

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **3 001 012**

51 Int. Cl.:

G01N 15/14

(2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **28.03.2019** **PCT/FR2019/050718**

87 Fecha y número de publicación internacional: **03.10.2019** **WO19186073**

96 Fecha de presentación y número de la solicitud europea: **28.03.2019** **E 19719349 (3)**

97 Fecha y número de publicación de la concesión europea: **23.10.2024** **EP 3775840**

54 Título: **Método de detección de células que presentan al menos una anomalía en una muestra citológica**

30 Prioridad:

29.03.2018 FR 1800251

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

04.03.2025

73 Titular/es:

**OFFICE NATIONAL D'ETUDES ET DE
RECHERCHES AÉROSPATIALES (100.00%)
Chemin de la Hunière
91120 Palaiseau, FR**

72 Inventor/es:

**HADDAD, ZEHIRA;
HERBIN, STÉPHANE y
CHAN-HON-TONG, ADRIEN**

74 Agente/Representante:

ARIAS SANZ, Juan

ES 3 001 012 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Método de detección de células que presentan al menos una anomalía en una muestra citológica

5 La presente invención se refiere a un método de detección de células que presentan al menos una anomalía en una muestra citológica, a partir de una imagen digitalizada o de varias imágenes digitalizadas de la misma muestra. La invención también se refiere a un dispositivo que pone en práctica este método.

10 La invención se refiere al campo de la citología, que es el estudio por microscopía de las células de una muestra citológica tras la fijación sobre un portaobjetos. Más precisamente, se trata del estudio de los aspectos morfológicos, radiométricos o bioquímicos de células normales o atípicas, es decir, que presentan al menos una anomalía. En el documento US 5 933 519 A se da a conocer un ejemplo.

15 El cribado del cáncer es uno de los campos de aplicación de la citología, y consiste en buscar y poner en evidencia un cáncer con la ayuda de tomas de muestras citológicas de un sujeto. El cribado se realiza en parte en un laboratorio de anatomopatología por citotécnicos y citopatólogos con formación en los métodos de detección de células atípicas. La citología se usa mucho para la detección de determinados cánceres, ya que permite diagnosticarlos en diferentes estadios, teniendo algunos cánceres, por ejemplo, un estadio precoz denominado de "grado bajo" y un estadio más avanzado denominado de "grado alto" (el grado también puede corresponder a un tipo de cáncer). Resulta evidente que el diagnóstico precoz de un cáncer es un gran desafío con importantes repercusiones asociadas a la tasa de supervivencia de los sujetos afectados y también al coste correspondiente para curar a estos últimos. Por ejemplo, cada año en el contexto del cáncer de vejiga, se realizan 10 millones de citologías de orina en todo el mundo para ayudar a diagnosticarlo, siendo una práctica habitual el análisis de portaobjetos de citología de muestras procedentes de micciones espontáneas para diagnosticar un cáncer de vejiga.

25 No obstante, este método de diagnóstico mediante análisis de muestras citológicas está limitado, por un lado, por el aspecto humano, es decir la experiencia y el número decreciente de citopatólogos y de citotécnicos, y, por otro lado, por una baja sensibilidad para determinados cánceres. Normalmente, existe una sensibilidad muy baja para el diagnóstico del cáncer de grado bajo de vejiga ya que las células atípicas o anómalas aún no han experimentado deformaciones morfológicas pronunciadas y, por tanto, son difíciles de detectar a simple vista. De ahí la necesidad de buscar procedimientos utilizables de manera rutinaria y poco costosos que permitan aumentar el rendimiento y también la fiabilidad del conjunto de esta práctica.

35 A partir de esta situación, uno de los objetivos de la invención consiste en proporcionar un método de detección de células que presentan al menos una anomalía celular en una muestra citológica que permita ayudar a continuación a la detección de cáncer, concretamente del cáncer de vejiga, a partir de imágenes de portaobjetos digitalizadas. La invención permite ventajosamente, con la ayuda de un módulo de procesamiento puesto en práctica por ordenador, un análisis automatizado y exhaustivo de las células de la muestra presente en el portaobjetos. Por tanto, el análisis de la muestra citológica y de todas las células posiblemente anómalas es completo, rápido, más preciso y permite aliviar los errores de análisis habituales de los citopatólogos y de los citotécnicos. Además, este método garantiza una normalización del resultado del análisis citológico de la muestra. Adicionalmente, este método permite además ayudar al diagnóstico precoz de cáncer denominado de grado bajo, concretamente del cáncer de grado bajo de vejiga, detectando células que presentan anomalías bajas.

45 La invención pretende resolver los inconvenientes del estado de la técnica citados anteriormente y está motivada por la ausencia de soluciones alternativas. Para alcanzar este u otros objetivos, un primer aspecto de la invención propone un método de detección de células que presentan al menos una anomalía en una muestra citológica a partir de al menos una primera imagen de microscopía electrónica digitalizada de la muestra, según la reivindicación 1. El método se ejecuta con la ayuda de un módulo de procesamiento puesto en práctica por ordenador, en concreto totalmente ejecutado por el módulo de procesamiento puesto en práctica por ordenador.

El método comprende una etapa de procesamiento de la primera imagen, que comprende las siguientes subetapas:

55 - a partir de la primera imagen, detectar cada elemento presente en la muestra, elegido de al menos una célula aislada o un grupo de células, con la ayuda de una detección colorimétrica,

- filtrar cada célula aislada y/o cada grupo de células con la ayuda de la detección colorimétrica,

60 - a partir de cada grupo de células detectado, individualizar al menos una célula con su núcleo, con la ayuda de al menos un método de segmentación basado en una detección de los contornos en el grupo de células,

Adicionalmente, el método comprende además una etapa de identificación de anomalía celular, que comprende las siguientes subetapas:

65 - para cada célula aislada o individualizada en la etapa anterior, calcular datos de al menos dos características de anomalías distintas denominadas "características expertas de anomalía", y determinar si dicha célula presenta al

menos una anomalía, comparando estos datos calculados con datos de referencias,

- después contar el número total de células en la muestra que presentan al menos una anomalía,

5 Y finalmente, comprende además una primera etapa de caracterización de la muestra según un nivel de anomalía de la muestra, que comprende las siguientes subetapas:

- comparar el número total de células que presentan al menos una anomalía con un valor umbral del nivel de anomalía, preferiblemente un valor umbral igual a al menos tres células que presentan, cada una, al menos una anomalía,

10 - si el número total de células que presentan al menos una anomalía es inferior al valor umbral del nivel de anomalía, clasificar la muestra en una clase representativa de una muestra con un nivel de anomalía nulo,

15 - si el número total de células que presentan al menos una anomalía es superior al valor umbral del nivel de anomalía, clasificar la muestra en una clase representativa de una muestra con un nivel de anomalía demostrado, es decir un nivel de anomalía bajo o intenso.

20 Según este aspecto de la invención, el método de detección clasifica, de manera fina, la muestra citológica según un nivel de anomalía celular a partir del análisis de al menos una primera imagen, preferiblemente con una segunda imagen. El módulo de procesamiento usado en la invención es adecuado para combinar datos calculados a partir de la primera imagen de microscopía electrónica de transmisión digitalizada y/o de una segunda imagen de microscopía de fluorescencia digitalizada de la muestra con datos de referencias conocidos en citología.

25 Se comprenderá que durante la etapa de procesamiento de la primera imagen citada anteriormente, se detectan varios elementos constituyentes de la primera imagen, concretamente células aisladas, grupos de células, elementos a rechazar o agrupaciones de células no analizables y que, por tanto, se apartan para la correcta continuación del método.

30 Esta solución permite beneficiarse de la clasificación de la muestra según dos clases y diagnosticar si la muestra tiene un nivel de anomalía celular nulo o promediado. A continuación, este resultado permitirá ayudar significativamente al diagnóstico de un cáncer o no por un especialista.

35 Por otro lado, se observará que, según la invención, una característica experta de anomalía de una célula puede considerarse como un buen indicador de la conformidad de una célula.

Preferiblemente, una de las características expertas de anomalía es una razón de superficie del núcleo en la célula o una irregularidad del núcleo o una colorimetría moteada o una hipercolorimetría del núcleo. Por colorimetría moteada se entiende la medición de la intensidad de coloración de determinados puntos o de manchas.

40 Más preferiblemente, la característica experta de anomalía es una característica del núcleo, en particular, una razón de tamaño del núcleo con respecto al tamaño de la célula, un posicionamiento del núcleo en la célula, o una característica de colorimetría del núcleo. Más precisamente, la característica experta de anomalía es una razón de superficie del núcleo en la célula o una irregularidad del núcleo o una colorimetría moteada o una hipercolorimetría del núcleo.

45 Ventajosamente, al menos una de las características expertas de anomalía es una irregularidad del núcleo, el cálculo de dicha irregularidad del núcleo es un cálculo de desviación correspondiente a la comparación entre cada núcleo de célula analizada y un núcleo tipo denominado de referencia de una célula denominada de "referencia" en citología. Este cálculo de desviación se elige de la siguiente lista: una diferencia de superficie, una diferencia de perímetro, una diferencia de la convexidad del núcleo, una diferencia de baricentro, una diferencia entre una longitud del contorno de cada núcleo detectado, una diferencia de irregularidad de un contorno del núcleo, una diferencia del número de píxeles en cada núcleo y una diferencia de excentricidad del núcleo en la célula.

50 Según una realización particular, al menos una característica experta de anomalía es una longitud del contorno del núcleo, el cálculo de esta característica comprende además las siguientes subetapas:

- calcular un contorno real del núcleo usando un método de extracción de contorno del núcleo,

60 - determinar un diámetro del núcleo, preferiblemente el diámetro más grande, trazando una envolvente de recubrimiento del núcleo,

- trazar una curva convexa cerrada de rango C2 cuya envolvente de recubrimiento es la de un núcleo de referencia,

65 - calcular una diferencia entre el contorno real del núcleo y la curva convexa trazada, que representa el contorno de un núcleo de referencia.

Según una realización particular, al menos una característica experta de anomalía es una irregularidad de un contorno del núcleo, el cálculo de esta característica comprende además las siguientes subetapas:

- 5 - detectar al menos un punto anguloso en el contorno del núcleo detectado,
- calcular un número total de puntos angulosos en el contorno del núcleo,
- medir un ángulo para cada punto anguloso correspondiente, después
- 10 - determinar un grado de irregularidad del contorno del núcleo combinando los datos del número de puntos angulosos detectado por núcleo y sus ángulos.

Según la invención, el método comprende además una etapa de selección tipológica, que comprende las siguientes subetapas:

- 15 - a partir de la primera imagen digitalizada de la muestra, medir el tamaño de las células, después clasificar las células según su tamaño,
- asociar cada clase de células a un tipo celular correspondiente, después seleccionar al menos un tipo celular elegido correspondiente a un tamaño de célula, tal como el tipo celular urotelial.
- 20

Según esta realización particular, la invención puede comprender además una segunda etapa de caracterización de la muestra según el nivel de anomalía de las células de la muestra. Esta segunda etapa comprende además las siguientes subetapas:

- 25 - a partir de una segunda imagen de fluorescencia digitalizada de la muestra, seleccionar todas las células de un tipo celular elegido, tal como urotelial, según su ubicación en la muestra predefinida previamente en la primera imagen digitalizada,
- 30 - detectar la presencia de un halo de fluorescencia alrededor de cada célula seleccionada anteriormente, preferiblemente medir una tasa de fluorescencia, más preferiblemente medir una tasa de dispersión de la fluorescencia,
- si se detecta un halo de fluorescencia, preferiblemente si se mide una tasa mínima de fluorescencia, más preferiblemente se mide una tasa mínima de dispersión de la fluorescencia, clasificar la muestra en una clase representativa de una muestra con un nivel de anomalía bajo.
- 35

Según otra realización particular, la invención comprende una etapa de recomendación de explotabilidad de la muestra según un número de células que comprenden un núcleo y según un valor umbral mínimo de explotabilidad, preferiblemente el valor umbral mínimo de explotabilidad es igual a al menos siete células que comprenden un núcleo en la primera imagen.

- 40
- 45 Según la invención, el método, con la ayuda del módulo de procesamiento, combina el conjunto de los datos obtenidos y clasifica la muestra citológica analizada a partir de la primera imagen y de la segunda imagen, en una de las tres clases asociadas al nivel de anomalía celular, nulo, bajo o intenso.

Un segundo aspecto de la invención propone un método de detección de las células que presentan al menos anomalía en una muestra citológica a partir de una primera imagen y/o de una segunda imagen con la ayuda de un método de aprendizaje profundo, con la ayuda de un módulo de procesamiento puesto en práctica por ordenador, para asignar la muestra a una de las tres clases asociadas a un nivel de anomalía celular de la muestra, nulo, bajo o intenso.

- 50
- Este método es ventajoso ya que usa los datos proporcionados por al menos una de las variantes del primer aspecto de la invención.

- 55 Preferiblemente, la muestra citológica es una muestra de orina de micción espontánea.

Un tercer aspecto de la invención propone un dispositivo que pone en práctica un método tal como se describió anteriormente.

- 60 Un cuarto aspecto de la invención propone un uso del método tal como se describe a continuación para ayudar al diagnóstico del cáncer de vejiga.

Otras particularidades y ventajas de la presente invención se desprenderán de la siguiente descripción de ejemplos de realización no limitativos, haciendo referencia a los dibujos adjuntos en los que:

- 65 - la figura 1 es un esquema que ilustra las diferentes etapas del método puesto en práctica según un modo preferido

de la invención,

- la figura 2 representa un primer ejemplo de una primera imagen de microscopía electrónica digitalizada de una muestra antes y después de una etapa de procesamiento. La primera imagen comprende en este caso, concretamente, células aisladas, agrupaciones celulares, grupos de células y elementos a rechazar,

- la figura 3 representa un segundo ejemplo de una primera imagen de microscopía electrónica digitalizada de una segunda muestra antes y después de una etapa de procesamiento de la primera imagen. La primera imagen comprende en este caso, concretamente, células aisladas, agrupaciones celulares, grupos de células y elementos a rechazar.

- La figura 4 representa dos imágenes de microscopía electrónica de una célula aislada y dos imágenes de microscopía electrónica de sus núcleos extraídos con la ayuda de un módulo de procesamiento puesto en práctica por ordenador,

- la figura 5 representa imágenes de microscopía electrónica de una célula aislada, de su núcleo extraído y de un ejemplo de trazo para medir la razón de tamaño de núcleo-célula, con la ayuda del módulo de procesamiento puesto en práctica por ordenador,

- la figura 6 es un esquema que ilustra las diferentes etapas del método puesto en práctica según otra realización según la cual comprende una segunda etapa de caracterización de la muestra a partir de una segunda imagen digitalizada,

- la figura 7 es un esquema que ilustra un segundo aspecto de la invención, más precisamente una metodología de implementación para un aprendizaje profundo aplicado a la clasificación de al menos una imagen de muestra citológica correspondiente a un conjunto de imágenes reducidas.

En primer lugar, debe indicarse que las figuras exponen la invención de manera esquemática. Estas figuras representan ejemplos de realización facilitados a título no limitativo y no respetan necesariamente una cronología particular.

La presente invención se refiere al contexto general de la aplicación de la visión por ordenador de imágenes digitalizadas de objetos complejos y numerosos. La invención se refiere a un método original para detectar las células que presentan al menos una anomalía en una muestra citológica a partir de una primera imagen digitalizada, y preferiblemente con la ayuda de una segunda imagen digitalizada de la misma muestra. La primera imagen se obtiene con la ayuda de un microscopio electrónico de transmisión y la segunda imagen se obtiene con la ayuda de una microscopía de fluorescencia. Las imágenes usadas proceden de la digitalización particular de una muestra citológica que se ha sometido a una preparación adaptada para el método, que a continuación se ha dispersado sobre un portaobjetos de citología. El método según la invención puede usar una muestra citológica procedente indistintamente de líquido o de extracción de elementos del cuerpo humano según el análisis deseado. Para la realización presentada en este caso, se toma una muestra de orina de micción espontánea. En este contexto, el resultado de este método de detección de células que presentan al menos una anomalía en una muestra citológica permitirá ayudar al diagnóstico de un cáncer de vejiga, concretamente en un estadio precoz o denominado de grado bajo del cáncer de vejiga.

A continuación en la descripción y tal como se ilustra en la figura 1, va a describirse una realización preferida que comprende varias etapas que pueden ser preliminar, intermedia o sucesiva. Se comprenderá que, según la realización elegida, podrán añadirse o eliminarse una o varias de estas etapas al método según la invención.

Ventajosamente, el método de detección según la invención comprende etapas ejecutadas con la ayuda de un módulo de procesamiento puesto en práctica por un ordenador adecuado para combinar datos calculados a partir de la muestra con datos de referencias conocidos en citología. Los datos denominados calculados, se calculan y se recopilan a partir de al menos una imagen digitalizada de un portaobjetos de al menos una muestra por el módulo de procesamiento. Por tanto, con la ayuda de un método de aprendizaje, el módulo de procesamiento puede combinar todos los datos calculados y recopilados a partir de la muestra o de varias muestras, y con datos de referencias conocidos en citología.

Según la realización preferida presentada en este caso, el método de detección según la invención comprende una etapa preliminar de procesamiento de imagen 110. Esta etapa preliminar de procesamiento de imagen 110, ilustrada en las figuras 1 a 3, tiene por objetivo afinar al máximo la selección de partida de las células para la detección de células, concretamente apartando elementos 4 constituyentes de una primera imagen digitalizada 1 que podrían interferir y reducir la finura del resultado del método de detección. Para ello, esta etapa preliminar de procesamiento 110 comprende una subetapa de detección 112 de los elementos 4 presentes en la primera imagen 1. Esta subetapa de detección 112 se basa al menos en parte en una detección colorimétrica que puede consistir al menos en parte en una segmentación en varios grupos del conjunto de los colores de la primera imagen 1, correspondiendo el grupo más claro y el más extendido al color del fondo de la imagen 6 y correspondiendo el grupo más oscuro al color de los núcleos celulares. Se considerará que cualquier píxel que no sea un píxel del fondo de la imagen 6 es un píxel de un elemento 4, es decir de un objeto biológico. Por tanto, se determina el fondo de la imagen 6 y todo lo que no forma parte del fondo de la imagen 6 corresponde a un elemento 4, es decir, una célula, un grupo de células, una agrupación

de células o un elemento a rechazar. Después, se extraen todos los elementos 4 del fondo de la imagen 6 creando varias imágenes reducidas 8 con la ayuda de un método de segmentación. Para categorizar los diferentes elementos 4 detectados, se usa además esta detección colorimétrica, detectando los núcleos, en este caso esta subetapa se denomina "basta".

Además, podrá combinarse esta primera subetapa de detección de los núcleos, es decir de detección colorimétrica con medidas de las dimensiones de cada uno de los elementos 4 detectados. Por ejemplo, si se detectan varios núcleos mediante detección colorimétrica, la medida de este elemento ayudará a identificar el tipo de elemento 4 detectado.

Por tanto, al final de la subetapa de detección 112, es posible identificar diferentes clases previas de los elementos 4 constituyentes de la primera imagen 1 con la ayuda de la detección colorimétrica de los núcleos, y eventualmente de la medida de estos elementos 4, tales como:

- Elementos a rechazar: es todo lo que no constituye una célula con núcleo. En efecto, cualquier elemento sin núcleo debe rechazarse, ya que no puede interpretarse, puede tratarse o bien de un desecho, bien de una falsa detección.

- Células aisladas: son las imágenes reducidas que sólo contienen un único núcleo y que, por consiguiente, son células aisladas.

- Grupo de células: son células agrupadas (que están en contacto o superpuestas) y que contienen por tanto varios núcleos.

- Agrupación: son agrupamientos particulares de células muy densos (muy oscuros) y en los que las células están superpuestas de una manera muy particular. Lo que diferencia a las agrupaciones de los grupos clásicos de células es el hecho de tener en las agrupaciones una superficie de la clase "núcleo" más importante.

De manera más detallada, la figura 3 ilustra un ejemplo de primera imagen 1 que se somete a la etapa preliminar de procesamiento 110. Tal como se describió anteriormente, durante esta etapa 110, la primera imagen 1 se somete a una segmentación para obtener las imágenes reducidas 8 que rodean a los elementos 4 que presentan un color oscuro, a continuación se clasifican previamente estas imágenes reducidas 8 en cuatro clases: elementos a rechazar 12, células aisladas 10, grupo de células 14 y agrupación de células 16 (que también se rechazarán). Para el correcto desarrollo de la invención, el método de detección, según la invención debe realizarse a partir de células necesariamente aisladas o individualizadas a partir de un grupo de células y que comprenden un núcleo. Las células aisladas 10 y los grupos de células 14 se filtran durante una subetapa de filtración 114 usando la detección colorimétrica.

Por consiguiente, la clase previa de grupo de células 14 se somete a una subetapa de individualización 116 tras la subetapa de filtración 114. Esta subetapa 116 consiste en individualizar las células del grupo de células 14. Por individualizar, se entiende extraer del grupo de células 14 al menos una célula que comprende un núcleo. Se observará que esta subetapa de individualización 116 no se ejecuta para la clase previa de las "agrupaciones de células" 16. En efecto, las agrupaciones de células 16 son estructuras demasiado complejas para individualizar y tenerlas en cuenta generará muchos más errores que información pertinente para usarse. Por eso es por lo que no se usan las agrupaciones 16 según la invención. La subetapa de individualización 116 de las células se realiza con la ayuda de una nueva segmentación que se basa en la detección de contornos mediante el cálculo del gradiente de la primera imagen o mediante el uso de una transformada matemática que tiene un aspecto multidireccional tal como Radon, Gabor o las transformadas de múltiples escalas multidireccionales. Tras haber determinado los contornos y eventualmente con la ayuda de un enfoque de morfología matemática, se extrae cualquier estructura comprendida en un contorno cerrado, correspondiente por tanto a una célula, y se define como que es una célula individualizada, interesante para la continuación del método según la invención.

Según la realización preferida en este caso, puede ejecutarse una etapa opcional de recomendación de explotabilidad 120 de la muestra para detectar células posiblemente cancerosas. Según las realizaciones, esta etapa de recomendación de explotabilidad 120 puede intervenir una o varias veces a varios niveles durante el desarrollo del método según la invención, preferiblemente al final de la etapa preliminar de procesamiento 110. Esta etapa de recomendación de explotabilidad 120 equivale a estimar si la primera imagen 1 contiene un número de células que comprenden un núcleo, suficiente para que la detección en la muestra no esté sesgada. Para ello, se determina el número total de células con núcleo presentes en la primera imagen 1. Este número total de células se comunicará a un usuario que validará o no la continuación del método según la invención. Preferiblemente, este número total de células con núcleo se compara con un valor de un umbral mínimo de explotabilidad que se determinará previamente por el usuario. Como variante, el usuario podrá introducir el valor del umbral mínimo de explotabilidad de la muestra deseado en el módulo de procesamiento para permitir continuar automáticamente o no el método de la invención. Por consiguiente, el método de detección puede continuarse aunque sólo haya muy pocas células con núcleo en la primera imagen 1, en la medida en que el número total de estas células sea superior o igual al valor del umbral mínimo de explotabilidad. Por ejemplo, el valor del umbral mínimo de explotabilidad es igual a al menos siete células con núcleo detectadas en la muestra, por tanto el método se continúa si el número total de células con núcleo es superior o igual

a siete. A la inversa, si el número de células con núcleo es insuficiente, la etapa de recomendación de explotabilidad 120 puede interrumpir el método de detección.

Según la realización preferida en este caso, puede ejecutarse una etapa opcional preliminar de selección tipológica 130 para detectar células de un tipo celular elegido en la muestra con al menos una anomalía celular, de manera más precisa, posiblemente cancerosas. Para ello, a partir de la primera imagen 1, se miden las células de la muestra y después se clasifican según sus tamaños, etapa 132. Con la ayuda del módulo de procesamiento, las células clasificadas según sus tamaños se asignan a tipos celulares, después se seleccionan según el tipo celular elegido, tal como urotelial o cualquier otro tipo celular del cual puede correlacionarse un nivel de anomalía celular a un tipo de cáncer, etapa 134. Según una variante, el usuario podrá definir por sí mismo el tipo celular preferido. Por ejemplo, para un método de detección aplicado a muestras de citología de orina para ayudar al diagnóstico del cáncer de vejiga, el tipo celular preferido será el tipo urotelial. De manera general, el estudio de las células uroteliales es un buen indicador para ayudar al diagnóstico del cáncer de vejiga, concretamente para determinar el estadio del cáncer en función de las anomalías encontradas.

Tras la etapa preliminar de procesamiento de imagen 110, en la etapa opcional de recomendación de explotabilidad de la muestra 120, y en la etapa opcional preliminar de clasificación tipológica 130 se obtiene una selección fina de células aisladas o individualizadas, que comprende un núcleo y un tipo celular elegido.

Como recordatorio, el método según la invención da lugar a la detección de células que presentan al menos una anomalía en la muestra citológica a partir de al menos una primera imagen. El nivel de anomalía celular de la muestra está correlacionado con el número y el tipo de anomalía celular encontrado en la muestra citológica. Dicho de otro modo, el método según la invención puede permitir clasificar la muestra o bien en una clase de nivel de anomalía celular denominado promediado, o bien en una clase de nivel de anomalía celular denominado nulo.

Para ello, la realización preferida ilustrada en la figura 1 comprende además una etapa de identificación de anomalía celular 140 en la muestra. Según la invención, una célula que presenta una anomalía se define como una célula que presenta una característica experta de anomalía divergente de un dato denominado de referencia, es decir divergente de la norma. Dicho de otro modo, se trata de una célula que presenta al menos una anomalía, es decir una característica denominada experta de anomalía que es anómala. Por tanto, el cálculo de una característica experta de anomalía permite obtener un dato que permite ayudar a determinar una desviación con la norma y, por tanto, la presencia o no de una anomalía. Por cálculo de características expertas de anomalía, se entiende la medición o cualquier otro tipo de métodos matemáticos. Según la invención, la característica experta de anomalía que se calcula es una característica del núcleo, en particular corresponde a una razón de tamaño entre el núcleo y el citoplasma de la célula o a una medida del posicionamiento del núcleo o a una medida colorimétrica del núcleo. Más precisamente, la característica experta de anomalía corresponde a una razón de superficie del núcleo en la célula, o a una irregularidad del núcleo, o a una colorimetría moteada o una hipercolorimetría del núcleo. Las diferentes características expertas de anomalía mencionadas anteriormente, a título no limitativo, pueden revelar anomalías según los valores calculados en la muestra. Se entenderá que podrán calcularse indistintamente otros tipos de características expertas de anomalías y usarse como parámetros para clasificar las células.

Las anomalías buscadas y detectadas en cada célula van a referirse, más particularmente, a su núcleo. Por tanto, para ello, hace falta una detección fina del núcleo para poder detectar las anomalías asociadas al mismo.

La detección fina del núcleo se realiza por una segmentación. Esta segmentación puede realizarse concretamente mediante afinamiento a partir de la segmentación basta usando únicamente la colorimetría. Esta técnica descrita anteriormente en la descripción consiste en partir de una envoltente de recubrimiento, que forma preferiblemente un rectángulo alrededor del núcleo. Esta envoltente de recubrimiento delimitará una zona alrededor del núcleo, siempre incluida en la célula, y que desborda del núcleo del 10 al 40% (de la superficie total del núcleo), preferiblemente entre el 20 y el 30%. Las figuras 4 y 5 ilustran, cada una, el resultado de una detección fina de núcleo.

Según la realización, el núcleo puede haberse detectado ya en la etapa preliminar de procesamiento de la imagen 110 durante la detección basta de los núcleos con la ayuda de una primera segmentación. Un ejemplo de realización consiste en tener en cuenta la envoltente de recubrimiento usada para la detección basta del núcleo y estimar en esta envoltente la distribución de color del núcleo así como la del fondo usando un modelo de mezclado de gaussianas. A continuación, se usa para construir un campo aleatorio de Markov en las etiquetas de los píxeles, con una función de energía que prefiere las regiones conectadas que tienen la misma etiqueta, y ejecutando una optimización basada en el corte de grafos para interferir con sus valores. Dado que esta estimación es susceptible de ser más precisa que la original extraída de la caja de recubrimiento, estas dos etapas se repiten hasta la convergencia.

La etapa de identificación de anomalía celular 140 en las células de la muestra se realiza en parte mediante el cálculo de al menos dos características expertas de anomalía distintas, etapa 142.

A continuación en la descripción, se describe el cálculo de cuatro tipos de características expertas de anomalías que ponen en evidencia, cada uno, la presencia o no de una anomalía celular morfológica.

En primer lugar, va a describirse el cálculo de la característica experta de anomalía de la razón núcleo/citoplasma en la célula.

Se propone calcular con la ayuda del módulo de procesamiento automatizado la razón RNC (razón núcleo-célula) que se define mediante las siguientes fórmulas:

$RNC = \text{longitud del diámetro más grande del núcleo} / \text{longitud del diámetro más grande de la célula correspondiente}$

La figura 5 ilustra un ejemplo de realización del cálculo de la razón RNC.

Las etapas del cálculo son las siguientes, con la ayuda del módulo de procesamiento:

1. Trazar el eje mayor correspondiente a la célula,
2. Calcular la longitud del eje delimitado por la célula,
3. Calcular la longitud de este mismo eje delimitado por el núcleo,

También se propone otra manera de calcular esta razón basándose en el cálculo de superficies.

$$RNC = (\text{superficie del núcleo} / \text{superficie de la célula})^{1/2}$$

Con:

Superficie de la célula = número de píxeles de la célula.

Superficie del núcleo = número de píxeles del núcleo.

Esta razón es una cifra inferior a 1 ya que el núcleo está en el interior de la célula. Si la RNC es muy próxima a 1, se considera que la célula es "un núcleo desnudo" y, por tanto, se rechazará durante el método por el módulo de procesamiento. Entonces, se propone ayudar al usuario para que introduzca un umbral a partir del cual se considera que la célula es "un núcleo desnudo". Lo mismo se aplica para los umbrales que indican una RNC normal, crítica o anómala. Por defecto, se inicializarán estos umbrales con los siguientes valores:

- a) $RNC < 0,5$ correspondiente a una RNC normal
- b) $0,5 < RNC < 0,7$ correspondiente a una RNC crítica
- c) $RNC > 0,7$ correspondiente a una RNC anómala
- d) $RNC > 0,9$ correspondiente a un núcleo desnudo que no se tiene en cuenta "a rechazar"

Más precisamente, la razón núcleo/célula se calcula a partir de la razón entre el diámetro más grande de la célula y el diámetro más grande del núcleo. En el caso en que la célula no es muy esférica, no se toma el diámetro real sino la distancia o longitud de extremo a extremo de la célula, correspondiente al diámetro más grande de la célula. Todas las razones comprendidas entre 0,9 y 1 no podrán interpretarse ya que esto corresponderá a núcleos desnudos. Las razones superiores a 0,7 corresponden a una intensa anomalía y las razones comprendidas entre 0,5 y 0,7 corresponden a una baja anomalía.

En segundo lugar, va a describirse el cálculo de otra característica experta de anomalía que es la irregularidad del núcleo.

Por núcleo irregular se entiende cualquier núcleo cuyo contorno se desvía de una forma clásica de contorno de núcleo denominado de referencia muy liso y regular. También puede considerarse como un contorno que tiene puntos angulosos o esquinas.

Para ello, se propone calcular la irregularidad del núcleo mediante tres métodos, de los cuales el tercer método descrito es la fusión de los otros dos métodos.

En un primer momento, va a describirse el primer método basado en el cálculo de diferencia con un núcleo de referencia de una célula de referencia.

La diferencia calculada puede ser una diferencia de superficie o una diferencia de perímetros o cualquier otra diferencia tal como la diferencia de baricentro, la diferencia de píxel en cada región del núcleo, la diferencia entre la longitud de los contornos de cada núcleo (de referencia y real), la diferencia lineal entre el centro de gravedad de cada núcleo, la diferencia entre el número de píxeles de las regiones convexas asociadas a cada núcleo, la diferencia entre la

excentricidad del núcleo real y la del núcleo de referencia. Como recordatorio, la excentricidad $e=0$ para un círculo, $0<e<1$ para una elipse, $e=1$ para una parábola, la diferencia en cuanto a regiones convexas.

Este método se basa, entre otras cosas, en el cálculo de la diferencia entre el contorno real del núcleo y el contorno de referencia del núcleo correspondiente a una curva convexa cerrada de rango C2 incluida en una envolvente de recubrimiento del núcleo. Para determinar el contorno de referencia del núcleo, se define la envolvente de recubrimiento del núcleo como un rectángulo de recubrimiento cuya longitud trazada corresponde al diámetro más grande del núcleo. Con la ayuda del rectángulo de recubrimiento, el módulo de procesamiento va a determinar la curva C2 en el interior de este último, formando el contorno del núcleo.

A modo de ejemplo, el óvalo es una curva convexa cerrada de rango C2. Asimismo, el círculo o la elipse son casos particulares de las curvas convexas cerradas C2 de tipo ovalado.

Después, se extrae el contorno del núcleo mediante cualquier método de extracción de contornos tal como: los detectores de gradiente mediante filtración lineal o por máscara (tales como Sobel, Prewitt o Roberts), los enfoques analíticos de extracción de contornos (tales como Canny o filtro de Deriche), los contornos activos tomando como contorno inicial la envolvente de recubrimiento vista anteriormente, los conjuntos de niveles, los contornos extraídos mediante transformadas de múltiples escalas y multidireccionales,... etc.

Resumen de las etapas que deben seguirse para calcular la característica experta de anomalía de irregularidad del núcleo con la ayuda del cálculo del contorno del núcleo:

1. calcular un contorno real del núcleo usando un método de extracción de contornos,
2. determinar el diámetro más grande del núcleo,
3. trazar una envolvente de recubrimiento del núcleo, es decir un rectángulo de recubrimiento el núcleo cuya longitud es igual al diámetro más grande del núcleo,
4. a partir de la envolvente de recubrimiento, trazar una curva convexa cerrada de rango C2, que forma el contorno estimado,
5. calcular un valor de diferencia entre el contorno real del núcleo y el contorno de referencia del núcleo de referencia correspondiente a la curva trazada en la etapa 4, calculando y combinando datos de la muestra y datos de referencia de diferentes características expertas de anomalía de formas,
6. usar el valor de diferencia entre el contorno de referencia y el contorno real del núcleo para determinar la presencia de una anomalía en función de valores umbrales predeterminados o introducidos por un usuario, o combinar este dato con otros datos de características expertas de anomalía y determinar la presencia de una anomalía en función de valores umbrales predeterminados o introducidos por un usuario.

Se comprenderá que pueden usarse estas características en la entrada de un clasificador de tipo separadores de margen amplio (técnica de aprendizaje supervisado que se basa en la maximización del margen. El margen es la distancia entre la frontera de separación y las muestras más próximas que se denominan vectores de soportes. La técnica consiste en encontrar la frontera de separación de margen máximo a partir de los ejemplos de aprendizaje) y/o mediante redes neuronales que determinarán la irregularidad del núcleo mediante aprendizaje.

En un segundo momento, va a describirse el segundo método del cálculo de la irregularidad de un núcleo, basándose en la detección de los puntos angulosos en un contorno de núcleo.

La idea en este caso es detectar los puntos angulosos detectando los puntos característicos de tipo "esquinas", calcular el ángulo formado correspondiente y estimar el grado de irregularidad del núcleo teniendo en cuenta el número de esquinas detectadas y sus ángulos correspondientes.

Una esquina se define en este caso como un ángulo de un contorno de membrana nuclear. Una esquina corresponde a la intersección de dos líneas o de dos contornos de la membrana nuclear, de orientación bastante diferente al entorno próximo de la esquina.

Las esquinas se detectan generalmente por el detector de Harris, salvo que este último no verifica la característica de invariancia con respecto a la escala y sólo puede detectar las esquinas en una única escala, lo cual puede llevar a detectar falsas esquinas o a omitir algunas esquinas importantes, en particular cuando se dispone de imágenes de bajo contraste.

Otro inconveniente principal del detector de Harris es que produce generalmente una sobredetección de las esquinas. Para resolver esto, se propone usar, tras el detector de Harris, el algoritmo FAST (*Features from Accelerated Segment Test*, características de la prueba de segmentos acelerada) inicialmente desarrollada por Edward Rosten y Tom

Drummond. En efecto, este último algoritmo puede resultar útil para suprimir las esquinas redundantes. Por consiguiente, se aplicará para eliminar las falsas esquinas detectadas por el detector Harris.

Por otro lado, la idea original e inventiva propuesta en el contexto de esta invención es combinar la transformada en tipo *contourlet* (transformada de múltiples escalas multidireccional), el detector de Harris y el algoritmo FAST (*Features from Accelerated Segment Test*, características de la prueba de segmentos acelerada) para detectar las esquinas correspondientes al contorno del núcleo y después extraer las orientaciones y estimar la irregularidad del núcleo. Uno de los objetivos es estimar correctamente la irregularidad del contorno del núcleo que define la membrana nuclear. La idea original es usar una transformada que destaque en la puesta en evidencia de los contornos y los resalte correctamente para poder estimar correctamente su irregularidad. Por tanto, la estimación no se realizará de manera clásica en el campo de la imagen inicial sino en un campo compuesto por diferentes escalas en el que los contornos estarán correctamente puestos en evidencia y correctamente resaltados para detectar correctamente su irregularidad. Por tanto, la transformada es una transformada de múltiples escalas y multidireccional de tipo *contourlet* que representa perfectamente las curvas y los contornos. Se trata de una herramienta de análisis geométrico de múltiples escalas que permite una descomposición según diferentes ángulos de orientación para cada escala. A diferencia de la transformada de ondículas clásica, los *contourlet* captan perfectamente las características anisotrópicas de las imágenes (información de contorno, de borde, de textura). Por tanto, su uso para poner en evidencia la irregularidad del contorno del núcleo es sin duda muy apropiado.

A continuación en la descripción se detallan las etapas del cálculo de la irregularidad del núcleo propuesto en el contexto de este segundo método:

1. Descomponer la primera imagen usando la transformada de tipo *contourlet*. Se dispondrá de varias imágenes correspondientes a las diferentes escalas y a las diferentes orientaciones,

2. Calcular la matriz de correlación M_θ (matriz relativa al detector de Harris) para cada píxel de cada escala,

3. Deducir la intensidad de esquinas (*corner response function*, función de respuesta de esquinas) de cada píxel correspondiente a cada escala que se estima mediante:

$$R_\theta = \det M_\theta - \alpha(\text{traza } M_\theta)^2 \text{ con } 0,04 < \alpha < 0,25$$

Más precisamente, se detecta la intensidad de las esquinas mediante la "función de respuesta de esquinas" que es una función que devuelve valores intensos para la detección de una esquina. Cuanto más grande o intenso es el valor de esta función en un punto, más probabilidades hay de que se trate de una esquina. El factor α define la sensibilidad del detector. Cuanto más grande es α más baja es la sensibilidad y más pequeño será el número de esquina. Generalmente, se fija el valor de α entre 0,04 y 0,06. Este valor puede ser más grande pero no supera 0,25.

4. Tras el cálculo de la intensidad de esquinas de cada píxel, guardar la posición de los puntos que tienen un valor alto de R_θ que corresponderán a las esquinas detectadas. Esta primera detección se afinará más adelante ya que puede contener tanto esquinas reales como esquinas erróneas (problema de sobredetección de esquinas),

5. Aplicar el algoritmo FAST con el objetivo de juzgar la densidad alrededor del círculo que rodea el punto detectado p. Verificar si existe un conjunto de n puntos contiguos en el círculo que sean todos ellos más luminosos que $\text{Intensidad}(p)+t$ o menos luminosos que $\text{Intensidad}(p)-t$ (siendo t un umbral). Si se da este caso, entonces p es una esquina,

6. Usar la supresión no máxima para eliminar las esquinas detectadas que son adyacentes entre sí:

6.1 Calcular una función de puntuación V para todas las esquinas detectadas, siendo V la suma de diferencias absolutas entre la esquina p y los 16 píxeles que la rodean,

6.2 Comparar la puntuación V de las esquinas adyacentes para eliminar aquella que tiene la puntuación V más baja.

7. Calcular el ángulo de cada esquina o punto anguloso detectado calculando la diferencia de orientación entre los dos segmentos que lo componen,

Por ejemplo, si el primer segmento tiene una orientación de 60° y el segundo de 45° entonces la diferencia es el valor absoluto de $60^\circ - 45^\circ = 15^\circ$.

Alternativamente, el cálculo de los ángulos puede realizarse no a partir de los segmentos estimados sino directamente con la ayuda de los valores de los píxeles en la primera imagen 1.

8. Calcular un grado de irregularidad del núcleo combinando los datos del número de puntos angulosos o de esquinas detectadas y sus ángulos respectivos. Ejemplo de cálculo del grado de irregularidad:

$$\text{Grado de irregularidad del núcleo} = \sum_{i=1:n} \text{esquinas} (1/\text{ángulo}_i)$$

9. A partir del resultado obtenido en la etapa 9, determinar el nivel de irregularidad del núcleo en función de valores umbrales predeterminados o introducidos por un usuario.

En un tercer momento, va a describirse el tercer método que consiste en fusionar el primera y el segundo métodos de cálculo de la irregularidad del núcleo.

Se propone fusionar los dos métodos anteriores. La combinación de los resultados de los dos métodos anteriores puede realizarse mediante enfoques de fusión de datos tales como, por ejemplo, la lógica difusa. A continuación se presenta un ejemplo de reglas que podrán aplicarse:

1. Si el primer método proporciona como resultado "núcleo regular" y el segundo método proporciona como resultado "núcleo con grado de irregularidad bajo" entonces la decisión final será "núcleo regular",

2. Si el primer método proporciona como resultado "núcleo irregular" y el segundo método proporciona como resultado "núcleo con grado de irregularidad medio" entonces la decisión final será "núcleo irregular",

3. Si el primer método proporciona como resultado "núcleo irregular" y el segundo método proporciona como resultado "núcleo con grado de irregularidad nulo" entonces la decisión final será "núcleo regular".

En tercer lugar, va a describirse el cálculo de otra característica experta de anomalía que es la hipercolorimetría del núcleo.

En este caso, se pretende calcular el grado de hipercolorimetría del núcleo. Las etapas son las siguientes:

1. Calcular para cada célula una intensidad luminosa media de su núcleo, correspondiente a la suma de la intensidad luminosa de los píxeles que componen el núcleo dividida entre el número de píxeles del núcleo,

2. Calcular una media I_{ref} de la intensidad luminosa media de los núcleos de todas las células presentes en la primera imagen 1. I_{ref} es la suma de las intensidades luminosas medias de todos los núcleos de las células dividida entre el número de células,

3. Calcular la intensidad de la hipercolorimetría como la diferencia entre la intensidad luminosa media del núcleo de la célula que va a controlarse e I_{ref} .

Se determinará un umbral de hipercolorimetría mediante un procedimiento de clasificación (mediante aprendizaje, por ejemplo de tipo SVM o redes neuronales). Por tanto, según el valor obtenido y el umbral de hipercolorimetría, se determinará la hipercolorimetría o no de la célula.

Indistintamente, para calcular la hipercolorimetría del núcleo, en el conjunto de las etapas descritas anteriormente podrá sustituirse el cálculo de la media por el cálculo de la mediana.

En cuarto lugar, va a describirse el cálculo de otra característica experta de anomalía que es la colorimetría moteada del núcleo.

Para ello, se proponen tres métodos para estimar la colorimetría moteada de un núcleo. El primer método se basa en los valores de la intensidad luminosa de los píxeles de la imagen, el segundo método se basa en una estimación en el campo de transformada y no en el campo de imagen (dicho de otro modo, no se aplica en los valores de los píxeles de la imagen sino en los valores de los coeficientes en el campo de la transformada), y el tercer método consiste en fusionar los dos métodos anteriores.

En un primer momento, va a describirse el primer método basado en la estimación en el campo de imagen:

1. Calcular una varianza V ,

2. Calcular un histograma de la imagen y detectar un número de picos (n .º picos) en el histograma, (si hay más de un pico, esto quiere decir que el núcleo no tiene un color homogéneo),

3. Calcular, para cada píxel de la imagen (o bloque), un tensor a partir de un gradiente para formar un mapa de coherencia. Si un píxel tiene una coherencia más intensa que la media de las coherencias de la zona en la que se encuentra, entonces es característico de una zona no homogénea:

Si $\text{coherencia}(x,y) > \text{media}(\text{zona circundante } (x,y)) + 3 \cdot \text{desviación estándar}(\text{zona circundante } (x,y))$ entonces (x,y) pertenece a una zona no homogénea,

Calcular una suma de los puntos (o bloques) no homogéneos detectados mediante la coherencia ($n.^{\circ}$ coh),

4. Fusionar estas características [V , $n.^{\circ}$ picos y $n.^{\circ}$ coh] y usarlas para detectar la colorimetría moteada del núcleo.

5 Se entenderá que pueden usarse estas características en la entrada de un clasificador de tipo SVM o redes neuronales.

En un segundo momento, va a describirse el segundo método basado en la estimación en el campo de la transformada:

10 Se pretende definir una colorimetría moteada por la presencia de tres manchas como mínimo en la zona afectada, es decir el núcleo. En este caso, la idea es detectar un mínimo de tres contornos cerrados en el interior de la zona. Para ello, se pretende detectar los contornos mediante una transformada apropiada tal como la transformada de tipo *contourlet*. Esta transformada tiene la particularidad de tener coeficientes de baja energía (próxima a 0) en las zonas homogéneas y coeficientes de energía muy intensa en las zonas anisotrópicas (no homogéneas) y que corresponden
15 siguientes etapas:

1. Realizar la descomposición de la imagen mediante una transformada de múltiples escalas multidireccional apropiada tal como la transformada de tipo *contourlet* o átomos de onda,
20

2. Detectar las zonas de intensa componente energética aplicando una determinación de umbrales para conservar únicamente los coeficientes correspondientes a los contornos de la imagen,

3. Aplicar una función no lineal de resalte de contraste a los coeficientes de *contourlet* conservados (sometidos a aplicación de umbral),
25

4. Reconstruir la imagen de contorno a partir de los coeficientes conservados usando la transformada inversa,

5. Detectar en la imagen de contorno los contornos cerrados y numerarlos,
30

6. Si hay más de tres contornos cerrados, se concluirá que el núcleo tiene una colorimetría moteada.

En un tercer momento, va a describirse el tercer método que consiste en fusionar el primer y el segundo métodos de cálculo de la colorimetría moteada.
35

Asimismo, para el cálculo de la irregularidad del núcleo, se propone, en este caso, combinar los resultados obtenidos mediante cada uno de los dos métodos anteriores para obtener un resultado fino para la característica experta de anomalía de colorimetría moteada del núcleo de cada célula.

40 Se comprenderá que, con la ayuda del módulo de procesamiento, el conjunto de las características expertas de anomalías calculadas (razón núcleo/célula, irregularidad del núcleo, hipercolorimetría y colorimetría moteada) se combinan con datos de referencia que consisten en datos de referencias conocidos en citología. En particular, estos datos calculados pueden usarse o bien tal cual, o bien por un clasificador, o bien fusionados con otras características calculadas y generadas mediante aprendizaje profundo para determinar si la célula es normal o presenta al menos
45 una anomalía, etapa 144. Si la muestra no comprende al menos una célula que presente al menos una anomalía, entonces el método de detección se detiene en este caso, y si la muestra presenta al menos una célula con al menos una anomalía, el método de detección continúa.

Al final de la etapa de identificación de anomalía celular 140, el módulo de procesamiento determina un número total de células que presentan al menos una anomalía en la muestra citológica, etapa 146.
50

Después, puede intervenir, en este caso, una primera etapa de caracterización de la muestra 150 según el nivel de anomalía de la muestra en su globalidad. Para ello, el número total de células detectadas que presentan al menos una anomalía se compara con un valor umbral de un nivel de anomalía de la muestra. Este valor umbral del nivel de anomalía de la muestra se define previamente por el usuario para el método. Corresponde a al menos tres células que presentan, cada una, al menos una anomalía. Este valor umbral del nivel de anomalía puede definirse indistintamente de manera previa por el usuario según la finura del resultado de clasificación deseada.
55

Según la invención, si el número total de células con al menos una anomalía es inferior al valor umbral del nivel de anomalía, la muestra citológica se clasifica en una clase representativa de una muestra con un nivel de anomalía nulo 152, si el número total de células con al menos una anomalía es superior o igual al valor umbral del nivel de anomalía, la muestra citológica se clasifica en una clase representativa de una muestra con una clase con un nivel de anomalía no despreciable 154, es decir un nivel de anomalía bajo y un nivel de anomalía intenso 154.
60

En resumen, según otro aspecto, la invención permite por tanto en esta fase clasificar la muestra citológica según dos clases: una clase con un nivel de anomalía celular nulo 152 y una clase con un nivel de anomalía celular no
65

despreciable 154. Este resultado permite poner en evidencia anomalías celulares que son características de células cancerosas. Por tanto, permite ayudar a los citotécnicos y citopatólogos para el diagnóstico del cáncer, aportando un resultado normalizado fiable y rápido.

5 Tal como se ilustra en la figura 6, el método de detección puede comprender una segunda etapa de caracterización de la muestra 160 según el nivel de anomalía celular de la muestra, que pone en concreto en evidencia específicamente un nivel de anomalía celular bajo en la muestra. Este nivel de anomalía corresponde a células que aún no ha experimentado deformaciones morfológicas pronunciadas. Preferiblemente, según esta realización particular, la invención permite clasificar de manera más precisa la muestra o bien en una clase de nivel de anomalía celular denominado "bajo", o bien en una clase de nivel de anomalía celular denominado "intenso", o bien en una clase de nivel de anomalía celular denominado "nulo".

15 Tras la detección de célula anómala, la clasificación de la muestra según el nivel de anomalía morfológico o radiométrico de estas células de la muestra proporciona una verdadera ayuda para el médico con respecto al diagnóstico del cáncer. En efecto, el nivel de anomalía celular denominado "bajo" puede estar asociado con el estadio precoz denominado "de grado bajo" del cáncer y el nivel de anomalía denominado "intenso" puede estar asociado con el estadio avanzado del cáncer denominado "de grado alto".

20 Para ello, para detectar las anomalías de un nivel denominado bajo, la invención usa una segunda imagen 2 digitalizada de obtención de imágenes por fluorescencia. A partir de esta segunda imagen 2, se seleccionan las células de al menos un tipo celular elegido, tal como urotelial, en una etapa de selección tipológica 162, gracias a sus ubicaciones predefinidas en la muestra en la primera imagen 1 digitalizada. Entonces se reajustan la primera imagen 1 y la segunda imagen, eventualmente con la ayuda de un método clásico de reajuste puesto en práctica por ordenador.

25 Se comprenderá que, para ejecutar esta segunda etapa de caracterización 160, la etapa preliminar de selección tipológica 130 a partir de la primera imagen 1 debe haberse ejecutado de manera necesaria previamente en el método según la invención para poder localizar las células del tipo celular elegido.

30 Esta subetapa de selección tipológica 162 de un tipo celular va seguida por una subetapa de detección de un halo de fluorescencia 164 alrededor de cada célula seleccionada anteriormente. El término "halo" se define como un fenómeno que se traduce en la presencia de una aureola fluorescente que rodea a la célula y que está bien contrastada. En una variante de realización, se mide la intensidad del halo, la tasa de fluorescencia o la tasa de dispersión de la fluorescencia. Por tanto, si se detecta un halo de fluorescencia, preferiblemente si se detecta una tasa mínima de fluorescencia, o más preferiblemente, una tasa mínima de dispersión de la fluorescencia, se clasifica la muestra en una clase representativa de una muestra con un bajo nivel de anomalía 168. A la inversa, las células que no presentan fluorescencia se clasifican en la clase de nivel de anomalía nulo e intenso 166.

40 Según la presente invención, se proponen tres métodos para realizar la etapa de segunda caracterización de la muestra 160 con la ayuda del módulo de procesamiento.

A continuación se detallan las subetapas de la etapa de la segunda caracterización de la muestra 160 en el contexto de un primer método:

45 - usar la primera imagen 1 y detectar el centro de cada célula de un tipo celular elegido, con la ayuda del cálculo del baricentro de la célula,

50 - trazar un perfil de intensidad a lo largo de cada segmento de recta que comienza en el centro de la célula hasta el límite de cada célula según diferentes ángulos. Por ejemplo, según 8 ángulos: 0°, 45°, 90°, 135°, 180°, 225°, 270° y 315°,

- detectar el pico P_{máx} (punto máximo) para cada perfil así como el valle P_{mín} que precede al máximo (el punto mínimo),

55 - trazar en la curva del perfil de intensidad una línea horizontal que parte del valle P_{mín} situado a la derecha del pico P_{máx} hasta la intersección Pint con la curva del perfil de intensidad a la izquierda del pico P_{máx},

60 - calcular la distancia lineal horizontal dx (en abscisas) entre X_{máx}, la abscisa del pico, y X_{int}, la abscisa del punto Pint,

- calcular la distancia vertical dy (en ordenadas) entre Y_{máx}, la ordenada del pico, e Y_{int}=Y_{mín}, la ordenada de Pint y de P_{mín},

65 - calcular para cada célula detectada Dx y Dy que son la media de dx y de dy según los 8 ángulos. Dy representará una primera estimación de la intensidad del halo de fluorescencia del halo y Dx una primera estimación de la dispersión del halo.

A continuación se detallan las subetapas de la etapa de la segunda caracterización de la muestra 160 en el contexto de un segundo método:

- 5 En este caso se propone medir y calcular la dispersión de un halo de fluorescencia a partir de los perfiles de Radon de las imágenes reducidas de las células.

La transformada de Radon representa una imagen como una colección de proyecciones en diversas direcciones. Cada proyección según una línea representa un perfil de Radon.

- 10 Las etapas que deben seguirse para el segundo método son las siguientes:

- marcar la posición de cada célula de un tipo celular elegido en la primera imagen 1,

- 15 - realizar la correspondencia en la segunda imagen 2 para extraer las imágenes reducidas correspondientes a cada célula del tipo celular elegido que supera del 10 al 30% de longitud y de anchura alrededor de cada núcleo,

- calcular los perfiles de Radon correspondientes a las imágenes reducidas para diferentes ángulos, preferiblemente los mismos ángulos usados en el primer método,

- 20 - para cada perfil, detectar los dos picos P_{\min} y P_{\max} de cada perfil de Radon de cada imagen reducida que supera del 10 al 30% de longitud y de anchura alrededor del núcleo y seguir a continuación las etapas del primer método para calcular una segunda estimación de la intensidad del halo y de su dispersión.

- 25 Alternativamente, en el contexto de un tercer método, la etapa de la segunda caracterización de la muestra 160 ejecutada por el módulo de procesamiento, consiste en un método que se basa en el cálculo de la irregularidad del contorno del "halo" de fluorescencia presente. Dicho de otro modo, si una célula presenta un nivel intenso de irregularidad del contorno del halo, esto significa que presenta un halo de fluorescencia. Para ello, este tercer método consiste en uno de los tres métodos del cálculo de irregularidad del núcleo presentado anteriormente.

- 30 Para el conjunto de los tres métodos de ejecución de la segunda caracterización de la muestra 160, descritos anteriormente, a título no limitativo, los resultados obtenidos para determinar la clase de la muestra se comparan con valores umbrales del nivel de intensidad de fluorescencia y/o de dispersión que están predeterminados o se introducen previamente por un usuario.

- 35 Se comprenderá que estos tres métodos descritos pueden combinarse y tratarse por método de aprendizaje para obtener un resultado fino.

- 40 Esta etapa de la segunda caracterización de la muestra 160 permite por tanto obtener, al final, la clasificación del portaobjetos de la muestra, por tanto de la muestra, en tres clases posibles, es decir: o bien en la clase de nivel de anomalía nulo, que puede asociarse a una muestra procedente de un sujeto sano; o bien en la clase de nivel de anomalía celular bajo, que puede asociarse a una muestra que presenta células cancerosas en un estadio precoz o denominado cáncer "de grado bajo"; o bien en la clase de nivel de anomalía celular intenso, que puede asociarse a una muestra que presenta células cancerosas en un estadio avanzado o denominado cáncer "de grado alto".

- 45 Un segundo aspecto de la invención propone usar un segundo método de detección de anomalía celular en la muestra citológica usado un método de aprendizaje profundo para realizar la clasificación directa de una imagen de portaobjetos de muestra citológica según un nivel de anomalía celular encontrado. De manera general, un método de aprendizaje profundo enseña a un modelo informático cómo realizar tareas tales como la clasificación directamente a partir de imágenes. A diferencia de lo que se realiza habitualmente en el campo del aprendizaje profundo y del procesamiento de imágenes, la originalidad de este segundo método es que clasifica un conjunto de imágenes reducidas sin realizar una clasificación previa de cada una de las imágenes reducidas que constituyen el conjunto de la imagen. Más precisamente, este método permite directamente la clasificación de la imagen de la muestra citológica del portaobjetos, es decir del conjunto de las células comprendidas en la imagen de la muestra sin pasar por una clasificación de cada imagen reducida o célula de la muestra. Se observará que este método de clasificación según la presente invención es particularmente útil cuando la detección de anomalías o de patologías sólo puede realizarse considerando el conjunto de un portaobjetos o la imagen del portaobjetos, y no considerando cada célula independientemente unas de otras. Para ello, este segundo método de detección de células que presentan al menos una anomalía en la muestra usa un modelo informático de aprendizaje profundo, que toma un conjunto de observaciones en la entrada y produce una clasificación de este conjunto en la salida. Estas observaciones pueden ser directamente imágenes reducidas brutas de la muestra citológica y/o una lista de valores que caracterizan la imagen reducida (o datos de descriptores), por ejemplo valores calculados con la ayuda del primer método (se comprenderá que estos últimos datos de descriptores tendrán la utilidad de enseñar al modelo informático). Para ello, este segundo método usa un método de aprendizaje profundo que usa una arquitectura de red neuronal profunda (NN). Cuando esta red neuronal profunda (NN) se usa con las imágenes reducidas brutas, se trata de una red neuronal profunda convolucional (CNN). La CNN somete a convolución las características aprendidas con los datos de entrada

del descriptor, y usa capas de convolución 2D, perfectamente adaptadas al procesamiento de los datos 2D tales como las imágenes. La red neuronal convolucional puede extraer representaciones expertas de las imágenes directamente mediante un aprendizaje. No obstante, el método de aprendizaje profundo necesita una potencia de cálculo considerable. Dado que los cálculos son altamente paralelizables, generalmente se realizan en un procesador gráfico o GPU. La GPU presenta grandes prestaciones y está dotada de una arquitectura paralela, que es eficaz para la puesta en práctica de este método de aprendizaje. La GPU está integrada en una unidad de procesamiento o CPU. No obstante, el uso de GPU no es posible tratando cada una de las entradas de manera independiente ya que una de las operaciones que más tiempo requieren sigue siendo la transferencia de los datos de la memoria a la CPU y a la GPU. De manera general, para aliviar este problema, los métodos de aprendizaje profundo usan tensores 4D (tablas de 4 dimensiones) para procesar de manera general las entradas por paquetes o en lote.

Gracias a un uso original del tensor 4D, este segundo método de detección de anomalía celular clasifica la imagen, es decir el conjunto de imágenes reducidas, sin detección previa de cada una de ellas. Dicho de otro modo, este segundo método de detección propone detectar las células que presentan al menos una anomalía celular y clasificar directamente no una imagen reducida sino un conjunto de imágenes reducidas (la imagen reducida bruta y/o caracterizada por valores calculados de otra manera) que representan la imagen de un portaobjetos de una muestra citológica (portaobjetos). Para realizar este segundo método, el entrenamiento del modelo informático de aprendizaje profundo se realiza con la ayuda de un gran conjunto de portaobjetos etiquetados. Sin este entrenamiento, el segundo método de clasificación no puede funcionar. En este caso, la entrada es un conjunto de imágenes reducidas (brutas y/o caracterizadas mediante el primer método) y la salida es una distribución de probabilidades según clases para asignar al final una clase a cada imagen del conjunto de las imágenes reducidas de la muestra analizada. Según la invención, las clases se elegirán de una clase de nivel de anomalía celular bajo, una clase de nivel de anomalía celular intenso y/o una clase de nivel de anomalía celular nulo.

Además, este método realiza un aprendizaje "de extremo a extremo": es decir que conecta directamente la entrada a la salida sin pasar por una fase intermedia de análisis de cada imagen reducida de una imagen de una muestra citológica. Más precisamente a partir de un conjunto de imágenes reducidas de células de la muestra citológica, a una red se le asignan tareas que deben completarse (una clasificación, por ejemplo) y aprende cómo automatizarlas. Otra ventaja principal de este método es el hecho de que los algoritmos del método de aprendizaje evolucionan con los datos y continúan mejorando al mismo tiempo que aumenta el volumen de los datos. Preferiblemente, estos datos proceden al menos en parte de los datos obtenidos gracias al método de detección según el primer aspecto descrito anteriormente. Más preferiblemente, este segundo aspecto del método de detección podrá acoplar ventajosamente los datos de una imagen de transmisión y de una imagen de fluorescencia, es decir de una primera y de una segunda imagen.

El segundo aspecto del método de detección propuesto en este caso pone en práctica un método de aprendizaje profundo, desviando la funcionalidad de lote para poder procesar un conjunto.

De manera habitual, el tensor en 4 dimensiones (4D) está compuesto por 2D para la componente espacial de la imagen (1D para la altura de la imagen y 1D para la anchura), 1D para los canales (RVB si se usa rojo, verde y azul), y 1D para el lote "B" (B correspondiente al número de imágenes para construir un lote o un paquete).

De manera general, el uso del tensor 4D tiene como única utilidad acelerar los cálculos. La originalidad del presente método es usar este tensor para otra finalidad que es la de clasificar poblaciones, en este caso conjuntos de imágenes reducidas. A diferencia de lo que hacen habitualmente los otros métodos de aprendizaje que toman un lote de B imágenes y producen B decisiones: 1 decisión para cada imagen (el hecho de tomar $B > 1$ sirve simplemente para ganar tiempo), aquí, se toma un lote de B imágenes reducidas (B: correspondiente al número de imágenes reducidas que constituyen el conjunto de la imagen) para dar 1 única decisión basada en el conjunto de imágenes reducidas. La originalidad propuesta en este caso es precisamente usar esta funcionalidad de realizar lotes que existe inicialmente por motivos de tiempo de cálculo y desviar su funcionalidad usándola para otro objetivo y con otra finalidad.

Se considera un conjunto de imágenes reducidas correspondiente a un número B (B deberá ser lo bastante pequeño como para que el lote pueda transferirse a la GPU que tiene una memoria limitada). De manera habitual, los métodos de aprendizajes usan un lote de 32 a 64 imágenes reducidas que van a elegirse para acelerar el tiempo de cálculo. En el contexto de la invención, no se elige un número B predefinido porque se usa este número como número de imágenes reducidas de interés, por ejemplo las imágenes reducidas de células uroteliales por imagen de portaobjetos, este número es diferente de una imagen de portaobjetos a otra (según el número de células uroteliales presentes en la muestra citológica en el portaobjetos). Por tanto, es posible formar un tensor 4D con estas B imágenes reducidas: $B \times RVB \times H \times W$ (B: número de imágenes reducidas, RVB: para los canales de tres colores rojo, verde y azul, H y W respectivamente para la altura y la anchura de las imágenes reducidas). También es posible considerar un número cualquiera de canal clásico RVB para imágenes de microscopía electrónica de transmisión pero también un número cualquiera de canal RVB asociado a F (F: fluorescencia) para un par de imágenes de microscopía electrónica de transmisión y de microscopía de fluorescencia, tal como RVB (rojo, verde y fluorescente). A continuación, si se transfiere este lote a la red neuronal hasta una capa completamente conectada (es decir, una capa de neuronas que no tiene en cuenta el aspecto espacial de la capa anterior y que debido a ello produce como salida únicamente un vector 1D por imagen reducida), la salida tendrá por tanto el tamaño $B \times F$ donde F es el tamaño de esta capa

completamente conectada (por tanto, el tensor 4D se ha transformado en un tensor 2D).

Ahora, es posible transponer este tensor 2D $B \times F$ como tensor $1 \times F \times B$ y aplicar una operación simétrica en B, por ejemplo una combinación 1D (máxima o media). Esta transposición es útil en la medida en que las operaciones simétricas están programadas de manera nativa para tener lugar en las dimensiones 3 y 4 y no en la dimensión 1, siendo la necesidad de usar una capa completamente conectada precisamente liberar estas dimensiones 3 y 4 para la transposición. Este resultado proporciona por tanto un tensor $1 \times F$, donde las características F deberán contener información referente al conjunto de las imágenes reducidas. Se comprenderá que este 1 resulta del hecho de que un conjunto debe conducir a 1 única decisión de clasificación asignada al lote de las B imágenes reducidas. De esta manera, es posible usar un método de aprendizaje profundo clásico para formar una red de extremo a extremo para un problema de clasificación de una imagen de muestra citológica, es decir de un conjunto de imágenes reducidas.

La figura 7 resume la metodología de implementación para un aprendizaje profundo aplicada a la clasificación de una imagen de muestra citológica correspondiente a un conjunto de imágenes reducidas.

La clase final se asigna a un conjunto de imágenes reducidas. Se comprenderá que este método de implementación se aplica en este caso a la clasificación de imagen de muestra citológica, pero puede aplicarse indistintamente a otros campos de manera no limitativa.

La invención propone, según un primer aspecto, un método de clasificación de muestra citológica que permite extraer y calcular características expertas de anomalía celular presente en una muestra para contribuir al diagnóstico del cáncer, con la ayuda de la visión por ordenador. La invención también propone en este caso, según un segundo aspecto, un segundo método de clasificación de muestra citológica según niveles de anomalía celular encontrados en la muestra, usando un método de aprendizaje profundo, que puede contribuir al diagnóstico del cáncer, con la ayuda de la visión por ordenador. La invención también se refiere a un dispositivo de clasificación de muestra citológica, a partir de al menos una imagen digitalizada, adaptada a la puesta en práctica de un método de este tipo. Cada usuario puede definir los umbrales que tiene en cuenta para realizar el método de clasificación. Además, se propone un enfoque general que se adapta a cada usuario según la necesidad, es decir, según el campo de aplicación y la finura del resultado de la clasificación deseada.

REIVINDICACIONES

1. Método de detección de células que presentan al menos una anomalía en una muestra citológica a partir de al menos una primera imagen (1) de microscopía electrónica digitalizada de la muestra, el método se ejecuta con la ayuda de un módulo de procesamiento puesto en práctica por ordenador,
el método comprende una etapa de procesamiento de la primera imagen (110), que comprende las siguientes subetapas:
- a partir de la primera imagen (1), detectar cada elemento (4) presente en la muestra (etapa 112), después filtrar cada elemento (4) elegido de al menos una célula aislada (10) o un grupo de células (14), con la ayuda de una detección colorimétrica,
- a partir de cada grupo de células (14) detectado, individualizar al menos una célula con su núcleo, con la ayuda de al menos un método de segmentación basado en una detección de los contornos en el grupo de células (14) (etapa 116),
el método comprende además una etapa de identificación de anomalía celular (etapa 140), que comprende las siguientes subetapas:
- para cada célula aislada (10) o individualizada en la etapa anterior, calcular datos de al menos dos características de anomalías distintas denominadas "características expertas de anomalía", y determinar si dicha célula presenta al menos una anomalía (etapa 144), comparando estos datos calculados con datos de referencias (etapa 142),
- después contar el número total de células en la muestra que presentan al menos una anomalía (etapa 146),
el método comprende además una primera etapa de caracterización de la muestra (etapa 150) según un nivel de anomalía de la muestra, que comprende las siguientes subetapas:
- comparar el número total de células que presentan al menos una anomalía con un valor umbral del nivel de anomalía, preferiblemente un valor umbral igual a al menos tres células que presentan, cada una, al menos una anomalía,
- si el número total de células que presentan al menos una anomalía es inferior al valor umbral del nivel de anomalía, clasificar la muestra en una clase representativa de una muestra con un nivel de anomalía nulo (152),
- si el número total de células que presentan al menos una anomalía es superior al valor umbral del nivel de anomalía, clasificar la muestra en una clase representativa de una muestra con un nivel de anomalía demostrado (154),
el método comprende además una etapa de selección tipológica (etapa 130), que comprende las siguientes subetapas:
- a partir de la primera imagen (1) digitalizada de la muestra, medir el tamaño de las células, después clasificar las células según su tamaño (etapa 132),
- asociar cada clase de células medidas a un tipo celular correspondiente, después seleccionar al menos un tipo celular elegido, tal como el tipo celular urotelial (etapa 134),
caracterizado porque el método comprende una segunda etapa de caracterización de la muestra (etapa 160) según un nivel de anomalía de las células de la muestra, que comprende además las siguientes subetapas:
- a partir de una segunda imagen (2) de fluorescencia digitalizada de la muestra, seleccionar todas las células del tipo celular elegido según su ubicación en la muestra predefinida (etapa 134) previamente en la primera imagen (1) digitalizada (etapa 162),
- detectar la presencia de un halo de fluorescencia alrededor de cada célula seleccionada anteriormente,
- si se detecta un halo de fluorescencia, clasificar la muestra en una clase representativa de una muestra con un nivel de anomalía celular bajo (168).
2. Método según la reivindicación 1, caracterizado porque al menos una de las características expertas de anomalía es una razón de superficie del núcleo en la célula o una irregularidad del núcleo o una colorimetría moteada o una hipercolorimetría del núcleo.

3. Método según la reivindicación 1 o 2, caracterizado porque al menos una de las características expertas de anomalía es una irregularidad del núcleo, el cálculo de dicha irregularidad del núcleo es un cálculo de desviación correspondiente a la comparación entre cada núcleo de célula analizada y un núcleo tipo de una célula denominada de "referencia", este cálculo de desviación se elige de la siguiente lista: una diferencia de superficie, una diferencia de perímetro, una diferencia de la convexidad del núcleo, una diferencia de baricentro, una diferencia entre una longitud del contorno de cada núcleo detectado, una diferencia de irregularidad de un contorno del núcleo, una diferencia del número de píxeles en cada núcleo y una diferencia de excentricidad del núcleo en la célula.
4. Método según una cualquiera de las reivindicaciones anteriores, en el que, en la segunda etapa de caracterización de la muestra (etapa 160), el tipo celular elegido de las células seleccionadas a partir de la segunda imagen (2) de fluorescencia digitalizada de la muestra es un tipo urotelial.
5. Método según una cualquiera de las reivindicaciones anteriores, en el que, en la segunda etapa de caracterización de la muestra (etapa 160), la subetapa que consiste en detectar la presencia de un halo de fluorescencia alrededor de cada célula seleccionada anteriormente, consiste en medir una tasa de fluorescencia (etapa 164), y en el que la muestra se clasifica en una clase representativa de una muestra con un nivel de anomalía celular bajo (168) si la tasa de fluorescencia medida es superior a una tasa mínima de fluorescencia de referencia.
6. Método según una cualquiera de las reivindicaciones anteriores, en el que, en la segunda etapa de caracterización de la muestra (etapa 160), la subetapa que consiste en detectar la presencia de un halo de fluorescencia alrededor de cada célula seleccionada anteriormente, consiste en medir una tasa de dispersión de la fluorescencia (etapa 164),

y en el que la muestra se clasifica en una clase representativa de una muestra con un nivel de anomalía celular bajo (168) si la tasa de dispersión de la fluorescencia medida es superior a una tasa mínima de dispersión de la fluorescencia.
7. Método según una cualquiera de las reivindicaciones anteriores, caracterizado porque comprende una etapa de recomendación de explotabilidad (etapa 120) de la muestra citológica según un número de células que comprenden un núcleo y según un valor umbral mínimo de explotabilidad, preferiblemente el valor umbral mínimo de explotabilidad es igual a al menos siete células que comprenden un núcleo detectadas en la primera imagen (1).
8. Método según una cualquiera de las reivindicaciones anteriores, caracterizado porque el método con la ayuda del módulo de procesamiento combina el conjunto de los datos obtenidos y clasifica la muestra citológica analizada a partir de la primera imagen (1) y de la segunda imagen (2), en una de las tres clases asociadas a un nivel de anomalía celular de la muestra, nulo, bajo o intenso.
9. Método según una cualquiera de las reivindicaciones anteriores, caracterizado porque dicha muestra citológica es una muestra de orina de micción espontánea.

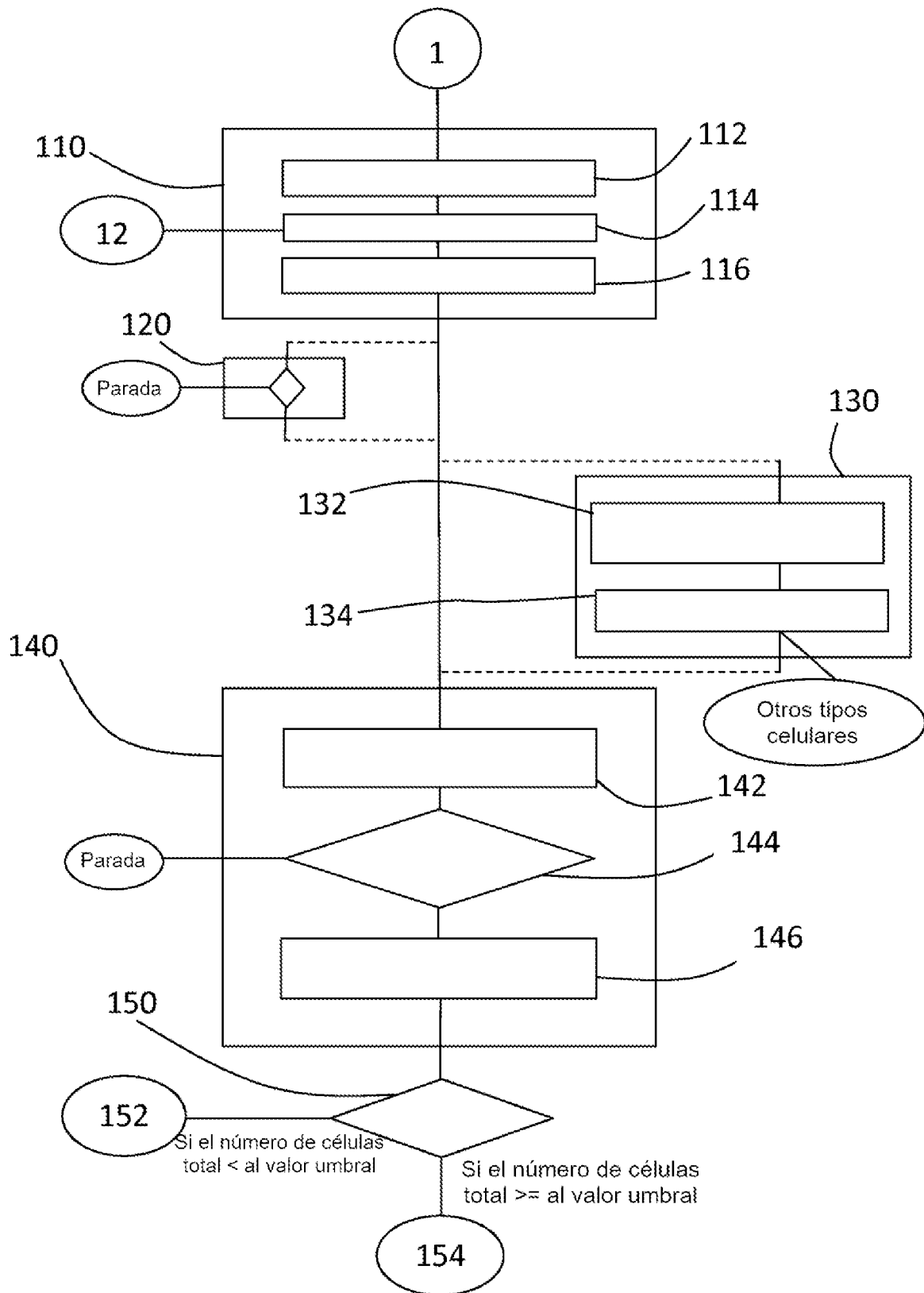


Figura 1

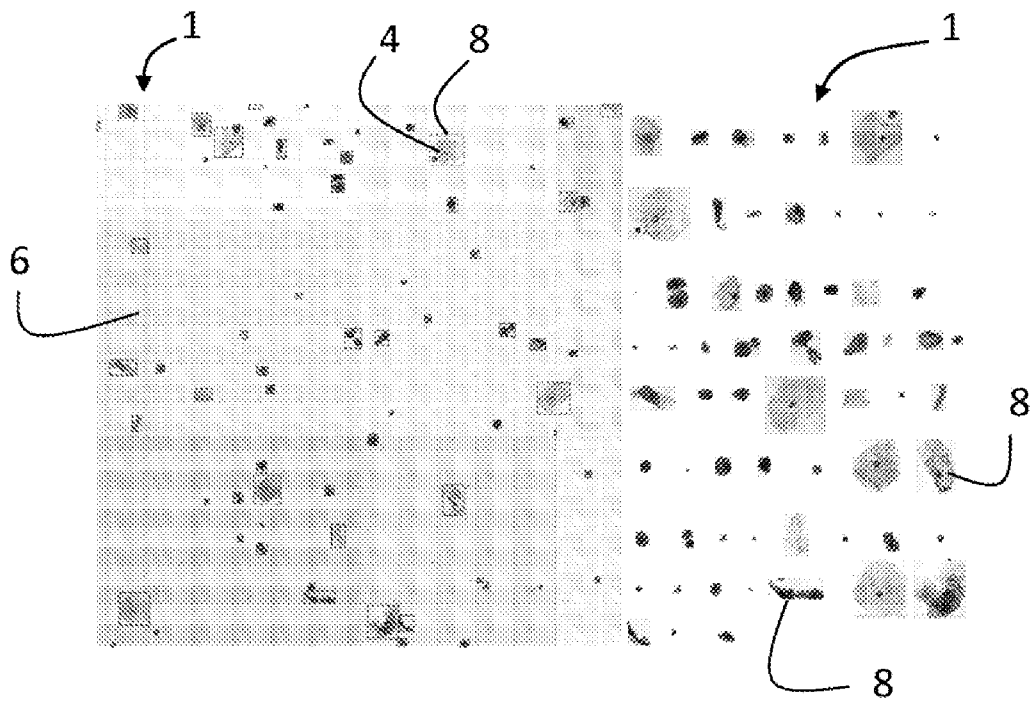


Figura 2

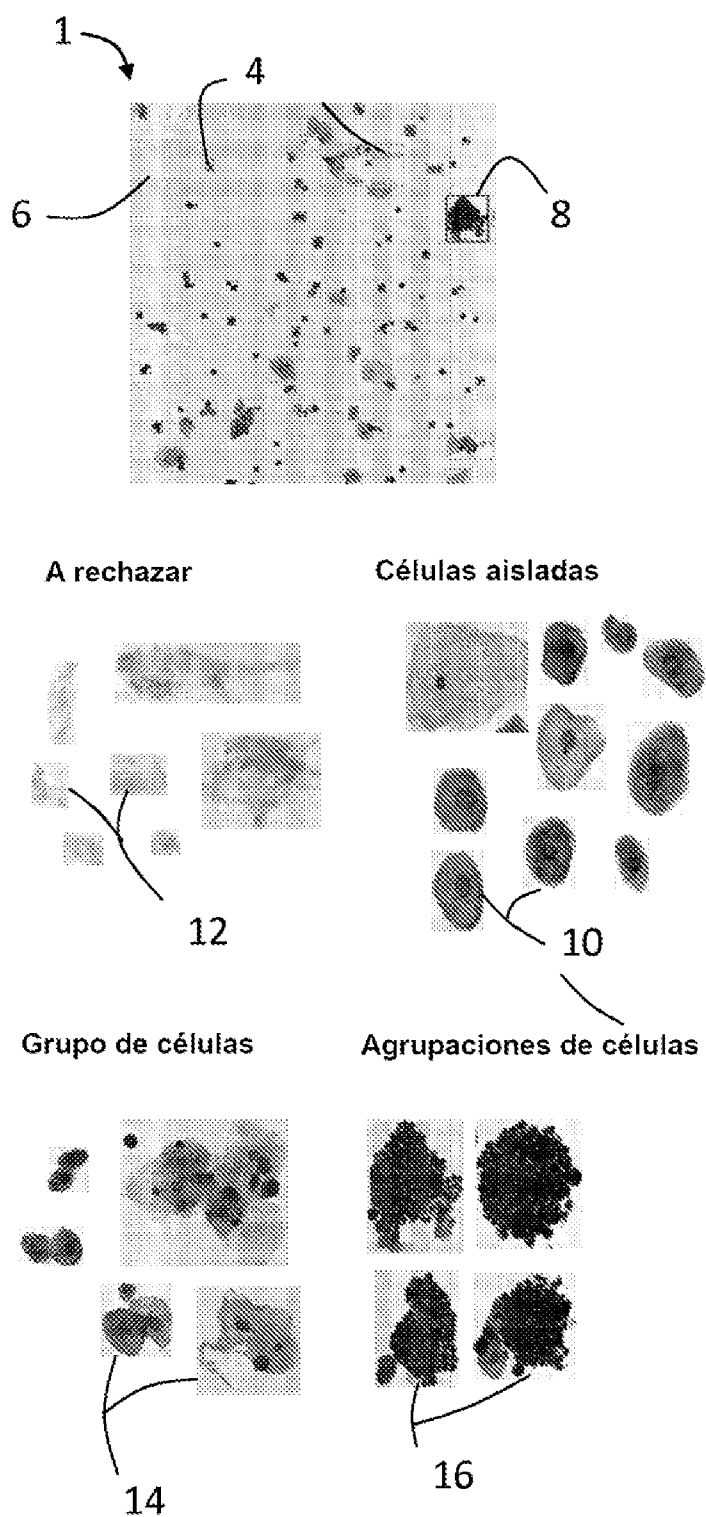


Figura 3

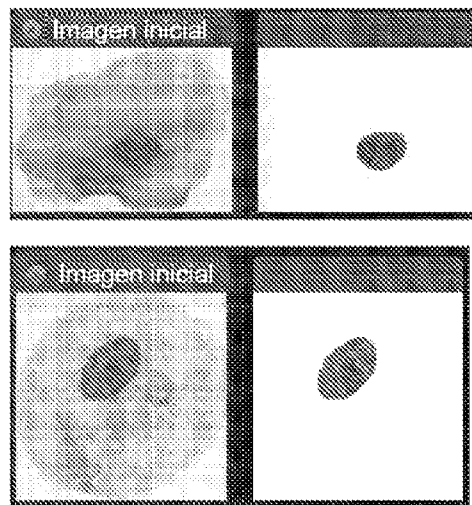


Figura 4

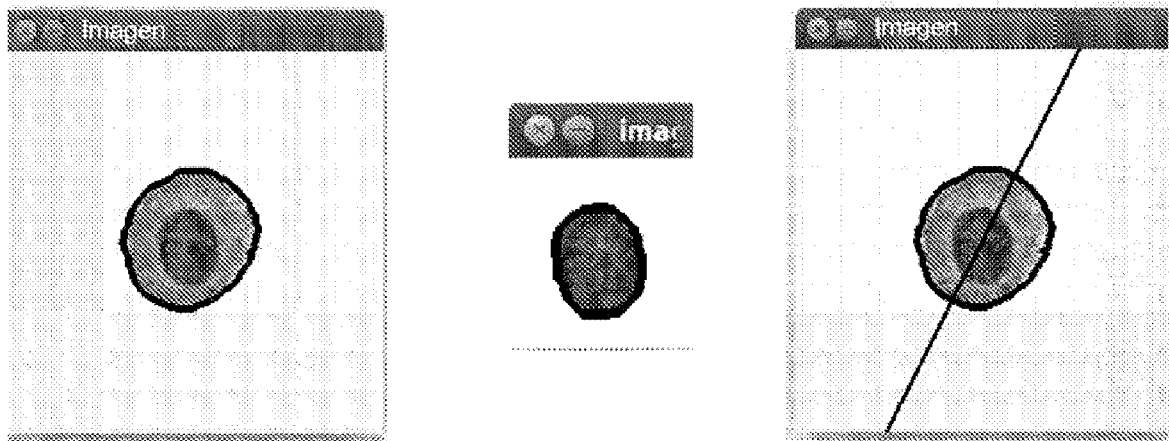


Figura 5

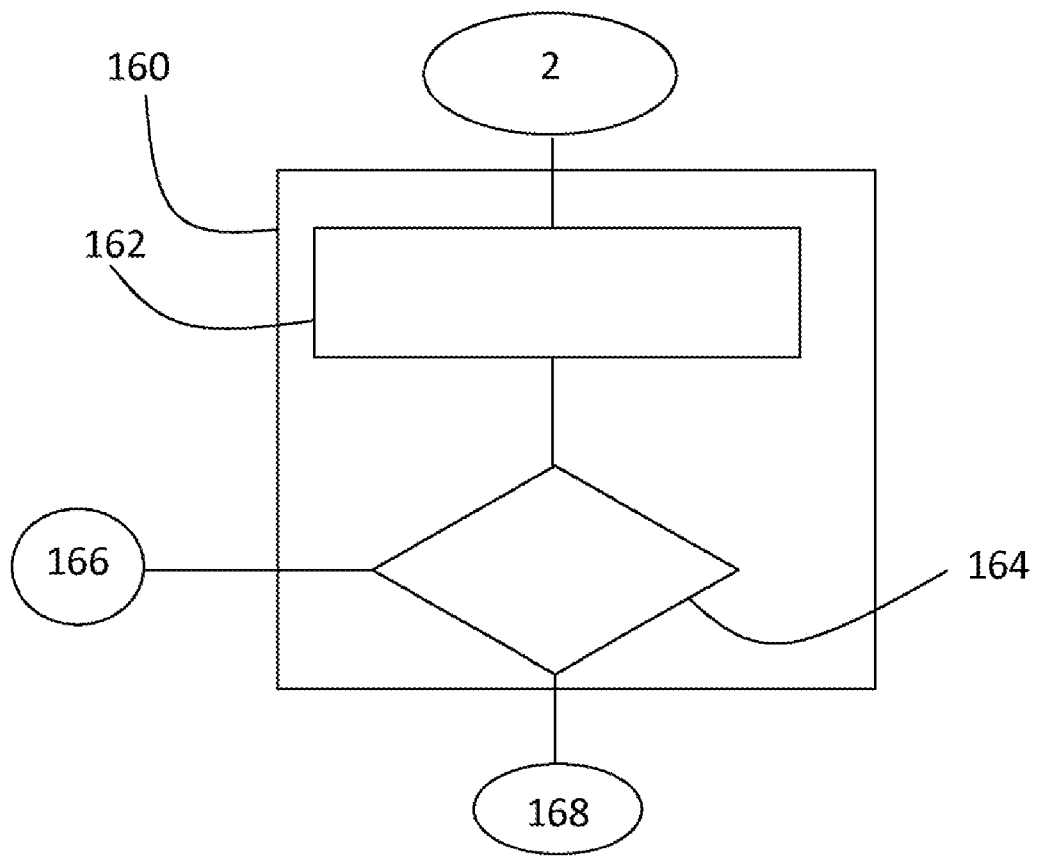


Figura 6

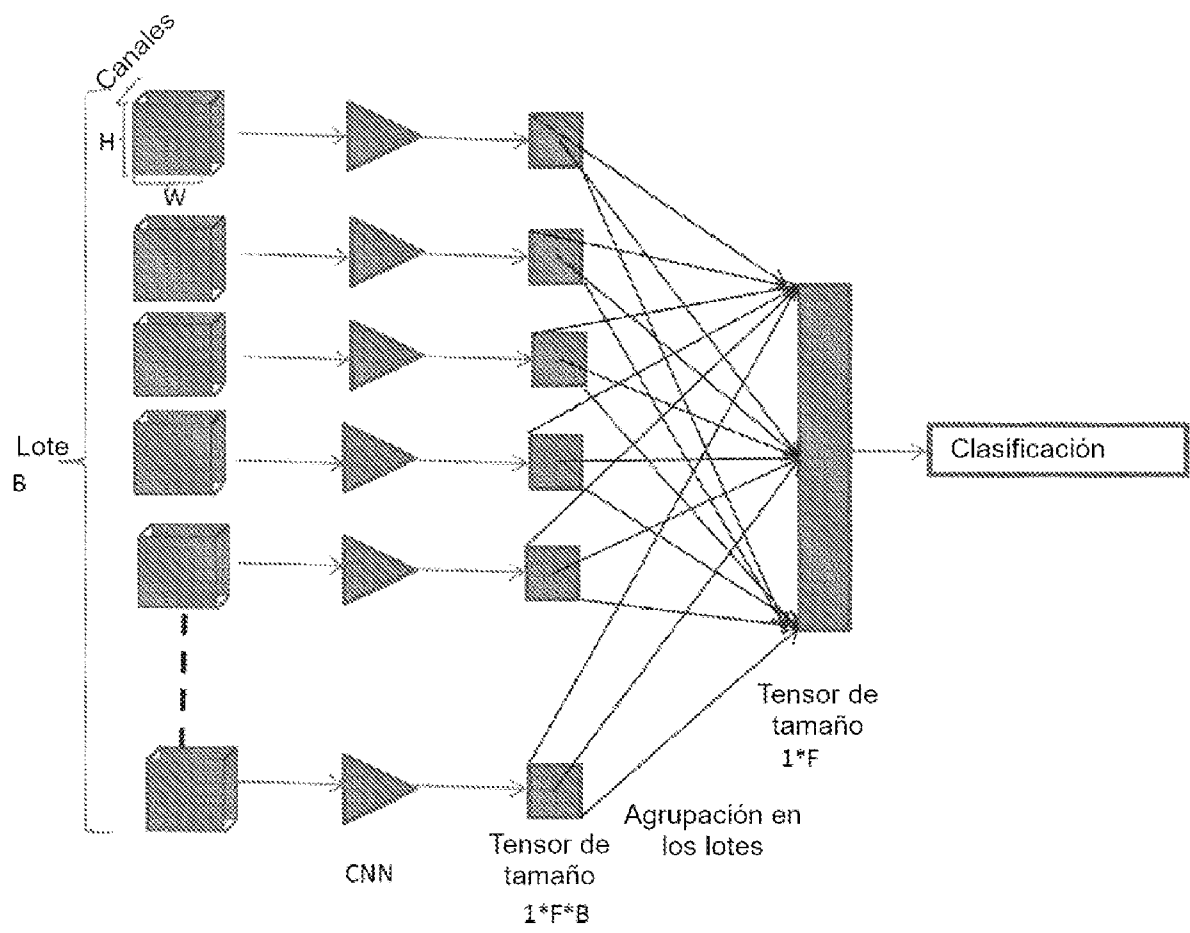


Figura 7