

Tokenización en Perfil Humano Digital para la Identificación Humana

Tokenization in Digital Human Profile for Human Identification

J. Sergio Ruiz Castilla¹, Dalia B. Luna Coria², Farid García Lamont³.

Resumen: La identidad incluye nombre, edad, género, etc. Sin embargo, dichos datos son datos confidenciales y requieren protección. Es necesario identificar a personas en procesos judiciales o personas desaparecidas. Una persona tiene múltiples características físicas como color de ojos, piel y cabello. Además, tipo de: la frente, la nariz, la boca o los labios. Para este trabajo, se han se han determinado 256 características que incluyen la dentadura, la ropa, el calzado y accesorios. Para cada característica se crean categorías. Luego, se genera un vector de tokens. Al vector de tokens de características se le ha denominado Perfil Humano Digital representado como una matriz de valores enteros, hasta llegar a un patrón tipo QR. Se pretende que, dichos Perfiles se almacenen en un repositorio para proyectos futuros como identificación de personas ante procesos judiciales, búsqueda de personas desaparecidas e identificación de LGBTs (Lesbianas, gays, bisexuales y Trans).

¹ Doctor en Ingeniería de Software, Universidad Popular Autónoma del Estado de Puebla, México, Puebla. Afiliación institucional: Universidad Autónoma del Estado de México, México. Correo electrónico personal e institucional e-mail: jsrgioruizc@gmail.com, jsruizc@uaemex.mx. ORCID: <https://orcid.org/0000-0001-7821-4912>

² Maestra en Ciencias de la Computación, Universidad Autónoma del Estado de México, México, Texcoco. Afiliación institucional: Universidad Autónoma del Estado de México, México. Correo electrónico personal e institucional e-mail: dalialuna928@gmail.com, dlunac136@alumno.uaemex.mx. ORCID: <https://orcid.org/0000-0003-3946-1499>

³ Doctor en Ciencias, Cinvestav, México, Ciudad de México. Afiliación institucional: Universidad Autónoma del estado de México, México. Correo electrónico personal e institucional e-mail: fgarcial@uaemex.mx ORCID: <https://orcid.org/0000-0002-9739-3802>

Palabras clave: Perfil digital humano, Características humanas, Identificación humana, Tokenización.

Abstract: Identity includes name, age, gender, etc. However, such data is considered sensitive and requires protection. It is necessary to identify people in judicial proceedings or missing persons. A person has multiple physical characteristics such as eye, skin, and hair color. In addition, the type of forehead, nose, mouth, or lips. For this work, 256 characteristics have been determined, including teeth, clothing, footwear, and accessories. Categories are created for each characteristic. A vector of tokens is then generated. The vector of characteristic tokens has been called a Digital Human Profile, represented as a matrix of integer values, until reaching a QR-type pattern. The intention is for these profiles to be stored in a repository for future projects such as identifying people in judicial proceedings, searching for missing persons, and identifying LGBTs (Lesbian, Gay, Bisexual, and Transgender).

Keywords: Human digital profile, Human characteristics, Human identification, Tokenization.

1. Introducción

Cada persona es diferente gracias a sus características físicas. Algunas características permanecen desde que la persona nace, mientras que, otras cambian a lo largo de la vida. Sin embargo, la mayoría de las características permanecen determinando la identidad. El ADN permite una identificación única, desafortunadamente, el ADN no es barato ni fácil para su obtención, almacenamiento y uso.

La identificación humana resulta muy importante en los siguientes casos: El primero, cuando una persona tiene homónimos, es decir comparte el nombre y origen con alguien más y es detenido por la policía. Segundo, cuando dos personas comparten la mayoría de las características físicas. Sin embargo, las características como la edad, estatura, peso, color de piel, cicatrices, lunares, tatuajes permiten determinar la identidad correcta.

En este caso, cuando una persona está desaparecida sus familiares inician su búsqueda. Para la búsqueda, se requieren sus características físicas como: género, edad, estatura, color de piel, lunares, tatuajes, etc. Las características ayudan a identificar a la persona y contactar a sus familiares. Otro caso especial, es cuando la persona es LGBT y decide hacer cambios físicos en su cuerpo y ahora su imagen ha cambiado. A pesar de los cambios, es posible identificar a la persona.

Generalmente, la identidad de una persona se basa en sus datos personales tales como el nombre, género, edad, domicilio, CURP, teléfono, correo electrónico. Sin embargo, dichos datos son confidenciales y requieren privacidad. Otros elementos pueden ser: el tono de voz, la forma de caminar o la forma de hablar. Finalmente, otros elementos son sus preferencias por: los colores, la música, la comida, etc. Esta última información se usa en plataformas digitales para crear perfiles de usuarios y mejorar su experiencia o bien para generar mercadotecnia digital enfocada.

La identificación de productos comerciales se inició con una serie de números, después se creó el código de barras. Actualmente, puede funcionar también con un código QR. Por lo que, dichos códigos como imágenes permiten identificar a un producto. El objetivo de este trabajo es buscar crear un patrón para representar la identidad de una persona. Para generar dichos patrones se busca usar sus características físicas.

Para este trabajo, se consideraron 256 características físicas, de las cuales podrían cambiar a lo largo de la vida, sin embargo, la persona es la misma. Observar que, si cambian cinco características, la persona sigue conservando el 98% de sus características matemáticamente.

Como producto final de este trabajo es un patrón que contiene las características de la persona. Por otro lado, la forma de cómo se puede salvaguardar y actualizar dicho patrón, si fuera necesario. Finalmente establecer los usos de los patrones.

1.1. Estado del arte

En la búsqueda de trabajos publicados, se encontraron múltiples formas para la Identificación humana. Desde la identificación dental [1] usando la carta o dentagrama que generan los dentistas [2] e incluso enfocándose en las caries [3]. Por otro lado, se han usado la biometría facial [4], así como la forma de la cara [5]. También se han usado los labios y la huella que dejan en objetos [6][7]. En otro trabajo, usaron las orejas [8][9], e incluso, usando las huellas de las plantas de los pies [10]. Pero no solo se han enfocado en la cara, los tatuajes han sido claves para la identificación de personas [11].

La fuente más recurrente para recabar las características físicas de una persona es a través de entrevistas a familiares [12]

En casos de identificar personas sin vida se ha recurrido a técnicas forenses. El médico forense genera un perfil de una persona que incluye características físicas y médicas [13].

Se resumen, los trabajos más significativos se muestran en la Tabla 1 para situar mejor el estado del arte que guarda la identificación humana basada en características físicas.

Autores	Año	Objetivo	Técnicas	Resultados
---------	-----	----------	----------	------------

Malpartida C. Annushka [1]	2019	Identificar a personas a través del registro del Dentagrama del dentista	Registro del Dentagrama	Similitudes erróneas por usar diferentes estándares usados por los Dentistas
Vázquez Díaz, Danny J. [2]	2022	Importancia y función de la odontología forense	Radiografías panorámicas y revisión física	Registro de estado de cada pieza dental
Hernández F., Carlos U., Huerta-Pacheco N. Sofía, Gil-Chavarria Ivett [3]		Usar la existencia de caries en las piezas dentales de la persona	Placas dentales	Las caries son más rampantes cuando se consumen sustancias psicoactivas, mala higiene y alimentación
Pedroza M. Rafael E. [4]	2019	Sistema basado en biometría facial para localizar a personas desaparecidas.	Comparación de rostros de personas desaparecidas	Funciona hasta el 92% aunque disminuye cuando hay accesorios en el rostro
Lescano M. Esther C. [5]	2021	Identificar fotografías por la forma de la cara a partir de múltiples fotografías	Clasificar las formas de las partes de la cara, como la nariz, boca, frentes, etc.	81% de precisión
Briem Stamm A.D., Velázquez G. H. [6]	2021	Crear una base de datos de huellas labiales para uso posterior de identificación	Queiloscopía. Huella de los labios	Propuesta de incluir las huellas labiales como parte biométrica
Díaz G. y otros [7]	2019	Proponer la huella labial como medio biométrico de 1020 imágenes	Patrón de 9 puntos característicos	95% de precisión
Ceballos A. A. G. y Rodes L. F. [8][9]	2024	Obtener la forma de la oreja y registrar en un Ortograma	Anatomía de la oreja	Comparar imágenes antes y después de una cirugía
Montoya-Flores Angie J. [10]	2024	Dividir en secciones las huellas de los pies y determinar características	Pelmatoscopia. Huella de los pies	Patrón de la huella de los pies
Melo S. G. y otros [11]	2022	Registrar tatuajes durante el proceso forense	Identificar tatuajes para identificar a personas	Clasificar tipos e tatuajes

Tabla 1 Resumen de trabajos del Estado del arte.

2. Metodología propuesta

Para lograr el objetivo de este trabajo se propone el método CTV (Caracterización-Tokenización-Vectorización). El método CTV permite obtener el PHD (Perfil Humano Digital) de una persona. El método CTV consiste en definir las 256 características para cada persona. Después, Generar un token a partir de cada categoría con un número entero. Luego,

generar un vector de tokens que puede manipularse como una matriz de 16 por 16 elementos. Con dicha matriz se busca una mejor forma de presentar el patrón del PHD. Ver la Figura 1.

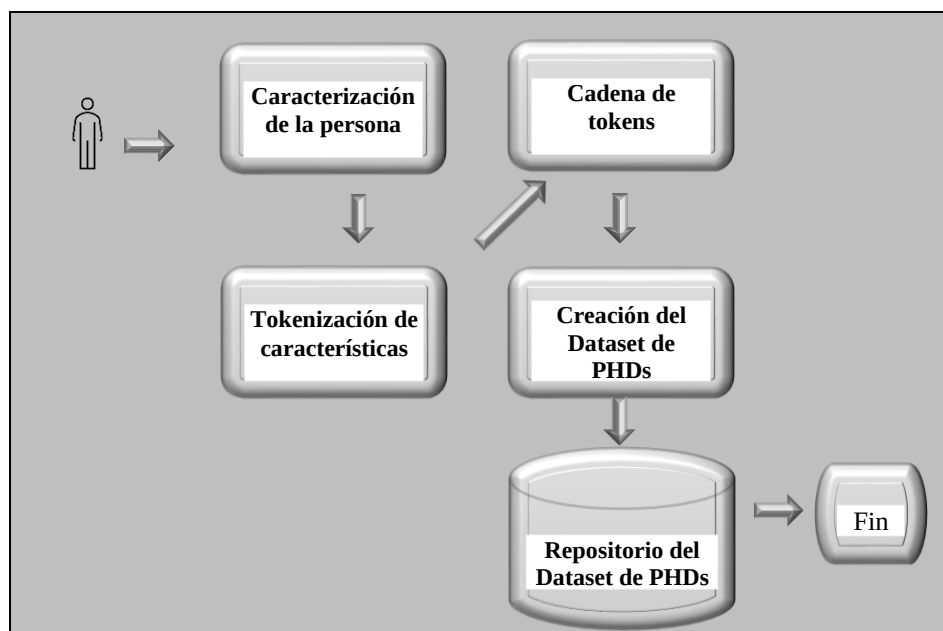


Figura 1 Método CTV (Caracterización-Tokenización-Vectorización) propuesto para obtener el PHD

3.1 Caracterización de una persona

La caracterización (CAR) consiste en elegir una categoría de una característica física. La caracterización está basada en características físicas como: la dentadura, los tatuajes, las cicatrices, los lunares, la ropa y los accesorios entre otros. Para obtener las características se obtienen de las siguientes fuentes: La primera, de un familiar cercano que proporcione las características. La segunda, a partir de una o más fotografías. La tercera, de un reporte forense. Y la cuarta, del expediente clínico tal como placas de la dentadura (el Dentagrama, expediente clínico generado por el Dentista).

3.2 Tokenización de características

Para la tokenización (TOK) basta contar con las diferentes posibilidades de cada característica física. Por ejemplo, si queremos registrar el género de la persona y el color de los ojos basta elegir un elemento de la galería, como se muestra en las Figuras 2 y 3.

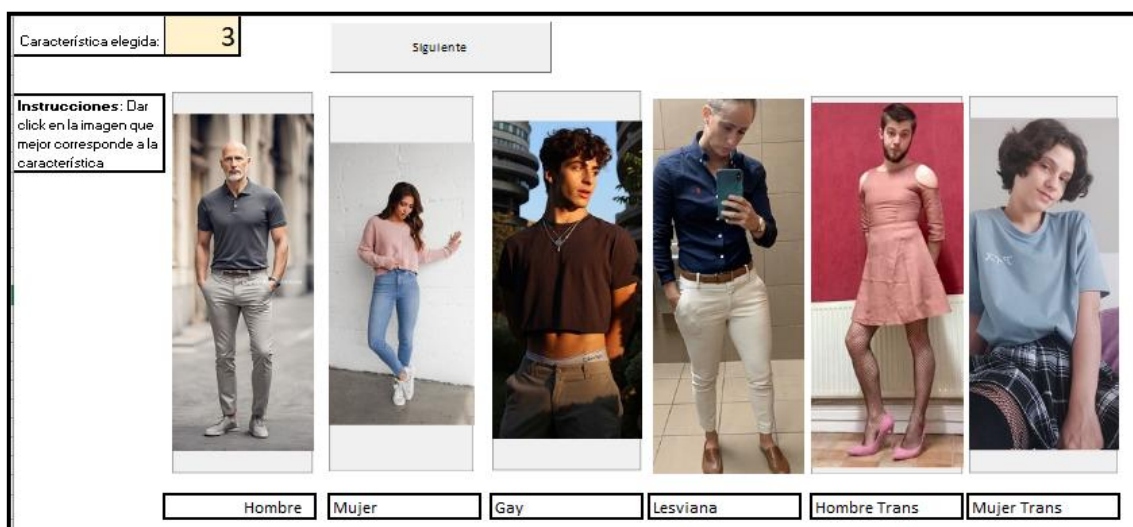


Figura 2 Galería de imágenes para elegir el género.

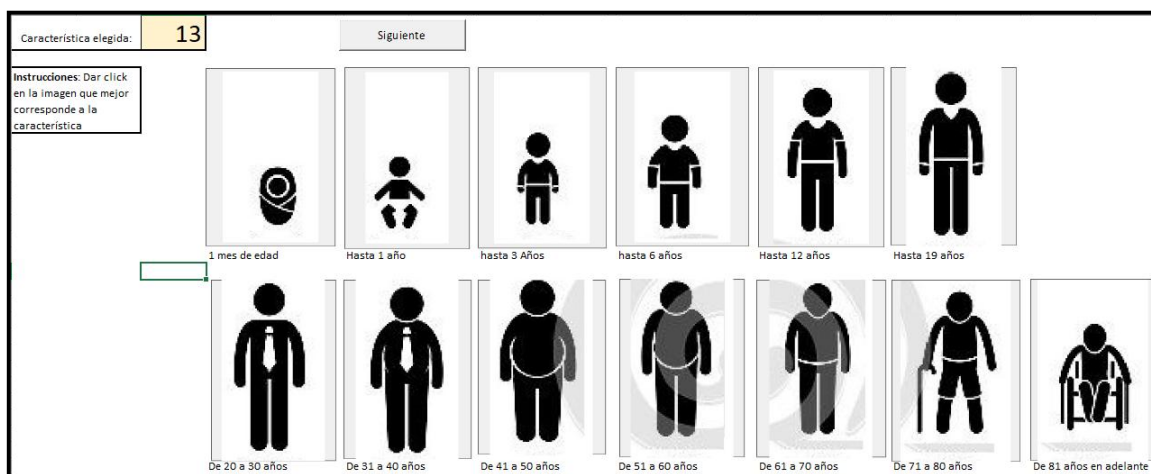


Figura 3 Galería de imágenes para elegir la edad.

El objetivo es que el usuario no capture datos, ni números. Se considera un entorno completamente grafico para la caracterización. El usuario, únicamente elegirá con un click o con un touch la imagen de la galería que corresponda. Considerando que, la aplicación podría estar en el celular, tableta o computadora.

Si contamos con hasta 16 categorías se deberá elegir una categoría. La categoría será un número entre 0 y 15. Se registra el número y se añadirá un nuevo token. En todos los casos se buscará usar imágenes. Las imágenes ayudan a que la caracterización sea más rápida y simple.

Para determinar la ubicación de un tatuaje, cicatriz o lunar se han usado las regiones anatómicas publicadas por el Dr. Banda. Considerando que un tatuaje, cicatriz o lunar puede abarcar más de una región. Ver la Figura 4 que muestra un cuerpo con las regiones anatómicas.

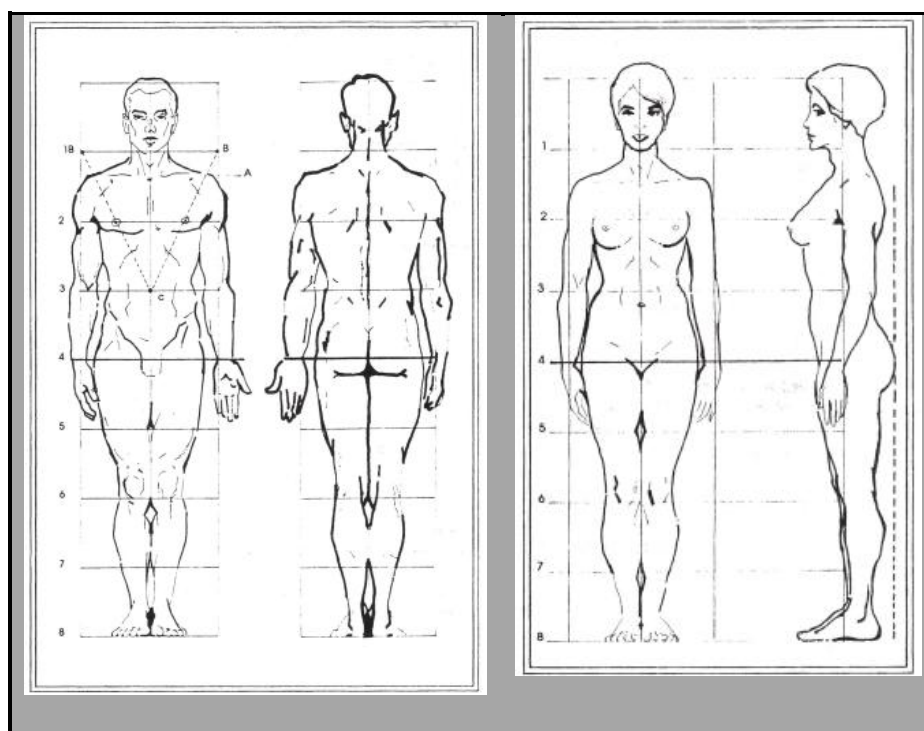


Figura 4 Regiones anatómicas del cuerpo humano.[16]

Las partes del cuerpo se usarán para determinar la existencia de cicatrices, lunas y tatuajes. Cada parte del cuerpo se trata por separado. Lo anterior obedece a que podría haber una cicatriz, un lunar e incluso un tatuaje en la misma área. Para este caso, se presenta la parte del cuerpo y la existencia de la característica y la respuesta solo es Si o No (Sin característica o Con característica).

3.3 Vectorización de tokens

Número de característica	001	002	003	004	005	006	007	008	009	...	256
Tokens	09	06	07	00	06	00	05	00	00	...	01

3.4 Mapa de identificación humana

[illegible]

00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00
00	00	00	00	01	00	00	00	00	00	00	00	00	00	00	00	00
00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00
00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00

Tabla 3 Matriz de 16 X 16 valores de Tokens.

3.5 Presentación y uso del PHD

El PHD es el vector de Tokens. Sin embargo, es posible almacenar y mostrar el PHD de una forma más prácticas para los usuarios. En la Figura 5 se muestran dos formas posibles de mostrar el PHD. El vector de categorías se puede mostrar como una matriz de 16 por 16 elementos. A partir de dicha matriz es posible presentar el PHD como un patrón QR. Cada PHD tendrá asociado un código de barras de los metadatos. Los metadatos pueden corresponder a la fecha de registro, estatus de la persona, ubicación, etc.

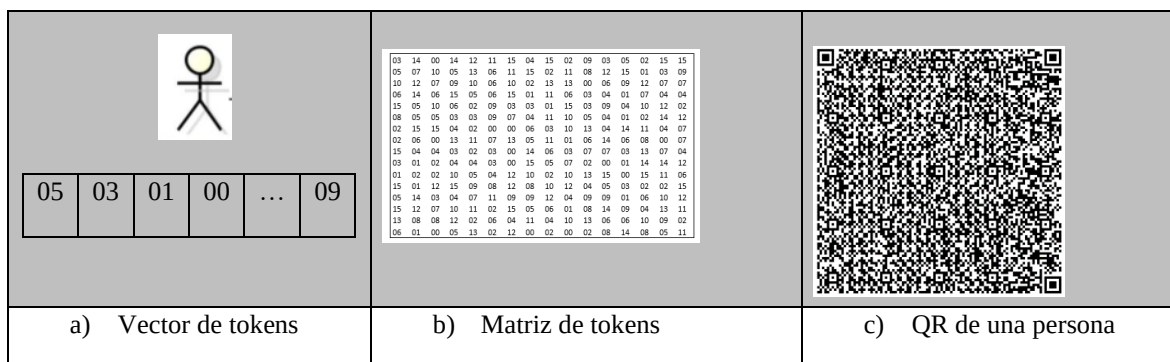


Figura 5 Proceso para convertir características a un QR.

3.6 Repositorio de PHDs

Los Vectores del PHD (VPHD) deben almacenarse para usos posteriores. Se propone que se disponga en un Sitio web con almacenamiento en la nube. Cada VPHD debe ser generado con la misma TCAR. Por lo tanto, el VCC permite obtener y almacenar los VPHD y el VCC en forma inversa (Vectorización-Categorización-Characterización) permite conocer las características de una persona.

3.7 Administración de los PHDs

Una vez que se ha generado y almacenado un PHD es precisa su administración. La administración de un PHD consiste en generar un número de registro. Dicho número es necesario para almacenar y buscar posteriormente. El Registro del PHD (RPHD) es un número que contiene metadatos. Los metadatos son los siguientes: Fecha de registro, estatus, si ha sido ubicada, el número del estado del País y número de dependencia. Cuando no exista algún dato solo habrá ceros. Ver Tabla 4.

Existen diversos métodos y estándares para generar códigos de barras. Se ha elegido Code 39. Dicho código puede codificar letras mayúsculas, números y algunos caracteres especiales. Cada carácter se representa por nueve barras, tres anchas y 6 delgadas. El código de barras Code 39 es muy utilizado en etiquetas de productos e inventarios [14].

RPHD	Descripción
01012025010112006 	01012025 corresponde a la fecha 01 corresponde al estatus (ubicado) 01 corresponde al país 12 corresponde al estado del País donde está físicamente 006 Número de dependencia de resguardo de cadáveres
01012025000100000 	01012025 corresponde a la fecha 00 corresponde al estatus (no ubicado) 01 corresponde al país 00 corresponde al estado del País donde está físicamente (no ha sido encontrado) 000 Número de dependencia de resguardo de cadáveres (no ha sido localizado)

Tabla 4 Códigos de barras de metadatos de cada PHD.

3. Resultados

Una vez obtenida la matriz de números de categorías es posible generar la presentación en dos formas posibilidades. La primera consiste en una matriz de números enteros. Ver la Figura 6 donde se muestra el PHD de dos personas.

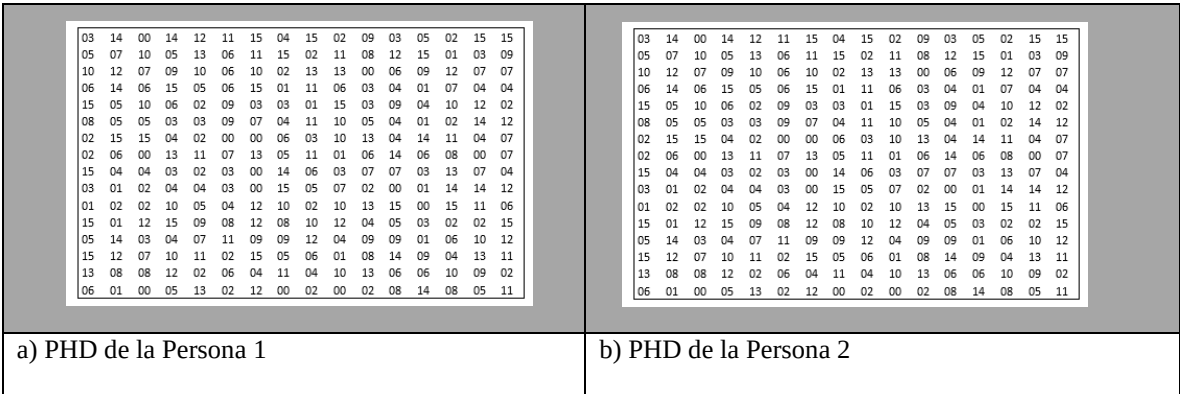


Figura 6 PHDs obtenidos de dos personas

En la Tabla de tokens muestran números y ceros. La mayoría de la tabla muestra ceros porque corresponden a características de tatuajes, cicatrices y lunares.

La otra posibilidad consiste en mostrar el MPHD en un código QR. Dicho código almacena los 512 dígitos. Mediante la aplicación correspondiente es posible leer todos los valores con una captura. Ver la Figura 7.

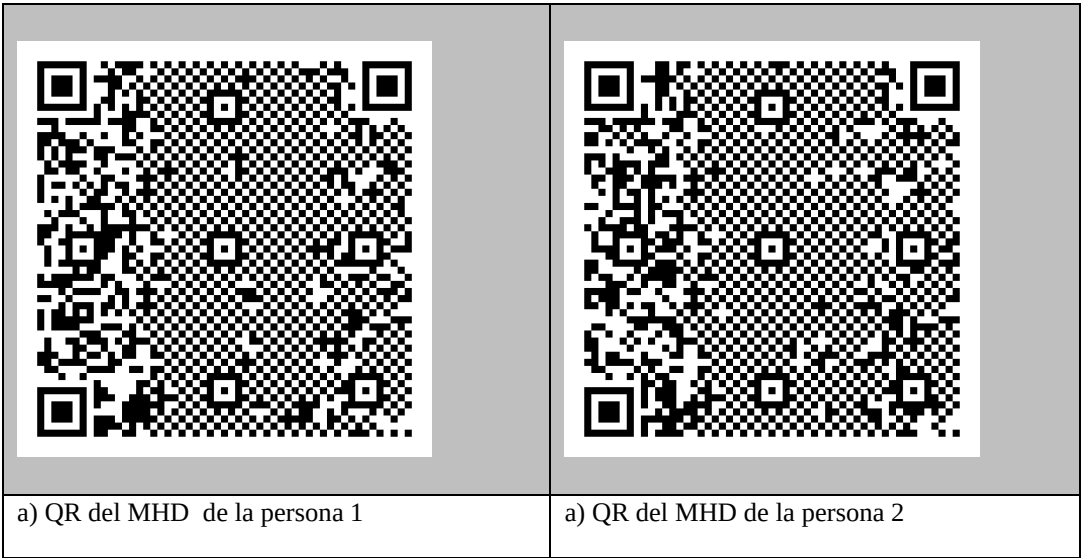


Figura 7 QR obtenidos a partir de la matriz de los PHDs. Groupdocs

(<https://products.groupdocs.app/es/generator/qrcode/excel>) [15]

Los usos posteriores que puede tener el PHD pueden ser: El primero, para clasificar y generar agrupamientos. El segundo, para buscar a personas desaparecidas. Lo anterior, generar un VCC de una persona buscada y otro de una persona encontrada sin vida y sin identificación. Luego, comparar los VCC en el conjunto de VCC almacenados en el Repositorio. Si existe un matching, existe una gran probabilidad de haberle identificado.

5. Conclusiones

La generación del PDH fue exitoso con el método CVT. La interfaz gráfica para la Caracterización resultó eficaz, debido a que solo se elige una imagen para cada característica. La aplicación genera el VCC completo. El VCC se propone sea un archivo de datos independiente. Cada VCC se almacena con su respectivo Código de metadatos. La generación de un QR resulta práctica porque se puede manejar como una etiqueta. Además, es posible capturar con el celular y obtener el VCC. No es necesario conocer las características de la persona, debido a que ha generado para un algoritmo. En trabajos futuros se está buscando que los VCC estén en un repositorio y un algoritmo haga las comparaciones para identificar personas. Igualmente, se busca que sea una plataforma de acceso público para búsquedas constantes desde cualquier parte de México o el mundo.

Referencias

- [1] A. Malpartida, Identificación de personas mediante la optimización de la odontograma de la norma técnica del Perú, Cusco 2018. Universidad Norbert Wiener, Perú, 2019.
- [2] D. J. Vázquez, La carta dental una técnica forense que da certeza de identificación, Universidad Nacional Abierta ya a Distancia, Escuela de –ciencias de la Salud, 2022.
- [3] C. U. Hernández, N. S. Huerta, I. Gil, Revisión sistemática de la Caries rampante como criterio de individualización en cuatro condiciones para el proceso de

identificación de personas en México, Revista digital de Ciencia forense, 1 (1): 38-58 pp. México, 2022.

- [4] R. E. Pedroza, Diseño e implementación de un sistema de biometría facial para la búsqueda e identificación de personas desaparecidas en Colombia. Universidad de Cartagena, Facultad de ingeniería, Colombia, 2019.
- [5] E. C. Lescano, Validación del método morfológico comparativo facial, para identificación de personas mediante registro fotográfico, Lima, 2021.
- [6] A. D. Briem A. D., H. Velázquez H., Propuesta de ficha queiloscópica para registro biométrico de huellas labiales con fines de identificación forense, Gaceta internacional de Ciencias Forenses, Argentina, 2021.
- [7] L. Cocco; G. Diaz; G. Di Girolamo; M. Alfaro; S; M. Brown; A. Elvira; J. Papasodaro; S. Degaetano; C. Guichon; G. Pezzuchi., La queiloscopía en la autenticación humana. Patrones de codificación aplicables en biometría, Instituto de Investigaciones en Educación Superior (IIES) Facultad de Odontología – UNLP, Argentina, 2019.
- [8] A. G. Ceballos. y L. Rodes, Análisis de las variaciones postquirúrgicas de la morfología de las orejas y su sutilidad como método de identificación forense, Boletín Galego de Medicina Legal e Forense no 35, 34-48 pp., España, 2024.
- [9] A. G. Ceballos. y L. Rodes., Tesis: Análisis en las variaciones de las orejas a causa de lesiones o traumatismos sufridos y su utilidad como método de identificación forense, Universidad de Alicante. España, 2024.
- [10] A. J. Montoya, La pelmatoscopia como sistema de identificación humana, Tecnológico de Antioquia, Colombia, 2021.
- [11] G. Melo., I. D. Jácome., P. Denis, J. I. Baltazar, A. Jiménez, E. Denis, J. Huesca. y S. Vaca, Identificación forense en cadáveres no identificados a partir de tatuajes reporte de un caso. Persep Med Legal Pericias Med., 2022. <https://dx.doi.org/10.47005/220611>.
- [12] P. S. S. Mejía y A. Ortega, La entrevista a familiares en la búsqueda e identificación de personas extraviadas o desaparecidas, Dimensión Antropológica, Año 25, Vol. 72, 141-172 pp., México, 2018.
- [13] A. Maldonado, D. Ruano, N. Herrera, M. Martínez, Un Sistema para la Identificación de Cadáveres NN en el Contexto de Búsqueda de Personas Desaparecidas, Repositorio institucional de la UNLP, Universidad Nacional de la Plata, Argentina, 2022.
- [14] ActiveBarcode, Descripción de Code 39. Consultado el 26 de agosto de 2024. Disponible en: <https://es.activebarcode.com/codes/code39>
- [15] Groupdocs. Generador de QR para Excel. Disponible en: Consultada el 26 de agosto de 2024. <https://products.groupdocs.app/es/generator/qrcode/excel>
- [16] M. G. Moreno. Dibujo de la figura humana. UAM Universidad Autónoma Metropolitana. 1991.

