The Progress of Speech Recognition in Health Care: Surgery as an Example

Bakheet ALDOSARI[a,b,1], Rana BABSAI[a,b], Ahmed ALANAZI[a,b], Hanan ALDOSARI[a,b], Abdullah ALANAZI[a,b]

[a] *King Saud bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia*
[b] *King Abdullah International Medical Research Center, Riyadh*

Abstract. Artificial Intelligence (AI) is a computer system that simulates intelligent human behavior. The use of AI is rapidly shifting Healthcare. Speech recognition (SR) is a type of AI physicians use to operate Electronic Health records (EHR). This paper aims to demonstrate the technological advancements made thus far concerning speech recognition in health care and explore multiple scholarly studies to generate a wide-ranging and detailed assessment of its current progress. The effectiveness of speech recognition is the heart of this analysis. This review investigates published papers on the progress and effectiveness of speech recognition in Healthcare. Eight research papers exploring the progress and effectiveness of speech recognition in Healthcare were thoroughly reviewed. Articles were identified from Google Scholar, PubMed, and the World Wide Web. The five relevant papers generally discussed the progress and current effectiveness of SR in Healthcare, implementing SR in the EHR, adapting healthcare workers to SR and the problems they face, developing an intelligent healthcare system based on SR and using SR systems in other languages. Conclusion: This report demonstrates the technological improvements realized concerning SR in Healthcare. It proved that SR could be a tremendous help to providers if every medical and health institution continued to progress in using this technology.

Keywords. Electronic Health Record (EHR), Artificial Intelligence (AI), Speech recognition (SR)

1. Introduction

Over the years, the healthcare sector relentlessly sought effective strategies to collect, sort, and store patient data. Artificial intelligence (AI) refers to computer systems that simulate and exhibit specific aspects of human intelligence or intelligent behavior, such as learning, reasoning, and problem-solving [1]. AI aims to mimic human cognitive functions. Its use is rapidly shifting healthcare and rapidly progressing analytics techniques [2,3].

Speech recognition (SR) is a type of AI physicians use to operate Electronic Health records (EHR). SR is used with documents, ward rounds, operations, and clinical results. [4] Despite speech recognition's ability to revolutionize healthcare, it is an emergent technology that remains to be studied and developed. [5] A speech recognition system is composed of a microphone that converts sound waves into electric signals. It also

---

[1] Corresponding Author, Prof. Bakheet ALDOSARI, King Saud Bin Abdul-Aziz University for Health Sciences, Saudi Arabia. Email: dosarib@gmail.com.

includes sound cards to digitalize the signals and a speech engine software that converts data into text. Using specially designed software, speech recognition systems use a device such as a telephone or a microphone to capture spoken words and convert the same into a digital format. To use the SR technology, healthcare workers speak into a microphone connected to a device where their words are translated into written text, appearing on their screen within seconds [4].

SR offers the potential to dramatically improve the cost and quality of healthcare services. Many developments have occurred in national and international standards of using SR, like e-prescription, clinical documentation, speech recognition in radiology, and pathological speech signal analysis, among others. SR technology has attained immense popularity in recent years and has been widely used for its many benefits, from individuals to large organizations. It helps reduce costs, increase productivity, and improve security. Other benefits include a noticeable decrease in document turnaround time, defined as the time taken for the entire process from creating the report to submitting it, and cost-effective reporting. SR documentation offers financial benefits related to increased staff savings per year. [6] Recently growing numbers of healthcare practitioners are increasingly using this medical technology.

Voice recognition is used in SR, which facilitates the use of the technology; however, faults remain. Spelling mistakes may occur, so users are encouraged to examine the accuracy of their dictation and correct any errors made while dictating on a routine basis. Another common problem is the omitted context recognition in speech; For example, "colon" may be mistaken for the organ, or conversely. These minor flaws need providers to revisit the recognized texts and make possible adjustments. These mistakes will decrease as AI improves voice recognition [4].

This review aims to investigate the technological advancements made by SR use in health care by exploring multiple scholarly studies which have researched this topic to generate a wide-ranging detailed assessment of the progress of Artificial intelligence and Speech recognition in healthcare.

2. Method

Papers related to the progress of speech recognition in healthcare were identified from Scholarly databases to find articles, research papers, conference proceedings, and reviews. Databases include PubMed, Embase, UpToDate, and Google scholar (up to 2022). The following keywords "Technology', "EHR," "Artificial Intelligence," "Speech recognition," and "Speech recognition AND Surgery" were searched as headings or text words found in the titles and abstracts. Furthermore, articles were revealed by using references from other pieces of literature. Any relative papers or published information were also included in the review.

3. Results

All the literature reviewed the progress and effectiveness of AI and SR in healthcare. Relevant papers were retrieved if they generally explained implementing speech recognition to Electronic Health Records, adapting healthcare workers to speech recognition, and developing an intelligent healthcare system based on speech recognition and recognition systems in other languages. Further, using SR in surgery is investigated.

According to the International Organization for Standardization (ISO), an EHR system is a data repository regarding the health and healthcare of a subject of care where all information is stored on electronic media. [7] A systemic review was done by Negro-Caldauch and colleagues. They provided an overview of the technological progress and optimization of EHR systems in healthcare that developed in the last decade. Their methods included a search strategy, study selection, data collection, and quality assessment. They retrieved 3,093 articles, but only seventy-two full-text articles were reviewed, and twenty-three systematic reviews were selected for their study. Their results showed that Speech recognition (SR) systems could improve workflow inefficiencies and assist documentation through dictation. They also included that SR technology reduced copy/paste behavior from 92.73% to 49.71%, leading to higher-quality reports. They concluded that enormous potential exists for EHR-integrated SR virtual assistants for data retrieval, command execution, and chart navigation [8].

On the other hand, to understand how revolutionary these new technologies are, they are evaluated from our present healthcare transformation perspective. Vase investigated healthcare practices as they adapted to using SR. Data for their study was collected during fieldwork and divided into two parts, each extending over three months of shadowing physicians, qualitative interviews, and focus group interviews. They also observed healthcare professionals in Southern Denmark and the Region of Zealand orthopedic departments [4].

To use SR, physicians need to have an account connected to the regional EHR system. Results showed that physicians had many opinions regarding using SR in their clinics. Some of the physicians' reported experiencing delays, and others complained of the slowness of the technology. These examples demonstrate the simplistic handling of complex technology and its contribution to workflow-related challenges. However, despite the slowness of the technology and lower recognition rate, physicians kept using SR when conducting EHRs. They concluded that workflows vary due to the complexity of the algorithm supporting SR. Forthcoming and more mature SR algorithms are made based on the current designs.[4] Understanding the practical solutions of speech recognition is imperative to provide physicians and patients with an easy control system. Ismail et al. aimed to build a low-cost system based on speech recognition. The system used pre-defined words with the device to speed up the recognition process. These words included: on, off, open, close, silent, degree/any number, lower, higher, access, above, down, channel number, and exit. A total of 10 male and ten female speakers were asked to record ten datasets to compare the performance and test the system. A total of 200 records were tested. The system achieved a high success rate of 97%. The researchers concluded that the system would be very effective in hospitals and smart homes to help patients become more independent [5].

3.1 Speech recognition in Surgery

Clinical speech recognition technology is a game-changer for documentation practice in the surgery specialty. It is touted to minimize the time spent typing or writing surgical records. Speech recognition technology reduces the administrative burden, supports the efficient achievement of professional goals of surgeons, ensures accurate patient data capture, and increases the uptake of digital transformation in facilities. According to Goss and others, speech recognition is a more effective solution for improving the efficiency and quality of documentation compared to typing, dictations, templates, or scribes. Thus, clinical speech recognition technology is expected to play a significant

role in clinical documentation in all medical specialties. [9] The numerous features and benefits of speech recognition technology make it appropriate for documentation in the surgery specialty. CPS Healthcare explains that in one minute, most physicians can speak 120 words but type less than 40 words. Hence, speech recognition is expected to be more than three times faster than typing or dictation. The technology also provides a robust medical vocabulary that can be used to customize personal medical records. Besides, features such as visual indicators, voice navigation commands, end-to-end security protocols, and shared speech profiles increase the efficiency of documentation by surgeons. Furthermore, strict adherence to the documentation practice ensures that patients have accurate information to make informed healthcare decisions. Strategic patient care planning also relies upon complete and accurate documentation of patient procedures. Similarly, documentation is the basis for financial planning and budgeting by the government and health facilities. Furthermore, reimbursements are based on patient medical records and affect the gross revenues of the facility.

3.2 Speech recognition in other languages

Does changing the language affect system efficiency? A study by Sayem et al. focused on the effect of spoken language on speech recognition. In their paper, they built an SR system for the Bangla alphabet. The job of the recognition system was to derive the information from the waveform needed to make the correct decision. In addition, they took the vocabulary of 10 consonants and test samples from 2 different speakers to observe the performance. They concluded that two key factors facilitated the implementation of machine translation: the degree of experience gained through previous exposure to other innovations and the degree to which individuals who interact are similar in terms of beliefs, educational attainment, social status, and other similar characteristics, among other things. Their results showed that their SR system's accuracy level was almost 90%. Ensuring the recording's perfection (noise-free) made it possible to increase recognition accuracy [10].

4. Discussion

Physicians and other healthcare practitioners have continued to deliver treatment and record outcomes similarly. Despite the existence of databases and servers capable of storing and transmitting health information, there is yet to be a widely acknowledged method of recording and obtaining the data into and from these systems [11].

Speech recognition has been available for quite some time as a technology, but it has only lately been shown to be worthwhile in terms of the effort required for widespread usage. Newer systems are more user-friendly, and technological advancements have made it simpler for consumers to operate the technology. They are more functional and may be utilized at all hospitals and nursing homes for patients and the elderly. The technology employs just specified phrases to manage EHRs in the hospital, with only accurate voice recognition.

Future verbal technologies may help providers who must write paperwork at the inpatient's bedside, such as those working in hospitals. The continuing usage of speech-recognition technology will be required to demonstrate its value.

5. Conclusion

Considering how the advancement of computers and information systems will be in the future, every medical and health institution must continue to progress in using this technology. As a result, a new field of use for such devices has opened. The review constructed a theoretical foundation and assessed the deployment and progress of verbal technology and its influence on performance.

References

[1]   McCarthy J. (2019). Artificial intelligence: past, present, and future. The Dartmouth summer research project on artificial intelligence.

[2]   Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, Wang Y, Dong Q, Shen H, Wang Y. Artificial intelligence in healthcare: past, present and future. Stroke and vascular neurology. 2017 Dec 1;2(4).

[3]   Rekha KP. Artificial Intelligence In Healthcare. KYAMC Journal. 2021 Feb 11;11(4):164-5.

[4]   Vase S. The maturing of automatic speech recognition in healthcare practices. Proceedings http://ceur-ws. org ISSN. 2021;1613:0073.

[5]   Ismail A, Abdlerazek S, El-Henawy IM. Development of a smart healthcare system based on speech recognition using a support vector machine and dynamic time warping. Sustainability. 2020 Mar 19;12(6):2403.

[6]   Hodgson T, Coiera E. Risks and benefits of speech recognition for clinical documentation: a systematic review. Journal of the American medical informatics association. 2016 Apr 1;23(e1):e169-79.

[7]   Boonstra A, Versluis A, Vos JF. Implementing electronic health records in hospitals: a systematic literature review. BMC health services research. 2014 Dec;14(1):1-24.

[8]   Negro-Calduch E, Azzopardi-Muscat N, Krishnamurthy RS, Novillo-Ortiz D. Technological progress in electronic health record system optimization: Systematic review of systematic literature reviews. International journal of medical informatics. 2021 Aug 1;152:104507.

[9]   Goss FR, Blackley SV, Ortega CA, Kowalski LT, Landman AB, Lin CT, Meteer M, Bakes S, Gradwohl SC, Bates DW, Zhou L. A clinician survey of using speech recognition for clinical documentation in the electronic health record. International journal of medical informatics. 2019 Oct 1;130:103938.

[10] Sayem A. Speech analysis for alphabets in Bangla language: automatic speech recognition.2014 3(2), 88-93. DOI: 10.17950/ijer/v3s2/211.

[11] Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, Cui C, Corrado G, Thrun S, Dean J. A guide to deep learning in healthcare. Nature medicine. 2019 Jan;25(1):24-9.