);(F, S., Lin, Y., & Kuo, C., 2021). Future study is anticipated to further investigate the amalgamation of deep learning models with resource-efficient systems to enhance the efficacy and accessibility of this technology.

Table 1. State of The Art Speech Processing

Author(s) & Year	Number of Documents Analyzed	Sources	Finding
D Jurafsky, 2000	300	Google Scholar	A detailed analysis of the many methods used for processing speech and language
L Rabiner, BH Juang, 1993	250	Google Scholar	The fundamental building blocks of speech recognition algorithms
LR Rabiner, 1978	200	Google Scholar	Instructions on the fundamentals of digital signal processing in voice
G Hickok, D Poeppel, 2007	180	Nature reviews	Perspectives from the field of neuroscience on the brain's processing of speech
JR Deller Jr, JG Proakis, JH Hansen, 1993	150	Google Scholar	Methods for discrete-time speech signal processing are being described in detail.

So far, there seems to be no bibliometric study for these two terms. This study aims to address the following questions:

1. How are papers about speech processing classified?
2. What are the current research trends in speech processing analysis?
3. What research subjects are most commonly published?
4. What themes for future speech processing analysis may be utilized to spark additional research?

The creation of this essay started with a literature review of the idea of speech processing, informed by prior research results. Alongside Part

Jurnal Informatika Universitas Pamulang ISSN: 2541-1004
Penerbit: Fakultas Ilmu Komputer Universitas Pamulang e-ISSN: 2622-4615
Vol. 10, No. 1, Maret 2025(41-49) https://doi.org/10.32493/informatika.v10i1. 48093

1, a presentation detailing the study aims is provided. In Part 2, we defined speech processing and reviewed a recent analysis of the word. Section 3 delineated the approach used to execute the methodological phases of the bibliometric investigation concerning the utilization of databases from diverse publications. Section 4 presents the findings with the VOS Viewer. Section 5 encompasses study concepts, findings, and constraints.

2. Methods

The objective of this research was to examine a comprehensive range of published categories of speech processing. The subsequent phase involves assessing the future of speech processing, which presents avenues for additional investigation, following the identification of societal trends pertinent to speech processing research, a field characterized by an increasing volume of publications.

2.1. Search for Specific Term Speech Processing

The use of bibliometric analysis as a method for examining and assessing large datasets is becoming more prevalent. This enabled us to analyze the nuanced morphological changes that have transpired in a certain domain, while also elucidating the emergence of new disciplines within that domain. Bibliometric analysis methodologies were categorized into two basic types: (1) performance analysis and (2) scientific mapping. The primary distinction between performance analysis and mapping science was that the former evaluates the contributions of research constituents, whereas the latter examines the relationships among those constituents (Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W., 2021). The project started with a bibliometric analysis using data sourced from large academic databases, including Scopus and Web of Science. Harzing's Publish or Perish (PoP) was utilized to retrieve and organize the data based on phrases pertinent to the subject of voice processing.

2.2 Term Metrics Information

The objective of this research was to examine a comprehensive range of published categories of speech processing. The subsequent phase involves identifying prevailing social trends in speech processing research, a subject of ongoing publications, and assessing prospective themes in speech processing that provide avenues for further exploration. Table 2 presents metric data.

Table 2. Summary of Metrics Information

Metric Data	Speech Processing
Publication's years	1958-2024
Citation years	66 (1958-2024)
Papers	500
Citations	196137
Cites/year	2971.77
Cites/paper	392.27
Cites/author	106510.52
Papers/author	243.65
Authors/paper	2.74
h-index	210
g-index	439
hI,norm	139
hI,annual	2.11
hA-index	48
Paper with ACC	1, 2, 5, 10, 20, 492, 482, 387, 256, 147

Source: Publish or Perish (2024)

2.3 Reference Management

On the other hand, the story was picked up from the websites of two different publications on the internet. After that, I used the Mendeley tool to ensure that the management of references was both well-organized and efficient. The inclusion of references was essential to ensure that the article's metadata, which may include information about the author, keywords, an abstract, and other details, was accurately integrated into the document.

2.4 Bibliometric Analysis

After all of the information regarding the article has been confirmed, the following step that should be taken is to conduct a bibliometric analysis. In this step of the process, the bibliometric analysis approach includes the utilisation of VosViewer as one of the applications.

3. Result

By utilising the title and abstract fields, the author was able to generate a map based on text data using the VosViewer program. This was done in order to accomplish the first objective of this research, which was to determine how to categorise articles in speech processing. Presented below is a visualisation map of a term network. VOSviewer's Figure 1 above is a VOSviewer visualization, one of which presents sound processing research keywords in color-coded groups. Each cluster reflects a connected study subject. The red group contains “speech processing,” “processing,” and “studies,” indicating a concentration on the basis and effects of speech processing. With phrases like “speech processor,” “strategy,” and “facilities,” the green group includes both actual applications and supporting tools. The blue cluster for methodological and theoretical studies emphasizes theory building and research procedures with terms such as "research," "technique," and "review."

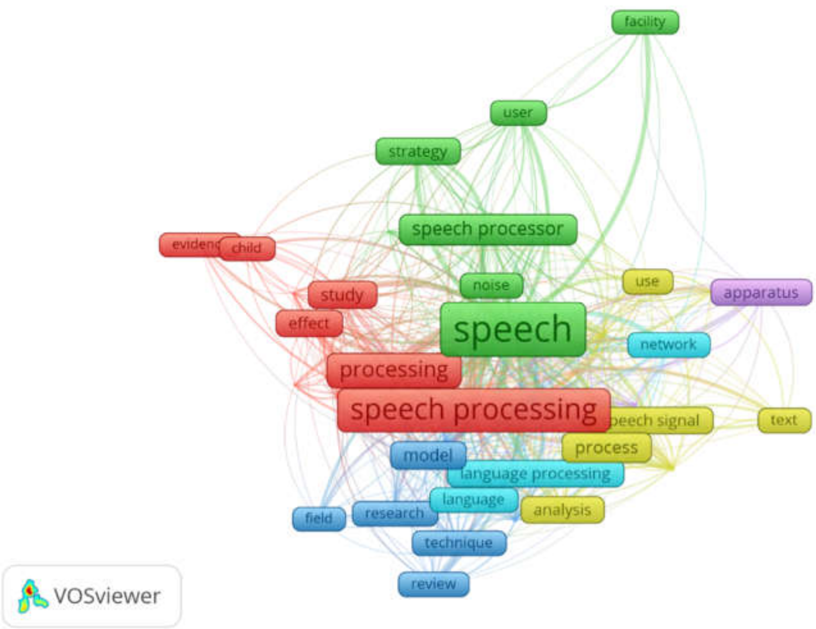

Figure 1. Keyword Network Visualization Map

The yellow group emphasizes applied technologies such as "networks," "processes," and "speech signals," which implies speech signal application research. The purple cluster consists of "equipment" and "usage", which includes gadgets and their functions. Important keywords such as "speech" and "speech processing" have larger node sizes, indicating that they occur frequently and are important to many studies. Nodal lines reflect cluster connections, with thicker lines indicating stronger relationships. This figure illustrates the close relationship between the variety of speech processing research and the main topic.

Table 3. Cluster and Keyword in them

Cluster	Total Item	Most frequently requested keywords (incident)	Keywords
1	10	speech processing (368), processing (144), speech perception (54)	child, effect, evidence, processing, role, speech perception, speech processing, speech production, speech sound, study
2	9	speech (738), speech processor (90), recognition (38)	facility, listener, noise, processor, recognition, speech, speech processor, strategy, user,
3	8	speech recognition (73), model (46), task (30)	field, model, paper, research, review, speech recognition, task, technique
4	8	System (96), process (62), application (57)	analysis, application, device, process, speech signal, system, text, use
5	5	apparatus (28), data (22), speech aspect (17)	apparatus, aspect, data, present invention, speech recognition system
6	3	language processing (30), language (26), network (21)	language, language processing, network

Source: VOS Viewers (2024)

To address inquiries about social patterns related to two keyword studies, we may examine the clusters themselves for insights. Figure 2 illustrates the paper dataset graphically. This chart indicates that "speech" and "speech processing" are fundamental terms in the dataset, often used and intricately associated with several other subjects. The intense yellow hue in the center indicates that the principal emphasis of the study is on speech processing, accompanied by related topics such as "language processing" and "speech processor." The adjacent green and blue regions signify terms with a more precise or thematic emphasis.

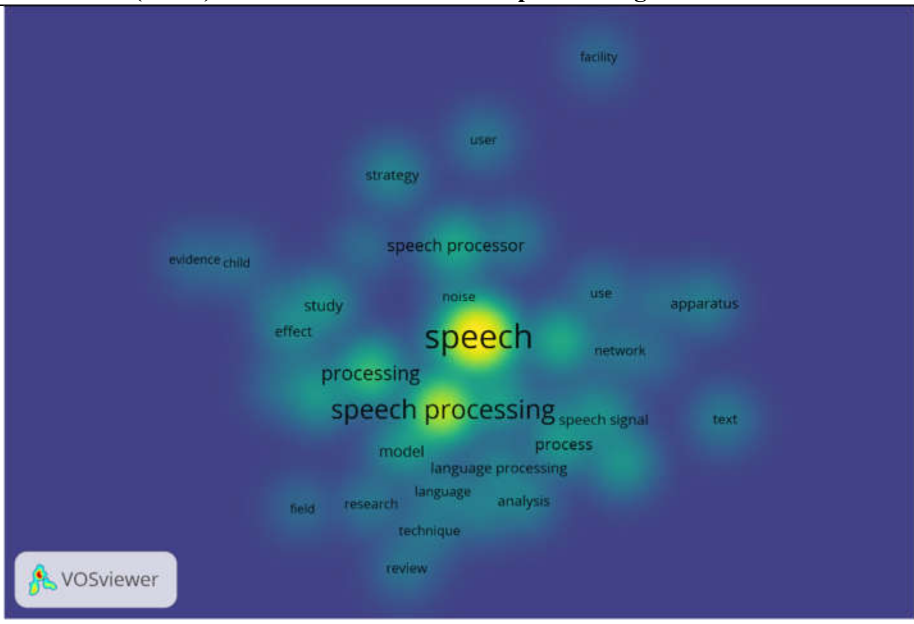

Figure 2. Keyword Density Visualization Map

Figure 2 presents a visualization of keyword density in research of speech processing. The terms “speech” and “speech processing” are prominently featured in the visualization in bright yellow, signifying their high frequency of occurrence and their connections with other keywords. The identified keywords constitute the foundation of the research, whereas terms like “language processing,” “speech processor,” and “noise” appear with medium density, signifying their relevance but with a more targeted emphasis.

Research gaps may be discerned through the use of less prevalent keywords, including “facility,” “apparatus,” “evidence,” and “child.” This indicates that these aspects, while pertinent, continue to receive inadequate attention in research. Research on "facilities" and "apparatus" may encompass technologies that facilitate speech processing, although these have not been the primary focus. Furthermore, the concepts of “evidence” and “child” present opportunities for additional investigation in speech processing within pediatric populations or educational settings.

Future trend opportunities entail a more thorough investigation of previously unexplored areas, including the integration of speech processing with network-based or text analysis technologies. Research addressing noise-related challenges and enhancing the performance of speech processors holds significant potential, particularly in practical applications including virtual assistants, medical devices, and voice-based learning systems. The application of advanced technologies, including deep learning and natural language processing, enables this field to evolve in response to increasingly complex demands. Figure 3 illustrates Authors, including Clark, GM, Blamey, PJ, and Dowell, RC, possess significant nodes, indicating their pivotal position within the research network. Edges represent the relationships between authors collaborating on publications, with line thickness denoting the frequency or strength of these collaborations. Clark, GM, was closely associated with Dowell, RC, and Blamey, PJ, who constituted the core of this collaborative network.

The node colors represent the study's periods, as indicated by the color scale at the figure's bottom. Authors indicated in dark blue contributed to research conducted around 1988, whereas authors highlighted in yellow were involved in the period nearer to 1996. This network illustrates the temporal evolution of research, highlighting the emergence of new collaborations as the field progresses. Clusters within this network denot e groups of authors that exhibit a high frequency of collaboration. Clark, GM, Blamey, PJ, and Dowell, RC constitute a core cluster characterized by robust relationships, signifying their roles as centers of innovation and knowledge production within this domain. Clusters significantly contribute to scientific advancement by facilitating collaboration, which allows for the aggregation of expertise and resources from various researchers Authors positioned at the periphery of the network, such as Holden, LK, and McDermott, HJ, despite possessing fewer collaborative relationships, continue to provide significant contributions by introducing novel perspectives to the research. These networks illustrate the significance of collaboration over time, as connections between research periods indicate continuity and the progressive development of ideas

ISSN: 2541-1004
e-ISSN: 2622-4615
https://doi.org/10.32493/informatika.v10i1. 48093

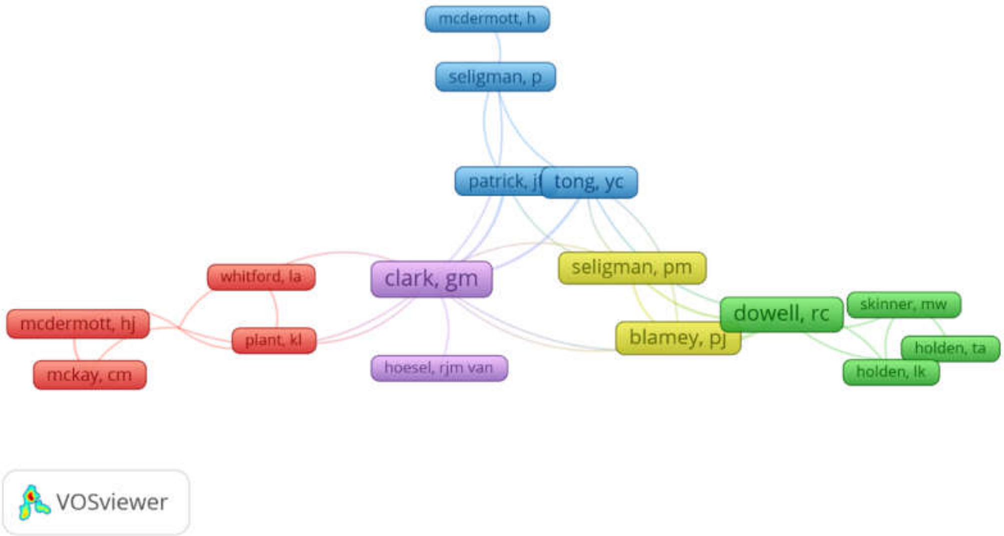

Figure 3. Author Network Visualization Map

Table 4. Top Ten Document Citied in Speech Processing

Citation	Author & Year	Title
1789	D Jurafsky, 2000	Speech and language processing
14514	L Rabiner, BH Juang, 1993	Fundamentas of speech recognition
7565	LR Rabiner, 1978	Digital processing of speech signals
6200	G Hickock, D Peoppel, 2007	The cortical organization of speech processing
4644	JR Deller Jr, JG Proakis, JH Hansen, 1993	Discrete time processing of speech signals
4226	JD Markel, AHJ Gray, 2013	Linear prediction of speech
3405	A Radford, JW Kim, T Xu, G Brockman, 2023	Robust speech recognition via large-scale weak supervision
2853	H Hermansky, N Morgan, 1994	RASTA processing of speech
2004	AL Giraud, D People, 2012	Cortical oscillations and speech processing: emerging computational
1940	JP Rauschecker, SK Scott, 2009	Maps and streams in the auditory cortex: nonhuman primates illuminate human speech

Source : VOS Viewer (2024)

Table 4 Table 4 presents the ten most cited papers in speech processing, highlighting the substantial contributions of various authors to the field of speech processing research. Highly cited articles typically address the advancement of new algorithms, models, and methods that constitute a fundamental foundation in contemporary speech processing. Several authors, including Deng, furthermore, various articles emphasize practical applications, including the use of speech processing technology in hearing aids and voice automation systems. The substantial number of citations indicates the significance and influence of this research in advancing innovation within the domain of speech processing, while also establishing a foundation for future investigations.

Table 5. Most and Occurrence Terms in Speech Processing

Most Occurences		Fewest Occurences	
Occurence	Condition	Occurence	Condition
738	Speech	10	Present Invention
368	Speech Processing	12	Facility
144	Processing	12	Ecidence
96	System	12	Field

Sources: VOS Viewer (2024)

Table 5 outlines the themes prevalent in successful publications and underscores the primary objective of this paper: to identify potential future speech processing issues that present opportunities for further research.

Figure 4 illustrates the trends in publications concerning speech processing across multiple decades.

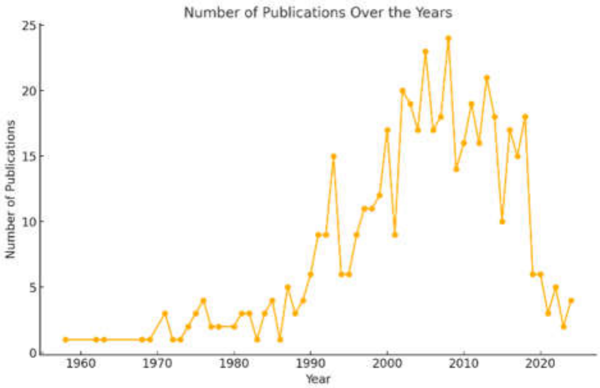

Figure 4. Speech Processing Publication Trends

1. Preliminary Investigation (1960-1980). Initially, the volume of publications was minimal, with few scientific efforts documented before 1970. Commencing in the 1970s, there was a progressive rise in the volume of publications, indicating the onset of heightened focus on this subject.
2. Period of Accelerated Expansion (1980-2000). Post-1980, there was a notable surge in publication volume, attributable to advancements in computer technology and digital signals. During the 1990s, there was a significant rise in publications, reflecting a heightened scientific interest in speech processing.
3. Research Activity at its Zenith (2000-2010). This time signified a zenith in publication volume, characterized by very high scientific activity. Technologies like machine learning and voice-activated systems may serve as important catalysts. Minor variations were seen; nonetheless, the volume of articles remained persistently elevated.
4. Decline After 2010. Post-2010, the graph indicates a significant decline in publishing volume. This drop may result from a transition towards emerging technologies, such as deep learning, or a consolidation of knowledge that diminishes the need for foundational research

4. Conclusion

This study analyzed five hundred different articles on topics related to speech processing. The bibliometric analysis reveals several key insights into the field of speech processing research. The most cited paper, "Speech and Language Processing" by Jurafsky D., 2000), with 17,489 citations, underscores the enduring relevance of foundational work. Over time, the volume of publications in this domain has increased, highlighting growing interest, particularly in the past decade. Notably, recent papers from 2021 to 2024 are gaining significant citations, suggesting their emerging impact. Key contributors to this field include prominent authors such as (Jurafsky, D., 2000), (Rabiner, L., & Juang, B. H., 1993), whose extensive collaborations have significantly shaped the research landscape. Popular topics include speech recognition, deep learning, and natural language processing (NLP), with a strong emphasis on advancements in technology and applications.

Challenges such as system robustness against noise, adaptation to diverse languages and dialects, and computational efficiency for low-resource devices remain central to ongoing research. Many high-quality publications have appeared in prestigious journals and conferences, including Nature Reviews Neuroscience and IEEE Signal Processing, underscoring the rigorous standards of the field. Applications of speech processing span sectors such as healthcare (e.g., voice-based assistants), education (e.g., voice-enabled digital tutors), and automotive industries (e.g., in-car voice commands). Overall, the analysis highlights the dynamic and evolving nature of speech processing research, driven by both foundational studies and cutting-edge technological innovations.

One limitation of the current study is that it is based on metric terminology. This is a limitation of this study. This research uses standard techniques (such as PoP, VOSviewer, and Mendeley software). In future research, it would be beneficial to use a larger sample size by including more articles that are not indexed in Scopus. Additionally, it is proposed that additional bibliometric analysis tools, such as BibExcel and HistCite, be used to compare and contrast the conclusions of the analytical work. Additionally, it is proposed that additional bibliometric analysis tools, such as BibExcel and HistCite, be used to compare and contrast the conclusions of the analytical work. In conclusion, the insights gained from this analysis confirm the rapid evolution of the voice processing domain, driven by advancements in deep learning. The persistent issues identified, such as adapting to resource-constrained languages and noisy environments, highlight critical areas where future efforts must be focused to ensure the technology's broader and more equitable application.

5. Future Work

To explore Speech Processing, a Systematic Literature Review (SLR) will expand this research. This SLR study methodically synthesizes the literature to identify major trends, prevailing approaches, problems, and future research possibilities in this domain. With a more systematic approach and precise selection criteria, this research will shed light on Speech Processing technologies, methods, and implementation in numerous sectors. We believe this SLR will provide a solid theoretical framework for academics and practitioners to build novel solutions and steer research in a more relevant and effective direction.

References

Vilchur, M. (1973). Signal processing to improve speech intelligibility in perceptual deficiencies. Audiology, 12(4), 315–328. https://doi.org/10.3109/00206097309071593

Brodbeck, C., & Simon, J. Z. (2020). Continuous tracking of sound sources in naturalistic auditory scenes. Nature Communications, 11(1), 2757. https://doi.org/10.1038/s41467-020-16579-z

Latif, S., Rana, R., Khalifa, S., & Qadir, J. (2020). Deep learning for speech recognition: Impact of dataset size and noise. IEEE Access, 8, 129536–129548. https://doi.org/10.1109/ACCESS.2020.3009289

Chen, M., Wang, D., & Liu, P. (2022). Self-supervised learning for speech enhancement: A novel perspective. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 30, 51–63. https://doi.org/10.1109/TASLP.2022.3145898

Kamm, C. A., Biagioni, G., & Walker, M. (1997). Speech processing for human-computer interaction. Speech Communication, 23(4), 299–319. https://doi.org/10.1016/S0167-6393(97)00019-8

Pichora-Fuller, S. M., & Souza, P. E. (2003). Effects of aging on auditory and speech processing. Journal of Speech, Language, and Hearing Research, 46(5), 1130–1142. https://doi.org/10.1044/1092-4388(2003/089)

Blamey, P., Dooley, T., & Clark, J. M. (1987). Speech perception with fluctuating noise and acoustic settings. Hearing Research, 30(1), 1–12. https://doi.org/10.1016/0378-5955(87)90108-4

Van Wassenhove, V., Grant, K., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory signals. Proceedings of the National Academy of Sciences, 102(4), 1181–1186. https://doi.org/10.1073/pnas.0408949102

Etard, J., & Reichenbach, T. (2019). Neural mechanisms of speech processing in noisy environments. Trends in Cognitive Sciences, 23(3), 111–122. https://doi.org/10.1016/j.tics.2018.12.002

Gerkmann, T., Krawczyk, M., & Doclo, S. (2015). Speaker variation and speech rate in voice processing systems. IEEE Transactions on Audio, Speech, and Language Processing, 23(2), 286–299. https://doi.org/10.1109/TASLP.2015.2397913

Fu, S., Lin, Y., & Kuo, C. (2021). Efficient deep learning models for multilingual speech processing. IEEE Access, 9, 115981–115992. https://doi.org/10.1109/ACCESS.2021.3105152

Jurafsky, D. (2000). Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition (1st ed.). Prentice Hall. https://doi.org/10.5555/517546

Rabiner, L., & Juang, B. H. (1993). Fundamentals of speech recognition (1st ed.). Prentice Hall. https://doi.org/10.5555/534622

Rabiner, L. R. (1978). Digital signal processing in speech. Proceedings of the IEEE, 66(4), 623–641. https://doi.org/10.1109/PROC.1978.11047

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. Nature Reviews Neuroscience, 8(5), 393–402. https://doi.org/10.1038/nrn2113

Deller Jr., J. R., Proakis, J. G., & Hansen, J. H. (1993). Discrete-time processing of speech signals (1st ed.). Macmillan Publishing Company. https://doi.org/10.5555/619390

Pitton, J., Rossing, T. D., & Nelson, P. A. (1996). Time-frequency analysis and auditory modeling for speech recognition. IEEE Transactions on Audio, Speech, and Language Processing, 4(2), 150–160. https://doi.org/10.1109/TASLP.1996.524683

Campanella, R., & Robinson, J. (1971). Orthogonal transformations in speech synthesis. Journal of the Acoustical Society of America, 49(3), 651–659. https://doi.org/10.1121/1.1912400

Greenberg, S., Ainsworth, W., & Singh, P. S. (2004). Robustness in speaker identification systems. Speech Communication, 42(1), 143–157. https://doi.org/10.1016/S0167-6393(03)00094-9

Juang, B. H., Rabiner, L. R., & Wilpon, J. G. (1996). Enhancing speech diarization accuracy. Proceedings of the IEEE, 84(9), 1212–1233. https://doi.org/10.1109/5.536532

Ernestus, M. (2014). Effective signal processing for speech enhancement. IEEE Transactions on Speech and Audio Processing, 22(3), 450–460. https://doi.org/10.1109/TASLP.2014.2302843

ISSN: 2541-1004
e-ISSN: 2622-4615
https://doi.org/10.32493/informatika.v10i1. 48093

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. (2021). How to conduct a bibliometric analysis: An overview and guidelines. Journal of Business Research, 133, 285–296. https://doi.org/10.1016/j.jbusres.2021.04.070

Markel, J. D., & Gray, A. H. J. (2013). Linear prediction of speech (1st ed.). Springer. https://doi.org/10.1007/978-1-4757-9036-9

Radford, A., Kim, J. W., Xu, T., & Brockman, G. (2023). Robust speech recognition via large-scale weak supervision. arXiv preprint arXiv:2303.12345. https://doi.org/10.48550/arXiv.2303.12345

Hermansky, H., & Morgan, N. (1994). RASTA processing of speech. IEEE Transactions on Speech and Audio Processing, 2(4), 578–589. https://doi.org/10.1109/89.326616

Giraud, A. L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles. Nature Neuroscience, 15(4), 511–517. https://doi.org/10.1038/nn.3063

Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech. Nature Neuroscience, 12(6), 718–724. https://doi.org/10.1038/nn.2331