

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 971 682**

51 Int. Cl.:

G06V 10/60 (2012.01)

G06V 40/12 (2012.01)

G06V 10/50 (2012.01)

G06V 10/141 (2012.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **09.04.2020** **E 20382290 (3)**

97 Fecha y número de publicación de la concesión europea: **03.01.2024** **EP 3893147**

54 Título: **Detección de vivacidad usando un dispositivo que comprende una fuente de iluminación**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
06.06.2024

73 Titular/es:
IDENTY INC. (100.0%)
8 The Green, Suite 7471
Dover, DE 19901, US

72 Inventor/es:
ARAGON, JESUS y
GUPTA, HARDIK

74 Agente/Representante:
MILTENYI, Peter

ES 2 971 682 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

Detección de vivacidad usando un dispositivo que comprende una fuente de iluminación

La presente invención se refiere a un método implementado por ordenador para identificar a un usuario de acuerdo con la reivindicación 1.

Técnica anterior

5 Los teléfonos inteligentes y otros dispositivos informáticos móviles se vuelven cada vez más parte de la vida diaria. En vista de esto, también es común hoy en día, por ejemplo, iniciar sesión en cuentas bancarias usando el teléfono inteligente o pagar en un restaurante con un teléfono inteligente.

10 Para facilitar el inicio de sesión del usuario en tales cuentas, la solicitud EP3540835 proporciona medios para identificar a un usuario usando una cámara de, por ejemplo, un dispositivo informático móvil y procesar la imagen respectiva obtenida para identificar al usuario al extraer características biométricas como una huella digital.

15 Si bien esto tiene ventajas significativas en vista de la seguridad, debido a que las características biométricas son muy difíciles de falsificar, es necesario garantizar que la cámara no sea engañada sosteniendo una imagen de las características biométricas respectivas (como la yema del dedo) frente a la cámara y tomar una foto de la misma por parte de un atacante malintencionado para obtener acceso a la respectiva información sensible como la cuenta bancaria.

20 A partir del documento US 2018/0165508 A1, son conocidos un sistema y método para realizar reconocimiento de huellas digitales. El documento se refiere a un sistema y método para capturar las características biométricas de un usuario y generar un identificador que caracteriza las características biométricas del usuario usando un dispositivo móvil tal como un teléfono inteligente. El identificador biométrico se genera usando imágenes capturadas de una pluralidad de dedos de un usuario para los propósitos de autenticación/identificación del usuario de acuerdo con las biométricas capturadas y determinar la vivacidad del usuario. El documento también describe técnicas adicionales para evitar la autenticación errónea provocada por la suplantación de identidad. En algunos ejemplos, las técnicas de anti-suplantación de identidad incluyen capturar una o más imágenes de los dedos de un usuario y analizar las imágenes capturadas para indicaciones de vivacidad.

30 Además, a partir del documento "Multi-Level Feature Abstraction from Convolutional Neural Networks for Multimodal Biometric Identification" de Soleymani et al, es conocido un enfoque que usa una red de fusión multimodal profunda para fusionar múltiples realizaciones para la identificación de usuario.

Objetivo

35 En vista de la técnica previa conocida, es un objeto de la presente invención proporcionar un método que permita distinguir entre un objeto vivo y una suplantación de aquel objeto.

Solución

40 Este problema se aborda mediante el método implementado por ordenador de acuerdo con la reivindicación independiente 1. Se proporcionan realizaciones preferidas en las reivindicaciones dependientes.

45 Se señala explícitamente que los ejemplos referidos como realizaciones en las que una imagen de diferencia se calcula y procesa según una red neuronal de una sola rama como la de la Fig. 3 no están dentro del alcance de la invención.

El método implementado por ordenador para identificar a un usuario de acuerdo con la invención comprende usar un dispositivo informático que comprende una fuente de iluminación que, cuando se activa, emite luz visible, comprendiendo el método tomar dos imágenes de una escena que comprende potencialmente una parte del cuerpo vivo que tiene una característica biométrica, en donde se toma una primera imagen sin que se active la fuente de iluminación y se toma la segunda imagen con la fuente de iluminación activada, transferir la primera imagen y la segunda imagen a una red neuronal y procesar, por medio de la red neuronal, la primera imagen y la segunda imagen, en donde el procesamiento comprende comparar la primera imagen y la segunda imagen, determinando de esta manera si la primera imagen y la segunda imagen son imágenes de una parte del cuerpo vivo, comprendiendo el método, además, si se determina que la primera imagen y la segunda imagen son imágenes de una parte del cuerpo vivo, realizar un algoritmo de identificación para encontrar una característica biométrica para identificar al usuario y si se determina que la primera imagen y la segunda imagen no son imágenes de una parte del cuerpo vivo, no realizar el algoritmo de identificación.

60 Una fuente de iluminación en el sentido de la invención se puede considerar, por ejemplo, la linterna de un teléfono inteligente. Sin embargo, la invención no está limitada a este respecto y cualquier fuente de iluminación que sea parte de un dispositivo informático correspondiente se considera adecuada siempre que se pueda activar selectivamente y

emita luz visible.

En este contexto, también es evidente que la segunda imagen no solo se toma con la fuente de iluminación activada, sino que la escena de la cual se toma la segunda imagen (y específicamente el objeto) también se ilumina con la fuente de iluminación.

La red neuronal puede ser cualquier red neuronal que está preferentemente entrenada para realizar la tarea respectiva de acuerdo con la invención. A este respecto, se pueden preferir realizaciones donde la red neuronal es una red neuronal de convolución que está adaptada específicamente para comparar las imágenes respectivas. Esto significa que se entrenó correspondientemente y también la estructura y arquitectura de red (por ejemplo, el número de capas) están adaptadas a esta tarea.

El algoritmo de identificación puede ser cualquier algoritmo que está adaptado para procesar información de imagen (independientemente de si esta información de imagen se toma de la primera imagen o de la segunda imagen) de tal manera que puede separar o al menos identificar una característica biométrica. Por ejemplo, en este caso se puede usar un algoritmo de recorte que separa una región de interés en la imagen, por ejemplo, la porción o región de la yema de un dedo de otra parte de la imagen.

Es un descubrimiento de la presente invención que una imagen tomada de una parte del cuerpo vivo, como un dedo, sin que se active una fuente de iluminación, se puede distinguir de una imagen correspondiente tomada de la misma parte del cuerpo con una fuente de iluminación adicional activada, tal como una linterna de un teléfono inteligente. Esto se debe a que las diferentes porciones de luz visible (es decir, diferentes longitudes de onda) se reflejan de manera diferente en las partes del cuerpo vivo. Esto significa, por ejemplo, que la luz roja se emite de manera diferente en comparación a la luz azul. Al comparar tales imágenes, se puede determinar esta diferencia y, si no se puede observar ninguna diferencia correspondiente, es muy probable que las dos imágenes tomadas provengan de una suplantación de identidad, como una imagen de la parte del cuerpo vivo real que se sostiene frente a la cámara cuando toma las imágenes.

Por lo tanto, con este método, se puede distinguir de manera fiable entre objetos suplantados y objetos vivos sin que esto requiera componentes adicionales que normalmente no están disponibles para un usuario de un teléfono inteligente.

En otra realización, un método implementado por ordenador para identificar a un usuario de acuerdo con la invención comprende usar un dispositivo que comprende una fuente de iluminación que, cuando se activa, emite luz visible, comprendiendo el método tomar dos imágenes de una escena que potencialmente comprende una parte del cuerpo vivo que lleva una característica biométrica, en donde se toma una primera imagen sin que se active la fuente de iluminación y se toma la segunda imagen con la fuente de iluminación activada, obtener una imagen de diferencia de la primera imagen y la segunda imagen, transferir la imagen de diferencia a una red neuronal y procesar, mediante la red neuronal, la imagen de diferencia, en donde el procesamiento comprende determinar, de la imagen de diferencia, información de diferencia de brillo y determinar, de la información de diferencia de brillo, si la primera imagen y la segunda imagen son imágenes de una parte del cuerpo vivo y, si se determina que la primera imagen y la segunda imagen son imágenes de una parte del cuerpo vivo, realizar un algoritmo de identificación para encontrar una característica biométrica para identificar al usuario y si se determina que la primera imagen y la segunda imagen no son imágenes de una parte del cuerpo vivo, no realizar el algoritmo de identificación.

La imagen de diferencia se puede considerar preferentemente una imagen de diferencia por píxeles de la primera imagen y la segunda imagen, lo que significa que preferentemente los valores de brillo y/o los valores de color de cada píxel se toman de la primera imagen y la segunda imagen y se restan de entre sí, obteniendo de esta manera nuevos valores de píxel para color y brillo para cada píxel en la primera imagen y la segunda imagen. A continuación, esta imagen "nueva" se considera como la imagen de diferencia en el contexto de la invención.

Como se ha explicado anteriormente, hay una diferencia en la característica de reflexión de luz cuando la linterna está encendida o apagada para una parte del cuerpo vivo en comparación con una suplantación de una parte del cuerpo vivo, como una imagen frente a la cámara. A partir del procesamiento respectivo con una red neuronal, a continuación, se puede determinar si la imagen de diferencia constituye una prueba de que la primera y la segunda imágenes se tomaron de una parte del cuerpo vivo o de una suplantación.

Este método tiene la ventaja de que no es necesario procesar ambas imágenes por la red neuronal y realizar una comparación correspondiente, sino que solo es necesario procesar una sola imagen que contiene la información respectiva.

En una realización, comparar la primera imagen y la segunda imagen comprende comparar las imágenes con respecto a una distribución de brillo de píxeles de las imágenes.

Dado que los valores de brillo obtenidos para la primera y la segunda imágenes difieren significativamente para una parte del cuerpo vivo, esto permite una mayor fiabilidad en la distinción entre suplantaciones y objetos vivos.

Además, comparar las imágenes con respecto a una distribución de brillo de píxeles de las imágenes puede comprender obtener valores de mediana de brillo rojo, azul y verde de los píxeles de la primera imagen y de los píxeles de la segunda imagen y determinar, si el valor de mediana de brillo de la segunda imagen excede los valores de brillo de mediana de los píxeles de la primera imagen en un umbral predeterminado, que las imágenes son imágenes de una parte del cuerpo vivo.

Este umbral puede ser un umbral entrenado que permite que la red neuronal identifique y distinga entre suplantaciones y objetos reales (vivos) debido a que, por ejemplo, el umbral se puede establecer a una cantidad para la que es casi seguro que el objeto debe ser un objeto real cuando se excede (o no se excede) el umbral.

En una realización más específica, la distribución de brillo de píxeles para los píxeles rojos, azules y verdes se obtiene para la primera imagen y la segunda imagen y la distribución de brillo de píxeles obtenida de cada color de la primera imagen se compara con la distribución de brillo de píxeles obtenida del color correspondiente de la segunda imagen.

Al separar los valores de brillo para cada uno de los colores y compararlos por separado, es posible un aumento adicional en la fiabilidad de la distinción entre el objeto real y la suplantación.

También se puede proporcionar que el procesamiento de las imágenes por la red neuronal comprenda, antes de la comparación, identificar, usando la red neuronal, una porción de la primera imagen y/o de la segunda imagen que es probable que contenga una parte del cuerpo y extraer información de píxeles de la primera y segunda imágenes correspondientes a esta porción, en donde comparar la primera y la segunda imágenes comprende comparar solo la información de píxeles extraída de la primera y segunda imágenes correspondiente a la porción.

Esto también se puede aplicar al método de acuerdo con la segunda realización descrita anteriormente donde, antes de calcular la imagen de diferencia, se realiza el método de esta realización, es decir, usando la red neuronal, una porción de la primera imagen y/o la segunda imagen que es probable que contenga una parte del cuerpo y la información de píxeles de esta porción se extrae de la primera imagen y la segunda imagen para calcular la diferencia de imagen. Esto reduce la cantidad de píxeles que ha de procesarse y, por tanto, reduce los recursos de cálculo necesarios sin que esto tenga un impacto negativo en la fiabilidad del método.

En una realización adicional, antes de tomar la primera imagen y la segunda imagen, se realiza una etapa de comprobación de iluminación, comprendiendo la etapa de comprobación de iluminación comprobar la condición de iluminación actual y el método comprende, además, si el brillo correspondiente a la condición de iluminación actual excede un umbral dado, no realizar las etapas posteriores del método.

Con esto, se puede garantizar que la detección de vivacidad puede proporcionar resultados fiables también en vista de las condiciones ambientales. De esta manera, se pueden evitar las identificaciones de falsos positivos o falsos negativos.

Más específicamente, si el brillo correspondiente a la condición de iluminación actual excede un umbral dado, presentar al usuario una información de que la identificación no se puede realizar bajo la condición de iluminación actual.

Por ejemplo, esta información puede ser indicativa para el usuario de que debe usar otro medio para distinguir entre objetos reales y suplantaciones de tales objetos o puede simplemente indicar que no se puede realizar ni la detección de vivacidad ni la identificación del usuario usando una característica biométrica obtenida desde la primera y/o segunda imagen. La primera alternativa permite seguir efectuando la identificación empleando otro método, mientras que la segunda garantiza que no sea posible una identificación falsa de un usuario malicioso usando una suplantación de la parte del cuerpo vivo.

También se puede proporcionar que cada una de las etapas del método se realice en un dispositivo informático móvil y en donde la red neuronal está completamente instalada en el dispositivo móvil.

Esto significa que no solo las imágenes se toman en el dispositivo informático móvil, sino que también el procesamiento posterior por la red neuronal y cada etapa posterior se realizan en el dispositivo móvil. De esta manera, no es necesaria una conexión permanente a Internet móvil, por ejemplo, lo que puede ser especialmente ventajoso en regiones del mundo en donde tal Internet móvil no siempre está disponible.

De acuerdo con la invención, la red neuronal comprende tres subredes, en donde una primera subred procesa la primera imagen para extraer primera información de la primera imagen y una segunda subred procesa la segunda imagen para extraer segunda información de la segunda imagen, en donde una tercera subred compara la primera información y la segunda información para determinar si la primera imagen y la segunda imagen son imágenes de una parte del cuerpo vivo.

Esta separación en la primera, segunda y tercera subredes permite emplear específicamente redes neuronales (en

forma de subredes) para las tareas respectivas en el procesamiento de imágenes y obtener una decisión en cuanto a si las imágenes tomadas muestran o no objetos reales o suplantación de ellos.

5 Además, en una realización, determinar si la primera imagen y la segunda imagen son imágenes de una parte del cuerpo vivo comprende decidir si la primera y la segunda imágenes son imágenes de imágenes en papel y/o si la primera y la segunda imágenes son imágenes de imágenes presentadas en una pantalla y/o si la primera y la segunda imágenes son imágenes de suplantaciones en 3D de partes del cuerpo vivo y/o si la primera y la segunda imágenes son imágenes de partes del cuerpo vivo.

10 Dado que, por ejemplo, las imágenes presentadas en papel a una cámara muestran características muy específicas cuando están iluminadas y no iluminadas, es posible que la red neuronal tome una decisión correspondiente, decidiendo de esta manera no solo si la imagen muestra un objeto real o una suplantación del objeto real, sino también decidiendo qué tipo de suplantación es en realidad.

15 En una realización de la segunda realización, la diferencia de brillo se obtiene de una sustracción por píxeles de valores de brillo de la primera imagen y la segunda imagen cuando se obtiene la imagen de diferencia.

20 Esta resta por píxel también se puede realizar en una resta por color y por píxel, lo que significa que para cada píxel y para cada valor de color (verde, azul y rojo) se obtienen los valores de brillo y se realiza la resta por píxel. De esta manera, se obtiene información específica y precisa en cuanto a las características reales de la primera imagen y la segunda imagen y esta información precisa se usa en la imagen de diferencia.

En una realización, la fuente de iluminación es una linterna del dispositivo informático.

25 Dado que tales fuentes de iluminación en forma de linternas se proporcionan hoy en día para dispositivos informáticos móviles tales como ordenadores portátiles, tabletas y teléfonos inteligentes, no es necesario emplear otros dispositivos para realizar los métodos de acuerdo con la invención, lo que facilita que un usuario aplique el método respectivo.

Breve descripción de los dibujos

30 La figura 1 muestra una representación esquemática de la toma de imágenes y contribuciones de brillo correspondientes.

35 La figura 2 muestra un diagrama de flujo de un método de acuerdo con una realización de la invención.

La figura 3 muestra una representación esquemática de cómo una red neuronal procesa la entrada de acuerdo con una realización.

40 La figura 4 muestra una realización del entrenamiento de la red neuronal.

La figura 5 muestra un dispositivo móvil ilustrativo para realizar el método de la invención.

Descripción detallada

45 La figura 1 muestra una representación esquemática de la toma de dos imágenes de un objeto. En la representación de la figura 1, se proporciona un dispositivo informático móvil, tal como un teléfono inteligente, 101 que comprende un sensor óptico (no mostrado por separado) para tomar imágenes de escenas, por ejemplo, de una mano 102 representada en este caso. Es claro que tal escena también puede comprender algún entorno. Sin embargo, para facilitar la explicación, no se muestran en este caso.

50 Cuando se emplean métodos para identificar a un usuario al tomar una imagen de una parte del cuerpo del usuario (como la cara, el ojo o la mano o la palma de la mano), es necesario saber, para la etapa posterior de identificación, si la imagen tomada en realidad es una imagen de la parte del cuerpo vivo o si es una imagen de una suplantación de la parte del cuerpo vivo. Tal suplantación podría ser, por ejemplo, una imagen obtenida por un usuario malintencionado de la persona real y colocada frente a la cámara. Si esta imagen tiene la calidad suficiente, podría potencialmente suplantar los mecanismos de identificación usados hoy en día para características biométricas y, de esta manera, se podría conceder acceso a información confidencial, tales como cuentas bancarias, a un usuario no autorizado mediante el uso de esta suplantación del objeto real.

60 De acuerdo con la invención, esto se puede impedir al tomar más de una imagen del objeto potencialmente real (siempre que no se sepa si es una imagen del objeto real o solo una imagen de la suplantación del objeto) y procesar estas imágenes mediante redes neuronales como se explicará más adelante.

65 Para dar a esto un contexto más general, el lado izquierdo de la figura 1 muestra el dispositivo informático móvil, tal como el teléfono inteligente, 101 tomando una imagen del objeto potencialmente real 102, en este caso una mano. En condiciones normales, esto dará como resultado en cualquier caso en una distribución de brillo como se muestra en

los diagramas 103 y 104. Esta distribución de brillo no es necesariamente la distribución de brillo que se obtendrá bajo cualquier condición de iluminación potencial, sino que solo es un diagrama ilustrativo que muestra las condiciones de iluminación de un objeto para el cual se toma una imagen sin que se active una linterna del dispositivo informático móvil (o cualquier otra fuente de iluminación adecuada). Solo con esto, no se puede determinar si la imagen es de un objeto vivo o una suplantación del objeto vivo.

A continuación, el lado derecho de la figura 1 muestra la misma escena en donde, esta vez, se toma la imagen con la fuente de iluminación del dispositivo informático móvil 101 activada e iluminando la escena con la luz 111.

Los diagramas 105 y 106 muestran la distribución de luz en los diferentes canales de color (como rojo, azul y verde) y los valores de brillo respectivos para esta situación.

Se ve claramente que el diagrama 105 refleja los diagramas 103 y 104 mientras que el diagrama 106 es claramente distinto del diagrama original 104.

Es un hallazgo de la presente invención que se obtiene una diferencia significativa en la distribución del brillo sobre los valores de color respectivos en una imagen si la imagen es una imagen del objeto real, es decir, la mano 102 es de hecho la mano real de un ser humano. En contraste con esto, una suplantación de tal objeto real, específicamente una suplantación en forma de una imagen del objeto real que se sostiene frente a la cámara en las escenas de la figura 1, dará como resultado los diagramas 103 y 105 en donde solo hay, en todo caso, diferencias menores entre la imagen tomada sin que se active una fuente de iluminación del dispositivo informático y una imagen tomada con el dispositivo de iluminación activado.

Esto se puede usar para distinguir entre situaciones donde solo se fotografía una suplantación del objeto real y situaciones donde se fotografía el objeto real. Sin embargo, los diagramas 103 a 106 dependen en gran medida de las condiciones ambientales (luz de fondo, fondo, distancia del objeto a la cámara y similares) y, por lo tanto, no se pueden procesar fácilmente por un algoritmo de cálculo para tomar una decisión correspondiente con respecto a si las imágenes constituyen o no imágenes de objetos reales o suplantaciones de objetos.

Es un hallazgo de la presente invención que una red neuronal que se entrena específicamente para procesar la primera y segunda imágenes correspondientes tomadas de una sola escena es más apropiada para tomar la decisión respectiva con alta precisión, es decir, lo que conduce a pocos falsos positivos y resultados falsos negativos.

La figura 2 muestra un diagrama de flujo esquemático de un método de acuerdo con la invención para procesar primeras y segundas imágenes correspondientes para tomar una decisión respectiva en cuanto a si las imágenes tomadas son o no imágenes de una parte real del cuerpo vivo o no.

En la figura 2, en las primeras etapas 201 y 202, se toman la primera y la segunda imágenes. Para la invención, no tiene relevancia si la primera imagen es la imagen que se toma con la fuente de iluminación activada o si es la segunda imagen que se obtiene con la fuente de iluminación activada. Lo que importa en el contexto de la invención es que una de las imágenes se toma con la fuente de iluminación activada y la otra se toma sin la fuente de iluminación activada. Solo a manera de ejemplo, se supondrá que la primera imagen tomada se toma sin que la fuente de iluminación se active en la etapa 201 y la segunda imagen tomada en la etapa 202 es la imagen que se toma con la fuente de iluminación activada. En vista de esto, es claro que la primera y la segunda imágenes se toman una después de la otra. Para evitar impactos negativos en la evaluación de las imágenes por condiciones cambiantes, el tiempo entre las etapas 101 y 102 debería ser preferentemente muy pequeño. Preferentemente, el tiempo entre la toma de las imágenes es menor que 1 s, preferentemente menor de 0,5 s.

Como ya se ha indicado anteriormente, la fuente de iluminación puede ser, por ejemplo, una linterna de los dispositivos informáticos móviles que se usan hoy en día, como teléfonos inteligentes y tabletas, pero que también se proporcionan potencialmente en ordenadores portátiles.

Las dos imágenes tomadas en las etapas 101 y 102 se pueden procesar previamente en la etapa 203. Este procesamiento previo puede incluir, por ejemplo, algunas etapas de normalización u otras etapas de procesamiento, tal como reducir el número original de píxeles tomados en la primera y la segunda imágenes a un número específico, como 256 x 256 píxeles o separar porciones de las imágenes para procesarlas más y descartar otras porciones de las imágenes. Sin embargo, la etapa 203 no es obligatoria y el proceso descrito en la figura 2 también se puede realizar sin aplicar ningún procesamiento previo a la primera y segunda imágenes. Además, también es posible procesar previamente solo una de las imágenes, por ejemplo, la que se toma con la fuente de iluminación activada o la otra.

En una siguiente etapa 204 también simplemente opcional, se puede realizar una comprobación de iluminación. Esta comprobación de iluminación se puede realizar al emplear un software específico que comprueba las condiciones de iluminación de las imágenes tomadas (o incluso de una tercera imagen tomada independientemente de la primera y la segunda imágenes). Como alternativa o, además, se puede usar hardware específico del dispositivo móvil, como un teléfono inteligente. Si el dispositivo móvil comprende (por ejemplo, como parte de una cámara) un sensor de brillo, la información obtenida de este sensor de brillo se puede usar para realizar la comprobación de iluminación que se

describe a continuación. La información del sensor de brillo se puede obtener sin haber tomado ninguna imagen. Como alternativa o, además, la comprobación de la iluminación se puede realizar usando el sensor de brillo mientras se toman las imágenes o después de que se hayan tomado las imágenes. Además, se puede usar una combinación de obtención de información del sensor de brillo y empleo de software para realizar la comprobación de iluminación (por ejemplo, basándose en las imágenes tomadas) para mejorar los resultados de la comprobación de iluminación.

Si bien esta comprobación de iluminación se realiza después que se hayan tomado las imágenes, esta comprobación de iluminación también se puede realizar antes de que se tomen las imágenes. La comprobación de iluminación puede comprender comprobar el brillo circundante o las condiciones de iluminación en el entorno, por ejemplo, si es de día, de noche o si la persona que sostiene el dispositivo informático móvil o actúa sobre él para realizar el método descrito en la figura 2 está dentro de un edificio con luz artificial y similares. Como se ha explicado anteriormente, las distribuciones de brillo realmente obtenidas (véase la figura 1) dependen en gran medida de las condiciones ambientales. Si las condiciones de iluminación en el entorno son tales que es muy probable que, si las dos imágenes se toman con una imagen tomada con la fuente de iluminación activada y la otra imagen tomada sin la fuente de iluminación activada, no da como resultado razonablemente valores de brillo distinguibles (por ejemplo, en el caso de que haya mucha luz diurna), es posible que no sea razonable realizar más el método de acuerdo con la figura 2, ya que no se puede obtener un resultado fiable. En tal caso, después de haber efectuado la comprobación de iluminación 204, se puede determinar que el método no se puede aplicar razonablemente. En tal caso, el método se puede abortar y/o se puede proporcionar información al usuario en la etapa 205, informando al usuario que las condiciones de iluminación no son tales que el método de identificación se pueda realizar de manera fiable. Esto se puede determinar, por ejemplo, basándose en un valor de brillo de las condiciones de iluminación actuales. Si este valor de brillo excede el umbral potencialmente preestablecido, la información se puede enviar en la etapa 205 al usuario y/o el método se puede cancelar.

Si la condición de iluminación y el valor de brillo correspondiente están por debajo del umbral respectivo, el método puede continuar con la etapa 206 o, en caso de que la comprobación de iluminación 204 se realice antes de que se tomen las primeras y segundas imágenes, el método podría continuar con las etapas 201 y 201. 202,

En la siguiente etapa, las imágenes son procesadas por una red neuronal para determinar si la primera imagen y la segunda imagen son imágenes de una parte del cuerpo vivo o no.

En una realización de la invención, esto comprende procesar la primera y la segunda imágenes al comparar, a través de la red neuronal, las imágenes con respecto a su distribución de brillo de píxeles y obtener, a partir de esto, un resultado que indica que las imágenes se tomaron del objeto real o no. Esto lo determina la red neuronal debido al hecho de que la red neuronal está entrenada específicamente para tomar esta decisión. Las contribuciones de brillo de píxeles de la primera imagen y la segunda imagen pueden procesarse por la red neuronal y durante este procesamiento, la red neuronal procesa cada una de las imágenes (o ambas simultáneamente) a través de su una o más capas, obteniendo de esta manera un resultado específico. Este resultado será casi idéntico para la primera y la segunda imágenes tomadas si las imágenes se toman de suplantaciones del objeto real (véase la figura 1) y diferirá, como también se explica con respecto a la figura 1, si las imágenes tomadas son imágenes del objeto real.

Estas diferencias, sin embargo, no siempre serán las mismas y es, por tanto, necesario disponer de una red neuronal específicamente entrenada (como se explicará con respecto a la figura 3) que, dependiendo de si las imágenes son imágenes de objetos reales o no, dará un resultado específico y bien definido

Específicamente, se puede considerar que cada una de las imágenes, la primera imagen y la segunda imagen se separan en sus respectivos valores de color rojo, azul y verde y/o se separan en la distribución de brillo en los respectivos colores rojo, azul y verde. A partir de esto, para cada una de las imágenes, se puede obtener un valor de brillo de mediana de los píxeles respectivos para cada uno de los colores. A partir de estos valores de brillo de mediana, se puede determinar si la imagen que se toma con la fuente de iluminación activada supera los valores de brillo de mediana de los píxeles de la imagen que se ha tomado sin que la fuente de iluminación esté activada. Si esto se determina de tal manera que el valor de brillo de mediana de la segunda imagen (la tomada con la fuente de iluminación activada) excede el valor de brillo de mediana de la primera imagen en un umbral específico que indica que la segunda y la primera imagen son imágenes tomadas de una parte del cuerpo vivo, la red neuronal puede determinar que la primera y la segunda imágenes son imágenes del objeto real y no una suplantación. Sin embargo, si los valores de brillo de mediana no difieren al menos en el umbral descrito anteriormente, la determinación puede ser negativa y se puede determinar que las primeras y segundas imágenes no son imágenes de la parte del cuerpo vivo.

Más específicamente, la distribución de brillo de píxeles para cada uno de los píxeles rojos, azules y verdes se puede obtener para la primera imagen y la segunda imagen y la distribución de brillo de píxeles obtenida de cada uno de los colores de la primera imagen se puede comparar con la distribución de brillo de píxeles del color correspondiente en la segunda imagen. Esto significa que, en realidad, hay tres etapas de comparación, de acuerdo con esta realización de la invención, donde la distribución de brillo rojo de la primera imagen se compara con la distribución de brillo rojo en la segunda imagen y la distribución de brillo verde en la primera imagen se compara con la distribución de brillo verde en la segunda imagen y la distribución de brillo azul en la primera imagen se compara con la distribución de brillo azul en la segunda imagen. Esto da como resultado tres determinaciones básicamente independientes que

pueden dar como resultado una evaluación más fiable de si la primera y la segunda imágenes muestran o no la parte del cuerpo vivo.

Para este procesamiento, también se puede considerar aplicar, por ejemplo, durante el procesamiento previo 203, una etapa donde se usa la red neuronal u otra red neuronal correspondientemente adaptada para separar una porción en la primera y segunda imágenes que muy probablemente contiene la parte del cuerpo de interés, por ejemplo, la yema del dedo, de la cual, como característica biométrica, se extraerá la huella digital para su posterior identificación. En este caso, esta denominada "región de interés" ROI se puede separar, por ejemplo, durante la etapa de procesamiento previo y a continuación se puede procesar por la red neuronal respectiva.

Si bien la realización anterior se ha descrito con referencia a una sola red neuronal, también se puede pensar en realizaciones en donde se emplee más de una sola red neuronal para realizar las etapas respectivas.

Específicamente, como se ha explicado anteriormente, antes de realizar realmente la determinación respectiva al comparar los valores de la primera imagen y la segunda imagen, se obtiene información, como la distribución del brillo, de cada una de las imágenes. Si bien la información de brillo también se puede obtener sin usar una red neuronal, otra información puede ser más difícil de obtener de las imágenes y puede requerir emplear una red neuronal. Para ello, se puede emplear una primera "subred" para extraer la respectiva información de la primera imagen mientras que se emplea una segunda "subred" para obtener la respectiva información de la segunda imagen. Las subredes son, por supuesto, también redes neuronales específicamente entrenadas. Más específicamente, la primera subred se puede entrenar para obtener información de una imagen que se toma sin que se active una fuente de iluminación del dispositivo informático móvil, mientras que la segunda subred se adapta y entrena específicamente para obtener la información correspondiente de una imagen que se toma con la fuente de iluminación del dispositivo informático que está activado.

Dado que las redes neuronales en general se entrenan usando imágenes ilustrativas de escenas que deben poder identificar, es muy posible que los parámetros de las respectivas redes neuronales (la primera subred y la segunda subred) difieran al menos ligeramente, dará como resultado una extracción más exacta de información correspondiente de las imágenes para las que se entrenaron las redes neuronales.

A continuación, se puede proporcionar una tercera subred (también una red neuronal entrenada específicamente) para realizar la determinación respectiva al comparar la información obtenida de la primera imagen y la información obtenida de la segunda imagen, determinando de esta manera si la primera y la segunda imágenes son imágenes de la parte real del cuerpo vivo o no.

Esta determinación también puede comprender una determinación más sofisticada, donde se determina que la primera y la segunda imágenes son imágenes de una parte del cuerpo vivo o también se puede decidir si la primera y la segunda imágenes son imágenes en papel que, a continuación, se fotografían con el dispositivo informático móvil y/o si la primera y la segunda imágenes son imágenes de imágenes presentadas en la pantalla y/o si la primera y la segunda imágenes son imágenes de suplantaciones 3D de partes del cuerpo vivo (como, por ejemplo, imitaciones de silicona). Esta determinación es posible con una red neuronal entrenada correspondientemente, debido a que cada una de las suplantaciones y también la parte del cuerpo realmente viva presenta características específicas de reflexión de la luz en las que se puede entrenar una red neuronal. El resultado de la respectiva determinación, en aquel caso, no sería "binario" en el sentido de que indica que las imágenes son imágenes de suplantaciones o imágenes de partes vivas del cuerpo, pero el resultado sería más complejo en el sentido de que también se identificaría el tipo de suplantación.

En una realización alternativa a los casos descritos anteriormente, también se puede proporcionar que la primera y la segunda imágenes no se procesen realmente en la etapa 206 por una red neuronal, sino que, antes de que se realice realmente el procesamiento por la red neuronal, se obtiene una imagen de diferencia de la primera y segunda imágenes. Esta imagen de diferencia se puede obtener, por ejemplo, mediante una sustracción por píxeles de los valores de color y/o valores de brillo de la primera imagen y la segunda imagen. Esto se puede realizar, por ejemplo, al restar el valor de color rojo de la primera imagen del valor de color rojo de la segunda imagen y correspondientemente para los otros valores de color, para cada uno de los píxeles de las imágenes. Adicional o alternativamente, los valores de brillo correspondientes se pueden restar para cada uno de los píxeles en la primera imagen de los valores de brillo correspondientes de los píxeles respectivos en la segunda imagen. Esto da como resultado una "imagen de diferencia" que tiene el mismo tamaño que la primera imagen y la segunda imagen pero que muestra solo las diferencias en los valores de brillo y/o color. Ya que son aquellas diferencias las que contienen la información más significativa a la hora de intentar identificar si la primera imagen y la segunda imagen constituyen imágenes de los objetos reales o no, es razonable hacer esta determinación antes de realizar el procesamiento con la red neuronal.

Una vez hecho esto, también se puede proporcionar en una realización de la invención que, de esta imagen de diferencia, se obtenga información de diferencia de brillo y esta información de diferencia de brillo se procese a continuación por la red neuronal para determinar, de esta información de diferencia de brillo, si la primera imagen y la segunda imagen son imágenes de una parte del cuerpo vivo. Esta determinación se obtiene básicamente mediante el

mismo procesamiento que se indicó anteriormente para la realización en donde las dos imágenes se comparan entre sí, pero sin requerir que la red neuronal realice la comparación real. Esto puede tener ventajas con respecto a las capacidades de procesamiento requeridas por el dispositivo informático y las capacidades reales de la red neuronal, lo que facilita su implementación también en dispositivos informáticos con menos recursos de cálculo disponibles y específicamente en dispositivos informáticos que no tienen acceso a recursos de cálculo externos tales como servidores externos a través de una conexión a Internet móvil.

Para esta realización, pero también para la realización descrita anteriormente en donde las dos imágenes se comparan entre sí, se puede proporcionar que todas las etapas del método respectivo se realicen en el dispositivo móvil, lo que también es válido para las etapas descritas a continuación. De esta manera, el método de la invención se vuelve ventajosamente aplicable a situaciones en donde el usuario del dispositivo informático móvil no tiene acceso a potencia de procesamiento externa. Sin embargo, también se puede proporcionar que no todas las etapas del método se realicen en un dispositivo móvil, sino que, por ejemplo, solo las imágenes se tomen realmente por el dispositivo móvil y, a continuación, se envíen a un servidor externo que efectúa el procesamiento adicional y, a continuación, solo proporcionar el resultado de este procesamiento de vuelta al dispositivo móvil donde, a continuación, se puede realizar, por ejemplo, la identificación del usuario o la extracción de la característica biométrica.

Independientemente de cómo se realice realmente la etapa 206 o dónde se realicen las etapas del método, el método procederá con una etapa 207 donde se determina realmente si la primera y la segunda imágenes son imágenes de una parte del cuerpo vivo o no. Si no lo son, el método puede simplemente terminar en la etapa 209 o se puede enviar información a alguna persona, por ejemplo, la persona que normalmente se identificaría con esta característica biométrica exhibida en la primera y segunda imágenes, proporcionándole información que alguien deseaba obtener acceso no permitido a su información confidencial usando suplantaciones de su característica biométrica.

En caso que se determine que la primera y la segunda imágenes constituyen imágenes de una parte real del cuerpo, el método puede proceder extrayendo (etapa 210), de las imágenes (si esto no se hizo ya durante el procesamiento previo), la característica biométrica. Esto se puede hacer, por ejemplo, al separar una porción de la imagen que contiene la característica biométrica, la llamada región de interés ROI, de otras porciones de la imagen y solo procesando esta porción de la imagen, por ejemplo, extraer las minucias de las yemas de los dedos o extraer información específica de la cara del usuario o el iris de un ojo del usuario o las líneas de la palma del usuario.

Habiendo extraído la característica biométrica en la etapa 210, el método puede continuar a continuación a una etapa opcional 211 en donde se usa realmente la característica biométrica extraída de las imágenes. Este uso de la característica biométrica puede comprender, por ejemplo, identificar al usuario para quien ahora se ha probado que las imágenes usadas para proporcionar la característica biométrica muestran la parte del cuerpo vivo real del usuario con su cuenta bancaria u otra información de acceso sensible restringida.

La figura 3 muestra el procesamiento interno de una entrada recibida en una capa 300 de una red neuronal que se puede emplear para decidir si un objeto en las dos imágenes es o no un objeto vivo o una suplantación de un objeto vivo. De acuerdo con una realización de la invención, la red (a la que también se hará referencia a continuación) es una red neuronal convolucional (Convolutional Neural Network, CNN). Sin embargo, también se pueden usar otras redes donde el procesamiento de la entrada para obtener una salida depende de la estructura de la red neuronal usada.

La capa mencionada anteriormente 300 puede ser una capa que es, en el orden de procesamiento de la entrada original a través de la red neuronal, la primera capa que recibe la entrada original después de, por ejemplo, algún procesamiento previo. La entrada recibida puede ser cualquiera de las dos imágenes o solo una imagen o la imagen de diferencia como se ha explicado anteriormente.

En cualquier caso, la capa 300 recibirá una entrada 330 que al menos de alguna manera se corresponde con las imágenes obtenidas originalmente o la imagen de diferencia (o alguna correspondencia procesada previamente). Esta entrada se proporciona preferentemente en forma de al menos una matriz que tiene la dimensión $N \times M$ en donde N y M son números enteros mayores que 0. La matriz puede, por ejemplo, representar los píxeles en la imagen para al menos un valor de color (por ejemplo, rojo) y/o valores de brillo correspondientes. Por lo tanto, las entradas en esta matriz pueden tener valores que correspondan al valor del color respectivo (en el caso del ejemplo, rojo) de este píxel específico y/o su brillo. Como será evidente a partir de lo siguiente, la entrada puede no ser idéntica a las imágenes obtenidas o la imagen de diferencia, pero puede ser una matriz P que se obtuvo de la matriz que representa las imágenes originales o la imagen de diferencia mediante algún procesamiento a través de capas precedentes en la red neuronal o incluso por algún procesamiento previo (por ejemplo, reducción en resolución como se ha explicado anteriormente).

Sin embargo, para facilitar el análisis, se supondrá que la entrada 330 corresponde a la matriz $N \times M$ que representa al menos una de las imágenes obtenidas originalmente y cada entrada en esta matriz $N \times M$ corresponde al valor de un color (por ejemplo, rojo) de un píxel en la imagen respectiva. La aplicación de esta enseñanza a cualquier otra matriz transformada que se origine de la matriz original $N \times M$ y se obtiene por medio de procesamiento de esta matriz en capas de la red neuronal o al caso en donde (solo) se proporcionan valores de brillo para cada píxel en la matriz

correspondiente es directa.

Siguiendo ahora el proceso ejemplificado en la figura 3, la entrada 330 se recibe por la capa convolucional en profundidad 311 para procesamiento. A continuación, se dará un ejemplo simple comparable con respecto a cómo la matriz de entrada 330 puede procesarse por la capa convolucional en profundidad. Esto implicará que se use un núcleo K para calcular productos internos con la matriz. El núcleo se ejecuta sobre la matriz en los llamados "pasos". Si bien el siguiente ejemplo usará valores para anchuras de paso horizontales y verticales de 1, se puede usar cualquier otro valor mayor que 1 siempre que las anchuras de paso sean números enteros mayores que 0. El núcleo K tiene un tamaño de $S \times T$, en donde S y T son números enteros y menores que N y M.

Además, se supondrá que sólo se usa la matriz de entrada original I (es decir, la matriz de entrada 330) de tamaño $N \times M$ para calcular el producto interno con el núcleo. Sin embargo, también se contempla que se pueda usar una matriz Z extendida para calcular los productos internos con el núcleo. Esta matriz ampliada Z se obtiene al "adjuntar", a la matriz original I, líneas y filas por encima de la primera línea y por debajo de la última línea, así como a la izquierda de la primera fila y a la derecha de la última fila.

Esto se llama "relleno". El relleno normalmente comprenderá que se añade un número P_w de líneas en dirección de la línea y un número P_h de filas en la dirección de fila. El número P_w puede ser igual a $S-1$ y el número P_h puede ser igual a $T-1$, de tal manera que cualquier producto interno calculado entre Z y el núcleo contiene al menos una entrada de la matriz original I. La matriz resultante Z será, por lo tanto, de tamaño $(N + 2P_w) \times (M + 2P_h)$. En vista de esto, la matriz Z tendrá las siguientes entradas:

$$Z_{cd} = \begin{cases} 0 & \forall c \leq P_w \\ 0 & \forall c > P_w + N \\ 0 & \forall d \leq P_h \\ 0 & \forall d > P_h + M \end{cases} \quad \{I_{ij} \text{ donde } c = i + P_w; d = j + P_h; i = 1 \dots N; j = 1 \dots M$$

En este contexto, se deduce que la nueva matriz obtenida al calcular todos los productos internos y disponerlos apropiadamente de acuerdo con líneas y filas será en general de tamaño $\left(\frac{N-S+2P_w}{W_w} + 1\right) \times \left(\frac{M-T+2P_h}{W_h} + 1\right)$ donde W_w y W_h definen la anchura de paso en dirección de las líneas y dirección de las filas, respectivamente. Es claro que solo se permiten aquellos rellenos y aquellas anchuras de paso para un núcleo K dado con tamaño $S \times T$ que dan como resultado números enteros para el tamaño de la nueva matriz. Además, las anchuras de paso W_w y W_h son preferentemente menores que S y T, respectivamente, ya que, de otra manera, el núcleo se movería sobre la matriz I de una manera que algunas líneas o filas de la matriz original quedarían fuera al calcular la nueva matriz.

Para facilitar el análisis, se supondrá a continuación que no se proporciona relleno a la matriz I original y que la anchura de paso es 1 para pasos horizontales y verticales. Además, se supondrá que el núcleo es una matriz de tamaño $S \times S$, es decir, el caso especial donde se supondrá $S=T$. La aplicación de las explicaciones que se dan a continuación al relleno arbitrario y anchura de paso, así como a cualquier tamaño de núcleo, es directa con la enseñanza que se proporciona a continuación.

En la capa convolucional profunda 311, la matriz de entrada recibida 330 se usa para formar un producto interno con el núcleo K que tiene el tamaño $S \times S$, donde $S < N, M$. El producto interno se calcula para cada matriz reducida de la matriz original $N \times M$ donde la matriz reducida es de tamaño $S \times S$ y contiene entradas coherentes en la matriz original $N \times M$. Por ejemplo, considerando $S=3$, la primera matriz reducida R de la matriz original $N \times M$ comprende las entradas $i = 1, 2, 3; j = 1, 2, 3$ de tal manera que la matriz reducida $(N \times M)_s$ está comprendida de nueve entradas y se calcula el producto interno con el núcleo K que da como resultado un solo número. La siguiente matriz reducida en las direcciones de las líneas de la matriz original $N \times M$ es la matriz donde i aumenta en 1, de tal manera que la siguiente matriz en esta dirección está constituida de los elementos de la matriz original $N \times M$ donde $i = 2, 3, 4; j = 1, 2, 3$. A continuación, esta matriz se puede usar para calcular el siguiente producto interno con el núcleo. Se observará que el ejemplo dado de la matriz $S \times S$ con $S = 3$ es solo un ejemplo y también se pueden usar otros núcleos.

Para calcular la siguiente matriz reducida R del tamaño $(N \times M)_s$ en dirección de las filas/columnas, el índice j de elementos en la matriz original $N \times M$ se incrementa en 1. Esto se hace hasta la última matriz reducida en dirección de las líneas, donde $i = N - S + 1, N - S + 2, N - S + 3$ en el caso para $S = 3$. Para las filas, esto se hace de manera correspondiente en donde $j = M - S + 1, M - S + 2, M - S + 3$. Calculando aquellos productos internos, una nueva matriz, la matriz P se calcula que tiene el tamaño $(N - S + 1) \times (M - S + 1)$. Sus entradas P_{ij} corresponden al producto interno respectivo calculado con la matriz reducida correspondiente de la matriz original $N \times M$ y el núcleo K. Se observará que una matriz de este tamaño, de hecho, se reenviará a la capa convolucional puntual de la capa 300.

El núcleo K constituye entradas que se obtienen a través de un proceso de aprendizaje en donde se entrena la red

neuronal con el fin de identificar apropiadamente los objetos previstos, en el caso de acuerdo con la invención, con el fin de averiguar si una parte del cuerpo vivo se muestra en las respectivas imágenes. El núcleo K usado en la capa 300 de la red neuronal no es necesariamente idéntico en tamaño y entradas a los núcleos usados en otras capas de la red neuronal respectiva. Adicionalmente, no es necesario que las entradas en el núcleo sean idénticas entre sí, sino que al menos constituyan números mayores o iguales a 0. Se puede considerar que las entradas representan "pesos" que se obtienen a través del aprendizaje de la red neuronal.

El resultado del procesamiento de la matriz 330 por la capa convolucional en profundidad es la matriz 331 que tiene, como se ha explicado anteriormente, tamaño $(N - S + 1) \times (M - S + 1)$ en caso de que el núcleo se mueva a pasos a través de la matriz original de $N \times M$ que tiene una distancia de $\Delta i = 1$ en la dirección de las líneas $\Delta j = 1$ en dirección de las filas. Sin embargo, en caso de que estos pasos tengan una distancia mayor como $\Delta i = 2$ o $\Delta i = 3$ (y, potencialmente, en correspondencia para las filas), la dimensión del resultado 331 cambiará correspondientemente como se ha explicado anteriormente.

En el procesamiento posterior, este resultado 331 se reenvía al primer lote normalizado 312 que sigue en el orden de procesamiento representado con las flechas en la figura 3 después de la capa convolucional en profundidad 311. Aquellos componentes se proporcionan para proporcionar amortiguación no deseada de información potencialmente relevante por el procesamiento de las imágenes a través de la red y específicamente a través del núcleo o núcleos usados en la capa o capas.

El normalizador de lotes intenta normalizar la matriz de resultados recibida 331. Esto se obtiene al calcular la suma de cada una de las entradas en la matriz $(N - S + 1) \times (M - S + 1)$ y dividiéndola por el número de entradas en la matriz $(N - S + 1) \times (M - S + 1)$. El valor medio V para $(N - S + 1) \times (M - S + 1)$ (indicado como P a continuación, con elementos correspondientes de la matriz P_{ij} se da como

$$V = \frac{\sum_{ij} P_{ij}}{n \cdot m}$$

donde n y m representan el número de líneas y columnas/filas en la matriz $N \times M$ o el número de líneas y columnas en la matriz P. Los elementos P_{ij} son las entradas de la matriz P donde un elemento P_{ij} es el elemento en la matriz en la línea i y columna j.

A continuación, el normalizador de lotes calcula una matriz reducida P' restando, de cada entrada P_{ij} en la matriz original, el valor medio V, de tal manera que $P'_{ij} = P_{ij} - V$. De esta manera, los valores en la matriz reducida P' se normalizan, de tal manera que se filtran las anomalías en una u otra dirección (valores extremadamente grandes o valores extremadamente bajos).

El resultado 332 creado por el primer normalizador de lotes 312 es una matriz que todavía tiene (en el ejemplo dado en la figura 4) el tamaño $(N - S + 1) \times (M - S + 1)$ ya que, hasta ahora, no se realizó ninguna reducción dimensional de la matriz.

A continuación, el resultado 332 se proporciona a la primera unidad lineal rectificadora 313 que sigue al primer normalizador de lotes 312.

La unidad lineal rectificadora modifica cada entrada en la matriz 332 calculando además nuevas entradas de matriz \bar{P}_{ij} , donde

$$\bar{P}_{ij} = \begin{cases} 0 & \forall P'_{ij} < 0 \\ P'_{ij} & \forall P'_{ij} \geq 0 \end{cases}$$

Esto da como resultado valores que serían menores que 0 después de haber pasado el normalizador de lotes para establecerse en 0, por lo que no tendrían más influencia en el procesamiento posterior en la capa convolucional en profundidad que se explicará a continuación. Esto significa que, por ejemplo, los valores de color que están por debajo del valor medio calculado en el normalizador de lotes no se consideran más y solo los valores que corresponden al menos al valor medio V tienen influencia en el resultado de la siguiente etapa en el cálculo.

Por lo tanto, el resultado 333 emitido por la primera unidad lineal rectificadora 313 sigue siendo una matriz de forma/tamaño $(N - S + 1) \times (M - S + 1)$ y esta matriz se reenvía a la capa convolucional puntual 321.

Esta capa convolucional puntual 321 crea un resultado 334. Este resultado 334 se crea por la capa convolucional puntual 321 tomando cada entrada en la matriz $(N - S + 1) \times (M - S + 1)$ 333 y multiplicando esta entrada por un peso α , α preferentemente es un número mayor que 0 en cualquier caso y este número es idéntico para cada entrada en la

matriz $(M - S + 1) \times (M - S + 1)$. El resultado 334 que se obtiene de la capa convolucional puntual 321, por lo tanto, es una matriz que tiene el mismo tamaño $(N - S + 1) \times (M - S + 1)$ pero donde cada entrada se multiplica por el peso α .

5 A continuación, el resultado 334 se proporciona al segundo normalizador de lotes 322, donde se normaliza de la manera explicada para el primer normalizador de lotes 312 y se calcula una matriz normalizada P' de la misma dimensión que el resultado 335 y esta matriz/resultado 335 se reenvía a la segunda unidad lineal rectificadora 323 en

donde se aplica una función de rectificación para obtener un resultado/matriz \overline{P} 236 que a continuación se reenvía a la siguiente capa en la red neuronal o, si no sigue ninguna otra capa en la red neuronal, el resultado 336 se proporciona como una salida.

10 A continuación, el resultado del procesamiento de las imágenes de entrada permite tomar una decisión en cuanto a si las imágenes de entrada constituyen o no imágenes de una parte del cuerpo vivo (esto se encontrará durante el procesamiento de las imágenes o la diferencia de imagen debido al diferente comportamiento de las partes del cuerpo vivo a la iluminación con la luz de la fuente de iluminación) o no, es decir, son suplantaciones de una parte del cuerpo vivo. Dependiendo de cómo de sofisticada sea la red neuronal usada, se puede hacer una distinción adicional con respecto a qué tipo de suplantación de identidad se usa realmente.

20 En la figura 4, ahora se explicará cómo se puede entrenar apropiadamente la red neuronal de tal manera que los pesos del núcleo K y el peso α explicados con respecto a la figura 3, así como los patrones que de hecho permiten determinar que las dos imágenes están mostrando partes reales del cuerpo vivo o solo suplantaciones de las mismas son aprendidas por la red neuronal.

25 El método de la figura 4 comienza con la provisión de datos de entrenamiento 401. Los datos de entrenamiento pueden estar constituidos por una pluralidad de imágenes de, por ejemplo, yemas de los dedos o una pluralidad de dedos representados en una imagen junto con otros objetos. Además, los datos de entrenamiento contendrán imágenes de dedos reales/otras partes del cuerpo vivo e imágenes que muestren suplantaciones de tales partes del cuerpo vivo, por ejemplo, imágenes de imágenes de dedos. Adicionalmente o como alternativa, los datos de entrenamiento pueden comprender imágenes diferentes como se ha analizado anteriormente para procesarse por la red neuronal.

30 Las imágenes se pueden multiplicar usando, de la misma imagen, copias giradas, resaltadas, oscurecidas, ampliadas o modificadas de otra manera que se introducen como datos de entrenamiento.

35 Además, se proporcionarán resultados 408 correspondientes a la entrada en donde aquellos resultados son los resultados que la red neuronal debería obtener y, por lo tanto, se usan en el aprendizaje para comparar los resultados del procesamiento por la red neuronal contra los resultados reales y para ajustar los parámetros de la red neuronal en consecuencia.

40 En la etapa siguiente, se proporciona una imagen de entrada específica 402 a la red neuronal en un entorno de entrenamiento donde, además de la red neuronal, se proporcionan un optimizador 407 y un calculador de función de pérdida 406.

45 La imagen de entrada (o un par de imágenes) se procesa, en una primera ronda, usando la capa convolucional en profundidad y el primer normalizador de lotes, así como la primera unidad lineal rectificadora 403, resumida como DCBR y, a continuación, se transfiere a la capa convolucional puntual, al segundo normalizador de lotes y a la segunda unidad lineal rectificadora, resumida como PCBR, donde se procesa de acuerdo con la descripción dada en la figura 3. Esto significa que las etapas o las secciones 403 y 404 representadas en la figura 4 se ejecutan usando, en cada sección 403 y 404 los pesos correspondientes para la capa convolucional puntual (PC) y el núcleo K de la capa convolucional profunda (DC). Los primeros y segundos normalizadores de lotes, así como las unidades lineales rectificadas de los elementos 403 y 404, funcionan de la manera explicada anteriormente con respecto a la figura 3.

50 Como resultado, de acuerdo con la figura 3, se proporcionará una salida. La salida puede tener la forma de un vector $\vec{v} = (a, b, c, d, \dots)$, donde a, b, c, d pueden representar números reales, donde $0 \leq a, b, c, d \leq 1$ y cada uno de a, b, c, d representa una probabilidad de que la entrada represente una parte del cuerpo vivo real (por ejemplo, valor a) o una suplantación de tal parte del cuerpo vivo (valores b, c y d , donde cada uno de b, c y d representa la probabilidad de un tipo diferente de suplantación de identidad). En lugar de 4 de estos valores, se podrían proporcionar más o menos. También se puede proporcionar que solo se produzca un valor de salida que indique (por ejemplo, si es menor que 1) que las imágenes muestran una parte del cuerpo vivo real y (por ejemplo, si es mayor que 1) que las imágenes muestran suplantaciones de una parte viva del cuerpo.

60 A continuación, este resultado 405 se proporcionará a la función de pérdida donde se comparará con los resultados 408 para identificar las diferencias entre el resultado 405 y el resultado correcto 408. Esta diferencia obtenida por la función de pérdida 406 se proporcionará a continuación al optimizador 407 que, a su vez, modificará los pesos de cada capa convolucional puntual y cada capa convolucional en profundidad, es decir, α y las entradas en el núcleo K . Esto significa que, ya sea para todas las capas en la red a la vez o para cada capa en aislamiento, se manipulan el

peso α de la capa convolucional puntual y las entradas en el núcleo K de la capa convolucional profunda.

5 Con aquellos nuevos valores, el ciclo se repite para la misma imagen o imágenes y el vector resultante $\vec{v} = (a, b, c, d, \dots)$ se proporciona a la función de pérdida y se compara con los resultados correctos 408, cuyo resultado se proporciona a continuación al optimizador 407 que, una vez más, modifica los pesos.

10 Este procedimiento se realiza siempre que la diferencia entre el vector resultante $\vec{v}^{(n)}$ y, específicamente, la determinación hecha si las imágenes muestran partes del cuerpo vivas reales o suplantaciones de las mismas comparadas con los resultados 408 exceden un umbral dado que, en esencia, corresponde a la precisión de identificación que se pretende.

15 Después de eso, la siguiente imagen o imágenes de entrada 402 se toman de los datos de entrenamiento 401 y se proporcionan los cuadros delimitadores correspondientes a la función de pérdida. A continuación, el proceso explicado se repite nuevamente para la nueva imagen y se obtienen los pesos óptimos para la capa convolucional puntual y la capa convolucional profunda. Esto se repite hasta que una combinación específica de pesos da como resultado una precisión de identificación apropiada para todas las imágenes de entrada. A continuación, la combinación de pesos que se obtiene se emite como pesos finales 410.

20 A continuación, estos pesos finales se introducen en la aplicación que ejecuta el método inventivo en el dispositivo móvil.

25 Por lo tanto, en el concepto de la presente invención, la red neuronal que se proporciona al dispositivo móvil ya está completamente adaptada a la identificación de imágenes de partes del cuerpo vivo y suplantaciones de tales partes del cuerpo vivo y, por lo tanto, se puede emplear sin que sea necesario ningún aprendizaje adicional, lo que reduce aún más los recursos de cálculo necesarios en los dispositivos móviles.

30 En total, mediante el uso de las capas convolucionales puntuales, las capas convolucionales profundas y los normalizadores de lotes, así como las unidades lineales rectificadas como se ha explicado anteriormente con referencia a las figuras y realizaciones anteriores, se puede proporcionar una aplicación de tamaño pequeño, que permite por lo tanto su uso en un dispositivo móvil de forma aislada, incluso sin ningún acceso a fuentes de datos adicionales a través de Internet o similares. Esto lo hace apropiado para aplicación en entornos en donde no es posible el acceso a redes inalámbricas o similares. Adicionalmente, la potencia del procesador necesaria para ejecutar esta aplicación se reduce al mínimo y, al mismo tiempo, proporciona resultados de detección de suplantación de identidad apropiados, lo que permite extraer las características biométricas de las partes del cuerpo vivo a partir de imágenes confirmadas (es decir, no suplantadas) que a continuación se pueden usar para fines posteriores en la identificación realizada del usuario por las características biométricas, como se ha explicado anteriormente.

40 Para dar un contexto donde se puede llevar a cabo el método inventivo, la figura 5 representa un dispositivo móvil en forma de teléfono inteligente de acuerdo con una realización de la invención.

45 El dispositivo móvil 500 se implementa como un teléfono inteligente como se conoce actualmente. Comprende un sensor óptico 520 (por ejemplo, una cámara) preferentemente en la parte trasera del dispositivo móvil que está enfrente del lado del dispositivo móvil 500 en el que se proporciona la pantalla 530. De acuerdo con la invención, se proporciona una fuente de iluminación (por ejemplo, como parte del sensor óptico 520) que puede activarse y desactivarse selectivamente.

50 La cámara puede ser una cámara con una resolución de 1 MP, 2 MP o incluso más, por ejemplo, una cámara HD. Puede proporcionarse con una linterna como fuente de iluminación, pero la invención no está limitada a este respecto y se puede usar cualquier fuente de iluminación razonable de tal dispositivo móvil. También se puede adaptar para tomar imágenes en tiempo real con una resolución reducida y una vez que se activa la cámara, la pantalla 530 puede mostrar una representación de lo que la cámara "ve" realmente. Esto puede ser, por ejemplo, una mano 510.

55 En una realización de la invención, una vez que se lleva a cabo el método inventivo para las imágenes tomadas, el resultado de la determinación se puede mostrar en la pantalla (no se muestra en la presente).

REIVINDICACIONES

1. Un método implementado por ordenador para identificar a un usuario, usando el método un dispositivo informático que comprende una fuente de iluminación que, cuando se activa, emite luz visible, comprendiendo el método tomar dos imágenes de una escena que comprende potencialmente una parte del cuerpo vivo que lleva una característica biométrica, en donde se toma una primera imagen sin que se active la fuente de iluminación y se toma la segunda imagen con la fuente de iluminación estando activada, transferir la primera imagen y la segunda imagen a una red neuronal y procesar, por la red neuronal, la primera imagen y la segunda imagen, en donde el procesamiento comprende comparar la primera imagen y la segunda imagen, determinando de esta manera si la primera imagen y la segunda imagen son imágenes de una parte del cuerpo vivo, comprendiendo el método además, si se determina que la primera imagen y la segunda imagen son imágenes de una parte del cuerpo vivo, realizar un algoritmo de identificación para encontrar una característica biométrica para identificar al usuario y, si se determina que la primera imagen y la segunda imagen no son imágenes de una parte del cuerpo vivo, no realizar el algoritmo de identificación; en donde la red neuronal comprende tres subredes, en donde una primera subred procesa la primera imagen para extraer primera información de la primera imagen y una segunda subred procesa la segunda imagen para extraer segunda información de la segunda imagen, en donde una tercera subred compara la primera información y la segunda información para determinar si la primera imagen y la segunda imagen son imágenes de una parte del cuerpo vivo.
2. El método de acuerdo con la reivindicación 1, en donde comparar la primera imagen y la segunda imagen comprende comparar las imágenes con respecto a una distribución de brillo de píxeles de las imágenes.
3. El método de acuerdo con la reivindicación 2, en donde comparar las imágenes con respecto a una distribución de brillo de píxeles de las imágenes comprende obtener valores de brillo de mediana de rojo, azul y verde de los píxeles de la primera imagen y de los píxeles de la segunda imagen y determinar, si los valores de brillo de mediana de la segunda imagen exceden los valores de brillo de mediana de los píxeles de la primera imagen en un umbral predeterminado, que las imágenes son de una parte del cuerpo vivo.
4. El método de acuerdo con la reivindicación 3, en donde la distribución de brillo de píxeles para los píxeles rojos, azules y verdes se obtiene para la primera imagen y la segunda imagen y la distribución de brillo de píxeles obtenida de cada color de la primera imagen se compara a la distribución de brillo de píxeles obtenida del color correspondiente de la segunda imagen.
5. El método de acuerdo con cualquiera de las reivindicaciones 1 o 2 a 4, en donde procesar las imágenes por la red neuronal comprende, antes de comparar, identificar, usando la red neuronal, una porción de la primera imagen y/o la segunda imagen que es probable que contenga una parte del cuerpo y extraer información de píxeles de la primera y segunda imágenes correspondiente a esta porción, en donde comparar la primera y segunda imágenes comprende comparar únicamente la información de píxeles extraída de la primera y segunda imágenes correspondientes a la porción.
6. El método de acuerdo con cualquiera de las reivindicaciones 1 a 5, en donde, antes de tomar la primera imagen y la segunda imagen, se realiza una etapa de comprobación de iluminación, comprendiendo la etapa de comprobación de iluminación comprobar la condición de iluminación actual y, comprendiendo el método, además, si el brillo correspondiente a la condición de iluminación actual excede un umbral dado, no realizar las etapas adicionales del método.
7. El método de acuerdo con la reivindicación 6, en donde, si el brillo correspondiente a la condición de iluminación actual excede un umbral dado, presentar al usuario una información de que la identificación no se puede realizar bajo la condición de iluminación actual.
8. El método de acuerdo con cualquiera de las reivindicaciones 1 a 7, en donde cada una de las etapas del método se realiza en un dispositivo informático móvil y en donde la red neuronal está completamente instalada en el dispositivo móvil.
9. El método de acuerdo con cualquiera de las reivindicaciones 1 a 8, en donde determinar si la primera imagen y la segunda imagen son imágenes de una parte del cuerpo vivo comprende decidir si la primera y segunda imágenes son imágenes en un papel y/o si la primera y segunda imágenes son imágenes de imágenes presentadas en una pantalla y/o si la primera y segunda imágenes son imágenes de suplantaciones 3D de partes del cuerpo vivo y/o si la primera y segunda imágenes son imágenes de partes del cuerpo vivo.
10. El método de acuerdo con cualquiera de las reivindicaciones 1 a 9, en donde la fuente de iluminación es una linterna del dispositivo informático.

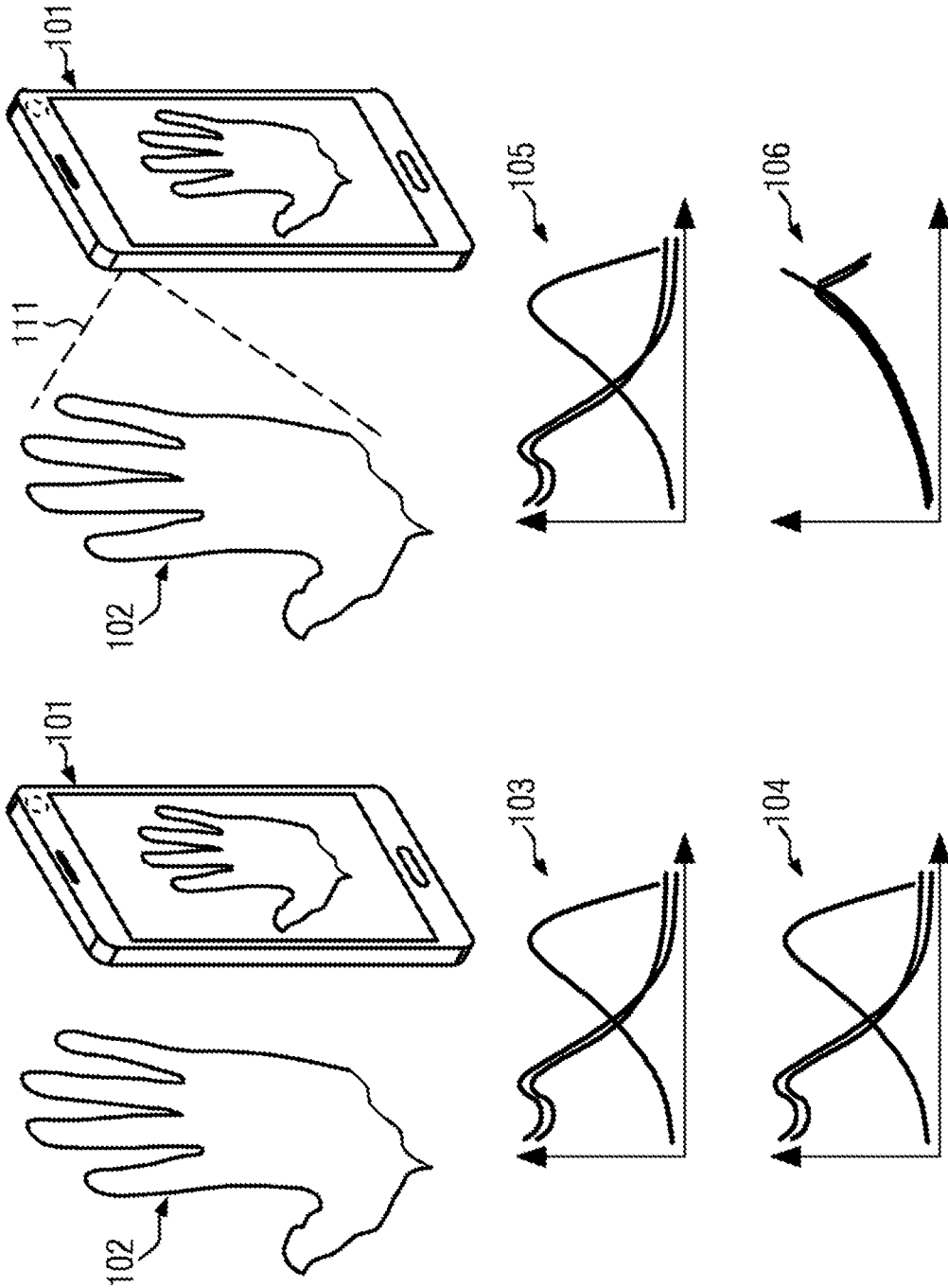


FIG. 1

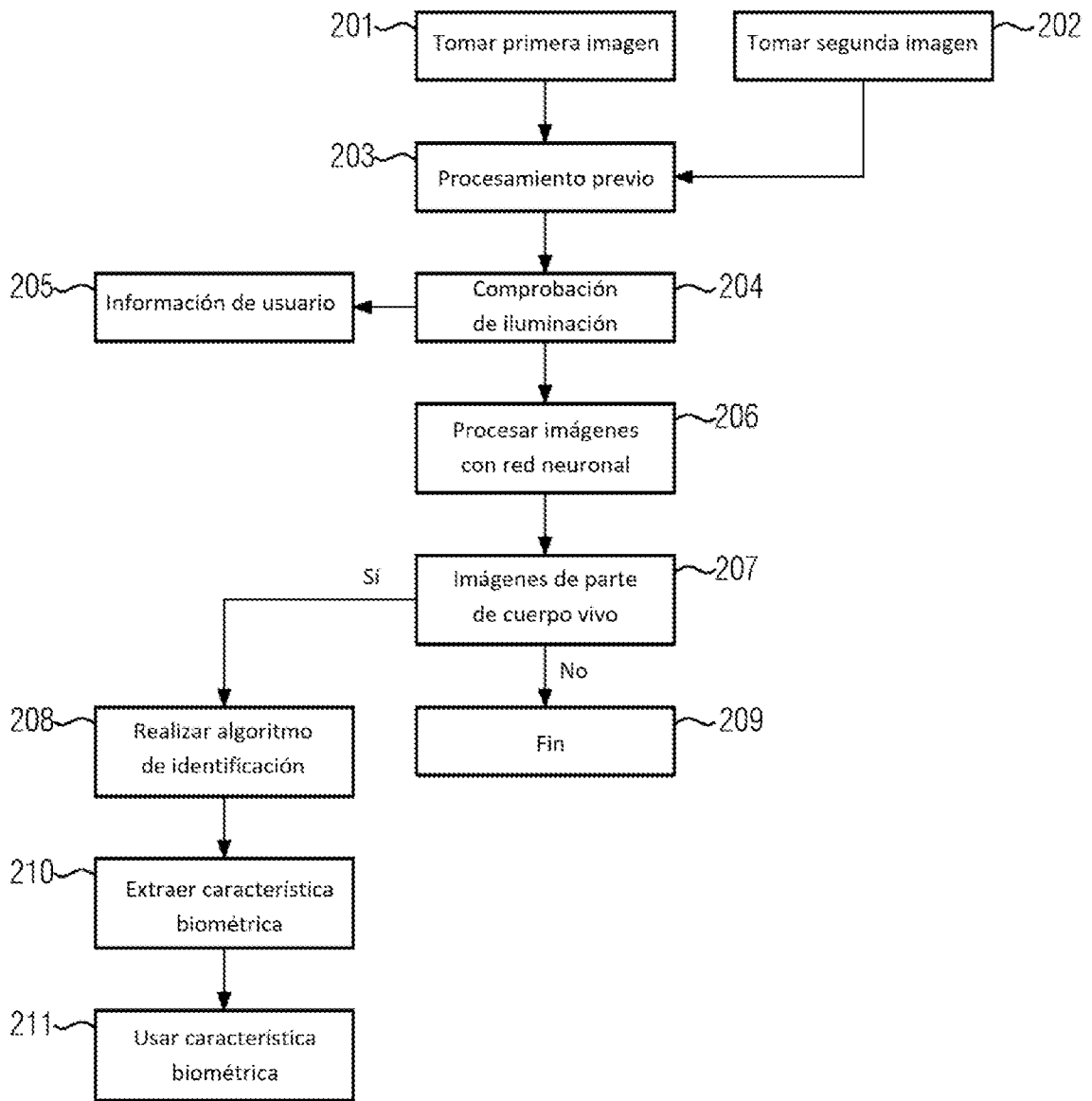


FIG. 2

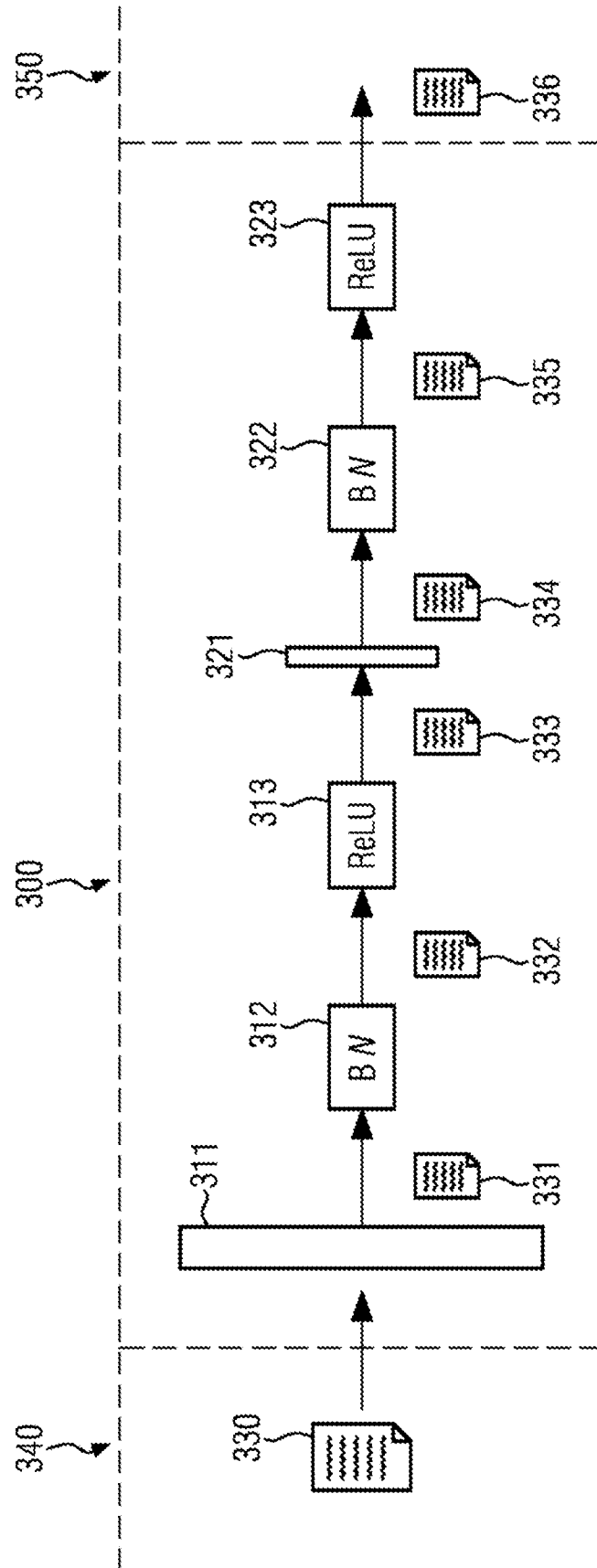


FIG. 3

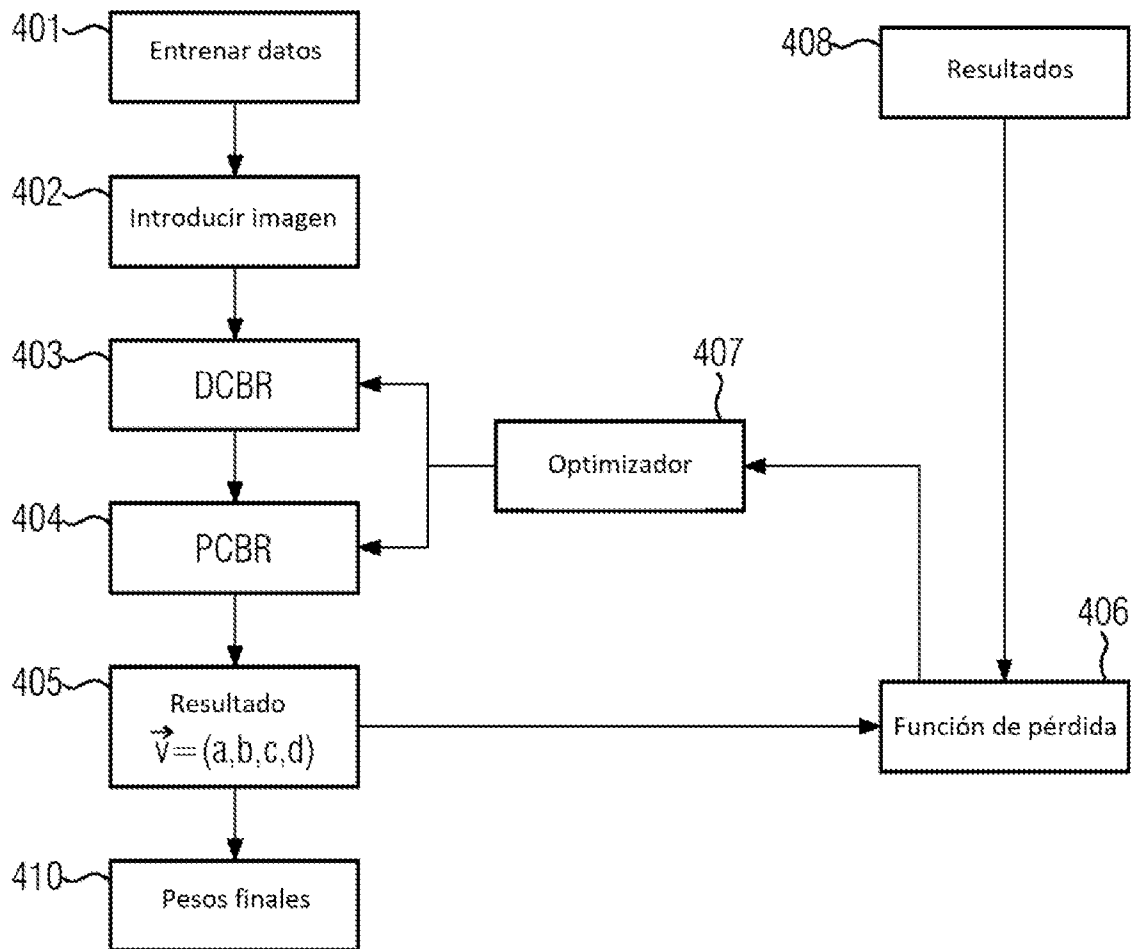


FIG. 4

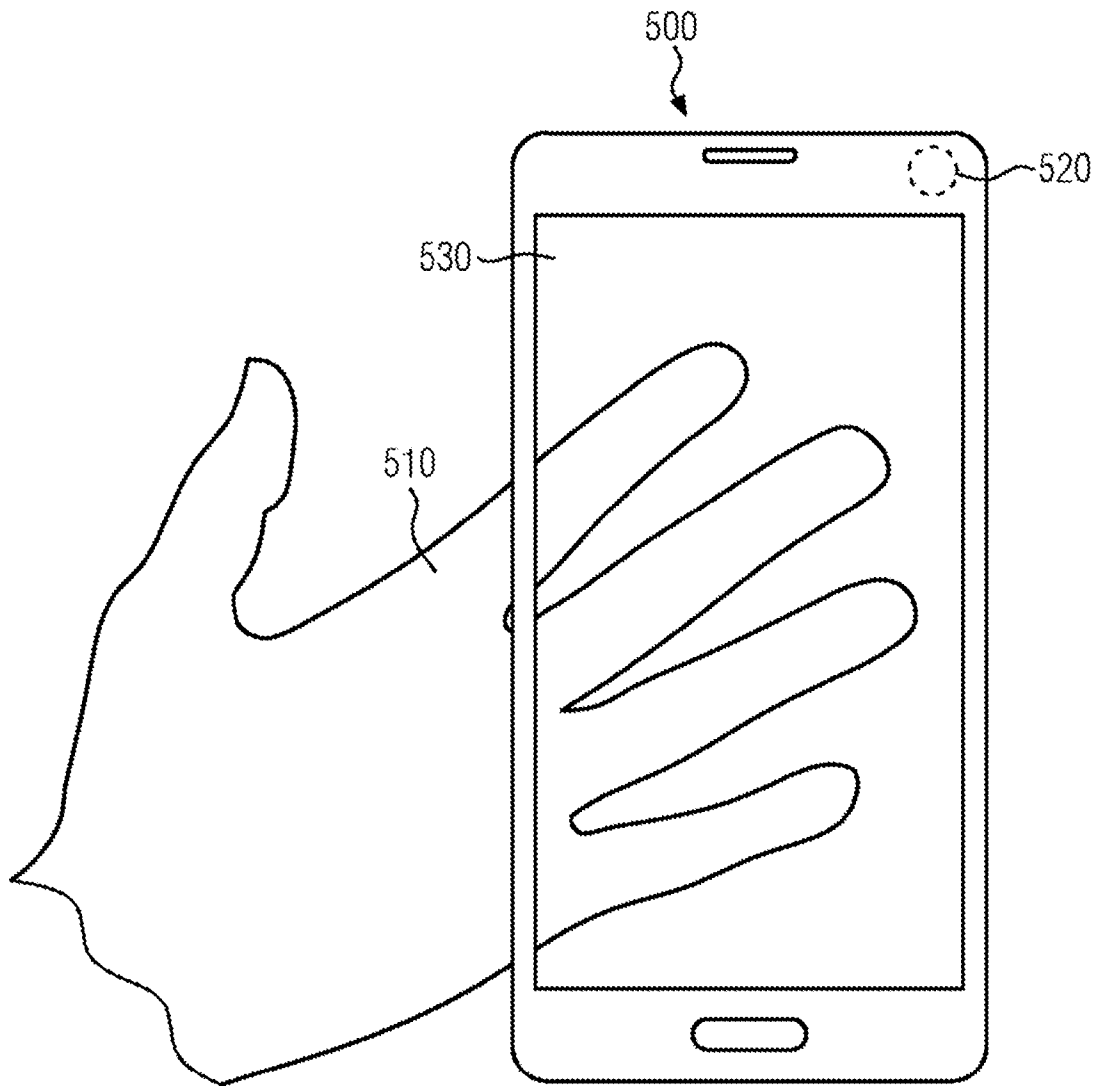


FIG. 5