

SISTEMA Y MÉTODO PARA LA IDENTIFICACIÓN BIOMÉTRICA

5 CAMPO DE LA INVENCION

[0001] La presente invención se refiere, en general, a sistemas y métodos para adquirir y determinar características biométricas y, en particular, a sistemas y métodos para adquirir y determinar características biométricas con el fin de identificar o autenticar a un usuario.

10

FUNDAMENTO DE LA INVENCION

[0002] La información de todo tipo sigue siendo almacenada y accesada de manera remota, tal como en dispositivos de almacenamiento accesibles a través de redes de comunicación de datos. Por ejemplo, muchas personas y empresas almacenan y acceden a información financiera, información médica y de salud, información de bienes y servicios, información de compras, información de entretenimiento, información multimedia a través de Internet u otra red de comunicación. Además de acceder a información, los usuarios pueden efectuar transferencias monetarias (*ej.*, compras, transferencias, ventas o similares). En un escenario típico, un usuario se registra para acceder a la información y luego envía un nombre de usuario y una contraseña para "iniciar sesión" y acceder a la información. Garantizar el acceso seguro a (y desde) tal información y datos almacenados en una red de datos/comunicación sigue siendo una preocupación primordial.

[0003] La conveniencia lleva a los consumidores hacia soluciones de control de acceso basado en biometría. Se cree que la mayoría de los usuarios de teléfonos inteligentes preferiría usar huellas digitales en lugar de una contraseña, y muchos prefiriendo el reconocimiento de iris en lugar del reconocimiento de huellas digitales. La biometría se está convirtiendo cada vez más en un método preferido y conveniente para la detección y verificación de la identidad, y para la autenticación.

[0004] La tecnología de cifrado a nivel de tránsito proporciona una protección relativamente fuerte de la transmisión de varios tipos de datos, incluidos los datos biométricos, y es compatible con los requisitos de confidencialidad, garantía y no repudio. Estándares, tales como el IEEE 2410-2016, brindan protección contra un adversario que escucha la comunicación y proporcionan mecanismos detallados para la autenticación con base en un dispositivo preinscrito y una identidad previa, incluyendo el almacenamiento de

un dato biométrico en forma cifrada. Esto se considera como un caso uno a uno e incluye pasos para enviar y recibir datos biométricos cifrados, en comparación con una muestra biométrica cifrada existente. En consecuencia, tal caso uno a uno se considera como un caso de uso de autenticación, debido a que un vector biométrico y una identidad dada pueden utilizarse como entrada y la autenticación puede ocurrir cuando el vector biométrico coincide con un vector biométrico existente que corresponde a la identidad respectiva.

RESUMEN

[0005] En una o más implementaciones, la presente solicitud proporciona sistemas y métodos implementados por computadora para cotejar un registro biométrico de entrada cifrado con al menos un registro biométrico cifrado almacenado, y sin descifrado de datos de entrada y un registro almacenado como mínimo. Se proporciona un vector biométrico inicial a una red neuronal, y la red neuronal traduce el vector biométrico inicial a un vector de características euclidianas medibles. El vector de características euclidianas medibles se almacena en un almacenamiento con otros vectores de características euclidianas medibles. Además, un vector biométrico actual que representa el registro biométrico de entrada cifrado se recibe desde un dispositivo informático móvil a través de una red de comunicación de datos, y el vector biométrico actual se suministra a la red neuronal. La red neuronal traduce el vector biométrico actual a un vector actual de características euclidianas medibles. Además, se realiza una búsqueda de al menos algunos de los vectores de características euclidianas medibles almacenados en una parte del almacenamiento de datos utilizando el vector actual de características euclidianas medibles. El registro biométrico de entrada cifrado se coteja con al menos un registro biométrico cifrado en el espacio cifrado en función de una distancia absoluta calculada entre el vector actual de características euclidianas medibles y un cálculo de cada uno de los vectores respectivos de características euclidianas medibles en la parte del almacenamiento.

[0006] En una o más implementaciones, la presente solicitud proporciona además la clasificación del vector de características euclidianas medibles y/o la clasificación del vector actual de características euclidianas medibles, en donde la clasificación se realiza al menos en parte utilizando una o más funciones de distancia.

[0007] En una o más implementaciones, la clasificación del vector de características euclidianas medibles y/o el vector actual de características euclidianas medibles devuelve

valores de punto flotante, y se utiliza un algoritmo de Frobenius para calcular una distancia absoluta entre cada punto flotante y su promedio.

[0008] En una o más implementaciones, la búsqueda se realiza en complejidad temporal $\log(n)$. Utilizando un algoritmo de Frobenius para clasificar los vectores

5 biométricos euclidianos medibles que atraviesan una jerarquía de los vectores biométricos euclidianos medibles clasificados en complejidad temporal; e identificar que un vector biométrico medible euclidiano respectivo es el vector actual de características euclidianas medibles.

[0009] En una o más implementaciones, la presente solicitud proporciona la

10 identificación, para cada vector biométrico medible respectivo euclidiano, de una pluralidad de valores de punto flotante, y el uso de un mapa de bits para eliminar a partir de un cálculo de distancias absolutas cualquier valor de una pluralidad de valores que no están presentes en cada vector.

[0010] En una o más implementaciones, la presente solicitud proporciona la

15 identificación, para cada vector biométrico medible euclidiano respectivo, de una pluralidad de valores de punto flotante, y el uso de un mapa de bits para eliminar a partir de un cálculo de distancias absolutas cualquier valor de la pluralidad de valores que no esté presente en cada vector.

[0011] En una o más implementaciones, la presente solicitud proporciona para

20 identificar, por cada vector biométrico euclidiano-medible, una variedad de valores de punto flotante; y definir una escala deslizante de importancia basada en el número de vectores en los que aparece un valor de punto flotante respectivo.

[0012] En una o más implementaciones, la red neuronal se configura con una variedad de capas convolucionales, junto con un rectificador (ReLU) y nodos de agrupación.

25 **[0013]** En una o más implementaciones, la red neuronal está configurada para usar la agrupación como una forma de muestreo descendente no lineal, y adicionalmente en donde uno o más nodos de agrupación reducen progresivamente el tamaño espacial de un vector Euclideano de características medibles representado para reducir la cantidad de parámetros y cálculos en la red neuronal.

30 **[0014]** En una o más implementaciones, la presente solicitud proporciona la identificación de cálculos, para cada uno de una pluralidad de vectores de características euclidiano-medibles almacenados, una diferencia de posición relativa entre un vector de cara

media y el respectivo vector de características euclidiano-medible, para elevar al cuadrado la diferencia de posición relativa, para sumar los valores y para calcular la raíz cuadrada.

[0015] En una o más implementaciones, el rendimiento de la red neuronal se determina con base en una función de costo, en la que varias capas dadas como un tamaño espacial de un volumen de salida se calculan en base de un tamaño de volumen de entrada W, tamaño de filtro (kernel o núcleo) del campo (receptivo), un paso con el que las capas se aplican (deslizan) S, y una cantidad de relleno cero (padding) P usada en un borde.

[0016] Estos y otros aspectos, características y ventajas pueden apreciarse a partir de la descripción adjunta de ciertas modalidades de la invención, las figuras y reivindicaciones adjuntas.

DESCRIPCIÓN DE LAS FIGURAS DE DIBUJO

[0017] La figura 1 ilustra un sistema de ejemplo para la identificación de un usuario de acuerdo con una o más modalidades;

[0018] La figura 2A es un diagrama de bloques que ilustra los componentes y características de un dispositivo informático de usuario de ejemplo, e incluye varios componentes de hardware y software que sirven para permitir el funcionamiento del sistema;

[0019] La figura 2B ilustra una pluralidad de módulos de ejemplo, tales como los que están codificados en un almacenamiento y/o en la memoria, de acuerdo con una o más modalidades;

[0020] La figura 2C es un diagrama de bloques que ilustra una configuración de ejemplo de un servidor del sistema.

[0021] La figura 3 es un diagrama del sistema que ilustra diversos aspectos del sistema de control y seguimiento de la salud de acuerdo con una o más modalidades;

[0022] Las FIGs. 4A y 4B muestran un ejemplo de una red neuronal de ejemplo en funcionamiento, de acuerdo con una implementación de ejemplo de la presente solicitud;

[0023] La figura 5 ilustra un proceso de ejemplo de acuerdo con la red neuronal; y

[0024] La figura 6 es un diagrama de flujo que ilustra los pasos en el proceso de ejemplo de acuerdo con una implementación.

DESCRIPCIÓN DETALLADA

[0025] El cifrado sigue siendo una forma muy popular y efectiva de proteger la información durante el tránsito de la información. La naturaleza de la información a menudo dicta el nivel y el tipo de cifrado utilizado para proteger la información, particularmente para evitar posibles peligros mientras la información está en tránsito. Desafortunadamente, no ha sido posible o práctico cifrar todos los datos almacenados a nivel de aplicación, por ejemplo, debido a la necesidad de buscar los datos. Ha sido inviable, al menos desde una perspectiva de rendimiento, buscar datos cifrados de manera efectiva sin requerir un proceso de búsqueda exhaustivo que incluya el descifrado de datos, registro por registro.

[0026] La información de identificación personal ("PII"), en particular, requiere mecanismos de cifrado y políticas y procesos adicionales para la protección de datos, debido a que varias operaciones en los datos exigen el descifrado de tales datos para su visualización y edición. Por ejemplo, la Ley de Portabilidad y Responsabilidad de Seguros de Salud ("HIPAA") requiere el cifrado de datos durante el tránsito y ofrece políticas para la divulgación y difusión de dichos datos. La política de seguridad criptográfica está destinada a proteger contra los posibles peligros de las bases de datos de PII, como en casos de robo. Al realizar operaciones tales como búsquedas, sin descifrado, los datos no necesitan estar expuestos a posibles peligros. Los datos biométricos requieren protecciones adicionales mediante procesos y políticas que introducen mecanismos adicionales que incluyen esquemas de cifrado más sofisticados, como la criptografía visual.

[0027] La presente solicitud incluye características y funcionalidades para proporcionar búsquedas cifradas, más allá de la comparación de registros uno a uno, dada una referencia biométrica y un nuevo registro de entrada. Además, la presente solicitud proporciona nuevos métodos y sistemas para buscar datos biométricos cifrados sin necesidad de descifrado de datos en medios de almacenamiento, tales como bases de datos, sistemas de archivos u otros mecanismos de persistencia. Además de una implementación uno a uno, las implementaciones de uno a varios son compatibles con las enseñanzas de este documento, en las que la búsqueda de registros biométricos cifrados puede ocurrir en función de un registro biométrico recién recibido. En tal caso, se puede realizar una búsqueda exhaustiva del grupo ortogonal $O(n)$, en la que se descifra y compara cada registro. De acuerdo con una o más implementaciones de la presente solicitud, se proporciona una solución $O(\log n)$ que no requiere descifrado y acepta la localización de un registro sin descifrado. En este documento, generalmente, como un caso de uso de identidad, se proporciona un vector biométrico dado

como entrada, y se puede buscar un dato biométrico respectivo para determinar si el dato biométrico está en una base de datos.

[0028] En una o más implementaciones, la presente solicitud proporciona una solución basada en polinomios para la identificación en un almacén grande de datos cifrados de datos biométricos, proporcionando así un sistema y mecanismo para proteger la privacidad. Además, se realiza una selección de un dato biométrico inicial con baja entropía. Posteriormente, una selección de una estructura de datos proporciona una complejidad $\log(n)$ para una búsqueda. Posteriormente, se realiza una selección de un algoritmo para nodos intermedios, de modo que no se pueda descubrir el dato biométrico o no se pueda revertir un algoritmo hash en el dato biométrico original. Por lo tanto, una implementación de acuerdo con la presente solicitud proporciona una técnica de extremo a extremo para que un dato biométrico se utilice para proporcionar una identificación en una base de datos de miles (o más) de individuos.

[0029] En una o más implementaciones, una red neuronal, que puede incluir una red neuronal convolucional, se utiliza para procesar imágenes y para determinar una función de costo correcta, tal como para representar la calidad de rendimiento de la red neuronal. En una o más implementaciones, se utiliza una red neuronal para procesar otros datos, que incluyen el contenido de audio, tal como una grabación o la representación de la voz de una persona. Aunque muchas de las implementaciones de ejemplo mostradas y descritas en este documento se refieren al procesamiento de uno o más archivos de imagen, un experto en la materia reconocerá que la presente solicitud es preferiblemente biométrica agnóstica, y que se puede usar cualquier dato biométrico adecuado de acuerdo con las enseñanzas de este documento. Varios tipos de redes neuronales son adecuados para recibir diversos formatos de información y generar un vector de características, como una red neuronal convolucional, una red neuronal recurrente ("RNN") o un sistema de aprendizaje automático profundo.

[0030] En una o más implementaciones, una red neuronal, que puede incluir una red neuronal convolucional, se utiliza para procesar imágenes y para determinar una función de costo correcta, tal como para representar la calidad de rendimiento de la red neuronal. En una o más implementaciones, se utiliza una red neuronal para procesar otros datos, que incluyen el contenido de audio, tal como una grabación o la representación de la voz de una persona. Aunque muchas de las implementaciones de ejemplo mostradas y descritas en este documento se refieren al procesamiento de uno o más archivos de imagen, un experto en la materia reconocerá que la presente solicitud es preferiblemente biométrica agnóstica, y que se

puede usar cualquier dato biométrico adecuado de acuerdo con las enseñanzas de este documento. Otros tipos de redes neuronales son adecuados para recibir diversos formatos de información y generar un vector de características, como una red neuronal recurrente ("RNN") o un aprendizaje automático profundo.

5 **[0031]** Un sistema de ejemplo para identificar un usuario se muestra como un diagrama de bloques en la Fig. 1, que se puede configurar para interactuar con una red neuronal (no mostrada). En una disposición de ejemplo, se pueden incluir un servidor de sistema 105, un dispositivo informático remoto 102 y dispositivos informáticos de usuario 101a y 101b. El servidor del sistema 105 puede ser prácticamente cualquier dispositivo informático y/o aparato de procesamiento de datos capaz de comunicarse con los dispositivos 101a, 101b y otros dispositivos informáticos remotos 102, que incluyen recibir, transmitir y almacenar la información electrónica y procesar solicitudes como se describe más adelante en este documento. El servidor del sistema 105, el dispositivo informático remoto 102 y los dispositivos de usuario 101a y 101b están destinados a representar diversas formas de computadoras, tales como computadoras portátiles, computadoras de escritorio, estaciones de trabajo informáticas, asistentes digitales personales, servidores, servidores blade, unidades centrales y otras computadoras y/o sistemas informáticos en red o basados en la nube.

15 **[0032]** En una o más implementaciones, el dispositivo informático remoto 102 puede asociarse con una organización empresarial, por ejemplo, una institución financiera, una compañía de seguros o cualquier entidad que mantenga cuentas de usuario empresariales (también conocidas como "cuentas de transacción"). Dichas organizaciones empresariales brindan servicios a los titulares de cuentas y requieren la autenticación del usuario antes de otorgar acceso a los sistemas y servicios empresariales. A modo de ejemplo adicional, el dispositivo informático remoto 102 puede incluir redes de pago y/o redes bancarias para procesar transacciones financieras, como se entiende por los expertos en técnica.

20 **[0033]** Los dispositivos informáticos de usuario 101a, 101b pueden ser cualquier dispositivo informático y/o aparato de procesamiento de datos capaz de incorporar los sistemas y/o métodos descritos en este documento, que incluyen, pero no se limitan a, una computadora personal, tableta, asistente digital personal, dispositivo electrónico móvil, teléfono celular o dispositivo de teléfono inteligente y similares. El terminal de transacciones 101b está destinado a representar varias formas de dispositivos informáticos, tales como estaciones de trabajo, sistemas de punto de venta especializados, terminales ATM, computadoras personales, computadoras portátiles, tabletas, dispositivos de teléfonos

inteligentes, asistentes digitales personales u otras computadoras apropiadas que puedan utilizarse para realizar transacciones electrónicas. Los dispositivos 101a, 101b también pueden configurarse para recibir entradas del usuario, así como capturar y procesar información biométrica, por ejemplo, imágenes digitales de un usuario, como se describen
5 adicionalmente en el presente documento.

[0034] En una o más implementaciones, el servidor del sistema 105 implementa reglas que gobiernan el acceso a la información y/o la transmisión de información entre dispositivos informáticos con los que los usuarios interactúan (*ej.*, el dispositivo 101a, 101b) y uno o más servidores back-end confiables (*ej.*, el dispositivo informático remoto 102).

[0035] Como se describe más adelante en este documento, los sistemas y métodos para identificar y/o autenticar un usuario pueden cumplir con los niveles de seguridad requeridos por un sistema empresarial mediante el uso de una API para integrarse con un sistema existente (*ej.*, un sistema de procesamiento de transacciones y gestión de datos de una institución financiera). En consecuencia, el servidor del sistema 105 no necesita saber si el
10 sistema subyacente (*ej.*, el dispositivo informático remoto 102) es un Sistema de gestión de bases de datos relacionales (RDBMS), un motor de búsqueda, un sistema de procesamiento de transacciones financieras y similares. En consecuencia, los sistemas y métodos para facilitar la autenticación segura ofrecen un mecanismo "point and cut" para agregar la seguridad adecuada a los sistemas empresariales existentes, así como a los sistemas en
15 desarrollo. En algunas implementaciones, la arquitectura del sistema es un lenguaje neutral que permite que REST, JSON y las capas de conexión segura (SSL) proporcionen la interfaz de comunicación entre los diversos dispositivos informáticos (*ej.*, 101a, 101b, 102 y 105). Además, en una o más implementaciones, la arquitectura se basa en la especificación de servlet, capas de puertos seguros abiertos, Java, JSON, REST y/o Apache Solr. En
20 consecuencia, los sistemas divulgados para autenticar a un usuario pueden implementar estándares abiertos, lo que permite una interoperabilidad significativa.

[0036] Debe entenderse además que, si bien los diversos dispositivos y máquinas informáticas a los que se hace referencia en el presente documento, que incluyen, pero sin limitarse a, los dispositivos de usuario 101a y 101b, el servidor del sistema 105 y el
30 dispositivo informático remoto 102 se denominan en este documento como dispositivos y/o máquinas individuales/sencillas. En ciertas implementaciones, los dispositivos y máquinas a los que se hace referencia, y sus operaciones, características y/o funcionalidades asociadas y/o complementarias pueden combinarse o arreglarse o emplearse de otra manera en

cualquier cantidad de dispositivos y/o máquinas, tales como a través de una conexión de red o conexión con cable, como es sabido por los expertos en la técnica.

[0037] La figura 2A es un diagrama de bloques que ilustra los componentes y características de un dispositivo informático de usuario 101a, e incluye varios componentes de hardware y software que sirven para permitir el funcionamiento del sistema, incluidos uno o más procesadores 110, una memoria 120, un micrófono 125, una pantalla 140, una cámara 145, una salida de audio 155, un almacenamiento 190 y una interfaz de comunicación 150. El procesador 110 sirve para ejecutar una aplicación de cliente en forma de instrucciones de software que se pueden cargar en la memoria 120. El procesador 110 puede ser un número de procesadores, una unidad de procesamiento central CPU, una unidad de procesamiento de gráficos GPU, un núcleo de multiprocesador o cualquier otro tipo de procesador, dependiendo de la implementación particular.

[0038] Preferiblemente, la memoria 120 y/o el almacenamiento 190 son accesibles por el procesador 110, permitiendo así que el procesador reciba y ejecute las instrucciones codificadas en la memoria y/o en el almacenamiento para hacer que el dispositivo y sus diversos componentes de hardware transporten operaciones para aspectos de los sistemas y métodos de ejemplo descritos en la presente. La memoria puede ser, por ejemplo, una memoria de acceso aleatorio (RAM) o cualquier otro medio de almacenamiento legible por computadora volátil o no volátil adecuado. El almacenamiento 190 puede adoptar varias formas, dependiendo de la implementación particular. Por ejemplo, el almacenamiento puede contener uno o más componentes o dispositivos, tales como un disco duro, una memoria flash, un disco óptico regrabable, una cinta magnética regrabable o alguna combinación de los anteriores. Además, la memoria y/o el almacenamiento pueden ser fijos o extraíbles.

[0039] Uno o más módulos de software 130 están codificados en el almacenamiento 190 y/o en la memoria 120. Los módulos de software 130 pueden comprender uno o más programas de software o aplicaciones que tienen un código de programa de computadora o un conjunto de instrucciones ejecutadas en el procesador 110. Como se representa en la Fig. 2B, se pueden incluir uno o más de un módulo de interfaz de usuario 170, un módulo de captura biométrica 172, un módulo de análisis 174, un módulo de inscripción 176, un módulo de base de datos 178, un módulo de autenticación 180 y un módulo de comunicación 182 entre los módulos de software 130 que son ejecutados por el procesador 110. Dicho código o instrucciones de programa de computadora configuran el procesador 110 para llevar a cabo operaciones de los sistemas y métodos descritos aquí y se pueden escribir en cualquier

combinación de uno o más lenguajes de programación.

[0040] El código del programa puede ejecutarse por completo en el dispositivo informático del usuario 101, como un dispositivo independiente, en parte en el dispositivo informático del usuario 101, en parte en el servidor del sistema 105, o por completo en el servidor del sistema 105 u otra computadora/dispositivo remoto 102. En el último escenario, la computadora remota se puede conectar al dispositivo informático del usuario 101 a través de cualquier tipo de red, incluida una red de área local (LAN) o una red de área amplia (WAN), una red de comunicaciones móviles, una red celular o la conexión puede hacerse a una computadora externa (por ejemplo, a través de Internet utilizando un proveedor de servicios de Internet).

[0041] También puede ser que el código de programa de los módulos de software 130 y uno o más dispositivos de almacenamiento legibles por computadora (tal como la memoria 120 y/o el almacenamiento 190) formen un producto de programa de computadora que pueda fabricarse y/o distribuirse de acuerdo con la presente invención, como es conocido por los expertos en la técnica.

[0042] Debe entenderse que en algunas modalidades ilustrativas, uno o más de los módulos de software 130 pueden descargarse a través de una red al almacenamiento 190 desde otro dispositivo o sistema a través de la interfaz de comunicación 150. Además, debe tenerse en cuenta que otra información y/o datos relevantes para el funcionamiento de los sistemas y métodos actuales (tales como la base de datos 185) también se pueden almacenar en el almacenamiento. En algunas implementaciones, tal información se almacena en un almacén de datos cifrados que se asigna específicamente para almacenar de forma segura la información recopilada o generada por el procesador 110 que ejecuta los módulos de software 130. Preferiblemente, las medidas de cifrado se usan para almacenar la información localmente en el almacenamiento del dispositivo informático del usuario y para transmitir información al servidor del sistema 105. Por ejemplo, tales datos se pueden cifrar utilizando un cifrado polimórfico de 1024 bits o, según los controles de exportación, un método de cifrado AES de 256 bits. Además, el cifrado se puede realizar utilizando una llave remota (semillas) o llaves locales (semillas). Se pueden usar métodos de cifrado alternativos como entenderan los expertos en la técnica, por ejemplo, SHA256.

[0043] Además, los datos almacenados en el dispositivo informático del usuario 101a y/o el servidor del sistema 105 se pueden cifrar utilizando una clave de cifrado generada a partir de la información biométrica, la información de signos de vida o la información del

dispositivo informático del usuario, como se describe más adelante en este documento. En algunas implementaciones, se puede utilizar una combinación de lo anterior para crear una clave única y compleja para el usuario que se puede cifrar en el dispositivo informático del usuario utilizando criptografía de curva elíptica, preferiblemente de al menos 384 bits de longitud. Además, esa clave se puede utilizar para proteger los datos del usuario almacenados en el dispositivo informático del usuario y/o el servidor del sistema.

[0044] También preferiblemente en el almacenamiento 190 se almacena la base de datos 185. Como se describirá con mayor detalle a continuación, la base de datos contiene y/o mantiene varios ítems y elementos de datos que se utilizan a lo largo de las diversas operaciones del sistema y método para autenticar a un usuario que realiza una transacción financiera en un terminal de transacciones. La información almacenada en la base de datos 185 puede incluir, pero sin limitarse a, un perfil de usuario, como se describirá con mayor detalle en este documento. Cabe señalar que aunque la base de datos 185 se representa como configurada localmente para el dispositivo informático del usuario 101a, en ciertas implementaciones la base de datos y / o varios de los elementos de datos almacenados en ella pueden, además o alternativamente, ubicarse de forma remota (como en un dispositivo remoto 102 o servidor del sistema 105 - no mostrado) y conectado al dispositivo informático del usuario a través de una red de una manera conocida por los expertos en la materia.

[0045] Una interfaz de usuario 115 también está operativamente conectada al procesador. La interfaz puede ser uno o más dispositivos de entrada o salida, tales como interruptores, botones, teclas, una pantalla táctil, micrófono, etc., como se entiende en la técnica de los dispositivos informáticos electrónicos. La interfaz de usuario 115 sirve para facilitar la captura de comandos del usuario, tales como comandos de encendido/apagado o información del usuario y configuraciones relacionadas con la operación del sistema para autenticar a un usuario. Por ejemplo, la interfaz sirve para facilitar la captura de cierta información del dispositivo informático del usuario 101, tal como la información personal del usuario para inscribirse en el sistema con el fin de crear un perfil de usuario.

[0046] El dispositivo informático 101a también puede incluir una pantalla 140 que también está operativamente conectada al procesador 110. La pantalla incluye una pantalla o cualquier otro dispositivo de presentación de este tipo que permite al dispositivo informático del usuario instruir o proporcionar comentarios al usuario sobre el funcionamiento del sistema 100. A modo de ejemplo, la pantalla puede ser una pantalla digital, como una pantalla de matriz de puntos u otra pantalla bidimensional.

[0047] A modo de ejemplo adicional, la interfaz y la pantalla se pueden integrar en una pantalla táctil. En consecuencia, la pantalla también se utiliza para mostrar una interfaz gráfica de usuario, que puede mostrar diversos datos y proporcionar "formularios" que incluyen campos que permiten la entrada de información por parte del usuario. Al tocar la pantalla táctil en lugares correspondientes a la visualización de una interfaz gráfica de usuario, el usuario puede interactuar con el dispositivo para ingresar datos, cambiar configuraciones, funciones de control, etc. De manera que cuando se toca la pantalla táctil, la interfaz de usuario 115 comunica este cambio al procesador 110, y las configuraciones pueden cambiarse o la información ingresada por el usuario puede capturarse y almacenarse en la memoria 120 y/o almacenamiento 190.

[0048] Los dispositivos 101a, 101b también pueden incluir una cámara 145 capaz de capturar imágenes digitales. La cámara puede ser uno o más dispositivos de imágenes configurados para capturar imágenes de al menos una parte del cuerpo del usuario, incluidos los ojos y/o la cara del usuario mientras se usa el dispositivo informático del usuario 101a.

La cámara sirve para facilitar la captura de imágenes del usuario para el análisis de imágenes por parte del procesador del dispositivo informático del usuario configurado, que incluye la identificación de características biométricas para autenticar (biométricamente) al usuario a partir de las imágenes. El dispositivo informático del usuario 101a y/o la cámara 145 también pueden incluir uno o más emisores de luz o señal (no mostrados), por ejemplo, un emisor de luz visible y/o un emisor de luz infrarroja y similares. La cámara se puede integrar en el dispositivo informático del usuario, tal como una cámara frontal o posterior que incorpora un sensor, por ejemplo, y sin limitación, un sensor CCD o CMOS.

Alternativamente, la cámara puede ser externa al dispositivo informático del usuario 101a. Los expertos en la materia entenderán las posibles variaciones de la cámara y los emisores de luz. Además, el dispositivo informático del usuario también puede incluir uno o más micrófonos 125 para capturar grabaciones de audio, como se entiende por los expertos en la técnica.

[0049] Una salida de audio 155 también se puede conectar operativamente al procesador 110. La salida de audio puede integrarse en el dispositivo informático del usuario 101 o puede ser externa al dispositivo informático del usuario y puede ser cualquier tipo de sistema de altavoces que esté configurado para reproducir archivos de audio electrónicos como se entiende por los expertos en la técnica.

[0050] Varios dispositivos/sensores de hardware 160 se pueden conectar

operativamente al procesador. Los sensores 160 pueden incluir: un reloj incorporado para rastrear la hora del día; un dispositivo habilitado con GPS para determinar una ubicación del dispositivo informático del usuario; un acelerómetro para rastrear la orientación y aceleración del dispositivo informático del usuario; un magnetómetro de gravedad para medir el campo magnético de la tierra, sensores de proximidad para medir la distancia desde el dispositivo informático del usuario a un objeto, sensores de radiación RF para medir la radiación y otros dispositivos similares para capturar información sobre el entorno del dispositivo informático del usuario, como se entenderá por los expertos en la técnica.

[0051] La interfaz de comunicación 150 también está conectada operativamente al procesador 110 y puede ser cualquier interfaz que permita la comunicación entre el dispositivo informático del usuario 101a y dispositivos, máquinas y/o elementos externos. Preferiblemente, la interfaz de comunicación puede incluir, entre otros, un módem, una tarjeta de interfaz de red (NIC), una interfaz de red integrada, un transmisor/receptor de radiofrecuencia (*ej.*, Bluetooth, celular, NFC), un transmisor/receptor de comunicación satelital, un puerto de infrarrojos, una conexión USB y/o cualquier otra interfaz para conectar el dispositivo informático del usuario a otros dispositivos informáticos y/o redes de comunicación tales como redes privadas e Internet. Tales conexiones pueden incluir una conexión por cable o una conexión inalámbrica (*ej.*, utilizando el estándar 802.11), aunque debe entenderse que la interfaz de comunicación puede ser prácticamente cualquier interfaz que permita la comunicación hacia/desde el dispositivo informático del usuario.

[0052] La figura 2C es un diagrama de bloques que ilustra una configuración de ejemplo del servidor del sistema 105. El servidor del sistema 105 puede incluir un procesador 210 que está conectado operativamente a diversos componentes de hardware y software que sirven para permitir el funcionamiento del sistema para, por ejemplo, autenticar a un usuario en conexión con una transacción en un terminal de transacciones. El procesador 210 sirve para ejecutar instrucciones para realizar diversas operaciones, incluidas las relacionadas con la autenticación de usuarios y el procesamiento/autorización de transacciones. El procesador 210 puede ser un número de procesadores, un núcleo de multiprocesador o algún otro tipo de procesador, dependiendo de la implementación particular.

[0053] En ciertas implementaciones, el procesador 210 puede acceder a una memoria 220 y/o un medio de almacenamiento 290, permitiendo así que el procesador 210 reciba y ejecute las instrucciones almacenadas en la memoria 220 y/o en el almacenamiento 290. La

memoria 220 puede ser, por ejemplo, una memoria de acceso aleatorio (RAM) o cualquier otro medio de almacenamiento legible por computadora volátil o no volátil adecuado.

Además, la memoria 220 puede ser fija o extraíble. El almacenamiento 290 puede adoptar varias formas, dependiendo de la implementación particular. Por ejemplo, el almacenamiento 290 puede contener uno o más componentes o dispositivos, tales como un disco duro, una memoria flash, un disco óptico regrabable, una cinta magnética regrabable o alguna combinación de los anteriores. El almacenamiento 290 también puede ser fijo o extraíble.

[0054] Uno o más módulos de software 130 (representados en la figura 2B) pueden codificarse en el almacenamiento 290 y/o en la memoria 220. Los módulos de software 130 pueden comprender uno o más programas de software o aplicaciones (denominadas colectivamente como la "aplicación de servidor de autenticación segura") que tienen un código de programa de computadora o un conjunto de instrucciones ejecutadas en el procesador 210. Tal código o instrucciones de programa de computadora para llevar a cabo operaciones para aspectos de los sistemas y métodos descritos en este documento se pueden escribir en cualquier combinación de uno o más lenguajes de programación, como entenderán los expertos en la técnica. El código del programa puede ejecutarse completamente en el servidor del sistema 105 como un paquete de software independiente, en parte en el servidor del sistema 105 y en parte en un dispositivo informático remoto, como un dispositivo informático remoto 102, un dispositivo informático de usuario 101a y/o un dispositivo informático de usuario 101b, o completamente en dichos dispositivos informáticos remotos. '

[0055] También preferiblemente en el almacenamiento 290 se almacena una base de datos 280. Como se describirá con mayor detalle a continuación, la base de datos 280 contiene y/o mantiene varios ítems de datos y elementos que se utilizan a lo largo de las diversas operaciones del sistema 100, incluidos, entre otros, los perfiles de usuario como se describirá con mayor detalle en el presente documento. Cabe señalar que aunque la base de datos 280 se representa como configurada localmente para el dispositivo informático 105, en ciertas implementaciones la base de datos 280 y/o varios de los elementos de datos almacenados en ella pueden almacenarse en una memoria legible por computadora o en un medio de almacenamiento que sea ubicado de forma remota y conectado al servidor del sistema 105 a través de una red (no mostrada), de una manera conocida por los expertos en la técnica.

[0056] Una interfaz de comunicación 250 también está operativamente conectada al procesador 210. La interfaz de comunicación puede ser cualquier interfaz que permita la

comunicación entre el servidor del sistema 105 y dispositivos, máquinas y/o elementos externos. En ciertas implementaciones, la interfaz de comunicación puede incluir, entre otros, un módem, una tarjeta de interfaz de red (NIC), una interfaz de red integrada, un transmisor/receptor de radiofrecuencia (*ej.*, Bluetooth, celular, NFC), un transmisor/receptor de comunicación satelital, un puerto de infrarrojos, una conexión USB y/o cualquier otra interfaz para conectar el dispositivo informático 105 a otros dispositivos informáticos y/o redes de comunicación tales como redes privadas e Internet. Dichas conexiones pueden incluir una conexión por cable o una conexión inalámbrica (*ej.*, utilizando el estándar 802.11) aunque debe entenderse que la interfaz de comunicación 255 puede ser prácticamente cualquier interfaz que permita la comunicación hacia/desde el procesador 210.

[0057] El funcionamiento del sistema para autenticar a un usuario y los diversos elementos y componentes descritos anteriormente se apreciarán adicionalmente con referencia a las figuras 3-4B y con referencia continua a las figuras 1 y 2A-2B. Los procesos ilustrados en las figuras 3 y 4 se muestran desde la perspectiva del dispositivo informático del usuario 101a así como del servidor del sistema 105, sin embargo, debe entenderse que los procesos pueden ser realizados, total o parcialmente, por el dispositivo informático del usuario 101a, el servidor del sistema 105 y/u otros dispositivos informáticos (*ej.*, dispositivo informático remoto 102) o cualquier combinación de los mismos.

[0058] La Fig. 3 es un diagrama simple que ilustra una serie de nodos X1, X2 y X3 en una red neuronal, y además ilustra una función de activación como una suma de una matriz de pesos multiplicada por los valores X, y que produce un valor Y. Un objetivo es encontrar el número ideal de neuronas y capas en la red neuronal. Una fórmula para calcular el número de neuronas que "caben" en un volumen dado es proporcionada por $(W - K + 2 P) / S + 1$. Por ejemplo, la función de costo es euclidiana y el número de capas dado como el tamaño espacial del volumen de salida se puede calcular como una función del tamaño de volumen de entrada W, tamaño de filtro (kernel o núcleo) del campo (receptivo), un paso con el que las capas se aplican (deslizan) S, y una cantidad de relleno cero (padding) P usada en un borde.

[0059] Por ejemplo, la red neuronal se entrena inicialmente como un clasificador utilizando datos biométricos etiquetados. Como parte del proceso de entrenamiento, cada persona se almacena y las imágenes, del entrenamiento, están presentes. Después de que se introduce un vector biométrico inicial en la red neuronal, el vector utilizado en la capa n-1 se puede utilizar como un vector único de características para identificar ese vector biométrico

inicial, aunque en el espacio encriptado homomórfico. Este vector de características es euclidiano medible y cifrado. Además, el vector de características reemplaza al vector biométrico. En una o más implementaciones, el vector de características es una lista de 256 números de punto flotante, y el dato biométrico se reconstruye a partir de la lista de los 256 números de punto flotante. El vector de características es, en consecuencia, un cifrado unidireccional. Los vectores de características tienen la propiedad de ser euclidianos medibles.

[0060] Se reconoce que la presente solicitud proporciona aplicabilidad en una variedad de mercados verticales que pueden requerir una búsqueda encriptada para la identificación utilizando entradas biométricas. Por ejemplo, las compañías de seguros y los bancos desean mecanismos para identificar al titular de una cuenta que ha perdido su número de cuenta. La solución proporcionada de acuerdo con la presente solicitud puede ejecutarse en tiempo $O(\log(n))$, que se cree que mejora tras los algoritmos $O(n)$ existentes para la búsqueda cifrada en una base de datos de registros biométricos cifrados.

[0061] En una o más implementaciones, por ejemplo, se emplea una transformación matemática para proporcionar una partición para una búsqueda lineal. En tal caso, esta transformación toma un vector biométrico y devuelve un vector de características. El vector de características es utilizable, posteriormente, para el acceso directo a un nodo. El nodo accedido puede escanearse linealmente para determinar una identidad respectiva. Por ejemplo, un solo usuario puede identificarse en función del nodo al que se accede en función del vector de características.

[0062] En una o más implementaciones, se puede emplear una función matemática que particiona la información (ej., los vectores) asociada con biometría para respaldar una estructura en árbol tal como un árbol B+. Alternativamente, o además, la información asociada con biometría (ej., vectores biométricos) se puede situar mediante una red neuronal, que es operable para crear una biométrica no reversible para coincidencia o proximidad.

[0063] En relación con una implementación que emplea una red neuronal, un algoritmo de distancia euclidiana funciona para convertir un vector biométrico facial u otro valor en una característica que se pueda comparar. Un valor biométrico inicial (plantilla) se procesa a través de la red neuronal y da como resultado un vector de características. Dicha implementación es útil, por ejemplo, en relación con un caso de uso "uno a uno" o con un caso de uno a varios. En cualquier caso, la red neuronal y el vector de características se pueden usar para realizar la comparación en el espacio cifrado. Además y después de un caso

uno a uno, una red neuronal puede operar para realizar un costo Euclidiano para acceder a un dato biométrico específico o a uno o más grupos de datos biométricos, después de aplicar múltiples niveles de la red neuronal. El vector resultante, seguido de las multiplicaciones de la matriz para cada capa respectiva, puede usar un algoritmo euclidiano de distancia, con base en la función euclidiana de costo, y se conoce en la presente, generalmente, como un vector de características.

[0064] El sistema informático puede incluir clientes y servidores. Un cliente y un servidor generalmente remotos entre sí y que interactúan típicamente a través de una red de comunicación. La relación del cliente y el servidor surge en virtud de los programas informáticos que se ejecutan en las respectivas computadoras y que tienen una relación cliente-servidor entre sí.

[0065] Las FIGs. 4A y 4B muestran un ejemplo de una red neuronal de ejemplo en funcionamiento, de acuerdo con una implementación de ejemplo de la presente solicitud. La imagen inicial se aplica a la red neuronal. Se proporciona el softmax resultante, que muestra a cuál de las "cajas" definidas pertenece la imagen. En la caja respectiva mostrada, se proporcionan las imágenes para el individuo dado. Por ejemplo, el vector de características para la persona X, de 256 valores, cuando se aplica a la función de clasificación da como resultado un número de punto flotante de Y. Para todos los valores cercanos a Y, el individuo puede identificarse, siempre que el vector de características sea Euclidiano medible y la función de clasificación sea estable.

[0066] Al continuar con la referencia a las Figuras 4A y 4B, una función softmax resultante muestra una imagen de la persona en cuestión en el conjunto de entrenamiento. Posteriormente, se muestran los cálculos matriciales para crear el softmax. Dos pasos convolucionales antes del softmax es fc7-conv, que tiene la salida de la capa 7 convolucional. La salida es el vector de características.

[0067] La figura 5 ilustra un proceso de ejemplo de acuerdo con la red neuronal. Se proporciona una variedad de capas convolucionales, junto con un rectificador (ReLU) y nodos de agrupación. En el ejemplo que se muestra en la Fig. 5, el rectificador es una función de activación adecuada para las redes neuronales profundas, y es una función de activación definida como $f(x) = \max(0, x)$.

[0068] Más allá de la ReLU, la red neuronal de la presente solicitud se puede configurar para usar la agrupación como una forma de muestreo descendente no lineal. Más particularmente, un tipo de algoritmo de agrupación, denominado generalmente en la presente

como maxpooling, particiona las imágenes de entrada en un conjunto de rectángulos no superpuestos y, para cada subregión, arroja el máximo. Una vez se encuentra un rasgo, su ubicación exacta puede ser menos importante que su relación con otros rasgos. La función de la capa de agrupación reduce progresivamente el tamaño espacial de la representación para reducir la cantidad de parámetros y el cálculo en la red. Por lo tanto, mediante el uso de convoluciones, ReLU y maxpooling, se proporciona un vector de tamaño razonable de 256 para su uso en la concordancia, clasificación y comparación. Además, los vectores de características Euclidianas medibles (EMFV) cifrados y resultantes de la red neuronal se pueden utilizar posteriormente en un algoritmo de clasificación para recibir un valor escalar que se puede comparar con los vectores de características de otros datos biométricos. En una o más implementaciones, la presente solicitud utiliza modelos previamente entrenados, tales como modelos de aprendizaje profundo entrenados en datos asociados con imágenes de caras.

[0069] Al continuar haciendo referencia a la Fig. 5, el vector de características es un índice para su uso en el almacenamiento. Para la indexación, se emplea un algoritmo de clasificación que, dado un vector de características de entrada, devuelve un número de punto flotante, que se puede usar como índice para el almacenamiento. El algoritmo de clasificación depende de un vector medio de alta calidad, que ayuda a crear distancias entre las personas. El algoritmo de clasificación permite buscar y encontrar a una persona inscrita en tiempo polinómico.

[0070] El número de punto flotante mencionado anteriormente permite una búsqueda de índice para un dato biométrico almacenado previamente. Además, la presente solicitud emplea un algoritmo de clasificación que proporciona la persona prevista en función del vector de características de entrada.

[0071] En una o más implementaciones, se utiliza un algoritmo en el que se almacenan datos y matrices para proporcionar un modelo aprendido. En tal caso, un vector de cara promedio se compara con un IBV respectivo, y se calculan las diferencias entre el vector de cara promedio y el IBV. Por ejemplo, se puede utilizar un algoritmo de Frobenius en el que se calcula una distancia absoluta entre cada punto flotante y su promedio. Además, la diferencia de posición relativa entre el vector de cara promedio y el IBV respectivo se calcula y, luego, se eleva al cuadrado. Todos los valores se suman posteriormente y se calcula la raíz cuadrada. Esto da como resultado valores relativamente cercanos, es decir, todas las distancias calculadas de Frobenius son grupos relativamente cercanos.

[0072] En una o más implementaciones, durante la inscripción se obtienen varios vectores faciales (0-12). Los vectores se procesan a través de la red neuronal y si alguno de los vectores de características procesados tiene un valor de 0, posteriormente el valor se degrada (cerca de 0 o tiene 0), puesto que el valor no es significativo porque no está presente en cada vector de característica.

[0073] Cada vez que se usa una imagen, en funcionamiento, se degrada, lo que resulta en una mayor separación dentro de los vectores de característica. Dado un vector de características de 256 enteros no negativos, el procedimiento descrito en este documento clasifica al vector en un grupo apropiado, y cada grupo es para una y solo una persona distinta. Si la persona no se encuentra, se asume que el algoritmo buscará personas próximas para completar la búsqueda.

[0074] Por lo tanto, la presente solicitud utiliza la distancia euclidiana para clasificar vectores no conocidos por la forma $x = (x_0, x_1, \dots, x_{255})$. A menos que se indique lo contrario, x_i es un número entero no negativo para todos $i \in \{0, 1, \dots, 255\}$. La distancia entre dos vectores se define como $d(x, y) = (\sum_{i=0}^{255} |x_i - y_i|^2)^{1/2}$. Dado un vector x , sea $\hat{x} = x/d(x, (0, 0, \dots, 0))$.

[0075] En funcionamiento, U se define como el conjunto de vectores conocidos y $|U|$ el número de vectores en U . $|U|$ is assumed ∞ . Posteriormente, la media de todos los vectores en U está dada por $m_U = 1/|U| \sum_{x \in U} x$ y puede calcularse explícitamente si $|U|$ es suficientemente pequeño. Si $|U|$ no es lo suficientemente pequeño, m_U puede ser aproximado. Tenga en cuenta que las coordenadas de m_U no son necesariamente enteros no negativos.

[0076] Posteriormente, considere una partición de U como $U = \coprod_j P_j$. Se ha observado experimentalmente que para cada P_j , existe $a_j, b_j \in (0, \infty)$ del mismo que $d(\widehat{m}_U, \hat{y}) \in [a_j, b_j]$ para todo $y \in P_j$. Por otra parte, $[a_j, b_j]$ es disjunto de $[a_k, b_k]$ para $j \neq k$. En otras palabras, se pueden usar bandas de distancias del vector medio para clasificar vectores desconocidos ..

[0077] Dado un vector $y \notin U$, $d(\widehat{m}_U, \hat{y})$ se calcula para determinar cómo extender de forma única particiones de U a particiones de $V = U \cup \{y\}$. Si $d(\widehat{m}_U, \hat{y}) \notin [a_j, b_j]$ para todo j , se elige el intervalo $[a_j, b_j]$ más cercano a y y se elige el subconjunto P_j asociado con este intervalo para incluir a y en la extensión de la partición original a una partición de V . Si sucede que y es equidistante a dos intervalos diferentes, , por lo tanto el subconjunto en el que se incluye y en la partición de V no está bien definido. En este caso, los resultados

numéricos deben ser reexaminados y debe hacerse una mejor elección para al menos uno de los dos intervalos equidistantes a y .

[0078] En una operación de ejemplo, se usan 257 imágenes para el entrenamiento. En la operación de entrenamiento de ejemplo, la red neuronal es una red de 32 nodos y 8 capas de convolución. El entrenamiento asigna los pesos apropiados a la red. Con esta red entrenada, se ingresó una nueva cara para el almacenamiento o la búsqueda. En cualquier caso, la cara se sometió a la red neuronal y en la capa convolucional 7, se recibió un vector de características.

[0079] Mediante el uso de un conjunto de imágenes y la aplicación de la red neuronal, se recibió el siguiente EMFV en la capa 7 convolucional. A continuación, se muestra un ejemplo de tal salida:

```
3 0 0 0 7 12 4 0 2 2 0 0 0 2 0 2 0 0 1 0 1 3 4 7 0 4 1 4 8
2 0 2 6 0 8 7 5 0 0 0 4 7 0 3 2 0 1 2 3 4 0 0 0 0 0 0 2 0
0 2 1 6 3 0 2 1 1 0 0 4 0 8 3 5 0 5 3 1 4 0 9 0 3 1 6 1 0
2 8 1 0 2 0 2 2 3 1 0 4 2 0 2 5 2 1 0 2 2 0 0 1 0 1 3 2 1
1 0 1 9 5 5 3 0 1 0 4 0 0 4 2 1 5 4 1 0 5 1 3 1 5 8 1 7 4
7 0 0 2 4 1 4 4 0 3 0 0 4 0 9 4 0 0 3 1 1 4 0 5 0 1 2 1 6
7 0 5 5 0 4 6 0 2 3 7 0 0 4 1 3 0 3 0 0 6 1 0 1 0 0 1 0 0
7 3 0 5 0 0 0 3 0 1 0 2 3 0 0 1 1 1 0 6 0 1 8 1 0 1 1 0 0 0
0 3 9 2 0 7 2 5 0 0 0 1 0 0 3 5 3 0 0 0 5 3 4 2
```

[0080] El vector de características clasifica a una persona específica dentro del tiempo de ejecución $O(1)$. Al utilizar el algoritmo descrito en la presente, se determinaron los rangos de tres individuos. Al utilizar los vectores normalizados, los rangos de imágenes resultantes son los siguientes. Para la Persona 1, las distancias normalizadas oscilan entre 0,85 y 1,12. Para la Persona 2, el rango es de 1.18 a 1.32. Para la Persona 3, las distancias normalizadas varían entre 0,39 a 0,68.

[0081] Como un ejemplo práctico, los IBV posteriores de la Persona 1, cuando se administran a la red neuronal y al algoritmo de clasificación, producen resultados entre 0,85 y 1,12. Del mismo modo, los IBV posteriores para la Persona 3 proporcionan resultados entre 0,39 y 0,68. Estos rangos aceptables ofrecen bandas para cada persona respectiva. Por lo tanto, la idea es que se observe formación bandas y no se observen colisiones para las bandas entre personas. Para pequeños conjuntos de datos, los resultados son exactos.

[0082] Por lo tanto, la presente solicitud proporciona tecnología para: (1) adquirir datos biométricos, (2) comparación de coincidencia biométrica no cifrada, (3) cifrar los datos biométricos, (4) realizar un cotejo Euclidiano medible y (5) realizar una búsqueda mediante el uso de un esquema de indexación de uno a varios. Estos se proporcionan de una manera basada en tiempo polinomial y protección de la privacidad y basada en el tiempo polinomial, que también es de biométrica agnóstica y se realiza en función del aprendizaje automático. La presente solicitud proporciona una solución de propósito general que produce información biométrica cifrada que es Euclidiano medible, incluyendo en función de la red neuronal convolucional. Esto se proporciona además en función de un algoritmo de clasificación que se emplea para la identificación de uno a varios, lo que maximiza la privacidad y se ejecuta entre el tiempo $O(1)$ y $O(\log(n))$.

[0083] Se recibe un vector biométrico inicial (IBV), y uno o más módulos aplican una red neuronal al IBV. El IBV se procesa, ya sea en el dispositivo informático del usuario o en un servidor, a través de un conjunto de operaciones matriciales, para crear un vector de características. Se puede almacenar tanto una parte como el vector de características completo en el cliente o servidor. Después de la multiplicación de la matriz en varias capas, el IBV se devuelve como un vector Euclidiano-medible. En una o más implementaciones, la misma acción puede ocurrir cuando el Vector biométrico actual (CBV) y los dos vectores coinciden.

[0001] La figura 6 es un diagrama de flujo que ilustra los pasos de proceso de ejemplo 600 de acuerdo con una implementación. Debe apreciarse que varias de las operaciones lógicas descritas en este documento se implementan (1) como una secuencia de actos implementados por computadora o módulos de programa que se ejecutan en un dispositivo de comunicación y/o (2) como circuitos lógicos de máquina interconectados o módulos de circuito dentro de un dispositivo de comunicación. La implementación depende de los requisitos del dispositivo (*ej.*, tamaño, energía, consumo, rendimiento, etc.). En consecuencia, las operaciones lógico descrito en el presente documento se denomina de manera diversa como funcionamiento, dispositivos estructurales, actos o módulos. Varias de estas operaciones, dispositivos estructurales, actos y módulos pueden implementarse en software, en firmware, en lógica digital para propósito especial y cualquier combinación de los mismos. También se aprecia que se pueden realizar más o menos operaciones que las mostradas en las figuras y las descritas en este documento. Estas operaciones también se pueden realizar en un orden diferente al descrito en este documento. Los pasos de ejemplo

600 ilustrados en la Fig. 6 proporcionan un método implementado por computadora para cotejar un registro biométrico de entrada cifrado con al menos un registro biométrico cifrado almacenado, y sin descifrado de datos de entrada y un registro almacenado como mínimo.

[0002] El proceso inicia en el paso 602, y se proporciona un vector biométrico inicial

5 a una red neuronal, y la red neuronal traduce el vector biométrico inicial a un vector de características Euclidianas-medibles (paso 604). El vector de características Euclidianas-medible se guarda en un almacenamiento con otros vectores de características medibles de manera euclidiana (paso 606). Además, un vector biométrico actual que representa el registro biométrico de entrada cifrado se recibe desde un dispositivo informático móvil a través de
10 una red de comunicación de datos, y el vector biométrico actual se suministra a la red neuronal (paso 608). La red neuronal traduce el vector biométrico actual a un vector actual de características Euclidianas-medible (paso 610). Además, se realiza una búsqueda de al menos algunos de los vectores de características Euclidianas-medibles almacenados en una parte del almacenamiento de datos utilizando el vector actual de características euclidianas-medibles (paso 614). El registro biométrico de entrada cifrado se coteja con al menos un
15 registro biométrico cifrado en el espacio de cifrado en función de una distancia absoluta calculada entre el vector actual de características Euclidianas-medible y un cálculo de cada uno de los vectores respectivos de características medibles de manera euclidiana en la parte del almacenamiento (paso 616).

20 **[0084]** El tema descrito anteriormente se proporciona solo a modo de ilustración y no debe interpretarse como limitante. Se pueden realizar diversas modificaciones y cambios al tema aquí descrito sin seguir las modalidades y aplicaciones de ejemplo ilustradas y descritas, y sin apartarse del verdadero espíritu y alcance de la presente invención, incluido lo establecido en cada una de las siguientes reivindicaciones.

25

REIVINDICACIONES

1. Un método implementado por computadora para cotejar un registro biométrico de entrada cifrado con al menos un registro biométrico cifrado almacenado, sin descifrado de
5 datos de entrada y un registro almacenado como mínimo, el método comprende:

proporcionar un vector biométrico inicial a una red neuronal, en donde la red neuronal traduce el vector biométrico inicial a un vector de características euclidiano-medibles,

almacenar el vector de características euclidiano-medibles en un almacenamiento con otros vectores de características euclidiano-medibles;

10 recibir, desde un dispositivo de computación móvil a través de una red de comunicación de datos, un vector biométrico actual que representa el registro cifrado de entrada biométrica;

proporcionar el vector biométrico actual a una red neuronal, en donde la red neuronal traduce el vector biométrico actual a un vector de características euclidiano-medibles:

15 realizar una búsqueda de al menos algunos de los vectores de características euclidiano-medibles almacenados en una parte del almacenamiento de datos utilizando el vector actual de características euclidiano-medibles,

en donde el registro biométrico de entrada cifrado se coteja con al menos un registro biométrico cifrado en el espacio de cifrado en función de una distancia absoluta calculada
20 entre el vector actual de características euclidiano-medibles y un cálculo de cada uno de los vectores respectivos de características euclidiano-medibles en la parte del almacenamiento.

2. El método de la reivindicación 1, que comprende, además:

clasificar el vector de características euclidiano-medibles; y/o

25 clasificar el vector de características euclidiano-medibles actual,

en donde la clasificación se realiza al menos en parte mediante el uso de una o más funciones de distancia.

3. El método de la reivindicación 2, en donde la clasificación de la característica
30 euclidiana medible y/o el vector actual de características euclidiano-medibles devuelve los valores de punto flotante, y se utiliza un algoritmo de Frobenius para calcular una distancia absoluta entre cada punto flotante y su promedio.

4. El método de la reivindicación 2, en donde la búsqueda se realiza con complejidad temporal $\log(n)$.

5. El método de la reivindicación 2, que comprende, además:
utilizar un algoritmo de Frobenius para clasificar los vectores biométricos euclidianomedibles;
atravesar una jerarquía de los vectores biométricos euclidianomedibles clasificados con complejidad temporal $\log(n)$; y
identificar que un vector biométrico euclidianomedible respectivo es el vector de características euclidianomedible actual.

6. El método de la reivindicación 1, que comprende, además:
identificar, por cada vector biométrico euclidianomedible respectivo, una pluralidad de valores de punto flotante; y
utilizar un mapa de bits para eliminar a partir de un cálculo de distancias absolutas cualquiera de la pluralidad de valores que no están presentes en cada vector.

7. El método de la reivindicación 1, que comprende, además:
identificar, por cada vector biométrico euclidianomedible respectivo, una pluralidad de valores de punto flotante; y
definir una escala deslizante de importancia basada en el número de vectores en los que aparece un valor respectivo del valor de punto flotante.

8. El método de la reivindicación 1, en donde la red neuronal se configura con una variedad de capas convolucionales, junto con un rectificador (ReLU) y nodos de agrupación.

9. El método de la reivindicación 1, en donde la red neuronal está configurada para usar la agrupación como una forma de muestreo descendente no lineal,
y además en donde uno o más nodos de agrupación reducen progresivamente el tamaño espacial de un vector de características euclidianomedibles representado para reducir la cantidad de parámetros y cálculos en la red neuronal.

10. El método de la reivindicación 8, que comprende, además:

calcular, para cada uno de una pluralidad de vectores de características euclidianomedibles almacenados, una diferencia de posición relativa entre un vector de cara media y el vector de características euclidianomedibles respectivas;

eleva al cuadrado la diferencia de posición relativa;

5 sumar los valores; y

calcular la raíz cuadrada.

11. El método de la reivindicación 1, en donde el rendimiento de la red neuronal se determina con base en una función de costo, en la que se calculan varias capas dadas como un tamaño espacial de un volumen de salida como una función de un tamaño de volumen de entrada W, tamaño de filtro (kernel o núcleo) del campo (receptivo), un paso con el que las capas se aplican (deslizan) S, y una cantidad de relleno cero (padding) P usada en un borde.

12. El método de la reivindicación 1, en donde la red neuronal traduce el vector biométrico inicial y el vector biométrico actual en función de las multiplicaciones de matriz para cada capa respectiva, y usa un algoritmo de distancia euclidiana basado en una función de costo euclidiana.

13. Un sistema implementado por computadora para cotejar un registro biométrico de entrada cifrado con al menos un registro biométrico cifrado almacenado, sin descifrado de datos de entrada y un registro almacenado como mínimo, el sistema comprende:

uno o más procesadores y un medio legible por computadora, en donde uno o más procesadores están configurados para interactuar con el medio legible por computadora para realizar operaciones que incluyen:

25 proporcionar un vector biométrico inicial a una red neuronal, en donde la red neuronal traduce el vector biométrico inicial a un vector de características euclidianomedibles.

almacenar el vector de características euclidianomedibles en un almacenamiento con otros vectores de características euclidianomedibles;

30 recibir, desde un dispositivo de computación móvil a través de una red de comunicación de datos, un vector biométrico actual que representa el registro cifrado de entrada biométrica;

proporcionar el vector biométrico actual a una red neuronal, en donde la red neuronal traduce el vector biométrico actual a un vector de características euclidianomedibles;

realizar una búsqueda de al menos algunos de los vectores de características euclidianas-medibles almacenados en una parte del almacenamiento de datos utilizando el vector actual de características euclidianas-medibles,

- 5 en donde el registro biométrico de entrada cifrado se coteja con al menos un registro biométrico cifrado en el espacio de cifrado en función de una distancia absoluta calculada entre el vector actual de características euclidianas-medibles y un cálculo de cada uno de los vectores respectivos de características euclidianas-medibles en la parte del almacenamiento.

- 10 14. El sistema de la reivindicación 13, en donde uno o más procesadores están configurados adicionalmente para interactuar con el medio legible por computadora para realizar operaciones que incluyen:

 clasificar el vector de características euclidianas-medibles; y/o

 clasificar el vector de características euclidianas-medibles actual,

- 15 en donde la clasificación se realiza al menos en parte mediante el uso de una o más funciones de distancia.

- 20 15. El sistema de la reivindicación 14, en donde la clasificación de la característica euclidiana medible y/o el vector actual de características euclidianas-medibles devuelve los valores de punto flotante, y se utiliza un algoritmo de Frobenius para calcular una distancia absoluta entre cada punto flotante y su promedio.

16. El sistema de la reivindicación 14, en donde la búsqueda se realiza en complejidad temporal $\log(n)$.

- 25 17. El sistema de la reivindicación 14, en donde uno o más procesadores están configurados para interactuar con el medio legible por computadora para realizar operaciones que incluyen:

 utilizar un algoritmo de Frobenius para clasificar los vectores biométricos euclidianos-medibles;

- 30 atravesar una jerarquía de los vectores biométricos euclidianos-medibles clasificados con complejidad temporal $\log(n)$; y

 identificar que un vector biométrico euclidiano-medible respectivo es el vector de características euclidianas-medibles actual.

18. El sistema de la reivindicación 13, en donde uno o más procesadores están configurados para interactuar con el medio legible por computadora para realizar operaciones que incluyen:

- 5 identificar, por cada vector biométrico euclidiano-medible respectivo, una pluralidad de valores de punto flotante; y
- utilizar un mapa de bits para eliminar a partir de un cálculo de distancias absolutas cualquiera de la pluralidad de valores que no están presentes en cada vector.

10 19. El sistema de la reivindicación 13, en donde uno o más procesadores están configurados para interactuar con el medio legible por computadora para realizar operaciones que incluyen:

- identificar, por cada vector biométrico euclidiano-medible respectivo, una pluralidad de valores de punto flotante; y
- 15 definir una escala deslizante de importancia basada en el número de vectores en los que aparece un valor respectivo del valor de punto flotante.

20 20. El sistema de la reivindicación 13, en donde la red neuronal se configura con una variedad de capas convolucionales, junto con un rectificador (ReLU) y nodos de agrupación.

21. El sistema de la reivindicación 13, en donde la red neuronal está configurada para usar la agrupación como una forma de muestreo descendente no lineal,

 y además en donde uno o más nodos de agrupación reducen progresivamente el

25 tamaño espacial de un vector de características euclidiano-medibles representado para reducir la cantidad de parámetros y cálculos en la red neuronal.

22. El sistema de la reivindicación 20, en donde uno o más procesadores están configurados para interactuar con el medio legible por computadora para realizar operaciones que incluyen:

30

 calcular, para cada uno de una pluralidad de vectores de características euclidiano-medibles almacenados, una diferencia de posición relativa entre un vector de cara media y el vector de características euclidiano-medibles respectivas;

elevar al cuadrado la diferencia de posición relativa;
sumar los valores; y
calcular la raíz cuadrada.

5 23. El sistema de la reivindicación 13, en donde el rendimiento de la red neuronal se determina con base en una función de costo, en la que se calculan varias capas dadas como un tamaño espacial de un volumen de salida en base de un tamaño de volumen de entrada W , tamaño de filtro (kernel o núcleo) del campo (receptivo), un paso con el que las capas se aplican (deslizan) S , y una cantidad de relleno cero (padding) P usada en un borde.

10

 24. El sistema de la reivindicación 13, en donde la red neuronal traduce el vector biométrico inicial y el vector biométrico actual en función de las multiplicaciones de la matriz para cada capa respectiva, y utiliza un algoritmo de distancia euclidiana basado en una función de costo euclidiana.

15

RESUMEN

Sistemas y métodos implementados por computadora para hacer coincidir un registro biométrico de entrada cifrado con al menos un registro biométrico cifrado

- 5 almacenado, y sin descifrado de datos de la entrada y el al menos un registro almacenado.