Title Page with ALL Author Contact Information

Tracking hyoid bone displacement during swallowing without videofluoroscopy using machine learning of vibratory signals

Cara Donohue, MA CCC-SLP[1], Shitong Mao, MS[2], Ervin Sejdić, PhD[2], James L. Coyle, PhD[1]

1Department of Communication Science and Disorders, School of Health and Rehabilitation Sciences, University of Pittsburgh, Pittsburgh, PA 15260 USA
2Department of Electrical and Computer Engineering, Swanson School of Engineering, Department of Bioengineering, Swanson School of Engineering, Department of Biomedical Informatics, School of Medicine Intelligent Systems Program, School of Computing and Information, University of Pittsburgh, Pittsburgh, PA 15260 USA

Reprint address:

Dr. James L. Coyle
Department of Communication Science and Disorders
School of Health and Rehabilitation Sciences
6035 Forbes Tower
University of Pittsburgh
Pittsburgh, PA 15260

E-mail: jcoyle@pitt.edu
Phone number: (412)-383-6608

Acknowledgements:
Funding:
Research reported in this publication was supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number R01HD092239, while the data was collected under Award Number R01HD074819. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or National Science Foundation.

People: Thanks are due to Amanda Mahoney, MA SLP, Aliaa Sabry, MD/PhD, Atsuko Kurosu, PhD, and Zhenwei Zhang, MS, for assistance with data collection and coding.

Abstract:

Identifying physiological impairments of swallowing is essential for determining accurate diagnosis and appropriate treatment for patients with dysphagia. The hyoid bone is an anatomical landmark commonly monitored during analysis of videofluoroscopic swallow studies (VFSSs). Its displacement is predictive of penetration/aspiration and is associated with other swallow kinematic events. However, VFSSs are not always readily available/feasible and expose patients to radiation. High resolution cervical auscultation (HRCA), which uses acoustic and vibratory signals from a microphone and triaxial accelerometer, is under investigation as a noninvasive dysphagia screening method and potential adjunct to VFSS when it's unavailable or not feasible. We investigated the ability of HRCA to independently track hyoid bone displacement during swallowing with similar accuracy to VFSS, by analyzing vibratory signals from a triaxial accelerometer using machine learning techniques. We hypothesized HRCA would track hyoid bone displacement with a high degree of accuracy compared to humans. Trained judges completed frame-by-frame analysis of hyoid bone displacement on 400 swallows from 114 patients and 48 swallows from 16 age-matched healthy adults. Extracted features from vibratory signals were used to train the predictive algorithm to generate a bounding box surrounding the hyoid body on each frame. A metric of relative overlapped percentage (ROP) compared human and machine ratings. The mean ROP for all swallows analyzed was 50.75%, indicating >50% of the bounding box containing the hyoid bone was accurately predicted in every frame. This provides evidence of the feasibility of accurate, automated hyoid bone displacement tracking using HRCA signals without use of VFSS images.

Key words: dysphagia, hyoid bone, videofluoroscopy, machine learning, cervical auscultation, swallow screening, deglutition, deglutition disorders

Introduction

The hyoid bone is an important physiological marker that is used in assessment of swallow function during videofluoroscopic swallow studies (VFSSs) [1]. Though it is true that the hyoid itself performs no important kinematic actions itself, it is a cardinal osseous structure that can be accurately tracked in frame by frame image analysis. In a recent systematic review, Molfenter & Steele cited no fewer than thirteen studies spanning from 1988 to 2010 in which hyoid displacements served as a dependent variable in investigations of swallow kinematics [2]. Hyoid displacement is produced by the summation of contractions of suprahyoid muscles originating on the mandible, tongue, and skull base. During complete hyolaryngeal excursion, the epiglottis is reoriented to a horizontal position while the larynx is displaced from the bolus pathway [3]. The resultant traction forces are delivered to the anterior wall of the upper esophageal sphincter, which aids in its distension to enable esophageal flow when adequate relaxation of UES resting tone is neurally attenuated at the onset of the pharyngeal response [4-6]. As a result, impaired hyolaryngeal displacement is associated with reduced airway protection and has been shown to be predictive of laryngeal penetration and tracheal aspiration during swallowing [7-9], as well as UES opening [9]. During treatment, efforts to evaluate the efficacy of compensatory or restorative treatments designed to increase hyolaryngeal excursion are dependent on measurement of hyoid displacement using VFSS frame by frame analysis. However, because VFSSs are an x-ray procedure, its feasibility in tracking swallow kinematic events during ongoing management is poor. Reasonably accurate alternatives to VFSSs that provide direct feedback regarding hyoid displacement would extend judgments of treatment efficacy beyond pre- and post-treatment VFSSs. Therefore, the ability to objectively and continuously quantify hyoid displacement over the course of treatment would provide great clinical utility given that improving hyoid bone displacement with behavioral compensatory or restorative therapy is a frequent dysphagia treatment target [10-11].

While hyoid bone displacement is ~~an~~ one important physiological event to measure, there are limitations to the tools currently available to assess hyoid bone movement in VFSS images. The Modified Barium Swallow Impairment Profile (MBSImP) is a standardized clinical rating tool that is used to assess 17 different physiological components of swallowing in the oral, pharyngeal, and esophageal phase [12]. This rating scale has been advantageous in establishing a standardized approach to conducting and clinically analyzing impairment severity from VFSSs using several ordinal, categorical rating scales. While this rating scale is a relatively efficient tool for rating VFSS images in the clinical setting, completion of the initial training is time-consuming (20-25 hours on average per the training website). In addition to this, the rating scale requires clinicians to use an element of subjective judgment to select from three categories of anterior hyoid displacement: absent, partial, or complete. Swallow kinematic analysis using frame by frame tracking of hyoid bone movement is an objective way to measure actual anterior and superior hyoid bone movement during swallowing. While this method has a higher degree of precision and is quantitative, it is typically performed solely in research studies because it requires training to ensure accurate and reliable measurements. Few clinicians are trained in making these measurements and moreover, few clinicians perform frame-by-frame kinematic analyses which is time-consuming and requires specialized image processing software to perform [13]. To examine how time-consuming hyoid frame-by-frame tracking is, we timed ourselves while we completed frame-by-frame tracking of the hyoid body on one swallow (46 frames). This took over 5 minutes to complete for one swallow. For a VFSS that contains ~15 swallows, this would mean that hyoid tracking alone would take 75 minutes, which is impractical within the clinical setting. While efforts have been made to improve the accessibility of swallow kinematic analysis to clinicians, there are no readily available courses to train clinicians and test their reliability in order to ensure accurate measurements. Because of these constraints, quantitative hyoid displacement measurements (if performed at all) are performed after the VFSS examination. This disables the clinician’s ability to utilize this information contemporaneously during the VFSS to identify impairments, test the efficacy of logical behavioral compensations, or assess the efficacy of other interventions to improve hyoid displacement.

While VFSSs remain one of the gold standard dysphagia assessment tools for identifying physiological swallowing impairments, they are expensive and expose patients to radiation. In addition to this, they are not always readily available or feasible in some clinical settings, in underserved regions of the world, or with some patients who prefer not to undergo VFSS. Clinical examination components are likewise limited in providing quantitative physiologic data needed to guide diagnosis and intervention. Likewise, dysphagia screening tools have gained popularity as a basis for intervention despite their limited scope, lack of any objective measures of swallow physiology, and poor specificity [14-16]. Therefore, there is growing enthusiasm for development of non-invasive and accurate dysphagia screening and diagnostic techniques that provide enhanced information about underlying swallowing physiology without imaging.

It should be underscored that the body of the hyoid bone is an exceptionally small anatomic feature within the overall VFSS image during swallow studies. Its actual height in adults (the sagittal cross section of which is viewed in the lateral plane during VFSS and tracked during image processing), ranges from approximately 6 mm to

14-15 mm depending on age and gender, based on physical measurements of cadaver hyoid bone dimensions and 3D geometric morphometric analyses [17-18]. Though its radiographic image is a relatively robust anatomic feature when visually tracked by human judges on each frame, clinical methods of estimating hyoid displacement are inherently inaccurate due to its small size and speed of movement. Additionally, noninvasive methods of hyoid tracking have yet to be developed, tested, and validated. The ability of a noninvasive hyoid tracking method that can follow a substantial proportion of this tiny structure during each swallow, would represent a useful innovation in dysphagia research and clinical work [19]. For example, it could be used as a dysphagia screening tool to provide insight into whether patients have reduced hyoid displacement that could increase their risk of penetration and/or aspiration if they are ~~is contributing to impaired swallow function in patients who are~~ unable to participate in a VFSS.~~s~~ I~~or i~~t could also be used to provide consistent biofeedback to patients ~~-~~with dysphagia who have had a VFSS that confirmed impaired hyolaryngeal excursion as they perform exercises or compensations to improve hyolaryngeal excursion.

High resolution cervical auscultation (HRCA) is a swallowing screening method that is currently under investigation as an adjunct to VFSS when unavailable or not feasible and as a potential biofeedback method during therapy. HRCA combines information (i.e. signal features) extracted from acoustic and vibratory signals obtained from a contact microphone and a tri-axial accelerometer affixed to the anterior neck overlying the cricoid cartilage during swallowing. In our ongoing research designed to investigate HRCA, we have performed more than 350 time-linked, concurrent videofluoroscopy and HRCA recordings on patients with dysphagia and healthy participants. Following data collection, we use standardized kinematic analysis of swallowing physiology as the input for machine-learning techniques or statistical models with HRCA signals to elucidate swallow physiology noninvasively [20-26]. Our previously published research findings have demonstrated that HRCA signals are strongly associated with a variety of VFSS kinematic measurements [22-23, 27-28], as well as the feasibility of machine learning in characterizing hyoid bone displacement [28-31]. The current investigation is building off of prior work in our lab, which established that features from HRCA signals are associated with hyoid bone displacement. In the present study we expanded upon this prior work by investigating the ability of HRCA to independently approximate frame-by-frame human measurements of hyoid bone movement and clinical (MBSImP) ratings of hyoid bone movement by using vibratory signals from a neck sensor and machine learning techniques using concurrently recorded VFSS images analyzed by trained raters. We hypothesized that 1. HRCA with machine learning would track frame-by-frame movement of the body of the hyoid bone with a high degree of agreement with human kinematic measurements and 2. HRCA signals combined with statistical methods would effectively dichotomize hyoid bone movement as “normal” or “reduced” based on MBSImP ratings, using vibratory signal features from neck sensors during swallowing. It should be emphasized that the purpose of this methodological study was not to characterize swallowing physiology as a function of participant age, diagnosis, posture in the sagittal plane during swallowing (i.e. flexion, extension), bolus characteristics, or any other variable. We sought solely to determine whether HRCA machine learning techniques for unsupervised tracking of the body of the adult hyoid bone during swallowing was comparable to that of a judge trained in human swallow kinematic analysis, and to test whether it can produce clinically relevant ratings of hyoid displacement (i.e. MBSImP component #9), regardless of the swallowing condition or participant characteristics.

Methods

The Institutional Review Board at the University of Pittsburgh approved this investigation and all participants provided written informed consent. The data analyzed in this study consisted of two datasets that were collected in a similar fashion at two different timepoints. Initially, data analysis was conducted on 400 swallows from 114 patients (65 males) between the ages of 19-94, selected from a larger prospectively accrued data set (n=3072 swallows from 244 patients) who were referred for and underwent VFSSs due to suspected or confirmed dysphagia at the University of Pittsburgh Medical Center Presbyterian University Hospital. All participants were imaged in the lateral plane. Swallows with high quality images were intentionally selected for analysis based on the visibility of the entire body of the hyoid bone on each frame of the video segment, and the absence of large-scale patient motion during swallowing segments. A variety of bolus conditions (volume, texture, mode of administration) were included in the analyzed data set. Table 1 contains the bolus characteristics of the original data set. Only single swallows performed in head neutral position (i.e. no flexion, hyperextension) and with a penetration aspiration scale score <3 were included in data analysis.

After conducting machine learning with the patient data set, we then analyzed a randomly selected set of 48 swallows from 16 adults from an ongoing HRCA clinical experiment with community dwelling healthy adults with no current or prior report of swallowing difficulties. Table 2 contains the bolus characteristics of the healthy community dweller data set. These participants had no reported history of neurological disorder, surgery to the head or neck region, or chance of being pregnant. Experimental procedures for the community dwelling adults were the same aside from bolus administration procedures, which were modified to minimize radiation exposure. Participants swallowed ten thin liquid boluses in a random presentation order (5 at 3mL by spoon, 5 unmeasured self-selected volume cup sips). For presentations by spoon, participants were instructed by the researcher to "Hold the liquid in your mouth and wait until I tell you to swallow it." For presentation by cup, participants were instructed by the researcher to "Take a comfortable sip of liquid and swallow it whenever you're ready."

A standard fluoroscopy system (Precision 500D system, GE Healthcare, LLC, Waukesha, WI) set at a pulse rate of 30 PPS was used for accruing swallow video segments and a frame grabber module (AccuStream Express HD, Foresight Imaging, Chelmsford, MA) was used to capture raw videos at a rate of 60 frames per second directly from the x-ray apparatus without compression or processing. The frame rate was set at 60 frames per second to accommodate the higher necessary sampling rate for the HRCA signals and is validated by Shannon's sampling theorem [32]. Following data collection, the videos were down sampled to 30 frames per second to eliminate duplicate frames prior to human judge kinematic analysis. HRCA signals were collected simultaneously during VFSSs by placing a tri-axial accelerometer (ADXL 327, Analog Devices, Norwood, Massachusetts) and contact microphone on the anterior neck region of patients. To obtain the best signals during swallowing, the accelerometer and contact microphone were housed in custom casings to ensure flat contact surfaces with the skin, and placed over the laryngeal framework at the level of the cricoid cartilage using tape, with the accelerometer at midline overlying the cricoid arch and the microphone to the right of midline and slightly inferior to the accelerometer so as not to interfere with imaging. Figure 1 shows the placement of the sensors in a single frame of one of the video segments. The three axes of the accelerometer were aligned with the anatomical anterior-posterior, superior-inferior, and medial-lateral directions axes of each patient's neck. A power supply with a 3V output (model 1504, BK Precision, Yorba Linda, California) was used to power the accelerometer. Once the raw signals from the accelerometer were obtained during data collection, they were bandpass filtered (model P55, Grass Technologies, Warwick, Rhode Island) from 0.1 to 3000 Hz and amplified ten times. Following this, the signal data from each axis of the accelerometer were fed into a data acquisition device (National Instruments 6210 DAQ) and recorded at a sampling rate of 20 kHz by the LabView program Signal Express (National Instruments, Austin, Texas). Thus, four separate signal data sets were generated from all swallows along with the VF images.

Kinematic analysis: First raters were trained and tested in swallow kinematic analyses and then they performed swallow segmentation and frame by frame plotting of hyoid bone movement using ImageJ software and a MatLab program. For ease of analysis, videos were segmented into individual swallows. Swallowing segment onset was defined as the frame in which the bolus head first passed the shadow of the ramus of the mandible. The offset of the swallow segment was defined as the frame in which the hyoid returned to its lowest position following clearance of the bolus tail through the upper esophageal sphincter. Note that these onset and offset moments are not identical to those used in the definitions of pharyngeal response duration or pharyngeal transit duration [33] as we sought solely to compare HRCA hyoid tracking predictions to human judges' measurements and did not seek to equate HRCA results to these durational parameters. To measure hyoid bone movement, the superior posterior and inferior posterior points of the cross-sectional area of the body of the hyoid bone were plotted in each frame of each swallowing segment. These landmarks of the hyoid bone were chosen rather than the center anterior aspect of the hyoid bone because they are necessary in capturing the entire height of the hyoid body and can provide information regarding rotational hyoid movement during swallowing, which will be analyzed in future work.

Prior to performing data analysis, all raters completed training and testing of inter and intra-rater reliability for swallow segmentation using videos that were not included in the investigated dataset. Their inter- and intra-rater reliability was assessed with intra-class correlation coefficients (ICCs) [34] and they produced greater than 0.99 for both measures. In order to control for rater drift during measurements within a large data set, intra-rater reliability for segmentation was maintained on a continual basis throughout data analysis by having raters randomly select one out of every ten swallows to re-analyze and compute ICCs. Inter-rater reliability for swallow segmentation was completed for 40 (10%) of the swallows analyzed in this study with an ICC of 0.998.

To avoid judgment bias, different raters that were blinded to each other's ratings, separately completed frame-by-frame hyoid tracking and MBSImP hyoid bone displacement ratings. An MBSImP certified clinician completed all MBSImP hyoid bone ratings. Inter-rater reliability was established prior to performing ratings for this study by completing the MBSImP reliability test with a score of 90% exact agreement and by a reliability test between all MBSImP certified clinicians in our lab with greater than 80% exact agreement. Intra-rater reliability for MBSImP ratings was completed for 10% of swallows with 80% exact agreement.

For hyoid bone tracking, one rater marked the superior posterior and inferior posterior points of the cross-sectional area of the body of the hyoid bone for all 400 swallows included in the dataset. Intra-rater reliability was maintained throughout hyoid frame-by-frame tracking measurements to control for rater judgment drift throughout measurements. A randomly selected subset of 10% of measured swallows were re-rated by the same judge, returning ICCs of 0.99 across all hyoid tracking measurements. Inter-rater reliability for hyoid bone tracking was completed on 10% of the swallows from the patient dataset, which were randomly selected. In order to compare human measurements with one another for hyoid bone displacement on each frame, bounding boxes of equal area (i.e. 35x35 pixels) were generated for each human rater hyoid plotted measurements (see Figure 2). The dimensions of the bounding boxes were determined based on the average length between the plotted superior-posterior and inferior-posterior points of the body of the hyoid bone. Three trained raters completed inter-rater reliability for 10% of the swallows. To do this, six two-way comparisons were made between each pair of human ratings of hyoid bone location in each frame of the swallow. The overlap between bounding boxes for human ratings of the hyoid bone location in each frame of a swallow was calculated for each two-way comparison and then averaged. The overall overlapped percentage of exact agreement across all human rater comparisons for all frames was 79.05%. This means that the overlap of the bounding boxes (i.e. exact pixel-level agreement of where the hyoid bone was located) in each frame between the human raters was 79.05%.

To predict the exact location of the hyoid bone in each frame based on HRCA signals, a second bounding box of 35x35 pixels was generated based on a structural recurrent neural network (SRNN) with tenfold cross validation, which is an advanced machine learning technique (see Figure 2). The SRNN was developed based on the feature extraction of hyoid bone movement from the HRCA accelerometer signals. These methods are previously described elsewhere [28-31]. We used ten-fold validation to train and test the algorithm for hyoid bone prediction, which means that the total data (400 swallows) was divided into ten groups (40 swallows each group). Nine groups (360 swallows) were used for training, and one group (40 swallows) was used for testing the ability of the SRNN to predict the location of the hyoid bone in each frame based on HRCA signals alone. This training and testing process was repeated ten times, so that each of the ten groups were used for testing one time. To determine the accuracy of the SRNN, a relative overlapping percentage (ROP) of the bounding boxes for the predicted hyoid bone location (based on the SRNN) and the gold standard measurement of hyoid bone location (based on human measurement) was calculated.

To determine whether hyoid bone movement equated with “normal” or “reduced MBSImP scores for component #9,” we examined the association between 27 different signal features from the HRCA signals and MBSImP ratings of hyoid bone displacement for a subset of the swallows (76). While the MBSImP has three ratings for hyoid bone movement (0-complete anterior movement, 1-partial anterior movement, and 2-no anterior movement), there were no swallows included in the analysis that had no anterior movement. For this reason, we dichotomized hyoid bone movement into “normal” (score of 0) or “reduced” (score of 1) (See Table 3).

Results

Table 4 summarizes the results and variability (45-57.6% ROP) of the accuracy of hyoid tracking of the SRNN across the ten groups of the original patient data set. Results of the testing period revealed that the predictive ability of the SRNN using accelerometry signals alone had a ROP of 51.6% on average for the original patient data set when compared to human ratings of the hyoid bone location based on frame by frame tracking. Likewise, the SRNN using accelerometry signals alone had similar performance on the healthy community dweller data set (ROP 49.9%) when compared to human ratings. This indicates that the SRNN algorithm was able to detect the exact location of 50% or more of the bounding box containing the hyoid bone on each frame during the swallow, and that the algorithm was able to generalize to an outside dataset that was not used during the training period.

Table 5 shows the statistically significant ($p<0.05$) results from examining the association between HRCA signal features and MBSImP ratings and Table 6 summarizes the definitions of the signal features that were extracted. Results from examining the association of 27 different signal features from HRCA signals and MBSImP scores of hyoid bone displacement revealed significant differences in the signal standard deviation feature for all three axes of the accelerometer (superior-inferior, medial-lateral, anterior-posterior) and in the signal spectral centroid feature in the superior-inferior axis. The average standard deviation values of all three accelerometer axes for “reduced” MBSImP scores was significantly smaller than the average standard deviation values for “normal” MBSImP scores and the average spectral centroid values for “reduced” MBSImP scores were significantly larger than the average spectral centroid values for “normal” MBSImP scores. These differences reflect systematic differences in frequency and amplitude characteristics of the HRCA signal features that separated “normal” from “reduced” hyoid displacement that aligned with human judgments using the MBSImP.

Discussion

This study demonstrated the feasibility of using machine learning of HRCA signal features to predict the position of the body of the hyoid bone on each frame of a VFSS study. The finding of this capability, can be

Formatted: Font: 10 pt

combined with ongoing and planned future analyses of HRCA signal features for other swallow kinematic/physiologic events (e.g., duration of UES opening, caliber of AP distension of the UES, penetration-aspiration scale scores) to assess the potential of using HRCA as a noninvasive dysphagia screening method that may provide enhanced insight into patients who are at increased risk of penetration and/or aspiration due to reduced hyoid bone movement when VFSS is not readily available or feasible. This study demonstrated the feasibility of using machine learning of HRCA signal features to predict the position of the body of the hyoid bone on each frame of a VFSS study and the potential of using HRCA as a noninvasive system for monitoring this swallow physiologic event when VFSS is not available or feasible. Our machine learning algorithms succeeded in locating approximately half (51% patient data set, 49.9% healthy data set) of the hyoid body on each frame. We acknowledge that 50.75% does not sound like a high level of accuracy. However, the hyoid is a very small structure. We plotted the superior-posterior and inferior-posterior points of the hyoid body on each VFSS frame. The distance between these two points in adult males ranges between 6.04 and 14.42mm, and 6.74-11.71mm in adult females [17-18]. Detection of the position of more than 50% of these tiny structures on each frame of swallowing segments without imaging could be considered as quite remarkable, though there is room for improvement by adding more training data. It's also important to note that our trained human judges that used VFSS images for hyoid tracking had a ROP of 79% between judges, indicating variability and error with the gold standard frame-by-frame tracking method. In addition to this, while efforts are being made to increase access and feasibility of swallow kinematic analysis within the clinical setting, the large majority of clinicians continue to use subjective measurements for swallow evaluation using VFSS. Vose and colleagues (2018) published survey data based on 303 speech language pathologist members of the ASHA dysphagia special interest group. Results indicated that 5% of clinicians performing VFSS studies performed frame-by-frame analysis of data 100% of the time, one-third of respondents admitted to never performing such analyses, and only one-third of respondents reported conducting these analyses more than half of the time [13]. Therefore, the ability to track hyoid bone displacement non-invasively with a high degree of accuracy has relevant clinical applications for dysphagia screening, assessment, and treatment purposes. Together with these results and results of other work we have published regarding the ability of HRCA to differentiate between safe and unsafe swallows based on the penetration-aspiration scale [20, 26] and preliminary studies that have demonstrated the ability of HRCA to detect other temporal swallow kinematic events such as the duration of upper esophageal sphincter opening and laryngeal vestibular closure [35-36], we are optimistic about the potential of the HRCA

Formatted: Font: 10 pt

Formatted: Font color: Text 1

system to be a valuable contributor to dysphagia screening and in the future, as a noninvasive adjunct to diagnosis of swallowing disorders when VFSS is not readily available or feasible.

There is a high demand to improve the sensitivity and specificity of dysphagia screening methods in order to non-invasively, quickly, and accurately identify patients with dysphagia to mitigate adverse events that occur secondary to dysphagia and to improve the efficiency of used resources in health care settings (i.e. clinician time, cost of unnecessary procedures). Screening methods that include more discrete information beyond simple observation of a patient swallowing and observing for coughing, that provide enhanced insight into physiological differences in swallow function, such as hyoid bone displacement, would be especially useful in settings that do not have access to instrumental swallow evaluations such as videofluoroscopy. For example, in the future with further validation and improved accuracy of HRCA, clinicians and patients in settings such as skilled nursing facilities and home care with limited access to VFSS may be able to gain information about the physiologic components~~y~~ of swallowing such as hyoid bone displacement by using this noninvasive system to ~~-~~reduce delays in the initiation of interventions (e.g. VFSS to further assess swallow function). While HRCA demonstrates promise in detecting aspects of swallowing physiology and safety, further development of this system and its accuracy is warranted before deploying it as a swallowing diagnostic or biofeedback tool in the clinical setting. While hyoid bone displacement is not the only important biomechanical event that occurs during swallowing, it is associated with other physiological events including laryngeal vestibular closure and UES opening. Anterior hyoid bone displacement has also been shown to be a predictor of the risk of penetration and aspiration in patients with dysphagia [7-9]. ~~In addition to this, i~~In the future (and with further validation), HRCA signals with non-invasive neck sensors may be used as a monitoring device during meals for patients that have already undergone imaging to detect physiological swallowing impairments in real time. HRCA may be ~~using~~ used for training patients to use compensatory strategies and for biofeedback purposes in dysphagia therapy by turning signal data into a visual display of hyoid bone displacement for patients to look at and aim to improve. This would provide clinicians and patients with an advantageous, objective method to characterize effectiveness of hyoid bone displacement augmentation while performing swallow maneuvers such as the Mendelsohn maneuver.

Machine learning is an iterative process of training computer algorithms using gold standard data, and then testing the precision of the algorithms with novel gold standard data to determine the algorithm’s ability to approximate the gold standard measurements. When successful, the algorithms can then stand alone. This

Formatted: Font color: Text 1

Formatted: Font color: Text 1

technology is the basis for consumer products such as smart watches, phones, and driver assisted technologies that have been widely adopted. Early in the development of the applications of machine learning, it is difficult to understand how a technology can perform acts previously accomplished either by humans or by different technologies. In our line of research, we are implementing machine learning techniques with HRCA signal feature extraction to determine if some swallow kinematic measurements can be performed by HRCA as accurately as a human judge using VFSS images. While this study has demonstrated HRCA's feasibility in hyoid tracking, future studies should focus on improving the SRNN algorithm to more accurately detect the exact location of the hyoid bone on each VFSS frame during swallowing. The algorithm will be improved by analyzing additional data for training and testing purposes, including swallowing data obtained from healthy community-dwelling adults across the lifespan, a process we have already initiated. We also intend to expand this research to include the many hundreds of VFSS swallows in our database that have PAS scores >3 and that do not have ideal visualization of the hyoid on every frame by using artificial intelligence to predict hyoid position based on trajectory and other displacement signal features. While this study aimed to determine the ability of deep learning (SRNN) to predict hyoid bone movement regardless of etiology of dysphagia, we also plan to determine the accuracy of deep learning across patient populations and the extent to which signal features of hyoid bone displacement vary based on the disease process causing dysphagia. Future work should also examine the ability of HRCA signals to predict MBSImP scores of anterior hyoid bone displacement and other MBSImP components based on the preliminary results in this study that revealed significant differences in signal features between impaired and normal MBSImP scores. These future research directions can increase patient and caregiver access to noninvasive swallowing monitoring that can be deployed in day to day settings and extend dysphagia screening and diagnostics beyond the bedside or x-ray room.

Limitations

While the main purpose of this study was not to characterize swallowing physiology based on patient characteristics or swallowing conditions during VFSSs, it's important to note that we did not control for these variables. We collected data in a manner consistent with standard clinical care. It could be argued that this limitation reduces the internal validity of our study results. However, because of this design component, the results of our study are generalizable and have direct applications in the clinical setting because they were generated from data produced in ordinary clinical settings with their inherently typical constraints against perfect methods of data collection. Additionally, as mentioned previously, the primary aim of this study was to establish the ability of HRCA signals to effectively track hyoid bone displacement and characterize MBSImP scores regardless of patient characteristics or swallowing conditions. We did not seek to determine whether hyoid bone displacement values were within the normal range based on an anatomical scalar, to characterize patients based on their diagnosis, or to characterize swallows based on bolus, posture, or other swallow-specific variables. In the future, we will explore the ability of HRCA signals to characterize swallowing physiology across different patient populations and testing conditions to gain additional insight into this methodology's potential value in dysphagia screening and diagnostics. Testing across these conditions is also a way to demonstrate the robustness of the machine learning algorithm. An additional limitation of this work is that we examined the ability of HRCA to detect one temporal kinematic event of swallowing (e.g. hyoid bone displacement). It should be emphasized that hyoid bone displacement, or any single measurement, alone should not be used to determine swallowing impairment. While hyoid bone displacement is associated with other important swallow kinematic events such as upper esophageal sphincter opening and laryngeal vestibular closure and while reduced hyoid bone displacement is associated with increased risk of penetration and/or aspiration [7-9], it is vital for clinicians to consider all kinematic swallow events that may contribute to impaired swallow function and/or airway protection. In the future, we plan to combine the machine learning algorithms we have developed in other studies for detecting other kinematic swallow events (e.g. upper esophageal sphincter opening, laryngeal vestibular closure) [35-36] with the machine learning algorithm we used for detecting hyoid bone displacement in this study in order to more accurately and robustly provide insight into swallowing physiology. This study provides substantial preliminary evidence regarding the ability of HRCA to track hyoid bone displacement and to independently provide information about MBSImP scores of anterior hyoid bone movement non-invasively using HRCA signals without human mediation. While we included a relatively large sample of swallows in our data analysis, the accuracy of machine learning improves with larger samples of data. As such, it will be important to continue adding fully-analyzed swallows to our growing data set in order to improve the accuracy of this non-invasive method. Likewise, we included a small preliminary analysis examining the association between HRCA signals and MBSImP scores, which would also be improved with a larger sample of swallows.

Conclusion

This study found that the position of more than half of the bounding box containing the hyoid body can be independently located on any given VFSS frame by our HRCA system via a SRNN using signal features obtained from non-invasive neck sensors without use of videofluoroscopy images or human judgment, and that HRCA signals combined with statistical methods can provide information about MBSImP ratings of anterior hyoid bone displacement. These preliminary results contribute to a growing body of literature that demonstrates that HRCA has future potential as an effective, non-invasive dysphagia screening system, and encouraging promise as an adjunct biofeedback modality during therapy.

1. Martin-Harris B. The VFS study. *Phys Med Rehabil Clin N Am*. 2008;19(4):769-785. doi:10.1016/j.pmr.2008.06.004.The

2. Molfenter SM, Steele CM. Physiological variability in the deglutition literature: Hyoid and laryngeal kinematics. *Dysphagia*. 2011;26(1):67-74. doi:10.1007/s00455-010-9309-x

3. Vandaele, D. J., Perlman, A. L., & Cassell, M. D. (1995). Intrinsic fibre architecture and attachments of the human epiglottis and their contributions to the mechanism of deglutition. *Journal of Anatomy, 186 (Pt 1)*(Pt 1), 1-15.

4. Kim Y, McCullough GH. Maximum hyoid displacement in normal swallowing. *Dysphagia*. 2008;23(3):274-279. doi:10.1007/s00455-007-9135-y

5. Kendall KA, Leonard RJ. Hyoid movement during swallowing in older patients with dysphagia. *Arch Otolaryngol - Head Neck Surg*. 2001;127(10):1224-1229. doi:10.1001/archotol.127.10.1224

6. Matsuo K, Palmer JB. Anatomy and physiology of feeding and swallowing: Normal and abnormal. *Phys Med Rehabil Clin N Am*. 2008;19(4):691-707. doi:10.1016/j.pmr.2008.06.001

7. Zhang Z, Perera S, Donohue C, et al. The prediction of risk of penetration–aspiration via hyoid bone displacement features. *Dysphagia*. 2019;(0123456789). doi:10.1007/s00455-019-10000-5

8. Perlman AL, Booth BM, Grayhack JP. Videofluoroscopic predictors of aspiration in patients with oropharyngeal dysphagia. *Dysphagia*. 1994;9(2):90-95. doi:10.1007/BF00714593

9. Molfenter SM, Steele CM. Kinematic and temporal factors associated with penetration-aspiration in swallowing liquids. *Dysphagia*. 2014;29(2):269-276. doi:10.1007/s00455-013-9506-5

10. McCullough GH, Kim Y. Effects of the mendelsohn maneuver on extent of hyoid movement and UES opening post-stroke. *Dysphagia*. 2013;28(4):511-519. doi:10.1007/s00455-013-9461-1

11. Wheeler-Hegland KM, Rosenbek JC, Sapienza CM. Submental sEMG and hyoid movement during mendelsohn maneuver, effortful swallow, and expiratory muscle strength training. *J Speech, Lang Hear Res*. 2008;51(5):1072-1087. doi:10.1044/1092-4388(2008/07-0016)

12. Martin-Harris B, Brodsky MB, Michel Y, et al. MBS measurement tool for swallow impairment-MBSimp: Establishing a standard. *Dysphagia*. 2008;23(4):392-405. doi:10.1007/s00455-008-9185-

13. Vose AK, Kesneck S, Sunday K, Plowman E, Humbert I. A Survey of Clinician Decision Making When Identifying Swallowing Impairments and Determining Treatment. *Journal of Speech, Language, and*

Field Code Changed

Hearing Research. 2018;61(11):2735-2756. doi:10.1044/2018_jslhr-s-17-0212

14. Suiter DM, Sloggy J, Leder SB. Validation of the yale swallow protocol: A prospective double-blinded videofluoroscopic study. *Dysphagia*. 2014;29(2):199-203. doi:10.1007/s00455-013-9488-3

15. Groves-Wright KJ, Boyce S, Kelchner L. Perception of wet vocal quality in identifying penetration/aspiration during swallowing. *J Speech, Lang Hear Res*. 2009;53(3):620-632. doi:10.1044/1092-4388(2009/08-0246)

16. Waito A, Bailey GL, Molfenter SM, Zoratto DC, Steele CM. Voice-quality abnormalities as a sign of dysphagia: Validation against acoustic and videofluoroscopic data. *Dysphagia*. 2011;26(2):125-134. doi:10.1007/s00455-010-9282-4

17. Ramagalla, A. R., Sadanandam, P., & Rajasree, T. K. (2014). Age related metric changes in the hyoid bone. *IOSR J Dent Med Sci*, *13*(7), 54-56. doi:10.9790/0853-13765456

18. Loth A, Corny J, Santini L, et al. Analysis of hyoid–larynx complex using 3D geometric morphometrics. *Dysphagia*. 2015;30(3):357-364. doi:10.1007/s00455-015-9609-2

19. Brates D, Molfenter SM, Thibeault SL. Assessing Hyolaryngeal Excursion: Comparing Quantitative Methods to Palpation at the Bedside and Visualization During Videofluoroscopy. *Dysphagia*. 2019;34(3):298-307. doi:10.1007/s00455-018-9927-2

20. Sejdic E, Steele CM, Chau T, Member S. Classification of penetration – aspiration versus healthy swallows using dual-axis swallowing accelerometry signals in dysphagic subjects. 2013;60(7):1859-1866.

21. Dudik JM, Coyle JL, Sejdic E. Dysphagia Screening : Contributions of cervical auscultation signals and modern signal-processing techniques. 2015;45(4):465-477.

22. Dudik JM, Jestrovic I, Luan B, Coyle JL, Sejdic E. A comparative analysis of swallowing accelerometry and sounds during saliva swallows. 2015:1-15.

23. Dudik JM, Kurosu A, Coyle JL, Sejdic E. A comparative analysis of DBSCAN , K-means, and quadratic variation algorithms for automatic identification of swallows from swallowing accelerometry signals. 2015;59:10-18. doi:10.1016/j.compbiomed.2015.01.007

24. Jestrovic I, Dudik JM, Luan B, Coyle JL, Sejdic E. Baseline characteristics of cervical auscultation signals during various head maneuvers. 2014;43(2013):2014-2020. doi:10.1016/j.compbiomed.2013.10.005

25. Movahedi F, Kurosu A, Coyle JL, Perera S, Sejdic E. Computer methods and programs in biomedicine: A

comparison between swallowing sounds and vibrations in patients with dysphagia. 2017;144:179-187. doi:10.1016/j.cmpb.2017.03.009

26. Dudik JM, Coyle JL, El-Jaroudi A, Mao ZH, Sun M, Sejdić E. Deep learning for classification of normal swallows in adults. *Neurocomputing*. 2018;285(April):1-9. doi:10.1016/j.neucom.2017.12.059

27. Kurosu A, Coyle JL, Dudik JM, Sejdic E. Detection of swallow kinematic events from acoustic high-resolution cervical auscultation signals in patients with stroke. *Arch Phys Med Rehabil*. 2019;100(3):501-508. doi:10.1016/j.apmr.2018.05.038

28. Rebrion C, Zhang Z, Khalifa Y, et al. High-resolution cervical auscultation signal features reflect vertical and horizontal displacements of the hyoid bone during swallowing. *IEEE J Transl Eng Heal Med*. 2019;7(January). doi:10.1109/JTEHM.2018.2881468

29. Zhang Z, Coyle JL, Sejdić E. Automatic hyoid bone detection in fluoroscopic images using deep learning. *Sci Rep*. 2018;8(1):1-9. doi:10.1038/s41598-018-30182-6

30. He Q, Perera S, Khalifa Y, Zhang Z, Mahoney A, Sabry A, Donohue C, Coyle J, & Sejdic E. The association of high-resolution cervical auscultation signal features with hyoid bone displacement during swallowing. *Transactions on Neural Systems & Rehabilitation Engineering.* 2019. doi: 10.1109/TNSRE.2019.2935302

31. Mao S, Zhenwei Z, Khalifa Y, Donohue C, Coyle J, Sejdic E. Neck sensor-supported hyoid bone movement tracking during swallowing. *Royal Society*. 2019; 6(7). doi.org/10.1098/rsos.181982

32. Oppenheim AV, Schafer RW. *Discrete-Time Signal Processing*. Harlow: Pearson; 2014.

33. Lof GL, Robbins JA. Test-retest variability in normal swallowing. *Dysphagia*. 1990;4(4):236-242. doi:10.1007/BF02407271

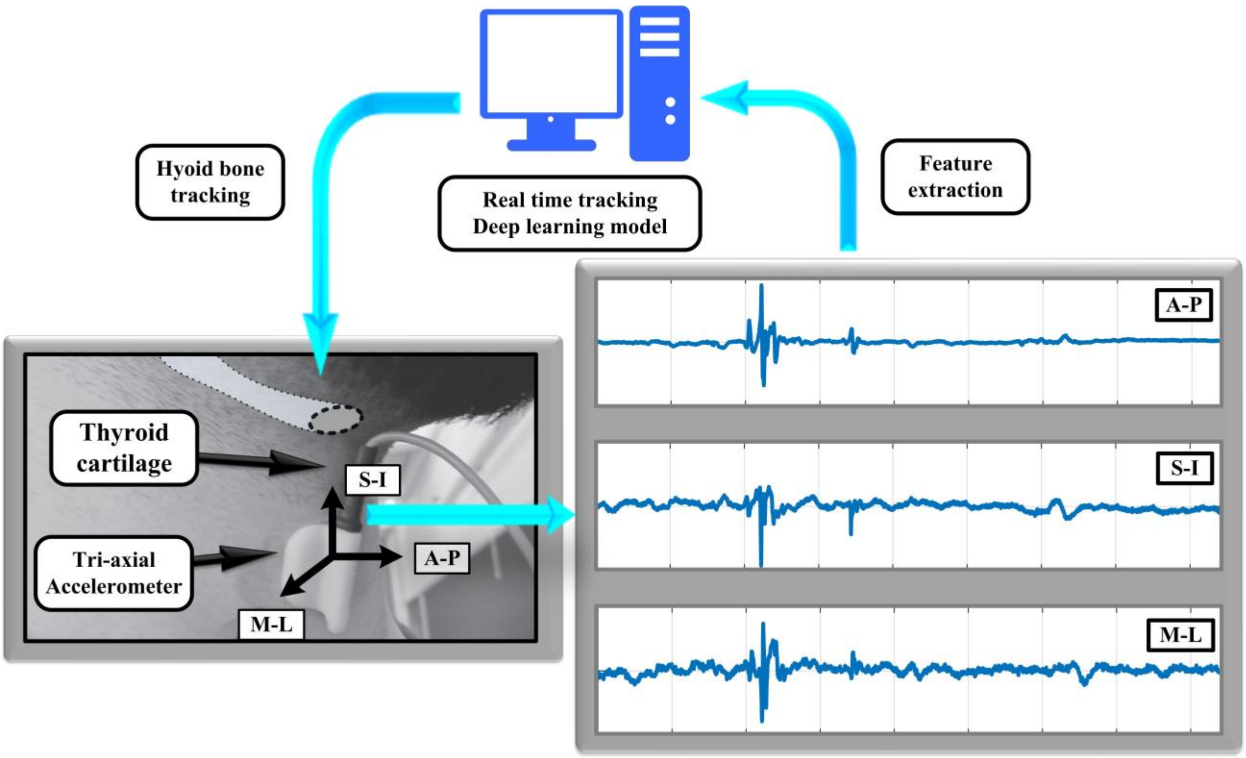
34. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin.* 2005;86(2):1-9.

35. Sabry A, Shitong M, Mahoney A, Khalifa Y, Sejdic E, Coyle J. Automatic estimation of laryngeal vestibular closure duration using high resolution cervical auscultation signals. Presentation at the *American Speech-Language Hearing Association Convention,* Orlando, FL. November 2019.

36. Donohue C, Khalifa Y, Sejdic E, Coyle J. (March 2019). How closely do machine ratings of duration of UES opening during videofluoroscopy approximate clinician ratings using kinematic analysis and the MBSImP? Presentation at the *Dysphagia Research Society Annual Meeting,* San Diego, CA.

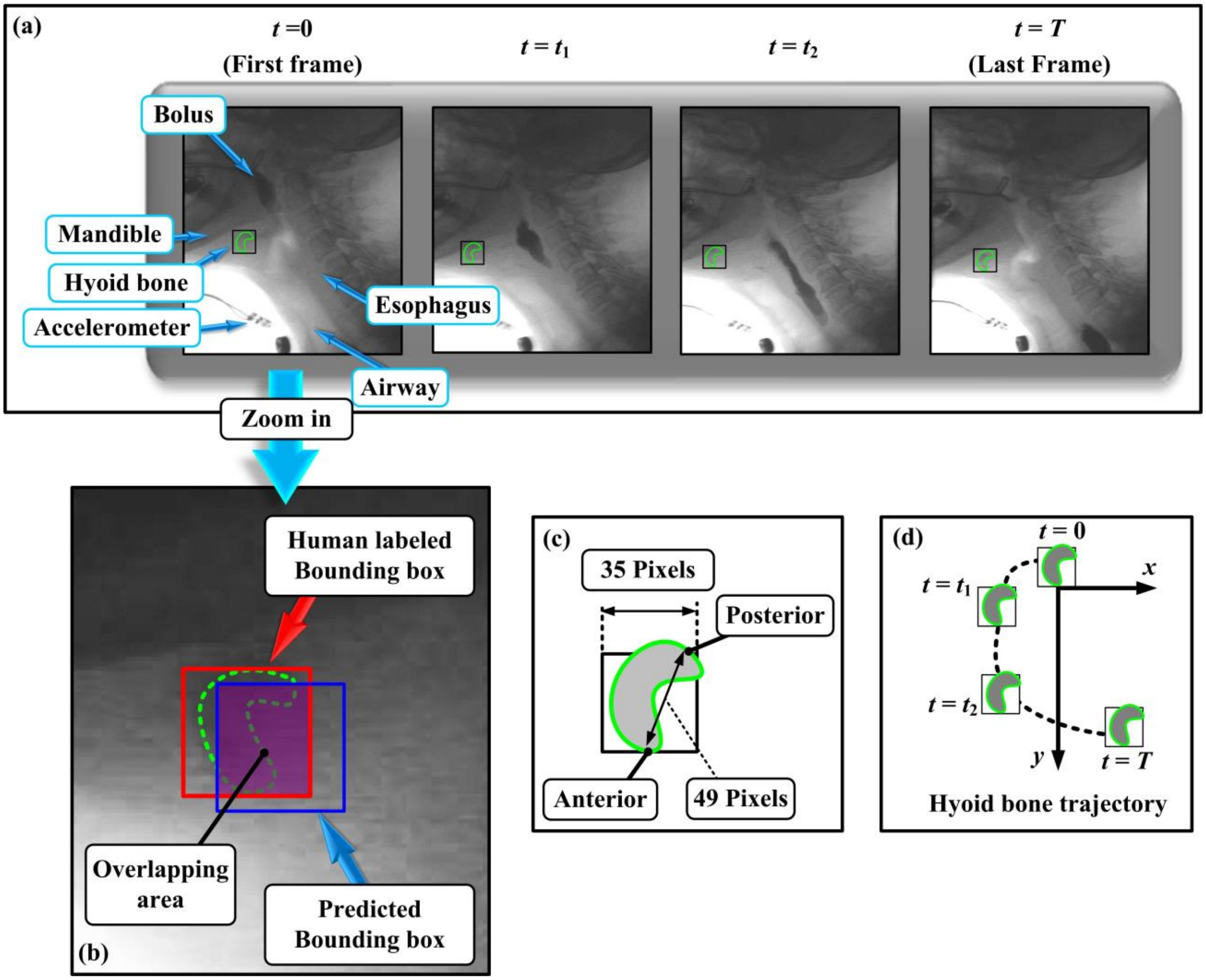
Figure

Figure 1: This shows the placement of the non-invasive neck sensors on the anterior laryngeal framework and the extraction of the acoustic and vibratory signals for the machine learning algorithm to track hyoid bone displacement.



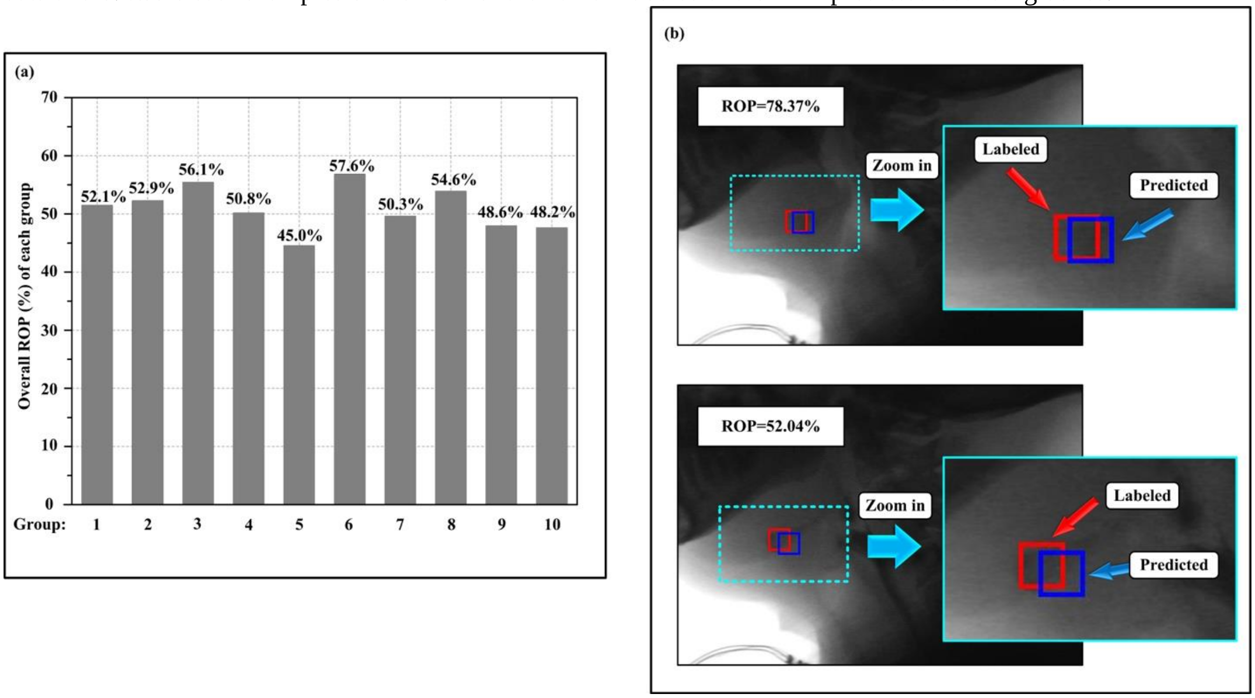
Figure

Figure 2: a. This shows the tracking of the hyoid bone over a period of time, b. the ROP of the human labeled and SRNN predicted bounding boxes, c. the dimensions of the bounding box, d. and the overall hyoid bone displacement over time.



Figure

Figure 3: a. This shows the ROP of the human labeled and SRNN predicted bounding boxes across the ten groups of data and b. two visual examples of the ROP of the human labeled and SRNN predicted bounding boxes.



Table

Table 1: *Bolus characteristics for all swallows included in the original patient data set.*

Bolus viscosity and utensil	Number of swallows	Percentage of swallows
Thin by spoon	73	18.25%
Thin by cup	121	30.25%
Thin by straw	35	8.75%
Saliva swallows	5	1.25%
Nectar thick liquid by spoon	40	10%
Nectar thick liquid by cup	44	11%
Nectar thick liquid by straw	10	2.5%
Pudding by spoon/cup	49	12.25%
Cookie	23	5.75%

Table

Table 2: *Bolus characteristics for all swallows included in the healthy community dweller patient data set.*

Bolus viscosity and utensil	Number of swallows	Percentage of swallows
Thin by spoon	24	50%
Thin by cup	24	50%

Note: Thin by spoon swallows were 3 mL and thin by cup swallows ranged from 3-60 mL.

Table

Table 3: *MBSImP Anterior Hyoid Bone Displacement Ratings for a Subset of Swallows (76).*

MBSImP score	Number of Swallows	Percentage of Swallows
Complete Anterior Movement (0)	39	51.3%
Partial Anterior Movement (1)	37	48.7%

Table

Table 4: *Average ROP % for each of the ten groups used during the training and testing of the SRNN.*

Group	One	Two	Three	Four	Five	Six	Seven	Eight	Nine	Ten
Overall ROP %	52.1%	52.9%	56.1%	50.8%	45.0%	57.6%	50.3%	54.6%	48.6%	48.2%

Table

Table 5: *Summary of the statistically significant HRCA signal features associated with MBSImP ratings of hyoid bone displacement.*

	Standard Deviation	Skewness	Kurtosis	Lempel-Ziv complexity	Entropy Rate	Peak Frequency	Spectral Centroid	Bandwidth	Wavelet entropy
Anterior-posterior	**0.0453***	0.3013	0.8234	0.9111	0.8386	.09484	0.9191	.06907	.09356
Superior-inferior	**0.0173***	0.9197	0.5516	0.7538	0.9701	0.5587	**0.0115***	0.6130	0.9566
Medial-lateral	**0.0117***	0.9774	0.6644	0.1817	0.1964	0.6105	0.3687	0.3709	0.4234

Note: *= $p<0.05$

Table

Table 6: *Summary of the features extracted from the HRCA signals.*

Domain	Feature	Significance
Time Domain		
	Standard deviation	Reflect the signal variance around its mean value.
	Skewness	Describe the asymmetry of amplitude distribution around mean.
	Kurtosis	Describe the peakness of the distribution relative to normal distribution.
Information-Theoretic domain		
	Lempel-Ziv Complexity	Describe the randomness of the signal.
	Entropy rate	Evaluate the degree of regularity of the signal distribution.
Frequency domain		
	Peak Frequency (Hz)	Describe the frequency of maximum power.
	Spectral Centroid (Hz)	Evaluate the median of the spectrum of the signal.
	Bandwidth (Hz)	Describe the range of frequencies of the signal.
Time-Frequency Domain	Wavelet Entropy	Evaluate the disorderly behavior for non-stationary signal.