



Revista EIA
ISSN 1794-1237
e-ISSN 2463-0950
Año XVIII/ Volumen 19/ Edición N.38
Junio-Diciembre de 2022
Reia3831 pp. 1-12

Publicación científica semestral
Universidad EIA, Envigado, Colombia

**PARA CITAR ESTE ARTÍCULO /
TO REFERENCE THIS ARTICLE /**
Betancur Rengifo, J. P.; Restrepo Uribe,
J. P.; Perez Giraldo, E.; Orozco Duque,
A. (2022)
Caracterización de sonidos
deglutorios adquiridos mediante
auscultación cervical en sujetos sanos
y con disfagia.. Revista EIA, 19(38),
Reia3831. pp. 1-12.
<https://doi.org/10.24050/reia.v19i38.1579>

✉ *Autor de correspondencia:*

Betancur Rengifo, J. P. (Juan Pablo)
Estudiante de ingeniería biomédica
Correo electrónico:
juanbetancur250634@correo.itm.edu.co

Recibido: 14-12-2021
Aceptado: 18-04-2022
Disponible online: 01-06-2022

Caracterización de sonidos deglutorios adquiridos mediante auscultación cervical en sujetos sanos y con disfagia.

✉ JUAN PABLO BETANCUR RENGIFO¹
JUAN PABLO RESTREPO URIBE¹
ESTEFANIA PEREZ GIRALDO¹
ANDRÉS OROZCO DUQUE¹

1. Instituto Tecnológico Metropolitano

Resumen

La deglución es un acto de alta complejidad neuromuscular debido a que intervienen más de 30 pares musculares y 5 pares craneales en un periodo corto de tiempo. Esta se divide en 4 etapas: pre oral, oral, orofaríngea y esofágica, una alteración en el desarrollo normal de alguna de estas fases puede desarrollar un síntoma secundario a enfermedades neuromusculares y neurogénicas que se conoce como disfagia, esta puede traer consigo muchas dificultades para quien la padece, entre esta neumonía bronquial, desnutrición, deshidratación o incluso la muerte por asfixia. La identificación de características que ayuden a reconocer dicho síntoma, además de describir correctamente el proceso deglutorio, es de gran importancia ya que los métodos existentes son invasivos. La auscultación cervical es una técnica mediante la cual se puede obtener información del cierre glótico en el proceso deglutorio por medio de señales de audio, y que puede ser analizada de manera off line. El objetivo de este estudio es evaluar diferentes métodos de caracterización de señales de auscultación cervical y desarrollar un modelo de aprendizaje automático con sonidos de eventos deglutorios segmentados de forma manual para clasificar entre sujetos de control y pacientes con disfagia orofaríngea. Los resultados mostraron, con una exactitud máxima de 75 %, que por medio de señales de auscultación cervical es posible identificar sujetos con disfagia, de igual manera se logró identificar que la potencia media de los segmentos deglutorios fue la característica con mejor rendimiento (curva ROC) y una distribución diferente entre clases según la prueba de U-Mann-Whitney, para discriminar entre sanos y pacientes en diferentes actividades deglutorias.

Palabras claves: Auscultación cervical, clasificación, disfagia, extracción de características, procesamiento de señales, exactitud, rendimiento, dificultades, invasivos, aprendizaje de máquina.

Characterization of swallowing sounds through Cervical Auscultation in healthy and dysphagic subjects.

Abstract

Swallowing is an act of high neuromuscular complexity due to the involvement of more than 30 muscle pairs and 5 cranial pairs that occurs in a short period of time. This activity is divided into 4 stages: pre-oral, oral, oropharyngeal and esophageal, an alteration in the normal development of this process can develop a symptom secondary to neuromuscular and neurogenic diseases that is known as dysphagia. Dysphagia can bring many difficulties for those who suffer from it including bronchial pneumonia, malnutrition, dehydration or even death by asphyxia. The identification of characteristics that help recognize this symptom, in addition to correctly describing the swallowing process is of great importance since the existing methods are invasive. Cervical auscultation is a technique by which information about the gothic closure can be obtained in the swallowing process using audio signals, and which can be analyzed offline. The aim of this study is to evaluate different methods of characterization of cervical auscultation signals and develop a machine learning model with manually segmented swallowing event sounds to classify between control subjects and patients with oropharyngeal dysphagia. The results showed, with a maximum accuracy of 75% that by means of signs is cervical auscultation it is possible to identify dysphageal subjects. In the same way it was possible to identify that the average potency of the swallowing segments was the feature with the best score (ROC curve) and that exists a different distribution between classes in this characteristic according to the U-Mann-whitney test to discriminate between healthy and pathologic subjects during different swallowing activities.

Index Terms: Cervical auscultation, classification, dysphagia, feature extraction, signal processing, accuracy, score, difficulties, invasive, machine learning.

1. Introducción

La deglución es el proceso mediante el cual se transportan sólidos y líquidos de la boca al estómago. Se considera una de las funciones neuromotoras más complejas, teniendo en cuenta que requiere un alto nivel de coordinación de más de 30 pares de músculos involucrados en un período de uno a dos segundos. Además, se estima que un adulto sano puede deglutir alrededor de 2400 a 2600 veces al día [1], [2].

La deglución se divide en cuatro etapas, pre-oral, oral, faríngea y esofágica. La etapa pre-oral incluye desde que se toma el alimento y es llevado a la boca, la etapa oral se da de manera voluntaria y consiste en la preparación del bolo en la boca, la masticación y finaliza con la propulsión que realiza la lengua sobre el bolo (masa de comida formada en la boca). Cuando se alcanza la fase faríngea, todas las otras vías como la nasofaringe, la cavidad oral y la laringe se cierran facilitando que el bolo llegue a la orofaringe. Finalmente, la fase esofágica, la cual es involuntaria, va desde la llegada del bolo al esfínter esofágico superior hasta la salida por el esfínter esofágico inferior [3].

En el proceso descrito anteriormente se pueden presentar alteraciones que se conocen como disfagia [4]. La disfagia son complicaciones asociadas al proceso de

deglución como neumonía bronquial, desnutrición, deshidratación o incluso la muerte súbita por asfixia, además de afectaciones psicológicas por la pérdida de la capacidad de alimentarse normalmente [3].

En la actualidad se calcula que la incidencia de la disfagia esta entre 16 % al 23 % en personas mayores de 76 años, debido al deterioro natural del envejecimiento o incluso en la funcionalidad neuromuscular [5]. Por otro lado, la disfagia es un síntoma que se encuentra asociado a condiciones clínicas como el Accidente Cerebrovascular (ACV), cáncer de cuello, Alzheimer, Parkinson entre otras enfermedades de etiología neurogénica. [4], [6].

Las técnicas de diagnóstico de la disfagia más aplicadas en la actualidad son la evaluación clínica, la videofluoroscopia (VFSS) y la evaluación fibroscópica de la deglución (FEES), aunque todas estas tienen ciertos inconvenientes. La primera técnica es altamente subjetiva y dependiente de la experiencia del evaluador. Por otro lado, la VFSS permite observarlas regiones anatómicas de interés, especialmente la cabeza y cuello, así como el tránsito del bolo durante la deglución, lo cual permite tener información directa de las estructuras que están afectando el proceso deglutorio, de aspiraciones o penetraciones. Sin embargo, para lograr dichas imágenes los pacientes deben ser sometidos a radiación ionizante, lo cual reduce la frecuencia con la que se practica dicho examen [7]–[11].

Debido a las complicaciones de los métodos tradicionales, ha sido de interés desarrollar técnicas que sirvan de apoyo para el diagnóstico de la disfagia de una manera objetiva, reproducible, económica y sin representar un riesgo para el paciente. Diversos estudios se han centrado en diferenciar momentos deglutorios, identificando cada una de las fases de la deglución descritas anteriormente. Para ellos usaron señales de electromiografía de superficie adquiridas sobre los músculos suprahioides, infrahioides y maseteros; implementaron una máquina de soporte de vectores (SVM) que alcanzo una exactitud del 90% [11]. Por otro lado, la deglución se ha analizado por medio de AC en compañía de señales de acelerometría adquiridas sobre el cartílago cricoides. Con estas señales se ha buscado clasificar degluciones, tos, tos fingida, o no degluciones, para se implementó una red neuronal profunda y se alcanzaron exactitudes de hasta el 95% [12].

La auscultación cervical (CA) por sí sola, es una técnica que se ha utilizado para escuchar los sonidos de la deglución durante la fase faríngea principalmente por personal médico [11]. Se ha reportado que estos sonidos pueden ser divididos en tres momentos: el primero asociado a la elevación de la laringe, el segundo con la apertura del esfínter esofágico superior; y el tercero a las estructuras volviendo a su posición anatómica normal luego de la deglución. Otros estudios han intentado verificar los tiempos de la deglución con sonidos deglutorios encontrando que en promedio pueden tener una duración de 1,19 segundos para saliva, 1,07 segundos para líquidos y hasta 17,33 segundos para sólidos, aunque a pesar de estos estudios, aún no se logra obtener una relación entre los sonidos deglutorios y los eventos fisiológicos [14]. Hoy en día, estos sonidos deglutorios pueden ser almacenados como señales digitales y analizados. Según [15] es posible clasificar las degluciones de personas sanas y las de personas con enfermedades neurodegenerativas mediante el aprendizaje automático y Auscultación Cervical de Alta Resolución (HRCA) la cual consiste en el uso de micrófonos para auscultar sonidos de la deglución en compañía de acelerómetros que permiten rastrear el movimiento del hueso hioides [8].

Aunque la AC es una técnica ampliamente utilizada por los especialistas para la evaluación de la deglución como se evidencia en [14], no llega a tener la misma

objetividad y exactitud que otras técnicas instrumentales como la FEES o la VFSS debido a su sesgo en cuanto a subjetividad de acuerdo al profesional que realice la evaluación, ya que la interpretación del sonido es realizada por el profesional que efectúe la auscultación, además se ve afectada por el ruido producido por diferentes estructuras anatómicas. Es por esto que el objetivo de esta investigación es implementar técnicas de caracterización y clasificación que permitan realizar una evaluación automática y cuantitativa de los sonidos deglutorios adquiridos mediante AC.

2. Materiales y métodos

2.1. Bases de datos

Se contó con la participación de 22 sujetos de control, 11 mujeres y 11 hombres, con edad máxima de 68 años y mínima de 27 años y con un promedio de $41,14 \pm 13,55$ años los cuales no presentaban dificultades deglutorias y eran considerados sanos para fines de la investigación. Por otro lado, se contó con 22 sujetos con diagnóstico confirmado de disfagia orofaríngea por causas como esclerosis múltiple, esclerosis lateral amiotrófica, enfermedad de párkinson, distrofia muscular de Duchene, TEC severo, entre otras, dando un total de 10 mujeres y 12 hombres, con edad máxima 65 años, mínima de 19 años y una edad promedio de $41,00 \pm 13,54$ años. Es importante resaltar que no todos los sujetos disfágicos lograron completar las tareas deglutorias por diversas complicaciones.

Todos los participantes de este proyecto fueron informados sobre el proceso y las implicaciones de dicha prueba. En el grupo de pacientes se consideraron criterios de inclusión como ser mayor de edad, comprensión de órdenes verbales, diagnóstico de disfagia funcional oral u orofaríngea secundaria a etiología neurogénica central, y de exclusión frente a existencia de disfagia esofágica estructural o mecánica, piel de región facial o cervical irradiada, edema o hematomas a nivel orofacial y cervical, disección quirúrgica en los últimos tres meses en piel de cuello, hipoxemia severa, estimulador cerebral profundo implantado o embarazo. Para el grupo de control se tuvo presente que la persona fuera mayor de edad y se aplicó un filtro médico en que se evaluaba que la persona no tuviera ningún tipo de sintomatología asociada a la disfagia.

A cada sujeto se le solicitó realizar actividades deglutorias controladas, las cuales consistieron en la ingesta de 5, 10 y 20 mL de yogur y agua, volúmenes recomendados y seguros según la reología [4], mientras se almacenaban las señales de audio por medio del estetoscopio electrónico en el software de adquisición desarrollado por [15].

2.2. Adquisición

Para la adquisición de señales de AC se utilizó un estetoscopio electrónico E-scope Cardionics conectado a una tarjeta de adquisición de datos DAQ NI USB-6215 de la National Instruments y usando el software MODAC desarrollado por [16] en trabajos anteriores para adquirir datos de auscultación cervical como de acelerometría. Las señales fueron muestreadas a una frecuencia de 10 kHz.

Tabla 1. Formulación matemática de las características utilizadas

Dominio	Característica	Descripción	Ecuación
Frecuencia	Ratio de frecuencia. (1)	Relación entre componentes de alta y baja frecuencia.	$\frac{\sum_{j=LLC}^{ULC} P_j}{\sum_{j=LHC}^{UHC} P_j}$
	Potencia media. (2)	Calcula la media del espectro de potencia de la señal	$\frac{\sum_{j=1}^M P_j}{M}$
	Frecuencia media. (3)	Calcula el promedio de frecuencias presentes en la señal.	$\frac{\sum_{j=1}^M f_j P_j}{\sum_{j=1}^M P_j}$
	Frecuencia mediana. (4)	Halla el punto medio del espectro de frecuencias dividiéndolo en igual amplitud.	$\sum_{j=1}^{f_m} P_j = \frac{1}{2} P_t$
	Pico de frecuencia. (5)	La frecuencia de mayor amplitud que se encuentra en la señal.	$\max(P_j)$ $j = 1, \dots, M$
Teoría de la información	Entropía de muestreo. (6)	Usada generalmente para evaluar la complejidad temporal de una señal fisiológica.	$\lim_{N \rightarrow \infty} (-\ln \frac{A^m(r)}{B^m(r)})$
	Complejidad de Lempel-Ziv. (7)	Se binariza la envolvente de la señal según un umbral y al final se cuenta la cantidad de diferentes patrones binarios encontrados.	$\frac{k * \log_{100} n}{n}$
	Entropía de Shannon. (8)	Mide la incertidumbre y cantidad de información que contiene una variable aleatoria.	$-\sum_{i=1}^n p(x_i) \log(p(x_i))$ <i>donde</i> $0 \leq p(x_i) \leq 1$ $\sum_{i=1}^n p(x_i) = 1$
Tiempo-frecuencia	Energía de los coeficientes de wavelet. (1 de aproximación y 8 de detalle)	Evalúa la energía contenida en distintos rangos	$\frac{1}{100} (E_{ADL} + \sum_{j=1}^{DL} E_{Dj})$

Donde: f_j = frecuencia del espectro en j ; M = tamaño de la densidad espectral de frecuencia; P_j = espectro de potencia; = frecuencia mediana; = acumulado del espectro de potencia; ULC = frecuencia de corte superior de la banda de baja frecuencia; LLC = frecuencia de corte inferior de la banda de baja frecuencia; UHC = frecuencia de corte superior de la banda de alta frecuencia; LHC = frecuencia de corte inferior de la banda de alta frecuencia; r = tolerancia de aceptación entre emparejamientos;) y) = probabilidad de que dos trayectorias se crucen para $m+1$; m = puntos; n = iteración del algoritmo; K = número de bloques en que es fragmentada la señal, =variable aleatoria; p (= probabilidad de que el evento x ocurra; DL = número de niveles de descomposición; = Energía relativa del coeficiente de aproximación final.

2.3. Segmentación

El análisis de las señales de AC se basó en la identificación de segmentos de señal donde sucedía la deglución. Este proceso se realizó de manera manual escuchando las señales e identificando principalmente el cierre glótico, proceso que hace parte de la etapa faríngea de la deglución. Se verifico de manera visual que el segmento identificado como deglución presentara características visuales acorde a lo escuchado, principalmente que se presenta cambios en la amplitud de la señal.

2.4. Procesamiento

Para el análisis computacional de las señales de AC se normalizaron las mismas buscando evitar sesgos producto de la amplitud de estas. Además, se utilizó un filtro pasa bandas de topología Butterworth de orden 4 con frecuencias de corte entre 80 Hz y 2000 Hz. Lo anterior considerando que algunas investigaciones reportan que la banda de interés de señales AC está comprendida aproximadamente entre los 50 a los 3000 Hz [15], [16].

2.5. Extracción de Características

Sobre los segmentos deglutorios identificados por medio de segmentación manual se extrajeron características en la frecuencia, tiempo-frecuencia y de teoría de la información, las cuales ha sido reportadas comúnmente en trabajos previos. La Tabla I muestra las características calculadas junto con el planteamiento matemático de las mismas [9], [14], [16]–[18].

2.6. Clasificación

La clasificación de sujetos sanos y patológicos consistió en la implementación de una SVM optimizada por medio de una búsqueda en cuadrícula. De igual manera, con la SVM se implementó una selección de características secuencial hacia adelante. La optimización consistió en determinar el mejor Kernel (lineal, radial y sigmoide), el Costo y el γ ($10e^{-05}$ hasta $10e^5$ en escala logarítmica). Igualmente se buscó determinar el número de características más adecuado para obtener los mejores resultados. Considerando que para algunas de las tareas deglutorias analizadas la cantidad de sujetos era limitada, se implementó una estrategia de validación cruzada uno contra todos, la cual permite una generalización del modelo analizado. El rendimiento de la SVM fue medido por medio del cálculo de la exactitud como métrica de rendimiento.

2.7. Análisis Individual de Características

Se analiza la capacidad de las características para discriminar entre clases según el cálculo del área bajo la curva ROC, y sobre aquellas que lograron una mejor discriminante, se realizó un diagrama de cajas y bigotes, seguido de una prueba estadística U-Mann Whitney para validar que las muestras adquiridas en sanos y pacientes tienen distribuciones diferentes.

3. Resultados y discusión

Los mejores resultados obtenidos de la implementación de la SVM se pueden ver en la Tabla II. En la que se muestra el desempeño del clasificador para discriminar entre sujetos patológicos o disfágicos en distintas actividades deglutorias.

Tabla 2: Resultados obtenidos por atarea deglutoria en función de la exactitud y las características seleccionadas

Tarea deglutoria	Exactitud	Características Seleccionadas	aucROC
Agua 5	70,00 %	Potencia media	0,658
		Entropía de Shannon	0,560
		Energía coeficiente 3	0,504
		Energía coeficiente 7	0,562
		Energía de coeficiente 8	0,569
Agua 10	75,00 %	Entropía de Shannon	0,516
		Potencia media	0,615
		Energía de coeficiente de aproximación wavelet	0,549
		Energía del coeficiente de wavelet de detalle 3	0,727
			0,476
Agua 20	68,00 %	Ratio de frecuencia	0,431
		Potencia media	0,814
		Frecuencia mediana	0,582
		Frecuencia promedio	0,559
		Pico de frecuencia	0,612
		Entropía de Shannon	0,434
		Energía coeficiente 5	0,508
		Energía coeficiente 6	0,526
		Energía coeficiente 7	0,464
		Energía coeficiente 8	0,599
Yogur 5	73,33 %	Entropía de Shannon	0,503
		Potencia media	0,653
		Energía de coeficiente de aproximación wavelet	0,529
Yogur 10	70,00 %	Complejidad de Lempel-Ziv	0,627
		Potencia media	0,569
		Entropía de Shannon	0,616
		Energía coeficiente 4	0,489
		Energía coeficiente 7	0,58
Yogur 20	72,66 %	Energía coeficiente 8	0,529
		Entropía de Shannon	0,525
		Potencia media	0,702
		Energía coeficiente 3	0,673
		Energía coeficiente 5	0,505
Yogur 20	72,66 %	Energía coeficiente 5	0,651
		Energía coeficiente 8	0,518
		Entropía de Shannon	0,587

Según los resultados que se pueden ver en la Tabla II las características que presentan un mayor aporte de información a la SVM en tareas deglutorias evaluadas a partir de señales de AC fueron la potencia media y la energía del coeficiente de aproximación wavelet, características que fueron seleccionadas para la mayoría de las actividades deglutorias. El mejor clasificador se obtuvo para la actividad de agua (líquido claro) con un volumen de 10 mL, donde se logró una exactitud del 75%, posiblemente porque la actividad deglutoria de agua de 5 mL es un volumen pequeño y el volumen de 20 mL suele ser fragmentado por pacientes y algunos sujetos sanos según lo observado en el estudio, por lo que no es posible asociar las características del sonido segmentado al total del volumen suministrado. Para el líquido espeso se obtuvo un desempeño muy similar entre bolos, con el mejor rendimiento en el volumen de 5 mL posiblemente porque este volumen en líquidos espesos es el que permite el ingreso del bolo en una sola deglución, mientras que en los demás se suele fragmentar.

Así mismo, se obtienen los resultados del área bajo la curva de cada una de estas características para encontrar aquellas que tienen una mayor relevancia para distinguir entre clases según la tarea deglutoria, encontrando que las mejores, independientemente del bolo, son: la energía de aproximación de wavelet en el bolo de agua de 10mL, que sugiere que existe una diferencia morfológica entre las señales, la potencia media de agua de 5mL y 20mL, y la potencia media del bolo de yogur de 20mL, lo que podría estar ligado a que existe una disparidad en la amplitud de las señales de sujetos sanos y pacientes. Siendo también importante recalcar que estas dos características hacen parte de las combinaciones que arrojó el clasificador en su mejor rendimiento para el bolo de agua de 10 mL, además de que están presentes en la mayoría de los otros volúmenes.

Las Figuras 1, 2, 3 y 4 muestran los diagramas de cajas y bigotes que representan el comportamiento de las características anteriormente descritas como las mejores en aucROC, donde 0 representa los sujetos de control, y 1 los sujetos con la patología.

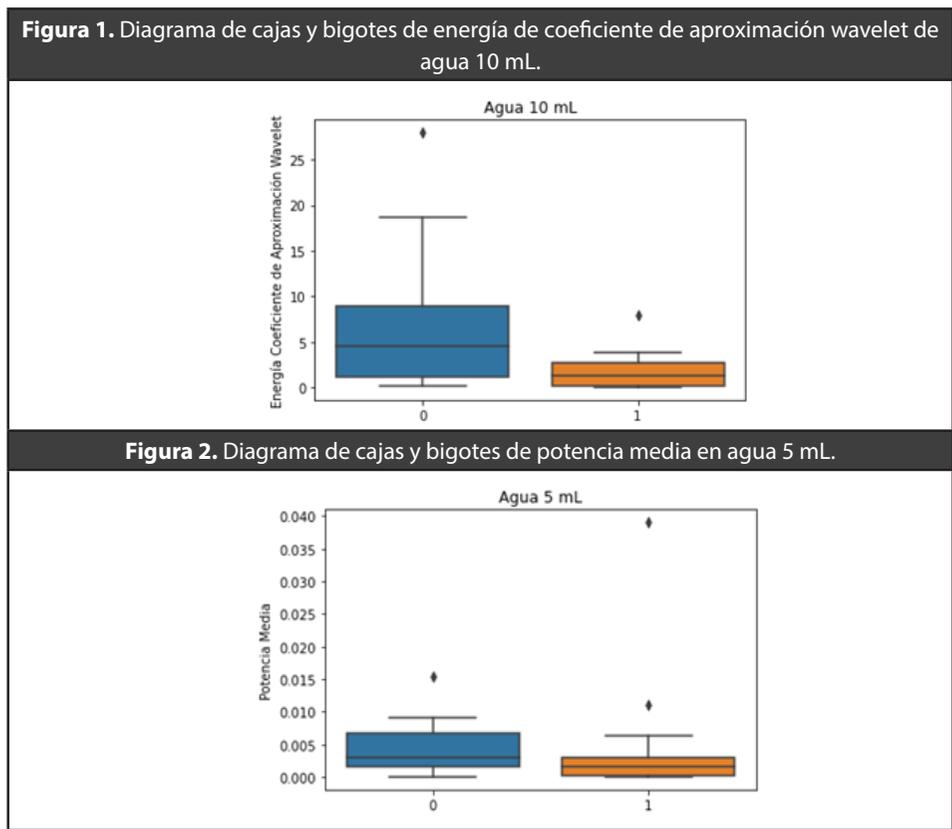


Figura 3. Diagrama de cajas y bigotes de potencia media en agua de 20 mL

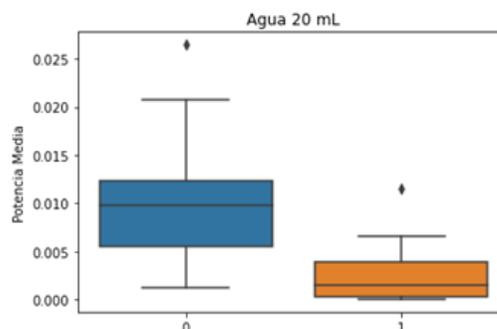
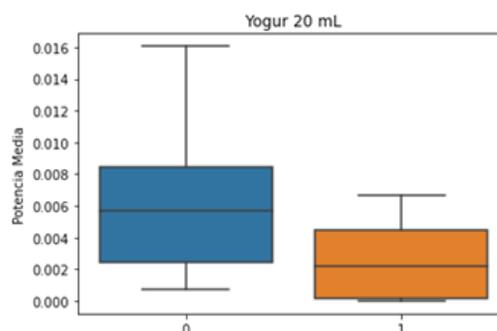


Figura 4. Diagrama de cajas y bigotes de potencia media en yogur de 20 mL.



Planteando la hipótesis nula de que no hay diferencias entre el valor de la característica medida entre sujetos sanos y patológicos, se realiza una prueba estadística de U-Mann-Whitney la cual tuvo como resultado un p-valor menor a 0,05, con lo cual se rechaza la hipótesis nula, en todas las características, excepto la potencia media para la tarea deglutoria de agua con un volumen de 5 mL. Si se rechaza la hipótesis nula, se concluye que hay diferencias en los valores de las características lo cual podría servir como herramienta para discriminar entre clases. El caso particular de 5 mL refuerza lo encontrado en los resultados de los clasificadores teniendo en cuenta que es un bolo muy pequeño para un líquido claro. Además, estos diagramas sugieren que existe una tendencia de mayor amplitud en las características de potencia media y energía de coeficiente de aproximación wavelet en los sujetos sanos en comparación de los sujetos con disfagia.

Una investigación previa presenta un enfoque basado en métodos de adquisición de señales no invasivas asociadas a la AC, como la unión de señales de acelerometría y audio captado por medio de micrófonos, cuyo fin fue diferenciar entre degluciones seguras y riesgosas utilizando regresión logística y árboles de decisión, obteniendo una clasificación acertada, pero sin realizar la diferencia entre sujetos sanos y patológicos [15].

También se han analizado señales de auscultación cervical propiamente con un estetoscopio electrónico, como en el presente trabajo, pero con la finalidad de discriminar momentos deglutorios en presencia de fuentes de ruido natural como la tos o la respiración [17].

Por otro lado, se encuentran trabajos similares al presente en que se evalúan exclusivamente señales de audio, con la diferencia de que la adquisición es con un

micrófono, tendencia que se evidencia en este tipo de estudios ya que estos dispositivos cuentan con la capacidad de detectar sonidos que la no linealidad de la escucha humana no permite [20]. En dicho estudio, se utiliza una máquina de soporte vectorial de base radial para clasificar entre sujetos con disfagia y de control con una exactitud similar a la conseguida por nosotros [19].

El presente estudio se diferencia a los anteriores en que las señales adquiridas son exclusivamente de audio empleando un estetoscopio electrónico. Lo anterior tiene como ventaja que se está utilizando la misma señal que es escuchada por los especialistas en la práctica clínica con lo cual se puede tener una mejor interpretabilidad de los resultados. Adicionalmente, se exploraron las características más relevantes para la discriminación entre sanos y pacientes mediante el uso de análisis estadístico y algoritmos de clasificación.

La principal limitación para el desarrollo de este trabajo fue el tamaño de la base de datos, y que no todos los sujetos disfágicos lograron completar todas las tareas deglutorias, lo cual, en algunos casos, generó un desbalance de clases. Para trabajos futuros es recomendable aumentar la cantidad de señales de AC utilizadas en ambos grupos poblacionales que permitan la generación de modelos más robustos.

4. Conclusiones

La implementación de clasificadores para la discriminación de sujetos sanos y disfágicos mediante el uso de señales de auscultación cervical es una técnica que debe ser más explorada, sin embargo, es prometedora, alcanzando exactitudes de hasta el 75 %. La auscultación cervical es una técnica usada en la práctica clínica, por lo que es de sumo interés crear un modelo capaz de apoyar en el diagnóstico de la disfagia sin incurrir en la subjetividad misma de la técnica, o en costo e invasividad de las demás herramientas diagnósticas utilizadas para este fin.

En este artículo se encontró que hay varias características en distinto dominios que logran extraer información valiosa de las señales de audio para ingresar a un clasificador. Como trabajo futuro se plantea el estudio de señales de audio junto a partir de las características con mejor desempeño encontradas en este estudio en conjunto con otro tipo de señales biológicas no invasivas para mejorar el rendimiento de los clasificadores.

5. Agradecimientos

Se agradece a la convocatoria del programa jóvenes investigadores e innovadores del ITM, y al proyecto 121077758144 financiado por el Ministerio de Ciencia, Tecnología e Innovación de Colombia y al proyecto PE20209 financiado por la convocatoria conjunta de proyecto de I+D+i en el marco de la agenda regional de I+D+i.

6. Conflictos de interés

Los autores declaran no tener conflictos de interés.

7. REFERENCIAS

- A. Sasegbon and S. Hamdy, "The anatomy and physiology of normal and abnormal swallowing in oropharyngeal dysphagia," *Neurogastroenterology and Motility*, vol. 29, no. 11, pp. 1–15, 2017, doi: 10.1111/nmo.13100.
- 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. 36, no. 2, pp. 259–269, 2021, doi: 10.1007/s00455-020-10124-z.
- C. Donohue, Y. Khalifa, S. Perera, E. Sejdić, and J. L. Coyle, "A Preliminary Investigation of Whether HRCA Signals Can Differentiate Between Swallows from Healthy People and Swallows from People with Neurodegenerative Diseases," *Dysphagia*, vol. 36, no. 4, pp. 635–643, 2021, doi: 10.1007/s00455-020-10177-0.
- C. Rebrion *et al.*, "High-Resolution Cervical Auscultation Signal Features Reflect Vertical and Horizontal Displacements of the Hyoid Bone during Swallowing," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 7, no. October 2018, 2019, doi: 10.1109/JTEHM.2018.2881468.
- D. G. Smithard, "Dysphagia : A Geriatric Giant ? The Normal Swallow," *iMedPub Journals*, vol. 2, no. 1:5, pp. 1–7, 2016, doi: 10.21767/2471-299X.1000.
- G. P. Andreu, "Del Envejecimiento," *Rev Cubana Invest Biomed*, vol. 22, no. 1, pp. 58–67, 2003.
- I. Cefac Brasil Bolzan *et al.*, "Contribution of the cervical auscultation in clinical assessment of the oropharyngeal dysphagia," *Revista CEFAC*, vol. 15, no. 2, pp. 455–465, 2013, [Online]. Available: <http://www.redalyc.org/articulo.oa?id=169326445023>
- I. Marchesan, "Deglución — Diagnóstico y Posibilidades Terapéuticas," *Espacio logopedico*, pp. 1–12, 2002, [Online]. Available: <https://s3.amazonaws.com/academia.edu.documents/34860642/deglucion.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1509723251&Signature=%2BUvBMiXkcg84moKyDMSSH1Or7LU%3D&response-content-disposition=inline%3Bfilename%3DDeglucion.pdf>
- J. M. Dudik, I. Jestrović, B. Luan, J. L. Coyle, and E. Sejdić, "A comparative analysis of swallowing accelerometry and sounds during saliva swallows," *BioMedical Engineering Online*, vol. 14, no. 1, pp. 1–15, 2015, doi: 10.1186/1475-925X-14-3.
- J. Pablo and R. Uribe, "Automatic swallowing analysis based on accelerometry and surface electromyography," 2021.
- L. J. Lazareck and Z. M. K. Moussavi, "Classification of normal and dysphagic swallows by acoustical means," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 12, pp. 2103–2112, 2004, doi: 10.1109/TBME.2004.836504.
- O. Ortega, A. Martín, and P. Clavé, "Diagnosis and Management of Oropharyngeal Dysphagia Among Older Persons, State of the Art," *Journal of the American Medical Directors Association*, vol. 18, no. 7, pp. 576–582, 2017, doi: 10.1016/j.jamda.2017.02.015.
- Q. He *et al.*, "The Association of High Resolution Cervical Auscultation Signal Features With Hyoid Bone Displacement During Swallowing," *IEEE transactions on neural systems and rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 27, no. 9, pp. 1810–1816, 2019, doi: 10.1109/TNSRE.2019.2935302.

- S. B. Leder, C. T. Sasaki, and M. I. Burrell, "Fiberoptic endoscopic evaluation of dysphagia to identify silent aspiration," *Dysphagia*, vol. 13, no. 1, pp. 19–21, 1998, doi: 10.1007/PL00009544.
- S. Roldan-Vasco, S. Restrepo-Agudelo, Y. Valencia-Martinez, and A. Orozco-Duque, "Automatic detection of oral and pharyngeal phases in swallowing using classification algorithms and multichannel EMG," *Journal of Electromyography and Kinesiology*, vol. 43, no. October, pp. 193–200, 2018, doi: 10.1016/j.jelekin.2018.10.004.
- T. Warnecke *et al.*, "Levodopa responsiveness of dysphagia in advanced Parkinson's disease and reliability testing of the FEES-Levodopa-test," *Parkinsonism and Related Disorders*, vol. 28, pp. 100–106, 2016, doi: 10.1016/j.parkreldis.2016.04.034.
- 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 Journal of Biomedical and Health Informatics*, vol. 25, no. 2, pp. 493–503, 2021, doi: 10.1109/JBHI.2020.3000057.
- Y. Khalifa, J. L. Coyle, and E. Sejdić, "Non-invasive identification of swallows via deep learning in high resolution cervical auscultation recordings," *Scientific Reports*, vol. 10, no. 1, pp. 1–13, 2020, doi: 10.1038/s41598-020-65492-1.
- Y. Sánchez-Cardona, A. Orozco-Duque, and S. Roldán-Vasco, "Characterization and classification of cervical auscultation signals acquired with stethoscope for automatic detection of swallowing sound," *Revista Mexicana de Ingeniería Biomedica*, vol. 39, no. 2, pp. 205–216, 2018, doi: 10.17488/RMIB.39.2.6.
- S. Miyagi, S. Sugiyama, K. Kozawa, S. Moritani, S. I. Sakamoto, and O. Sakai, "Classifying dysphagic swallowing sounds with support vector machines," *Healthcare (Switzerland)*, vol. 8, no. 2, pp. 1–12, 2020, doi: 10.3390/healthcare8020103.