



ORIGINAL

Inteligencia Artificial en Antropología Forense: estado del arte, retos y oportunidades.

ARTIFICIAL INTELLIGENCE IN FORENSIC ANTHROPOLOGY: STATE OF THE ART, CHALLENGES AND OPPORTUNITIES.

Ibañez O.^{1,2}, Martos R.^{1,3}, Mesejo P.^{1,2,4}

1 Panacea Cooperative Research S. Coop., España.

2 Instituto Andaluz Interuniversitario en Data Science and Computational Intelligence, Universidad de Granada, España.

3 Departamento de Antropología Física, Universidad de Granada, España.

4 Departamento de Ciencias de la Computación e Inteligencia Artificial y Centro de Investigación Tecnologías de Información y Comunicaciones de la Universidad de Granada, España.

RESUMEN: La Inteligencia Artificial (IA) representa un vasto campo de investigación de las ciencias de la computación. Este campo científico, creado en los años 40 y 50 del siglo pasado, está compuesto por disciplinas como el procesamiento del lenguaje natural, la representación del conocimiento, el razonamiento automatizado, el aprendizaje máquina, la visión por ordenador y la robótica. En las últimas décadas, la IA ha permitido automatizar tareas repetitivas o tediosas para los seres humanos (por ejemplo, la automatización de procesos industriales o tareas de limpieza), así como superar a los seres humanos en la realización de tareas complejas (por ejemplo, el procesamiento de cantidades masivas de datos para extraer nuevos conocimientos o superar a los campeones humanos jugando al Ajedrez o al Go). Recientemente, los avances relacionados con el aprendizaje automático y la visión por ordenador, bajo el paraguas terminológico del aprendizaje profundo, han proporcionado avances sorprendentes en el reconocimiento de imágenes, la restauración de imágenes, la generación de imágenes, el reconocimiento del habla y la traducción automática, entre otros. El campo de la medicina no ha sido una excepción: la IA ha proporcionado herramientas tremendamente útiles para los profesionales en la estimación de parámetros, la segmentación de imágenes, la clasificación de patologías o la mejora de imágenes, por nombrar sólo algunas. Por lo tanto, es sorprendente la forma en que la Antropología Forense (AF) se ha mantenido al margen de estos avances y, aún hoy, en términos generales, es una disciplina esencialmente manual y precaria a nivel tecnológico. Por consiguiente, este artículo se centra en mostrar los beneficios de la utilización de las técnicas de IA en la AF, centrándose en aquellas técnicas de IA que son más potencialmente útiles en las aplicaciones de la AF. Por otra parte, también se presenta la literatura existente sobre la hibridación entre la IA y la AF, así como posibles líneas de trabajo futuro.

PALABRAS CLAVE: Inteligencia Artificial, Aprendizaje automático, Visión por Computador, Soft Computing, Antropología Forense, Perfil Biológico, Radiografía Comparativa, Identificación Craneofacial, Superposición Craneofacial. Reconstrucción facial

ABSTRACT: Artificial Intelligence (AI) represents a vast field of research in computer science. This scientific field, developed in the 40s and 50s of the last century, is comprised of disciplines such as natural language processing, knowledge representation, automated reasoning, machine learning, computer vision, and robotics. In the last decades, AI has allowed to automate repetitive or tedious tasks for human beings (e.g., the automation of industrial processes or cleaning tasks), as well as surpassing human beings in performing complex tasks (e.g., processing massive amounts of data to extract new knowledge or overcome human champions playing Chess or Go). Recently, advances related to machine learning and computer vision, under the terminological umbrella of deep learning, have provided astonishing advances in image recognition, image restoration, image generation, speech recognition, and machine translation, among others. The medical field has not been an exception: AI has provided useful tools for practitioners in parameter estimation, image segmentation, pathology classification, and image enhancement, to name a few. Therefore, it is surprising how forensic anthropology (FA) has remained behind from these advances and, still today, in general terms, it is an essential manual and precarious discipline at the technological level. This paper focuses, therefore, on showing the benefits of using AI techniques in FA, focusing on those AI techniques that are more potentially useful in FA applications. On the other hand, the existing literature on the hybridization between AI and FA, as well as possible lines of future work, is also presented.

KEY WORDS: Artificial Intelligence, Machine Learning, Computer Vision, Soft Computing, Forensic Anthropology, Biological Profiling, Comparative Radiography, Craniofacial Identification, Craniofacial Superimposition.

CONTACTO: Óscar Ibañez. Email: oscar.ibanez@panacea-coop.com

1. INTRODUCCIÓN.

La identificación humana (ID) es de gran importancia en nuestra sociedad. No sólo resuelve los graves problemas legales y sociales, sino que proporciona una solución a las familias afligidas que necesitan cerrar su tristeza. El número irracional (y vergonzoso) de personas no identificadas en todo el mundo es asombroso y sigue creciendo debido a los repetidos desastres masivos, conflictos bélicos, desapariciones forzadas e investigaciones criminales de personas muertas y desaparecidas, etc. Todos estos factores contribuyen a la creciente necesidad de contar con instrumentos precisos y eficaces que permitan establecer y verificar la identidad humana. El número de ejemplos es espantoso. Los desastres naturales han causado más de 1,3 millones de muertes en los últimos dos decenios [1]. Por ejemplo, el reciente ciclón en Mozambique (2019) causó más de 700 muertes [2], y el devastador terremoto de 2010 en Haití mató a más de 160.000 personas que aún no han sido identificadas. Los países que han sufrido guerras recientes, como Chile, Argentina, Panamá y Sudáfrica, han establecido comisiones de la verdad [2], lo que constituye uno de sus principales objetivos para la identificación de sus víctimas. Dentro de la UE, existen iniciativas similares en España (Asociación para la Recuperación de la Memoria Histórica, ARMH), donde según la ONU, se estima que hay más de 100.000 personas no identificadas (2º país del mundo, después de Camboya) en fosas comunes de la guerra civil [3] o en los Balcanes (Comisión Internacional de Personas Desaparecidas, ICMP). En ciertos países de América Latina, la situación es particularmente dramática debido a sus altos niveles de muertes violentas y desapariciones forzadas [4]. Sólo en México, según la Comisión Nacional de Búsqueda (CNB) del país, hay más de 26.000 cadáveres no identificados y más de 40.000 personas desaparecidas [5], y los movimientos sociales y políticos están exigiendo soluciones a esta epidemia y al problema de identificación de las víctimas. Los países vecinos cuentan por miles el número de personas desaparecidas también, Colombia (más de 83.000), Guatemala (45.000), Argentina (>30.000), Perú (>20.000), El Salvador (>8.000), Chile (>3.200). Por otra parte, han surgido nuevos retos en el ámbito de la identificación de personas vivas, en el que la existencia de técnicas de identificación precisas puede ser decisivo. En los últimos años, la UE se enfrenta a una ola de migración masiva sin precedentes, que está creando una tensión en nuestras fronteras, y un número cada vez mayor de muertos no identificado en el Mediterráneo [6]. Además, uno de los retos asociados más delicados es la estimación de la edad de los migrantes menores no acompañados: más de 20.000 durante 2017, alrededor del 90% de 15 a 17 años, en su mayoría indocumentados [7]. La estimación forense de la edad es solicitada por los tribunales y otras autoridades gubernamentales, para que los inmigrantes -en la mayoría de los casos, refugiados no acompañados que pueden ser

menores y cuya edad real se desconoce- no sufran desventajas injustas debido a su edad estimada. Según el Alto Comisionado de las Naciones Unidas para los Refugiados (ACNUR), casi la mitad de las personas desplazadas por la fuerza de sus hogares son niños, lo que representa un grupo demográfico que corre un riesgo especial de "abuso, abandono, violencia, explotación, trata o reclutamiento militar forzoso" y que necesita protección internacional. Además, otra serie de métodos que tienen una relación con las estructuras anatómicas subyacentes, los métodos de identificación basados en el reconocimiento facial pueden utilizarse en delitos con presencia de cámaras de videovigilancia u otras pruebas de imagen para identificar criminales, detectar la presencia de menores en material pedopornográfico o en el control de fronteras, entre otros.

Los gobiernos se esfuerzan por abordar esta tarea titánica, ya que están totalmente desbordados y carecen de los recursos humanos y la tecnología forense adecuados que se necesitan. Aunque todas las cuestiones jurídicas y sociales son difíciles de resolver, la aplicación de sistemas de identificación eficientes y asequibles contribuiría a cerrar el duelo de muchas familias y a proteger los sectores más vulnerables de la sociedad.

El ADN y las huellas dactilares son los métodos de identificación más comunes. Sin embargo, el ADN requiere una muestra contra la que comparar que puede no estar siempre disponible, además de ser una técnica costosa y que requiere mucho tiempo. Las principales limitaciones de las huellas dactilares son la necesidad de que se conserven los surcos interpapilares en la yema de los dedos del cadáver, la necesidad de una huella dactilar AM en una base de datos consultable junto con un costoso sistema AFIS. De ahí que la eficiencia de estas técnicas de identificación de referencia disminuya drásticamente en los escenarios de identificación de desastres de masas.

La antropología forense (AF) representa una alternativa y una contrapartida con un rango de aplicación mucho más amplio. Según el Grupo de Trabajo Científico de Antropología Forense [8], los antropólogos forenses contribuyen a la identificación en dos niveles. El primer nivel es a través de métodos que establecen una identificación. El segundo nivel es a través de métodos que contribuyen a la identificación limitando las posibles coincidencias con el individuo. Entre el primer grupo, el SWGANTH incluye la radiología comparativa (RC) y la comparación de implantes quirúrgicos, mientras que deja para el segundo nivel las siguientes técnicas o líneas de evidencia: Superposición craneofacial (SCF), perfil biológico (PB), registros médicos y/o dentales, anomalías congénitas y condiciones patológicas, y fotografía comparativa. Véase la figura 1 para una visualización esquemática del rango de aplicabilidad de las técnicas de AF.

La identificación a partir de implantes quirúrgicos implica la localización del símbolo del fabricante junto con el número de serie único del dispositivo, que, si bien no tiene ninguna complicación técnica, se encuentra con grandes dificultades ya que los implantes más antiguos no son fáciles de rastrear. Los métodos restantes son en gran parte complejos de aplicar. A continuación, resumimos las principales

dificultades que experimentan los médicos / antropólogos forenses al aplicar estas técnicas. A veces no se consideran dificultades, ya que se suponen como limitaciones de la naturaleza del problema en cuestión. Sin embargo, hay pruebas empíricas en campos científicos cercanos de que esas limitaciones pueden transformarse en desafíos que pueden abordarse siguiendo un enfoque interdisciplinario.

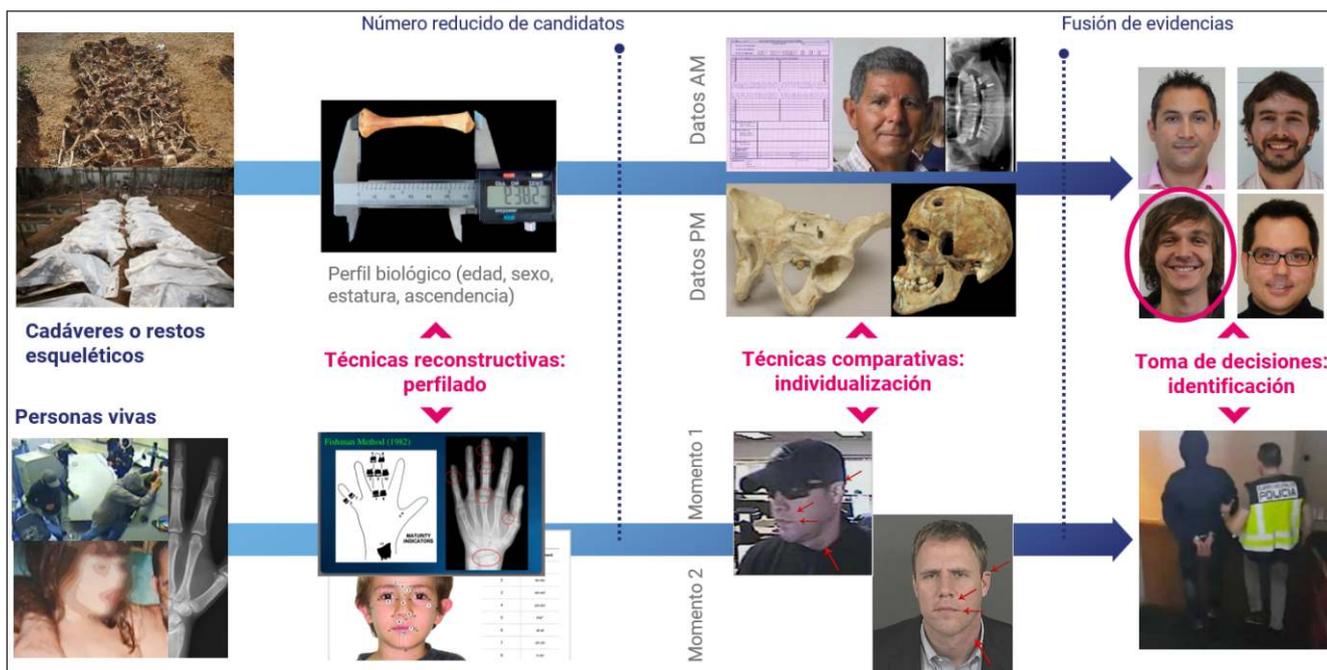


Figura 1: Aplicación de las técnicas de antropología forense a los casos de personas vivas y muertas.

Técnicas reconstructivas (perfilado):

- Mediciones manuales (basadas en el uso del calibre), variables (dispersión intra- e inter-observador) y repetitivas.
- Comparación subjetiva e inexacta de formas (categorización basada en el atlas).
- Patrones desconocidos y/o relaciones ocultas en una gran cantidad de datos multidimensionales.
- Falta de consenso y buenas prácticas en los estudios de validación.

- El filtrado manual dificulta la aplicación en desastres masivos y grandes bases de datos.

Algunos de los métodos de estimación del PB más utilizados se desarrollaron hace 50 a 100 años (o más), como es el caso del método basado en el atlas (GP) de Greulich y Pyle [9] o el de Tanner-Whitehouse (TW2) [10]. Se elaboraron con un número limitado de muestras¹ (de unos pocos cientos a unos pocos miles), normalmente sesgadas por una ascendencia común y un rango de edad limitado en comparación con las poblaciones modernas. En la actualidad, gracias al avance tecnológico, el acceso a TACs y a las autopsias virtuales, resulta más sencillo para los antropólogos obtener información sobre las poblaciones actuales superando esas limitaciones del pasado.

1. Aunque un conjunto de datos de unos pocos miles de muestras se considera probablemente una situación ideal (raramente alcanzada) en el campo de la AF, otros campos de investigación forense y de ciencias biológicas desarrollan sus modelos utilizando conjuntos de datos mucho más grandes, como es el caso de las huellas dactilares (hasta 500.000 pares se emplean para validar el sistema en [11]), el reconocimiento facial (3,7 millones de imágenes de 58.207 sujetos en [12]) o la clasificación de imágenes médicas (se emplearon dos millones de radiografías de tórax en [13]) para la clasificación de patologías; 129.450 imágenes clínicas en [14] para la clasificación del cáncer de piel a nivel de dermatológico).

Además, las fases o patrones fueron observados y documentados (categorización basada en el atlas) por uno o unos pocos investigadores sobre la base de su experiencia, conocimientos y capacidad de reconocimiento de patrones. Por una parte, se limita a la capacidad humana de identificar patrones y de trabajar con un enorme conjunto de datos de muestras al mismo tiempo. La aplicación de estos métodos basados en el atlas está vinculada a la subjetividad (por la naturaleza del problema de hacer coincidir holísticamente un patrón/forma con un dibujo del patrón/forma promedio) y a la dispersión intra- e inter-observador [15,16]. Un gran grupo de métodos se basan en una serie de mediciones lineales tomadas directamente del hueso limpio con el calibre o en imágenes radiográficas. Por último, el número de métodos de estimación del PB publicados puede contarse por cientos. Sin embargo, en la actualidad no hay un consenso claro sobre cómo calcular, expresar e interpretar el error al validar los métodos de estimación de la edad, mientras que al mismo tiempo existen pruebas empíricas sobre la inadecuación de algunos métodos de estimación de la edad para calcular, expresar e interpretar los errores obtenidos [17].

Técnicas comparativas (individualización):

- Observaciones visuales complejas y subjetivas.
- Necesidad de adquirir la imagen del hueso post-mortem (PM) en la misma postura y condiciones que el hueso ante-mortem (AM).
- Muy dependiente de los conocimientos y la experiencia del antropólogo.
- Muy alta variabilidad en la precisión y poca o ninguna validación.
- Falta de estandarización y acuerdos internacionales.
- Uso de software genérico de procesamiento de imágenes, muy limitado, introduce nuevos errores.

Las características implicadas en el emparejamiento físico y el emparejamiento visual de patrones existentes en el estado de la naturaleza no son eventos susceptibles de ser contados; en cambio, se trata de un conjunto de unidades visibles que ocupan toda el área del patrón, extendiendo la probabilidad de que vuelva a ocurrir un patrón de verosimilitud hasta el infinito, ofreciendo un apoyo epistémico a la unicidad [18]. La aplicación de técnicas comparativas solía reducirse a la comparación subjetiva de los tipos de patrones anteriores mediante métodos de observación como el emparejamiento físico, la superposición o la evaluación visual lado a lado.

La normalización es una cuestión importante en las técnicas comparativas y en la etapa de toma de decisiones. Si bien hay ejemplos meritorios de propuestas individuales o de equipo para la normalización, por ejemplo, en la nomenclatura de antropometría craneofacial[19] o la identificación basada en la RC [20], existe una laguna de iniciativas internacionales/multicentro, como por ejemplo la iniciativa MEPROCS de la UE[21].

La principal dificultad y fuente de errores son los diferentes momentos de adquisición (de meses a años), los diferentes entornos de adquisición (distancia cámara-sujeto, distancia focal, pose relativa, etc.), los diferentes dispositivos de adquisición (diferentes modelos de cámara o de rayos X) e incluso la diferente naturaleza de los materiales AM y PM que se deben comparar, como es el caso de la técnica SCF en la que se comparan una o más fotografías de la cara con un cráneo. En cualquier caso, la comparación lado a lado debe ir precedida de un proceso de adquisición de datos PM que intente imitar las condiciones de adquisición AM. Esta tarea se aborda normalmente en un proceso de ensayo y error, que lleva mucho tiempo y el error no puede medirse ni estimarse, es decir, su validez depende en gran medida de los conocimientos y la experiencia del antropólogo [22].

Toma de decisiones (identificación):

- Decisiones basadas en un conocimiento parcial y aproximado con una alta presencia de incertidumbres y conjeturas.
- Difícil visualización de los datos e interpretación de los resultados.
- Diferentes grados categóricos en la toma de decisiones.
- Fuertemente dependiente del experto y difícil de defender en procedimientos legales.
- Falta de reproducibilidad. Ausencia de estudios sólidos de fiabilidad
- Limitaciones de tiempo: inabordable para múltiples comparaciones / desastres masivos / fosas comunes.

Por último, el proceso de toma de decisiones que conduce a la decisión final de identificación se expresa normalmente en diversos grados lingüísticos (apoyo fuerte, moderado o limitado; presunto o positivo, etc.) que explican la presencia de incertidumbre e información incompleta. La forma en que el experto combina y pondera las diferentes pruebas no puede expresarse cuantitativamente, depende de los conocimientos y la experiencia del experto. Todo el proceso de identificación carece de reproducibilidad y es casi imposible de aplicar a un escenario de comparación múltiple

(a menos que se disponga de un equipo de expertos durante un período de trabajo prolongado).

Los grandes estudios de fiabilidad no son comunes en el campo de la AF, ya que todavía existen barreras éticas y jurídicas no superadas para la adquisición y el intercambio de conjuntos de datos públicos. Además, la falta de normalización y de reproducibilidad implica conclusiones contradictorias en cuanto a la fiabilidad de las técnicas. Este es, al menos, el caso de la SCF, en el que los diferentes estudios de fiabilidad publicados hasta la fecha están llenos de limitaciones. La ausencia de una medida objetiva del grado de emparejamiento cráneo-cara, las limitaciones técnicas del software utilizado, la falta de consideración de la localización precisa de los landmarks al realizar los métodos basados en estos, la ausencia de datos sobre el grosor del tejido blando de la población examinada, la calidad insuficiente de los modelos craneales tridimensionales, las fotografías post mortem, las muestras reducidas, la ausencia de análisis estadísticos apropiados, la ausencia de estudios inter- e intra- observador son sólo algunas de ellas[23].

A modo de conclusión, resumimos todas las limitaciones identificadas en los métodos existentes en un solo desafío: la necesidad de métodos automáticos, objetivos y más precisos. Este parece ser el entorno perfecto para los sistemas computarizados basados en técnicas de Inteligencia Artificial (IA) [24], debido a su capacidad para modelar el comportamiento humano en la resolución de problemas específicos, y su capacidad para procesar cantidades de datos extremadamente grandes y descubrir patrones ocultos.

A pesar de la conveniencia teórica y empírica de la aplicación de las técnicas de IA a la AF, en el último decenio sólo se han publicado un número limitado de trabajos basados en IA. Sin embargo, han demostrado su precisión y solidez, de ahí su viabilidad para las aplicaciones de identificación forense en humanos. El presente artículo tiene por objeto dar visibilidad y aumentar la conciencia de la comunidad de AF sobre la potencialidad de la IA para resolver problemas de este ámbito, mediante la realización de la primera revisión bibliográfica. De hecho, consideramos que existe un problema de falta de investigación interdisciplinar. Esta afirmación se origina en el hecho de que sólo hay unos pocos grupos de investigación multidisciplinarios en todo el mundo donde los médicos forenses o antropólogos trabajan junto con especialistas en IA. Por otro lado, pensamos también que hay una falta de desarrollo tecnológico en la comunidad de AF que, unido a la alta complejidad de algunas de las técnicas de antropología, ha desembocado en errores e imprecisiones. Como consecuencia de estos, el uso de técnicas como la SCF ha sido relegado y desalentado. A pesar de los beneficios demostrados, el uso de las técnicas de IA en la casuística forense diaria es prácticamente nulo.

El contenido de este artículo está organizado de la siguiente manera. La sección 2 supone una introducción a un importante grupo de técnicas de IA, aquellas que nosotros consideramos más adecuadas para abordar los principales desafíos del campo de la AF y en las que se basan la mayoría de las propuestas existentes hasta la fecha. La sección 3 representa la parte central de este trabajo, el estudio del estado del arte relativo a la aplicación de técnicas de la IA en el ámbito de la AF. Este se divide a su vez en cuatro partes que recogen, de manera casi exhaustiva, los enfoques basados en técnicas de IA para abordar: la estimación del PB a partir del estudio del esqueleto, el análisis traumatológico y de patologías, la RC, la SCF y finalmente la aproximación facial (o reconstrucción facial). Para cada uno de estos métodos de AF se incluye también un resumen de los fundamentos y trabajos de referencia (que para nada pretende ser exhaustivo ni preciso), con el objetivo de que la presente publicación sirva, no solo para acercar la AF a la IA, sino también el camino inverso, que sea capaz de atraer el interés de la comunidad de IA a la problemática de los AF. Por último, la sección 4 incluye algunas observaciones finales, y una discusión sobre las acciones que los autores consideran más importantes de cara a avanzar en el campo de la AF.

2. INTELIGENCIA ARTIFICIAL: TÉCNICAS Y FUNDAMENTOS

Las siguientes son algunas ramas específicas de la IA, que, en nuestra opinión, son las herramientas más adecuadas para abordar los principales desafíos del campo de la FA. Es importante señalar que estos campos de investigación de la IA no corresponden a categorías desarticuladas o herméticas, pues hay un alto grado de permeabilidad e hibridación entre ellos.

2.1. Visión Artificial.

La Visión por Ordenador, o Visión Artificial (VA) es la disciplina científica que se ocupa de la interpretación automática de las imágenes [25]. Cubre diferentes niveles de abstracción. En primer lugar, la visión por ordenador de bajo nivel, también llamada procesamiento de imágenes, implica operaciones primitivas que consideran las propiedades locales de una imagen. Este nivel toma como entrada una imagen y da como salida la misma imagen transformada de acuerdo con la operación mencionada (por ejemplo, la restauración y el realce de la imagen). En segundo lugar, la visión por computador de nivel medio, también llamada análisis de imágenes, entraña etapas intermedias previas a la comprensión semántica de la imagen. En este nivel se toma

como entrada una imagen y se da como salida un conjunto de rasgos que caracterizan la imagen (por ejemplo, la segmentación de la imagen y la detección de características). Por último, la visión por computador de alto nivel tiene por objeto "dar sentido" a la imagen/vídeo de entrada y desarrollar las funciones cognitivas relacionadas con el sistema de visión humana (por ejemplo, el reconocimiento de objetos y el registro de imágenes).

Dentro de la VA, entre las técnicas concretas con más potencial de aplicación en el campo de la AF, destacamos las siguientes:

La segmentación de imágenes (SI), que se define como la división de una imagen en regiones que son homogéneas con respecto a algún rasgo visual [26]. Más precisamente, es el proceso de asignación de una etiqueta (clase) a cada píxel de la imagen de forma que los píxeles que compartan la misma etiqueta también tendrán ciertas características visuales similares, como el color, la intensidad o la textura. Además de las propiedades individuales de cada píxel, en problemas de segmentación más complejos será necesario

tener en cuenta información a distintos niveles jerárquicos, desde el píxel a regiones e incluso la imagen entera, donde se analizarán cambios en las características de los píxeles, distancias o formas. La figura 2 muestra un ejemplo de segmentación multiclase, es decir, se segmentan diferentes clases de objetos, en concreto, pulmones, clavículas y corazón en una radiografía. La segmentación precisa de órganos y huesos en radiografías es especialmente difícil debido al solapamiento que existe entre estos y por lo difuso de los contornos de los "objetos" a segmentar en muchas zonas de la imagen. De hecho, en este problema concreto ha sido necesario un proceso de aprendizaje (más adelante se explican los métodos de aprendizaje automático) en donde, a partir de cientos de radiografías segmentadas previamente por un radiólogo de manera manual (usando un software de tratamiento de imágenes, fue cuidadosamente delimitando los contornos de pulmones, corazón y clavículas), se aprenden características de cada uno de los "objetos" de interés (forma típica y posibles variaciones, localización dentro de la imagen y respecto a otros objetos, rango de intensidades y texturas).

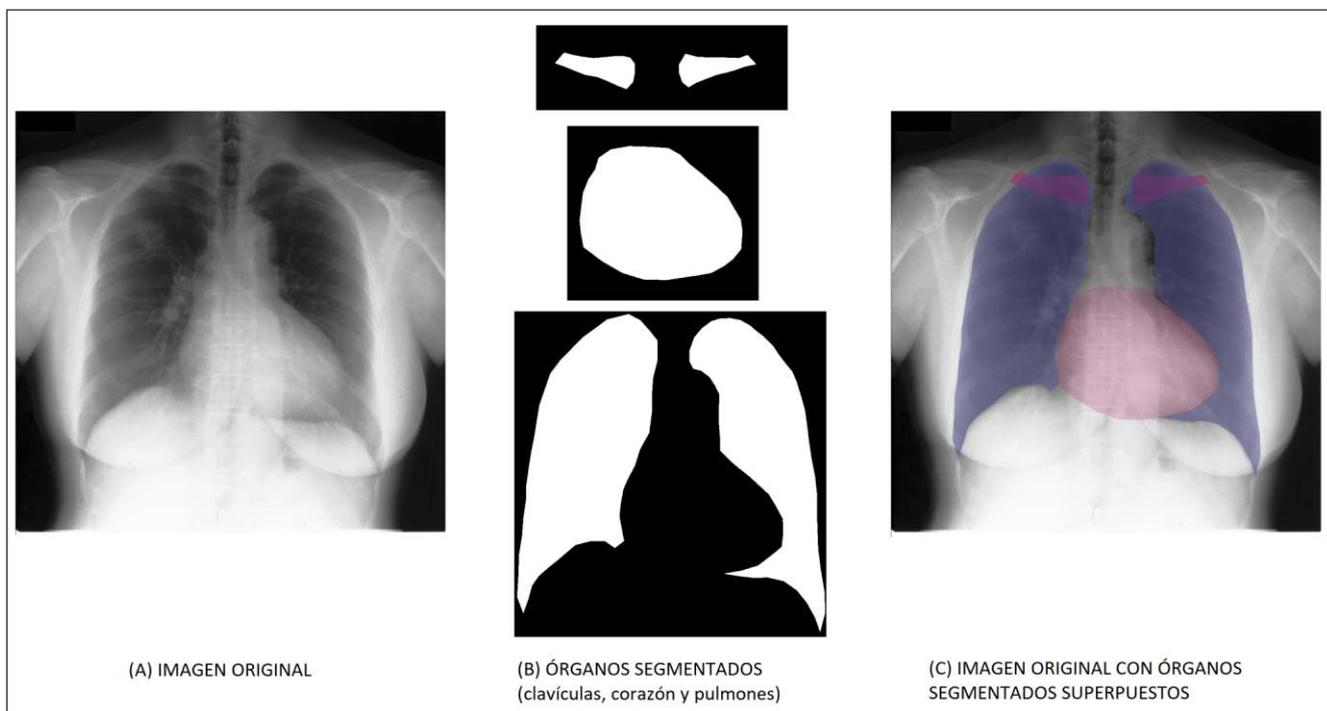


Figura 2: ejemplo de SI. La imagen original es una radiografía de tórax (A). Mediante el algoritmo automático de segmentación de imágenes en se extraen las clavículas, el corazón y los pulmones (B). La imagen de la derecha (C), representa sobre la imagen original, los órganos y huesos segmentados con diferentes colores.

La detección, clasificación y comprensión de imágenes: fuertemente relacionada y dependiente de la SI, consiste en detectar/localizar objetos dentro de una imagen (dibujando, por ejemplo, un rectángulo alrededor de ellos), categorizar esos objetos e incluso, describir matemáticamente propiedades de los objetos detectados[25]. La comprensión de la imagen es el proceso de interpretación real de esas

regiones/objetos para averiguar lo que realmente está sucediendo en la imagen. Esto puede incluir, en última instancia, la toma de alguna decisión para la acción futura. En la figura 3, se muestra un ejemplo de esto último, un sistema que detecta armas de fuego a partir de las imágenes de video-vigilancia y activa una alarma cuando eso ocurre.



Figura 3: en la imagen de la izquierda podemos ver un ejemplo de detección de armas . En este ejemplo el algoritmo desarrollado es capaz de detectar armas y clasificarlas asignando una probabilidad (como puede verse en la imagen con la clase pistola) en tiempo real, es decir, sobre fotogramas de un video a la vez que este se visualiza. En la imagen de la derecha se puede observar el resultado de un algoritmo de detección y clasificación automática de landmarks cefalométricos.

El registro de imágenes (RI), que tiene como objetivo encontrar una transformación geométrica que superponga dos imágenes tomadas en diferentes momentos, desde diferentes puntos de vista, y/o por diferentes sensores [30]. Según sea la naturaleza los sistemas de adquisición, podemos distinguir métodos de RI de diferente modalidad (monomodalidad / multimodalidad) o dimensionalidad. Dentro de este último grupo podemos distinguir: 2D-2D (es decir, entre 2 fotos o 2 radiografías, por ejemplo), 3D-3D (entre dos modelos 3D, por ejemplo, dos escaneos intraorales de la misma persona con un mes de diferencia), 3D-2D (entre un modelo 3D y una imagen 2D, por ejemplo, un modelo 3D de un cráneo y una fotografía de la cara de una persona). Otras posibles clasificaciones de métodos de RI incluyen: el registro intersujeto (imágenes de diferentes personas, por ejemplo, para estudiar la morfometría del cráneo y relacionarla posteriormente con variables como sexo o ascendencia) y diferentes métodos según la

información de referencia usadas (extrínseca / intrínseca) y las métricas de similitud empleadas para medir el error de registro (basados en intensidad / basado en características).

La VA podría ser sumamente valiosa en la AF. Tomemos como ejemplo la comparación de una imagen AM de rayos X y un modelo 3D PM para determinar si el hueso que aparece en ambas pertenece a la misma persona. Primero, necesitaríamos aislar la estructura anatómica de interés en los datos AM y PM. Esto implica un proceso de segmentación de la imagen (ver figura 2). A continuación, podríamos alinear automáticamente la proyección de la segmentación resultante de los datos PM sobre la imagen segmentada de rayos X AM. Este proceso se conoce como registro de la imagen (ver figura 4).

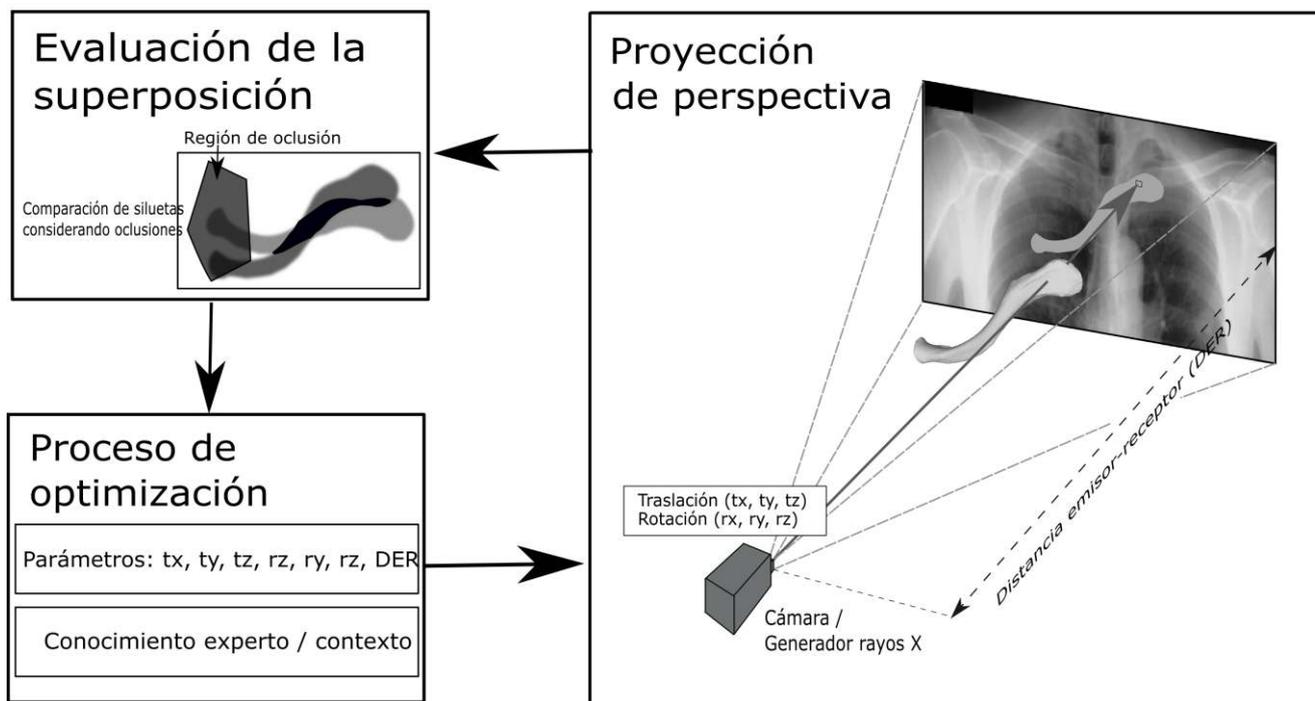


Figura 4: ejemplo de proceso iterativo de IA en donde un modelo 3D de una clavícula es registrado sobre una radiografía (2D) de tórax. En este proceso iterativo, un algoritmo de optimización busca reproducir los parámetros de adquisición de la radiografía AM, es decir, busca los valores de traslación, rotación y distancia del emisor de rayos X a la placa. Este proceso de búsqueda está guiado por el error (proceso de minimización del error) de solapamiento entre el contorno del modelo 3D (una vez proyectado sobre la radiografía) y el contorno de la clavícula (obtenido mediante un proceso de SI) en la radiografía.

2.2. Aprendizaje Automático.

El Aprendizaje Automático (AA) es la rama de la IA cuyo objetivo es el desarrollo de técnicas que permitan a los ordenadores aprender directamente de los datos [31], es decir, aprender modelos generales a partir de ejemplos particulares de datos. El AA representa un vasto campo de investigación que incluye, sin pretender ser exhaustivo:

- El aprendizaje supervisado: en el que el conjunto de datos empleados para entrenar los modelos incluye entradas y salidas deseadas.
- El aprendizaje no supervisado (en el que el sistema, según una noción de similitud, tiene que ser capaz de reconocer patrones basados únicamente en los ejemplos de entradas)
- El aprendizaje de refuerzo (en el que el sistema aprende mediante un procedimiento de ensayo y error, interactuando con el entorno, basado en una señal genérica de retroalimentación sobre lo bien o mal que se está realizando la tarea en cuestión).

Los datos son baratos y abundantes mientras que el conocimiento es caro y escaso, de ahí que se utilicen las técnicas de AA para resolver aquellos problemas en los que no existe experiencia humana (por ejemplo, navegar en Marte), donde los humanos son incapaces de explicar su experiencia (el reconocimiento de voz), cuando la solución debe adaptarse a casos particulares (biometría de usuario) o el tamaño del problema es demasiado grande para nuestras limitadas capacidades de razonamiento (calculando los rangos de las páginas web).

Desde un punto de vista práctico, el AA se ocupa principalmente de la resolución de problemas de regresión (es decir, la estimación de valores continuos), problemas de clasificación (es decir, la estimación de valores o clases discretas) y agrupación (que consiste en agrupar un conjunto de objetos de tal manera que los miembros del mismo grupo, o *cluster*, sean lo más parecidos posible según determinados criterios).

En los más de 70 años de existencia del campo del AA se han desarrollado multitud de métodos, entre los que podemos destacar las redes neuronales artificiales, las redes bayesianas, los clasificadores estadísticos, los sistemas

basados en reglas, los árboles de decisión o *Support Vector Machines*. Hoy en día, el campo más prolífero y revolucionario de la IA es el denominado Aprendizaje Profundo, o *Deep Learning* (DL) [32]. Este representa una familia de algoritmos de aprendizaje automático que, durante el último decenio, han mejorado drásticamente los resultados proporcionados por las técnicas existentes en tareas como el reconocimiento del habla[33], la traducción automática [34] y la clasificación de imágenes [35], o que han superado directamente la capacidad humana frente a actividades complejas[36].

La capacidad demostrada por estos algoritmos para resolver problemas de todo tipo presupone su potencial para resolver también problemas en la AF, como la estimación del perfil biológico, la localización de landmarks cefalométricos, o la automatización de ciertas tareas implicadas en un proceso de toma de decisiones en el ámbito forense.

2.3. Soft Computing.

A diferencia de los métodos informáticos tradicionales (*hard computing*), las técnicas de *Soft Computing* (también llamadas Inteligencia Computacional) no se basan en soluciones de forma cerrada (por ejemplo, una fórmula que resuelve una ecuación específica), sino principalmente en procedimientos de búsqueda/optimización (heurística) y otras técnicas "aproximadas" capaces de resolver problemas en presencia de datos cualitativos, inciertos, imprecisos e incompletos[37]. Esas técnicas tratan de emular la capacidad de los seres humanos para hacer frente a esas situaciones de manera eficaz. Los principales paradigmas del *soft computing* se analizan a continuación.

Las redes neuronales artificiales (RNA) son enfoques de AA poderosos y flexibles, capaces de aproximar funciones complejas. Son modelos bioinspirados, es decir, se inspiran en las redes de neuronas del cerebro (ver figura 5), aunque no necesariamente buscan imitar su comportamiento.

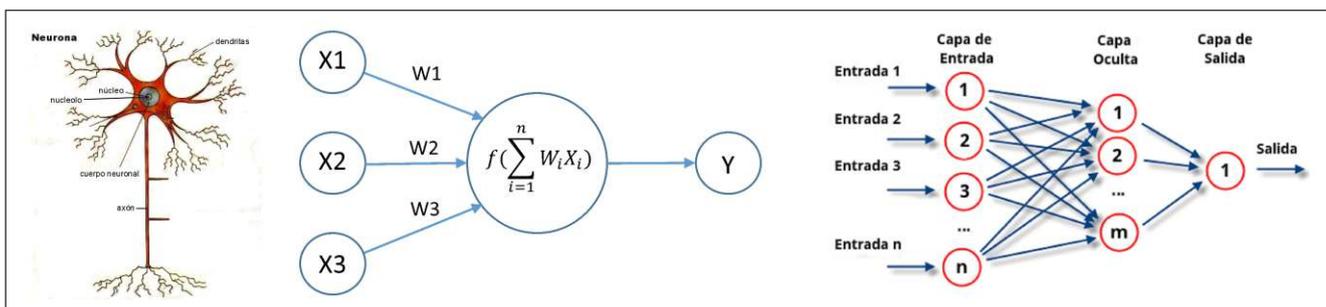


Figura 5: de izquierda a derecha, representación de una neurona biológica (dendritas, núcleo y axón, una neurona artificial (diferentes variables de entrada X, pesos w, asociado a las entradas, y una función de procesamiento de la información f y salida Y) y una RNA (capas de entrada, salida y oculta con diferente número de neuronas artificiales y conectividad entre estas).

Los primeros modelos de RNAs son del año 1943[38], y desde entonces han surgido centenas de modelos y aplicaciones, pero no ha sido hasta el último decenio cuando su uso se ha popularizado gracias al aumento de las capacidades de los ordenadores, lo que ha permitido desarrollar modelos con muchas más capas y neuronas. Son las denominadas redes neuronales profundas, o DL, paradigma de AA del que hemos hablado anteriormente. Dentro de estas redes, cabe destacar

una familia de modelos concretos, las redes neuronales convolucionales (o *ConvNets*), modelos de RNAs profundas que asumen explícitamente que las entradas tienen una topología similar a la de una cuadrícula (por ejemplo, imágenes), y que son especialmente adecuadas para analizar imágenes visuales sin ninguna etapa de preprocesamiento o extracción de características (ver figura 6).

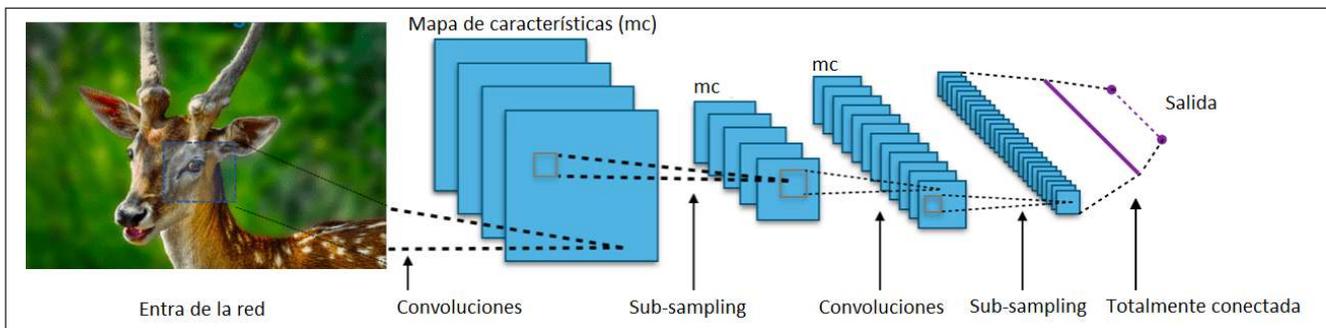


Figura 6: representación simplificada de una *ConvNets* en donde pueden verse múltiples capas que actúan (convolucionan) en cada momento sobre una ventana de $n \times n$ de la imagen de entrada.

Los algoritmos evolutivos (AE) son algoritmos de optimización, búsqueda y aprendizaje inspirados en los procesos de selección natural y evolución genética. Destacan por su capacidad de resolver eficazmente problemas NP-duros y multimodales y no requieren información a priori de la función de error. El modo de funcionamiento de estos algoritmos es muy sencillo, evolucionan poblaciones de individuos que representan soluciones potenciales a problemas. En este proceso de

evolución se seleccionan individuos (padres) en base a un modelo que combina aleatoriedad y adaptación al medio (la cual se mide con una función matemática de lo bueno que es un individuo resolviendo el problema, función de error), se cruzan para generar nuevos descendientes (intercambio de información genética) y estos se mutan (generación de nueva información), con una cierta probabilidad (ver figuras 7 y 8).

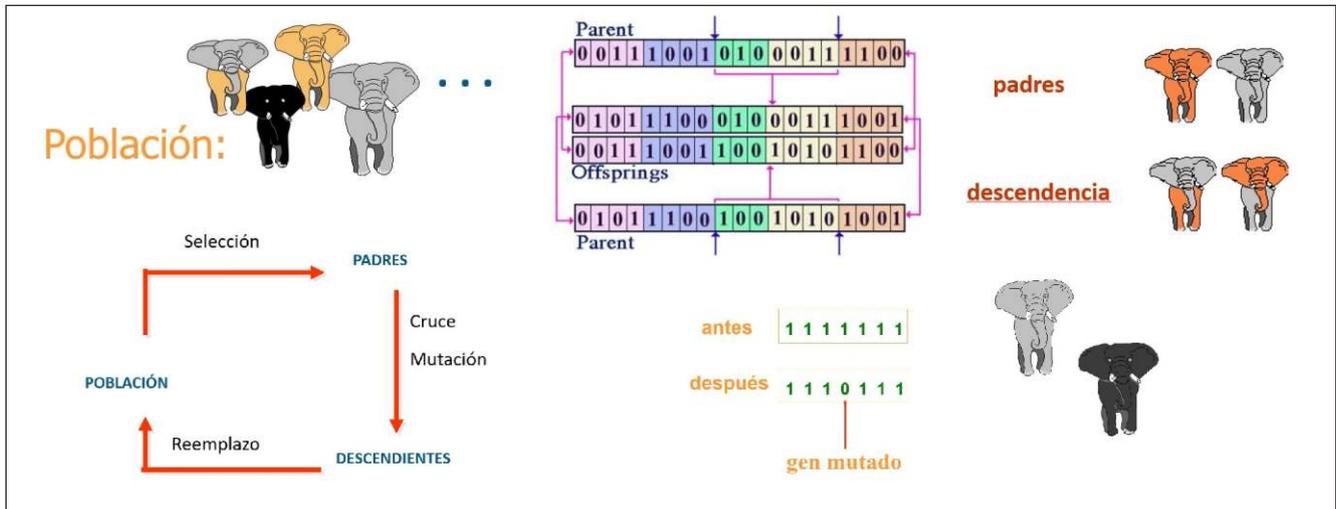


Figura 7: representación del proceso de evolución de un AE, a partir de una población inicial, se seleccionan, cruzan y mutan individuos (cromosomas) para generar una nueva población que reemplaza a la anterior, y así sucesivamente.

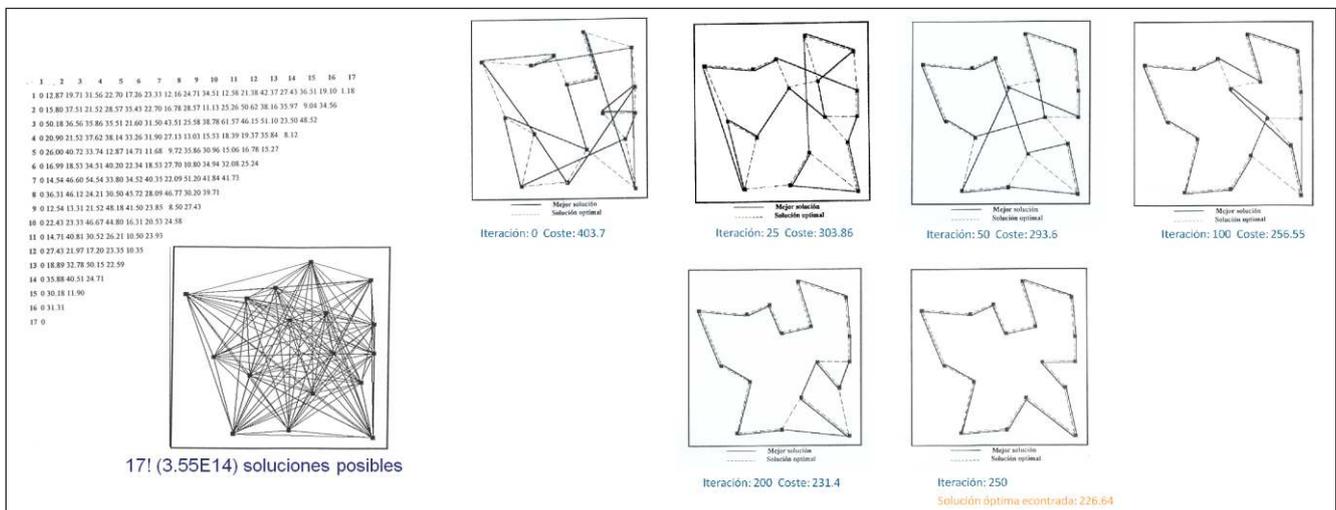


Figura 8: ejemplo de resolución del problema del vendedor ambulante mediante un AE. Dado un conjunto de ciudades (nodos de un gráfico) conectadas por diferentes caminos (bordes o arcos), encontrar el camino del coste mínimo (menor número de Kms) pasando por todas ellas una sola vez y volviendo al punto de partida. Cada individuo del AE, cromosoma, codifica en sus genes el orden en el que se visitan las ciudades. La función de adaptación de cada individuo es número de Kms recorridos para pasar por todas las ciudades. En la izquierda, coste (Kms) de desplazamiento entre cada ciudad/nodo y mapa de conexiones. A la derecha, mejor solución del AE en distintas generaciones (inicial, 25, 50, 100, 200 y 250, donde se encuentra la solución ideal).

La lógica difusa extiende la lógica booleana al dominio continuo [0, 1], donde una expresión difusa no es necesariamente verdadera o falsa (se aceptan varios grados de verdad/pertenencia a una clase). Esto permite procesar conceptos que son esencialmente cualitativos, de acuerdo con una teoría matemática rigurosa. Fue diseñada para representar y razonar sobre conocimiento expresado de forma lingüística o verbal, el cual es generalmente “vago”, “impreciso”, “difuso”, o “borroso”. Nos permite modelar matemáticamente fenómenos (de los que se desconoce una distribución de probabilidad) tan variados como el grado de soporte en la toma de decisiones (por ejemplo, limitado, moderado o fuerte) o como afecta el índice de masa corporal a la fiabilidad de un determinado criterio morfológico (por ejemplo, fiable hasta un índice de masa corporal menor de 25, no fiable a partir de 35, pero con distinto grado de fiabilidad entre ambas medidas).

Por último, el razonamiento probabilístico, que se basa en una estructura de datos denominada red bayesiana (también llamada red probabilística, red causal o red de creencias), que permite tomar un acontecimiento ocurrido y predecir la probabilidad de que cualquiera de varias posibles causas conocidas sea el factor contribuyente.

Tanto las técnicas de Soft Computing entre sí, como estas con el resto de los paradigmas de la IA mencionados en las páginas anterior, suelen combinarse (hibridarse) para abordar la resolución de problemas más complejos. En el caso concreto de la AF, los AE pueden ser de gran utilidad para resolver los problemas de optimización que suelen plantearse en las tareas VA, como la SI o el RI (por ejemplo, en la SCF o en la RC); mientras que la lógica difusa puede ser el paradigma subyacente empleado para aumentar la explicabilidad e interpretabilidad de los sistemas de apoyo a las decisiones asistidas por ordenador que analizan datos e imágenes mediante técnicas de AA y VA.

La siguiente sección recoge las publicaciones en las que se desarrollan enfoques basados en alguna o varias de estas técnicas de IA para abordar algún problema de AF concreto.

3. ESTADO DEL ARTE RELATIVO A LA APLICACIÓN DE ENFOQUES BASADOS EN INTELIGENCIA ARTIFICIAL A TÉCNICAS DE ANTROPOLOGÍA FORENSE.

Durante las últimas dos décadas, técnicas como el ADN o las huellas dactilares se han empleado en muchos escenarios de identificación. Sin embargo, la aplicación de estos métodos fracasa cuando no se dispone de suficiente información AM o PM debido a la falta de datos (segunda muestra de ADN) o al

estado de conservación del cadáver. Aunque el esqueleto suele sobrevivir a procesos de descomposición tanto naturales como no naturales (fuego, sal, agua, etc.), el tejido blando se degrada progresivamente y se pierde. Este es el caso más frecuente en los escenarios de identificación de víctimas en desastres de masas. De hecho, la experiencia de varios profesionales en estos escenarios indica la menor eficacia del análisis de ADN (alrededor del 3% de las identificaciones) y la dactiloscopia (15-25%) frente a las técnicas de identificación forense basada en el esqueleto (70-80%)[39], que son la práctica diaria de los expertos en FA. Estas técnicas implican la evaluación de restos óseos humanos para determinar la identidad y causa de la muerte de la persona fallecida [40]. Los métodos empleados por antropólogos forenses, odontólogos y patólogos son cruciales en la recopilación de datos de AM y PB, y aún más importante, representan la última oportunidad de la víctima para la identificación mediante técnicas como la RC (incluida la comparación dental), la SCF o la aproximación facial.

Recientemente han surgido desafíos en la identificación de los seres vivos en los que la precisión y la robustez son fundamentales. Los métodos de imagen facial —[41] son indispensables en muchas circunstancias, ya que proporcionan medios para la investigación policial y pruebas valiosas en las instancias judiciales. Con el creciente número de cámaras de vigilancia, hay una mayor demanda de sistemas de reconocimiento facial. Sin embargo, las diferencias en las condiciones fotográficas (por ejemplo, la iluminación, el ángulo de la cámara y el objetivo), la posición de la cabeza, la expresión facial y la resolución entre las imágenes de referencia y las de prueba son comunes y complican enormemente todo el procedimiento [42]. Otro problema reciente relacionado con la identificación de personas vivas es que el aumento de los movimientos migratorios transfronterizos también ha dado lugar a un aumento del número de informes de expertos sobre la evaluación de la edad de los jóvenes y adultos jóvenes en numerosos países. Estos menores, que por lo general llevan documentos de identidad inválidos, pueden dar información inexacta sobre su edad, lo que es de suma importancia en los procedimientos judiciales [43]. Los límites de edad legalmente relevantes están entre los 12 y los 21 años, estableciendo la edad legal en 18 años en la mayoría de los países[44]. Según las recomendaciones del Grupo de Estudio sobre el Diagnóstico Forense de la Edad [45], para la evaluación de la edad debe utilizarse un examen físico, un examen de rayos X de la mano y un examen dental (incluido el ortopantomograma). Si se ha completado el desarrollo esquelético de la mano, se debe realizar una tomografía computarizada adicional de la clavícula medial [43, 46]. Es importante señalar que, para evitar la violación de los derechos del menor, al comienzo del procedimiento de diagnóstico se realiza una encuesta de anamnesis (en la que se pregunta sobre las enfermedades que influyen en el

crecimiento o los medicamentos) y un examen físico (altura y peso corporal y signos reconocibles de madurez sexual). El principal objetivo de este examen físico inicial es determinar los trastornos del crecimiento y el desarrollo. La conclusión desde la edad biológica (edad esquelética, edad dental) hasta la edad cronológica sólo está justificada para las personas que no llaman la atención a este respecto[43].

En las siguientes subsecciones se resumen los enfoques basados en IA para resolver problemas de la AF más relevantes y novedosos. Hemos incluido en este grupo los métodos tradicionales basados en el análisis directo del esqueleto como el PB, la SCF, la RC, la aproximación facial o la descripción de traumatismos y patologías, junto con métodos relacionados con la AF como son la estimación del Sexo y la Edad en Imágenes Faciales.

3.1. Visión general de los enfoques de inteligencia artificial existentes para la estimación del perfil biológico.

La estimación del PB ha sido estudiada durante más de 300 años, y hoy en día juega un papel crucial en la reducción del rango de posibles coincidencias durante el proceso de identificación. Este paso es previo a la aplicación de cualquier técnica de identificación e implica el estudio de los restos óseos con el objetivo de encontrar rasgos característicos que apoyen la determinación de la identidad del individuo. Estos rasgos incluyen: (1) estimación de la edad, mediante el análisis de la apariencia, los patrones de osificación, el desarrollo dental y la morfología de las costillas y la pelvis [47]; (2) estimación del sexo, basada en estudios morfométricos y morfológicos, es decir forma del cráneo y el hueso de la cadera [48]; (3) la ascendencia, mediante el análisis del cráneo tanto morfológico como métrico[49]; y (4) la estatura, mediante la medición de la longitud de los huesos largos (fémur) y la extrapolación de las mediciones para estimar la altura en vida.

La estimación del PB a partir de restos óseos es una tarea compleja, especialmente cuando se estima la edad de un sujeto, ya que hay diferentes métodos a aplicar dependiendo de la etapa de desarrollo del individuo. La mayoría de los rasgos individualizantes se pueden identificar en varias estructuras como las suturas del cráneo, la sínfisis del pubis y la extremidad de las costillas externas. Las variaciones en la morfología de los huesos son bien conocidas, pero no se producen al mismo tiempo en diferentes individuos, ya que están sujetos a condiciones genéticas y ambientales. Por todas estas razones, el PB se evalúa como un procedimiento secuencial que comienza con la estimación de la ascendencia, seguida a su vez por el sexo biológico, la edad y, por último, la estatura.

Con el objetivo de reproducir los métodos previamente consolidados en la AF, la imagen radiológica, principalmente la tomografía computarizada (TAC), y en menor medida la resonancia magnética (RM), está ganando popularidad a la hora de estimar el PB, ya que es un enfoque no invasivo (en el sentido de que el experto no necesita tener acceso visual directo al hueso), y permite mejores posibilidades de observación y cálculos métricos [48]. La estimación del PB a partir de imágenes de rayos X se ha abordado desde tres comunidades científicas diferentes con tres propósitos distintos. En medicina clínica, la edad biológica es importante para determinar las enfermedades endocrinológicas en los adolescentes o para planificar de manera óptima el momento de las intervenciones de cirugía ortopédica pediátrica. No obstante, en medicina legal se utiliza para aproximar la edad cronológica desconocida, cuando se determina la edad en casos de investigaciones penales o para procedimientos de solicitud de asilo, en los que faltan documentos de identificación de niños o adolescentes. Por último, los antropólogos físicos y forenses se interesan por determinar el sexo, la edad, la estatura y la ascendencia de cualquier resto humano.

Los mayores esfuerzos internacionales se han centrado en la elaboración de enfoques automáticos precisos y objetivos para evaluar si las personas vivas han alcanzado la edad umbral que significa la mayoría de edad legal. Hay diferentes propuestas que se apoyan en programas o algoritmos informáticos de cara a facilitar el procedimiento de estimación de la edad mediante el uso de métodos de imágenes radiológicas [50]. No obstante, hasta donde sabemos, no existe un enfoque automático para una evaluación forense completa de la estimación de la edad teniendo en cuenta las diferentes etapas del procedimiento y todo el rango de edades posibles. La mayoría de los métodos automáticos publicados funcionan sólo sobre radiografías de sujetos caucásicos menores de 10 años, y unos pocos enfoques tratan con sujetos menores de 18 años. No existen métodos automáticos que realicen la estimación de la edad de los adultos, y sólo unos pocos métodos automáticos de estimación del sexo a partir de los rayos X. Por esta razón, las siguientes subsecciones se centrarán en los principales métodos de estimación del PB (estimación del sexo y estimación de la edad), aplicados principalmente a individuos vivos y abordados mediante técnicas de IA.

3.1.1. Enfoques de inteligencia artificial para la determinación del sexo a partir de imágenes óseas.

La estimación del sexo es un pilar fundamental dentro del PB. Si la estimación del sexo es incorrecta, se pueden producir

una serie de errores en cadena en el resto de parámetros del PB a estimar. Los métodos actuales para la determinación del sexo se basan principalmente en criterios morfométricos o morfológicos [51, 52]. El enfoque morfométrico incluye las mediciones de las manos, los pies y las extremidades correspondientes a los huesos superiores, inferiores y largos. Por otra parte, el enfoque morfológico encuentra su fundamento en el dimorfismo sexual, que se presenta en ciertos grados en la mayoría de los huesos del cuerpo humano. En particular, las partes más comunes y populares del cuerpo humano para la estimación del sexo son el cráneo y la zona de la pelvis [53]. Sin embargo, estos métodos están sujetos al sesgo del analista, requieren un alto grado de pericia, son complejos, llevan mucho tiempo [54] y no siempre son adecuados, principalmente debido a la presencia de un daño significativo (químico y/o físico) en los restos óseos.

Entre los pocos trabajos publicados al respecto, destacan enfoques de AA, ya sea combinados con otros métodos de VA [53, 55, 56], o de manera aislada, mediante técnicas de DL [57, 58]. Las técnicas de DL ofrecen una solución flexible y potente para la estimación del sexo a partir de imágenes de estructuras esqueléticas. Una de las ventajas del DL es que descubren de forma automática y autónoma las características de una imagen que son más útiles para la clasificación [59, 60].

De estos cinco trabajos que aplican técnicas de IA a la estimación del sexo, solo uno se centra en niños [58]. El cuerpo de un adulto es lo suficientemente maduro para señalar pistas significativas que pueden ayudar a distinguir su sexo, pero puede no ser adecuado para la estimación del sexo en los niños.

En [55] los autores utilizaron una combinación de RNAs con un Enjambre de Partículas (una técnica de *soft computing* para optimización) para la estimación del sexo. Utilizaron un conjunto de datos de imágenes de rayos X de la mano izquierda de la población asiática. Su conjunto de datos era pequeño y sus resultados sugieren una precisión diferente para los distintos grupos de edad. Pinto et al. [53] introdujeron una metodología para la cuantificación objetiva de los rasgos sexualmente dimórficos en imágenes del cráneo y la pelvis utilizando la Transformada de *Wavelet*, que es una herramienta matemática de escala múltiple que permite medir las variaciones de forma que están ocultas a diferentes escalas de resolución. Los autores afirmaron que su método se había aplicado con éxito para la estimación del sexo en una muestra piloto de mallas tridimensionales de la morfología supraorbital. Esta información puede ser utilizada por los expertos para mejorar la precisión de la evaluación del BP y para describir las variaciones geográficas y temporales dentro de las poblaciones y entre

ellas. En [56] los autores presentaron un sistema automatizado de detección del canal de Haversian basado en la histomorfología, que utiliza sólo fragmentos de hueso para estimar la edad y el sexo. Dividieron su sistema de detección en dos partes. En la primera parte, analizaron manualmente y observaron las diferencias en los parámetros de la muestra de huesos masculinos y femeninos. En la segunda parte, aplicaron técnicas de procesamiento de imágenes microestructurales para identificar el sexo. Bewes et al. [57] abordaron el problema de la determinación del sexo de los restos óseos adaptando una red profunda muy popular, llamada GoogLeNet [61], y diseñada para analizar imágenes de entrada en 2D. Entrenaron la red con imágenes de 900 cráneos virtualmente reconstruidos a partir de TACs de hospitales. Cuando se probó en imágenes no vistas previamente de cráneos, la red profunda mostró un 95% de precisión en la estimación del sexo.

Por otro lado, solo los autores de [58] se han centrado en la estimación del sexo en los niños. Ellos propusieron una técnica llamada GDCNN (Determinación del Sexo con ConvNets), donde las radiografías de la mano izquierda de los niños entre un amplio rango de edades (que van desde 1 mes a 18 años) son examinadas para estimar el sexo. Para identificar el área de atención (parte de la mano) utilizaron Mapas de Activación de Clases, descubriendo que la parte inferior de la mano alrededor de los carpos (muñeca) es más importante que otros factores para la estimación del sexo del niño. Obtuvieron una precisión del 98%, identificando el género de un niño incluso con la mitad de la parte inferior de la mano, lo cual es impresionante considerando el que esqueleto de los niños no se ha desarrollado completamente.

3.1.2. Enfoques basados en inteligencia artificial para la estimación de la edad a partir de estructuras óseas.

Es bien sabido por los antropólogos forenses que ciertas zonas óseas muestran una secuencia definida de modificación que está estrictamente asociada con la edad. En los años veinte del siglo pasado, Todd y col. [47, 62] propusieron un método de estimación de la edad basado en el análisis de la sínfisis púbica que ha sido ampliamente utilizado en el área. Definió diez categorías de edad, y también proporcionó una descripción muy completa de los aspectos morfológicos de la sínfisis púbica en cada una de ellas. Sin embargo, las descripciones son a veces imprecisas y pueden dar lugar a confusión entre los grupos de edad, especialmente para los nuevos profesionales.

Es bien sabido por los antropólogos forenses que ciertas zonas óseas muestran una secuencia definida de

modificación que está estrictamente asociada con la edad. Las estructuras óseas comúnmente utilizadas para la estimación de la edad en adultos comprenden la sínfisis púbica [47, 62, 63], la cuarta costilla [64, 65], las suturas craneales [66] o los dientes [67]. En los años veinte del siglo pasado, los cambios en la sínfisis púbica se estudiaron por primera vez por Todd y col. [47, 62] en una colección de esqueletos varones de edad conocida. En este trabajo, que ha sido referencia en el área durante más de 70 años, se definieron diez categorías de edad junto con una descripción muy completa de los aspectos morfológicos de la sínfisis púbica en cada una de ellas. A principios de los 90, Brooks y Suchey [63] utilizaron una muestra de 739 varones y 273 mujeres, cuya sínfisis obtuvieron de una muestra moderna bien documentada de individuos autopsiados en el condado de los Ángeles, para establecer 6 fases con sus características (tanto para varones como para mujeres), con la media, la desviación estándar y el rango de edades. Otra de las estructuras utilizadas para la estimación de la edad en adultos es la cuarta costilla. Iscan y col. [64, 65] estudiaron 230 costillas derechas extraídas de individuos de edad conocida, demostrando ser un método fiable y relativamente fácil de aplicar, aunque es dependiente de la población. Otro de los métodos empleados, esta vez en el cráneo, es el cierre de las suturas craneales y su relación con la edad del individuo. Meindl y Lovejoy [68], propusieron un método basado en el estudio de 10 puntos situados en la calota craneal, estableciendo cuatro estadios de cierre. Por último, los dientes son comúnmente utilizados para la estimación de la edad, Lamendin y col. [67] realizaron su trabajo sobre 360 dientes unirradiculares, obtenidos de 208 individuos de edades entre los 22 y los 90 años y valoraron la periodontitis y la transparencia radicular.

Es importante señalar que cuando se trata de estimar la edad en un caso antropológico-forense, se han de tener en cuenta los condicionantes de cada método y sus variantes en cuanto a la población de origen que ha dado lugar al método y las condiciones particulares de cada caso. Además, las descripciones proporcionadas por algunos de estos métodos a veces pueden ser imprecisas y dar lugar a confusión entre los grupos de edad, especialmente para los nuevos profesionales.

La gran mayoría de propuestas basadas en IA utilizan técnicas de *soft computing*, pero podemos distinguir dos enfoques muy distintos. Por un lado, aquellas propuestas que toman como punto de partida un modelo clásico de AF e intentan automatizarlo y/o mejorarlo, como puede ser el caso de [69] y [70], en donde se proponen diferentes enfoques basados en lógica difusa para automatizar los métodos de Tanner y col. (TW2)[71], (TW3)[72] o de Todd [47, 62]. Por otro lado, y de manera mayoritaria, sobre todo en los últimos 5-6 años, enfoques denominados *end-to-end*, es decir

enfoques donde la única entrada del algoritmo es una imagen (radiografía o RM de la mano/muñeca por normal general) y la salida es la edad cronológica que predice el algoritmo. Estos enfoques son solo posibles gracias a la utilización de ConvNets entrenadas con cientos de imágenes cuya edad cronológica precisa es conocida [7378].

Dentro del primer grupo, algunos estudios utilizan la integral difusa para la estimación de la edad en adultos [79, 80]. Estos métodos consideran la aplicación manual de diferentes métodos de evaluación de la edad basados en diferentes huesos: el método de Todd que utiliza la sínfisis púbica y otros dos que utilizan la superficie auricular y el cierre de la sutura craneal. Estas tres estimaciones independientes se combinan mediante una integral difusa para obtener una estimación más precisa. Villar y col. [81], proponen un sistema preliminar de clasificación basado en reglas difusas [82] para la estimación de la edad a partir del estudio de la sínfisis púbica, siguiendo el método de Todd. Su objetivo era simplificar el proceso, detectando las principales características del problema y aprender un sistema compacto para clasificar los nuevos casos en función de dichas características. Generaron un conjunto de datos etiquetados con resultados muy prometedores. A parte de estos dos trabajos basados exclusivamente en técnicas provenientes de la lógica difusa, cabe destacar un último trabajo que hibrida estas últimas con técnicas de aprendizaje supervisado, en concreto, árboles de decisión. En [69], los autores proponen el uso de árboles de decisión difusos para automatizar el método de Tanner y col. (TW2)[71].

También en este primer grupo de propuestas, pero centrados en estimar la edad en subadultos, los enfoques automatizados que reproducen el método TW2 pueden clasificarse principalmente en función de si utilizan el procesamiento de imágenes o técnicas basadas en el conocimiento, y se puede encontrar un examen exhaustivo en [50]. La mayoría de los métodos basados en el procesamiento de imágenes se remontan a la década del 2000. Estos métodos utilizan radiografías de mano de personas vivas como fuente de conocimiento para el entrenamiento de clasificadores. En [69] se propone un clasificador para la evaluación de la madurez del esqueleto basado en computación con palabras. En [83] la propuesta se basa en un clasificador (una RNA) y en la salida de un filtro difuso. En [50], se utiliza un sistema de inferencia difusa para la evaluación de la edad.

Además de la AF, hay varias contribuciones en la estimación de la edad para aplicaciones tanto de medicina clínica como legal. En medicina legal, cuando faltan documentos de identificación de niños o adolescentes, como puede ser el caso en los procedimientos de solicitud de asilo o en las investigaciones penales. La estimación de la maduración

física se utiliza como aproximación para evaluar la edad cronológica desconocida. Los métodos radiológicos establecidos para estimar la edad desconocida en niños y adolescentes se basan en exámenes visuales de la osificación en imágenes de rayos X de la mano [9, 71]. La osificación se observa mejor en la mano debido al gran número de huesos evaluables que son visibles en las imágenes de rayos X, junto con el hecho de que el progreso del envejecimiento no es simultáneo para todos los huesos de la mano. A partir del nivel de osificación evaluado por el radiólogo, los métodos más comunes para la estimación de la maduración física de un individuo son el método GP [84] y el método TW3 [85]. El método GP es el enfoque utilizado por la mayoría de los radiólogos debido a su simplicidad y rapidez. Este método se basa en la comparación entre la imagen de rayos X de la mano y un atlas de mano y muñeca normales a varias edades cronológicas. La radiografía del paciente se compara con la imagen adecuada del atlas. El método TW2 analiza huesos específicos, en lugar de toda la mano como en el método GP. En particular, este método tiene en cuenta un conjunto de regiones de interés específicos divididos en regiones de epífisis/metáfisis y regiones del carpo.

Es en este marco en donde podemos encontrar a la mayoría de los pocos laboratorios con una composición y orientación multidisciplinaria que han desarrollado enfoques *end-to-end* basados en las técnicas de IA más recientes. De esta manera, y mediante el uso principal de *ConvNets*, han demostrado las grandísimas posibilidades de este paradigma para la estimación de la edad ósea.

Un primer grupo de publicaciones se han centrado en la estimación de la edad en sujetos subadultos partiendo de los fundamentos del método de GP en cuanto a la fuerte correlación entre los procesos de osificación de la mano y la edad cronológica [74, 76, 77, 86, 87].

En [86], los autores utilizaron una *ConvNet* sobre un total de 14.036 radiografías clínicas de mano y sus correspondientes informes, obtenidos de dos hospitales infantiles para entrenar y validar el modelo. Para el primer conjunto de pruebas, compuesto por 200 exámenes, se utilizó como norma de referencia la media de las estimaciones de edad ósea del informe clínico y tres revisores humanos adicionales. El rendimiento general del modelo se evaluó comparando el error cuadrático medio (RMS) y la desviación absoluta media (MAD) entre las estimaciones del modelo y las edades óseas estándar de referencia. Se comparó el RMS de un segundo conjunto de pruebas compuesto por 1.377 exámenes provenientes del Atlas Digital de la Mano de GP – [88, 89] con los informes publicados de un modelo automático existente [90]. La diferencia media entre las estimaciones de edad ósea del modelo y las de los revisores

fue de 0 años, con una media de RMS y MAD de 0,63 y 0,50 años, respectivamente. Las estimaciones del modelo, el informe clínico y los tres revisores estuvieron dentro de los límites del 95% de acuerdo. El RMS para el conjunto de datos del Atlas de la Mano Digital fue de 0,73 años, en comparación con los 0,61 años de un modelo informado anteriormente.

En [74], los autores utilizaron una técnica de DL basada en el método GP para desarrollar un sistema de software automático para la determinación de la edad ósea. Utilizando ese programa informático, estimaron la edad ósea a partir de radiografías de la mano izquierda de 200 pacientes (de 3 a 17 años de edad) utilizando la edad ósea de primer rango (sólo el programa informático), la edad ósea asistida por ordenador (dos radiólogos con asistencia de programa informático) y la edad ósea asistida por el atlas GP (dos radiólogos con la ayuda del atlas del GP). La edad ósea de referencia fue determinada por el consenso de dos radiólogos experimentados. Las edades óseas de primer rango determinadas por el sistema de software automático mostraron un 69,5% de tasa de concordancia y correlaciones significativas con la edad ósea de referencia ($r = 0,992$; $p < 0,001$). Las tasas de concordancia también aumentaron con el uso del software automático para ambos revisores (una mejora de alrededor del 8,0% respecto a los resultados guiados por el atlas GP) y el tiempo de evaluación de las imágenes de rayos X requerido por los radiólogos se redujo entre el 18,0% y el 40,0%. Sus resultados sugirieron que el sistema de software automático mostraba de forma fiable estimaciones precisas de la edad ósea, y parecía mejorar la eficiencia al reducir los tiempos de evaluación sin comprometer la precisión del diagnóstico.

En [76], los autores crearon un sistema DL para detectar y segmentar automáticamente la mano y la muñeca. Realizan una evaluación automatizada de la edad ósea con una *ConvNet* entrenada sobre un conjunto de 4.278 radiografías de mujeres y 4.047 de hombres (con una edad cronológica de 5 a 18 años). Las imágenes fueron etiquetadas por edad esquelética, y generan informes radiológicos estructurados con la decisión final de un radiólogo. Este sistema estandariza automáticamente todas las radiografías de mano de diferentes formatos, proveedores y calidad para ser utilizadas como un conjunto de datos de capacitación para la mejora del modelo en el futuro. Su modelo alcanza una precisión del 57,32% y 61,40% para las cohortes femeninas y masculinas, respectivamente. A las radiografías de prueba de las mujeres se les asignó una edad ósea en con un margen de error de 1 año en el 90,39% de los casos y con un margen de error de 2 años en el 98,11% de los casos. En el caso de los hombres, el porcentaje de radiografías a las que se le asignó una edad con un margen de error dentro del año fue de un 94,18%, y del 99,0% para márgenes de error de dos años.

En [87], se presenta una forma de utilizar la DL para la estimación de la edad a partir de las imágenes de rayos X de la mano de un sujeto, empleando un conjunto de puntos de características en la mano. Estos puntos deben definirse para que sirvan de referencia para centrarse en una determinada región que sea informativa en cuanto a los cambios morfológicos inducidos por el desarrollo esquelético. Obtuvieron un error absoluto medio (MAE) de 18,9 meses sobre un conjunto de test de 200 radiografías (el modelo lo entrenaron con otro conjunto distinto formado por 400).

Lo más común en el ámbito del DL es utilizar una RNA profunda diseñada y pre-entrenada (sobre un conjunto de cientos de miles de imágenes a millones) para un problema similar y, tras una fase de entrenamiento sobre imágenes del problema a resolver (en este caso radiografías de la mano y edad cronológica), medir la precisión de la red sobre un conjunto de imágenes de test del problema concreto. Este ha sido el enfoque seguido en los trabajos revisados anteriormente. Sin embargo, Mutasa y col. [77], desarrollaron su propia arquitectura de RNA profunda y la entrenaron de cero sobre 10.289 radiografías de diferente edad esquelética (0 a 15 años, mitad niños y mitad niñas). Sus buenos resultados, con un MAE sobre el conjunto de test de 0,536, apoyan la hipótesis de que las RNAs profundas entrenadas expresamente para esta tarea proporcionan un mejor rendimiento frente a redes pre-entrenadas en problemas relacionados.

La aparición y proliferación de *ConvNets* utilizando radiografías de la mano ha facilitado nuevas aplicaciones a la evaluación de la edad utilizando otros huesos. Spampinato y col. [91] propusieron y probaron varios enfoques de DL. En particular, emplearon varios *ConvNets* pre-entrenados existentes para evaluar automáticamente la edad ósea del esqueleto, basados en el método TW2 y utilizando un conjunto de datos de unas 1.400 imágenes de rayos X. Los resultados mostraron una discrepancia media entre la evaluación manual y la automática de unos 0,8 años. También diseñaron y entrenaron desde cero un *ConvNet* personalizada, que demostró ser la solución más eficaz y robusta para evaluar la edad ósea en todos los grupos étnicos, rangos de edad y género. Además, este fue el primer trabajo de evaluación automatizada de la edad ósea esquelética probado en un conjunto de datos públicos.

A parte de la sínfisis púbica, la apófisis de la cresta ilíaca constituye una excelente estructura ósea para la aplicación de métodos de estimación de la edad en vivos, en particular para determinar los umbrales de edad de 14, 16 y 18 años. Por esta razón, Li y col. [78], desarrollaron un sistema de DL utilizando técnicas de aprendizaje por transferencia para realizar una estimación automática de la edad ósea basada

en 1.875 imágenes radiológicas pélvicas clínicas, particularmente para individuos de entre 10 y 25 años de edad. Puede manejar todos los casos posibles de evaluación automatizada de la edad ósea esquelética, incluso para muestras de individuos de 19, 20 y 21 años de edad. Sin embargo, puede no ser práctico para determinar las edades superiores a 22 años debido al poco cambio en la puntuación media de la osificación. En comparación con el modelo de regresión cúbica existente, su modelo *ConvNet* logra un mejor rendimiento medio (MAE= 0,89 y RMSE= 1,21), pero estos resultados son ligeramente peores a los estudios previos con arquitecturas DL basadas en imágenes de rayos X de la mano izquierda, en las que los valores MAE oscilan entre 0,54 y 0,80 años [74, 77, 86, 91]. Sin embargo, aunque su análisis estadístico indica una alta correlación positiva entre la edad estimada y la real ($r= 0,916$; $p < 0,05$), este número es menos exacto que los métodos de imágenes radiográficas de mano ($r= 0,992$; $p < 0,001$).

Por otra parte, como alternativa a los métodos de estimación de la edad basados en imágenes de rayos X, la investigación en la estimación de la edad mediante RM a ha ganado un enorme interés en los últimos años. El interés en desarrollar métodos automáticos basados en la RM para la estimación de la edad tiene su origen en los problemas de la exposición a las radiaciones ionizantes, la necesidad de definir nuevos sistemas de estadificación específicos de la RM y la influencia subjetiva del examinador [73]. En [70], los autores utilizaron una técnica de AA denominada *Random Forest* para hacer una regresión de la edad cronológica a partir de las características de intensidad de las resonancias de los huesos de la mano de 11 sujetos adolescentes. Un árbol de decisión que excluía la información del metacarpo y la falange de los sujetos de más edad sirvió como estrategia de fusión heurística para la estimación de la edad. En [92], se exploró nuevamente de las técnicas de *Random Forest* para la fusión de información, permitiéndole decidir internamente de qué huesos aprender la edad cronológica de un sujeto. Así, trataron el desarrollo esquelético como un proceso de desarrollo global sin necesidad de esquemas de fusión heurística, como en [70], o de funciones no lineales predefinidas, como en [71] o [90].

Al igual que explicamos anteriormente en el caso de radiografías, la tendencia actual en el área del AA a partir de imágenes ha pasado de ser la de extraer características y usar un clasificador para relacionar esas características con la clase de la imagen, a un enfoque con un único elemento, el uso de DL, capaces de hacer ambas tareas. Siguiendo este nuevo paradigma, los autores de [92] propusieron una arquitectura *ConvNet* para combinar la información sobre la edad de los huesos individuales de forma automática, dejando que la arquitectura aprendiera directamente las características más relevantes para la estimación de la edad.

Más recientemente, en [73] los autores presentan una solución para la estimación automática de la edad a partir de RMs de la mano. Los autores evalúan diferentes métodos de AA, como *Random Forest* y *ConvNets*, con diferentes variantes de la información de la imagen utilizada como entrada para el aprendizaje. Entrenados sobre un conjunto de datos de 328 imágenes de RM, comparan el rendimiento de las diferentes estrategias de entrada y demuestran resultados sin precedentes logrando la precisión más avanzada en comparación con los métodos anteriores basados en RMs. Para estimar la edad biológica, obtienen un MAE de $0,37 \pm 0,51$ (desviación estándar) años para el rango de edad de los sujetos ≤ 18 años, es decir, donde la osificación ósea aún no se ha saturado. Por último, adaptaron su método más eficaz a las imágenes en 2D y lo aplicaron a un conjunto de datos de imágenes de rayos X para validar sus hallazgos, lo que demuestra que su método está en consonancia con los métodos más avanzados desarrollados específicamente para los datos de rayos X.

3.2. Enfoques basados en inteligencia artificial para el análisis de patologías y traumatología

En los últimos años, el éxito del DL para clasificar imágenes ha suscitado un gran interés en su aplicación al análisis de imágenes médicas en varios campos, entre ellos la detección de cáncer de piel[14], la retinopatía diabética[93], las lesiones mamográficas[94] o la detección y clasificación de nódulos pulmonares [95]. Sin embargo, un ejemplo de patología de gran interés, pero también poco frecuente, son las lesiones óseas (Franchi, 2012). Hasta donde sabemos, sólo hay unos pocos trabajos en el campo de la ortopedia relacionados con la aplicación de DL para detectar lesiones o patologías óseas en las imágenes de rayos X.

Olczack y col. [96] extrajeron 256.000 radiografías de muñecas, manos y tobillos del Hospital de Danderyd y las etiquetaron en base a las siguientes 4 clases: la presencia o no de fracturas, lateralidad izquierda o derecha, parte del cuerpo (tobillo, muñeca y mano, esta última también tenía subgrupos de escafoides, pulgar o dedo) y vista de examen (anteroposterior/frontal, lateral, oblicua, y, en el caso de los escafoides, proximal, distal, cubital y radial). Luego, evaluaron la precisión diagnóstica de 5 RNAs profundas. Todas las redes mostraron una precisión de al menos el 90% al identificar la lateralidad, la parte del cuerpo y la vista del examen. La precisión final a la hora de identificar las fracturas se estimó en un 83% para la red de mejor rendimiento. La red se comportó de manera similar a dos cirujanos ortopédicos muy experimentados cuando se les presentaron imágenes con la misma resolución que la red.

Chung y col.[97] evaluaron la capacidad de las técnicas de IA (*ConvNet*) para detectar y clasificar las fracturas proximales de húmero utilizando radiografías anteroposteriores de hombro. El conjunto de datos evaluados estaba compuesto por 1.891 imágenes (1 imagen por persona) de hombros normales ($n = 515$) y 4 tipos de fracturas proximales de húmero (tuberosidad mayor, 346; cuello quirúrgico, 514; 3 partes, 269; 4 partes, 247) clasificadas por 3 especialistas. La precisión de la *ConvNet* se comparó con las opiniones de 28 médicos generales, 11 ortopedistas generales y 19 ortopedistas especializados en el hombro para detectar y clasificar las fracturas proximales de húmero. Sus resultados mostraron una precisión del 96%, para distinguir los hombros normales de las fracturas proximales de húmero, y una precisión del 65-86% para clasificar el tipo de fractura.

En[98], los autores abordan el problema de la clasificación de las lesiones óseas a partir de imágenes de rayos X. Como ocurre en la mayoría de los trabajos basados en DL, es necesario hacer un aumento artificial del conjunto de datos de entrenamiento, pues los pocos cientos o miles de imágenes de las que se disponen suelen ser insuficientes para que la red sea capaz de generalizar adecuadamente. Este aumento de dato suele implicar una tarea sencilla de generación de miles de imágenes nuevas, rotando, haciendo *zoom* o añadiéndole ruido a las imágenes originales. En este caso, los autores proponen un enfoque de aumento de los datos basado en una red adversarial generativa (un tipo de RNA profunda que es capaz de crear imágenes sintéticas con un gran parecido a imágenes reales) consistente en ciclos que sintetiza las lesiones óseas en imágenes sin patología. Plantean la tarea generativa como un problema de traducción de imágenes que optimizan específicamente para los distintos huesos (húmero, tibia y fémur). En los resultados experimentales, confirman que el método descrito mitiga el problema de desbalanceo de clases en la tarea de clasificación binaria de detección de lesiones óseas. Demuestran que los conjuntos de entrenamiento aumentados permiten el entrenamiento de clasificadores que logran una mayor precisión. Además, demuestran la viabilidad de transferir el aprendizaje y aplicar un modelo generativo que fue entrenado en una parte del cuerpo a otra. Los resultados de su mejor modelo en cuanto a sensibilidad y especificidad son de 0,84 y 0,80 en el caso del húmero, 0,75 y 0,46 en el caso de la tibia y de 0,66 y 0,67 en el caso del fémur.

3.3. Enfoques basados en inteligencia artificial para la Radiología Comparativa.

La RC implica la comparación directa de radiografías AM, adquiridas con fines clínicos, con radiografías PM, adquiridas específicamente para la identificación. Ambas

radiografías se evalúan visualmente para comprobar la similitud de las formas y densidades óseas, a fin de determinar si representan a un mismo sujeto[99]. Los huesos más comunes y valiosos para individualizar son los dientes, los huesos frontales del cráneo, las vértebras y las clavículas, aunque la identificación basada en los dientes es la técnica más empleada y discriminatoria. Además, las imágenes AM y PM más comúnmente empleadas con la técnica de RC [100102] incluyen radiografías, TACs y modelos de superficie tridimensionales.

Los enfoques metodológicos para realizar la identificación basada en RC se dividen en tres grupos según la dimensionalidad de los datos empleados: 2D-2D (radiografía-radiografía), 2D-3D (radiografía-TAC o modelos 3D de superficie) y 3D-3D (TAC-TAC o modelo 3D). Cuanto mayor sea la dimensionalidad, mayor será la precisión y la solidez de los métodos. Dentro de cada uno de estos grupos, los métodos pueden clasificarse además en enfoques manuales y enfoques semiautomáticos. En esta sección nos centraremos en los métodos semiautomáticos, aquellos en los que algunas tareas del proceso de identificación se automatizan mediante algunas de las técnicas basadas en la IA ya mencionadas.

3.3.1. Enfoques 2D-2D para Radiografía Comparativa.

Dentro de la RC, la comparación de radiografías AM y PM es el enfoque más extendido. A fin de destacar la aplicabilidad de la identificación forense basada en la RC, es importante señalar que las imágenes de rayos X representan la modalidad de imágenes médicas más comúnmente empleada[103]. En particular, las radiografías de tórax (CXR) son el examen radiológico más comúnmente realizado en todo el mundo [104] porque pueden producir imágenes del corazón, los pulmones, las vías respiratorias, la sangre, los vasos, la columna vertebral y el tórax [105], así como por su capacidad de diagnóstico y potencial de tratamiento [104, 106]. Para cuantificar la importancia del análisis de CXR, es importante observar que el Servicio Nacional de Salud del Reino Unido realizó 2,02 millones de CXR en 2015/16[107], y que sólo en los Estados Unidos se adquieren anualmente 150 millones de CXR [13].

Existen varios trabajos que comparan de forma semiautomática diferentes estructuras esqueléticas entre radiografías AM y PM. Estas estructuras esqueléticas incluyen los senos frontales [108, 109], la bóveda craneal [110] y los dientes [111, 112]. Estos métodos se basan en la comparación de las siluetas de las estructuras esqueléticas mediante técnicas de morfometría geométrica. En concreto, utilizan la elíptica de Fourier "[113] para comparar automáticamente las siluetas AM y PM obteniendo una lista

de las coincidencias PM más probables para cada AM.

La segmentación de las estructuras esqueléticas (marcado de la silueta) en las radiografías AM y PM es necesaria en todos estos métodos y se realiza de forma manual. No obstante, existen algunos enfoques computacionales que automatizan parte de este proceso mediante métodos ad-hoc basados en reglas, como el algoritmo automático de segmentación de los senos frontales presentado en[114].

Otros enfoques no requieren la etapa de segmentación de la imagen, ya que se realizan a través de la comparación directa de las intensidades, como es el caso del sistema CADI [115] para la comparación de vértebras. Sin embargo, este último enfoque se ve afectado por el tiempo transcurrido entre la adquisición de las radiografías AM y PM, y el consiguiente cambio en las intensidades de las estructuras esqueléticas. El sistema CADI reduce su impacto mediante la selección manual de una región de interés alrededor de cada vértebra, la ecualización de los píxeles dentro de estas áreas (por ejemplo, con un filtro de ecualización de histograma) y, por último, la comparación de las vértebras AM y PM utilizando la métrica de similitud Jaccard.

3.3.2. Enfoques 3D-2D para Radiografía Comparativa.

En el enfoque manual, la metodología de comparación requiere la adquisición de radiografías PM tratando de reproducir los parámetros de adquisición de la radiografía AM, ya sea directamente sobre el cadáver con un dispositivo de adquisición de radiografías [116], o simuladas a partir de un TAC [101, 117, 118]. La adquisición de estas radiografías es una tarea subjetiva que requiere mucho tiempo y es propensa a errores.

Sin embargo, hay sólo unos pocos enfoques automáticos para la comparación de radiografías AM e imágenes 3D [99, 119, 120]. Estos enfoques se basan en el uso de escáneres láser para la adquisición de modelos 3D de la superficie de las estructuras esqueléticas de interés. En particular, las clavículas en [99, 120], y las rótulas en [119]. En ambos casos siguen un procedimiento en el que se obtiene un conjunto de imágenes 2D a partir de proyecciones de estos modelos de superficie 3D mediante la rotación de este en unos ángulos predeterminados. Estas proyecciones 2D sólo contienen la silueta de la estructura esquelética objetivo. Finalmente, este conjunto de proyecciones PM se compara automáticamente con la silueta segmentada manualmente de la estructura esquelética en las radiografías AM utilizando descriptores de la Elíptica de Fourier. La mayor limitación de estos métodos radica en el conjunto de proyecciones 2D predefinidas, y la presunción de que los parámetros que modulan las distorsiones de perspectiva son conocidas.

Por otra parte, [121], desarrollaron un método de RI mediante AE capaz de resolver con éxito este problema. Los datos AM son radiografías clínicas (imágenes 2D) de un hueso determinado que deben compararse con el hueso PM real (un modelo 3D). Los prometedores resultados obtenidos los llevaron a diseñar recientemente una metodología para automatizar totalmente el proceso de identificación mediante RC (véase la figura 9). Este método se probó con TACS y radiografías simuladas en senos frontales, clavículas y rótulas obteniendo una alta precisión. El promedio de error de superposición en los casos positivos sintéticos es inferior al 2% de los píxeles de la silueta según la métrica DICE. Los peores resultados los obtuvieron con las rótulas debido a la menor singularidad de la región visible con un error promedio de 0,761 mm y 1,5% de píxeles mal superpuestos según las métricas mRPD y DICE, respectivamente. El error medio de registrado para la clavícula fue de 0,055 mm y 0,2% de píxeles mal superpuestos. Los mejores resultados se obtuvieron con los senos frontales, con un error promedio de 0,028 mm y 0,2% de píxeles mal superpuestos. Además, llevaron a cabo dos estudios para medir la capacidad discriminatoria del método de RI desarrollado (sin ningún tipo de análisis posterior, solo basándose en el error de solapamiento). En el primer experimento, en el que se usaron 10 radiografías simuladas contra 10 TACs de cabeza, los casos positivos siempre se clasificaron en primer lugar en el ranking resultante de ordenar cada proyección AM contra todos los modelos 3D PM según la métrica del DICE. Sin embargo, en un segundo experimento usando 10 radiografías reales, el porcentaje de casos positivos en la primera posición del ranking bajó a un 58%. Esto se debe principalmente a la baja visibilidad de los senos frontales en cinco de las diez radiografías AM que no obtienen la primera posición en el

ranking. Al respecto, los autores concluyen que un protocolo para la segmentación de la región de hueso/cavidad y oclusión podría ayudar a mejorar los resultados de los casos con baja visibilidad.

Además de la dependencia de un buen marcado de los contornos, el método mostró los siguientes inconvenientes: 1) ninguna de las proyecciones proyectivas consideradas reproducía la distorsión de perspectiva de las radiografías en las que el generador de rayos X no era perpendicular al receptor de la imagen (por ejemplo, radiografías de senos frontales obtenidas mediante el protocolo Waters [122]); 2) la robustez del AE empleado, *Diferential Evolution*, especialmente con clavículas y rótulas, que en algunas ejecuciones da como resultado malas superposiciones debido a su naturaleza estocástica y al espacio de búsqueda altamente multimodal aboradado; y 3) la gran cantidad de tiempo necesario para obtener una superposición (1.800 segundos en promedio). Estos tres inconvenientes han sido abordados recientemente por los autores, que presentaron un nuevo algoritmo en el último congreso de la AEAOF[123].

Por último, el mismo grupo de investigadores abordó el problema de la segmentación de múltiples órganos (corazones, pulmones y clavículas) en imágenes de rayos X de tórax utilizando ConvNets [27]. En este ámbito han propuesto varias arquitecturas profundas, logrando situar su algoritmo como el estado del arte en segmentación de clavículas en radiografías de tórax, con errores de 0,0884, 0,939 y 18,022 para el índice Jaccard, el índice de similitud DICE y las métricas de distancia de Hausdorff, respectivamente. La precisión de este enfoque automático es muy similar a la de expertos radiólogos.

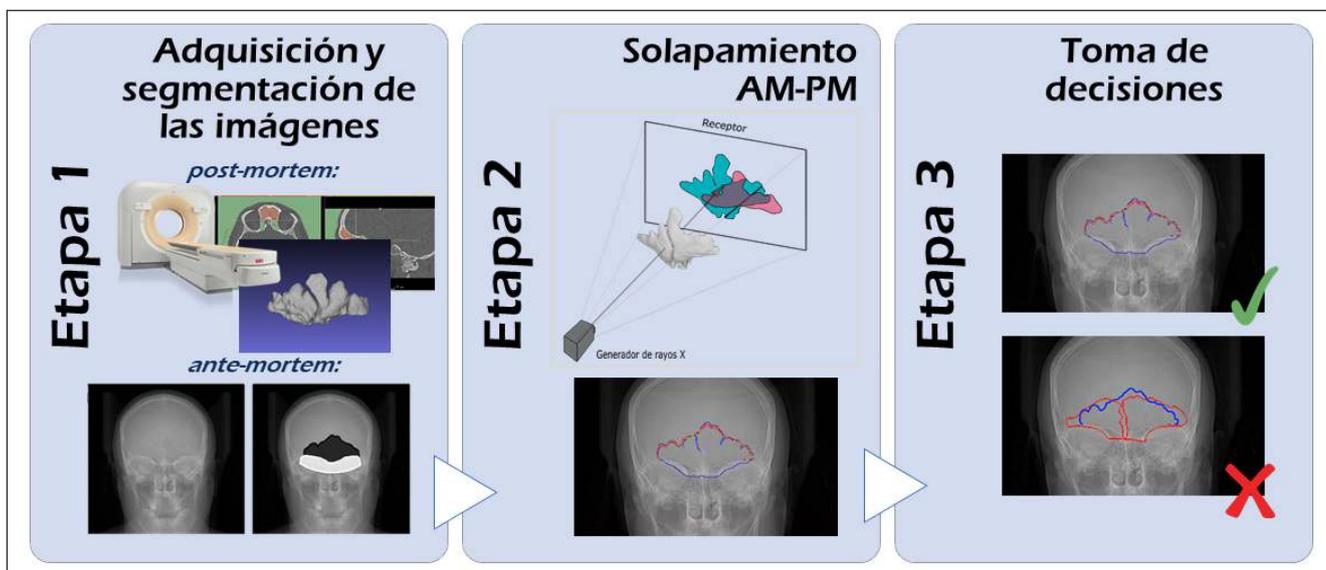


Figura 9: División del proceso de identificación forense mediante RC en tres etapas [121].

3.3.3. Enfoques 3D-3D para Radiografía Comparativa.

El enfoque de comparación TAC-TAC es el más preciso y fiable y no tiene ninguna de las limitaciones anteriores ya que las formas 3D pueden ser comparadas directamente [124127]. Así pues, cuando se dispone de TACs AM y PM, este enfoque se recomienda sobre los dos últimos debido a su mayor fiabilidad y potencial forense [128].

En los enfoques manuales (aunque apoyados en software), la identificación se realiza a través de la comparación visual de sus formas 3D, evitando oclusiones o distorsiones de perspectiva, o a través de mediciones antropológicas, donde las distancias pueden ser comparadas directamente ya que los TACs mantienen las unidades físicas originales. De hecho, en [129] se estudiaron 15 medidas en TACs de cabeza de 119 cadáveres coreanos. Se observó que el volumen total de los senos frontales, algunas características no métricas y el índice de asimetría bilateral en los hombres diferían significativamente de los de las mujeres ($p < 0,05$). Con una combinación de estas medidas fueron capaces de identificar un 98% de los individuos.

Se han propuesto algunos enfoques computarizados para la comparación de los datos tridimensionales AM y PM de diferentes estructuras esqueléticas como dientes [130, 131], senos frontales [132], o vértebras lumbares [133]. La segmentación de las estructuras esqueléticas tridimensionales tanto en los TAC AM como PM (aunque los datos PM podrían adquirirse alternativamente con un escáner láser 3D), su registro (superposición) automático y la medición de la calidad de la coincidencia son necesarios para la aplicación de estos métodos. Sin embargo, la disponibilidad de datos AM tridimensionales (como los TACs) es escasa en comparación con el número de radiografías AM disponibles, lo que reduce considerablemente su aplicabilidad.

En [132], treinta sujetos que se sometieron a dos TACs maxilofaciales (intervalo: 1 mes a 5 años) fueron seleccionados de una base de datos del hospital. Los senos frontales se segmentaron automáticamente mediante el software de código abierto ITK-SNAP y los modelos tridimensionales pertenecientes al mismo paciente se superpusieron automáticamente de acuerdo con la menor diferencia punto a punto entre las dos superficies. Doscientos pacientes fueron seleccionados al azar de la misma base de datos y sometidos al mismo procedimiento para realizar 200 superposiciones de senos frontales pertenecientes a individuos diferentes, divididos por igual entre hombres y mujeres (casos negativos). Se evaluaron las diferencias estadísticamente significativas del error cuadrático medio (RMS) punto a punto entre el grupo de casos positivos, por un lado, y el de negativos por otro, así como las posibles diferencias según el sexo. En el grupo de

casos positivos, el RMS osciló entre 0,07 y 0,96 mm (RMS medio $0,35 \pm 0,23$ mm), mientras que, en el grupo de casos negativos, osciló entre 0,96 y 10,29 mm (RMS medio $2,59 \pm 1,79$ mm), con una diferencia estadísticamente significativa ($p < 0,0001$).

En [133], el objetivo era documentar la eficacia de la comparación parte a parte de los modelos 3D de la columna lumbar obtenidos mediante TACs en la identificación forense. Este estudio pretendía proporcionar una nueva técnica de identificación positiva cuantificable (a través de una coincidencia porcentual) que cumpliera los requisitos explícitos de la sentencia Daubert y los retos establecidos en el informe de la NAS de 2009. Se compararon los modelos AM y PM simulado PM de las vértebras lumbares (L1-L5) de 30 individuos mediante análisis de comparación de partes. Un umbral de $\pm 0,5$ mm, con al menos un 90% de coincidencia, se consideró una identificación positiva. Usando este umbral, los resultados de la comparación de partes tuvieron una tasa de identificación perfecta sin falsos positivos ni negativos. Se generó una curva ROC con una puntuación de 1, lo que significa una sensibilidad y especificidad "perfectas", con un valor de corte de 65,5%. En promedio, las identificaciones positivas tuvieron una coincidencia del 94,7% dentro del umbral establecido, mientras que las identificaciones negativas tuvieron un promedio del 21,4%. Al observar el impacto de los diferentes componentes del perfil biológico, la edad y el sexo del individuo desconocido desempeñaron un papel mínimo en el porcentaje de coincidencia tanto para una identificación positiva como para una negativa.

3.4. Enfoques basados en inteligencia artificial para la Superposición Craneofacial.

La SCF es probablemente el método de identificación más desafiante de todos los basados en el estudio del esqueleto [134, 135]. Implica la superposición de una imagen de un cráneo con una serie de imágenes de la cara AM de un individuo y el análisis de su correspondencia morfológica. Este proceso de solapamiento cráneo-cara (SCC) se realiza normalmente mediante los correspondientes puntos de referencia anatómicos (antropométricos) situados en el cráneo (craneométricos) y en la cara (cefalométricos). Así, a diferencia de la RC, se comparan dos objetos de distinta naturaleza (una cara y un cráneo). LA SCF se ha utilizado durante un siglo, pero no es una técnica madura y plenamente aceptada debido a la ausencia de enfoques científicos sólidos, estudios de fiabilidad significativos y normas internacionales. Por otra parte, esta técnica se emplea ampliamente en los países en desarrollo porque su aplicación es poco costosa y los únicos datos AM requeridos son una o más fotografías del rostro.

En los estudios exhaustivos más recientes del campo de la SCF se diferencian tres etapas consecutivas (véase la figura 10) –[23, 136]: 1) la adquisición y el procesamiento de los materiales, es decir, el cráneo (o modelo 3D del cráneo) y las fotografías faciales AM con la correspondiente ubicación de landmarks antropométricos en ambos; 2) el proceso de SCC, que se ocupa de lograr el mejor solapamiento posible del cráneo sobre una sola fotografía AM de la persona desaparecida. Este proceso se repite para cada fotografía disponible, obteniéndose diferentes SCCs; 3) la toma de

decisiones que evalúan el grado de apoyo de que cráneo y cara correspondan a la misma persona o no (exclusión) basándose en los anteriores SCCs. En esta decisión influye la correlación morfológica entre el cráneo y la cara, la coincidencia entre los puntos antropométricos correspondientes considerando el grosor del tejido blando existente entre estos y la coherencia entre las asimetrías. Estos criterios pueden variar según la región anatómica y la pose de la cara en la fotografía[137].

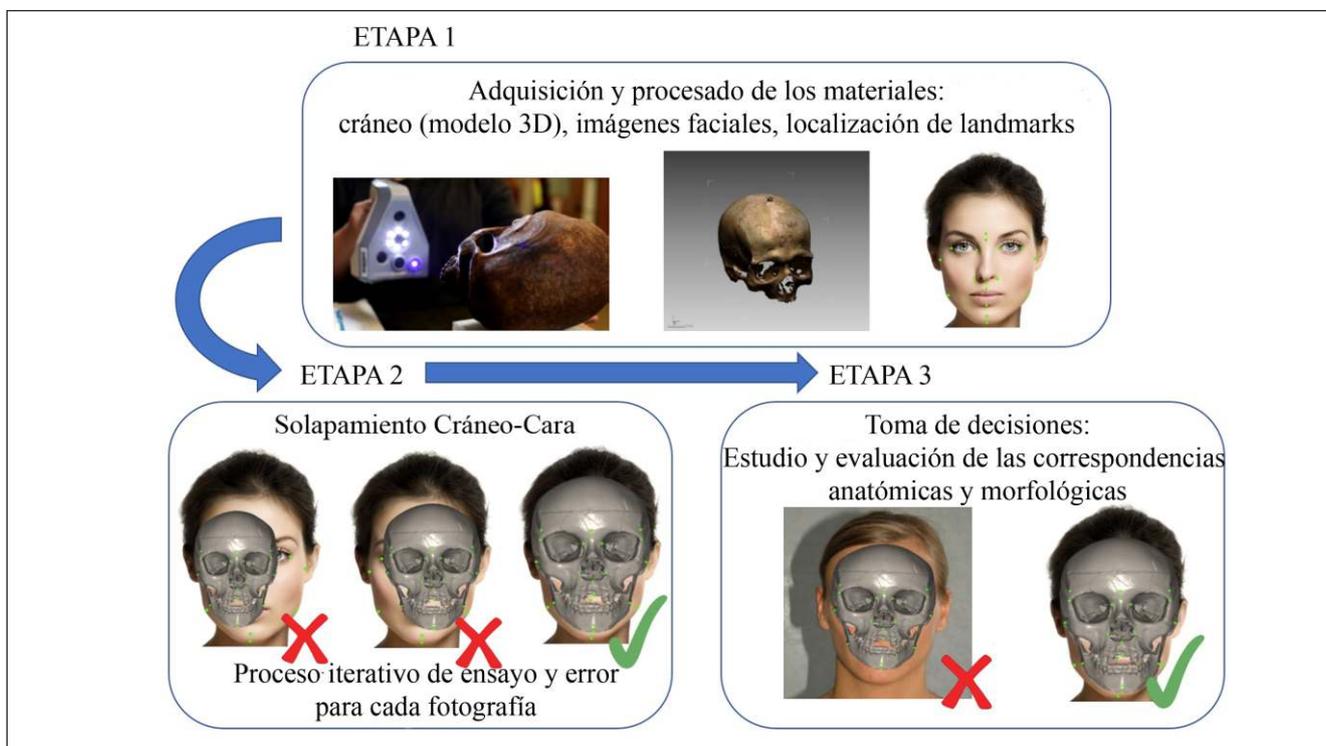


Figura 10: etapas en el proceso de identificación forense mediante SCF.

En las siguientes subsecciones se examinan los principales enfoques de SCF asistida por ordenador existentes para cada una de las tres etapas, haciendo hincapié en los métodos automáticos desarrollados en los últimos diez años.

3.4.1. Adquisición y procesamiento de los materiales.

Los sistemas computarizados desarrollados para la primera etapa de la SCF están relacionados con los procedimientos de mejora de la calidad de las imágenes de la cara y del cráneo. Mientras que el formato de los datos AM se limitan a fotografías (digitales o impresas), las posibilidades en cuanto

al cráneo van desde fotografías a modelos 3D (lo más frecuente en la actualidad), pasando por la grabación del mismo mediante una videocámara. Con el uso de dispositivos de escaneo, como los escáneres láser, el antropólogo forense puede obtener un modelo 3D del cráneo con una precisión inferior a un milímetro en un tiempo razonable [138]. El uso de un modelo 3D en lugar de una imagen 2D del cráneo es más recomendable, porque es una representación más precisa y completa del cráneo. De hecho, el uso de modelos 3D en antropología es ya hoy en día algo bastante común y que ha demostrado ser un soporte más informativo "[139]. En el campo biomédico, las imágenes de exploración por TAC son los datos de partida para reconstruir el cráneo[140, 141].

Desde la primera propuesta que utilizaba un modelo tridimensional del cráneo para abordar el problema de SCF [142], el software de reconstrucción de imágenes 3D se convirtió en una necesidad real, dando lugar a una serie de trabajos centrados específicamente en la reconstrucción 3D de un cráneo completo a partir de varias imágenes adquiridas mediante escáneres láser [143,145]. La tecnología actual, es decir, los actuales dispositivos de adquisición en 3D y el software correspondiente, resuelven automáticamente la alineación de las diferentes vistas de adquisición sin necesidad de plataformas giratorias o cualquier otro dispositivo adicional. Relacionado con la adquisición de modelos 3D del cráneo, en [146] los autores propusieron un nuevo algoritmo para abordar el problema de la simplificación de la malla (estructura matemática que almacena la información del modelo 3D), permitiendo generar diferentes modelos considerando dos objetivos en conflicto, la precisión del modelo y la simplicidad (y tamaño en disco).

Por otro lado, esta etapa también engloba las tareas de marcado de landmarks cefalométricos y craneométricos. Un problema muy similar al primero, la localización de landmarks faciales en fotografías, ha sido ampliamente abordado en la comunidad de VA, y más recientemente, con técnicas de DL [147]. Estas últimas son capaces de automatizar totalmente el proceso, con una precisión similar a la de un experto, incluso en fotografías “in the wild”, es decir, fotografías totalmente realistas, sin ningún tipo de restricción en cuanto pose, iluminación, oclusiones, etc. Sin embargo, los conjuntos de landmarks que se han utilizado en estos trabajos no son los mismos que se utilizan en AF, y solo podemos encontrar una única publicación que se centre en la localización automática de landmarks cefalométricos en fotografías [148] (hay varias en radiografías [149, 150]). En [148] los autores usan técnicas de VA y AA, sobre un conjunto de 1000 fotos y el marcado manual de los landmarks en estas por parte de un experto, para detectar un total de 28 landmarks cefalométricos. La metodología propuesta logra un error de distancia media normalizada (en píxeles) de 0,014, similar a la dispersión media entre expertos (0,009). El principal inconveniente de este método es que está diseñado para imágenes frontales que cumplan el estándar ICAO 9303 [151] (el mismo estándar que se usa para la adquisición de fotografías para el DNI). Además, en el caso de los Gonion, Zygion, Nasion y Glabella, el landmark localizado no es el estándar y los autores dan una definición propia.

3.4.2. Solapamiento cráneo-cara.

La automatización de la etapa de solapamiento cráneo-cara ha recabado el interés de un grupo muy reducido de investigadores, destacando entre todos, por el volumen de

publicaciones y el avance que estas representan, el grupo de investigación *Soft Computing and Intelligent Information Systems* (SOCCER) de la Universidad de Granada. Del análisis de la literatura no queda duda que la forma más natural de tratar el problema de SCF es replicar el escenario original de la fotografía AM en la que la persona viva estaba en una determinada pose en algún lugar dentro del campo de visión de la cámara. Por lo tanto, el objetivo de estos métodos asistidos por el ordenador es replicar en un cráneo la pose y el resto de los parámetros de adquisición de una fotografía facial dada. El escenario descrito es muy similar a un problema de Visión por Computador, replicar la pose de un objeto 3D a partir de una foto basada en algunos puntos de referencia. Técnicamente, se nos dan n puntos que tienen posiciones 3D a_1, \dots, a_n y posiciones 2D objetivo b_1, \dots, b_n . El objetivo es encontrar una proyección P que minimice la distancia entre cada par de puntos $P(a_i) - b_i$. Este problema se denomina calibración de la cámara, y es análogo al problema de RIs explicado en la sección 2 de este trabajo [152].

El primer enfoque asistido por computador para la tarea de SCC fue propuesto por [142]. En este trabajo se utilizan un conjunto de landmarks localizados en un modelo 3D del cráneo y sus homólogos en la fotografía para calcular la transformación de perspectiva (rotación, escalado, traslación y campo de visión) utilizando AEs. El siguiente trabajo que se centró en automatizar la tarea de SCC se publicó 10 años más tarde [153]. En este caso el enfoque es totalmente distinto pues por un lado la información que tenía en cuenta eran fotografías tanto del cráneo como de la cara. Por otro lado, hacía uso de una RNA para implementar una evaluación objetiva de la simetría entre dos imágenes 2D casi frontales (imagen del cráneo y la cara) que se guiaba nuevamente por dos conjuntos de landmarks homólogos.

Durante la última década la gran mayoría de propuestas provienen del grupo de investigación SOCCER y se centran en la utilización de AEs y conjuntos difusos [154,160]. Estos enfoques se basan en la superposición de un modelo 3D del cráneo sobre una fotografía facial, minimizando la distancia entre pares de landmarks craneométricos – cefalométricos mientras se modela tanto la imprecisión introducida por la localización de los mismos [161, 162] como la incertidumbre relativa a la cantidad del grosor de tejido blando entre estos [154]. Más recientemente, Valsecchi y col. [163] propusieron un novedoso algoritmo automático de SCC llamado POSEST-SFO. Los trabajos publicados hasta ese momento [154, 159] resolvieron el problema mediante un proceso de optimización iterativo y estocástico, en el que se evalúan múltiples soluciones en cada paso, que finalmente convergen en una solución de alta calidad. A diferencia de estos métodos, el algoritmo POSEST-SFO resuelve un sistema de ecuaciones polinómicas de manera determinística que relacionan las distancias entre los puntos antes y después de la proyección. Este último algoritmo se

probó en un conjunto de datos sintéticos compuesto por 9 CBCT de 9 sujetos diferentes y 60 fotografías simuladas, 30 laterales y 30 frontales (véase la figura 11), de cada uno de ellos, es decir, 540 SCCs. Este método es extraordinariamente rápido, pues proporciona resultados en 78 milisegundos frente a los 2-5 minutos que eran necesarios en los métodos precursores. En el escenario más realista, considerando la media del estudio estadístico de referencia [164] para el grosor del tejido blando y un error uniforme de ± 5 píxeles en la localización de los landmarks faciales, el error medio de cometido por POSEST_SFO fue de 2,0 mm y 3,2 mm en las fotos frontales y laterales, respectivamente. Sin embargo, este algoritmo, contrariamente a publicaciones anteriores [154, 165], no aborda las fuentes de incertidumbre, es decir, la articulación de la mandíbula, la estimación del grosor de los tejidos blandos y el error o dispersión intra- e inter-experto en la localización de landmarks. Queda claro entonces que todavía hay un gran margen de mejora.

Figura 11: Dos escenarios simulados de SCC (usando modelos 3D adquiridos mediante un con-beam CT) que involucran a uno de los sujetos con una pose casi frontal y otra lateral. Las imágenes (a) y (c) muestran la solución ground-truth (GT) mientras que las imágenes (b) y (d) muestran los resultados proporcionados por el algoritmo POSEST-SFO.

3.4.3. Evaluación del solapamiento cráneo-cara y toma de decisiones.

En esta tercera etapa del proceso de SCF (véase la Figura 10), el experto forense tiene que tomar una decisión para determinar el grado de apoyo a que las fotografías de la cara y el modelo 3D del cráneo pertenezcan a la misma persona o no. Esta decisión se toma mediante el análisis de los SCC obtenidos en la etapa anterior sobre los que se analizan varios criterios que evalúan la correspondencia anatómica cráneo-cara. Diferentes autores han definido y clasificado estos criterios en cuatro familias diferentes:

- 1) Análisis de la consistencia de los contornos/curvas morfológicas óseas y faciales
- 2) Evaluación de la consistencia anatómica por relación posicional
- 3) Localización de líneas y comparación para analizar la consistencia anatómica
- 4) Evaluación de la consistencia del grosor de los tejidos blandos entre los landmarks craneales y faciales correspondientes.

Este proceso de evaluación es subjetivo y se basa en las habilidades del experto forense y en la cantidad y calidad de los materiales utilizados. Sin embargo, son pocos los trabajos que abordan la automatización del análisis de las correspondencias craneofaciales en el marco de la identificación por SCF [166, 167]. La mayor parte de la bibliografía existente se publicó hace más de 20 años y consisten en trabajos muy básicos y limitados. Además, no consideran el uso de modelos tridimensionales del cráneo o de técnicas informáticas para realizar el SCC. Además, la técnica empleada para el análisis de la forma implica una interacción manual, proporcionan un valor que no tiene en cuenta la relación espacial real entre el cráneo y la cara, ya que los métodos empleados son invariables a la traslación, la escala y la rotación. Por último, estos sistemas sólo aplican un único grupo de criterios para evaluar la correspondencia craneofacial, los de la primera familia, pues lo que hacen es extraer características matemáticas de los contornos, mediante la Elíptica de Fourier, para su posterior cuantificación y comparación.

Recientemente, en [168170], integrantes del grupo de investigación SOCCER, presentaron un sistema jerárquico para la automatización de la etapa de toma de decisiones, así como algoritmos de visión por computador para evaluar la consistencia anatómica en SCCs. Esto significa que a partir de una serie de SCCs de un mismo individuo el sistema de apoyo a la decisión (DSS, del inglés *Decision Support System*) proporciona al experto forense un valor de salida cuantitativo indicativo de la consistencia morfológica de un determinado problema de SCF. Este valor cuantitativo se basa en la utilización de varios criterios anatómicos relativos a la correspondencia cráneo-cara combinados en tres niveles diferentes por medio de funciones de agregación difusa: evaluación de criterios (nivel 3), evaluación de los SCCs (nivel 2) y evaluación de la SCF (nivel 1). Se modelaron las fuentes de incertidumbre y los grados de confianza que intervienen en el proceso (estado de conservación de los huesos y calidad del modelo 3D, calidad de la imagen, poder de discriminación de cada criterio individual, influencia dI PB) y se tuvieron en cuenta en cada nivel del sistema [169]. Con el fin de estudiar la precisión del sistema DSS, se utilizó sobre el mismo conjunto de datos experimentales de [22]. En ese estudio, se pidió a 26 participantes de 17 instituciones diferentes que se ocuparan de 14 escenarios de identificación, algunos de los cuales implicaban la comparación de múltiples candidatos y cráneos desconocidos. Se abordó un total de 60 problemas de SCF. El porcentaje de aciertos medio de los resultados de los 26 expertos, el de los tres mejores expertos y el del sistema DSS automático se muestran en el cuadro 3-1. Como puede observarse, el DSS diseñado puede considerarse el primer sistema automático para clasificar parejas de fotografías faciales y cráneos desconocidos como casos positivos o negativos con una precisión similar a la del mejor experto

forense [170]. Sin embargo, una de las principales conclusiones de este estudio fue que los resultados de identificación obtenidos usando el DSS podían verse fuertemente influenciados por la mala calidad de algunos de

los SCCs obtenidos también automáticamente (nótese que este estudio es anterior a la publicación del algoritmo de SCC automático POSEST-SFO del que hablamos en la anterior sección) por el método desarrollado en [154].

Método	Decisiones correctas	Ground Truth	Decisión		Decisión (%)	
			P	N	P	N
Media expertos	78.99%	P	100	90	52.63%	47.3%
		N	152	810	15.80%	84.20%
Mejor expert	93.33%	P	8	2	80.00%	20.00%
		N	2	48	4.00%	96.00%
2° mejor experto	88.14%	P	6	3	66.67%	33.33%
		N	4	46	8.00%	92.00%
3° mejor experto	86.21%	P	5	3	62.50%	37.50%
		N	5	45	10.00%	90.00%
Sistema automático	90.00%	P	6	4	60.00%	40.00%
		N	2	48	4.00%	96.00%

Tabla 1: La tabla muestra el valor medio de los resultados de los 26 expertos y los resultados de los tres mejores expertos según el estudio publicado en ". Además, añade una nueva fila que muestra el porcentaje de acierto del DSS automático. En la tabla se detallan, además del porcentaje de decisiones correctas, los falsos positivos y negativos, así como los verdaderos positivos y negativos. La columna Ground Truth contiene es indicativa de los casos verdaderos (o positivos, P) o falsos (o negativos, N) en cada fila.

Por último, en [171, 172] los autores propusieron un modelo de reconocimiento automatizado del cráneo sin necesidad de llevar a cabo un SCC. En este trabajo los autores utilizaron un conjunto de datos formado por 464 fotografías de personas fallecidas y el mismo número de fotografías de sus respectivos cráneos, para aprender un modelo matemático (aprendizaje semi-supervisado y no supervisados) que fuera capaz de relacionar una foto con su respectivo. Los resultados alcanzados posicionan a este enfoque como una alternativa prometedora para poder filtrar un número importante de posibles candidatos usando únicamente una foto frontal del cráneo.

3.4.4. Enfoques 3D-3D asistidos por ordenador para la superposición craneofacial.

En los últimos años, algunos autores han propuesto un enfoque consistente en la comparación de modelos 3D tanto del cráneo como de la cara.

Duan et al. [173] propusieron un novedoso método de identificación basado en una medida de correlación morfológica entre el cráneo 3D y la cara 3D. El mapeado entre el cráneo y la cara se obtiene mediante el análisis de

correlación canónica. A diferencia de las técnicas existentes, este método no necesita la relación exacta entre el cráneo y la cara, y sólo mide la correlación entre ellos. A fin de medir la correlación entre el cráneo y la cara de manera más fiable y mejorar la capacidad de identificación del modelo de análisis de correlación, se adopta una estrategia de fusión de regiones. Los resultados experimentales validan el método propuesto y muestran que el método basado en la región aumenta significativamente la exactitud de la correlación. La tasa de reconocimiento correcta alcanza el 100% utilizando un conjunto de datos de TACs. Sin duda el mayor problema de este enfoque es su escasa aplicabilidad al requerir de un modelo 3D facial AM.

El método propuesto en [174] hace hincapié en el cumplimiento de dos parámetros importantes a la hora de hacer una superposición: (1) el mantenimiento del tamaño real de la imagen de la cara en relación con el tamaño del cráneo, y (2) la orientación del cráneo sobre una base antropológica utilizando puntos de características específicos. El método propuesto comienza por reconstruir el modelo facial tridimensional a partir de una imagen facial 2D dada utilizando un modelo elástico medio (modelo estadístico de una cara en 3D). Siguiendo una metodología de "leave-one-out" estudiaron la precisión del método de generación de un modelo 3D facial a partir de una foto sobre

la base de modelos faciales pública “USF 3D face database”, la cual contiene 102 individuos. El error cuadrático medio de los modelos 3D generados se situó cercano a los 4mm. A continuación, se lleva a cabo una etapa de registrado 3D-3D del modelo facial con el craneal considerando la línea de la mandíbula mediante la utilización un modelo de B-spline. El error de registrado se evalúa, mediante 6 métodos diferentes, para sugerir el grado en que el modelo facial 3D (obtenido a partir de una fotografía) se corresponde con el modelo 3D del cráneo. Con el error de registrado se calcula un índice de precisión entre 0 y 1, resultado de comparar todos los casos entre si y ordenar para cada uno de los casos en que posición queda el caso positivo. Usando un conjunto de 29 modelos 3D de cráneos (CTs) y comparando cada uno de ellos contra 131 fotografías (en realidad contra el modelo 3D que se genera a partir de cada una de estas), el índice de precisión alcanzó un valor medio de 0,8, o lo que es lo mismo, el método propuesto sitúa al candidato correcto en la posición 27 de media. De estos resultados, y a falta de un análisis estadístico más completo, se puede concluir que este método podría servir para filtrar un número indeterminado de casos y nunca para identificar.

3.5. Enfoques basados en inteligencia artificial para la Aproximación Facial.

La aproximación facial forense, también llamada reconstrucción facial, es una técnica de identificación de restos óseos desconocidos, o cadáveres encontrados en investigaciones penales, basada en la estimación de un rostro a partir de un cráneo con el fin de obtener información sobre la identidad de la persona fallecida. En [99] pueden distinguirse cuatro enfoques diferentes para llevar a cabo la estimación facial con diversas técnicas:

- 1) La representación bidimensional (2D) del rostro sobre una fotografía del cráneo[175178].
- 2) Escultura del rostro en arcilla o masilla sobre el cráneo o el moldeado del cráneo[179182].
- 3) Reconstrucción computarizada del rostro utilizando dispositivos de retroalimentación háptica y una exploración tridimensional del cráneo[183187].
- 4) Reconstrucción computarizada del rostro utilizando algoritmos de IA[188195].

Una característica común de todos estos enfoques es su dependencia de los estudios del grosor del tejido blando de la cara.

Dentro del cuarto y último grupo, en los últimos 10 años se han propuesto varios trabajos de automatización de esta técnica, que han dado lugar a métodos de aproximación facial completamente informatizados y en gran parte automatizados. En la mayoría de los casos se han servido de imágenes médicas, TACs, para desarrollar y entrenar los modelos mediante técnicas de AA –[192194, 196, 197].

Vandermulen y col. [194] presentaron un procedimiento totalmente automático para la reconstrucción craneofacial, utilizando una base de datos de TACs de la cabeza. Todas las imágenes de referencia se segmentan automáticamente en volúmenes de cabeza (encerrados por la superficie cutánea externa, la cara) y volúmenes de hueso/cráneo, ambos representados por un mapa de transformación de distancia (sDT). Los sDT del cráneo de referencia se deforman de forma no lineal al sDT del cráneo objetivo y esta misma deformación se aplica a todos los sDT de la cara de referencia. Se propone una combinación lineal de los sDTs de la cara de referencia deformados como la reconstrucción de la cara del sujeto objetivo. Los resultados sobre un conjunto de 20 casos mostraron la viabilidad de este enfoque. En primer lugar, es necesario eliminar de las imágenes el ruido añadido por la presencia de artefactos metálicos (principalmente en los dientes), ya que es posible que distorsionen las reconstrucciones hasta un grado inaceptable. En segundo lugar, el procedimiento de deformación debe examinarse más cuidadosamente, prestando atención, por una parte, a que la referencia se ajuste mejor al cráneo objetivo y, por otra, a que la extrapolación de la deformación se haga sin problemas. En tercer lugar, es necesario explorar otras combinaciones lineales además de la mera media. Por último, es necesario llevar a cabo un marco de validación cuantitativo más amplio para las reconstrucciones.

Tu y col. [198] propusieron la automatización del proceso de reconstrucción mediante un modelo generativo tridimensional de la cara construido a partir de una base de datos de TACs. La reconstrucción puede limitarse sobre la base de conocimientos previos como la edad y/o el peso. Para validar la precisión del método proponen métodos geométricos con los que comparan las reconstrucciones obtenidas con una galería de imágenes faciales. En primer lugar, se utilizan *Active Shape Models* (una técnica de visión por computador que consiste en definir mediante un conjunto de puntos la forma media y la variabilidad estadística de un objeto determinado) para detectar automáticamente un conjunto de landmarks faciales en cada imagen de la galería. Estos landmarks se asocian con puntos 3D en la reconstrucción. La comparación directa de la reconstrucción es problemática ya que, en general, la geometría de la cámara utilizada para la captura de imágenes es desconocida y existen incertidumbres asociadas a los procesos de reconstrucción y detección de landmarks. El

primer método de comparación utiliza una optimización con restricciones para determinar la proyección óptima de la reconstrucción 3D sobre la imagen de la galería. Los residuos (errores entre pares de puntos) son entonces analizados resultando un ranking con de todas las fotos de la galería para cada reconstrucción. El segundo método utiliza la técnica estadística de *boosting* para aprender qué puntos son fiables y discriminantes. Esto resulta en un clasificador binario que clasifica entre emparejamiento positivo y negativo.

Claes y col. [189] describieron la metodología común del software moderno de aproximación facial:

- Primero, es necesario que un experto examine el cráneo desconocido para determinar el PB.
- A continuación, se produce una réplica virtual del cráneo en base a los parámetros de modelado de la etapa anterior.
- Una plantilla (modelo estadístico) craneofacial que codifica la información de la cara, del cráneo y del tejido blando se deriva de una base de datos de cabeza.
- Esta plantilla se deforma (teniendo en cuenta una serie de restricciones en cuanto las transformaciones geométricas admisibles) para adaptarse al cráneo objetivo según la "proximidad" entre los cráneos.
- Como resultado de la anterior etapa, la plantilla de la cara se deforma en base a las mismas transformaciones geométricas que dieron lugar a la deformación de la plantilla del cráneo, dando como resultado un modelo facial de la geometría de la cara reconstruida.
- Finalmente, solo queda añadir a este modelo facial una textura de piel y vello.

Guyomarc'h y col. [192] desarrollaron un método computarizado para estimar la forma facial basado en 500 TACs de personas francesas. El grosor del tejido blando en cada landmark craneométrico se estima a partir de un modelo de regresión que tiene en cuenta edad, sexo, corpulencia y la craneometría (coordenadas x, y, z de todos los landmarks craneales). El grosor estimado se proyecta, a partir de cada landmark craneométrico, utilizando planos de referencia para obtener el correspondiente landmark facial. La posición y la forma de los ojos, la nariz, la boca y las orejas se infieren a partir de puntos de referencia craneales mediante morfometría geométrica. Los 100 landmarks faciales estimados se utilizan luego para deformar un rostro genérico (modelo estadístico) hasta obtener la geometría de la aproximación facial objetivo. Se llevó a cabo una validación mediante remuestreo a partir de una submuestra de los 500

TACs, dando como resultado un error medio de ~4 mm para geometría de la cara. La aproximación resultante es una forma facial objetiva probable, pero también es sintética (es decir, sin textura), y por lo tanto debe mejorarse artísticamente antes de su utilización en los casos forenses. Este enfoque de aproximación facial está integrado en el programa informático TIVMI, el cual está disponible gratuitamente en la web de los autores.

De Buhan y Nardoni [193] presentaron un nuevo método numérico para la reconstrucción facial que combina características clásicas, como el uso de una base de datos de cráneos y caras, y aspectos más originales que se detallan a continuación: 1) utilizan un método original de comparación de formas para vincular el cráneo desconocido con las plantillas de la base de datos; y 2) la cara final se ve como una máscara 3D elástica que se deforma y adapta al cráneo desconocido. En este método, el cráneo se considera como una superficie completa y no restringida a algunos landmarks anatómicos, lo que permite una descripción densa de la relación cráneo/cara. Además, el enfoque está totalmente automatizado.

En Liu y Li [196], a diferencia de la mayoría de las estrategias existentes que reconstruyen directamente el rostro a partir del cráneo, los autores utilizan una base de datos de fotografías faciales para crear muchos candidatos (modelos faciales 3D) para el rostro, y luego realizan una superposición con el cráneo del que se quiere hacer la reconstrucción facial. Para apoyar este proceso, emplean RNAs profundas (DL): primero, construyen un autocodificador para la reconstrucción facial 3D a partir de una fotografía y, segundo, utilizan un modelo generativo para adaptar el modelo facial al cráneo una vez realizada la superposición y dentro de unas restricciones geométricas.

Imaizumi y col. [195], desarrollaron una solución de software para la aproximación facial tridimensional del cráneo basada en TACs craneales de 59 voluntarios japoneses adultos (40 hombres, 19 mujeres). Se analizó la relación posicional entre la forma del cráneo y la superficie de la cabeza creando modelos de forma anatómicamente homóloga. Antes del modelado, las formas del cráneo se simplificaron ocultando las estructuras huecas (órbitas, fosas nasales, etc.) del cráneo. El grosor del tejido blando, representado por la distancia entre los vértices correspondientes de la superficie simplificada del cráneo y la cabeza, se calculó para cada individuo, y se promedió para cada sexo. Aunque las formas aproximadas de la cabeza de los individuos conocidos mostraban un parecido relativamente bueno tanto en la forma de toda la cabeza como en las partes faciales, se identificaron algunos errores, particularmente en áreas con mayor grosor de tejido blando en las mejillas, y tejido más grueso en la glabella, la nariz, la boca y el mentón. Además, crearon modelos craneofaciales a partir de modelos

promedio de la forma del cráneo y la superficie de la cabeza para cada sexo con el objetivo de que sirvieran como material de referencia con el que guiar la evaluación de la correspondencia anatómica en problemas de SCF.

Más allá de la reconstrucción facial a partir de un cráneo, y a la par relacionada con esta, nos parece importante señalar los últimos avances en la reconstrucción facial a partir de información del perfil genético [41, 199]. Denominados "fotoajuste molecular [196] o fenotipado basado en el ADN, estos métodos pueden complementar los métodos de aproximación facial para predecir rostros y son especialmente útiles para morfologías con relaciones tangibles limitadas con la estructura esquelética. En cuanto a los rasgos específicos de la cara que se pueden predecir con gran exactitud está el color del pelo rojo y los colores azul y marrón del iris [201]. Se ha registrado una precisión de aproximadamente el 70% para la predicción del cabello rojo [202], mientras que los intervalos de predicción positiva de los colores del iris oscilaron entre el 66% y el 100% para los ojos azules y el 70% y el 100% para los ojos marrones [203,206]. Normalmente, los valores predictivos positivos para los ojos marrones eran más altos (> 85%) que para los azules (> 75%), con una drástica reducción de la misma estadística para los llamados colores intermedios de los ojos. También se están investigando, probando y validando modelos de predicción del color de la piel [207,208].

4. CONCLUSIONES Y DISCUSIÓN.

La IA surgió a mediados del siglo pasado como resultado del trabajo de investigación de un reducido grupo de científicos de diferentes campos (principalmente, matemáticos, psicólogos e ingenieros). Actualmente, está presente en un sinfín de aplicaciones y servicios de todo tipo, donde las técnicas de IA se aplican con notable éxito en el reconocimiento del habla, la traducción automática, los videojuegos, el transporte y la robótica, entre muchos otros, y muy notablemente en la sanidad y la imagen médica. No hay duda de que la IA ha llegado a nuestra sociedad para quedarse. Por lo tanto, es sorprendente la escasa penetración de la IA en la medicina forense, más teniendo en cuenta el notable éxito de estas técnicas en otras disciplinas médicas en cuanto al uso de sistemas inteligentes para ampliar las capacidades de los expertos humanos. De hecho, las técnicas de IA pueden utilizarse para automatizar tareas tediosas para los humanos, como la segmentación de estructuras anatómicas o regiones faciales, el registro de imágenes para estudiar la correspondencia anatómica, o el marcado de información relevante en las imágenes (landmarks, por ejemplo). La IA también proporciona herramientas para integrar y estructurar la información

disponible sobre un determinado problema, permitiendo la inferencia y la exploración para el descubrimiento de conocimientos. En consecuencia, las técnicas de IA tienen la potencialidad de mejorar los enfoques manuales existentes en términos de reproducibilidad, robustez, precisión y tiempo de ejecución. Las máquinas son mejores que los humanos en tareas repetibles/bien definidas, y en tareas que requieren alta precisión/velocidad de ejecución. A su vez, los humanos son mejores que las máquinas en todo lo demás. Pero, en particular, los humanos pueden lidiar con la incertidumbre, aprender de los ejemplos y generalizar el conocimiento, mejorar su desempeño basado en la experiencia, evolucionar y cooperar para realizar tareas difíciles. Los sistemas actuales de IA están mejorando su eficiencia en los escenarios mencionados anteriormente (el AA, las técnicas de VA y el *soft computing* son campos de investigación dedicados a hacer que las máquinas emulen las habilidades de los humanos).

Dado que la IA tiene actualmente una presencia residual en la AF, en este artículo se ha presentado la primera revisión bibliográfica del estado de la técnica en la aplicación de enfoques basados en IA a la AF. La aplicación de la IA reduciría el tiempo necesario para realizar la identificación forense de personas. Ello permitiría resolver más casos en menos tiempo, permitiendo, por ejemplo, la realización de múltiples comparaciones en problemas de SCF o CR (algo impensable hasta la fecha debido a la falta de esta tecnología). Por otra parte, las decisiones podrían apoyarse en resultados objetivos y reproducibles que, tal vez, dotarías a estas técnicas de identificación de un mayor reconocimiento a la par que facilitarían su aplicación. Los estudios mencionados a lo largo de este capítulo del libro muestran que se pueden entrenar RNAs para estimar el PB, describir las condiciones patológicas o clasificar el tipo de traumatismos de un individuo a partir de restos óseos o imágenes radiográficas. Son capaces de estimar la edad, el sexo o identificar a partir de fotografías de personas vivas con gran precisión y tienen el potencial de eliminar el sesgo humano en todas estas tareas. Sin embargo, es importante destacar una de las principales características, y al mismo tiempo limitaciones, de estos modelos de AA: el resultado será tan bueno como los datos disponibles, por lo que, si el conjunto de datos utilizado para entrenar estos sistemas no es lo suficientemente amplio y diverso, los resultados serán seguramente subóptimos.

En nuestra opinión, hay dos cuestiones principales que limitan la aplicación y el impacto positivo de las técnicas de IA al campo de la AF:

- 1 **La existencia de un número muy reducido de grupos de investigación multidisciplinares que incluyan a investigadores en AF y disciplinas afines (odontología,**

patología, anatomía) y expertos en IA. Esta interacción es esencial para establecer una colaboración fluida y productiva entre las diferentes disciplinas científicas. Por una parte, permite unificar la terminología y facilitar la transferencia útil de conocimientos entre los científicos. Por otra parte, permite unir fuerzas en una dirección común hacia proyectos de investigación que supongan un reto (incluido el reconocimiento de las posibilidades y limitaciones de los diferentes campos de conocimiento involucrados). Esta colaboración interdisciplinaria debería dar fruto en el reconocimiento de que las máquinas no están aquí para reemplazar a los seres humanos, sino para complementarlos, ayudarlos y asistirlos en aquellas tareas que los humanos no quieren realizar o quieren realizar más rápidamente.

- 2 La falta de conjuntos de datos públicos amplios y abiertos para fines de investigación en AF.** El intercambio de conocimientos y la posibilidad de mejorar los enfoques existentes, y proponer nuevos métodos para resolver problemas, son las principales fuerzas que impulsan la evolución de la ciencia. En la actualidad, cuando los enfoques que requieren gran cantidad de datos, como el DL, están rompiendo barreras cada semana en la resolución de problemas complejos, la disponibilidad de datos es el punto de partida y el requisito para la validación de muchas técnicas de IA. La garantía de procedimientos objetivos para evaluar el rendimiento de esas técnicas es esencial. Para comparar el rendimiento de los métodos recientemente desarrollados, debería disponerse de un conjunto de datos forenses común de estudios de casos conocidos. Se trata de una práctica muy extendida en otros campos tan próximos como la medicina clínica y la investigación biomédica, con ejemplos representativos como la clasificación y localización de enfermedades comunes a partir de radiografías de tórax [209], la clasificación del melanoma en fotografías de la piel [210], el archivo de imágenes del cáncer [211], el atlas del cerebro del ratón y humano (Allen) para combinar la genómica y la neuroanatomía [212], la clasificación de las lesiones gastrointestinales mediante colonoscopia regular [213], la base de datos MIMIC de cuidados críticos que contiene datos de salud asociados con aproximadamente 40.000 pacientes de cuidados críticos (Johnson et al., 2016), la iniciativa de neuroimágenes de la enfermedad de Alzheimer [214], o la iniciativa OpenNeuro para compartir datos de IRM y RMf [215], entre otros. También es importante reconocer el hecho de que no se trata sólo de adquirir y almacenar datos, sino de cómo se compilan esos datos para cumplir los requisitos éticos y evitar un posible sesgo en los mismos [216] (por ejemplo, que debido a que hay muchos más ejemplos de hombres que de mujeres, la

clasificación del sistema funciona mejor con los primeros que con los segundos), entre otros.

Una novedosa base de datos pública, “*The New Mexico Decent Image Database*” [217], podría ser el mejor ejemplo existente del tipo de recurso que puede impulsar la investigación y la multidisciplinariedad en el campo de la AF. Esta base de datos recoge TACs de cuerpo entero de más de 15.000 personas del estado norteamericano de Nuevo México con una resolución de 0,5 mm y los metadatos incluyen casi 60 variables sobre la demografía, la vida y la muerte de los individuos. Abre la posibilidad de desarrollar nuevos modelos (por ejemplo, relacionando los parámetros extrínsecos y la morfología ósea para la estimación de la edad) y la validación de los existentes sobre un número sin precedentes de muestras.

En el contexto de las ciencias forenses, los avances en la potencia de procesamiento de las computadoras y la magnitud de las técnicas de IA emergentes durante la última década, como el AA, han mostrado una mejora en las capacidades de los ordenadores para comprender imágenes y reconocer patrones con una precisión similar o mayor a la de un experto humano. Su gran potencial reside en la capacidad de ejecutar automáticamente tareas que para el experto son tediosas, requieren mucho tiempo y son propensas a errores (como la etapa de SCC en el método de SCF). La potencia de computación y los algoritmos automáticos permitirán a los expertos forenses abordar y resolver un mayor número de casos en menos tiempo, lo que permitirá realizar múltiples comparaciones de casos en las bases de datos de personas desaparecidas, restos óseos no identificados, víctimas de desastres masivos, terrorismo, etc. También facilitan el análisis numérico de tareas subjetivas al hacer más objetivos los métodos de identificación, apoyando al experto en la toma de decisiones y sirviendo como una mejor garantía jurídica en los procedimientos judiciales.

La motivación principal de este artículo es dar a conocer lo que la IA ya es capaz de hacer, y presentar algunas contribuciones que puede aportar a corto plazo en el campo de la IA. La tragedia humana de las personas desaparecidas y los restos no identificados necesita soluciones a la altura de las circunstancias, en las que los expertos forenses tienen que apoyarse en la tecnología para poder abordar de manera fiable el enorme número de casos a los que se enfrentan. Como conclusión, consideramos que la AF puede beneficiarse enormemente de la ciencia y la tecnología. Después de todo, como dijo Garry Kasparov después de perder contra la computadora de ajedrez Deep Blue en 1997, es probable que un buen humano más una máquina sea la mejor combinación.

5. AGRADECIMIENTOS.

Este trabajo ha sido cofinanciado por el Ministerio de Ciencia, Innovación y Universidades, a través de las subvenciones SNEO-20191236 (programa Neotec del CDTI), PTQ-17-09306 (programa Torres Quevedo) y PGC2018-101216-B-I00 (proyecto EXASOCO), todos ellos incluyendo fondos FEDER.

6. BIBLIOGRAFÍA.

1. RITCHIE H, ROSER M. Natural Disasters. *Our World in Data*, <https://ourworldindata.org/natural-disasters> (2014, accessed 2 May 2020).
2. HAYNER PB. Truth commissions: a schematic overview. *International Review of the Red Cross* 2006; 88: 295-310.
3. Refugees UNHC for. Refworld | Report of the Working Group on Enforced or Involuntary Disappearances, Addendum?: Mission to Spain. *Refworld*, <https://www.refworld.org/docid/53eb35814.html> (accessed 2 May 2020).
4. JARAMILLO K. Latin America: The region with the highest number of enforced disappearances - LatinAmerican Post, <https://latinamericanpost.com/23187-latin-america-the-region-with-the-highest-number-of-enforced-disappearances> (accessed 2 May 2020).
5. Mexico's missing people agency searching for trafficking victims. *Reuters*, 15 March 2019, <https://www.reuters.com/article/us-mexico-trafficking-government-intervi-idUSKCN1QW2CG> (15 March 2019, accessed 2 May 2020).
6. Missing Migrants Project, <https://missingmigrants.iom.int/> (accessed 2 May 2020).
7. Latest statistics and graphics on refugee and migrant children, <https://www.unicef.org/eca/emergencies/latest-statistics-and-graphics-refugee-and-migrant-children> (accessed 2 May 2020).
8. Scientific Working Group for Forensic Anthropology (SWGANTH). 6.
9. IANNACCONE G. W. W. GREULICH AND S. I. Pyle: Radiographic atlas of skeletal development of the hand and wrist. 2nd edition. I volume-atlante di 256 pagine. Stanford University Press, Stanford, California, 1959. *Acta geneticae medicae et gemellologiae: twin research* 1959; 8: 513-513.
10. TANNER JM, WHITEHOUSE RH, MARSHALL WA, et al. Prediction of adult height from height, bone age, and occurrence of menarche, at ages 4 to 16 with allowance for midparent height. *Arch Dis Child* 1975; 50: 14-26.
11. PERALTA D, TRIGUERO I, GARCÍA S, et al. DPD-DFF: A dual phase distributed scheme with double fingerprint fusion for fast and accurate identification in large databases. *Information Fusion* 2016; 32: 40-51.
12. RANJAN R, SANKARANARAYANAN S, CASTILLO CD, et al. An All-In-One Convolutional Neural Network for Face Analysis. In: *2017 12th IEEE International Conference on Automatic Face Gesture Recognition (FG 2017)*. 2017, pp. 17-24.
13. LASERSON J, LANTSMAN CD, COHEN-SFADY M, et al. TextRay: Mining Clinical Reports to Gain a Broad Understanding of Chest X-Rays. In: Frangi AF, Schnabel JA, Davatzikos C, et al. (eds) *Medical Image Computing and Computer Assisted Intervention - MICCAI 2018*. Cham: Springer International Publishing, 2018, pp. 553-561.
14. ESTEVA A, KUPREL B, NOVOA RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017; 542: 115-118.
15. CAMERIERE R, DE LUCA S, DE ANGELIS D, et al. Reliability of Schmelzing's stages of ossification of medial clavicular epiphyses and its validity to assess 18 years of age in living subjects. *Int J Legal Med* 2012; 126: 923-932.
16. RITZ-TIMME S, GABRIEL P, OBERTOVÀ Z, et al. A new atlas for the evaluation of facial features: advantages, limits, and applicability. *Int J Legal Med* 2011; 125: 301-306.
17. VALSECCHI A, IRURITA OLIVARES J, MESEJO P. Age estimation in forensic anthropology: methodological considerations about the validation studies of prediction models. *Int J Legal Med* 2019; 133: 1915-1924.
18. JAYAPRAKASH PT. Practical relevance of pattern uniqueness in forensic science. *Forensic Science International* 2013; 231: 403.e1-403.e16.
19. CAPLE J, STEPHAN CN. A standardized nomenclature for craniofacial and facial anthropometry. *Int J Legal Med* 2016; 130: 863-879.
20. ROSS A, LANFEAR A, MAXWELL A. Establishing Standards for Side-by-Side Radiographic Comparisons. *The American Journal of Forensic Medicine and Pathology* 2016; 37: 86-94.
21. DAMAS S, WILKINSON C, KAHANA T, et al. Study on the performance of different craniofacial superimposition approaches (II): Best practices proposal. *Forensic Science International* 2015; 257: 504-508.
22. IBÁÑEZ O, VICENTE R, NAVEGA DS, et al. Study on the performance of different craniofacial superimposition approaches (I). *Forensic Science International* 2015; 257: 496-503.
23. HUETE MI, IBÁÑEZ O, WILKINSON C, et al. Past, present, and future of craniofacial superimposition: Literature and international surveys. *Leg Med (Tokyo)* 2015; 17: 267-278.

24. RUSSELL SJ, NORVIG P. *Artificial intelligence: a modern approach*. Englewood Cliffs, N.J: Prentice Hall, 1995.
25. FORSYTH DA, PONCE J. *Computer Vision: A Modern Approach*. Prentice Hall Professional Technical Reference, 2002.
26. ZHANG H, FRITTS JE, GOLDMAN SA. *Image Segmentation Evaluation: A Survey of Unsupervised Methods*. 2008.
27. GÓMEZ O, MESEJO P, IBÁÑEZ O, et al. Deep architectures for high-resolution multi-organ chest X-ray image segmentation. *Neural Comput & Applic*. Epub ahead of print 21 October 2019. DOI: 10.1007/s00521-019-04532-y.
28. OLMOS R, TABIK S, HERRERA F. Automatic handgun detection alarm in videos using deep learning. *Neurocomputing* 2018; 275:66-72.
29. GÓMEZ G, MESEJO P, IBÁÑEZ O, et al. Automatic Cephalometric Landmarks Localization Using Deep Convolutional Neural Networks. In: *Proceedings of the 18th Meeting of the International Association for Craniofacial Identification (IACI)*. Baton Rouge, USA, 2019, p. 48.
30. MARKELJ P, TOMAŽEVIC D, LIKAR B, et al. A review of 3D/2D registration methods for image-guided interventions. *Medical Image Analysis* 2012; 16: 642-661.
31. Bishop C. *Pattern Recognition and Machine Learning*. New York: Springer-Verlag, <https://www.springer.com/gp/book/9780387310732> (2006, accessed 2 May 2020).
32. LECUN Y, BENGIO Y, HINTON G. Deep learning. *Nature* 2015; 521: 436-444.
33. HINTON G, DENG L, YU D, et al. Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups. *IEEE Signal Processing Magazine* 2012; 29:82-97.
34. SUTSKEVER I, VINYALS O, LE QV. Sequence to Sequence Learning with Neural Networks. In: Ghahramani Z, Welling M, Cortes C, et al. (eds) *Advances in Neural Information Processing Systems 27*. Curran Associates, Inc., pp. 3104-3112.
35. KRIZHEVSKY A, SUTSKEVER I, HINTON GE. ImageNet classification with deep convolutional neural networks. *Commun ACM* 2017; 60:84-90.
36. SILVER D, HUANG A, MADDISON CJ, et al. Mastering the game of Go with deep neural networks and tree search. *Nature* 2016; 529: 484-489.
37. ENGELBRECHT AP. *Computational Intelligence: An Introduction, 2nd Edition | Wiley*. <https://www.wiley.com/en-us/Computational+Intelligence%3A+An+Introduction%2C+2nd+Edition-p-9780470035610> (accessed 2 May 2020).
38. MCCULLOCH WS, PITTS W. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* 1943; 5: 115-133.
39. BEAUTHIER J-P, VALCK E, LEFEVRE P, et al. Mass Disaster Victim Identification: The Tsunami Experience. *The Open Forensic Science Journal*; 2, <https://benthamopen.com/ABSTRACT/TOFORSJ-2-54> (2009, accessed 3 May 2020).
40. BETHARD JD. *Forensic Anthropology: An Introduction*. Edited by MariaTeresa A. Tersigni-Tarrant and Natalie R. Shirley. Boca Raton, FL: CRC Press. 2013. 462 pp. ISBN 978-1-4398-1646-2. \$99.95 (hardcover). *American Journal of Physical Anthropology* 2014; 154: 166-167.
41. STEPHAN CN, CAPLE JM, GUYOMARC'H P, et al. An overview of the latest developments in facial imaging. *Forensic Sci Res* 2018; 4: 10-28.
42. VALENTINE T, DAVIS JP. *Forensic Facial Identification: Theory and Practice of Identification from Eyewitnesses, Composites and CCTV | Wiley*. <https://www.wiley.com/en-us/Forensic+Facial+Identification%3A+Theory+and+Practice+of+Identification+from+Eyewitnesses%2C+Composites+and+CCTV-p-9781118469583> (accessed 3 May 2020).
43. SCHMELING A, DETTMAYER RB, RUDOLF E, et al. Forensic Age Estimation. *Deutsches Arzteblatt international*. Epub ahead of print 2016. DOI: 10.3238/arztebl.2016.0044.
44. JAYARAMAN J, ROBERTS GJ, WONG HM, et al. Ages of legal importance: Implications in relation to birth registration and age assessment practices. *Med Sci Law* 2016; 56: 77-82.
45. Forensische Altersdiagnostik - DGRM, <https://www.dgrm.de/arbeitsgemeinschaften/forensische-altersdiagnostik/> (accessed 2 May 2020).
46. SCHMELING A, GRUNDMANN C, FUHRMANN A, et al. Criteria for age estimation in living individuals. *Int J Legal Med* 2008; 122: 457.
47. TODD TW. Age changes in the pubic bone. *American Journal of Physical Anthropology* 1921; 4: 1-70.
48. UBELAKER DH. *Enterramientos humanos: excavación, análisis, interpretación*. Sociedad de Ciencias Aranzadi, 2007.
49. BYERS SN. *Introduction to Forensic Anthropology*. Pearson Education, 2011.
50. MANSOURVAR M, ISMAIL MA, HERAWAN T, et al. Automated Bone Age Assessment: Motivation, Taxonomies, and Challenges. *Computational and Mathematical Methods in Medicine* 2013; 2013: e391626.
51. BECK LA. Standards for data collection from human skeletal remains. Edited by Jane E. Buikstra and Douglas H. Ubelaker. 272 pp. Fayetteville: Arkansas Archeological Survey Research Series No. 44, 1994. \$25.00 (paper). *American Journal of Human Biology* 1995; 7: 672-672.

52. KRISHAN K, CHATTERJEE PM, KANCHAN T, et al. A review of sex estimation techniques during examination of skeletal remains in forensic anthropology casework. *Forensic Science International* 2016; 261:165.e1-165.e8.
53. PINTO SCD, URBANOVA P, CESAR-JR RM. Two-Dimensional Wavelet Analysis of Supraorbital Margins of the Human Skull for Characterizing Sexual Dimorphism. *IEEE Transactions on Information Forensics and Security* 2016; 11:1542-1548.
54. SIERP I, HENNEBERG M. The Difficulty of Sexing Skeletons from Unknown Populations. *Journal of Anthropology* 2015; 2015: e908535.
55. DARMAWAN MF, YUSUF SM, ROZI MA, et al. Hybrid PSO-ANN for sex estimation based on length of left hand bone. In: *2015 IEEE Student Conference on Research and Development (SCoReD)*. 2015, pp. 478-483.
56. ABDULLAH H, JAMIL MMA, NOR FM. Automated haversian canal detection for histological sex determination. In: *2017 IEEE Symposium on Computer Applications Industrial Electronics (ISCAIE)*. 2017, pp. 69-74.
57. BEWES J, LOW A, MORPHETT A, et al. Artificial intelligence for sex determination of skeletal remains: Application of a deep learning artificial neural network to human skulls. *Journal of Forensic and Legal Medicine* 2019; 62: 40-43.
58. KALOI MA, HE K. Child Gender Determination with Convolutional Neural Networks on Hand Radio-Graphs. *arXiv:181105180 [cs]*, <http://arxiv.org/abs/1811.05180> (2018, accessed 4 May 2020).
59. GOODFELLOW I, BENGIO Y, COURVILLE A. Deep Learning, <https://mitpress.mit.edu/books/deep-learning> (2016, accessed 4 May 2020).
60. OAKDEN-RAYNER L, CARNEIRO G, BESSEN T, et al. Precision Radiology: Predicting longevity using feature engineering and deep learning methods in a radiomics framework. *Scientific Reports* 2017; 7: 1-13.
61. SZEGEDY C, WEI LIU, YANGQING JIA, et al. Going deeper with convolutions. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2015, pp. 1-9.
62. TODD TW. Age changes in the pubic bone. I. The male white pubis. *American Journal of Physical Anthropology* 1920; 3: 285-334.
63. BROOKS S, SUCHEY JM. Skeletal age determination based on the os pubis: A comparison of the Acsádi-Nemeskéri and Suchey-Brooks methods. *Hum Evol* 1990; 5: 227-238.
64. ISCAN MY, LOTH SR, WRIGHT RK. Age estimation from the rib by phase analysis: white males. *J Forensic Sci* 1984; 29: 1094-1104.
65. ISCAN MY, LOTH SR, WRIGHT RK. Age estimation from the rib by phase analysis: white females. *J Forensic Sci* 1985; 30: 853-863.
66. MEINDL RS, LOVEJOY CO. Ectocranial suture closure: a revised method for the determination of skeletal age at death based on the lateral-anterior sutures. *Am J Phys Anthropol* 1985; 68: 57-66.
67. LAMENDIN H, BACCINO E, HUMBERT JF, et al. A simple technique for age estimation in adult corpses: the two criteria dental method. *J Forensic Sci* 1992; 37: 1373-1379.
68. MEINDL RS, LOVEJOY CO. Ectocranial suture closure: a revised method for the determination of skeletal age at death based on the lateral-anterior sutures. *Am J Phys Anthropol* 1985; 68: 57-66.
69. AJA-FERNÁNDEZ S, DE LUIS-GARCÍA R, MARTÍN-FERNÁNDEZ MÁ, et al. A computational TW3 classifier for skeletal maturity assessment. A Computing with Words approach. *Journal of Biomedical Informatics* 2004; 37: 99-107.
70. STERN D, EBNER T, BISCHOF H, et al. Fully automatic bone age estimation from left hand MR images. *Med Image Comput Assist Interv* 2014; 17: 220-227.
71. TANNER JM, WHITEHOUSE RH, MARSHALL WA, et al. *Assessment of Skeletal Maturity and Prediction of Adult Height*. Subsequent edition. London?; New York: Academic Pr, 1983.
72. TANNER JM, HEALY MJR, CAMERON N, et al. *Assessment of Skeletal Maturity and Prediction of Adult Height (TW3 Method)*. W.B. Saunders, 2001.
73. ŠTERN D, PAYER C, URSCHLER M. Automated age estimation from MRI volumes of the hand. *Medical Image Analysis* 2019; 58: 101538.
74. KIM JR, SHIM WH, YOON HM, et al. Computerized Bone Age Estimation Using Deep Learning Based Program: Evaluation of the Accuracy and Efficiency. *American Journal of Roentgenology* 2017; 209: 1374-1380.
75. LARSON DB, CHEN MC, LUNGREN MP, et al. Performance of a Deep-Learning Neural Network Model in Assessing Skeletal Maturity on Pediatric Hand Radiographs. *Radiology* 2018; 287: 313-322.
76. LEE H, TAJMIR S, LEE J, et al. Fully Automated Deep Learning System for Bone Age Assessment. *J Digit Imaging* 2017; 30: 427-441.
77. MUTASA S, CHANG PD, RUZAL-SHAPIRO C, et al. MABAL: a Novel Deep-Learning Architecture for Machine-Assisted Bone Age Labeling. *J Digit Imaging* 2018; 31: 513-519.
78. LI Y, HUANG Z, DONG X, et al. Forensic age estimation for pelvic X-ray images using deep learning. *Eur Radiol* 2019; 29: 2322-2329.
79. ANDERSON MF, ANDERSON DT, WESCOTT DJ. Estimation of adult skeletal age-at-death using the Sugeno fuzzy integral. *American Journal of Physical Anthropology* 2010; 142: 30-41.

80. ANDERSON DT, KELLER JM, ANDERSON M, et al. Linguistic description of adult skeletal age-at-death estimations from fuzzy integral acquired fuzzy sets. In: *2011 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2011)*. 2011, pp. 2274-2281.
81. VILLAR P, ALEMÁN I, CASTILLO L, et al. A first approach to a fuzzy classification system for age estimation based on the pubic bone. In: *2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. 2017, pp.1-6.
82. ISHIBUCHI H, NAKASHIMA T, NII M. *Classification and Modeling with Linguistic Information Granules: Advanced Approaches to Linguistic Data Mining*. Springer Science & Business Media, 2004.
83. HSIEH C-W, JONG T-L, CHOU Y-H, et al. Computerized geometric features of carpal bone for bone age estimation. *Chin Med J* 2007; 120: 767-770.
84. BREEN MA, TSAI A, STAMM A, et al. Bone age assessment practices in infants and older children among Society for Pediatric Radiology members. *Pediatr Radiol* 2016; 46: 1269-1274.
85. CARTY H. Assessment of skeletal maturity and prediction of adult height (TW3 method). *The Journal of Bone and Joint Surgery British volume* 2002; 84-B: 310-311.
86. LARSON DB, CHEN MC, LUNGREN MP, et al. Performance of a Deep-Learning Neural Network Model in Assessing Skeletal Maturity on Pediatric Hand Radiographs. *Radiology* 2017; 287: 313-322.
87. LEE JH, KIM KG. Applying Deep Learning in Medical Images: The Case of Bone Age Estimation. *Healthc Inform Res* 2018; 24: 86-92.
88. GERTYCH A, ZHANG A, SAYRE J, et al. Bone Age Assessment of Children using a Digital Hand Atlas. *Comput Med Imaging Graph* 2007; 31: 322-331.
89. ZHANG A, SAYRE JW, VACHON L, et al. Racial Differences in Growth Patterns of Children Assessed on the Basis of Bone Age. *Radiology* 2009; 250: 228-235.
90. THODBERG HH, NEUHOF J, RANKE MB, et al. Validation of bone age methods by their ability to predict adult height. *Horm Res Paediatr* 2010; 74: 15-22.
91. SPAMPINATO C, PALAZZO S, GIORDANO D, et al. Deep learning for automated skeletal bone age assessment in X-ray images. *Medical Image Analysis* 2017; 36: 41-51.
92. ŠTERN D, PAYER C, LEPETIT V, et al. Automated Age Estimation from Hand MRI Volumes Using Deep Learning. In: Ourselin S, Joskowicz L, Sabuncu MR, et al. (eds) *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2016*. Cham: Springer International Publishing, 2016, pp. 194-202.
93. GULSHAN V, PENG L, CORAM M, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA* 2016; 316: 2402-2410.
94. KOOI T, LITJENS G, VAN GINNEKEN B, et al. Large scale deep learning for computer aided detection of mammographic lesions. *Medical Image Analysis* 2017; 35: 303-312.
95. HUA K-L, HSU C-H, HIDAYATI SC, et al. Computer-aided classification of lung nodules on computed tomography images via deep learning technique. *Onco Targets Ther* 2015; 8: 2015-2022.
96. OLCZAK J, FAHLBERG N, MAKI A, et al. Artificial intelligence for analyzing orthopedic trauma radiographs. *Acta Orthop* 2017; 88: 581-586.
97. CHUNG SW, HAN SS, LEE JW, et al. Automated detection and classification of the proximal humerus fracture by using deep learning algorithm. *Acta Orthop* 2018; 89: 468-473.
98. GUPTA A, VENKATESH S, CHOPRA S, et al. Generative Image Translation for Data Augmentation of Bone Lesion Pathology. In: *International Conference on Medical Imaging with Deep Learning*, pp. 225-235.
99. STEPHAN CN, AMIDAN B, TREASE H, et al. Morphometric comparison of clavicle outlines from 3D bone scans and 2D chest radiographs: a shortlisting tool to assist radiographic identification of human skeletons. *J Forensic Sci* 2014; 59: 306-313.
100. CHRISTENSEN AM, SMITH MA, GLEIBER DS, et al. The Use of X-ray Computed Tomography Technologies in *Forensic Anthropology*. *Forensic Anthropology* 2018; 1: 124-140.
101. HATCH GM, DEDOUIT F, CHRISTENSEN AM, et al. RADid: A pictorial review of radiologic identification using postmortem CT. *Journal of Forensic Radiology and Imaging* 2014; 2: 52-59.
102. THALI MJ, BRAUN M, DIRNHOFER R. Optical 3D surface digitizing in forensic medicine: 3D documentation of skin and bone injuries. *Forensic Science International* 2003; 137: 203-208.
103. SANDSTRÖM S, OSTENSEN H, PETTERSSON H, et al. *The WHO manual of diagnostic imaging?: radiographic technique and projections*, <https://apps.who.int/iris/handle/10665/42720> (2003, accessed 4 May 2020).
104. DAFFNER RH, HARTMAN MS. *Clinical Radiology: The Essentials*. Wolters Kluwer Lippincott Williams & Wilkins Health, 2013.
105. RIGBY D-M, HACKING L. Interpreting the chest radiograph. *Anaesthesia & Intensive Care Medicine* 2018; 19: 50-54.
106. VAN GINNEKEN B, TER HAAR ROMENY BM, VIERGEVER MA. Computer-aided diagnosis in chest radiography: a survey. *IEEE Transactions on Medical Imaging* 2001; 20: 1228-1241.

107. Statistics>> Diagnostic Imaging Dataset 2015-16 Data, <https://www.england.nhs.uk/statistics/statistical-work-areas/diagnostic-imaging-dataset/diagnostic-imaging-dataset-2015-16-data/> (accessed 4 May 2020).
108. CHRISTENSEN AM. Assessing the variation in individual frontal sinus outlines. *Am J Phys Anthropol* 2005; 127: 291-295.
109. CHRISTENSEN AM. Testing the reliability of frontal sinuses in positive identification. *J Forensic Sci* 2005; 50: 18-22.
110. MAXWELL AB, ROSS AH. A Radiographic Study on the Utility of Cranial Vault Outlines for Positive Identifications. *Journal of Forensic Sciences* 2014; 59: 314-318.
111. JAIN AK, CHEN H. Matching of dental X-ray images for human identification. *Pattern Recognition* 2004; 37: 1519-1532.
112. NOMIR O, ABDEL-MOTTALEB M. Human Identification From Dental X-Ray Images Based on the Shape and Appearance of the Teeth. *IEEE Transactions on Information Forensics and Security* 2007; 2: 188-197.
113. CAPLE J, BYRD J, STEPHAN CN. Elliptical Fourier analysis: fundamentals, applications, and value for forensic anthropology. *Int J Legal Med* 2017; 131: 1675-1690.
114. TABOR Z, KARPISZ D, WOJNAR L, et al. An Automatic Recognition of the Frontal Sinus in X-Ray Images of Skull. *IEEE Transactions on Biomedical Engineering* 2009; 56: 361-368.
115. DERRICK SM, HIPPI JA, GOEL P. Chapter 23 - The Computer-Assisted Decedent Identification Method of Computer-Assisted Radiographic Identification. In: Latham KE, Bartelink EJ, Finnegan M (eds) *New Perspectives in Forensic Human Skeletal Identification*. Academic Press, pp. 265-276.
116. KAHANA T. *El aporte de la radiología al avance de la antropología forense perspectiva profesional*. Editorial de la Universidad de Granada, 2009.
117. PFAEFFLI M, VOCK P, DIRNHOFER R, et al. Post-mortem radiological CT identification based on classical ante-mortem X-ray examinations. *Forensic Science International* 2007; 171: 111-117.
118. SHINKAWA N, HIRAI T, NISHII R, et al. Usefulness of 2D fusion of postmortem CT and antemortem chest radiography studies for human identification. *Jpn J Radiol* 2017; 35: 303-309.
119. NIESPODZIEWANSKI E, STEPHAN CN, GUYOMARC'H P, et al. Human Identification via Lateral Patella Radiographs: A Validation Study. *Journal of Forensic Sciences* 2016; 61: 134-140.
120. D'ALONZO SS, GUYOMARC'H P, BYRD JE, et al. A Large-Sample Test of a Semi-Automated Clavicle Search Engine to Assist Skeletal Identification by Radiograph Comparison. *Journal of Forensic Sciences* 2017; 62: 181-186.
121. GÓMEZ O, IBÁÑEZ O, VALSECCHI A, et al. 3D-2D silhouette-based image registration for comparative radiography-based forensic identification. *Pattern Recognition* 2018; 83: 469-480.
122. BROGDON BG. The Scope of Forensic Radiology. *Clinics in Laboratory Medicine* 1998; 18: 203-240.
123. GÓMEZ O, IBÁÑEZ O, MESEJO P, et al. Towards a Computer-aided Decision Support System for Comparative Radiography. In: *Proceedings of the XI Reunión Científica Internacional de la Asociación Española de Antropología y Odontología Forense (AEAOF)*. Pastrana, Spain, 2019.
124. IINO M, FUJIMOTO H, YOSHIDA M, et al. Identification of a jawless skull by superimposing post-mortem and ante-mortem CT. *Journal of Forensic Radiology and Imaging* 2016; 6: 31-37.
125. RUDER TD, BRUN C, CHRISTENSEN AM, et al. Comparative radiologic identification with CT images of paranasal sinuses - Development of a standardized approach. *Journal of Forensic Radiology and Imaging* 2016; 7: 1-9.
126. HAEL A, COSTA ALF, MAYARA OLIVEIRA J, et al. Three-dimensional volumetric analysis of frontal sinus using medical software. *Journal of Forensic Radiology and Imaging* 2017; 11: 1-5.
127. DELOIRE L, DIALLO I, CADIEU R, et al. Post-mortem X-ray computed tomography (PMCT) identification using ante-mortem CT-scan of the sphenoid sinus. *Journal of Neuroradiology* 2019; 46: 248-255.
128. GASCHO D, PHILIPP H, FLACH PM, et al. Standardized medical image registration for radiological identification of decedents based on paranasal sinuses. *Journal of Forensic and Legal Medicine* 2018; 54: 96-101.
129. KIM D-I, LEE U-Y, PARK S-O, et al. Identification Using Frontal Sinus by Three-Dimensional Reconstruction from Computed Tomography*. *Journal of Forensic Sciences* 2013; 58: 5-12.
130. ZHONG X, YU D, FOONG KWC, et al. Towards automated pose invariant 3D dental biometrics. In: 2011 *International Joint Conference on Biometrics (IJCB)*. 2011, pp. 1-7.
131. ZHONG X, YU D, WONG YS, et al. 3D dental biometrics: Alignment and matching of dental casts for human identification. *Computers in Industry* 2013; 64: 1355-1370.
132. GIBELLI D, CELLINA M, CAPPELLA A, et al. An innovative 3D-3D superimposition for assessing anatomical uniqueness of frontal sinuses through segmentation on CT scans. *Int J Legal Med* 2019; 133: 1159-1165.
133. DECKER SJ, FORD JM. Forensic personal identification utilizing part-to-part comparison of CT-derived 3D lumbar models. *Forensic Science International* 2019; 294: 21-26.
134. ISCAN MY, HELMER RP. *Forensic analysis of the skull:: craniofacial analysis...*. Wiley-Liss, 1993.
135. YOSHINO M. Craniofacial superimposition. *Craniofacial Identification* 2012; 238-253.

136. DAMAS S, CORDÓN O, IBÁÑEZ O, et al. Forensic identification by computer-aided craniofacial superimposition: A survey. *ACM Comput Surv* 2011; 43: 27:1-27:27.
137. JAYAPRAKASH PT, SRINIVASAN GJ, AMRAVANESWARAN MG. Cranio-facial morphanalysis: a new method for enhancing reliability while identifying skulls by photo superimposition. *Forensic Science International* 2001; 117: 121-143.
138. PARK H-K, CHUNG J-W, KHO H-S. Use of hand-held laser scanning in the assessment of craniometry. *Forensic Science International* 2006; 160: 200-206.
139. DE ANGELIS D, SALA R, CANTATORE A, et al. A new computer-assisted technique to aid personal identification. *Int J Legal Med* 2009; 123: 351-356.
140. SINGARE S, LIAN Q, PING WANG W, et al. Rapid prototyping assisted surgery planning and custom implant design. *Rapid Prototyping Journal* 2009; 15: 19-23.
141. FANTINI M, DE CRESCENZIO F, PERSIANI F, et al. 3D restitution, restoration and prototyping of a medieval damaged skull. *Rapid Prototyping Journal* 2008; 14: 318-324.
142. NICKERSON BA, FITZHORN PA, KOCH SK, et al. A methodology for near-optimal computational superimposition of two-dimensional digital facial photographs and three-dimensional cranial surface meshes. *J Forensic Sci* 1991; 36: 480-500.
143. SANTAMARÍA J, CORDÓN O, DAMAS S. Evolutionary Approaches for Automatic 3D Modeling of Skulls in Forensic Identification. In: Giacobini M (ed) *Applications of Evolutionary Computing*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 415-422.
144. SANTAMARÍA J, CORDÓN O, DAMAS S, et al. Performance evaluation of memetic approaches in 3D reconstruction of forensic objects. *Soft Comput* 2009; 13: 883-904.
145. BALLERINI L, CORDÓN O, DAMAS S, et al. Automatic 3D Modeling of Skulls by Scatter Search and Heuristic Features. In: Avineri E, Köppen M, Dahal K, et al. (eds) *Applications of Soft Computing*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 149-158.
146. CAMPOMANES-ÁLVAREZ BR, DAMAS S, CORDÓN Ó. Mesh simplification for 3D modeling using evolutionary multi-objective optimization. In: *2012 IEEE Congress on Evolutionary Computation*. 2012, pp. 1-8.
147. JOHNSTON B, CHAZAL P de. A review of image-based automatic facial landmark identification techniques. *J Image Video Proc* 2018; 2018: 86.
148. PORTO LF, CORREIA LIMA LN, PINHEIRO FLORES MR, et al. Automatic cephalometric landmarks detection on frontal faces: An approach based on supervised learning techniques. *Digital Investigation* 2019; 30: 108-116.
149. LINDNER C, WANG C-W, HUANG C-T, et al. Fully Automatic System for Accurate Localisation and Analysis of Cephalometric Landmarks in Lateral Cephalograms. *Scientific Reports* 2016; 6: 1-10.
150. SONG Y, QIAO X, IWAMOTO Y, et al. Automatic Cephalometric Landmark Detection on X-ray Images Using a Deep-Learning Method. *Applied Sciences* 2020; 10: 2547.
151. ISO/IEC 19794-5:2011(en), Information technology - Biometric data interchange formats - Part 5: Face image data, <https://www.iso.org/obp/ui/#iso:std:iso-iec:19794:-5:ed-2:vl:en> (accessed 7 May 2020).
152. CAMPOMANES-ÁLVAREZ BR, IBÁÑEZ O, NAVARRO F, et al. Computer vision and soft computing for automatic skull-face overlay in craniofacial superimposition. *Forensic Science International* 2014; 245: 77-86.
153. GHOSH AK, SINHA P. An economised craniofacial identification system. *Forensic Science International* 2001; 11.
154. CAMPOMANES-ÁLVAREZ BR, IBÁÑEZ O, CAMPOMANES-ÁLVAREZ C, et al. Modeling Facial Soft Tissue Thickness for Automatic Skull-Face Overlay. *IEEE Transactions on Information Forensics and Security* 2015; 10: 2057-2070.
155. IBÁÑEZ O, CAVALLI F, CAMPOMANES-ÁLVAREZ BR, et al. Ground truth data generation for skull-face overlay. *Int J Legal Med* 2015; 129: 569-581.
156. CAMPOMANES-ÁLVAREZ BR, IBÁÑEZ O, NAVARRO F, et al. Computer vision and soft computing for automatic skull-face overlay in craniofacial superimposition. *Forensic Science International* 2014; 245: 77-86.
157. IBÁÑEZ O, CORDÓN O, DAMAS S, et al. An advanced scatter search design for skull-face overlay in craniofacial superimposition. *Expert Systems with Applications* 2012; 39: 1459-1473.
158. IBÁÑEZ O, CORDÓN O, DAMAS S. A cooperative coevolutionary approach dealing with the skull-face overlay uncertainty in forensic identification by craniofacial superimposition. *Soft Comput* 2012; 16: 797-808.
159. IBÁÑEZ Ó, CORDÓN Ó, DAMAS S, et al. Modeling the Skull-Face Overlay Uncertainty Using Fuzzy Sets. *IEEE Transactions on Fuzzy Systems* 2011; 19: 946-959.
160. IBÁÑEZ O, BALLERINI L, CORDÓN O, et al. An experimental study on the applicability of evolutionary algorithms to craniofacial superimposition in forensic identification. *Information Sciences* 2009; 179: 3998-4028.
161. CUMMAUDO M, GUERZONI M, MARASCIUOLO L, et al. Pitfalls at the root of facial assessment on photographs: a quantitative study of accuracy in positioning facial landmarks. *Int J Legal Med* 2013; 127: 699-706.
162. CAMPOMANES-ÁLVAREZ BR, IBÁÑEZ O, NAVARRO F, et al. Dispersion assessment in the location of facial landmarks on photographs. *Int J Legal Med* 2015; 129: 227-236.

163. VALSECCHI A, DAMAS S, CORDÓN O. A Robust and Efficient Method for Skull-Face Overlay in Computerized Craniofacial Superimposition. *IEEE Transactions on Information Forensics and Security* 2018; 13:1960-1974.
164. STEPHAN CN, MEIKLE B, FREUDENSTEIN N, et al. Facial soft tissue thicknesses in craniofacial identification: Data collection protocols and associated measurement errors. *Forensic Science International* 2019; 304:109965.
165. BERMEJO E, CAMPOMANES-ÁLVAREZ C, VALSECCHI A, et al. Genetic algorithms for skull-face overlay including mandible articulation. *Information Sciences* 2017; 420:200-217.
166. RICCI A, MARELLA GL, APOSTOL MA. A New Experimental Approach to Computer-Aided Face/Skull Identification in Forensic Anthropology. *The American Journal of Forensic Medicine and Pathology* 2006; 27: 46-49.
167. YOSHINOM, MATSUDA H, KUBOTA S, et al. Computer-assisted skull identification system using video superimposition. *Forensic Science International* 1997; 90:231-244.
168. CAMPOMANES-ALVAREZ C, IBÁÑEZ O, CORDÓN O. Design of criteria to assess craniofacial correspondence in forensic identification based on computer vision and fuzzy integrals. *Applied Soft Computing* 2016; 46:596-612.
169. CAMPOMANES-ALVAREZ C, IBÁÑEZ O, CORDÓN O, et al. Hierarchical information fusion for decision making in craniofacial superimposition. *Information Fusion* 2018; 39: 25-40.
170. CAMPOMANES-ÁLVAREZ C, MARTOS-FERNÁNDEZ R, WILKINSON C, et al. Modeling Skull-Face Anatomical/Morphological Correspondence for Craniofacial Superimposition-Based Identification. *IEEE Transactions on Information Forensics and Security* 2018; 13:1481-1494.
171. NAGPAL S, SINGH M, JAIN A, et al. On matching skulls to digital face images: A preliminary approach. In: *2017 IEEE International Joint Conference on Biometrics (IJCB)*. 2017, pp. 813-819.
172. SINGH M, NAGPAL S, SINGH R, et al. Learning A Shared Transform Model for Skull to Digital Face Image Matching. *arXiv:180804571 [cs]*, <http://arxiv.org/abs/1808.04571> (2018, accessed 7 May 2020).
173. DUAN F, YANG Y, LI Y, et al. Skull Identification via Correlation Measure Between Skull and Face Shape. *IEEE Transactions on Information Forensics and Security* 2014; 9:1322-1332.
174. TAN JS, LIAO IY, VENKAT I, et al. Computer-aided superimposition via reconstructing and matching 3D faces to 3D skulls for forensic craniofacial identifications. *Vis Comput.* Epub ahead of print 30 October 2019. DOI: 10.1007/s00371-019-01767-7.
175. TAYLOR KT. *Forensic Art and Illustration*. CRC Press. Epub ahead of print 15 September 2000. DOI: 10.1201/9781420036954.
176. STEPHAN CN, HENNEBERG M. Building faces from dry skulls: are they recognized above chance rates? *J Forensic Sci* 2001; 46:432-440.
177. CLEMENT JG, MARKS MK. *Computer-Graphic Facial Reconstruction*. Elsevier, 2005.
178. IOS Press Ebooks - Computer-Aided Forensics: Facial Reconstruction, <http://ebooks.iospress.nl/publication/10639> (accessed 7 May 2020).
179. ULLRICH H, STEPHAN CN. On Gerasimov's Plastic Facial Reconstruction Technique: New Insights to Facilitate Repeatability*: GERASIMOV'S PLASTIC FACIAL RECONSTRUCTION TECHNIQUE. *Journal of Forensic Sciences* 2011; 56:470-474.
180. PRAG J, NEAVE R. *Making Faces: Using Forensic and Archaeological Evidence*. Trustees of the British Museum, 1997.
181. GATLIFF BP. Facial sculpture on the skull for identification. *Am J Forensic Med Pathol* 1984; 5: 327-332.
182. WILKINSON C. Forensic Facial Reconstruction. *Cambridge Core*. Epub ahead of print May 2004. DOI: 10.1017/CBO9781107340961.
183. WILKINSON C. Computerized forensic facial reconstruction?: A review of current systems. *Forensic Sci Med Pathol* 2005; 1: 173-177.
184. WILKINSON C, RYNN C, PETERS H, et al. A blind accuracy assessment of computer-modeled forensic facial reconstruction using computed tomography data from live subjects. *Forensic Sci Med Pathol* 2006; 2:179-187.
185. LEE W-J, WILKINSON CM, HWANG H-S. An Accuracy Assessment of Forensic Computerized Facial Reconstruction Employing Cone-Beam Computed Tomography from Live Subjects. *Journal of Forensic Sciences* 2012; 57: 318-327.
186. WILKINSON C, RYNN C. *Craniofacial Identification*. Cambridge University Press, 2012.
187. EVENHOUSE RJ, RASMUSSEN M, SADLER LL. Computer-aided forensic facial reconstruction. In: *Biostereometric Technology and Applications*. International Society for Optics and Photonics, pp.147-156.
188. CLAES P, VANDERMEULEN D, DE GREEF S, et al. Craniofacial reconstruction using a combined statistical model of face shape and soft tissue depths: Methodology and validation. *Forensic Science International* 2006; 159: S147-S158.
189. CLAES P, VANDERMEULEN D, DE GREEF S, et al. Computerized craniofacial reconstruction: Conceptual framework and review. *Forensic Science International* 2010; 201: 138-145.
190. PARKS CL, RICHARD AH, MONSON KL. Preliminary performance assessment of computer automated facial approximations using computed tomography scans of living individuals. *Forensic Sci Int* 2013; 233: 133-139.

191. QUATREHOMME G, COTIN S, SUBSOL G, et al. A fully three-dimensional method for facial reconstruction based on deformable models. *J Forensic Sci*1997;42: 649-652.
192. GUYOMARC'H P, DUTAILLY B, CHARTON J, et al. Anthropological Facial Approximation in Three Dimensions (AFA3D): Computer-Assisted Estimation of the Facial Morphology Using Geometric Morphometrics. *Journal of Forensic Sciences*2014;59: 1502-1516.
193. DE BUHAN M, NARDONI C. A facial reconstruction method based on new mesh deformation techniques. *Forensic Sci Res* 2018; 3:256-273.
194. VANDERMEULEN D, CLAES P, LOECKX D, et al. Computerized craniofacial reconstruction using CT-derived implicit surface representations. *Forensic Sci Int*2006;159 Suppl1: S164-174.
195. IMAIZUMI K, TANIGUCHI K, OGAWA Y, et al. Development of three-dimensional facial approximation system using head CT scans of Japanese living individuals. *Journal of Forensic Radiology and Imaging*2019;17: 36-45.
196. LIU C, LI X. Superimposition-guided Facial Reconstruction from Skull. *arXiv:181000107 [cs]*, <http://arxiv.org/abs/1810.00107> (2018, accessed 7 May 2020).
197. Development of three-dimensional facial approximation system using head CT scans of Japanese living individuals - *S c i e n c e D i r e c t* , <https://www.sciencedirect.com/science/article/pii/S2212478019300152> (accessed 7 May 2020).
198. TU P, BOOK R, LIU X, et al. Automatic Face Recognition from Skeletal Remains. In: *2007 IEEE Conference on Computer Vision and Pattern Recognition*. 2007, pp. 1-7.
199. CLAES P, LIBERTON DK, DANIELS K, et al. Modeling 3D Facial Shape from DNA. *PLOS Genetics*2014;10: e1004224.
200. *Molecular Photofitting*. Elsevier. Epub ahead of print 2008. DOI: 10.1016/B978-0-12-088492-6.X5001-4.
201. KAYSER M, SCHNEIDER PM. DNA-based prediction of human externally visible characteristics in forensics: Motivations, scientific challenges, and ethical considerations. *Forensic Science International: Genetics*2009;3: 154-161.
202. SULEM P, GUDBJARTSSON DF, STACEY SN, et al. Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nat Genet*2007;39: 1443-1452.
203. KASTELIC V, PO?PIECH E, DRAUS-BARINI J, et al. Prediction of eye color in the Slovenian population using the IrisPlex SNPs. *Croat Med J*2013;54: 381-386.
204. PNEUMAN A, BUDIMLIJA ZM, CARAGINE T, et al. Verification of eye and skin color predictors in various populations. *Legal Medicine*2012;14: 78-83.
205. WALSH S, LIU F, BALLANTYNE KN, Et al. IrisPlex: A sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. *Forensic Science International: Genetics*2011;5: 170-180.
206. WALSH S, WOLLSTEIN A, LIU F, et al. DNA-based eye colour prediction across Europe with the IrisPlex system. *Forensic Science International: Genetics*2012;6: 330-340.
207. SPICHENOK O, BUDIMLIJA ZM, MITCHELL AA, et al. Prediction of eye and skin color in diverse populations using seven SNPs. *Forensic Science International: Genetics*2011;5: 472-478.
208. MAROÑAS O, PHILLIPS C, SÖCHTIG J, et al. Development of a forensic skin colour predictive test. *Forensic Sci Int Genet*2014; 13: 34-44.
209. WANG X, PENG Y, LU L, et al. TieNet: Text-Image Embedding Network for Common Thorax Disease Classification and Reporting in Chest X-Rays. In: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2018, pp. 9049-9058.
210. BRINKER TJ, HEKLER A, ENK AH, et al. Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task. *European Journal of Cancer*2019;113: 47-54.
211. TCIA Collections - The Cancer Imaging Archive (TCIA), <https://www.cancerimagingarchive.net/collections/> (accessed 7 May 2020).
212. LEIN ES, HAWRYLYCZ MJ, AO N, et al. Genome-wide atlas of gene expression in the adult mouse brain. *Nature* 2007; 445: 168-176.
213. MESEJO P, PIZARRO D, ABERGEL A, et al. Computer-Aided Classification of Gastrointestinal Lesions in Regular Colonoscopy. *IEEE Transactions on Medical Imaging* 2016; 35: 2051-2063.
214. ADNI | Alzheimer's Disease Neuroimaging Initiative, <http://adni.loni.usc.edu/> (accessed 7 May 2020).
215. POLDRACK RA, BARCH DM, MITCHELL JP, et al. Toward open sharing of task-based fMRI data: the OpenfMRI project. *Front Neuroinform*; 7. Epub ahead of print 8 July 2013. DOI: 10.3389/fninf.2013.00012.
216. DRESSEL J, FARID H. The accuracy, fairness, and limits of predicting recidivism. *Science Advances*2018;4: eaao5580.
217. New Mexico Decedent Image Database - Welcome Home, <https://nmdid.unm.edu/> (accessed 7 May 2020).