

DISPOSITIVO DE OLFATO ELECTRÓNICO PARA EL DIAGNÓSTICO PRECOZ DE CÁNCER GÁSTRICO

5

DESCRIPCIÓN DE LA INVENCION

SECTOR TECNOLÓGICO

10

La siguiente invención aporta una herramienta eficaz, económica y efectiva para la detección del cáncer gástrico a través de la detección y análisis del aliento del paciente; por este motivo la invención pertenece al sector de la salud (la medicina), especialmente debido al uso de dispositivos médicos. Esta invención está dentro de la clasificación internacional de patentes con el código A61B5, correspondiente a las medidas encaminadas a establecer un diagnóstico.

ESTADO DE LA TÉCNICA

20 El cáncer de estómago o cáncer gástrico es un tipo de crecimiento celular maligno producido con capacidad de invasión y destrucción de otros tejidos y órganos, en particular el esófago y el intestino delgado, causando cerca de un millón de muertes en el mundo anualmente.

A pesar de la disminución de la incidencia del carcinoma en los últimos años, esta enfermedad todavía es la causa de muerte más común por cáncer en todo el mundo. La incidencia es muy variable en todos los países. Por ejemplo en EUA se presenta una incidencia de 11.1, en Inglaterra de 22.1 y en Japón de 100.2 por 100,000. Lo que demuestra la gran variabilidad que existe entre distintas regiones. A nivel

mundial es la cuarta causa de casos nuevos de cáncer por año según un reporte del año 2000, con 945,000 casos nuevos.

En Colombia, el CG es una enfermedad de alta prevalencia y alta morbilidad, siendo la principal causa de muerte por cáncer en ambos sexos, representado fundamentalmente por el adenocarcinoma de tipo intestinal.

La etiología del CG es multifactorial, en la cual intervienen factores genéticos del individuo, incluyendo antecedentes familiares, factores medioambientales tales como el alto consumo de sal, tabaquismo y la infección por *Helicobacter pylori* (*H.pylori*) cuyo riesgo es variable dependiendo de diversos factores de virulencia (Martínez, J. Revista de Gastroenterología, 2010).

La manera más efectiva para el tratamiento de esta enfermedad, radica en la detección temprana y oportuna de la misma; actualmente el cáncer de estómago se puede detectar en la consulta médica mediante la identificación de signos o síntomas, en este punto la enfermedad ya se encuentra en estado avanzado; si el doctor sospecha de cáncer de estómago autoriza realizar la prueba confirmatoria, es decir, endoscopia con respectiva biopsia.

20

Para optimizar el tiempo de detección se plantea un dispositivo que mediante prueba de aliento permita identificar la presencia o no de cáncer gástrico; BAXTER plantea en la solicitud de patente WO2004006766, "*METHOD AND APPARATUS FOR THE DETECTION OF THE PRESENCE OF A BACTERIA IN THE GASTROINTESTINAL TRACT OF A SUBJECT*", el cual describe un método y un aparato para detectar la presencia de una bacteria en un sujeto. Más particularmente, el método se dirige a la detección de una bacteria que se asocia con catalizar urea en dióxido de carbono y amoníaco cuando está presente en el

tracto gastrointestinal de un sujeto. El método comprende administrar en el sujeto urea, para la obtención de una muestra de fluido del sujeto después de la administración de la urea, posterior a esto determinar la presencia o cantidad de gas amoníaco en la muestra de fluido. Un método algo incómodo para el paciente.

5

En el documento de patente US2003216660 “Breath test apparatus and methods” del grupo de investigación israelí (BEN-OREN ILAN et al) describen aparatos para aumentar la precisión y reducir el tiempo necesario para obtener resultados útiles para el diagnóstico. Con el fin de determinar cuándo hay un aumento clínicamente significativa de la relación isotópica de la respiración exhalada, además se describen métodos para el uso de una variable y el nivel de umbral múltiple; para reducir el tiempo necesario para determinar un nivel de línea de base precisa; y para evitar los efectos de la actividad oral al hacer mediciones. El método de medición de estas muestras es mediante curvas de absorción del espectrómetro de gas. Este es un documento bastante interesante, sin embargo es un equipo complejo y la espectrometría de gases es costosa.

Las anteriores dos soluciones conforman en si el estado de la técnica más cercano a nuestra invención, pero tienen sus limitaciones en cuanto tiempo, comodidad y costo, algo sencillamente vital para los servicios clínicos de cualquier parte del mundo. Por tal motivo, surge la necesidad de desarrollar sistemas más flexibles para que pueda ser usado con mayor facilidad en el sector de la salud.

A partir de algunos trabajos similares desarrollados previamente por el instituto Technion de Israel, donde los investigadores de esa institución han trabajado en el desarrollo de sensores y aplicaciones para la detección de diferentes enfermedades, y el trabajo realizado por Xu, Broza, Ionescu et al, “A nanomaterial-based breath test for distinguishing gastric cancer from benign gastric conditions”, publicado en el British Journal of Cancer (2013). Donde utilizan nanopartículas de

oro para la fabricación de los sensores. Usando como medio de almacenamiento de las muestras Tubos Tenax, para almacenar las muestras, Se llegó a realizar la formulación de la propuesta aterrizada a una problemática en Colombia como es la detección temprana del cáncer gástrico y definiendo nuevas estrategias para la

5 toma de muestras y análisis de la información proveniente de las señales de los sensores. Los sensores basados en previos estudios fueron fabricados en el Instituto de la Universidad de Uppsala Suecia, y en la Universidad Rovira I Virgili de España los cuales posteriormente fueron caracterizados en la universidad de Pamplona Colombia, permitiendo desarrollar el sistema completo con base al

10 concepto de los sistemas de olfato electrónico y acoplando las etapas más relevantes como son: La toma de muestras, adquisición de las señales y procesamiento de la información utilizando técnicas de reconocimiento de patrones, siendo validado el sistema en diferentes fases (prueba con fumadores, y directamente con pacientes de cáncer, previa colaboración de una institución

15 médica colombiana).

DESCRIPCIÓN DETALLADA DE LA INVENCION.

La presente invención se refiere a un dispositivo preventivo visto como un

20 mecanismo de examen rutinario y frecuente, ya que no implica invasión al paciente.

El paciente sopla o deposita su aliento exhalado en un contenedor de aliento (14), este aliento se transfiere a una cámara de sensores (18), es allí donde una serie de sensores químico-resistivos (33) identifican patrones característicos de los

25 volátiles, y los transforma en señales eléctricas, las cuales son adquiridas por medio de un sistema de adquisición de datos (3). Los sensores requieren de unos voltajes de alimentación que son suministrados por la fuente de precisión (4), donde la información es adquirida para ser llevada y almacenada en un computador con fines de realizar las respectivas etapas de procesamiento y análisis de la información (9).

A diferencia del documento del estado de la técnica, todo el proceso de adquisición y tratamiento de datos no toma más de 20 minutos, por lo que no necesita de un procesamiento previo (p.e., pasar la muestra desde el contenedor del aliento a un preconcentrador de volátiles, para luego liberar los volátiles a través de desorción térmica), un equipo de química analítica (p.e. espectrómetro de masas) y no es invasivo para el paciente.

Una vez explicado el proceso general, entraremos a explicar el detalle del proceso y el dispositivo.

10

1- Toma de muestra

El usuario toma el contenedor de aliento (14) por la boquilla (17) exhala su aliento a totalidad, gracias a la presencia de un tapón cheque (13), el aire entra pero no sale de la cámara del contenedor (14), una vez el paciente termine de exhalar, el especialista toma el contenedor (14) y por el tapón inferior (15) inserta el embolo (16) quita el cheque (13) y conecta el contenedor a la cámara de sensores (33) mediante acople (23) y luego empuja lentamente el embolo para vaciar el contenedor. Todo este proceso en un tiempo de 1 a 2 minutos.

La cámara de medición (33) lleva en su interior una serie de sensores (26) con una capa de nano partículas de oro funcionalizadas con distintos materiales orgánicos (35) que representa el material activo del sensor (34), los sensores usados en el dispositivo, presentan un comportamiento tipo resistivo, varían su resistencia en función de los Compuestos Orgánicos Volátiles (COV's) con los que reacciona la capa activa del sensor.

El comportamiento de los sensores depende de la fuente de precisión (4) la cual es una fuente reguladora de voltaje como lo puede ser la B2902A de Keysight, que lo que hace es suministrar un voltaje constante entre las dos terminales de cada sensor (26) como entrada, adquiriéndose los valores de corriente generados por los mismos.

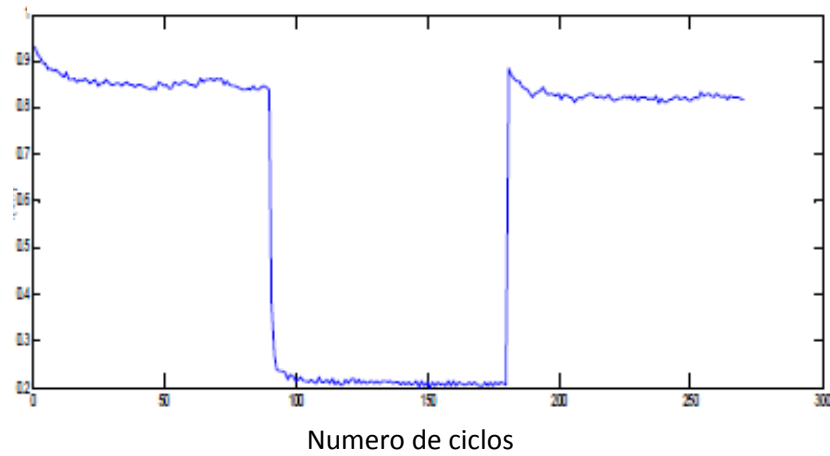
La placa de circuito impreso (PCB) (29) recoge las señales eléctricas de los sensores (26) y se conecta a la unidad que contiene la tarjeta de adquisición (3), esta tarjeta recibe señales eléctricas analógicas de los sensores (26) y las transforma en señales digitales que enviara al computador (9) para su posterior análisis.

2- Adquisición de datos digitales

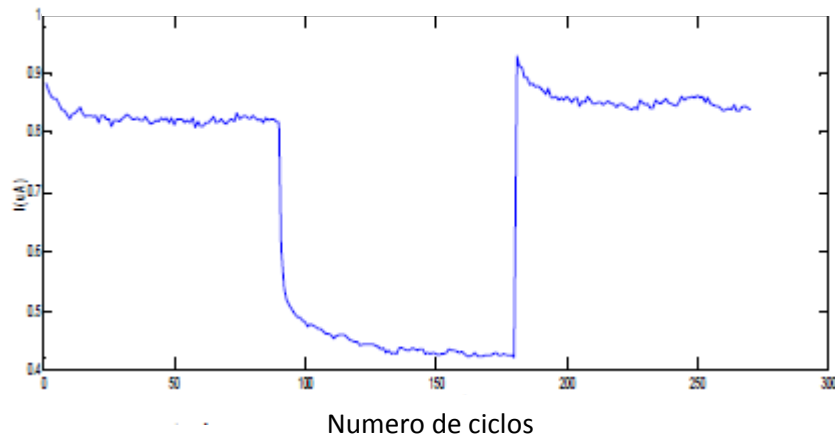
La adquisición de datos de los sensores (26) se realizó en forma secuencial con el equipo de medida 34972A LXI/Data Acquisition (Keysight) es posible conmutar un canal a la vez y con esto se asegura el buen funcionamiento de los sensores.

Se realizó un barrido ascendente de los sensores (26), con un tiempo de adquisición de 10 segundos por sensor, este barrido se repitió cíclicamente, por lo que conocemos esta acción como un ciclo, la cantidad de ciclos necesarios son definidos en el software desarrollado en Labview, el cual nos genera un fichero .txt con los datos de la corriente de cada uno de los sensores.

Respuesta del sensor ante el aliento de un paciente control



Respuesta del sensor ante el aliento de un paciente con Cáncer



5

3- Procesamiento de datos

Una vez la unidad de adquisición (3) suministra las señales (mediante conexión alámbrica tipo BUS de datos (10) o inalámbrica como Bluetooth, Wi-Fi, ZigBee o Li-Fi) y almacenadas en el computador (9) se procede al análisis y procesamiento de la información. Dicho procesamiento consta de tres fases: La normalización de datos, la reducción de la dimensionalidad de la respuesta dada por los sensores (26), y posteriormente una fase de procesamiento y reconocimiento de patrones.

10

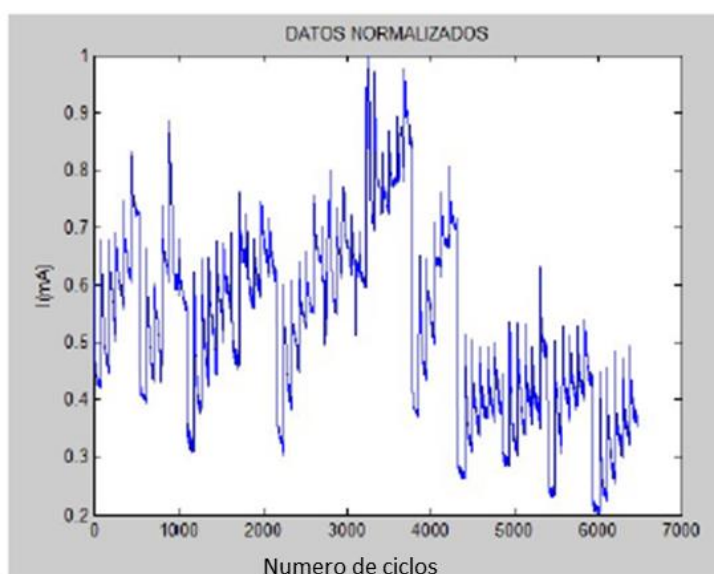
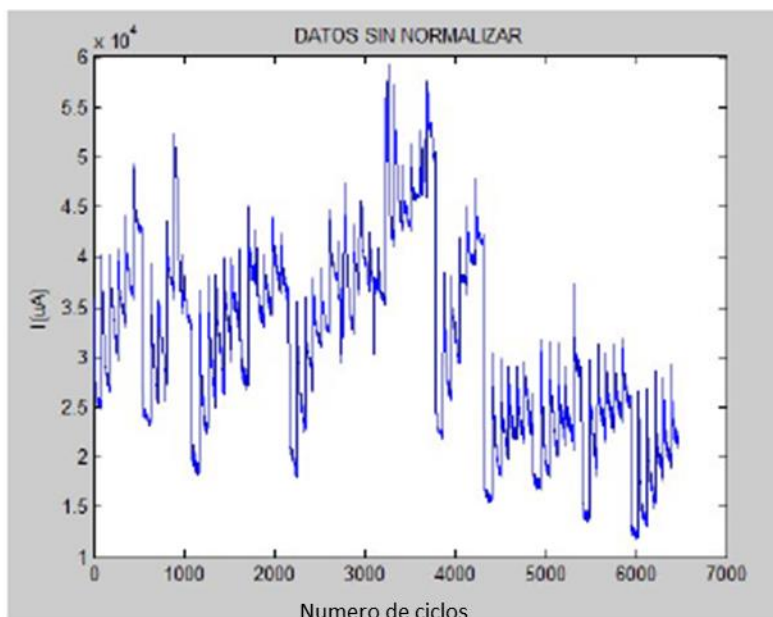
A. Normalización de datos

La normalización consiste en ajustar a una norma, una regla o un modelo común los datos adquiridos. Mediante la normalización es posible estandarizar un conjunto de datos y hacer comparable resultados que no lo son directamente. Al normalizar los datos se buscó que estos estuviesen dentro de un mismo rango y así integrar todas las fases, para obtener un único resultado con todos los datos en conjunto. (Sc, Buena, & Pr, 1970),(García et al., 2014).

- 10 Para normalizar los datos, se buscó el máximo valor dentro del vector de datos de cada fase y se dividió el vector de datos entre dicho valor, visto matemáticamente en la ecuación 1.

$$D_n = \frac{D}{\max(D)} \quad (\text{Ec. 1})$$

D_n : vector de datos normalizados, D : vector de datos, $\max(D)$: máximo del vector de datos.



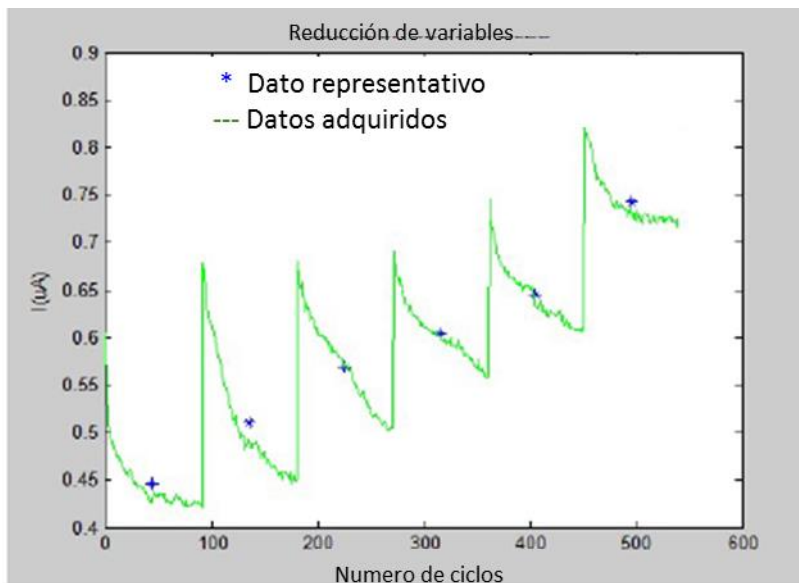
En la figura anterior se pueden apreciar una comparación entre dos señales, la primera sin normalizar y la segunda ya normalizada, verificándose el cambio de rango de valor en el eje vertical, pero manteniendo el comportamiento de la señal original.

La normalización permitió llevar todos los datos a un rango entre [0-1] sin afectar el comportamiento de las señales. El proceso de normalización se realizó con cada una de las fases y posteriormente se unificaron para hacer el análisis en conjunto de todos los datos.

5

B. Reducción de variables

Con la finalidad de facilitar el aprendizaje de los algoritmos de reconocimiento de patrones utilizados y obtener mejores resultados en la clasificación, se realizó una reducción significativa del número de datos con los cuales se entrenan los algoritmos de clasificación.



En la anterior figura se puede apreciar la respuesta de un sensor con los datos ya normalizados, donde se identifican 6 series resultantes de la adquisición en los tiempos definidos.

Para reducir el número de datos se determinó que cada uno de las series se constituiría por un único valor, representando así la señal original de 540 puntos o datos adquiridos con solo 6 valores. Para obtener el valor representativo de cada serie se calculó el promedio de los 90 puntos originales tal y como se muestra en la figura, símbolo color azul.

C. Procesamiento de datos (Análisis cualitativo)

Una vez finalizada la etapa de pre-procesamiento se continuo con la última fase encargada de realizar los procesos de clasificación y reconocimiento de patrones, proporcionando información de los resultados finales en función de porcentajes de clasificación o de acierto de las muestras recopiladas en la toma de datos, ya sea de forma no supervisada como de forma supervisada.

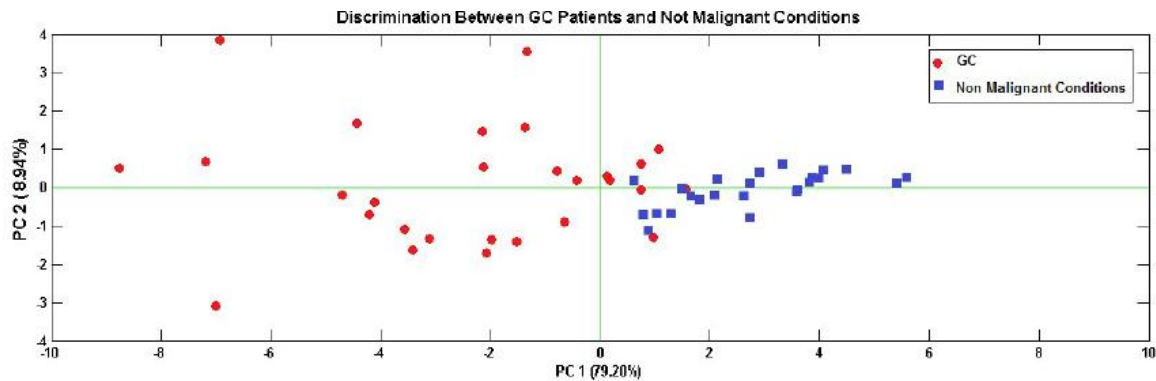
i. Análisis no supervisado (PCA)

El método estadístico de reconocimiento de patrones PCA es una técnica de uso frecuente que se basa en el análisis multivariado; es un algoritmo lineal no supervisado en el que se proyectan los datos de varios sensores en un plano de dos o tres dimensiones (Cadima, 2004). El método PCA determina los componentes que mejor representan los datos de acuerdo con el error mínimo cuadrado. El algoritmo PCA se ejecuta con una interfaz de usuario dando la posibilidad de seleccionar los sensores y las categorías que se muestran en forma gráfica, lo que permite obtener las agrupaciones deseadas en forma rápida.

Una vez normalizados los datos descritos en la sección de pre-procesamiento, se aplica la técnica PCA con centrado de datos, la cual discrimina los dos grupos de estudio, Pacientes CA (Cáncer Gástrico)- Pacientes con gastritis (Control), resaltando la importancia de su uso, ya que sin necesidad de dar a conocer

previamente la clase a la cual pertenece la muestra, ésta la agrupa en el clúster correspondiente.

- Es importante mencionar que los resultados analizados con PCA fueron aplicados a las variables originales proveniente de las señales suministradas por los sensores, donde se obtuvo como mejor respuesta o resultado el siguiente gráfico.



- En la figura anterior se aprecia que las muestras simbolizadas en color rojo, corresponden a pacientes con cáncer gástrico, y las muestras color azul a pacientes control.

- En la gráfica anterior se representan dos muestras por paciente. Este resultado demostró la buena repetitividad y selectividad del sistema al momento de la realización de las diferentes pruebas.

ii. Análisis Supervisado (Métodos de reconocimiento de patrones)

Como paso posterior al análisis no supervisado empleando técnicas multivariable como son los análisis de componente principales (PCA), con resultados prometedores, se procedió a emplear otras técnicas supervisadas de clasificación y reconocimiento de patrones específicamente las redes neuronales y las máquinas de soporte vectorial (SVM), (Chen, Yuan, He, & Luo, 2016),(Linn et al., 2016) para determinar los porcentajes de acierto en la clasificación entre las dos categorías determinadas (Cáncer Gástrico y control). Las diferentes técnicas utilizadas tomaron como referencia los datos obtenidos del proceso de pre-procesamiento, es decir, luego de pasar por las etapas de normalización y reducción de variables, explicados anteriormente. La plataforma utilizada para la ejecución de los modelos clasificadores fue MATLAB versión 2014.

Inicialmente se realizó el uso de las redes neuronales como estrategia de clasificación, estas redes se inspiran en la estructura y funcionamiento de los sistemas nerviosos, en la actualidad son altamente utilizadas como modelos alternos de cómputo para solución de problemas complejos (Ricardo, Patricia, Gerardo, Francisco, & Arturo, 2016).

- Clasificación con redes neuronales perceptrón multicapa (MLP) (Rezai, Keshavarzi, & Mahdiye, 2014), función de base radial (RBF) (Science, Phenomena, Guan, Zhu, & Song, 2016), red neuronal probabilística (PNN) (Iounousse, Er-raki, Motassadeq, & Chehouani, 2015), empleando en la fase de entrenamiento la totalidad de las muestras, dando un resultado del 100 % de clasificación, pero es importante mencionar, que dicho resultado no es conveniente tomarlo como referencia ya que el hecho de entrenar y validar con la misma información no nos da confiabilidad en el momento de analizar los resultados.

- Clasificación con las mismas redes neuronales anteriores pero realizando una distribución por partes, es decir, un setenta por ciento de los datos fueron empleados para entrenar los modelos neuronales y el 30 % restante

utilizados para la validación. Para tener un mayor uso de esta distribución se crearon 4 bloques, aplicando a cada uno de ellos la técnica en mención, seleccionando de forma aleatoria las muestras para el 70 y 30 % respectivamente, dando resultados entre el 82.27 y el 92.3 % por ciento de acierto, dando en promedio una relación de 12 aciertos y un fallo para el mejor de los casos, el cual fue utilizando el modelo neuronal con topología PNN.

- Se realizó una prueba uniendo la técnica PCA, como estrategia de reducción de la dimensión de las variables, para luego realizar la técnica por partes mencionada anteriormente, dando un resultado no tan deseado llegando al 76.9 % de clasificación.
- Otra de las estrategias empleadas para evaluar la clasificación, es la técnica conocida como *leave one out* (dejar uno por fuera) (Series & Techniques, 2013) la cual consiste en generar un ciclo de evaluación en donde secuencialmente se va dejando una muestra del conjunto a evaluar por fuera y con las restantes se realiza el proceso de entrenamiento, para luego evaluar y validar el modelo entrenado con la muestra retirada, realizando n ciclos similares según el número de muestras totales siendo al final validadas todas las muestra retiradas, dando un porcentaje de acierto en función de la totalidad de dichas muestras, en nuestro caso, del total de 52 muestras, el ciclo de leave-one-out se realizó 52 veces, evaluándose en cada uno de los casos con las diferentes redes neuronales definidas (MLP,RBF,PNN). Evaluando todas las posibilidades el mejor resultado se obtuvo con la red neuronal probabilística con un porcentaje de acierto de 92.3%, con 48 aciertos de las muestras correspondientes a cada una de las dos clases, por consiguiente, 4 fallos del total de 52 muestras.

En recientes años, el uso de las máquinas de soporte vectorial se ha convertido en una herramienta ideal en los procesos de clasificación, generando resultados en muchos casos mejores que los obtenidos con las mismas redes neuronales (Kazemian, Adnan, White, & Maxime, 2016), ello es debido a factores como

5 optimizar el mínimo riesgo estructural, entre otros, a través del manejo y control de elementos como la función kernel o el parámetro de costo C , en otras palabras, un SVM permite construir hiperplanos en un espacio de dimensionalidad que puede ser utilizado en problemas de clasificación o regresión.

10 En nuestro caso se decidió explorar esta técnica e intentar mejorar los resultados obtenidos con las diferentes redes neuronales implementadas, evaluando en cada caso diferentes modelos SVM, variando en cada uno de ellos las diferentes opciones de funciones kernel (Polinomial, lineal, gaussiano, spline, entre otros).

15 Inicialmente se realizó la misma estrategia de distribución de la información por partes usadas con las redes neuronales (70 para entrenar y 30 para validar), empleando en este caso, como modelo de clasificación las máquinas de soporte vectorial con una función kernel polinomial y un valor de C de 1000, generando unos resultados promedio de 92.3 % de clasificación, es decir, 12 muestras con aciertos

20 y una muestra no clasificada correctamente.

Finalmente, se dejó como ultima estrategia por mostrar la que se consideró como el mejor resultado obtenido una vez valoradas todas las estrategias implementadas, ella es el uso de las máquinas de soporte vectorial en acople a la técnica de

25 evaluación leave-one-out utilizada igualmente con las redes neuronales y que fue explicada anteriormente, en este caso se logró llegar a un porcentaje de acierto en la clasificación del 94.23 %, llevando a un total de 49 muestras bien clasificadas y únicamente 3 muestras no clasificadas correctamente.

Según referencias tomadas de los expertos el porcentaje patrón definido para la valoración por endoscopia, es cercano al 95 por ciento, si se compara con los obtenidos con los modelos evaluados, específicamente con los SVM entrenados empleando leave-one-out (94.23 %) se puede concluir que los resultados con el sistema desarrollado de detección de cáncer gástrico empleando el sistema multisensorial es prometedor y podrá seguramente consolidarse como estrategia paralela a las técnicas utilizadas en la actualidad.

La tabla, representa de forma resumida y sintetizada los resultados explicados y analizados anteriormente.

Técnica supervisada	Porcentaje de clasificación
Red MLP entrenamiento por partes (distribución 4 bloques)	82.27%
Red RBF entrenamiento por partes (distribución 4 bloques)	90.38%
Red PNN entrenamiento por partes (distribución 4 bloques)	92.30%
PCA unido PNN por partes	76.9%
Leave one out red MLP	88.46%
Leave one out red RBF	90.38%
Leave one out red PNN	92.30%
SVM entrenamiento por partes (kernel polinomial)	92.30%
SVM entrenamiento leave one out (kernel lineal)	94.23 %

4- Limpieza del dispositivo

Una vez finalizada la toma de una muestra y el uso del dispositivo cada día, es importante realizar la limpieza de la cámara de sensores (18) con un flujo de aire sintético. Se utiliza aire sintético que mantiene aproximadamente un 21% de oxígeno y de 79% de nitrógeno, lo cual garantiza la recuperación y estabilidad de

los sensores de la cámara de medidas. Para este procedimiento se cuenta con un cilindro de aire sintético (7), un regulador de flujo (5) y de presión (6) y una manguera de limpieza (8) que se conecta a la cámara (18). Se recomienda que este proceso se haga en un tiempo de 5 minutos.

5

Detalles de los sensores

La fabricación de los sensores (26) está conformado por dos fases la primera que se encarga de asegurar la conducción eléctrica en la cual se emplea la técnica “advanced gas deposition (AGD) para depositar nano partículas dispersas de oro (capa de oro (35)) (AuNPs) ultra puras, y como segundo paso un proceso “dip coating” para la funcionalización de los AuNPs con distintos compuestos orgánicos (Capa de sustrato base (36) – Patente ES2566067A1

Las capas depositadas de AuNPs fueron funcionalizadas en los sensores con varias moléculas ligadas de (4-methoxy- α -toluenethiol, methyl-3-mercaptopropionate, y 1-decanethiol, entre otras). (Los ligandos fueron seleccionados basados en estudios previos, “Xu, Broza, Ionescu et al, “A nanomaterial-based breath test for distinguishing gastric cancer from benign gastric conditions”, publicado en el British Journal of Cancer, 2013). El material orgánico reacciona físico-químicamente con los compuestos del aliento, produciendo un cambio en el valor de la resistencia del sensor).

Las nanopartículas de oro modificadas molecularmente han demostrado un excelente potencial para detectar concentraciones muy bajas de COV en muestras biofluidas.

Los sensores capacitivos pueden ser, QCM, QMB, cantilevel, FET, diodos Schottky, u otro tipo se componen de una monocapa de nanopartículas de oro enlazadas mediante compuestos orgánicos ligados (34) (35): nanotubos de carbono, óxidos metálicos, nanohilos, grafeno y encima de una capa de sustrato base (36) sobre una capa de oro o cobre o platino o cualquier otro material (37) en una matriz de silicio, vidrio, alumina, cualquier otro tipo sustrato (38).

Ventajas del dispositivo

Como se mostró anteriormente, este dispositivo cuenta con un diseño exclusivo y eficiente para la detección del cáncer gástrico, con una precisión mínima del 96%, solo necesita del aliento exhalado del paciente, por lo que lo convierte en un método no invasivo e indoloro, el tiempo de análisis por muestra no es mayor a 20 minutos, es portable dado que no pesa más de 25 kilos, siendo la bala de aire sintético la que mayor peso conlleva, lo que nos permitirá llegar a regiones de difícil acceso y no necesita herramientas computacionales complejas, lo que convierte en un dispositivo accesible a la comunidad médica.

DESCRIPCIÓN DE LAS FIGURAS

Fig 1. Procedimiento operativo del dispositivo de detección precoz de cáncer gástrico

Fig 2. Vista completa del dispositivo de detección.

Fig 3. Detalle del contenedor de aliento (14), se puede observar cómo se acopla con el tapón cheque (13) y con el embolo en su extremo inferior (16)

Fig 4. Detalle de la cámara de sensores o cámara de concentración (18) se observa los ductos de entrada y la salida de volátiles (19) y (20)

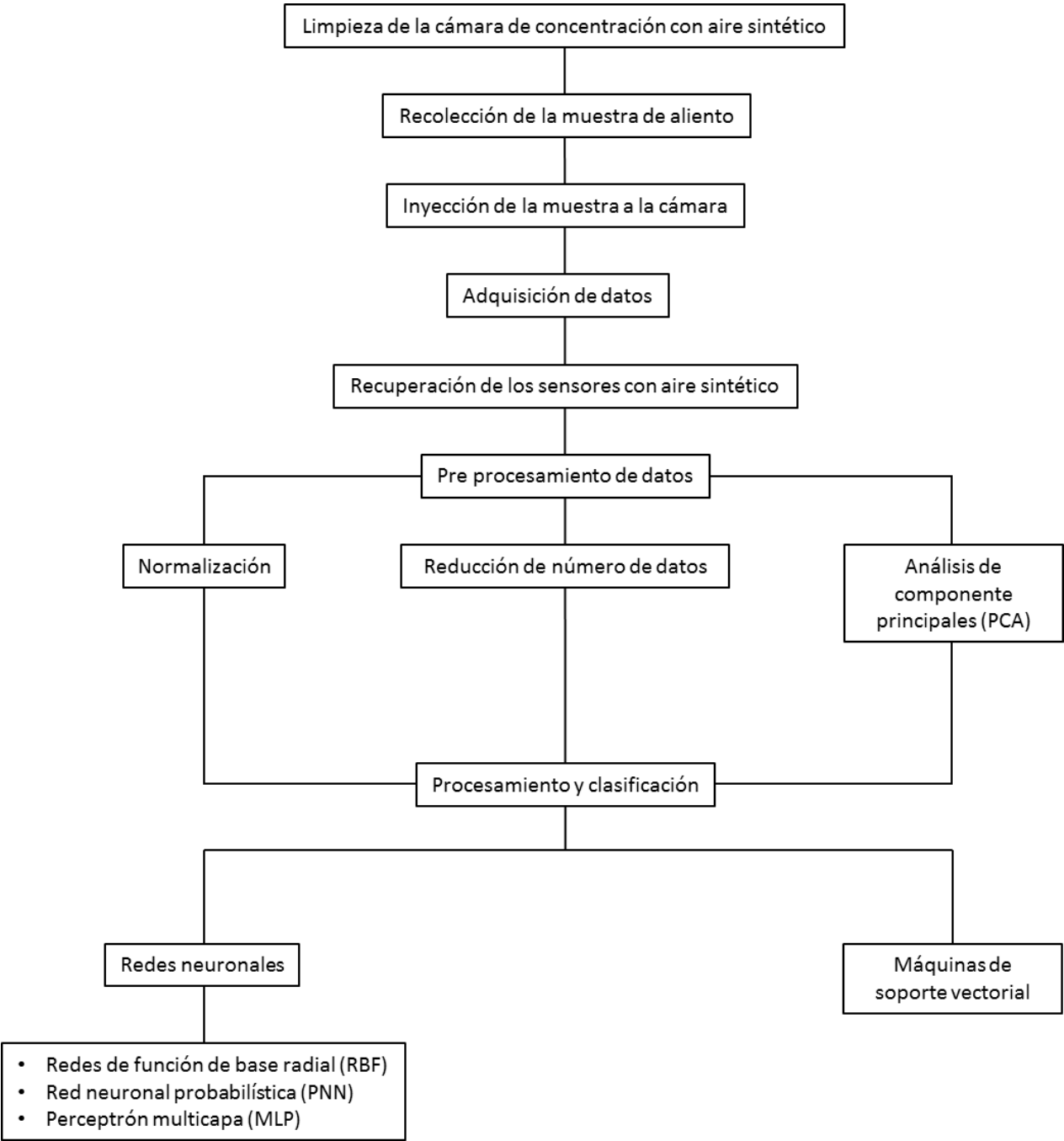
Fig 5. Despiece de la cámara de concentración (18), se observa los sensores (26) y el circuito de recolección (29), para este diseño inicial se aprecian tornillos (22, 30, 32) y elementos constitutivos de la cámara (24,25, 27, 31)

Fig 6. Sensor (26)

Fig 7. Despiece del sensor (26) y sus diferentes capas.

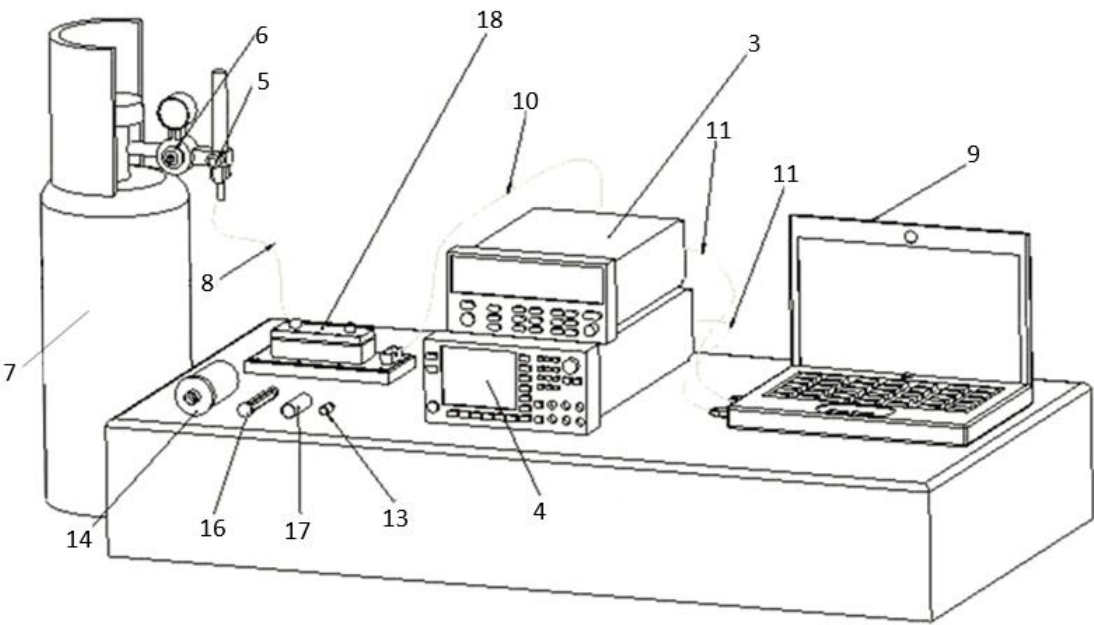
FIGURAS

Fig 1.



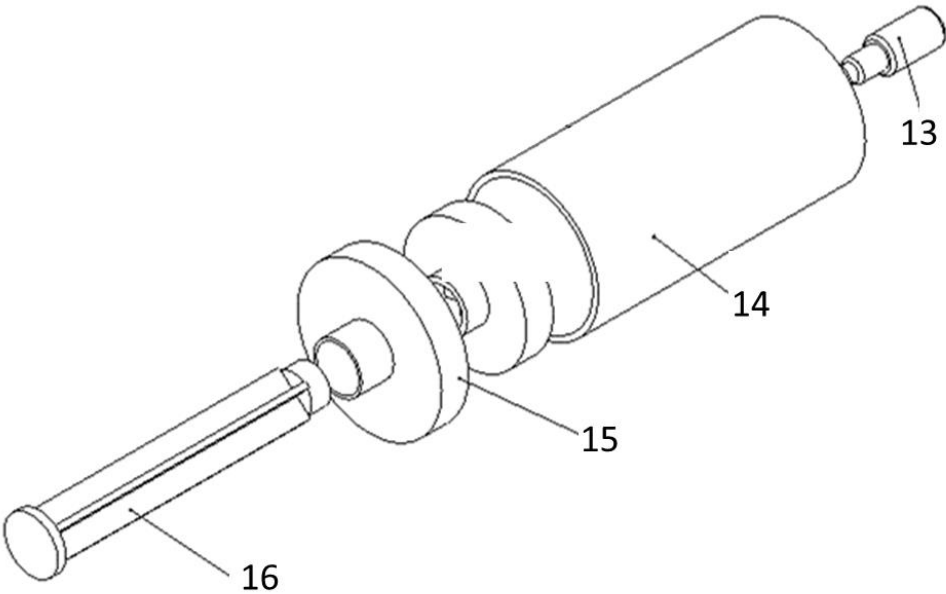
5

Fig 2.



10

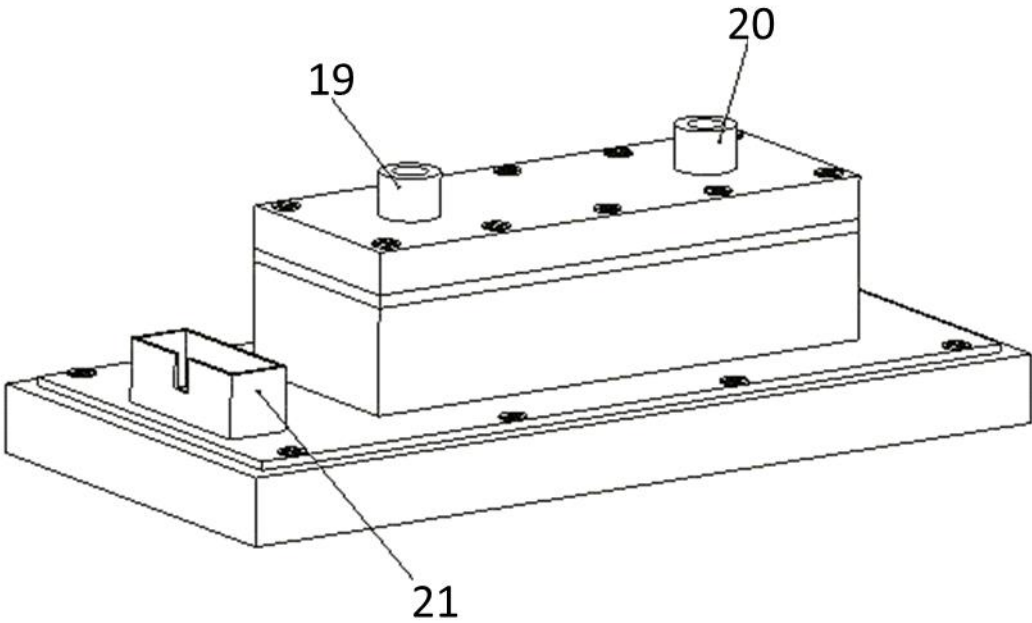
Fig 3.



5

10

Fig 4.



5

10

Fig 5.

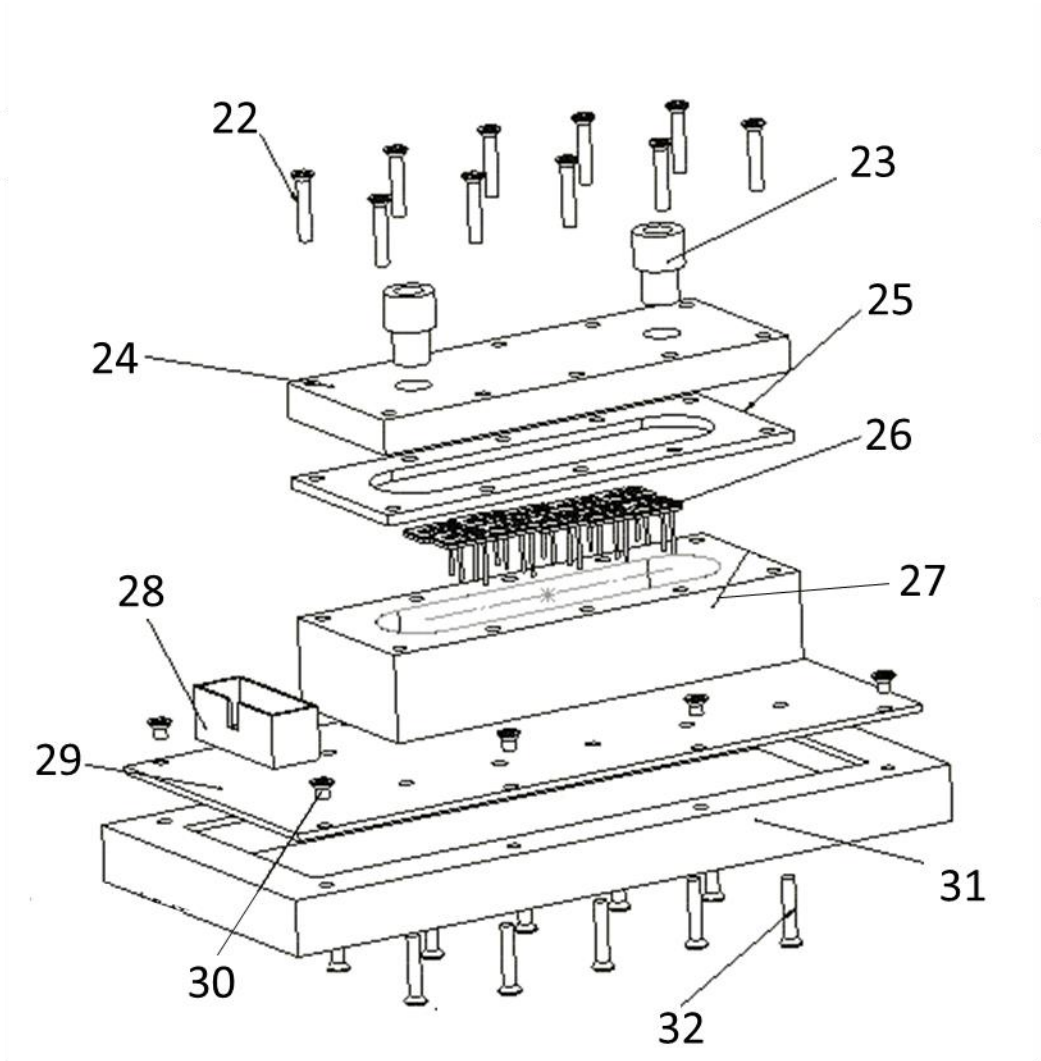
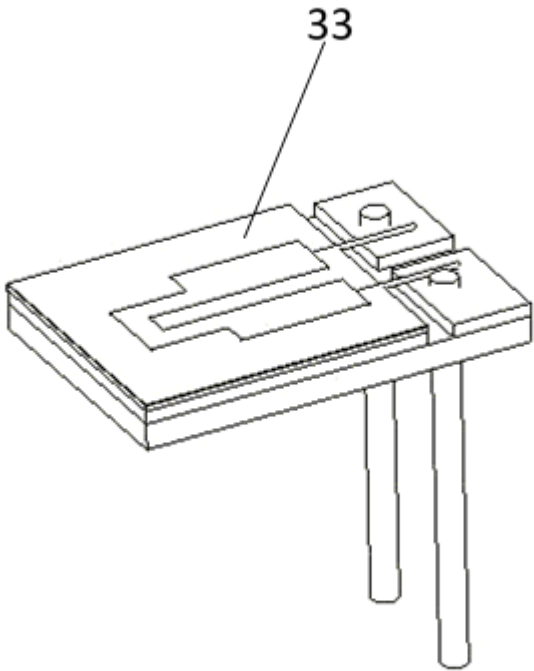


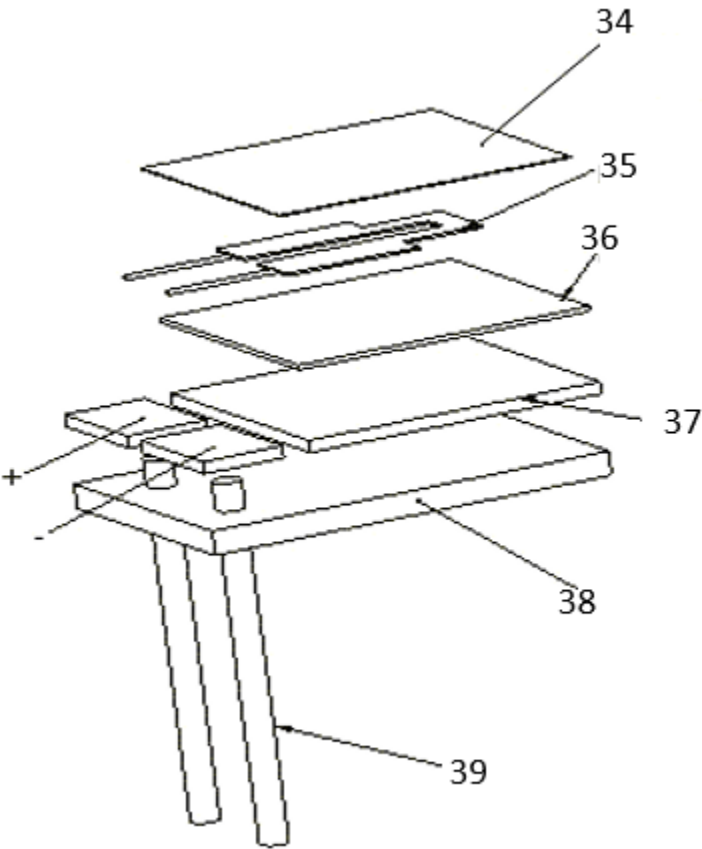
Fig 6.



5

10

Fig 7.



5

10