



**UNIVERSIDAD POLITÉCNICA SALESIANA**  
**SEDE QUITO**  
**CARRERA DE COMPUTACIÓN**

**ESTADO DEL ARTE PARA LA DETECCIÓN DE OLORES MEDIANTE  
EL SENSOR CYRANOSE 320 EN ECUADOR**

Trabajo de titulación previo a la obtención del  
Título de Ingeniero en Ciencias de la Computación

AUTOR: JONATHAN JOEL GUERRA CUJI  
TUTOR: MANUEL RAFAEL JAYA DUCHE

Quito - Ecuador  
2023

# **CERTIFICADO DE RESPONSABILIDAD Y AUTORÍA DEL TRABAJO DE TITULACIÓN**

Yo, Jonathan Joel Guerra Cuji con documento de identificación N.º 1720064169; manifiesto que:

Soy autor y responsable del presente trabajo; y, autorizo a que sin fines de lucro la Universidad Politécnica Salesiana pueda usar, difundir, reproducir o publicar de manera total o parcial el presente trabajo de titulación.

Quito, 22 de Agosto del año 2023

Atentamente,



---

Jonathan Joel Guerra Cuji  
1720064169

## **CERTIFICADO DE CESIÓN DE DERECHOS DE AUTOR DEL TRABAJO DE TITULACIÓN A LA UNIVERSIDAD POLITÉCNICA SALESIANA**

Yo, Jonathan Joel Guerra Cuji con documento de identificación No. 1720064169, expreso mi voluntad y por medio del presente documento cedo a la Universidad Politécnica Salesiana la titularidad sobre los derechos patrimoniales en virtud de que soy autor del Artículo Académico: “Estado del arte para la detección de olores mediante el Sensor Cyranose 320 en Ecuador”, el cual ha sido desarrollado para optar por el título de: Ingeniero en Ciencias de la Computación, en la Universidad Politécnica Salesiana, quedando la Universidad facultada para ejercer plenamente los derechos cedidos anteriormente.

En concordancia con lo manifestado, suscribo este documento en el momento que hago la entrega del trabajo final en formato digital a la Biblioteca de la Universidad Politécnica Salesiana.

Quito, 22 de Agosto del año 2023

Atentamente,



---

Jonathan Joel Guerra Cuji  
1720064169

## **CERTIFICADO DE DIRECCIÓN DEL TRABAJO DE TITULACIÓN**

Yo, Manuel Rafael Jaya Duche con documento de identificación N° 1710631035, docente de la Universidad Politécnica Salesiana, declaro que bajo mi tutoría fue desarrollado el trabajo de titulación: ESTADO DEL ARTE PARA LA DETECCIÓN DE OLORES MEDIANTE EL SENSOR CYRANOSE 320 EN ECUADOR, realizado por Jonathan Joel Guerra Cuji, con documento de identificación N.º 1720064169, obteniendo como resultado final el trabajo de titulación bajo la opción Artículo Académico que cumple con todos los requisitos determinados por la Universidad Politécnica Salesiana.

Quito, 22 de Agosto del año 2023

Atentamente,



---

Ing. Manuel Rafael Jaya Duche, MSc  
1710631035

# Estado del arte para la Detección de Olores mediante el Sensor Cyranose 320 en Ecuador

1<sup>st</sup> Jonathan Joel Guerra Cuji  
*Ingeniería en las Ciencias de la Computación*  
*Universidad Politécnica Salesiana*  
Quito, Ecuador  
jguerrac3@est.ups.edu.ec

2<sup>nd</sup> Manuel Rafael Jaya Duche  
*Ingeniería en las Ciencias de la Computación*  
*Universidad Politécnica Salesiana*  
Quito, Ecuador  
mjaya@ups.edu.ec

**Resumen**—Las narices electrónicas (e-noses) prometen ser un método de cribado a gran escala, ofreciendo ventajas como: i) bajo coste, ii) no invasivo, iii) análisis en tiempo real, entre otras. Motivo por el cual se ha propuesto un estado del arte con la finalidad de identificar los tipos de sensores y e-noses que se han empleado en el análisis de Compuestos Orgánicos Volátiles (COVs) de pacientes con cáncer, aplicando dos metodologías, mapeo sistemático (MS) para recopilar las revistas científicas con más relevancia desde el año 2015 hasta el 2023; y revisión de la literatura sistemática (RLS), en la cual mediante la construcción de una taxonomía se proporcionó para este documento: i) compuestos y concentraciones ii) métodos de preprocesamiento de datos, iii) métodos de compresión de datos, iv) métodos de selección y extracción de características, v) algoritmos de aprendizaje automático, vi) métodos de validación cruzada y vii) métricas de evaluación. Obteniendo resultados que impulsan esta tecnología en la práctica clínica, siendo de Leon-Martínez *et al* [1] el estudio con los mejores resultados cuantitativos, presentando una sensibilidad, especificidad, precisión y área bajo la curva (AUC) de 100%, 100%, 100% y 1, respectivamente. Sin embargo, se requieren artículos científicos que realicen validación externa, con la finalidad de que este dispositivo se posicione como una forma de cribado fiable dentro del campo de la medicina.

**Palabras Clave**—Compuestos Orgánicos Volátiles, Nariz electrónica, Sensores, Cromatografía de Gases-Espectrometría de Masas, Cáncer, Mapeo sistemático, Revisión de la literatura sistemática

**Abstract**—Electronic noses (e-noses) promise to be a large-scale screening method, offering advantages such as: i) low cost, ii) non-invasive, iii) real-time analysis, among others. Motivation for which a state of the art has been proposed in order to identify the types of sensors and e-noses that have been used in the analysis of Volatile Organic Compounds (VOCs) of cancer patients, applying two methodologies, systematic mapping (SM) to collect the scientific journals with more relevance from 2015 to 2023; and systematic literature review (SLR), in which by building a taxonomy was provided for this paper: (i) composites and concentrations (ii) data preprocessing methods, (iii) data compression methods, (iv) feature selection and extraction methods, (v) machine learning algorithms, (vi) cross-validation methods, and (vii) evaluation metrics. Obtaining results that boost this technology in clinical practice, de Leon-Martínez de Leon-Martínez said [1] the study with the best quantitative results, presenting sensitivity, specificity, precision and area under the curve (AUC) of 100%, 100%, 100% and 1, respectively. However, scientific articles are required to perform external validation in order to position this device as a reliable form of screening in the medical field.

**Keywords**—Volatile Organic Compounds, Electronic nose, Sen-

sors, Gas Chromatography-Mass Spectrometry, Cancer, Systematic mapping, Systematic literature review

## I. INTRODUCCIÓN

El análisis de los Compuestos Orgánicos Volátiles (COVs) endógenos, exógenos e incluso locales [2] es un proceso fundamental al momento de detectar y clasificar olores que pueden estar relacionadas con enfermedades, patologías e infestaciones [3]. Los COVs son químicos orgánicos [4] producidos a partir de procesos carcinogénicos que incluyen el i) estrés oxidativo, ii) citocromo p450, iii) peroxidación lipídica, iv) metabolismo lipídico, v) mutación o reordenación genética, vi) efecto Warburgetc, etc [5], [6], o pueden ser absorbidos del medio ambiente por medio del contacto con la piel, la ingestión e inhalación [7]; y excretados a través de la piel y fluidos biológicos como: i) aliento, ii) sangre, iii) saliva, iv) heces, v) orina, entre otros [4]. La presencia de estos compuestos puede ser un indicador de ciertos mecanismos fisiopatológicos que están ocurriendo en el cuerpo humano [7].

Es por este motivo que, en la actualidad se necesita una herramienta de cribado i) no invasiva, ii) de bajo coste, iii) segura, iv) portátil y v) rápida [2], [5], que ayude a diagnosticar tempranamente a sujetos con cualquier tipo de cáncer, a fin de, poder determinar la *huella química* de los pacientes mediante técnicas de análisis de datos y reconocimiento de patrones para clasificarlos de manera correcta y así evitar procedimientos invasivos y costosos. Una nariz electrónica (e-nose) es un dispositivo capaz de reaccionar a características estereoquímicas específicas de una o varias moléculas odorante presentes en el aire [8], a través de una matriz de sensores de gas. Se ha convertido en una tecnología novedosa y atractiva, la cual presenta más ventajas que las mencionadas previamente, frente a otros procedimientos que existen en la actualidad [7]. Su aplicación en la medicina ha resultado muy beneficiosa al momento de presentarse escenarios de cribado a gran escala, lo que acelera el proceso de detección y análisis [3].

El cáncer es la principal causa de muerte en todo el mundo [6], especialmente el cáncer de pulmón afecta tanto a hombres como mujeres y su alta mortalidad esta ligada a su diagnóstico tardío, cuando lamentablemente se presentan síntomas, como: i) tos, ii) disnea, iii) fatiga y iv) dolor en el tórax [9].

Los métodos diagnósticos i) invasivos, ii) de alto costo, iii) incómodos, iv) con un alto consumo de tiempo y v) con una elevada tasa de falsos positivos y falsos negativos en los resultados [9]–[12], son motivos suficientes para dejar a un lado estos métodos y aprovechar al máximo las capacidades de las narices electrónicas. Que como resultado de la aplicación de estas se tiene: i) Análisis de COV de manera más rápida y en tiempo real, ii) económica y sobre todo iii) no invasivo [2], [5], lo que incrementa la capacidad de detectar enfermedades tempranamente, con el fin de dar tratamiento. En un estudio publicado en 2018 por Madara Tirzite *et al* [13], quienes analizaron el aliento exhalado de pacientes con cáncer de pulmón mediante una e-nose Cyranose 320 y análisis de regresión logística (LRA), obtuvieron una sensibilidad de 95.8% y especificidad de 92.3% para fumadores y una sensibilidad de 96.2% y especificidad de 90.6% para no fumadores.

A nivel continental se han realizado distintos estudios de varios tipos de cáncer, tales como: i) el de mama, ii) vejiga, iii) cabeza y cuello, iv) ovario, v) colorrectal, vi) páncreas, vii) próstata, entre otros; utilizando sensores químicos, sistemas e-noses diseñados y desarrollados o mediante e-noses comerciales. Estos enfoques se basan comúnmente en el análisis de COV en muestras de orina, sangre y aliento exhalado. Simian Zhu *et al* [11], analizaron muestras de orina de pacientes con cáncer de vejiga y controles (personas sanas) utilizando matrices ópticas de sensores de fluorescencia de gases, métodos de análisis de datos y análisis discriminante por mínimos cuadrados parciales (PLS-DA), en donde obtuvieron una sensibilidad y especificidad de 84.21% y 87.80%, respectivamente con Leave-one out cross-validation y una sensibilidad y especificidad de 77.42% y 85.82%, respectivamente con Random Monte Carlo cross validation. Las e-noses tienen aplicaciones en diversos campos más allá de la medicina, como en la: i) agricultura, ii) comida, iii) bebida, iv) transporte aéreo, v) aplicaciones militares, vi) control ambiental, vii) industria cosmética [14] lo que las convierte en una herramienta de gran utilidad.

A nivel Nacional son escasos los artículos académicos que se enfocan en el uso de narices electrónicas, no obstante, los pocos que existe están comprobados científicamente, aquí algunos ejemplos. Electrónicos de la Universidad de las Fuerzas Armadas ESPE [15] desarrollaron un prototipo económico para la detección de explosivos mediante una nariz electrónica junto con un modelo PCA y como clasificador el 3-NN. Este dispositivo tuvo la capacidad de detectar e identificar distintas sustancias como: i) 2,4 dinitrotolueno ii) vinagre iii) etanol, obteniendo una sensibilidad del 80% y una especificidad del 90% por medio de Leave-one out cross-validation. Además de lo mencionado, un ejemplo más son los miembros del Departamento de Electricidad y Electrónica de la ESPE [16], quienes desarrollaron un modelo de clasificación de aprendizaje profundo (*LeNet-5*) preciso, rápido y ligero, que en conjunción con sensores químicos fue capaz de detectar y clasificar sustancias explosivas con un 100% de precisión.

Al momento de revisar la literatura científica sobre el tema en cuestión, se observó que aunque existían artículos

científicos en donde utilizaban la e-nose Cyranose 320, estos no eran suficientes para poder obtener información consistente. Es por este motivo que, en el presente artículo se ha propuesto un estado del arte mediante las metodologías de Mapeo Sistemático y Revisión de la Literatura Sistemática con la finalidad de realizar una búsqueda extensa sobre los tipos de sensores y narices electrónicas que se han empleado con la finalidad de analizar COV de pacientes con cáncer y controles, y realizar una comparación de los resultados cuantitativos obtenidos en cada estudio.

## II. MATERIALES Y MÉTODOS

Dada la naturaleza del artículo, no se utilizó ningún apartado tanto de hardware como de software, motivo por el cual, no se requieren Materiales.

Según [17], la metodología de Mapeo Sistemático (MS) se trata de una exploración exhaustiva de estudios primarios con la finalidad de determinar la evidencia que se encuentra a disposición sobre un tópico en concreto. Por otro lado, Revisión de la Literatura Sistemática (RLS) tiene por objetivo aplicar un procedimiento estricto a fin de identificar, analizar y combinar la evidencia disponible de estudios primarios [17]. Sin embargo, estas dos metodologías se complementan, razón por la cual, han sido implementadas en el presente estado del arte. Tres bloques fundamentales fueron definidos en el proceso de esta investigación [17].

### 1) Definición para la búsqueda:

- Definición de las preguntas de investigación, tanto de MS como de RLS
- Alcance de la revisión
- Criterios de inclusión y exclusión de MS y RLS
- Conducta de la búsqueda

### 2) Ejecución de la búsqueda:

- Selección de estudios primarios
- Definición de los criterios de análisis

### 3) Discusión de los resultados:

- Análisis de los resultados

## A. Bloque 1: Definición para la búsqueda

### 1) Preguntas de investigación para MS:

- PMS1: ¿En qué revistas científicas se publican trabajos que utilizan narices electrónicas y/o sensores para la detección de olores de cáncer?
- PMS2: ¿Cuántas investigaciones se han publicado sobre el tema de detección de olores de cáncer mediante el uso de narices electrónicas y/o sensores entre el 2015 a 2023?

Las revistas científicas consideradas para esta investigación fueron: *Science Direct*, *IOPscience*, *Pubmed*, *Scientific Reports*, *JAMA Network*, *Springer*, *Taylor Francis*, *IEEE Xplore*, debido a la necesidad de contar con una búsqueda avanzada.

### 2) Preguntas de investigación para RLS:

- PRLS1: ¿Cuáles son los sensores más utilizados en la detección de olores para cáncer?
- PRLS2: ¿Cuáles son las ventajas y desventajas de utilizar narices electrónicas y/o sensores en el campo de la salud para detectar el cáncer?

- PRLS3: ¿Cuáles son los Compuestos Orgánicos Volátiles (COVs) más comunes que se han identificado en el cáncer?
- PRLS4: ¿Qué compuestos y concentraciones químicas se han identificado en los diferentes tipos de cáncer?
- PRLS5: ¿Qué resultados cuantitativos han proporcionado las métricas utilizadas para la clasificación de olores en el cáncer?
- PRLS6: ¿Qué métodos se utilizan comúnmente en el preprocesamiento de los datos generados por las narices electrónicas y/o sensores?
- PRLS7: ¿Cuáles son las técnicas más utilizadas que permiten mejorar la calidad de los datos previo al reconocimiento de patrones?

3) *Alcance de la revisión:* PICOC es un enfoque de solución en el cual se organizan los datos de tal manera que se identifican las palabras clave o términos de búsqueda, que están estrechamente relacionados con las preguntas de investigación [18]. Los criterios PICOC fueron definidos como se puede visualizar en la Tabla I.

Tabla I: CRITERIOS CLAVE DE PICOC

PICOC	Descripción
Population (P): ¿Quién?	Pacientes con cáncer o posible diagnóstico de cáncer
Intervention (I): ¿Qué?, ¿Cómo?	Uso de narices electrónicas y/o diferentes tipos de sensores, para detectar olores relacionados con cáncer
Comparasion (C): ¿Con qué comparar?	No existe una comparación específica, por lo queda excluida
Outcomes (O): ¿Qué se busca conseguir/mejorar?	Métricas de evaluación
Context (C): ¿En qué tipo de organización y bajo qué circunstancias?	Revisar estudios existentes sobre la detección de olores de cáncer mediante narices electrónicas y/o diferentes tipos de sensores

Una vez que los criterios en PICOC hayan sido definidos, se procede a realizar una exploración en el buscador de Google Scholar, de las palabras clave que se encuentran en artículos afines al tema de estudio. Mediante este proceso se obtuvieron las terminologías adecuadas y con más relevancia a fin de armar la cadena de búsqueda ideal, como se puede observar en la Tabla II. Sin embargo, para llegar a obtener estos términos finales, el proceso tuvo que ser a base de prueba y error, en donde varias palabras clave se quedaron atrás a razón de los pocos resultados que se obtenían utilizando estos términos, motivo por el cual fueron descartados. Es importante señalar que, con el fin de obtener la cadena de búsqueda adecuada, se procedió a revisar 10 artículos científicos de manera aleatoria en los cuales si 6 contenían información relevante, quería decir

que la cadena era la adecuada, por lo cual podía ser utilizada en las Revistas Científicas.

Tabla II: PALABRAS CLAVE Y SINÓNIMOS

Palabras clave	Sinónimos
Sensors	Electronic-nose OR E-nose OR Artificial nose
Volatile Organic Compounds	VOC OR VOCs
Cáncer	Carcinoma

Fruto de este proceso, se obtuvo la siguiente cadena de búsqueda primaria: ("sensors" OR "electronic-nose" OR "e-nose" OR "artificial nose") AND ("volatile organic compounds" OR "VOC" OR "VOCs") AND ("cancer" OR "carcinoma")

4) *Criterios de inclusión y exclusión:* Con la finalidad de filtrar los estudios se aplicaron criterios de inclusión y exclusión para MS y RLS, que se pueden observar en la Tabla III.

Tabla III: CRITERIOS DE INCLUSIÓN Y EXCLUSIÓN DE MS Y RLS

Criterios de inclusión	Criterios de exclusión
CIMS1: Artículos a partir del año 2015 a 2023	CEMS1: Artículos sin DOI
CIMS2: Artículos publicados en inglés	CEMS2: Exclusión de libros
CIRLS1: MS3: Artículos que se centran en el diagnóstico de cáncer	CERLS1: Artículos con datos cualitativos
CIRLS2: Artículos que incluyen pacientes diagnosticados con cáncer y pacientes sanos como grupo de control.	CERLS2: Artículos que no contengan métricas, técnicas y datos cuantitativos
CIRLS3: Artículos enfocados en la detección de olores de cáncer mediante narices electrónicas y/o sensores	CERLS3: Artículos duplicados
	CERLS4: Artículos que mencionen sensores pero no contengan información relevante como: métricas, técnicas y datos cuantitativos

5) *Conducta de la búsqueda:* Varios filtros de revisión se aplicaron con la finalidad de seleccionar artículos primarios, como se detalla a continuación:

**Primer filtro (F1):** Dado que el cáncer es uno de los puntos fuertes en esta investigación se filtraron los artículos donde mencionaban las palabras clave: i) cáncer o ii) carcinoma en los títulos, haciendo alusión a los *CIRLS2* y *CIRLS3*. Además, se excluyeron los artículos que no contenían: i) métricas, ii) técnicas, iii) resultados cuantitativos y iv) métodos en el abstract, haciendo alusión a los *CERLS1* y *CERLS2*. Cabe mencionar que todas las palabras se introdujeron en el idioma inglés.

**Segundo filtro (F2):** Las e-noses/sensores, son el segundo punto fuerte de esta investigación, motivo por el cual se filtraron los artículos que contenían e-nose y sensor en su abstract, haciendo alusión al *CIRLS3*.

**Tercer filtro (F3):** Fueron eliminados los artículos duplicados, haciendo alusión al *CERLS3*.

**Cuarto filtro (F4):** Una vez los artículos hayan pasado por estos tres filtros, se procedió a descargar sus citas en las revistas científicas. Algunas de estas eran de acceso privado, motivo por el cual se requirió de la página web *schi-hub.se* con la finalidad de obtener artículos de paga de forma gratuita, a través del ingreso del DOI del estudio requerido en dicha página. Aunque esto es moralmente reprochable, sin embargo, se lo realizó con fines netamente académicos para que la presente investigación siguiera adelante.

**Quinto filtro (F5):** Finalmente, se eliminaron los artículos que no contenía información relevante para la investigación, haciendo alusión al *CERLS4*. Cabe mencionar que, se necesitó un poco más de información sobre los compuestos químicos y sus concentraciones en diferentes tipos de cáncer, por lo cual, se optó por el método de bolas de nieve, aumentando así la cantidad de estudios primarios finales.

Es importante destacar que:

- El *CIMS1* se realizó en cada una de las revistas científicas, seleccionando un rango específico de años.
- El *CIMS2* se realizó mediante las palabras clave ingresadas en la cadena de búsqueda, las cuales fueron en inglés.
- El *CEMS1* se realizó de manera implícita, ya que al momento de revisar los artículos, estos si contaban con DOI.
- El *CEMS2* se realizó a través de los otros criterios de inclusión y exclusión. Eliminando así todos los libros.

## B. Bloque 2: Ejecución de la búsqueda

1) *Selección de estudios primarios:* Por medio de la cadena de búsqueda, previamente definida, se llevo a cabo la indagación de artículos en cada una de las revistas científicas a través de los operadores lógicos *OR* y *AND*. Una vez obtenidos los resultados se procedió a descargar las citas, como se hizo mención previamente, y abrirlas con el software de trabajo *EndNote 20*, con la finalidad de, que la información de la cita se traslade directa y automáticamente a esta herramienta y así poder filtrar los artículos más relevantes como se mencionó en la sección anterior *Conducta de la búsqueda*. Los resultados se puede visualizar en la Tabla IV.

2) *Definición de los criterios de análisis:* Se definieron criterios de inclusión: i) *CIRLS1*, ii) *CIRLS2* y iii) *CIRLS3*, con la finalidad de analizar una serie de filtros a los cuales fueron sometidos todos los artículos. Para de esta manera obtener resultados relevantes, enfocados en el análisis de COVs de pacientes con cáncer mediante e-noses y/o sensores, aquí se encuentran:

- Año: Se señala el año de las publicaciones con el objetivo de descubrir si existió una evolución en el transcurso de ese tiempo.
- Tipo de publicación: Artículos científicos
- Número de estudios primarios: Se tomará en cuenta como estudios primarios, aquellos que cuenten con técnicas,

métodos o métricas que ayuden con el análisis de datos en los COVs de pacientes con cáncer.

Además, se plantea la siguiente taxonomía, observable en la Fig. 1.

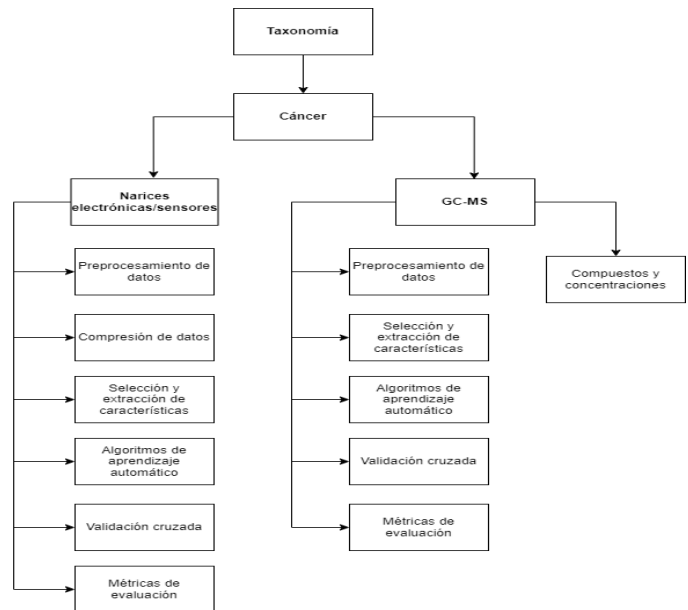


Fig. 1: Taxonomía

**Cáncer:** Esta es una de las enfermedades que presentan la mayor mortalidad al rededor del mundo, teniendo cifras en 2012 alrededor de 14.1 millones de nuevos casos y 8.1 millones de muertes [2]. Según un informe de la Organización Mundial de la Salud (OMS) en 2018, el continente Asiático reportó aproximadamente el 48% de los 18.1 millones de casos en todo el mundo [6]. Sin embargo, un diagnóstico temprano de la enfermedad ayudaría a reducir la tasa de mortalidad y consecuentemente aumentaría la tasa de supervivencia [2]. En esta investigación se pudo encontrar varios tipos de cáncer, siendo el de pulmón el que predomina sobre los otros tipos de cáncer, como se puede observar en la Fig. 2. Esto se debe a que las e-noses/sensores se emplean mayormente para analizar el VOC de los pacientes con este cáncer. Sin embargo, su aplicación también ha sido inmiscuida para los tipos de cáncer.

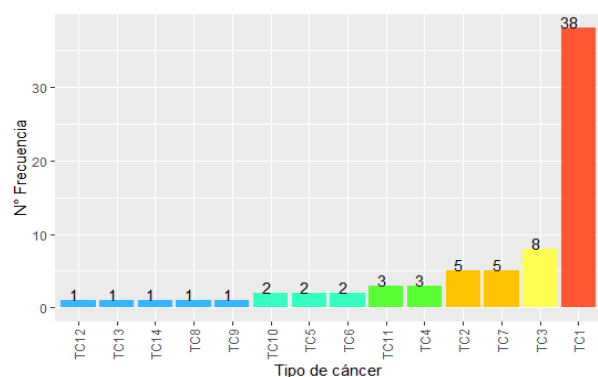


Fig. 2: Tipos de cáncer



Tabla IV: REVISTAS CIENTÍFICAS, CADENAS DE BÚSQUEDA Y RESULTADOS DE BÚSQUEDA

Revista Científica	Cadena de búsqueda	BDB	F1	F2	F3	F4	F5
IEEE Xplore	("All Metadata":sensors OR "All Metadata":electronic-nose OR "All Metadata":e-nose OR "All Metadata":artificial nose) AND ("All Metadata":volatile organic compounds OR "All Metadata":VOC OR "All Metadata":VOCs) AND ("All Metadata":cancer OR "All Metadata":carcinoma)	7	2	3	3	2	1
Springer	(sensor OR electronic-nose OR e-nose OR artificial nose) AND (volatile organic compounds OR VOC OR VOCs) AND (cancer OR carcinoma)	336	26	8	7	7	6
Science Direct	(sensor OR electronic-nose OR e-nose OR artificial nose) AND (volatile organic compounds OR VOC OR VOCs) AND (cancer OR carcinoma)	72	33	30	30	23	18
ACM	[[Abstract: volatile organic compounds] OR [Abstract: VOC] OR [Abstract: VOCs]] AND [[Abstract: "cancer"] OR [Abstract: carcinoma]] AND [[Abstract: sensors] OR [Abstract: "electronic-nose"] OR [Abstract: "e-nose"] OR [Abstract: artificial nose]] AND [E-Publication Date: (01/01/2015 TO 12/31/2023)]	36	8	1	1	1	0
Pubmed	((sensor[Title/Abstract]) OR (electronic-nose[Title/Abstract]) OR (e-nose[Title/Abstract]) OR (artificial nose[Title/Abstract])) AND ((volatile organic compounds [Title/Abstract]) OR (VOC[Title/Abstract]) (VOCs)) AND ((cancer[Title/Abstract]) OR (carcinoma[Title/Abstract]))	76	32	29	21	19	17
Hindawi	("sensor" OR "electronic-nose" OR "e-nose" OR "artificial nose") AND ("volatile organic compounds" OR "VOC" OR "VOCs") AND ("cancer" OR "carcinoma")	32	4	2	2	2	0
Taylor & Francis Online	[[All: sensor] OR [All: electronic-nose] OR [All: e-nose] OR [All: artificial nose]] AND [All: volatile organic] AND [[All: compounds] OR [All: voc] OR [All: vocs]] AND [[All: cancer] OR [All: carcinoma]] AND [Article Type: Article] AND [Publication Date: (01/01/2015 TO 12/31/2023)]	145	9	3	3	3	3
IOPscience	(sensor OR electronic-nose OR e-nose OR artificial nose) AND (volatile organic compounds OR VOC OR VOCs) AND (cancer OR carcinoma)	220	22	7	7	7	7
Scientific Reports	(sensor OR electronic-nose OR e-nose OR artificial nose) AND (volatile organic compounds OR VOC OR VOCs) AND (cancer OR carcinoma)	54	11	4	4	4	3
JAMA Network Open	(sensor OR electronic-nose OR e-nose OR artificial nose) AND (volatile organic compounds OR VOC OR VOCs) AND (cancer OR carcinoma)	6	2	1	1	1	1
Método de bolas de nieve	Science Direct (5 artículos); IOPscience(1 artículo); Pubmed (2 artículo); Springer (2 artículos); Wiley Online Library (2 artículo); British Journal of Cancer (1 artículo); BMJ Journals (1 artículo)	-	-	-	-	-	14
<b>Total</b>		<b>984</b>	<b>149</b>	<b>88</b>	<b>79</b>	<b>69</b>	<b>70</b>

TC1:Cáncer de pulmón; TC2:Cáncer de mama; TC3:Cáncer colorectal; TC4:Carcinoma de células escamosas de cabeza y cuello; TC5:Cáncer de ovario; TC6:Cáncer de próstata; TC7:Cáncer de vejiga; TC8:Cáncer de colon; TC9:Cáncer de cuello de útero; TC10:Adenocarcinoma pancreático; TC11:Cáncer gástrico; TC12:Carcinoma bronquial de células escamosas; TC13:Carcinoma laríngeo de células escamosas; TC14:Carcinoma de células renales claras

Continuando con la línea del cáncer, las principales causas de que la ocurrencia de esta misma se deben a sus fuentes primarias de COV, endógenas y exógenas. La fuente endógena es producto de la alteración de COV a través de procesos carcinogénicos, los más importantes se mencionan en la Tabla V. Por otro lado la fuente exógena tienen su origen por medio de factores externos, que se pueden observar en la Tabla VI.

Tabla V: FUENTE ENDÓGENA

Causas de COVs endógenos	Referencias
Estrés oxidativo	[6] [9] [12] [10] [19] [20] [21] [1] [22] [12] [23] [24] [4] [25] [26] [27] [28] [29] [30] [31] [32] [33] [34]
Peroxidación lipídica	[11] [22] [35] [27] [36] [29] [30] [31] [37] [33] [34]
Citocromo p450	[9] [5] [10] [23] [24] [26] [27] [29] [30] [34]
Pérdida del equilibrio de las especies reactivas del oxígeno	[5] [10] [20] [11] [1]
Actividad inflamatoria excesiva	[5] [1] [25]
Metabolismo alterado	[19] [38] [26]
Cambios genéticos	[4] [35] [28]

Tabla VI: FUENTE EXÓGENA

Causas de COVs exógenas	Referencias
Consumo de tabaco	[19] [20]
Consumo de alcohol	[20]
Consumo de alimentos	[36]
Exposición a la radiación	[20]
Inhalación de sustancias tóxicas	[20] [36]

Narices electrónicas/sensores: Esta investigación presenta tres enfoques que están estrechamente relacionados, pero a la vez difieren en algunos aspectos. En primer lugar, se tiene los sensores, los cuales pueden ser utilizados tanto de forma individual como de forma colectiva, es decir, formando un conjunto de sensores. En segundo lugar, se tiene los sistemas e-nose diseñados y desarrollados por un grupo de investigadores. Y por último, se tiene los e-noses comerciales, los cuales son productos que se encuentran disponibles en el mercado y listos para su uso. Como se puede visualizar en la Fig. 3, las e-noses comerciales fueron las más utilizadas, debido a que los investigadores se focalizaron en el análisis de COV de pacientes con cáncer a través de esta herramienta. Despreocupándose de el diseño y desarrollo de la misma. Sin embargo, el 27% de estudios si se dedicaron a desarrollar un nuevo sistema e-nose, que pueda ser comparable con los que se encuentran en el mercado. Y el 23% de estudios solamente utilizaron diferentes tipos sensores.

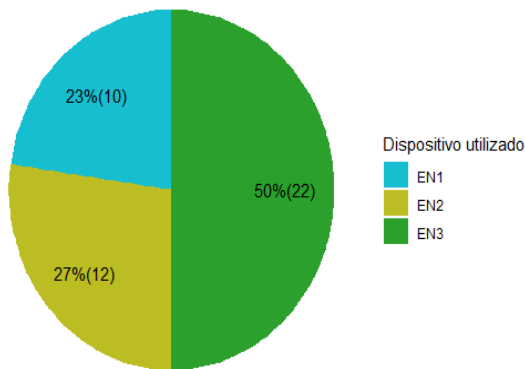


Fig. 3: Enfoque utilizado

EN1:Solamente sensores; EN2:Sistema e-nose desarrollado; EN3:E-nose comercial

Además, como se observa en la Fig. 4, las e-noses comerciales más utilizadas fueron i) Aeonose y ii) Cyranose 320, esto se debe a que son las más conocidas dentro del mercado y los investigadores las utilizaron con más frecuencia que las otras. Sin embargo, esto no sugiere que las demás tengan menores cualidades que las mencionadas.

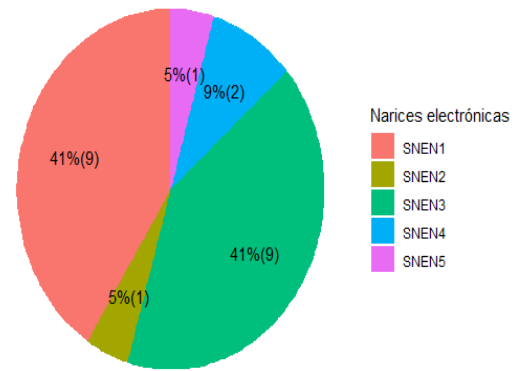


Fig. 4: Narices electrónicas utilizadas  
SNEN1: Cyranose 320; SNEN2: BIONOTE; SNEN3: Aeonose; SNEN4: PEN3; SNEN5: MCWbioG

Pero ¿Qué es una e-nose? Una e-nose es un dispositivo que incorpora un conjunto de sensores químicos de alta sensibilidad [39] y especificidad parcial, junto con un sistema de reconocimiento de patrones, con un potencial para identificar olores simples y complejos [12]. Es un símil de la nariz humana, salvando las distancias, en donde su respuesta a un determinado olor produce una llamada "huella olfativa", que es analizada y reconocida [8]. Funciona bajo el principio de sensibilidad cruzada, esto quiere decir que los sensores interactúan con una variedad de COVs, en lugar de uno en específico [5], [8], produciendo así en los sensores un cambio físico que es transformado en una señal eléctrica, la cual da lugar a la ya mencionada "huella olfativa" [12], obteniendo así datos multidimensionales para su posterior análisis estadístico [8].

A estos datos multidimensionales se los analiza con métodos de: i) preprocesamiento de datos, ii) compresión de datos y iii) selección y extracción de características. Por otra parte, para el reconocimiento de patrones se utilizan: i) Algoritmos de aprendizaje automático capaces de proporcionar resultados cuantitativos, en donde su rendimiento es evaluado a través de métodos de validación cruzada. Para finalmente, utilizar métricas de evaluación a fin de comparar los resultados obtenidos en cada estudio.

Con el objetivo de, determinar exitosamente una enfermedad, en este caso el cáncer, los datos originales que previamente fueron medidos por los sensores deben ser analizados y preprocesados, aplicando diferentes métodos de preprocesamiento de datos [40]. Según [1], el rendimiento en la etapa de reconocimiento de patrones se verá influenciado por la selección de él o los métodos de preprocesamiento de datos. Estos métodos se presentan en la Tabla VII.

Tabla VII: MÉTODOS DE PREPROCESAMIENTO DE DATOS

Métodos	Referencias
Normalización	[11] [1] [22] [41] [42] [43] [4] [27] [8]
Estandarización	[40] [44] [28]
Procesamientos de la línea de base	[45] [40] [46] [26] [8]
Eliminación de la retrodispersión	[11]
De-noising	[11] [4]
Serialización	[11]
Autoescalado	[5] [1] [22] [8]
Filtro de la ventana-media	[41]
Filtrado Savitzky-Golay	[4] [8]

Una vez realizado el procesamiento de los datos, existes dos posibilidades: i) la compresión de datos y ii) la selección y extracción de características, las cuales se explican a continuación. El único método de compresión de datos encontrado en esta investigación fue *Tucker3*, que se puede observar en la Tabla VIII, en donde los datos obtenidos por la medición de los sensores de la e-nose se descomponen en tensores [47], generando así un vector unidimensional de tamaño limitado por participante [28], que tiene por objetivo evitar i) correlaciones espurias [38], ii) datos redundantes y iii) el ruido [44]. Un aspecto a destacar es que, en los estudios donde aplicaron este método de compresión de datos, también utilizaron Artificial Neural Network (ANN). Siendo los datos comprimidos la entrada para dicho clasificador.

Tabla VIII: MÉTODOS DE COMPRESIÓN DE DATOS

Métodos	Referencias
Tucker3	[19] [38] [2] [21] [44] [47] [28]

Por otra parte, los métodos de selección y extracción de características permiten, como su nombre mismo lo dice, seleccionar las características más discriminatorias de los datos multidimensionales [9], esto quiere decir que se reduce la complejidad de los datos con la finalidad de evitar i) redundancia y ii) ajuste excesivo al momento de la clasificación [45]. Y como resultado de este proceso se obtiene un conjunto de datos capaz de discriminar entre grupos previamente definidos [22]. Los métodos de selección y extracción de características se presentan en la Tabla IX.

Tabla IX: MÉTODOS DE SELECCIÓN Y EXTRACCIÓN DE CARACTERÍSTICAS

Métodos	Referencias
Principal Component Analysis (PCA)	[9] [5] [14] [48] [1] [22] [23] [40] [43] [46] [49] [8]
Canonical Discriminant Analysis (CDA)	[1] [4]
Linear Discriminant Analysis (LDA)	[50] [40]
Kernel Principal Component Analysis (KPCA)	[45]
Spare Group Lasso (SGL)	[41]
Laplacian Eigenmap (LE)	[40]
Local Linear Embedding (LLE)	[40]
t-Stochastic Neighbor Embedding (tSNE)	[40]

Después que el conjunto de datos haya pasado por esta serie de procesos, uno o varios algoritmos de Aprendizaje automático, serán entrenados a partir de un conjunto de datos de entrada [14] y tendrán la capacidad de reconocer patrones y discriminar entre grupos predefinidos. Cabe mencionar que, existen algoritmos supervisados y no supervisados, donde el supervisado es entrenado con ejemplos semejantes y su clase correspondiente; mientras que el no supervisado, tiene la capacidad de aprender a distinguir las diferentes clases sin la inclusión de la misma [14]. Como se puede visualizar en la Fig. 5, se encontró una gran variedad de algoritmos, en donde ANN es el más utilizado, superando casi el doble al segundo, el cual es Support Vector Machine (SVM).

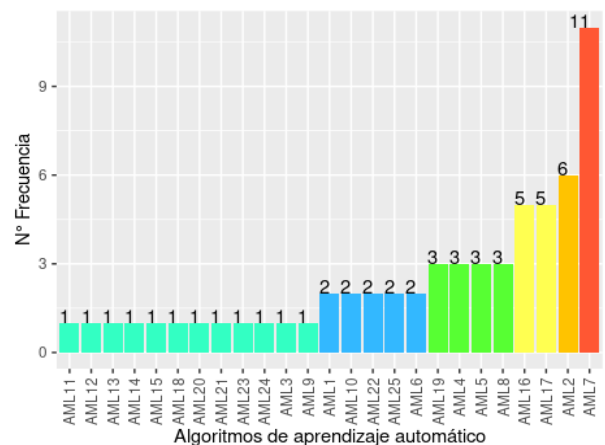


Fig. 5: Algoritmos de aprendizaje automático AML1:Linear Discriminant Analysis (LDA); AML2:Support Vector Machine (SVM); AML3:Multilayer Perceptron (MLP); AML4:Random forest; AML5:Logistic regression; AML6:Convolutional Neural Network (CNN); AML7:Artificial Neural Network (ANN); AML8:K-nearest neighbour algorithm (K-NN);

AML9:Multivariable logistic regression analysis; AML10: Naive bayesian; AML11:Classification and regression tree (CART) analysis; AML12:Probabilistic Neural Network (PNN); AML13:XGBoost; AML14:AdaBoost; AML15:LogitBoost; AML16:Discriminant Factor Analysis (DFA); AML17:Partial least squares Discriminant Analysis (PLS-DA); AML18:Canonical Discriminant Analysis (CDA); AML19:Discriminant Analysis (DA) classification; AML20:Fuzzy k-NN; AML21:Bootstrap forest; AML22:Canonical Analysis of Principal Coordinates (CAP); AML23:Principal Component Analysis (PCA); AML24:Receiver Operating Characteristic (ROC) curves AML25:Discriminant Function Analysis (DFA)

Los métodos de validación cruzada básicamente permiten dividir el conjunto de datos en dos subconjuntos, entrenamiento y prueba, con la finalidad de entrenar un algoritmo de aprendizaje automático con un subconjunto y probar su rendimiento con el otro [51]. Existen diferentes métodos de validación cruzada los cuales trabajan bajo la misma lógica mencionada. En la Tabla X, se pueden observar los métodos identificados.

Tabla X: MÉTODOS DE VALIDACIÓN CRUZADA

Métodos	Referencias
Cross-validation	[9] [52] [53] [8] [34]
Leave-one out cross-validation	[5] [54] [19] [55] [11] [1] [22] [24] [41] [42] [4] [56] [49] [27] [29] [31] [32]
K-fold cross-validation	[14] [45] [23] [41] [40] [57] [51] [43] [46] [26] [58]
Root Mean Square Error (in Cross-Validation) (RMSECV)	[54]
Leave-10%-out cross-validation	[38] [21] [39] [44] [47] [28]
Double cross-validation	[2]
20% Random Monte Carlo cross-validation	[11]
Leave-two-out cross-validation	[58]

Finalmente, las métricas de evaluación son utilizadas con el fin de determinar el rendimiento obtenido por el o los algoritmos de aprendizaje automático con los datos de prueba. Los resultados cuantitativos extraídos a partir de las métricas, dependerá en gran medida de la dimensión del conjunto de datos [59]. Cabe mencionar que mientras más alto sea el valor, mejor será el rendimiento del algoritmo. Pero ¿Que es alto y que es bajo? En todas las métricas, salvo Matthews correlation coefficient (MCC) donde su rango es de -1 a 1, existe un rango entre 0 y 1, siendo esta la base para poder determinar el concepto de alto y bajo. Como se puede visualizar en la Fig. 6, las métricas más utilizadas son sensibilidad y especificidad.

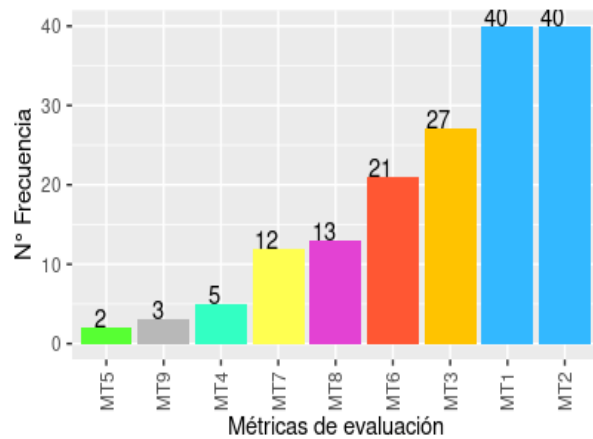


Fig. 6: Métricas de evaluación

MT1:Sensibilidad; MT2:Especificidad; MT3:Precisión; MT4:Tasa de clasificación correcta; MT5:Tasa de error de clasificación; MT6:Área bajo la curva (AUC); MT7:Valor predictivo positivo (VPP); MT8:Valor predictivo negativo (VPN); MT9:Coeficiente de correlación de Matthews (CCM)

Dada la cantidad de resultados cuantitativos extraídos de cada uno de los estudios, se pudo identificar uno especial [1], el cual cuenta con el mejor rendimiento, obteniendo sensibilidad, especificidad, precisión y AUC de 100%, 100%, 100% y 1, al momento de discriminar entre pacientes con cáncer de mama y controles, mediante la e-nose Cyranose 320 y CDA. Sin embargo, existen más resultados cuantitativos de diferentes estudios, los cuales pueden ser observados en el Material Suplementario, en la Tabla S1, presentando el tipo de cáncer a identificar, el dispositivo utilizado (sensores/e-nose), factores varios que son tomados en cuenta al momento del reconocimiento de patrones, el o los algoritmos de aprendizaje automático utilizados y finalmente las métricas de evaluación que fueron empleadas. Además, se incluyó la referencia a cada uno de los estudios para más información al respecto.

GC-MS: Debido a la naturaleza de esta investigación, adicionalmente se pudo identificar otra forma de análisis de COVs. Se trata de Cromatografía de Gases acoplada a la Espectrometría de Masas (GC-MS), cuyo objetivo es la identificación y cuantificación de "biomarcadores" específicos de una enfermedad [20]. Este es considerado el método de referencia para la detección de COVs, proporcionando: i) alta sensibilidad, ii) especificidad y iii) precisión [51]. Sin embargo, esta técnica presenta numerosas desventajas en comparación de las ventajas mencionadas, algunas de estas se muestran en la Tabla XI.

Tabla XI: DESVENTAJAS DE GC-MS

Deventajas	Referencias
Alto costo	[6] [9] [14] [60] [45] [61] [62] [20] [2] [12] [41] [40] [42] [39] [51] [25] [26] [47] [8] [30] [34]
Conlleva mucho tiempo	[6] [9] [14] [20] [2] [12] [39] [51] [43] [25] [26] [47] [49] [30] [34]
Requiere el análisis de expertos	[20] [2] [26] [47] [8]
Requieren un operador formado para utilizar el dispositivo	[9] [14] [45] [62] [12] [41] [40] [42] [39] [51] [43]
Requiere la preconcentración de la muestra	[6] [40] [34]
Es voluminoso	[6] [42] [30]
No es portátil	[60] [45] [61]
Es complejo	[26] [49]
No puede utilizarse en tiempo real	[4]

La cantidad de estudios focalizados en el análisis de COVs mediante GC-MS fue baja, aunque si identificaban compuestos, las concentraciones pertenecientes a cada uno de estos no eran mostradas, motivo por el cual se utilizó el método de bolas de nieve, con la finalidad de suplir este problema. A través de la Fig. 7, se puede observar la cantidad de estudios que utilizaron: i) e-noses/sensores, ii) GC-MS y iii) e-noses/sensores y GC-MS. Se acuño como "Procedimiento de análisis de COVs", a la forma en que estos fueron analizados.

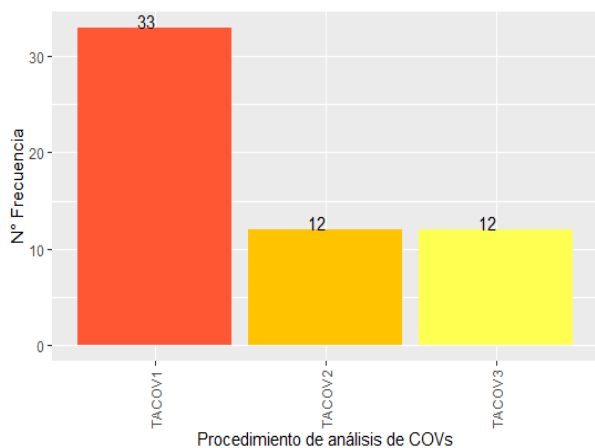


Fig. 7: Procedimiento de análisis de COVs  
TACOV1:E-noses/sensores; TACOV2:GC-MS;  
TACOV3:E-noses/sensores y GC-MS

Por otro lado, mediante los datos obtenidos a través de GC-MS, es posible analizarlos y reconocer patrones, tal y como se mencionó anteriormente con las e-noses/sensores. Existen: i) métodos preprocesamiento de datos, ii) métodos selección y extracción de características, iii) algoritmos de aprendizaje automático, iv) métodos de validación cruzada y v) métricas de

evaluación. En la Tabla XII, se pueden observar los métodos de preprocesamiento de datos que fueron identificados. Solamente los métodos i) normalización y ii) estandarización están presentes cuando se utilizan e-noses/sensores y GC-MS.

Tabla XII: MÉTODOS DE PREPROCESAMIENTO DE DATOS

Métodos	Referencias
Normalización	[35] [63] [36] [64]
Filtrado de cultivos	[35]
Detección de picos	[35]
Creador de cromatogramas	[35]
Deconvolución	[35]
Alineación	[35]
Estandarización	[35]

Los métodos de selección y extracción de características también fueron identificados y se pueden apreciar en la Tabla XIII.

Tabla XIII: MÉTODOS DE SELECCIÓN Y EXTRACCIÓN DE CARACTERÍSTICAS

Métodos	Referencias
Mann-Whitney U test	[35]
Principal component analysis (PCA)	[63]
Discriminant analysis (DA)	[36]
Discriminant function analysis (DFA)	[64]

Los algoritmos de aprendizaje de maquina que fueron recopilados, se centraban en el reconocimiento de patrones de grupos predefinidos, en este caso, paciente vs control, más no, para la discriminación de compuestos (COVs). Estos algoritmos pueden visualizarse en la Tabla XIV.

Tabla XIV: ALGORITMOS DE APRENDIZAJE AUTOMÁTICO

Algoritmos	Referencias
Principal Components Analysis (PCA)	[52]
Convolutional Neural Network (CNN)	[52]
Partial least squares Discriminant Analysis (PLS-DA)	[24] [35]
Random forest	[51] [43]
Artificial Neural Network (ANN)	[51] [43]
Logistic regression	[43]
Support Vector Machine (SVM)	[43]
Standard binary logistic regression	[30]

Al igual que los i) métodos de preprocesamiento de datos, ii) métodos de selección y extracción de características y iii) algoritmos de aprendizaje automático, se encontraron varios métodos de validación cruzada, los cuales se mencionan en la Tabla XV.

Tabla XV: MÉTODOS DE VALIDACIÓN CRUZADA

Métodos	Referencias
Cross-validation	[52]
K-fold cross-validation	[51] [43] [35]
Leave-one-out cross-validation	[30]

Y finalmente las métricas de evaluación identificadas se encuentran para su visualización en la Fig. 8. En donde la sensibilidad y especificidad fueron las más utilizadas.

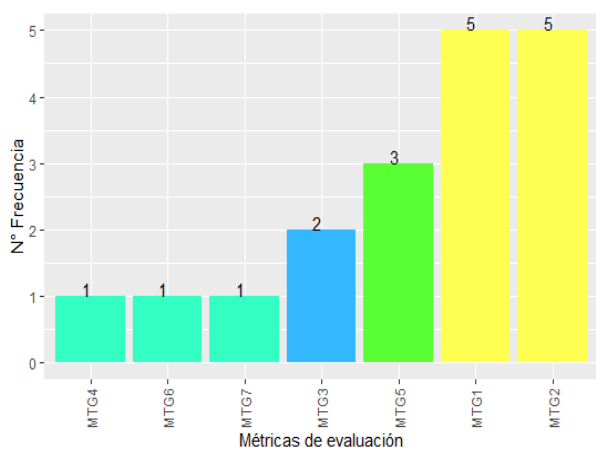


Fig. 8: Métricas de evaluación

MTG1:Sensibilidad; MTG2:Especificidad; MTG3:Precisión; MTG4:Tasa de clasificación correcta; MTG5:Área bajo la curva (AUC); MTG6:Valor predictivo positivo (VPP); MTG7:Valor predictivo negativo (VPN)

En el Material Suplementario, en la Tabla S2, se muestran los resultados cuantitativos que se obtuvieron de los estudios que utilizaron GC-MS para el análisis de COVs. En donde se observa el tipo de cáncer a identificar, factores varios que son tomados en cuenta al momento del reconocimiento de patrones, el o los algoritmos de aprendizaje automático utilizados y las métricas de evaluación empleadas. Destacando a [51], ya que fueron capaces de discriminar a paciente con cáncer colorrectal y controles, con una sensibilidad, especificidad y AUC de 86%, 81% y 0.93, respectivamente, utilizando Random forest.

Como se mencionó anteriormente, la identificación y cuantificación de COVs en la muestra de un paciente se hace posible mediante GC-MS, motivo por el cual se utilizó el método de bolas de nieve con la finalidad de encontrar estos apartados. A continuación, se mostrarán una serie de figuras que representan la frecuencia de detección de diferentes COVs para cada tipo de cáncer.

En la Fig. 9, se pueden observar los COVs más representativos del cáncer de pulmón. Siendo i) *acetone*, ii) *benzene* y iii) *toluene*, los compuestos más relevantes para este tipo de cáncer, debido a la frecuencia de repetición que tuvieron estos COVs en distintos estudios.

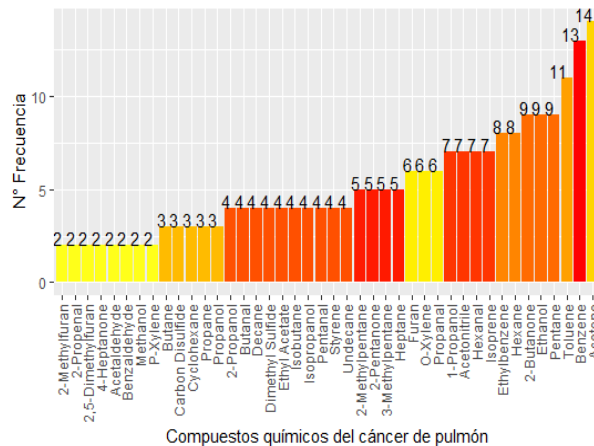


Fig. 9: COVs del cáncer de pulmón

En la Fig. 10, se pueden observar los COVs identificados en el cáncer de mama. Cabe mencionar que, solo dos compuestos fueron encontrados en diferentes artículos, estos son: i) *2-Nonanone* y ii) *4-Methyl-2-Heptanone*.

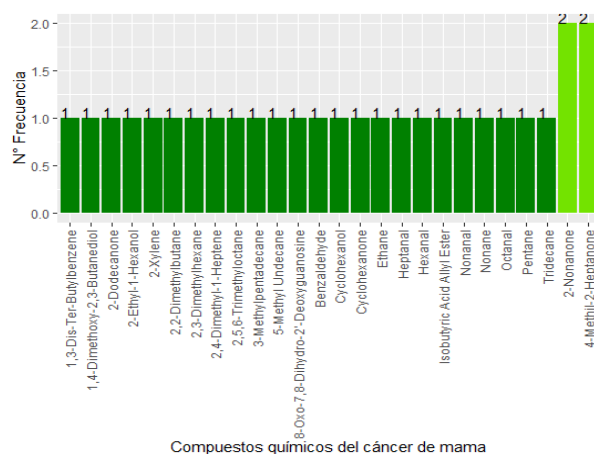


Fig. 10: COVs del cáncer de mama

En la Fig. 11, se pueden observar los COVs encontrados en el cáncer de ovario.

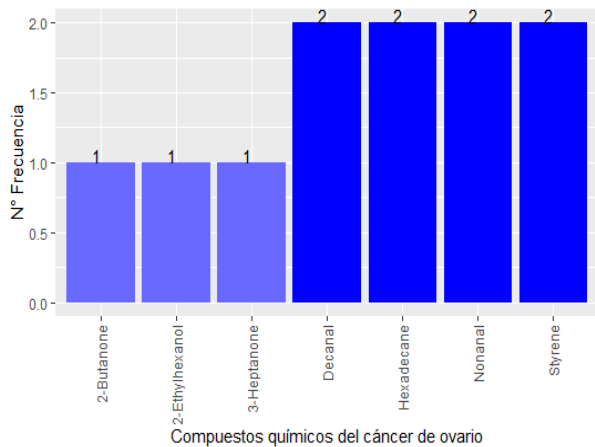


Fig. 11: COVs del cáncer de ovario

En la Fig. 12, se muestran los COVs identificados en el cáncer de vejiga.

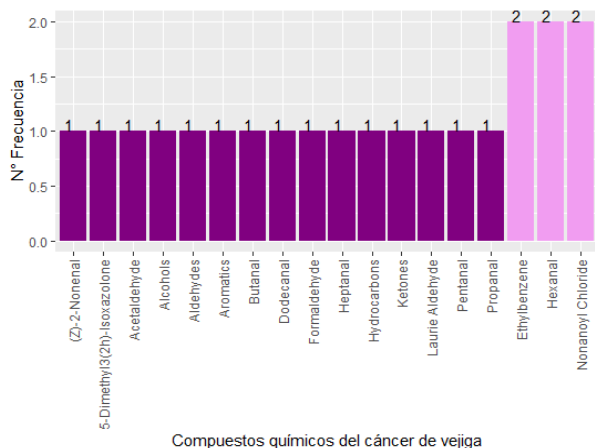


Fig. 12: COVs del cáncer de vejiga

En la Fig. 14, se pueden observar los COVs encontrados en el cáncer de cuello uterino.

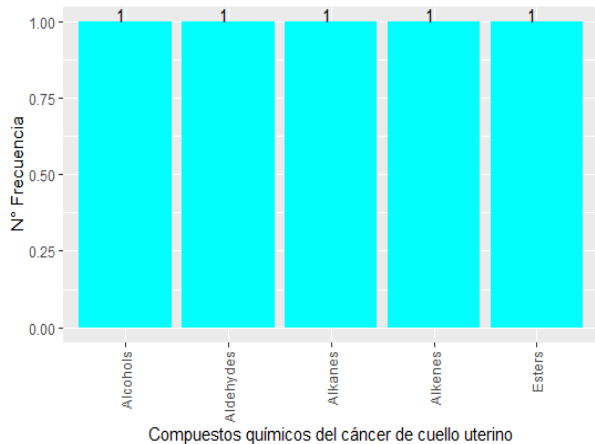


Fig. 13: COVs del cáncer de cuello uterino

En la Fig. 14, se pueden observar los COVs encontrados en el cáncer colorectal.

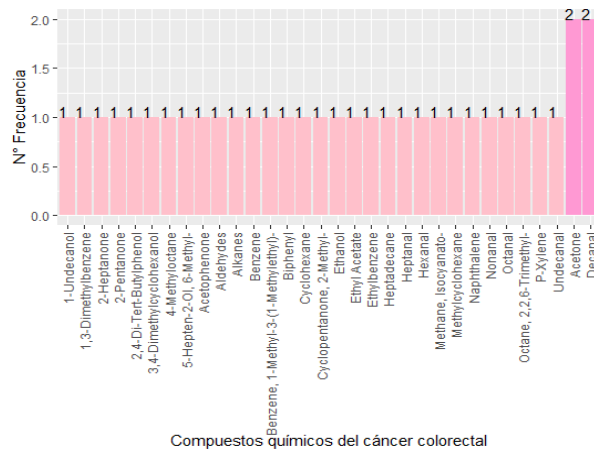


Fig. 14: COVs del cáncer colorectal

En la Fig. 15, se muestran los COVs identificados en el cáncer gástrico

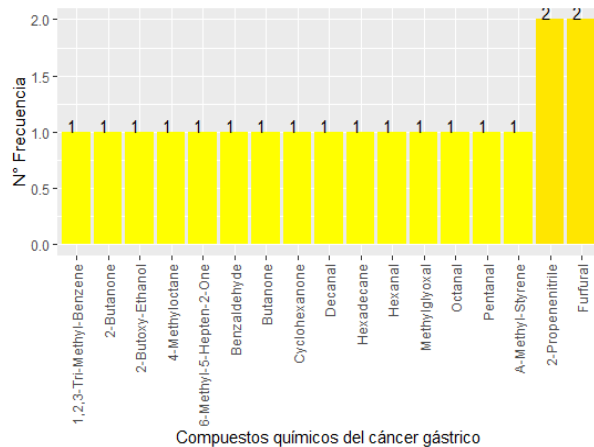


Fig. 15: COVs del cáncer gástrico

En la Fig. 16, se puede apreciar los COVs que fueron identificados para el carcinoma bronquial de células escamosas.



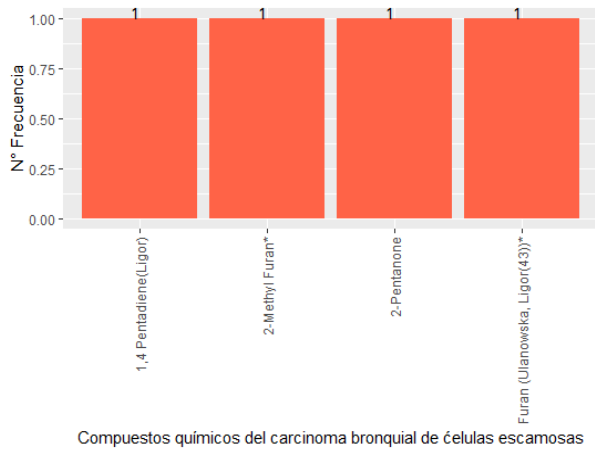


Fig. 16: COVs del carcinoma bronquial de células escamosas

En la Fig. 17, se puede observar los COVs encontrados en el carcinoma laríngeo de células escamosas.

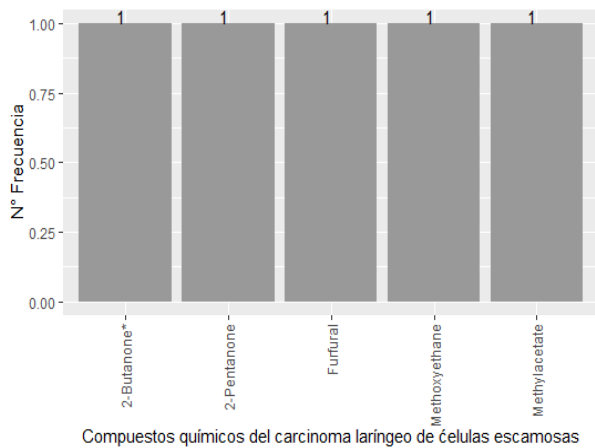


Fig. 17: COVs del carcinoma laríngeo de células escamosas

En la Fig. 18, se muestran los COVs identificados en el carcinoma de células renales claras.

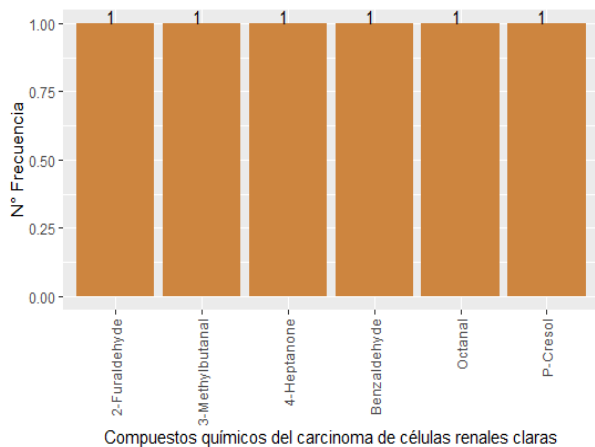


Fig. 18: COVs del carcinoma de células renales claras

En la Fig. 19, se puede observar los COVs encontrados en el carcinoma de células escamosas de cabeza y cuello.

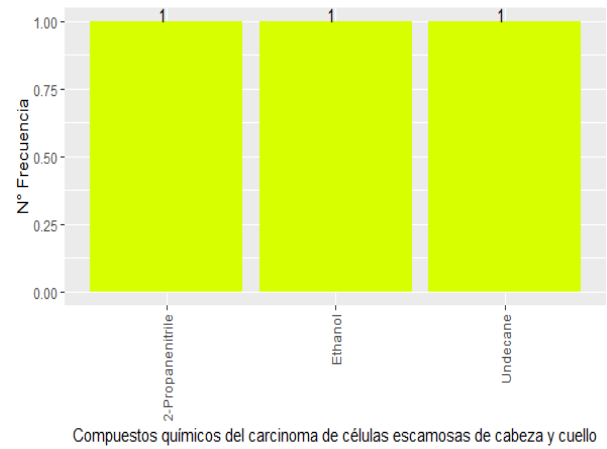


Fig. 19: COVs del carcinoma de células escamosas de cabeza y cuello

En la Fig. 20, se presentan los COVs identificados para el cáncer de próstata.

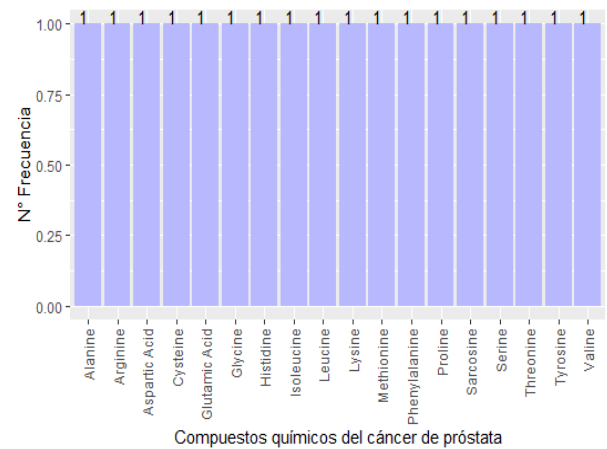


Fig. 20: COVs del cáncer de próstata

Finalmente, en la Fig. 21, se presentan los COVs encontrados para el adenocarcinoma pancreático.



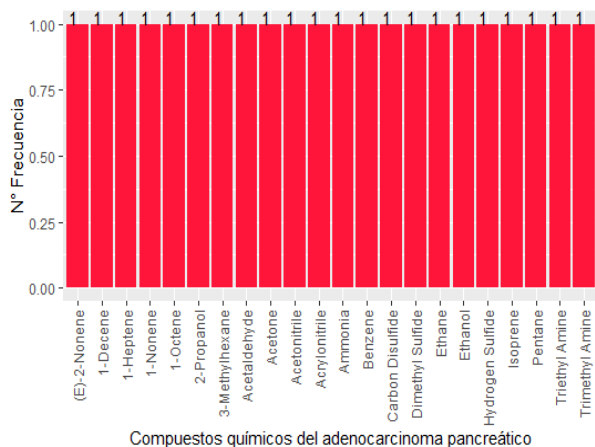


Fig. 21: COVs del cáncer de adenocarcinoma pancreático

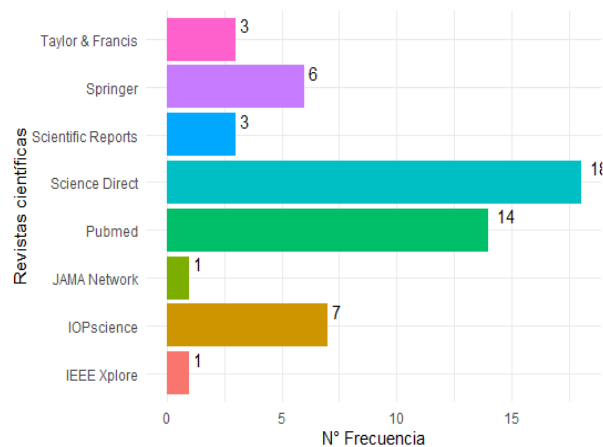


Fig. 22: Revistas científicas

Existen COVs de ocurrencia repetida en varios tipos de cáncer, es por este motivo que, fue fundamental encontrar estudios que investiguen la concentración de estos mismos, con la finalidad de determinar si las concentraciones de estos compuestos repetidos varían dependiendo el tipo de cáncer. Esta información puede ser observada en la Tabla XVI. Cabe mencionar que, no todos los COVs cuentan con valores de concentración disponibles.

*PMS2: ¿Cuántas investigaciones se han publicado sobre el tema de detección de olores de cáncer mediante el uso de narices electrónicas y/o sensores entre el 2015 a 2023?*

Mediante la Fig. 23, se puede observar los años en los cuales se publicaron estudios sobre el tema, siendo 2021 el año con más artículos publicados.

### C. Etapa3: Discusión de resultados

1) *Análisis de resultados:* Finalizado el proceso de recolección, organización y etiquetado de la información relevante de cada uno de los estudios en base a las preguntas de mapeo sistemático y revisión de la literatura sistemática, se obtienen los siguientes resultados.

*PMS1: ¿En qué revistas científicas se publican trabajos que utilizan narices electrónicas y/o sensores para la detección de olores de cáncer?*

En la Fig. 22, se puede observar las revistas científicas con más artículos publicados acerca del uso de e-noses y/o sensores para la detección de olores de cáncer, siendo *Science Direct* y *Pubmed* las revista con más artículos publicados acerca del tema. Cabe mencionar que, las revistas: i) *ACM* y ii) *Hindawi* fueron eliminadas debido a los filtros de revisión.

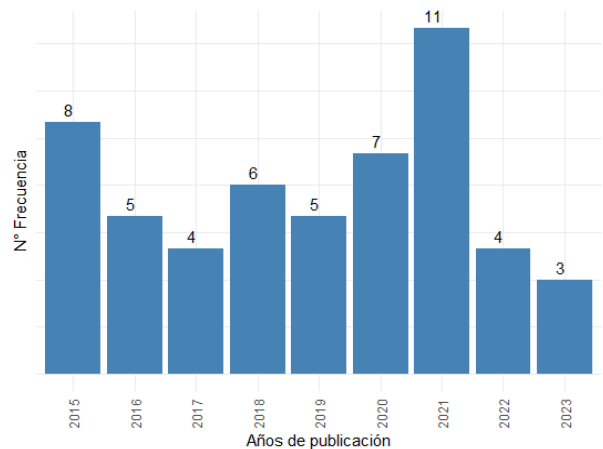


Fig. 23: Años de publicación general

Además, en este periodo la revista científica "*Science Direct*" fue la tuvo mayor aporte, con un total de 6 artículos, como se puede apreciar en la Fig. 24.

Tabla XVI: COMPUESTOS Y SUS CONCENTRACIONES PARA DIFERENTES TIPOS DE CÁNCER

Compuesto	Tipo de cáncer	Concentración	Referencias
Benzene	Cáncer de pulmón	1.29 - 3.82 ppb	[65]
	Cáncer colorrectal	-	[66]
	Adenocarcinoma pancreático	5.6 - 11.1 ppb	[33]
Hexanal	Cáncer de pulmón	3.8 - 5.3 ppb	[36]
	Cáncer de mama	34.49 ± 39.00 ppb	[30]
	Cáncer de vejiga	1.74 ± 1.45 µ/mL	[37]
	Cáncer colorrectal	-	[51]
	Cáncer gástrico	-	[67]
Pentanal	Cáncer de pulmón	4.8 - 7.2 ppb	[36]
	Cáncer de vejiga	0.206 ± 0.117 µ/mL	[37]
	Cáncer gástrico	-	[67]
Pentane	Cáncer de pulmón	1.73 - 17.50 ppb	[65]
	Cáncer de mama	-	[5]
	Adenocarcinoma pancreático	39 - 101.9 ppb	[33]
Styrene	Cáncer de pulmón	-	[20]
	Cáncer de ovario	249.83 ± 120.81 ppb	[32]
Isoprene	Cáncer de pulmón	19.2 - 295.5 ppb	[36]
	Adenocarcinoma pancreático	11.2 - 29.2 ppb	[33]
Ethylbenzene	Cáncer de pulmón	1.45 - 3.16 ppb	[65]
	Cáncer de vejiga	-	[12]
	Cáncer colorrectal	-	[51]
Ethanol	Cáncer de pulmón	12.8 - 1520.1 ppb	[36]
	Cáncer colorrectal	95.9 ± 48.1 ppb	[29]
	Carcinoma de células escamosas de cabeza y cuello	1198.3 ppb	[31]
	Adenocarcinoma pancreático	67.9 - 142.2 ppb	[33]
Ammonia	Cáncer de pulmón	-	[45]
	Adenocarcinoma pancreático	53.2 - 140.5 ppb	[33]
Acetone	Cáncer de pulmón	34.57 - 390.60 ppb	[65]
	Cáncer colorrectal	999.6 ± 116.8 ppb	[29]
	Adenocarcinoma pancreático	327.6 - 1779.4 ppb	[33]
Butanal	Cáncer de pulmón	1.32 - 2.55 ppb	[65]
	Cáncer de vejiga	0.071 ± 0.050 µ/mL	[37]
Ethyl acetate	Cáncer de pulmón	3.98 - 22.89 ppb	[65]
	Cáncer colorrectal	128.4 ± 37.0 ppb	[29]
Propanal	Cáncer de pulmón	1.56 - 3.74 ppb	[65]
	Cáncer de vejiga	0.195 ± 0.111 µ/mL	[37]
2-Propanol	Cáncer de pulmón	3.32 - 7.19 ppb	[65]
	Adenocarcinoma pancreático	54.5-88.8 ppb	[33]
Acetonitrile	Cáncer de pulmón	10.96-23.60 ppb	[65]
	Adenocarcinoma pancreático	39.1 - 62.6 ppb	[33]
Cyclohexane	Cáncer de pulmón	3.1 ppb	[68]
	Cáncer colorrectal	-	[69]
Nonanal	Cáncer de pulmón	-	[43]
	Cáncer demama	47.27 ± 49.23 ppb	[30]
	Cáncer de ovario	426.93 ± 116.20 ppb	[32]
	Cáncer colorrectal	-	[51]
2-pentanone	Cáncer de pulmón	3.25 - 8.77 ppb	[65]
	Cáncer colorrectal	-	[51]
	Carcinoma bronquial de células escamosas	-	[27]
	Carcinoma laríngeo de células escamosas	-	[27]
Benzaldehyde	Cáncer de pulmón	-	[43]
	Cáncer de mama	-	[52]
	Cáncer gástrico	-	[67]
	Carcinoma de células renales claras	-	[35]

Compuesto	Tipo de cáncer	Concentración	Referencias
Octanal	Cáncer de pulmón	-	[43]
	Cáncer de mama	25.22 ± 31.36 ppb	[30]
	Cáncer colorrectal	-	[51]
	Cáncer gástrico	-	[67]
	Carcinoma de células renales claras	-	[35]
3-Heptanone	Cáncer de pulmón	-	[43]
	Cáncer de ovario	-	[55]
2-Heptanone	Cáncer de pulmón	-	[43]
	Cáncer colorrectal	-	[51]
2-Butanone	Cáncer de pulmón	1.35 - 2.86 ppb	[65]
	Cáncer de ovario	186.51 ± 136.72 ppb	[32]
	Cáncer gástrico	68.3 ± 49.0 ppb	[34]
	Carcinoma laríngeo de células escamosas	-	[27]
4-Heptanone	Cáncer de pulmón	4.0-60.0 ppb	[68]
	Carcinoma de células renales claras	-	[35]
Octane, 2,2,6-trimethyl-	Cáncer de pulmón	-	[24]
	Cáncer colorrectal	-	[51]
Undecane	Cáncer de pulmón	-	[24]
	Carcinoma de células escamosas de cabeza y cuello	219.7 ppb	[31]
Dimethyl sulfide	Cáncer de pulmón	6.3 - 18.5 ppb	[36]
	Adenocarcinoma pancreático	6.6 - 8.5 ppb	[33]
Acetaldehyde	Cáncer de pulmón	4.1 - 272.7 ppb	[36]
	Cáncer de vejiga	0.421 ± 0.221 µ/mL	[37]
	Adenocarcinoma pancreático	69.1 - 159.1 ppb	[33]
Carbon disulfide	Cáncer de pulmón	1.0 - 19.0 ppb	[36]
	Adenocarcinoma pancreático	6.6 - 17.3 ppb	[33]
p-Xylene	Cáncer de pulmón	1.6 - 32.7 ppb	[36]
	Cáncer colorrectal	-	[51]
4-Methyloctane	Cáncer de pulmón	2.2 - 4.5 ppb	[36]
	Cáncer colorrectal	16.0 ± 0.6 ppb	[29]
	Cáncer gástrico	18.8 ± 4.5 ppb	[34]
Methylacetate	Cáncer pulmón	3.0-39.0 ppb	[68]
	Carcinoma laríngeo de células escamosas	-	[27]
3-Methylhexane	Cáncer de pulmón	2.0 ppb	[68]
	Adenocarcinoma pancreático	49 - 130.7 ppb	[33]
Ethane	Cáncer de mama	-	[5]
	Adenocarcinoma pancreático	97.1 - 105 ppb	[33]
Cyclohexanone	Cáncer de mama	-	[1]
	Cáncer gástrico	-	[67]
Heptanal	Cáncer de mama	22.52 ± 22.86	[30]
	Cáncer de vejiga	0.219 ± 0.098 µ/mL	[37]
	Cáncer colorrectal	-	[51]
Decanal	Cáncer de ovario	176.73 ± 48.12 ppb	[32]
	Cáncer colorrectal	-	[51]
	Cáncer gástrico	-	[67]
Hexadecane	Cáncer de ovario	3.0 ± 1.34 ppb	[32]
	Cáncer gástrico	10.7 ± 12.3 ppb	[34]
Aldehydes	Cáncer de vejiga	-	[11]
	Cáncer de cuello uterino	-	[22]
	Cáncer colorrectal	-	[66]
Alcohols	Cáncer de vejiga	-	[11]
	Cáncer de cuello uterino	-	[22]
Alkanes	Cáncer de cuello uterino	-	[22]
	Cáncer colorrectal	-	[66]

Compuesto	Tipo de cáncer	Concentración	Referencias
Furfural	Cáncer gástrico	5.6 ± 3.5 ppb	[34]
	Carcinoma laríngeo de células escamosas	-	[27]

ppb:pares por billón;  $\mu/mL$ :microgramos por mililitro;  $\pm$ :desviación estándar

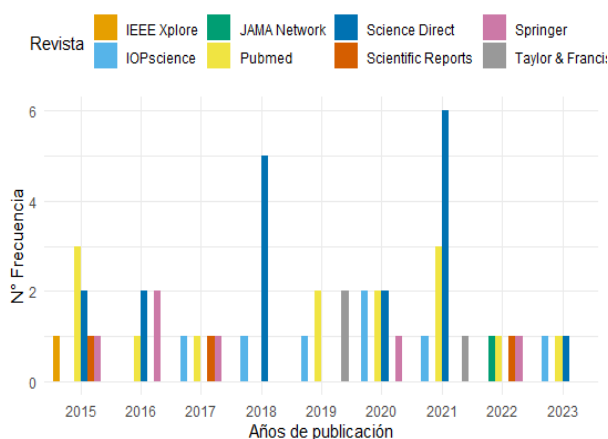


Fig. 24: Años de publicación por revista científica

*PRLS1: ¿Cuáles son los sensores más utilizados en la detección de olores para cáncer?*

En la presente investigación se pudo identificar quince tipos de sensores, mediante la Fig. 25, se pueden observar los sensores más utilizados en la detección del cáncer. Destacan especialmente los sensores de óxido metálico, los cuales muestran una diferencia significativa frente a los demás tipos. Esto se debe a las numerosas ventajas que ofrecen, las cuales serán evidenciadas más adelante. Debido al amplio número y diversidad de sensores comerciales encontrados se recomienda consultar los estudios previos [40], [42], [46], [49], [41], [51] y [66] para obtener más detalles sobre las características específicas de cada sensor.

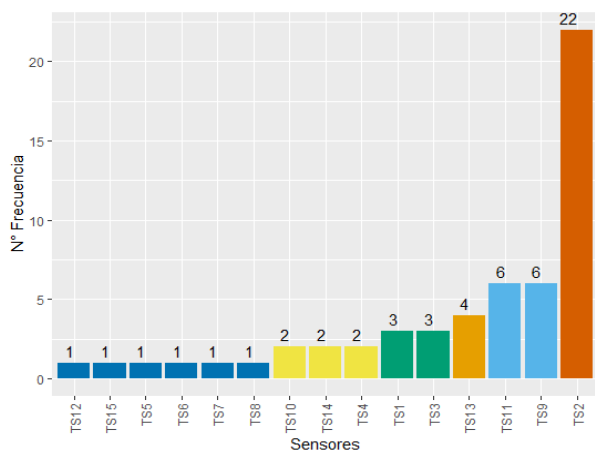


Fig. 25: Sensores más utilizados

TS1:sensores de microbalanza de cristal de cuarzo (QCM);

TS2:sensores de óxido metálico; TS3:sensores químicos;

TS4:sensores de gas de hilo caliente; catalítica

TS5:sensores de electrolito sólido; TS6:sensores plasmónicos de color; TS7:sensores ópticos infrarrojos; TS8:sensores de fotoionización; TS9:sensores químicos basados en nanopartículas de oro (GNP) modificadas molecularmente; TS10:matrices ópticas de sensores de fluorescencia de gases; TS11:sensores basados en polímeros; TS12:sensores con transistores de efecto de campo de nanocables de silicio modificados molecularmente (SiNW FET); TS13:sensores de nanotubos de carbono de pared simple (SWCNT); TS14:sensores electroquímicos; TS15:sensores de gas de tipo combustión

Además, en la Tabla XVII, se pueden observar los costos de los sensores comerciales que se emplearon con más frecuencia. Con la finalidad que, trabajos futuros puedan tener conocimiento del precio de estos sensores al momento de comprarlos y construir el dispositivo. Cabe mencionar que, el precio de los sensores que sobrepasan los \$ 60 dolares se debe a varios factores como: i) tecnología más avanzada, ii) precisión en la detección de gases y iii) su aplicación en diversas industrias. Sin embargo, esto no quiere decir que los demás sensores sean pésimos en sus funciones, ya que estos han sido utilizado en varios estudios, obteniendo buenos resultados.

Tabla XVII: COSTOS DE SENSORES COMERCIALES

Nombre del sensor	Costo
<b>Sensores de óxido metálico</b>	
TGS 2600	\$16.00
TGS 822	\$7.00
TGS 826	\$76.28
TGS 2620	\$5.70
TGS 2610	\$5.50
TGS 2602	\$4.16
TGS 2444	\$112.80
WSP 2110	\$8.83
MQ-3	\$1.13
<b>Sensores de gas de hilo caliente</b>	
MR 156	\$62.62
<b>Sensores electroquímicos</b>	
ME4-C6H6	\$123.00

*PRLS2: ¿Cuáles son las ventajas y desventajas de utilizar narices electrónicas y/o sensores en el campo de la salud para detectar el cáncer?*

Mediante la Tabla XVIII, se pueden apreciar las ventajas y desventajas de: i) los tipos de sensores, ii) e-noses diseñadas y desarrolladas y ii) e-noses comerciales.

*PRLS3: ¿Cuáles son los Compuestos Orgánicos Volátiles (COVs) más comunes que se han identificado en el cáncer?*

Tabla XVIII: VENTAJAS Y DESVENTAJAS DE SENSORES Y E-NOSES

Ventajas	Referencias	Desventajas	Referencias
<b>Sensores de microbalanza de cristal de cuarzo (QCM)</b>			
Alta sensibilidad	[6] [70] [11]	Cuestiones de respuesta cruzada	[11]
Respuesta rápida	[6] [70]		
Buena estabilidad	[6]		
Bajo coste	[70]		
Portátil	[70]		
Procedimiento en tiempo real	[70]		
Simplicidad de la estructura	[11]		
<b>Sensores de óxido metálico</b>			
Alta sensibilidad	[9] [52] [60] [45] [71] [46]	Poco selectivo a temperatura ambiente con alta presencia de vapor de agua	[52]
Respuesta rápida	[9] [60]	Desplazamiento de la línea de base	[11]
Bajo coste	[9] [52] [60] [71] [48] [11] [46] [72] [66]	La sensibilidad se activa a alta temperatura	[11] [23]
Estabilidad a largo plazo	[9]	Poco específico	[52]
Fiabilidad de la señal	[9]		
Portátil	[52]		
No invasivo	[72] [66]		
Fácilmente disponibles en el mercado	[52]		
Resultados notablemente buenos	[52]		
Sensible a una amplia gama de COV	[52] [23]		
Bajo consumo de energía	[60] [45]		
Pequeño tamaño	[45] [71]		
Resistentes	[45]		
Amplia gama de respuesta	[11]		
Larga vida útil	[46]		
Estabilidad química a lo largo del tiempo	[60] [46]		
Circuitos asociados sencillos	[46]		
Sencillos	[66]		
Buena reproducibilidad	[60]		
Rápidos	[23]		
Fácil de usar	[23]		
Posibilidad de realizar pruebas en el punto de atención	[23]		
<b>Sensores químicos</b>			
Alta sensibilidad	[5]		
Respuesta rápida	[5]		
Alta especificidad	[5]		
<b>Sensores plasmónicos de color</b>			
Características de detección rápida	[57]		
Facilidad de uso	[57]		
Bajo coste	[57]		
<b>Sensores químicos basados en nanopartículas de oro (GNP) modificadas molecularmente</b>			
Pequeños	[55]		
Fácil de usar	[55]		
Bajo costo	[55]		
Simple	[55]		
Portátil	[55]		
No invasivo	[55]		
<b>Matrices ópticas de sensores de fluorescencia de gases</b>			
Bajo costo	[11]		
No invasivo	[11] [4]		
Resultados en tiempo real	[11]		
Intuitivo	[11]		
Rentables	[11] [4]		

Ventajas	Referencias	Desventajas	Referencias
<b>Sistemas e-nose diseñado y desarrollado</b>			
No invasivo	[14] [40] [42]	Un procedimiento de prueba completo lleva demasiado tiempo	[41]
Bajo costo	[14] [45] [40] [4]		
Portátil	[14] [45] [46]		
Simple	[14] [42]		
Respuesta rápida	[14] [45]		
Proporciona una mayor precisión	[14]		
Proporciona una mayor sensibilidad	[14]		
Proporciona una mayor especificidad	[14]		
Resistente	[45] [46]		
Rentable	[45] [42] [46]		
Pueden miniaturizarse	[4]		
Requiere menos tiempo	[4]		
<b>E-nose comercial</b>			
No invasivo	[5] [73] [48] [53] [38] [21] [22] [12] [39] [51] [44] [25] [26] [47] [58] [28] [8]	Costoso	[14]
Bajo costo	[5] [13] [62] [38] [2] [21] [22] [12] [39] [51] [47] [58] [28] [8]	Difícil de usar	[14]
Seguro	[5]	Incapacidad para diferenciar entre compuestos endógenos y exógenos	[73]
Fácil de usar	[5] [62] [12] [51] [47] [49] [58] [28]	La precisión diagnóstica depende, en gran medida, del tamaño del conjunto de entrenamiento y de la representatividad de la población de la muestra	[38] [47]
Portátil	[5] [62] [38] [2] [51] [8]	Los sensores de la placa deben calentarse para distinguir los distintos COV con mayor precisión	[38] [2] [21]
Análisis en tiempo real	[5] [51] [26]	Incapacidad de para identificar un COV específico	[62] [12]
Accesible a la población	[5]	Estrictos controles de temperatura de los sensores	[28]
Reduce los periodos de espera de los resultados	[5]	Posibles problemas de reproducibilidad	[28]
Simple	[13] [21]	Desviación de los sensores	[28]
Ampliamente disponible	[13]	Variabilidad de los instrumentos	[28]
Herramienta para el punto de atención	[38] [39] [28]	Pérdida de sensibilidad en presencia de alcohol	[28]
Rápido	[2] [21] [22] [12] [39] [51] [47]	No está pensada como herramienta de diagnóstico de enfermedades avanzadas	[12]
Proporcionar análisis rápidos	[12]	La precisión se ve afectada por factores endógenos y exógenos	[62]
No depende del operador	[12]		
Repetible	[12]		
Potencialmente rentable	[5] [12]		
Preciso	[12]		
Alta sensibilidad	[22] [49] [8]		
Ofrece la oportunidad de transferir el modelo de calibración	[39] [47] [28]		
Proporcionar resultados a un coste inferior por prueba	[51]		
Fácil medición	[44]		
Rápida administración	[47]		
Fiable	[49]		
Resultados rápidos	[73] [62]		
Cómodo de usar	[73]		
Facilidad de administración de la prueba	[8]		

En los diferentes tipos de cáncer se han encontrado varios COVs que pueden estar relacionados con estas enfermedades, Sin embargo, en el cáncer de pulmón se han identificado COVs que parecen tener una fuerte relación con este tipo de cáncer debido a su frecuencia de aparición, estos son: i) acetone, ii) benzene, iii) toluene, iv) pentane, v) ethanol, vi) 2-butanone, vii) hexane y viii) ethylbenzene, esta información puede ser observada en la Fig. 9. Además, como se mencionó previamente, se han encontrado varios COVs en los demás tipos de cáncer y aunque su frecuencia de aparición fue baja,

es importante mencionarlos a través de las Fig. 10-21, con la finalidad de poner en conocimiento estos COVs.

*PRLS4: ¿Qué compuestos y concentraciones químicas se han identificado en los diferentes tipos de cáncer?*

Debido a la gran variedad de tipos de cáncer encontrados en esta investigación, se pudo observar la presencia de ciertos compuestos en múltiples categorías de esta enfermedad, motivo por el cual, se identificaron dichos compuestos junto con sus concentraciones como se puede observar en la Tabla XVI, este ultimo es de gran importancia ya que se observó que su valor varía dependiendo el tipo de cáncer, dando a

entender que un compuesto en una concentración o rango de concentración específica puede ser un indicador de la afección (en este caso el cáncer).

*PRLS5: ¿Qué resultados cuantitativos han proporcionado las métricas utilizadas para la clasificación de olores en cáncer?*

Dada la numerosa cantidad de estudios que utilizaron e-noses y/o sensores para el análisis de COVs en pacientes y controles, se pudo extraer los resultados cuantitativos de dichos estudios, los cuales pueden ser observados en la Tabla S1, del Material Suplementario. Donde se destacan los siguientes resultados:

- de Leon-Martínez *et al* [1], utilizaron la e-nose Cyranose 320 y CDA para detectar el cáncer de mama mediante el aliento exhalado de pacientes y controles, obteniendo una sensibilidad, especificidad, precisión, tasa de clasificación correcta, AUC, VPP y VNP de 100%, 100%, 100%, 98%, 1, 100% y 100%, respectivamente. Siendo este estudio el que obtuvo el mejor rendimiento.
- Liu, B. *et al* [41], diseñaron y desarrollaron una e-nose con sensores: i) de óxido metálico, ii) químicos, iii) de hilo caliente y iv) de electrolito sólido, junto con SVM, obteniendo una sensibilidad, especificidad, precisión, y AUC de 95.65%, 80.49%, 88.51% y 0.95, respectivamente, cuando no se seleccionaron las características, en cambio, cuando este proceso fue realizado mediante el método SGL, obtuvieron una sensibilidad, especificidad, precisión, y AUC de 97.83%, 90.24%, 94.25% y 0.99, respectivamente. Lo que quiere decir que, si se realiza el proceso de selección y extracción de características el rendimiento del algoritmo de aprendizaje automático mejorará.
- Amal, H *et al* [34], utilizaron sensores GNP y SWCNT y DFA, obteniendo una sensibilidad, especificidad, precisión y AUC de 73%, 98%, 92% y 0.92, respectivamente.

Y por parte, los estudios que utilizaron GC-MS con el mismo propósito, también arrojaron resultados cuantitativos, lo cuales son presentados en la Tabla S2, del Material Suplementario. Destacando el siguiente resultado:

- Tyagi, H. *et al* [51], obtuvieron una sensibilidad, especificidad, AUC, VPP y VPN de 89%, 75%, 0.93, 85% y 92%, respectivamente, cuando utilizaron GC-MS y Random forest.

*PRLS6: ¿Qué métodos se utilizan comúnmente en el preprocesamiento de los datos generados por las narices electrónicas y/o sensores?*

Como se puede observar en la Tabla VII, existen varios métodos de preprocesamiento de datos, estos son: i) normalización, ii) estandarización, iii) procesamiento de la línea de base, iv) eliminación de la retrodispersión, v) de-noising, vi) serialización, vii) autoescalado, viii) filtro de la ventana-media y ix) filtrado Savitzky-Golay. Cada uno de estos métodos da un tratamiento al conjunto de datos con la finalidad de reducir efectos de respuesta ajenos a los COVs de una muestra, como por ejemplo el efecto ambiental [22]. Además, mediante la

Tabla XVI, se puede decir que, la *normalización* es el método más usado para procesar los datos generados a partir de las e-noses/sensores y por GC-MS.

*PRLS7: ¿Cuáles son las técnicas más utilizadas que permiten mejorar la calidad de los datos previo al reconocimiento de patrones?*

Mediante las Tablas VIII, IX y XIII, se pueden visualizar los métodos que mejoran la calidad de los datos, siendo: i) Tucker3 y ii) PCA, los métodos más utilizados. Por un lado *Tucker3* evita i) correlaciones espurias, ii) datos redundantes y iii) ruido. Y por otro lado *PCA* evita i) redundancia y ii) ajuste excesivo al momento de la clasificación. Sin embargo, ambos métodos son fundamentales para mejorar la calidad de un conjunto de datos, y así poder obtener una clasificación más precisa.

### III. CONCLUSIONES

En la presente investigación se ha realizado un estado del arte mediante las metodologías de MS y RLS con la finalidad de recopilar información acerca del análisis y detección de olores de cáncer mediante e-noses y/o sensores, utilizando como referencia primaria 70 artículos de investigación. Varios procesos/bloques, fueron los que facilitaron la estructura de la investigación, ofreciendo al lector bloques como: B1 Definición de las preguntas de investigación de MS y RLS, alcance de la revisión, criterios de inclusión y exclusión de MS y RLS y conducta de la búsqueda; B2 Selección de estudios primarios y definición de los criterios de análisis y B3 Análisis de resultados.

En la metodología de MS las revistas con mayor información en relación al tema de investigación, fueron Pubmed y Science Direct, sin embargo, este último presenta más resultados en el año 2018 y 2021. Cabe mencionar que, en la actualidad solo se han publicado 3 artículos sobre el tema en cuestión. En la metodología de RLS se estructuró una taxonomía con la finalidad de organizar la información en base a: i) cancer (como tema principal) y dos ramas secundarias; en la primera: i) e-noses/sensores, ii) preprocesamiento de datos iii) compresión de datos, iv) selección y extracción de características, v) algoritmos de aprendizaje automático, vi) validación cruzada y vii) métricas de evaluación; y la segunda se divide en dos subramas, la primera: i) GC-MS, ii) preprocesamiento de datos, iii) selección y extracción de características, iv) algoritmos de aprendizaje automático, v) validación cruzada y vi) métricas de evaluación; y en la segunda subrama: i) concentraciones y compuestos.

Los métodos más utilizados para el preprocesamiento de datos son: i) normalización, ii) estandarización, iii) procesamiento de la línea de base y iv) autoescalado. En cuanto a los métodos de compresión de datos, se identificó uno, el cual es i) Tucker3. Por otro lado, el método de selección y extracción de características más utilizado fue: i) PCA. En cambio, los algoritmos de aprendizaje automático que más se utilizaron fueron: i) ANN y ii) SVM. Los métodos de validación cruzada más utilizados fueron: i) leave-one out cross-validation, ii) k-fold cross-validation y iii) leave-10%-out cross-validation.

Finalmente las métricas de evaluación más utilizadas fueron: i) sensibilidad ii) especificidad y iii) AUC.

Además, se pudo identificar resultados cuantitativos con valores altos, presentando sensibilidad, especificidad, precisión y AUC de: 100%, 100%, 100% y 1, respectivamente en [1]; 95.65%, 80.49%, 88.51% y 0.95, respectivamente en [41]; 73%, 98%, 92% y 0.92, respectivamente en [34] y sensibilidad, especificidad y AUC de 89%, 75% y 0.93, respectivamente en [51].

Finalmente, el estado del arte actual servirá de apoyo para el análisis y comprensión de cada uno de los procesos realizados, desde la definición de las preguntas de investigación hasta el análisis de resultados. Y además, será un sustento científico que ayude a investigadores, analistas, estudiantes de las ciencias de la computación para la creación de aplicaciones de hardware y software, que proporcionen beneficios a usuarios finales vinculados con esta área de investigación.

## REFERENCES

- [1] L. D. de Leon-Martínez, M. Rodríguez-Aguilar, P. Gorocica-Rosete, C. A. Domínguez-Reyes, V. Martínez-Bustos, J. A. Tenorio-Torres, O. Ornelas-Rebolledo, J. A. Cruz-Ramos, B. Balderas-Segura, and R. Flores-Ramírez, "Identification of profiles of volatile organic compounds in exhaled breath by means of an electronic nose as a proposal for a screening method for breast cancer: A case-control study," *Journal of Breath Research*, vol. 14, no. 4, p. 046009, 2020.
- [2] R. Van De Goor, N. Leunis, M. Van Hooren, E. Francisca, A. Masclee, B. Kremer, and K. Kross, "Feasibility of electronic nose technology for discriminating between head and neck, bladder, and colon carcinomas," *European Archives of Oto-Rhino-Laryngology*, vol. 274, pp. 1053–1060, 2017.
- [3] B. Behera, R. Joshi, G. A. Vishnu, S. Bhalerao, and H. J. Pandya, "Electronic nose: A non-invasive technology for breath analysis of diabetes and lung cancer patients," *Journal of breath research*, vol. 13, no. 2, p. 024001, 2019.
- [4] S. Zhu, S. Corsetti, Q. Wang, C. Li, Z. Huang, and G. Nabi, "Optical sensory arrays for the detection of urinary bladder cancer-related volatile organic compounds," *Journal of Biophotonics*, vol. 12, no. 10, p. e201800165, 2019.
- [5] M. Rodríguez-Aguilar, L. D. de León-Martínez, P. Gorocica-Rosete, R. Pérez-Padilla, C. A. Domínguez-Reyes, J. A. Tenorio-Torres, O. Ornelas-Rebolledo, G. Mehta, B. N. Zamora-Mendoza, and R. Flores-Ramírez, "Application of chemoresistive gas sensors and chemometric analysis to differentiate the fingerprints of global volatile organic compounds from diseases. preliminary results of copd, lung cancer and breast cancer," *Clinica Chimica Acta*, vol. 518, pp. 83–92, 2021.
- [6] N. L. L. Hekiem, A. A. M. Ralib, F. B. Ahmad, A. N. Nordin, R. Ab Rahim, N. F. Za'bah *et al.*, "Advanced vapour sensing materials: Existing and latent to acoustic wave sensors for vocs detection as the potential exhaled breath biomarkers for lung cancer," *Sensors and Actuators A: Physical*, vol. 329, p. 112792, 2021.
- [7] C. Baldini, L. Billeci, F. Sansone, R. Conte, C. Domenici, and A. Tonacci, "Electronic nose as a novel method for diagnosing cancer: a systematic review," *Biosensors*, vol. 10, no. 8, p. 84, 2020.
- [8] R. F. Machado, D. Laskowski, O. Deffenderfer, T. Burch, S. Zheng, P. J. Mazzone, T. Mekhail, C. Jennings, J. K. Stoller, J. Pyle *et al.*, "Detection of lung cancer by sensor array analyses of exhaled breath," *American journal of respiratory and critical care medicine*, vol. 171, no. 11, pp. 1286–1291, 2005.
- [9] J.-E. Chang, D.-S. Lee, S.-W. Ban, J. Oh, M. Y. Jung, S.-H. Kim, S. Park, K. Persaud, and S. Jheon, "Analysis of volatile organic compounds in exhaled breath for lung cancer diagnosis using a sensor system," *Sensors and actuators B: chemical*, vol. 255, pp. 800–807, 2018.
- [10] G. Rocco, G. Pennazza, M. Santonico, F. Longo, R. Rocco, P. Crucitti, and R. A. Incalzi, "Breathprinting and early diagnosis of lung cancer," *Journal of thoracic oncology*, vol. 13, no. 7, pp. 883–894, 2018.
- [11] S. Zhu, Z. Huang, and G. Nabi, "Fluorometric optical sensor arrays for the detection of urinary bladder cancer specific volatile organic compounds in the urine of patients with frank hematuria: a prospective case-control study," *Biomedical Optics Express*, vol. 11, no. 2, pp. 1175–1185, 2020.
- [12] P. Bassi, L. Di Gianfrancesco, L. Salmaso, M. Ragonese, G. Palermo, E. Sacco, R. A. Giancristofaro, R. Ceccato, and M. Racioppi, "Improved non-invasive diagnosis of bladder cancer with an electronic nose: A large pilot study," *Journal of Clinical Medicine*, vol. 10, no. 21, p. 4984, 2021.
- [13] M. Tirzite, M. Bukovskis, G. Strazda, N. Jurka, and I. Taivans, "Detection of lung cancer with electronic nose and logistic regression analysis," *Journal of breath research*, vol. 13, no. 1, p. 016006, 2018.
- [14] V. Binson and M. Subramoniam, "Artificial intelligence based breath analysis system for the diagnosis of lung cancer," in *Journal of Physics: Conference Series*, vol. 1950, no. 1. IOP Publishing, 2021, p. 012065.
- [15] P. López, R. Triviño, D. Calderón, A. Arceales, and A. V. Guamán, "Electronic nose prototype for explosive detection," in *2017 CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON)*. IEEE, 2017, pp. 1–4.
- [16] J. Torres-Tello, A. V. Guaman, and S.-B. Ko, "Improving the detection of explosives in a mox chemical sensors array with lstm networks," *IEEE Sensors Journal*, vol. 20, no. 23, pp. 14 302–14 309, 2020.
- [17] D. Carrizo and C. Moller, "Estructuras metodológicas de revisiones sistemáticas de literatura en ingeniería de software: un estudio de mapeo sistemático," *Ingeniare. Revista chilena de ingeniería*, vol. 26, pp. 45–54, 2018.
- [18] A. B. Marques and L. da Silva Monte, "How are software technologies being evaluated with autistic users? a systematic mapping," *Universal Access in the Information Society*, vol. 21, no. 3, pp. 587–597, 2022.
- [19] M. R. van Hooren, N. Leunis, D. S. Brandsma, A.-M. C. Dingemans, B. Kremer, and K. W. Kross, "Differentiating head and neck carcinoma from lung carcinoma with an electronic nose: a proof of concept study," *European Archives of Oto-Rhino-Laryngology*, vol. 273, pp. 3897–3903, 2016.
- [20] I. Nardi-Agmon, M. Abud-Hawa, O. Liran, N. Gai-Mor, M. Ilouze, A. Onn, J. Bar, D. Shlomi, H. Haick, and N. Peled, "Exhaled breath analysis for monitoring response to treatment in advanced lung cancer," *Journal of Thoracic Oncology*, vol. 11, no. 6, pp. 827–837, 2016.
- [21] E. Steenhuis, I. Schoenaker, J. de Groot, H. Fiebrich, J. de Graaf, R. Brohet, J. van Dijk, H. van Westreenen, P. Siersema *et al.*, "Feasibility of volatile organic compound in breath analysis in the follow-up of colorectal cancer: A pilot study," *European Journal of Surgical Oncology*, vol. 46, no. 11, pp. 2068–2073, 2020.
- [22] L. D. de León-Martínez, R. Flores-Ramírez, C. M. López-Mendoza, M. Rodríguez-Aguilar, G. Metha, L. Zúñiga-Martínez, O. Ornelas-Rebolledo, and L. E. Alcántara-Quintana, "Identification of volatile organic compounds in the urine of patients with cervical cancer. test concept for timely screening," *Clinica Chimica Acta*, vol. 522, pp. 132–140, 2021.
- [23] L. Lavra, A. Catini, A. Ulivieri, R. Capuano, L. Baghermajad Salehi, S. Sciacchitano, A. Bartolazzi, S. Nardis, R. Paolesse, E. Martinelli *et al.*, "Investigation of vocs associated with different characteristics of breast cancer cells," *Scientific reports*, vol. 5, no. 1, pp. 1–12, 2015.
- [24] R. Capuano, M. Santonico, G. Pennazza, S. Ghezzi, E. Martinelli, C. Roscioni, G. Lucantoni, G. Galluccio, R. Paolesse, C. Di Natale *et al.*, "The lung cancer breath signature: a comparative analysis of exhaled breath and air sampled from inside the lungs," *Scientific Reports*, vol. 5, no. 1, p. 16491, 2015.
- [25] R. Vadala, B. Pattnaik, S. Bangaru, D. Rai, J. Tak, S. Kashyap, U. Verma, G. Yadav, R. Dhaliwal, S. Mittal *et al.*, "A review on electronic nose for diagnosis and monitoring treatment response in lung cancer," *Journal of Breath Research*, 2023.
- [26] A. McWilliams, P. Beigi, A. Srinidhi, S. Lam, and C. E. MacAulay, "Sex and smoking status effects on the early detection of early lung cancer in high-risk smokers using an electronic nose," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 8, pp. 2044–2054, 2015.
- [27] D. Fielding, G. Hartel, D. Pass, M. Davis, M. Brown, A. Dent, J. Agnew, G. Dickie, R. S. Ware, and R. Hodge, "Volatile organic compound breath testing detects in-situ squamous cell carcinoma of bronchial and laryngeal regions and shows distinct profiles of each tumour," *Journal of breath research*, vol. 14, no. 4, p. 046013, 2020.
- [28] K. E. van Keulen, M. E. Jansen, R. W. Schrauwen, J. J. Kolkman, and P. D. Siersema, "Volatile organic compounds in breath can serve as a non-invasive diagnostic biomarker for the detection of advanced adeno-



- mas and colorectal cancer," *Alimentary pharmacology & therapeutics*, vol. 51, no. 3, pp. 334–346, 2020.
- [29] H. Amal, M. Leja, K. Funka, I. Lasina, R. Skapars, A. Sivins, G. Ancans, I. Kikuste, A. Vanags, I. Tolmanis *et al.*, "Breath testing as potential colorectal cancer screening tool," *International journal of cancer*, vol. 138, no. 1, pp. 229–236, 2016.
- [30] J. Li, Y. Peng, Y. Liu, W. Li, Y. Jin, Z. Tang, and Y. Duan, "Investigation of potential breath biomarkers for the early diagnosis of breast cancer using gas chromatography–mass spectrometry," *Clinica chimica acta*, vol. 436, pp. 59–67, 2014.
- [31] M. Gruber, U. Tisch, R. Jerjes, H. Amal, M. Hakim, O. Ronen, T. Marshak, D. Zimmerman, O. Israel, E. Amiga *et al.*, "Analysis of exhaled breath for diagnosing head and neck squamous cell carcinoma: a feasibility study," *British journal of cancer*, vol. 111, no. 4, pp. 790–798, 2014.
- [32] H. Amal, D.-Y. Shi, R. Ionescu, W. Zhang, Q.-L. Hua, Y.-Y. Pan, L. Tao, H. Liu, and H. Haick, "Assessment of ovarian cancer conditions from exhaled breath," *International journal of cancer*, vol. 136, no. 6, pp. E614–E622, 2015.
- [33] U. Navaneethan, M. A. Parsi, D. Lourdasamy, D. Grove, M. R. Sanaka, J. P. Hammel, J. J. Vargo, and R. A. Dweik, "Volatile organic compounds in urine for noninvasive diagnosis of malignant biliary strictures: a pilot study," *Digestive Diseases and Sciences*, vol. 60, pp. 2150–2157, 2015.
- [34] H. Amal, M. Leja, K. Funka, R. Skapars, A. Sivins, G. Ancans, I. Liepniece-Karele, I. Kikuste, I. Lasina, and H. Haick, "Detection of precancerous gastric lesions and gastric cancer through exhaled breath," *Gut*, vol. 65, no. 3, pp. 400–407, 2016.
- [35] J. Pinto, F. Amaro, A. R. Lima, C. Carvalho-Maia, C. Jerónimo, R. Henrique, M. d. L. Bastos, M. Carvalho, and P. Guedes de Pinho, "Urinary volatiles unveils a candidate biomarker panel for noninvasive detection of clear cell renal cell carcinoma," *Journal of Proteome Research*, vol. 20, no. 6, pp. 3068–3077, 2021.
- [36] A. Ulanowska, T. Kowalkowski, E. Trawińska, and B. Buszewski, "The application of statistical methods using vocs to identify patients with lung cancer," *Journal of breath research*, vol. 5, no. 4, p. 046008, 2011.
- [37] Y. Wei, M. Wang, H. Liu, Y. Niu, S. Wang, F. Zhang, and H. Liu, "Simultaneous determination of seven endogenous aldehydes in human blood by headspace gas chromatography–mass spectrometry," *Journal of Chromatography B*, vol. 1118, pp. 85–92, 2019.
- [38] C. G. Waltman, T. A. Marcelissen, and J. G. van Roermund, "Exhaled-breath testing for prostate cancer based on volatile organic compound profiling using an electronic nose device (aeonose™): a preliminary report," *European urology focus*, vol. 6, no. 6, pp. 1220–1225, 2020.
- [39] S. Kort, M. Tiggeloven, M. Brusse-Keizer, J. Gerritsen, J. Schouwink, E. Citgez, F. De Jongh, S. Samii, J. van der Maten, M. van den Bogart *et al.*, "Multi-centre prospective study on diagnosing subtypes of lung cancer by exhaled-breath analysis," *Lung Cancer*, vol. 125, pp. 223–229, 2018.
- [40] W. Li, H. Liu, D. Xie, Z. He, and X. Pi, "Lung cancer screening based on type-different sensor arrays," *Scientific reports*, vol. 7, no. 1, pp. 1–12, 2017.
- [41] B. Liu, H. Yu, X. Zeng, D. Zhang, J. Gong, L. Tian, J. Qian, L. Zhao, S. Zhang, and R. Liu, "Lung cancer detection via breath by electronic nose enhanced with a sparse group feature selection approach," *Sensors and Actuators B: Chemical*, vol. 339, p. 129896, 2021.
- [42] D. Marzorati, L. Mainardi, G. Sedda, R. Gasparri, L. Spaggiari, and P. Cerveri, "A metal oxide gas sensors array for lung cancer diagnosis through exhaled breath analysis," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2019, pp. 1584–1587.
- [43] E. Gashimova, A. Temerdashev, V. Porkhanov, I. Polyakov, D. Perunov, and E. Dmitrieva, "Non-invasive exhaled breath and skin analysis to diagnose lung cancer: Study of age effect on diagnostic accuracy," *ACS omega*, vol. 7, no. 46, pp. 42 613–42 628, 2022.
- [44] H. Uslu, A. Dölle, H. Dulleman, H. Aktas, J. Kolkman, and N. Venne-man, "Pancreatic ductal adenocarcinoma and chronic pancreatitis may be diagnosed by exhaled-breath profiles: a multicenter pilot study," *Clinical and experimental gastroenterology*, pp. 385–390, 2019.
- [45] V. Binson, M. Subramoniam, and L. Mathew, "Detection of copd and lung cancer with electronic nose using ensemble learning methods," *Clinica Chimica Acta*, vol. 523, pp. 231–238, 2021.
- [46] B. VA, M. Subramoniam, and L. Mathew, "Noninvasive detection of copd and lung cancer through breath analysis using mos sensor array based e-nose," *Expert Review of Molecular Diagnostics*, vol. 21, no. 11, pp. 1223–1233, 2021.
- [47] R. Van de Goor, M. van Hooren, A.-M. Dingemans, B. Kremer, and K. Kross, "Training and validating a portable electronic nose for lung cancer screening," *Journal of Thoracic Oncology*, vol. 13, no. 5, pp. 676–681, 2018.
- [48] W. Zhao, L. F. Al-Nasser, S. Shan, J. Li, Z. Skeete, N. Kang, J. Luo, S. Lu, C.-J. Zhong, C. J. Grausgruber *et al.*, "Detection of mixed volatile organic compounds and lung cancer breaths using chemiresistor arrays with crosslinked nanoparticle thin films," *Sensors and Actuators B: Chemical*, vol. 232, pp. 292–299, 2016.
- [49] D. Altomare, F. Porcelli, A. Picciariello, M. Pinto, M. Di Lena, O. Caputi Iambrenghi, I. Ugenti, A. Guglielmi, L. Vincenti, and G. De Genaro, "The use of the pen3 e-nose in the screening of colorectal cancer and polyps," *Techniques in coloproctology*, vol. 20, pp. 405–409, 2016.
- [50] E. Westenbrink, R. P. Arasaradnam, N. O'Connell, C. Bailey, C. Nwokolo, K. D. Bardhan, and J. A. Covington, "Development and application of a new electronic nose instrument for the detection of colorectal cancer," *Biosensors and Bioelectronics*, vol. 67, pp. 733–738, 2015.
- [51] H. Tyagi, E. Daulton, A. S. Bannaga, R. P. Arasaradnam, and J. A. Covington, "Non-invasive detection and staging of colorectal cancer using a portable electronic nose," *Sensors*, vol. 21, no. 16, p. 5440, 2021.
- [52] J. Giró Benet, M. Seo, M. Khine, J. Gumà Padró, A. Pardo Martnez, and F. Kurdahi, "Breast cancer detection by analyzing the volatile organic compound (voc) signature in human urine," *Scientific Reports*, vol. 12, no. 1, p. 14873, 2022.
- [53] S. Kort, M. Brusse-Keizer, H. Schouwink, E. Citgez, F. H. de Jongh, J. W. van Putten, B. van den Borne, E. A. Kastelijn, D. Stolz, M. Schurbiers *et al.*, "Diagnosing non-small cell lung cancer by exhaled breath profiling using an electronic nose: a multicenter validation study," *Chest*, vol. 163, no. 3, pp. 697–706, 2023.
- [54] R. Rocco, R. A. Incalzi, G. Pennazza, M. Santonico, C. Pedone, I. R. Bartoli, C. Vernile, G. Mangiameli, A. La Rocca, G. De Luca *et al.*, "Bionote e-nose technology may reduce false positives in lung cancer screening programmes," *European Journal of Cardio-Thoracic Surgery*, vol. 49, no. 4, pp. 1112–1117, 2016.
- [55] N. Kahn, O. Lavie, M. Paz, Y. Segev, and H. Haick, "Dynamic nanoparticle-based flexible sensors: diagnosis of ovarian carcinoma from exhaled breath," *Nano letters*, vol. 15, no. 10, pp. 7023–7028, 2015.
- [56] N. Shehada, G. Bronstrup, K. Funka, S. Christiansen, M. Leja, and H. Haick, "Ultrasensitive silicon nanowire for real-world gas sensing: noninvasive diagnosis of cancer from breath volatolome," *Nano letters*, vol. 15, no. 2, pp. 1288–1295, 2015.
- [57] T. M. Nguyen, J. H. Chung, G.-H. Bak, Y. H. Kim, M. Kim, Y.-J. Kim, R. J. Kwon, E.-J. Choi, K. H. Kim, Y. S. Kim *et al.*, "Multiarray biosensor for diagnosing lung cancer based on gap plasmonic color films," *ACS sensors*, 2022.
- [58] W. Biehl, A. Hattesoehl, R. A. Jörres, T. Duell, U. Althöhn, A. R. Koczulla, and H. Schmetzer, "Voc pattern recognition of lung cancer: a comparative evaluation of different dog-and enose-based strategies using different sampling materials," *Acta Oncologica*, vol. 58, no. 9, pp. 1216–1224, 2019.
- [59] H. Dalianis and H. Dalianis, "Evaluation metrics and evaluation," *Clinical text mining: secondary use of electronic patient records*, pp. 45–53, 2018.
- [60] G. Gregis, J.-B. Sanchez, I. Bezverkhy, W. Guy, F. Berger, V. Fierro, J.-P. Bellat, and A. Celzard, "Detection and quantification of lung cancer biomarkers by a micro-analytical device using a single metal oxide-based gas sensor," *Sensors and Actuators B: Chemical*, vol. 255, pp. 391–400, 2018.
- [61] Y. Sakumura, Y. Koyama, H. Tokutake, T. Hida, K. Sato, T. Itoh, T. Akamatsu, and W. Shin, "Diagnosis by volatile organic compounds in exhaled breath from lung cancer patients using support vector machine algorithm," *Sensors*, vol. 17, no. 2, p. 287, 2017.
- [62] M. H. Scheepers, Z. Al-Difaie, L. Brandts, A. Peeters, B. van Grinsven, and N. D. Bouvy, "Diagnostic performance of electronic noses in cancer diagnoses using exhaled breath: A systematic review and meta-analysis," *Jama network open*, vol. 5, no. 6, pp. e2219 372–e2219 372, 2022.
- [63] S. Kischkel, W. Miekisch, A. Sawacki, E. M. Straker, P. Trefz, A. Amann, and J. K. Schubert, "Breath biomarkers for lung cancer detection and assessment of smoking related effects—confounding variables,

- influence of normalization and statistical algorithms,” *Clinica Chimica Acta*, vol. 411, no. 21-22, pp. 1637–1644, 2010.
- [64] J. Rudnicka, T. Kowalkowski, T. Ligor, and B. Buszewski, “Determination of volatile organic compounds as biomarkers of lung cancer by spme-gc-tof/ms and chemometrics,” *Journal of Chromatography B*, vol. 879, no. 30, pp. 3360–3366, 2011.
- [65] B. Buszewski, T. Ligor, T. Jezierski, A. Wenda-Piesik, M. Walczak, and J. Rudnicka, “Identification of volatile lung cancer markers by gas chromatography–mass spectrometry: comparison with discrimination by canines,” *Analytical and bioanalytical chemistry*, vol. 404, pp. 141–146, 2012.
- [66] M. Astolfi, G. Rispoli, G. Anania, E. Artioli, V. Nevoso, G. Zonta, and C. Malagù, “Tin, titanium, tantalum, vanadium and niobium oxide based sensors to detect colorectal cancer exhalations in blood samples,” *Molecules*, vol. 26, no. 2, p. 466, 2021.
- [67] L. Huang, Y. Zhu, C. Xu, Y. Cai, Y. Yi, K. Li, X. Ren, D. Jiang, Y. Ge, X. Liu *et al.*, “Noninvasive diagnosis of gastric cancer based on breath analysis with a tubular surface-enhanced raman scattering sensor,” *ACS sensors*, vol. 7, no. 5, pp. 1439–1450, 2022.
- [68] J. Rudnicka, M. Walczak, T. Kowalkowski, T. Jezierski, and B. Buszewski, “Determination of volatile organic compounds as potential markers of lung cancer by gas chromatography–mass spectrometry versus trained dogs,” *Sensors and Actuators B: Chemical*, vol. 202, pp. 615–621, 2014.
- [69] D. Bhattacharyya, P. Kumar, S. K. Mohanty, Y. R. Smith, and M. Misra, “Detection of four distinct volatile indicators of colorectal cancer using functionalized titania nanotubular arrays,” *Sensors*, vol. 17, no. 8, p. 1795, 2017.
- [70] N. A. A. A. Rahman, A. H. Ma’Radzi, and A. Zakaria, “Determination of non-invasive lung cancer biomarker by quartz crystal microbalance coated with pegylated lipopolymer,” in *IOP Conference Series: Materials Science and Engineering*, vol. 458, no. 1. IOP Publishing, 2018, p. 012020.
- [71] G. Gregis, J.-B. Sanchez, S. Schaefer, V. Fierro, F. Berger, I. Bezverkhyy, G. Weber, J.-P. Bellat, and A. Celzard, “Detection of lung cancer bio-markers in human breath using a micro-fabricated air analyzer,” *Materials today: proceedings*, vol. 2, no. 9, pp. 4664–4670, 2015.
- [72] M. Salimi and S. M. R. M. Hosseini, “Smartphone-based detection of lung cancer-related volatile organic compounds (vocs) using rapid synthesized zno nanosheet,” *Sensors and Actuators B: Chemical*, vol. 344, p. 130127, 2021.
- [73] S. Kort, M. Brusse-Keizer, J.-W. Gerritsen, and J. van der Palen, “Data analysis of electronic nose technology in lung cancer: generating prediction models by means of aethena,” *Journal of breath research*, vol. 11, no. 2, p. 026006, 2017.