# Artificial intelligence and dysphagia: Novel solutions to old problems

Ervin Sejdić[1,2,3,4]
Yassin Khalifa[1]
Amanda S. Mahoney[5]
James L. Coyle[5,6]

[1]Department of Electrical and Computer Engineering, Swanson School of Engineering
[2]Department of Bioengineering, Swanson School of Engineering
[3]Department of Biomedical Informatics, School of Medicine
[4]Intelligent Systems Program, School of Computing and Information
[5]Department of Communication Science and Disorders, School of Health and Rehabilitation Sciences
[6]Department of Otolaryngology, School of Medicine
University of Pittsburgh, Pittsburgh, PA, USA

## Abstract

Dysphagia management, from screening procedures to diagnostic methods and therapeutic approaches, is about to change dramatically. This change is prompted not solely by great discoveries in medicine or physiology, but by advances in electronics and data science and close collaboration and cross-pollination between these two disciplines. In this editorial, we will provide a brief overview of the role of artificial intelligence in dysphagia management.

## Introduction

In the last five years artificial intelligence (AI) has emerged as a promising, though controversial, construct in many clinical communities. Published papers propose various AI solutions for clinical questions. Editorials in leading medical journals propose AI as a magical solution to medical problems, while others declare it is a sham. Similarly, many workshops and tutorials at clinical and engineering/data science conferences discuss either the value, or the danger, of AI in healthcare. Though conflicting opinions exist, we live in the midst of a new industrial revolution, this time driven by AI, and it's role is becoming more and more clear within healthcare.

AI will play a major role in our lives going forward; however, AI is not a new concept. We have been mesmerized by AI for years. In the 1980s, the beloved *Star Trek: The Next Generation* character, Data, was an android with AI. AI in movies continued with Will Smith's movie *I, Robot*. Then, in 1997, IBM produced a chess-playing computer named Deep Blue. This AI system played several chess matches against the world champion, Gary Kasparov, eventually beating him in the series. Around the same time, industrial applications of AI and robotics emerged, especially for use in assembly lines, thus completely revolutionizing manufacturing in several sectors. The advances in industry prompted rapid economic development of many countries, and steadily decreased product costs. Though AI in the past may have

been seen as science fiction, it's early use prompted the development of AI tools we use today. Many of our homes are vacuumed by an electronic robot vacuum. Our "smart" appliances perform functions and possess features that we could not have imagined 20 years ago and the apps on our TVs have transformed our viewing experiences by enhancing streaming quality through bandwidth usage prediction and personalization of content recommendations based on watching history. A cell phone in early 2000s could make a call and send a short text message, but cell phones today can literally make interactive phone calls and book appointments on our behalf. These handheld devices pack more AI-based solutions in your hand than what was available to a few select governments in the world 20-30 years ago.

## A brief history of AI

AI was officially born in its current form at a workshop at Dartmouth College in 1956, during which the term "artificial intelligence" was coined by John McCarthy, an American computer scientist [1]; however, its beginnings go back to scientists such as Descartes, Leibniz, Pascal and many before them. Contemporary scientists such as John von Neumann, Claude Shannon, Isaac Asimov, Alan Turing and Herbert Simon also played a major role in the birth of this field.

Although the next four decades (1960 – 1999) presented theoretical contributions to the field, most did not result in solutions that affect our daily lives. Even the American Department of Defense, AI's primary funder in the early days, had a lukewarm relationship with the field. The Department fluctuated from phases of excitement and funding to stopping or having negligible funds for the field.

The development of computer technologies in 2000s, primarily driven by the development in semiconductors, enabled us to pack more transistors on ever-smaller areas and thus more computing power in smaller devices. These advancements prompted the current renaissance of AI. Increasing available computational resources enabled researchers to develop AI solutions, many of which we use today (e.g., your smartphone is more computationally powerful than computers used by NASA in 1969 to send astronauts on the moon).

## What is AI?

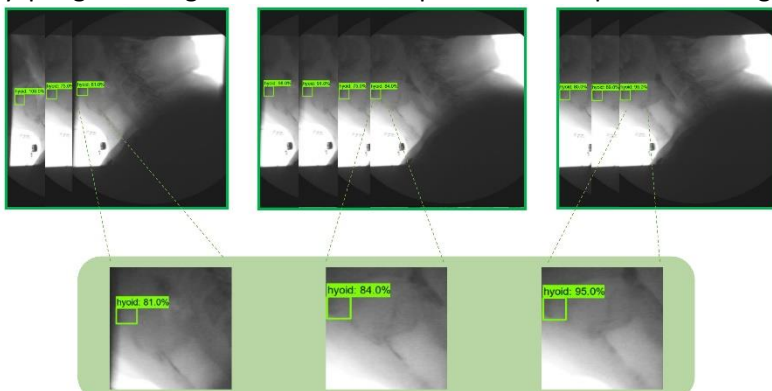AI is the study of intelligent agents, or autonomous entities that learn from their environment, which
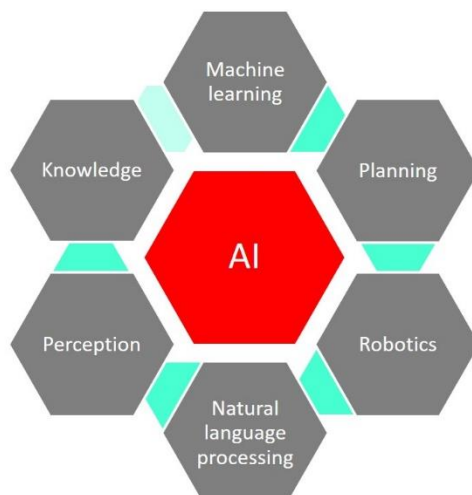


**Figure 1** – *AI compromises of several major areas, which are interconnected.*

attempt to maximize their chances of successfully achieving their goals [2]. In other words, a machine, or computer, takes information from the external world, learns from that information, and then uses that information to achieve a goal or complete a task. AI encompasses various technological aspects as depicted in Figure 1. **Planning** is an important aspect of AI solutions. Just like humans construct plans to achieve a goal, intelligent systems must also perform this task. **Robotics** is a multidisciplinary study involving engineering, computer science, and mathematics; its primary goal is to design, construct, and use robots that should do no harm, but rather, assist humans in their daily lives. Some robots can resemble humans and perform humans movements like walking, reaching, lifting, or even talking. N**atural language processing** is the interaction

between computers and human language; the primary goal is for computers to understand, interpret, and manipulate human speech and language. For example, autocomplete helps you finish the rest of the word you are typing, and spelling and grammar check keeps your documents (mostly) error-free. **Perception** is the ability to use sensors to automatically deduce various aspects of the world. Sensors range from microphones and cameras to radar systems. Examples of applications using machine perception are facial recognition, which is used to open the home screen on some smart phones, and speech recognition, which allows you to ask your phone to call someone in your contacts.   **Knowledge** is the ability of intelligent systems to gather knowledge possessed by experts in a field. Although these systems do not yet possess "general knowledge", or a wide breath of knowledge as possessed by humans, their knowledge is more narrowly focused. For example, a system might possess knowledge *only* necessary for self-driving cars. Lastly, and perhaps familiar to many, is machine learning**. Machine learning** represents a study of algorithms that can improve through experience in the form of training data. Simplified, a large body of data is given to the computer/machine to use for solving a problem. The machine will use a human provided algorithm to train itself with the provided data. The machine then works autonomously and will become more accurate and effective over time.

## AI for dysphagia

AI offers numerous possibilities for dysphagia management: from detecting dysphagia, even in patients without symptoms, to facilitating delivery of advanced rehabilitation treatments to patients worldwide who may  not have had access to treatment otherwise [3]. AI has already had a major influence on the field of dysphagia. In a position paper published in 2019, we proposed a novel field called Computational Deglutition aimed at developing translational algorithms that can address clinical shortcomings in dysphagia management.  AI developed from 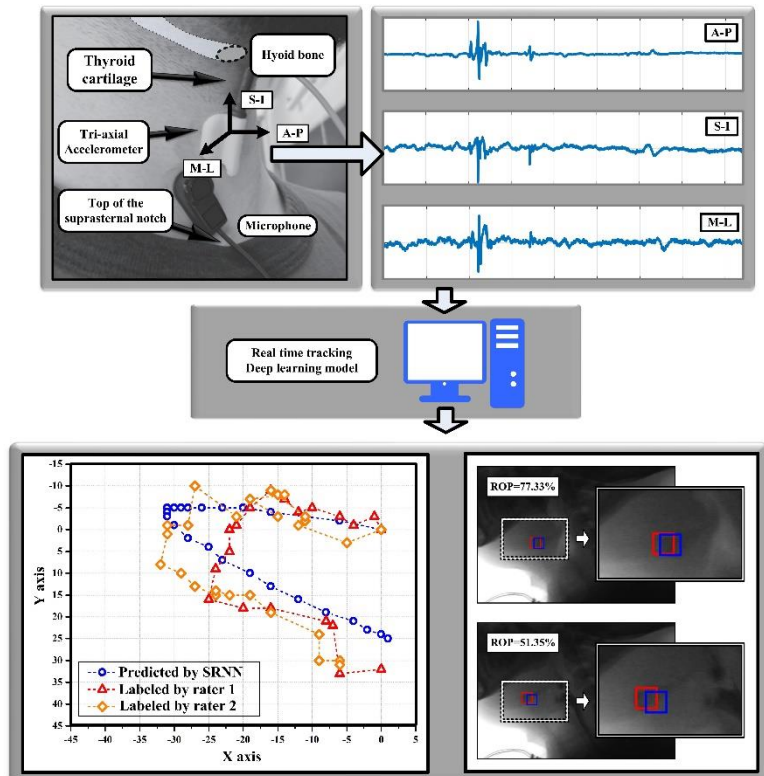computational deglutition has already had an impact on various aspects of dysphagia management. First, we aimed to track hyoid bone displacement during swallowing in real-time in videofluoroscopic images without human input. Hyoid bone tracking  is a daunting task considering  a videofluoroscopic exam consists of upwards of 20 swallows and each swallow is approximately one to three seconds long. X-ray machines typically acquire 30 frames per



**Figure 2** – *An AI-based solution developed by our research team to automatically track hyoid bone displacement in x-ray frames without input from humans.*

second so even for a short exam, annotation of hyoid bone placement would be required for at least 200 to 300 frames. This task takes a trained human judge between 30 and 60 minutes depending on the number of frames and clinician's expertise. In our recent paper, we proposed a machine learning method that can automatically and accurately detect the location of a hyoid bone on a frame-by-frame basis (Figure 2). On average, the algorithm is over 89% accurate [4]. Most importantly, it only takes about 20 to 30 seconds to analyze hyoid displacement in a swallow! In other words, we can now analyze many aspects of swallow physiology with comparable accuracy as trained human judges, from all swallows acquired during a videofluoroscopy exam within a matter of minutes. Hyoid tracking is just one of the algorithms we are developing to automatically analyze videofluoroscopy images. Within the next few years, we hope

to have a collection of algorithms that can extract information from physiological swallowing events captured during videofluoroscopy, including characterization of airway protection, and opening of the entrance to the esophagus. The information from the exam will be summarized and reported for the clinician by the AI system, extending the diagnostic availability of dysphagia clinicians to many people with limited access to traditional testing.

Videofluoroscopy is not widely available across the globe, and repeated exposure to x-ray could result in harmful ionizing radiation effects. Although the videofluoroscop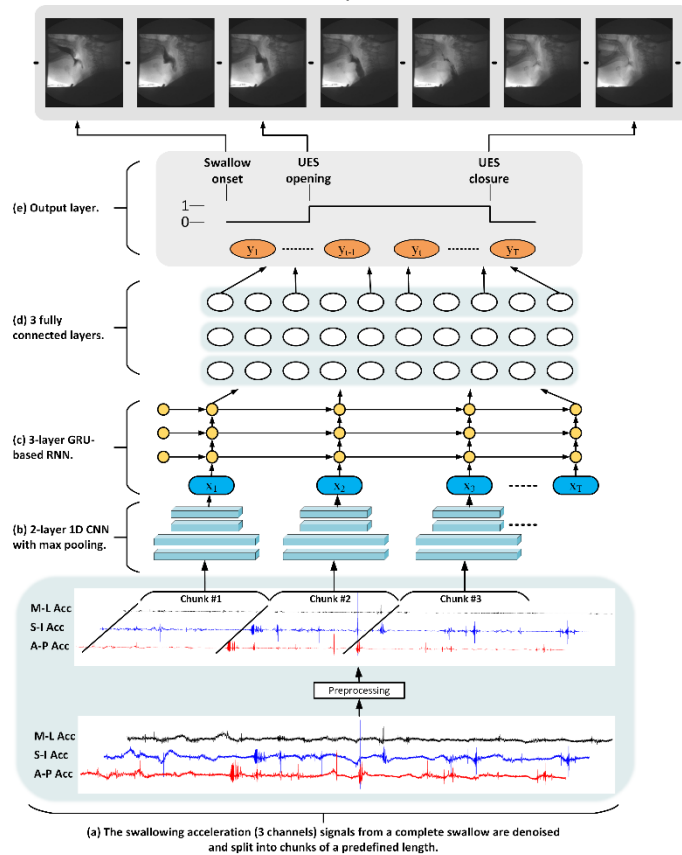ic exam is still the gold standard diagnostic approach for dysphagia, recent developments in sensor technology and AI provide us with a way to rethink this approach. In our recent research investigation, we mounted an inexpensive sensor (accelerometer) on the neck of patients to acquire neck vibrations during swallowing. We concurrently acquired vibration signals and videofluoroscopy images during the swallowing exam. We used the ground truth from the trained-human ratings of the videofluoroscopic images to develop a machine learning model that can infer about hyoid bone displacement from neck vibrations during swallowing (Figure 3) [5], [6]. Trained raters marked the location of the hyoid bone on every frame of a swallow and this information was given to the computer model for training. We first trained the model using x-ray images and neck vibrations, but once the model was accurately trained, we checked the model's accuracy using only neck vibrations. It was astonishing to witness the machine learning model accurately infer about the hyoid bone displacement using only neck vibrations (no x-rays needed!) [5].



**Figure 3** – *Tracking hyoid bone displacement during swallowing without x-rays. It is possible using a simple sensor and advanced machine learning algorithms.*

This exciting finding opened many interesting questions, which we are pursuing in our lab. Our most recent contribution used non-invasive neck vibrations (through an accelerometer) to approximate upper esophageal sphincter (UES) opening and closure obtained from videofluoroscopy images (Figure 4). We call these recordings high-resolution cervical auscultation (HRCA) recordings. To determine the UES opening and closure through HRCA, we trained a deep neural network [7], [8]. Neural networks are meant to simulate activity in the human brain. The network is repeatedly trained on datasets of HRCA signals with UES opening and closing marked by human raters on the simultaneously collected videofluoroscopy images. After training, the network recognizes patterns and can determine the UES opening and closing by HRCA signals alone. Our results demonstrated that we can detect these two events non-invasively (i.e. with HRCA) with over 90% accuracy [7].

Our research moved from simple observations of "safe" versus "unsafe" swallows via neck vibrations to demonstrating that through AI, we could potentially offer non-invasive non-ionizing radiation approaches for dysphagia management.



Figure 4 – *Using advanced machine learning algorithms, upper esophageal sphincter openings and closures can be predicted from neck vibrations without using x-rays.*

Will AI replace speech-language pathologists and other clinicians working with patients with dysphagia? NO. AI will improve services provided by clinicians by extending their expert reach to a broader range and larger number of patients, but will never replace them. Ideally, AI-based technologies will automate some clinical services and improve others, by performing many necessary measurements and providing preliminary interpretations that enable the clinician to focus on advanced differential diagnosis and treatment methods. This argument is analogous to the projected reduced need for bank tellers with the introduction of the ATM; conversely, the automated money disbursement machine enabled bank tellers to offer more services to their customers. Similarly, clinicians will always play a major role in dysphagia management but with AI technology, they could provide enhanced services for patients. For example, imagine we have a handheld non-invasive device that can infer about physiological swallowing events as accurately as videofluoroscopy. In such a case, the diagnosing clinician could devote more time to each patient. Without the complete reliance on videofluoroscopy, clinicians could test the efficacy of rehabilitation treatments without the typical time, environmental, and financial constraints of traditional testing. Treating clinicians would have the benefit of a non-invasive biofeedback device and could provide patients with real-time feedback about their swallowing that is currently not possible without imaging. It is clear that such a device powered by AI could tremendously increase the productivity and service level of speech-language pathologists and other clinicians treating patients with dysphagia.

## Conclusions

Though not a new branch of science, AI has been become a focus for many researchers in recent years and pervades our daily lives. No longer seen, to most, as science fiction, AI-based technologies, when properly developed and implemented, will become valuable tools for many clinicians. The goal of AI in dysphagia management is to offer a non-invasive tool to help support clinical decision making and provide patients with a safer and simpler path to recovery.

# Acknowledgement

# References

[1] J. McCarthy, M. L. Minsky, N. Rochester, and C. E. Shannon, "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence," Dec. 2006. doi: 10.1609/AIMAG.V27I4.1904.

[2] J. Andreu-Perez, F. Deligianni, D. Ravi, and G.-Z. Yang, "Artificial Intelligence and Robotics," *Artif. Intell.*, vol. 26, no. 1, pp. 79–121, Mar. 2018, Accessed: Sep. 24, 2020. [Online]. Available: http://arxiv.org/abs/1803.10813.

[3] J. L. Coyle and E. Sejdić, "High-Resolution Cervical Auscultation and Data Science: New Tools to Address an Old Problem," *Am. J. Speech-Language Pathol.*, vol. 29, no. 2S, pp. 992–1000, Jul. 2020, doi: 10.1044/2020_AJSLP-19-00155.

[4] Z. Zhang, J. L. Coyle, and E. Sejdić, "Automatic hyoid bone detection in fluoroscopic images using deep learning," *Sci. Rep.*, vol. 8, no. 1, pp. 1–9, Dec. 2018, doi: 10.1038/s41598-018-30182-6.

[5] S. Mao, Z. Zhang, Y. Khalifa, C. Donohue, J. L. Coyle, and E. Sejdic, "Neck sensor-supported hyoid bone movement tracking during swallowing," *R. Soc. Open Sci.*, vol. 6, no. 7, p. 181982, Jul. 2019, doi: 10.1098/rsos.181982.

[6] C. Donohue, S. Mao, E. Sejdić, and J. L. Coyle, "Tracking Hyoid Bone Displacement During Swallowing Without Videofluoroscopy Using Machine Learning of Vibratory Signals," *Dysphagia*, vol. 1, p. 3, May 2020, doi: 10.1007/s00455-020-10124-z.

[7] Y. Khalifa, C. Donohue, J. L. Coyle, and E. Sejdic, "Upper Esophageal Sphincter Opening Segmentation with Convolutional Recurrent Neural Networks in High Resolution Cervical Auscultation," *IEEE J. Biomed. Heal. Informatics*, pp. 1–1, Jun. 2020, doi: 10.1109/jbhi.2020.3000057.

[8] C. Donohue, Y. Khalifa, S. Perera, E. Sejdić, and J. L. Coyle, "How Closely do Machine Ratings of Duration of UES Opening During Videofluoroscopy Approximate Clinician Ratings Using Temporal Kinematic Analyses and the MBSImP?," *Dysphagia*, vol. 1, p. 3, Sep. 2020, doi: 10.1007/s00455-020-10191-2.