

Extracción de características estadísticas en el dominio frecuencial de fonemas en pacientes de covid-19

Extraction of Statistical Features in the Frequency Domain of Phonemes in covid-19 Patients

Bryan Josue Lopez Zavala¹, Carlos Rodríguez Doñate¹, Luis Manuel Ledesma Carrillo¹,

¹Departamento de Estudios Multidisciplinarios, División de Ingenierías, Campus Irapuato-Salamanca, Universidad de Guanajuato
bj.lopezzavala@ugto.mx¹, lm.ledesma@ugto.mx¹

Resumen

La pandemia del covid-19 fue un antes y después en nuestra sociedad actual, pues complico la capacidad de atención, detección y tratamiento de la enfermedad en los sistemas de salud a nivel mundial.. Desde su aparición, numerosos métodos para la detección de esta enfermedad han sido propuestos e implementados como una medida de contingencia para sosegar la propagación de la enfermedad. Algunos de los métodos actuales como la prueba de Polymerease chain reaction (PCR) han demostrado una efectividad significativa y han sido de gran ayuda para el propósito mencionado. Actualmente, se ha continuado investigando nuevos métodos que resulten eficaces en la detección de la enfermedad y al mismo tiempo propongan una opción no invasiva para los pacientes. Numerosos métodos enfocados en técnicas computacionales para el análisis de diferentes factores biológicos han sido puestas a prueba con este propósito, de entre ellas, algunas relacionadas con el análisis de la voz son algunas de las mas prometedoras. En este trabajo se realiza un análisis frecuencia por medio de la transformada discreta de Fourier a un conjunto de muestras de voz separadas por fonemas, recabadas de pacientes con covid-19 para posteriormente realizar un análisis de sus características estadísticas intentando encontrar un patrón dentro de estas que pueda delimitar una medida para la detección de la enfermedad. Como resultado se obtuvo una mayor dispersión de los datos para las medidas de kurtosis y skewness, lo cual a primera instancia nos da indicios para pensar en estas como un posible factor de detección.

Palabras clave: covid-19, Fonemas, Análisis frecuencial, Extracción de Características.

Introducción

El covid-19 ha sido una de las enfermedades más mortales de los últimos tiempos, la cual llevo a que en marzo del 2020, debido al significativo avance de la enfermedad, y el elevado número de contagios, la organización mundial de la salud declarara un estado global de pandemia. La pandemia originalmente comenzó en Wuhan, China, a partir de donde comenzó a expandirse al resto del mundo. El método de esparcimiento de le enfermedad es por medio de fluidos respiratorios de una persona infectada producidos a través de tos y estornudos. A partir de este método, es muy fácil contaminar cualquier superficie, generando un posible foco de infección para otras personas. (WHO,2023)

La detección de esta enfermedad fue un factor clave para la disminución de casos de contagio. Durante la pandemia, doctores, junto con personal médico, así como investigadores pusieron sus esfuerzos en la realización de nuevos métodos para la detección del virus. El estándar actual para la detección del virus se basa en un análisis molecular conocido como PCR (Polymerease chain reaction), que ha sido ampliamente aceptado como uno de los métodos más eficaces para la detección del virus. Junto con lo anterior, numerosas técnicas que involucran la detección por otros medios como lo son los niveles de oxigeno en la sangre o las detecciones por medio de tomografías en los casos mas avanzados. Sin embargo, no han existido pruebas favorables que puedan acreditar a la mayoría de estos métodos como eficaces al momento de generar un patrón biológico que permita garantizar su efectividad en la detección, además de que juega también en contra el factor de invasión al paciente, así como los costos de la realización de los procedimientos. (Secretaria de Salud, 2022)



Debido a lo anterior, se han buscado nuevos métodos que involucren procedimientos menos invasivos con los pacientes y de costo menos elevado que puedan ofrecer a su vez un resultado fidedigno a los análisis. Uno de los focos principales en los que se centra la investigación de dichos métodos es en la implementación de sistemas computacionales tales como el Deep learning para el reconocimiento de ciertos marcadores biológicos. Uno de los marcadores biológicos mas prometedores para la detección de la enfermedad son las señales de voz; pues, no sería la primera vez en proponer este tipo de patrones en la detección de enfermedades, habiendo sido utilizado anteriormente para la detección de enfermedades como parkinson, enfermedades de la arteria coronaria y obstrucciones pulmonares crónicas. Esta última aplicación puede darnos una idea de la utilidad que tiene la detección por voz para la detección de ciertas anomalías relacionadas con las vías respiratorias; y ya que, el covid-19 afecta el sistema respiratorio, causando síntomas como la disminución en la presión de la faringe e inflamación de las vías nasales, ambos, factores que afectan a la vibración de la voz. (Ubaya COVID-19 Voice Dataset, 2023) (Subramanian et al., 2022) (Pah et al., 2022)

De lo anterior y en el mismo contexto en este proyecto se propone el análisis frecuencial de fonemas en pacientes diagnosticados con covid-19 como una posible medida para determinar la presencia de la enfermedad en pacientes bajo la sospecha del contagio del virus.

Metodología Propuesta

Se siguió la metodología presente en la figura 1, a continuación, se describe mas a detalle cada una de las etapas.

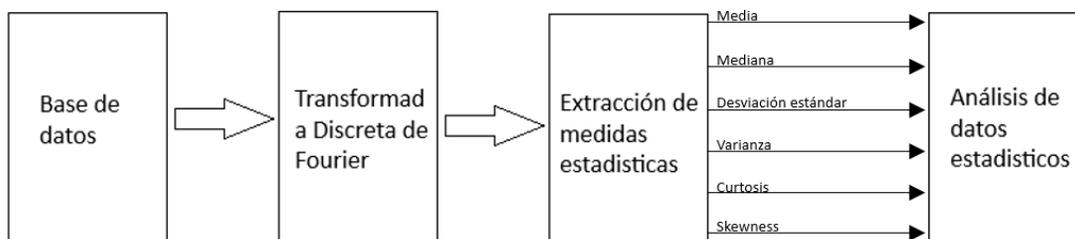


Figura 1. Diagrama de la metodología propuesta.
 Fuente: Elaboración propia

a) Base de datos

Se partió de la base de datos llamada “Ubaya covid-19 Voice Dataset”, una base de datos que compila archivos de voz en formato .wav de los fonemas /a/e/i/o/u/m/ de un total de 40 pacientes diagnosticados con covid-19 a lo largo de los meses de junio y julio del 2021, durante la segunda oleada de la variante Delta. [3]

Para el procesamiento del dataset se hizo uso del lenguaje Python, junto con sus librerías soundfile, numpy y pandas, para el procesamiento de señales, funciones matemáticas y estructuras de datos, respectivamente. El dataset fue organizado, de acuerdo al tipo de fonema, en dataframes dentro del proyecto de Python.

b) Transformada discreta de Fourier

La Transformada de Fourier es una herramienta de análisis de señales que transforma una señal representada en el dominio del tiempo discreto (n) al dominio de la frecuencia discreta ($k = 2\pi f$). De manera matemática se define como:

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-\frac{j2\pi kn}{N}} \tag{1}$$



Al ser aplicada a cada una de nuestras muestras de voz, se obtuvo como resultado las frecuencias fundamentales de cada audio. Los resultados de cada una de las transformaciones fueron almacenados de manera gráfica, en formato de imagen para futura referencia. La figura 2 muestra un ejemplo de las representaciones en tiempo y frecuencia de los fonemas.

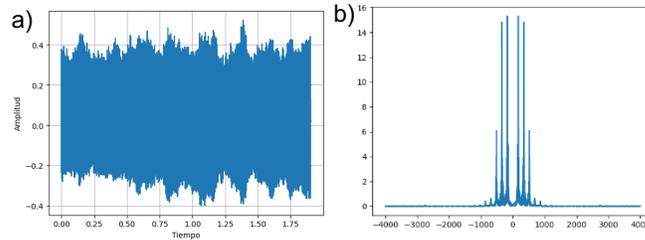


Figura 2. a) representación gráfica de un fonema en el dominio del tiempo. b) Representación gráfica del fonema en el dominio frecuencial.
 Fuente: Elaboración propia

Derivado de los datos de las frecuencias obtenidas para cada caso, se rescataron únicamente las frecuencias de los tres primeros armónicos de cada una de las muestras de voz. Dichos datos fueron almacenados de manera ordenada, por fonema y armónico, en estructuras de Excel.

c) Medidas estadísticas.

A partir de las frecuencias rescatadas de los armónicos, se extrajeron las siguientes características estadísticas: media, mediana, desviación estándar y varianza, así como los cuatro primeros momentos estadísticos a partir de los cuales se determinó la medida de curtosis y skewness de las distribuciones.

1. Media

La media aritmética, también llamada primer momento estadístico, es el parámetro de centralización más común y se define como:

$$\bar{X} = \sum_{k=0}^N \frac{X(K)}{N} \quad (2)$$

Donde N es el número de muestras y $X(k)$ es el espectro de la señal $x(n)$.

2. Mediana

La mediana de un conjunto de N datos ordenados de forma ascendente es el valor central del conjunto de datos y se define como:

$$M = X\left(\frac{N+1}{2}\right) \text{ para } N = \text{par} \\ M = \frac{1}{2} \left[X\left(\frac{N}{2}\right) + X\left(\frac{N+1}{2}\right) \right] \text{ para } N = \text{impar} \quad (3)$$

3. Varianza

La varianza es una medida de variabilidad que utiliza todos los datos de la muestra. La varianza está basada en la diferencia entre el valor de cada observación y la media.



$$s^2 = \frac{\sum_{k=0}^N (X(k) - \bar{X})^2}{N - 1} \quad (4)$$

4. Desviación estándar

Se define como la raíz cuadrada positiva de la varianza. Es utilizada para comprender cuan dispersos están los datos de un conjunto con respecto a la media.

$$\sigma = \sqrt{s^2} \quad (5)$$

5. Curtosis

La curtosis denota la deformación vertical de la curva de la distribución normal en el conjunto de datos. Se pueden identificar tres tipos de curtosis para un conjunto de datos y se define como:

$$\text{Kurt} = \frac{\frac{1}{k} \sum (X[k] - \bar{X})^4}{\left(\frac{1}{k} \sum (X[k] - \bar{X})^2 \right)^2} \quad (6)$$

- Una curtosis mayor al centro denota un apuntamiento a cero denotara un apuntamiento alto de la curva de distribución de frecuencias.
- Una curtosis igual a cero denotara un apuntamiento normal de la curva de distribución de frecuencias.
- Una curtosis menor a cero denotara un apuntamiento bajo de la curva de la distribución de frecuencias.

6. Skewness

La skewness o coeficiente de asimetría determina la medida en que los datos de una distribución normal se alejan de la media aritmética, que, vista de manera gráfica, representa el eje de simetría de la distribución, y se define como:

$$\text{skew} = \frac{\frac{1}{k} \sum (X[k] - \bar{X})^3}{\sqrt{\frac{1}{k} \sum (X[k] - \bar{X})^2}} \quad (7)$$

Dependiendo del coeficiente de asimetría, podemos clasificar las distribuciones en tres tipos:

- Si la asimetría es mayor a 0, hablamos de una asimetría positiva a la derecha.
- Si la asimetría es menor a 0, hablamos de una asimetría negativa hacia la izquierda
- Si la asimetría es igual a 0, hablamos de una gráfica simétrica.



Resultados

La tabla 1 muestra los resultados obtenidos de las medidas estadísticas para los armónicos de cada fonema. Analizando los valores de las medidas de Media y Mediana, podemos destacar que no existe una diferencia significativa entre los valores obtenidos para estas cuatro medidas, ya que, para cada armónico, los valores siempre se mantuvieron dentro de un mismo margen sin presentar alguna variación significativa. Por su parte, en la desviación estándar, podemos ver una variación mayor en los terceros armónicos de los fonemas /e/ y /m/, misma que se ve reflejada en la distribución normal, lo cual nos indica una mayor dispersión de los datos con respecto a la media. De la misma manera, en la varianza, podemos notar una mayor variación en los terceros armónicos de los fonemas anteriormente mencionados.

Tabla 1. Valores obtenidos de las medidas estadísticas.

FONEMA	ARMONICO	MEDIA (HZ)	MEDIANA (HZ)	DESV. EST. (HZ)	VARIANZA (HZ) ²
/a/	Primero	541.2414	304.5261	430.5042	188897.9
	Segundo	558.6244	449.9448	375.4561	143678.2
	Tercero	739.4034	700.8061	347.393	123002.7
/e/	Primero	448.3377	454.5985	175.2837	31273.0536
	Segundo	308.0326	258.0122	137.8884	19352.72
	Terceto	797.5735	571.1205	773.8768	609579.6
/i/	Primero	276.9993	279.5876	73.33937	5495.591
	Segundo	361.0498	388.5071	158.0529	25523.78
	Tercero	1281.836	595.1194	1110.203	1259345
/o/	Primero	442.7686	461.0861	108.8569	12065.29
	Segundo	367.2787	293.6481	201.3664	41285.67
	Tercero	646.6974	682.5215	264.5014	71233
/u/	Primero	323.1757	302.668	100.561	10332.36
	Segundo	385.839	374.6499	188.7149	36387.51
	Tercero	600.9279	621.1794	217.9049	48514.77
/m/	Primero	236.8053	213.9846	80.86565	6678.386
	Segundo	381.0657	371.0481	149.2355	22745.08
	Tercero	735.3508	555.6933	498.4658	253754.7

Fuente: Elaboración propia.

De la misma manera, para contrastar los resultados obtenidos a manera de poder generar una conclusión más concisa, se analizó la distribución normal de los datos, contrastándolos por armónico, obteniendo tres graficas de distribución normal. La figura 3 muestra las gráficas de las distribuciones normales obtenidas para cada armónico.



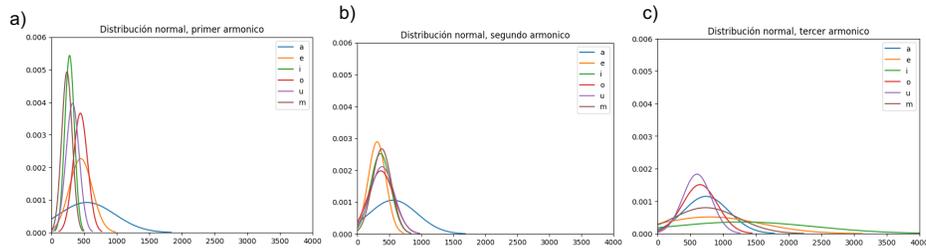


Figura 3. a) Distribución normal de los datos del primer armónico. b) Distribución normal de los datos del segundo armónico. c) Distribución normal de los datos del tercer armónico.
 Fuente: Elaboración propia

Así mismo, se utilizó como punto de análisis la comparativa de la kurtosis y skewness, donde nuevamente se hizo una comparativa entre los armónicos para identificar alguna variación significativa en alguno de los fonemas. La figura 4 muestra los resultados gráficos obtenidos para ambas medidas.

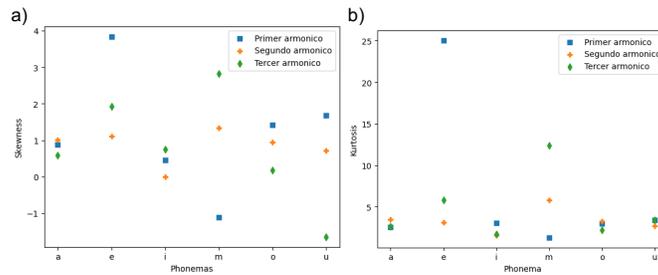


Figura 4. a) Skewness de los tres armónicos. b) Kurtosis de los tres armónicos.
 Fuente: Elaboración propia

Analizando las medidas de kurtosis y skewness, podemos notar las singularidades antes mencionadas, donde los fonemas /m/ y /e/ presentan una mayor variación entre sus armónicos; los resultados obtenidos mostraron un mejor desempeño en la característica de kurtosis para los fonemas /e/ y /m/, mientras que, para la característica de la Skewness, destacamos los fonemas /e/, /m/ y /u/ como aquellos en donde existe una mayor dispersión entre armónicos.

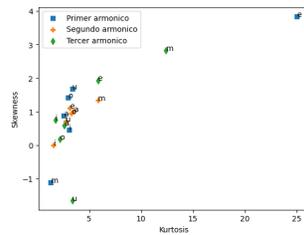


Figura 5. Grafica Skewness contra Kurtosis.
 Fuente: Elaboración propia

Adicionalmente, la comparativa realizada entre estas dos características, visible en la figura 5, para conocer su dispersión entre los fonemas, termino de proponer a estos fonemas como aquellos con mayor potencial para indicar como aquellos con mayor potencial para indicar una anomalía causada por el covid-19 los fonemas /e/ y /m/.

Conclusiones

En este trabajo se realizó un análisis de las características estadísticas de muestras de voz rescatadas de pacientes con covid-19. Durante el desarrollo de este, se pudo apreciar un cierto comportamiento inusual de los fonemas /e/ y /m/ con respecto a los demás. Dicho comportamiento, como se discutió durante el trabajo, podría llegar a incorporarse como una medida para la detección de la enfermedad mediante el análisis de la voz a manera de ofrecer una alternativa de prueba para los pacientes menos invasiva. A futuro, será necesario un análisis más exhaustivo, enfocado en la comparación con muestras de pacientes saludables, a fin de poder establecer si existe alguna relación entre las anomalías presentadas en los fonemas mencionados y la enfermedad.

Referencias

- Pah, N. D., Indrawati, V., & Kumar, D. (2022). Voice features of sustained Phoneme as COVID-19 biomarker. *IEEE Journal of Translational Engineering in Health and Medicine*, 10, 1-9. <https://doi.org/10.1109/jtehm.2022.3208057>
- Secretaria de Salud. (2022, 30 enero). *Comunicados técnicos diarios COVID 19*. gob.mx. Recuperado 2 de octubre de 2023, de <https://www.gob.mx/salud/documentos/comunicados-tecnicos-diarios-covid19>
- Subramanian, N., Elharrouss, O., Al-Maadeed, S., & Chowdhury, M. E. H. (2022). A review of deep learning-based detection methods for COVID-19. *Computers in Biology and Medicine*, 143, 105233. <https://doi.org/10.1016/j.combiomed.2022.105233>
- Ubaya COVID-19 Voice Dataset. (2023, 15 febrero). Kaggle. Recuperado 12 de junio de 2023, de <https://www.kaggle.com/datasets/nemueldanielpah/ubaya-covid-voice-dataset>
- WHO. (2023, 25 septiembre). *Coronavirus (covid-19) dashboard*. covid19.who.int. Recuperado 2 de octubre de 2023, de <https://covid19.who.int/>

