



Benemérita Universidad Autónoma de Puebla

FACULTAD DE CIENCIAS DE LA ELECTRÓNICA

MAESTRÍA EN INGENIERÍA ELECTRÓNICA, OPCIÓN
INSTRUMENTACIÓN ELECTRÓNICA

Tesis para obtener el grado de:

MAESTRO EN INGENIERÍA ELECTRÓNICA

Implementación de un sistema automático para la detección de “landmarks” en imágenes radiográficas

Presenta:

Jamid Israel Saenz Giron*

Asesores:

Dr. Salvador Eugenio Ayala Raggi

Dr. Aldrin Barreto Flores

Dr. Leopoldo Altamirano Robles

Becario CONACYT*

Puebla, Pue. Enero 2021

Resumen

La localización automática de *landmarks* en imágenes médicas ha ganado interés en los últimos años, debido a que procesos posteriores se han visto beneficiados de una correcta localización. Un ejemplo de aplicación importante es la estimación de la edad ósea, ya que al marcar *landmarks* en regiones que presentan un mayor cambio de osificación permiten obtener una mejor estimación de dicha edad.

Por lo tanto, en esta tesis se presenta un nuevo método totalmente automático para la localización de *landmarks* en imágenes radiográficas de manos, el cual es robusto a los cambios en escala, traslación, postura angular, contraste de las imágenes radiográficas y artefactos que se encuentren alrededor de la región de la mano. Se propone un método de estimación de la posición de la mano en una imagen radiográfica, basado en una mejora iterativa del cálculo del centro de masa. Se propone, además, un algoritmo basado en cambios de densidad para estimar el tamaño de la mano en la imagen. De igual forma, se propone un método que corrige el contraste de las radiografías de manos permitiendo realizar una mejor segmentación de la mano y encontrar de manera correcta la región de la mano. Así mismo, se presenta un método fundado en la dirección de los gradientes de la imagen para realizar la estimación de la postura angular de la mano. Igualmente, se presenta un método que localiza en su totalidad las *landmarks* a pesar de que alrededor de la región de la mano se encuentren artefactos que pudieran imposibilitar dicha localización, esto se logra al encontrar de manera correcta la región que solo contenga la mano. Finalmente, se presenta un enfoque multiescala, que utiliza *eigenespacios*, para realizar la localización automática de las *landmarks*. Los resultados experimentales demuestran que el método propuesto consigue un error similar al de los métodos actuales, pero con la ventaja de la robustez a las variaciones en posición, postura angular, escala, contraste de la región de interés y artefactos alrededor de la región de la mano.

Agradecimientos

A Dios por permitirme llegar hasta este punto de mi vida.

A mi mamá Alicia Girón y mi hermano Brayan Girón por todo su apoyo, paciencia y quienes desde siempre han estado conmigo en los buenos y malos momentos.

A mi papá Tomás Saénz por todo su apoyo, que ha pesar de sus dificultades siempre me lo pudo brindar.

A mi asesor de tesis el Dr. Salvador Eugenio Ayala Raggi por todas sus enseñanzas, consejos y paciencia otorgados durante la maestría.

A mis revisores de tesis: la M.C. Ana María Rodríguez Domínguez, el M.C. Rodrigo Lucio Maya Ramírez y el M.C. Nicolás Quiroz Hernández por todas sus observaciones y comentarios que permitieron mejorar mi trabajo de tesis.

A todos los profesores que conforman el posgrado; por brindarme sus conocimientos y consejos que me ayudaron a crecer tanto personal como académicamente.

A mi compañero Mauricio por toda su ayuda, además de compartir conmigo sus conocimientos que fueron muy valiosos en esta etapa. También a todos mis compañeros de la maestría por su apoyo y buenos momentos que pasamos juntos.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo económico brindado durante el periodo agosto 2018 a julio 2020.

A la Benemérita Universidad Autónoma de Puebla (BUAP), Facultad de Ciencias de la Electrónica (FCE) que me permitió desde la licenciatura y ahora en la maestría el curso de mis estudios profesionales.

Finalmente a todos los profesores que desde niño hasta este nivel me motivaron y ayudaron en continuar mis estudios.

Índice general

Resumen	I
Índice general	III
Índice de figuras	V
Índice de tablas	VIII
Lista de acrónimos	IX
1. Introducción	1
1.1. <i>Landmarks</i>	1
1.2. Ejemplo de una aplicación importante: Estimación de la Edad Ósea	3
1.3. Método de reconocimiento de características	4
1.3.1. <i>Eigenfaces</i>	5
1.4. Problemática a resolver	5
1.5. Objetivos	6
1.5.1. Objetivo General	6
1.5.2. Objetivos Específicos	6
1.6. Estructura de la tesis	7
2. Trabajo relacionado	8
3. Descripción del método	17
3.1. Descripción general del sistema	17
3.2. Descripción General del proceso de Alineación de Imagen	19
3.3. Alineación de la imagen	19
3.3.1. Segmentación de la mano	19
3.3.1.1. Detección del centro geométrico de la mano	22
3.3.1.2. Obtención de la región de la mano	25
3.3.1.3. Corrección de contraste	29
3.3.1.4. Corrección angular de la imagen de la mano	30
3.4. Etapa de entrenamiento	34

3.5. Etapa de evaluación	36
3.5.1. Búsqueda de la <i>landmark</i> en diferentes niveles de resolución	36
3.5.1.1. Proyección de parches en el subespacio de las <i>eigenfaces</i>	38
3.5.2. Métricas de evaluación	39
4. Resultados experimentales	41
4.1. Interfaz de usuario desarrollada	44
4.1.1. Descripción de cada elemento de la interfaz	48
4.2. Discusión de resultados	49
5. Conclusiones y trabajo futuro	50
5.1. Conclusiones	50
5.2. Trabajo futuro	51
A. Artículos publicados	52
Bibliografía	74

Índice de figuras

1.1. Los puntos con el número 1, corresponden a las <i>landmarks</i> Tipo I, los puntos con el número 2, pertenecen al Tipo II y con el número 3 al Tipo III (Reproducida de [1])	3
1.2. Visión general de la estimación de la edad ósea.	4
1.3. Al tener un conjunto grande de imágenes se reduce al aplicar PCA y obtener un subespacio lineal $\mathbf{Q}$, en el cual se puede reconstruir una imagen $\bar{\mathbf{X}}$ al proyectarlo en $\mathbf{Q}$ y sumarle la imagen promedio $\bar{\Psi}$	5
2.1. Ejemplos de imágenes médicas marcadas con <i>landmarks</i> , donde el a) corresponde a una imagen cefálica [2], b) cerebro [3], c) rodilla [4] y d) fémur y pelvis [5].	8
2.2. En el a) se observa la segmentación de los huesos de la mano con el algoritmo <i>Watershed</i> y en b) la localización de <i>landmarks</i> en la misma imagen (Reproducida de [6]).	9
2.3. Resultados donde se ubican y comparan la localización de <i>landmarks</i> , de forma manual y automática en la tesis de [7] (Reproducida de [7]).	10
2.4. Puntos localizados en [8] a) imagen 2D y b) imagen 3D (Reproducida de [8]).	11
2.5. En la parte izquierda se observan las <i>landmarks</i> localizadas en [9]. Los puntos rojos son los inicializados y los de color verde los encontrados por su algoritmo. También en la parte derecha se muestran los modelos de forma que obtienen en el entrenamiento y los ocupados para su evaluación (Reproducida de [9]).	12
2.6. De lado izquierdo se muestra una imagen cefálica lateral y en la parte derecha una radiografía con puntos localizados en [10] (Reproducida de [10]).	13
2.7. Diagrama general de [11], que muestra su arquitectura llamada Spatial Configuration-Net (Reproducida de [11]).	14
2.8. Ejemplos de imágenes marcadas en [12], donde a) es una radiografía de mano, b) MRI de mano, c) imagen cefálica lateral y d) espina dorsal (Reproducida de [12]).	14
3.1. Descripción general de la localización automática de <i>landmarks</i> en el tercer hueso metacarpiano	18

3.2.	A partir de $I(m, n)$ se modificaron sus dimensiones, donde la dimensión menor es de 256 píxeles y la segunda a un valor p . Después se localizó el baricentro b_n y la región de la mano para obtener el radio r_c . Lo anterior se reubicó en la imagen $I(m, n)$. Se continuó aplicando un corte dentro de $I(m, n)$ creando $C(m, n)$ y cambiando sus dimensiones a 256×256 píxeles obteniendo $F(m, n)$. Por último se realizó una corrección de contraste $I^g(m, n)$ y postura angular obteniendo $I^s(m, n)$	21
3.3.	Se detectaron los bordes de la imagen de entrada $I(m, n)$ obteniendo la imagen $E_1(m, n)$ y localizar el baricentro (m_{b1}, n_{b1}) . Después se corrigió el contraste de $I(m, n)$. Finalmente se detectaron los bordes de la imagen anterior con el Algoritmo de <i>Canny</i> y obtener una nueva imagen $E_2(m, n)$ que corresponde a los bordes de los huesos de la mano.	23
3.4.	Localización del baricentro en la región de la mano. Se inició localizando el punto en la imagen $E_2(m, n)$ y siendo este el centro se dibujó una circunferencia para ajustar la ubicación del baricentro. Así el punto (m_{b2}, n_{b2}) representa el baricentro anteriormente ubicado y (m_{b3}, n_{b3}) el nuevo, notándose como se va reduciendo el área de la circunferencia.	23
3.5.	Creación de las coronas circulares con la imagen de bordes $E(m, n)$, la cual se inició con el baricentro b_n siendo el centro y comenzar con el radio $\mathcal{M}_{E1}$ para formar el primer elemento H_1 , se continuó con las coronas formadas por $H_2 \dots H_{18}$	28
3.6.	Gráfica con dos densidades que corresponden a la mano más chica y más grande del conjunto utilizado. Por lo que se pudo determinar el valor del umbral igual a 0.2 para la selección de las circunferencias en las demás imágenes.	28
3.7.	Corrección de contraste de la imagen $I(m, n)$ al tomar una región $W(j, k)$ y obtener los valores W_{min} y W_{mas} respectivos a su histograma, obteniendo la imagen $I^c(m, n)$	30
3.8.	Se realizó la sumatoria ponderada de los vectores comprendidos entre una ventana de i a $i + 90^\circ$ a), obteniendo los vectores ponderados b). Este proceso se realiza por todo el plano polar (para $i = 1, 2, \dots, 360$), obteniendo así a c).	31
3.9.	A partir de b_f se dividió en dos mitades la región de la mano siendo $C_1(m, n)$ y $C_2(m, n)$ respectivamente. Para cada mitad $C_{1,2}(m, n)$, en el primer píxel (m, n) se tomó una ventana $P(m, n)$ y si el píxel (m, n) es mayor al valor $\bar{g}$, entonces se obtiene el valor p que corresponde a la sumatoria de los valores de gris contenidos en $P(m, n)$	32
3.10.	Se marcaron de manera manual las dos <i>landmarks</i> en la parte superior e inferior del tercer hueso metacarpiano para crear parches en los 3 niveles de resolución de búsqueda y crear con ellos los <i>eigenespacios</i>	34

3.11. La imagen $\Psi(m, n)$ corresponde a la mano promedio obtenida al aplicar PCA y ubicar los puntos marcados manualmente, para obtener los límites: derecho, izquierdo, superior e inferior, y crear las regiones de búsqueda para la parte superior a) y para la parte inferior b).	35
3.12. Al localizar la posición de la <i>landmark</i> l_3 en la imagen $I_3^s(m, n)$ se redimensionó el punto en la imagen I_2^s obteniendo una región de búsqueda $R_2(m, n)$, agregando dos píxeles a la región siendo de dimensiones 6×6 píxeles repitiendo la búsqueda y localizar el punto l_2 , el cual se redimensionó en la imagen $I_1^s(m, n)$ obteniendo una nueva región de búsqueda $R_1(m, n)$, también agregando dos píxeles y repitiendo la búsqueda del punto en $R_1(m, n)$, para localizar la <i>landmark</i> final l_1	37
3.13. A partir de $R(m, n)$ (como ejemplo se tienen 54 píxeles), se inicia en el píxel 1 y siendo éste el centro se toma un parche que se convierte en un vector columna Φ proyectándolo en su <i>eigenespacio</i> $\mathbf{Q}$ correspondiente, obteniendo el parche reconstruido $\hat{\Phi}$ y calcular la norma e entre los dos vectores la cuál se guarda en el mapa de errores E para cada píxel de $R(m, n)$. Una vez reconstruida $R(m, n)$ se busca la posición mínima en E que corresponderá a la <i>landmark</i> l	38
3.14. Interpolación de la longitud del tercer hueso metacarpiano en hombres a) y mujeres b), a partir de los datos proporcionados en [13].	40
4.1. Histograma de las imágenes utilizadas en el conjunto de evaluación, mostrando la cantidad y los rangos de edad utilizados.	41
4.2. Ejemplos de la alineación de la imagen y localización automática de <i>landmarks</i> en diferentes radiografías de manos.	43
4.3. Comparación de <i>landmarks</i> marcadas manual y automáticamente.	44
4.4. Diagrama con los casos de uso para la localización automática de <i>landmarks</i>	44
4.5. Diagrama con los casos de uso para el cálculo de la longitud del tercer hueso metacarpiano.	45
4.6. Interfaz que se desarrolló para mostrar la localización automática de <i>landmarks</i> y calcular la longitud del tercer hueso metacarpiano.	48

Índice de tablas

2.1. Comparación de métodos para ubicar <i>landmarks</i> en distintas imágenes de manos.	16
4.1. Comparación del desempeño de los métodos para localización de <i>landmarks</i> investigados en el estado del arte. Aunque el método desarrollado en esta tesis tiene un error de localización ligeramente mayor al reportado por Lindner et al [9], se tiene que hacer notar que nuestro método para obtener dicho error es mucho más preciso y confiable, puesto que sí se tomó en cuenta la variación en el tamaño de los huesos en función de la edad, mientras que en [9] se supuso un tamaño fijo de ancho de muñeca (50mm) para todas las imágenes, a pesar de haber utilizado un rango amplio de edades.	43
4.2. Caso de uso de Seleccionar imagen.	45
4.3. Caso de uso de Alineación de la imagen.	46
4.4. Caso de uso de Localización automática de <i>landmarks</i>	46
4.5. Caso de uso de Ingresar la edad ósea.	47
4.6. Caso de uso de Seleccionar sexo.	47
4.7. Caso de uso de Calcular la longitud del tercer hueso metacarpiano.	47

Lista de acrónimos

CNNs Convolutional Neural Networks (Redes neuronales convolucionales).

KNN K-Nearest Neighbors (K-vecinos más cercanos).

MIP Medical Image Processing (Procesamiento de imágenes médicas).

MRI Magnetic Resonance Imaging (Imagen de resonancia magnética).

PCA Principal Component Analysis (Análisis de componentes principales).

RF Random Forests (Bosques aleatorios).

Capítulo 1

Introducción

Con la invención de las imágenes digitales se ha podido tener una mejor apreciación y análisis de lo que sucede en el entorno o en el interior de los seres vivos. Especialmente en el cuidado, prevención y tratamiento de los seres humanos. Así, es inevitable que se busquen aplicaciones en el área médica como son las radiografías, tomografías, imágenes de resonancia magnética (MRI), ultrasonidos, etc. Con ello se da paso al procesamiento de imágenes médicas (MIP) [14] que ha permitido crear múltiples métodos computacionales para su análisis y brindar ayuda al médico al proporcionar información de gran utilidad de acuerdo a su área de especialidad. El procesamiento permite al médico mostrar la información más relevante y clara posible de una manera conveniente. Por tal motivo se han diseñado diversos algoritmos que aplican un pre-procesamiento al conjunto de imágenes a usar como son: segmentación, registro, etc, con el fin de poder tener una mejor vista y estudio. También los procesos pueden ser semi-automáticos o automáticos. Sin embargo, debido a la gran variedad de características que tienen, como es: la forma, contraste, color, entre otros, hace que sea muy difícil crear un solo método que permita realizar un análisis para todos los tipos de imágenes, por lo que se han creado diversas áreas que se especializan en las diferentes propiedades de las imágenes. Es por esta razón que una sub-área del MIP como es la localización automática de *landmarks* ha ganado interés en investigación en los últimos años, ya que procesos posteriores se han podido beneficiar de una correcta y confiable localización. Es en este tema en el que se centrará esta tesis.

1.1. *Landmarks*

La detección automática de *landmarks* o puntos clave de interés es un campo muy importante tanto en interpretación de rostros, como en análisis de imágenes biológicas y radiológicas, pues posee básicamente dos tipos de utilidad:

1. Ubicar puntos de interés específicos que ayudan a localizar o enmarcan zonas de interés en imagenología médica.

2. La morfología en la estructura subyacente en el conjunto de puntos, puede aportar información.

Las *landmarks* permiten capturar los contornos o puntos de referencia de una estructura, por lo que en sus inicios tomó relevancia en la morfología de los seres vivos, que es uno de los estudios más importantes en el área de la biología, ya que se puede analizar diferencias entre la misma o diferentes especies.

Por lo que a principios del siglo XX se desarrolló la morfometría tradicional que corresponde “al estudio cuantitativo de la variación de las formas biológicas” [15], donde se utilizaban variables lineales como distancias y ángulos. Después se obtenían conjuntos de datos que al analizarlos se mostraban en gráficas, pero el tamaño y forma eran bastante complejos de interpretar.

Debido a las dificultades de la morfometría tradicional, a finales del siglo XX se creó la morfometría geométrica definida como “una fusión empírica de la geometría con la biología” [15], ya que se analiza a los organismos tanto en su forma como en sus estructuras, además de considerar su espacio geométrico empleando métodos estadísticos.

En morfometría un concepto importante es la “forma”, que es una propiedad geométrica de un objeto y no toma en cuenta la escala, rotación y traslación [15, 16]. Aunque no se puede analizar la forma y tamaño por separado, debido a que dichas características difieren en los organismos, a pesar de pertenecer a la misma especie, lo cual es causado por crecimiento y desarrollo natural de cada organismo.

La morfometría geométrica utiliza las siguientes fuentes de información para su análisis:

- Homología biológica.
- La localización geométrica.

La primera realiza una conexión biológica de partes de los seres vivos, mientras que la segunda trata de la configuración espacial en dos o tres dimensiones de estructuras o partes [17].

Para localizar estructuras homólogas, se utilizan dos variables: *Outlines* o contornos y *landmarks*.

Las *landmarks* son puntos en un espacio tanto 2D o 3D ubicado en una posición de un rasgo en particular de acuerdo a un objeto de interés, por lo tanto, son puntos anatómicos que no alteran su posición topológica con respecto a otras *landmarks*, contienen información de la forma y de acuerdo a la experiencia se pueden ubicar fácil y rápidamente de un organismo a otro.

Las *landmarks* en morfometría geométrica, son utilizadas para cuantificar la forma, variación y covarianza biológica, junto con otras variables bióticas o abióticas.

Las *landmarks* ideales cumplen con los siguientes requisitos [18]:

1. Proveen una adecuada cobertura morfológica.
2. Pueden ser identificadas rápidamente y con precisión.

3. Deben ubicarse en el mismo plano.

Se pueden clasificar 3 tipos de *landmarks*:

- Tipo I: corresponden a yuxtaposiciones entre los tejidos, además de ser respaldadas por un significado biológico.
- Tipo II: se ubican en zonas donde existe una máxima o mínima curvatura.
- Tipo III: Estos son localizados a lo largo de los contornos o puntos externos.

En la Figura 1.1 se visualizan las 3 tipos de *landmarks*, en una hoja.

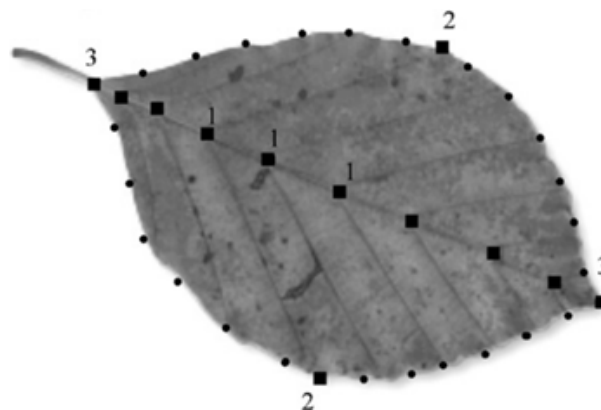


Figura 1.1. Los puntos con el número 1, corresponden a las *landmarks* Tipo I, los puntos con el número 2, pertenecen al Tipo II y con el número 3 al Tipo III (Reproducida de [1])

Con respecto al uso de ubicar *landmarks* en imágenes médicas se busca extraer información que contenga un significado específico para el médico. De manera que existen diversas radiografías de diferentes partes del cuerpo que el doctor puede marcar con *landmarks* como son radiografías de: rodillas, cráneos, dentaduras, manos, etc. Esto es con el fin de que en un proceso posterior se pueda extraer información alrededor de cada punto que sea de utilidad como puede ser la estimación de la edad ósea.

1.2. Ejemplo de una aplicación importante: Estimación de la Edad Ósea

La estimación de la madurez esquelética también conocida como edad ósea, es un procedimiento clínico el cual es usado para determinar la edad o desarrollo esquelético tanto en niños como en adultos jóvenes. Ya que resuelve problemas de estatura en niños inusualmente bajos. También estudia las alteraciones de la pubertad como el adelanto o retraso de esta misma.

La estimación de la edad ósea, normalmente hecha por un médico radiólogo, es de extrema utilidad para médicos en otras especialidades, como pediatras o endocrinólogos infantiles.

Así uno de los problemas más importantes es que la edad esquelética no coincide con la edad cronológica, debido a que en ocasiones existen trastornos en el desarrollo del paciente y la detección oportuna de una diferencia drástica entre la edad esquelética y la edad cronológica puede ser de gran utilidad para la administración de algún tratamiento. Por lo tanto, se han desarrollado métodos manuales para la estimación de la edad ósea [19, 20]. Con esta información el médico podría elegir los medicamentos que tendría que tomar el paciente de acuerdo a su diagnóstico. Sin embargo, no siempre existe homogeneidad en diagnósticos hechos por distintos médicos para el mismo paciente, es decir para la misma imagen. Esta subjetividad puede ser eliminada al utilizar métodos automáticos basados en aprendizaje de máquina capaces de determinar la edad ósea [21].

En el caso de la edad ósea, se ha visto que no todas las zonas dentro de la imagen radiológica muestran cambios con la edad. Por tanto, resulta mucho más eficiente trabajar únicamente con aquellas regiones que presentan diferencias en apariencia relacionadas con el cambio en la edad ósea. La localización automática de *landmarks* permite justamente segmentar dichas regiones de interés. Por lo que la localización de *landmarks* se ha convertido en un paso obligado para encontrar regiones de interés y determinar la edad ósea del paciente [22]. Este proceso se ilustra en la Figura 1.2.

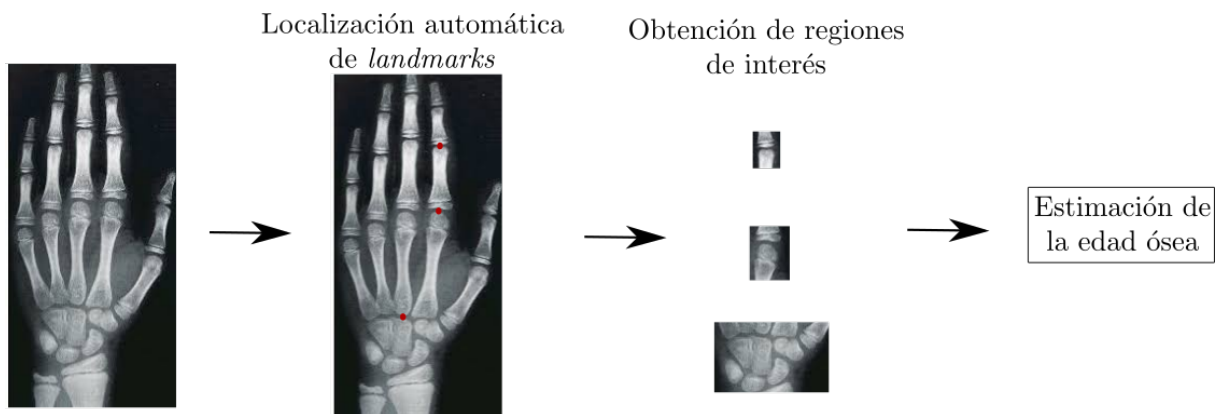


Figura 1.2. Visión general de la estimación de la edad ósea.

1.3. Método de reconocimiento de características

Para saber que zonas se desean marcar con las *landmarks* se deben aplicar algoritmos que sean capaces de extraer características que permitan reconocer en donde se encuentra el área de interés. Para ello se presenta el siguiente método.

1.3.1. *Eigenfaces*

El método de las *eigenfaces* fue creado por Sirovich y Kirby en [23] y Turk y Pentland en [24] el cual hace uso del análisis por componentes principales PCA. Tiene como objetivo reducir la dimensionalidad del conjunto de imágenes utilizado y representar los datos originales en una dimensión mucho menor a la original [25], debido a que, con el conjunto que se trabaja cuenta con demasiadas variables a analizar, además de conseguir las características base de todo el conjunto. El método busca combinaciones lineales, de las variables originales que puedan representar la varianza de la mejor manera posible, esto quiere decir que teniendo pocas combinaciones lineales (componentes principales), se puede entender gran parte de la información en el conjunto de datos. Ver Figura 1.3.

Para generar las *eigenfaces* se requiere tener una considerable cantidad de imágenes que tengan las mismas condiciones como son: tener una pose del rostro frontal, una iluminación uniforme con respecto al de los demás rostros, básicamente que estén “normalizadas”. Lo cual resulta un gran problema en el uso de esta técnica. Por lo tanto, es necesario modificar las características de las imágenes de manera que sean similares entre ellas para obtener un buen desempeño de este proceso.

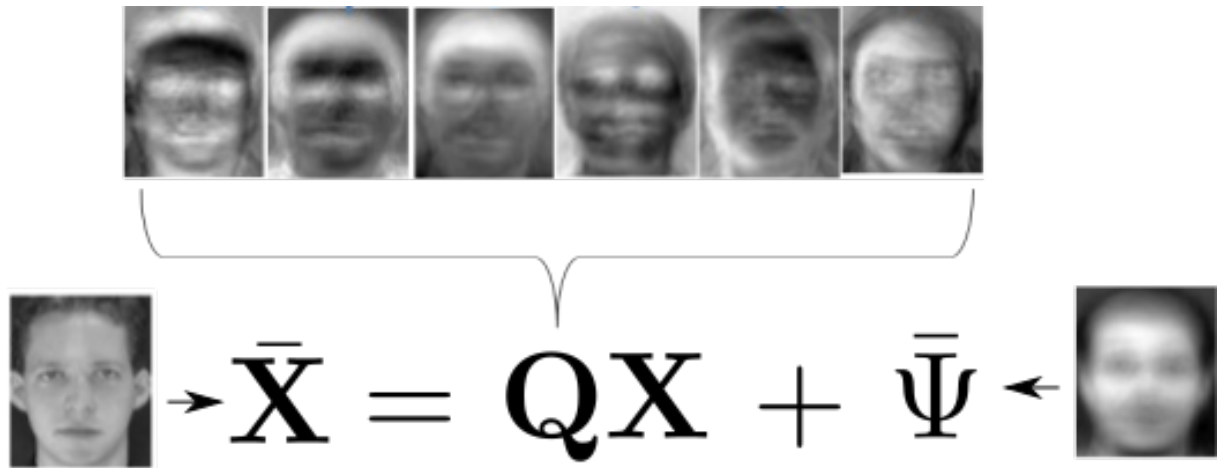


Figura 1.3. Al tener un conjunto grande de imágenes se reduce al aplicar PCA y obtener un subespacio lineal Q , en el cual se puede reconstruir una imagen $\bar{X}$ al proyectarlo en Q y sumarle la imagen promedio $\bar{\Psi}$.

1.4. Problemática a resolver

Como se mencionó anteriormente los puntos son marcados manualmente por un médico o experto, lo cual causa que sea además de costoso, monótono y cansado. Debido a que en ocasiones se busca que las *landmarks* sean ubicadas en grandes cantidades de conjuntos de imágenes. Esto puede causar a que se tengan grandes posibilidades de errores al localizarlos. Por tanto, resulta imperativo dar atención a la investigación y desarrollo

de técnicas para la detección precisa de puntos de interés.

El problema de la localización automática de *landmarks* ha sido extensamente estudiado en rostros [26]. Por lo que ¿Es posible adaptar los algoritmos de localización automática de *landmarks* en rostros a las imágenes radiográficas de manos?, ¿Es necesario realizar un pre-procesamiento al conjunto de imágenes médicas obtenidas para la localización de *landmarks*?, ¿Qué zonas se deben marcar con *landmarks* en los huesos de la mano?.

Por tal motivos en esta tesis se propone diseñar un algoritmo implementado en *Python* que encuentre un conjunto de *landmarks* o puntos clave de forma automática, con aplicación y utilización directa en otro sistema para la estimación de la edad ósea.

1.5. Objetivos

1.5.1. Objetivo General

Implementación de un sistema automático de detección de *landmarks* en imágenes radiográficas de manos, con utilidad en la estimación automática de la edad ósea.

1.5.2. Objetivos Específicos

- Estudiar métodos paramétricos y no paramétricos utilizados para detección de *landmarks* en rostros y evaluar su aplicabilidad en imágenes radiográficas.
- Plantear un método de localización automática de *landmarks* basada en 2 fuentes de información: la posición individual de cada *landmark* y su entorno local, y el patrón existente en las posiciones relativas entre sí de todas las *landmarks*.
- Diseñar una disposición o *layout* de puntos o *landmarks* sobre la imagen radiográfica de la mano no dominante que incluya las regiones entre huesos del metacarpo y falanges proximales, así como la región de los huesos del carpo.
- Proponer la utilización de características invariantes a la rotación y al escalamiento y aplicar dichas características en el desarrollo del algoritmo de localización automática de *landmarks*.
- Ubicar y regresar regiones alrededor de las *landmarks*, de tal manera que estén normalizadas en escala y rotación.
- Implementar en *Python* el algoritmo de detección automática de *landmarks*.
- Desarrollar en *Python* una interfaz de usuario adecuada para su utilización por un médico radiólogo.

1.6. Estructura de la tesis

En este capítulo se presentó la introducción, la problemática presentada, además de los objetivos que se tienen para la conclusión de la tesis.

En el capítulo 2 se describen los trabajos relacionados acerca de la localización automática de *landmarks* tanto en imágenes digitales como son: radiografías e imágenes de resonancia magnética MRI.

En el capítulo 3 se presenta la descripción del método. Se detallan los pasos para lograr la localización automática de *landmarks*.

En relación con el capítulo 4 se muestran los resultados experimentales y el diseño de la interfaz gráfica.

Finalmente el capítulo 5 contiene las conclusiones y trabajo futuro de acuerdo a los resultados obtenidos en el capítulo anterior.

Capítulo 2

Trabajo relacionado

La localización automática de *landmarks* es utilizada en la búsqueda de regiones anatómicas, lo cual es ampliamente utilizado en el campo del diagnóstico asistido por computadora [27]. Por ello existen distintos métodos que localizan puntos de interés en imágenes radiográficas o 2D y en MRI o 3D. Esto se realiza para diferentes zonas del cuerpo humano como son en imágenes cefálicas laterales [2], donde el autor hace uso de su algoritmo en otro banco de imágenes 2D (moscas de fruta y peces cebra). También se buscan marcar *landmarks* en imágenes radiográficas de fémur y pelvis [5], cerebro [3], rodilla [4], entre otros, como se observa en la Figura 2.1. Pero debido a la gran diferencia de una imagen a otra resulta ser un problema desafiante, por lo que se han tenido que crear diversos métodos en localizar *landmarks* en imágenes médicas.

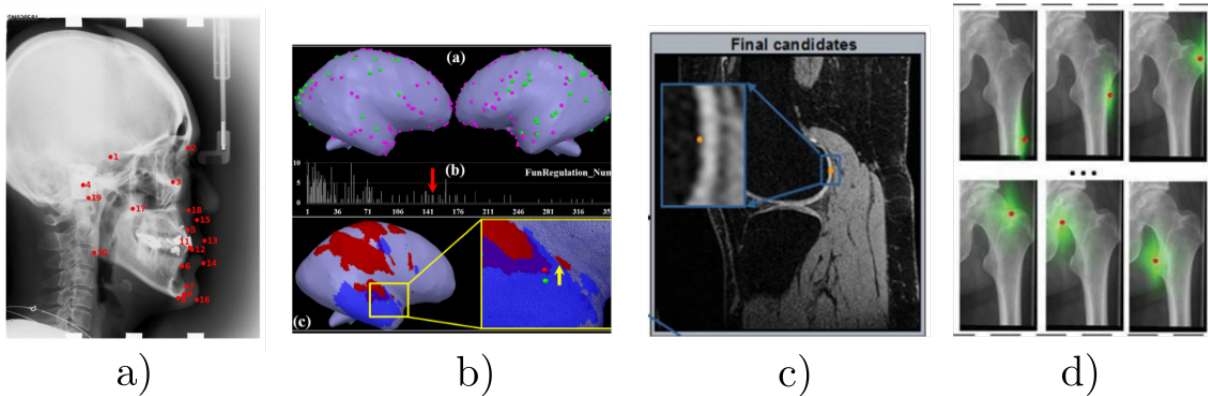


Figura 2.1. Ejemplos de imágenes médicas marcadas con *landmarks*, donde el a) corresponde a una imagen cefálica [2], b) cerebro [3], c) rodilla [4] y d) fémur y pelvis [5].

Se comenzará a profundizar en como localizar los puntos directamente en imágenes médicas de manos. En [28] se obtienen *landmarks* en imágenes MRI, por medio de la técnica de esquemas ponderados, acorde a métodos desarrollados en rostros por el mismo autor. Dicho trabajo utiliza regression random forests (RRF). Tal procedimiento escoge

estructuras, las cuáles fueron elegidas por la distancia de una manera probabilística, donde la posición estimada, se pondera por la distancia a la estimación. Como paso posterior, se obtuvo información adicional, sobre la posición del punto, conectando múltiples pasos de estimación, así las salidas de cada paso individual limita el área para seguir buscando más *landmarks*. Para realizar dichas uniones se deben usar varias etapas RRF, cada una de ellas se encargan de reducir el área de búsqueda.

En [6] se localizan *landmarks* utilizando técnicas de procesamiento digital de imágenes. En primer lugar se aplicó un filtro gaussiano para eliminar el ruido de las imágenes, luego se detectan los bordes para segmentar zonas de interés como son los huesos de la mano, esto se realizó con el algoritmo de *Canny*. El siguiente paso es segmentar los huesos con el algoritmo *Watershed*, ver Figura 2.2, para etiquetar los huesos de acuerdo a un color que se asigna y pasar la imagen a formato binario, de acuerdo a un umbral específico. A continuación se extraen ejes de los dedos mediante el algoritmo de adelgazamiento y se obtiene un esqueleto grueso de la mano formando ramas que corresponden a ejes longitudinales aproximados. Estos se refinan con el fin de posicionarlos en el centro de cada dedo analizando secciones transversales sucesivas. El último proceso es marcar *landmarks* estudiando los cambios de intensidad a lo largo de las líneas, ubicando un total de 20 puntos.

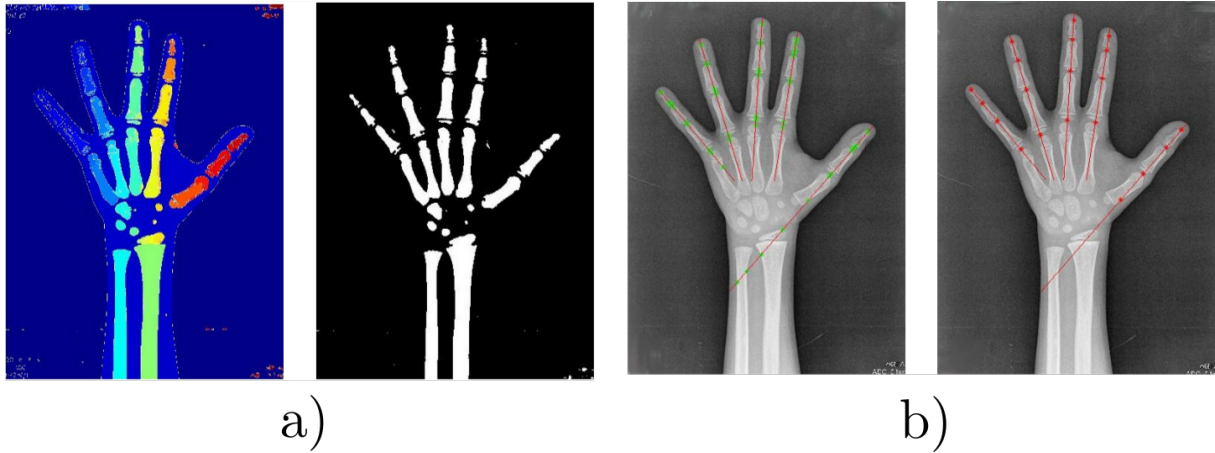


Figura 2.2. En el a) se observa la segmentación de los huesos de la mano con el algoritmo *Watershed* y en b) la localización de *landmarks* en la misma imagen (Reproducida de [6]).

Igualmente en [7], se creó un sistema automático con el propósito de estimar la edad ósea, pero se centrará en el método desarrollado en la localización automática de *landmarks*. El procedimiento para detectar *landmarks*, consiste de tres pasos: entrenamiento, detección inicial y refinamiento; como pasos anteriores se procesa la radiografía para quitar diferencias de contraste, alineación de la forma de la mano y marcas alrededor de la radiografía. Continuando con el primer paso que es el entrenamiento, se generaron conjuntos de parches, los cuales, se utilizaron para determinar la ubicación de *landmarks* en la etapa de evaluación, conjuntamente se crearon “formas” para considerar la zona de los

puntos de acuerdo a su posición relativa y de todo el conjunto de *landmarks*. El siguiente punto es la detección inicial, iniciando con la reducción de las dimensiones en la imagen para comenzar con la búsqueda de puntos, de manera independiente, se van extrayendo regiones para comparar con los conjuntos de entrenamiento y determinar si en el parche se localizó una *landmark*. En la última etapa, de refinamiento, a partir de un conjunto de puntos, se obtuvo una región reducida para encontrar cada *landmark*, este paso se repite hasta volver a las dimensiones originales de la imagen, como se muestra en la Figura 2.3.

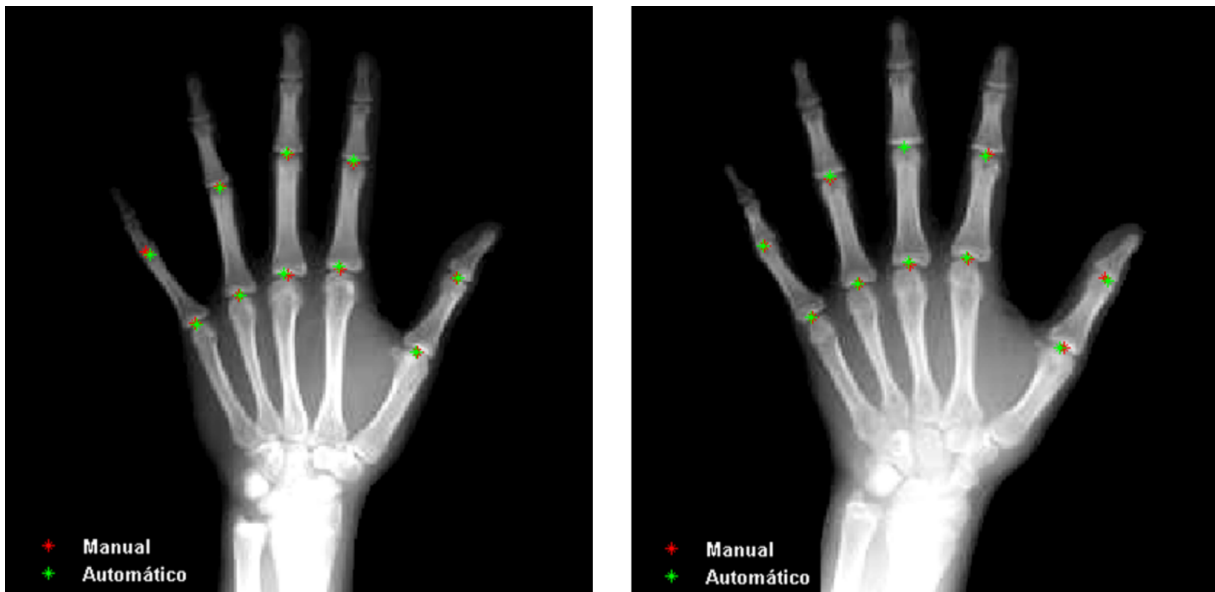


Figura 2.3. Resultados donde se ubican y comparan la localización de *landmarks*, de forma manual y automática en la tesis de [7] (Reproducida de [7]).

En lo que respecta a la localización de puntos de interés en radiografías de manos, en [8] se propuso un algoritmo que localiza un total de 24 *landmarks* en imágenes 2D y 3D, observar la Figura 2.4, el cuál está compuesto por dos etapas una de entrenamiento y la otra de evaluación. Para el primero, los autores generaron un “codebook” que consiste en *landmarks* que marcaron manualmente para crear parches cuadrados alrededor de cada punto y extraer la información sobre el desarrollo de los huesos, esto lo realizaron con el algoritmo PCA. Para cada parche utilizaron diferentes escalas donde la primera s_0 consiste en la imagen completa de la mano en menor resolución, hasta la escala s_7 que corresponde a la zona específica que se desea encontrar el punto de interés y de una resolución mayor a las anteriores.

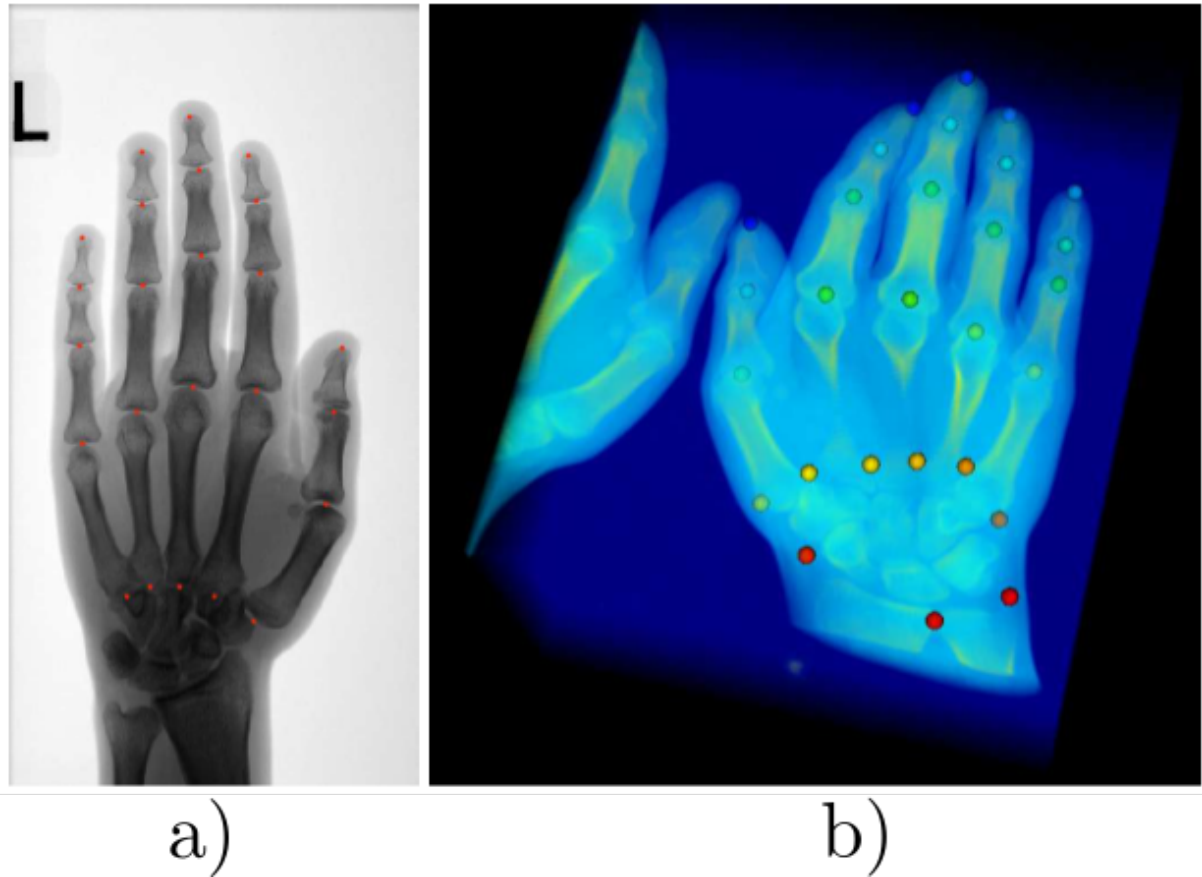
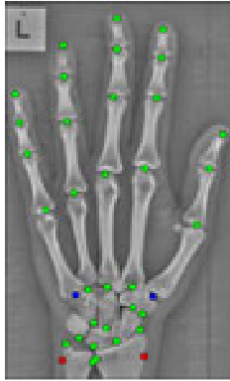


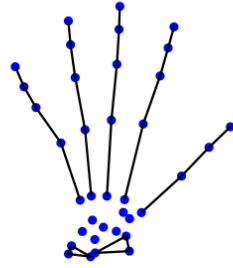
Figura 2.4. Puntos localizados en [8] a) imagen 2D y b) imagen 3D (Reproducida de [8]).

En la etapa de evaluación, se comenzó con la imagen de menor resolución s_0 , donde van extrayendo parches en el centro de la imagen, proyectándolos en el “codebook” creado en el entrenamiento. Dado que no se conoce una referencia de donde se encuentran las *landmarks* hacen uso del algoritmo KNN para comparar el parche con el conjunto de entrenamiento. Así, crean múltiples predicciones para la *landmark* y calculan la *landmark* con la posición media de todas las predicciones. Este proceso se repitió en todas las resoluciones posteriores.

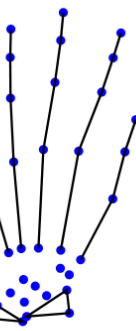
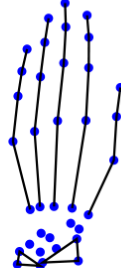
Por otra parte, en [9] se hace uso de distintos algoritmos para localizar las *landmarks* en radiografías de manos e imágenes de caras. Para ello se inicia con el uso del algoritmo llamado Hough Forest [29] para estimar la posición, orientación y escala de la imagen. Así mismo, obtienen la imagen promedio. Con respecto a su entrenamiento, marcaron un total de 37 puntos alrededor de los huesos de la mano en conjunto con 2 modos de forma al aplicar el algoritmo Active Shape Models (ASMs) Figura 2.5, en un total de 564 radiografías. Para la localización automática, primero inicializaron dos puntos en la parte inferior de la mano y con su método llamado Random Forest Regression-Voting and Constrained Local Models (RFRV-CLM) encuentran 9 *landmarks* y siendo ellos nuevas referencias vuelven a ejecutar RFRV-CLM para localizar 37 puntos restantes.



Radiograph



Shape mode 1



Shape mode 2

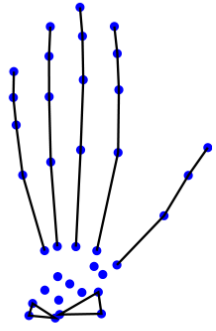


Figura 2.5. En la parte izquierda se observan las *landmarks* localizadas en [9]. Los puntos rojos son los inicializados y los de color verde los encontrados por su algoritmo. También en la parte derecha se muestran los modelos de forma que obtienen en el entrenamiento y los ocupados para su evaluación (Reproducida de [9]).

Siguiendo con el uso del algoritmo Random Forest (RF) en [30] se propuso que la localización de *landmarks* sea de acuerdo a las zonas que están cerca de ellas. Pero está limitado a los objetos que se encuentran cerca de la región de la mano. Para ello, primero en su etapa de entrenamiento usaron Random Regressing Forests (RRF) en dos fases (siguiendo las ideas de [8, 9]). Para la primera en un radio R centrado en la *landmark*, el RRF se entrenó localmente con entidades cuyos rectángulos tienen el tamaño máximo en cada dirección y el desplazamiento máximo. Para la segunda fase realizaron una segmentación de acuerdo al artículo [31] optimizando el tamaño del nodo del bosque y con el uso de back-projection se fue refinando la localización de la *landmark*, ya que proporciona una mejor predicción en los RF. Para su evaluación el punto de interés es calculado por un vector de posición relativa y el píxel evaluado, almacenándolo en una imagen de acumulación. Utilizaron un total de 895 imágenes de manos, localizando un total de 37 puntos.

En años más recientes los autores en [10] siguieron haciendo uso de RF para localizar las *landmarks* en radiografías de manos, pero ahora en toda la imagen teniendo en cuenta donde están las otras *landmarks* ya localizadas. Esto fue posible al tomar la información geométrica de los huesos de la mano buscando en varias regiones de manera simultanea y encontrarlas de manera más exacta. A este método lo nombraron “estrategia de optimización de descenso de coordenadas”, que es un esquema iterativo que divide la búsqueda simultánea de todos los puntos de referencia mediante la optimización sucesiva de puntos individuales. Algo que hacen destacar los autores es el tiempo que tarda en encontrar los puntos siendo de 56 segundos para 43 *landmarks* localizadas. Probaron su algoritmo en otro tipo de conjuntos como son: MRI de cuerpo completo, de dientes, manos y radiografías cefálicas laterales. Se muestran ejemplos de su ubicación de puntos en la Figura 2.6.

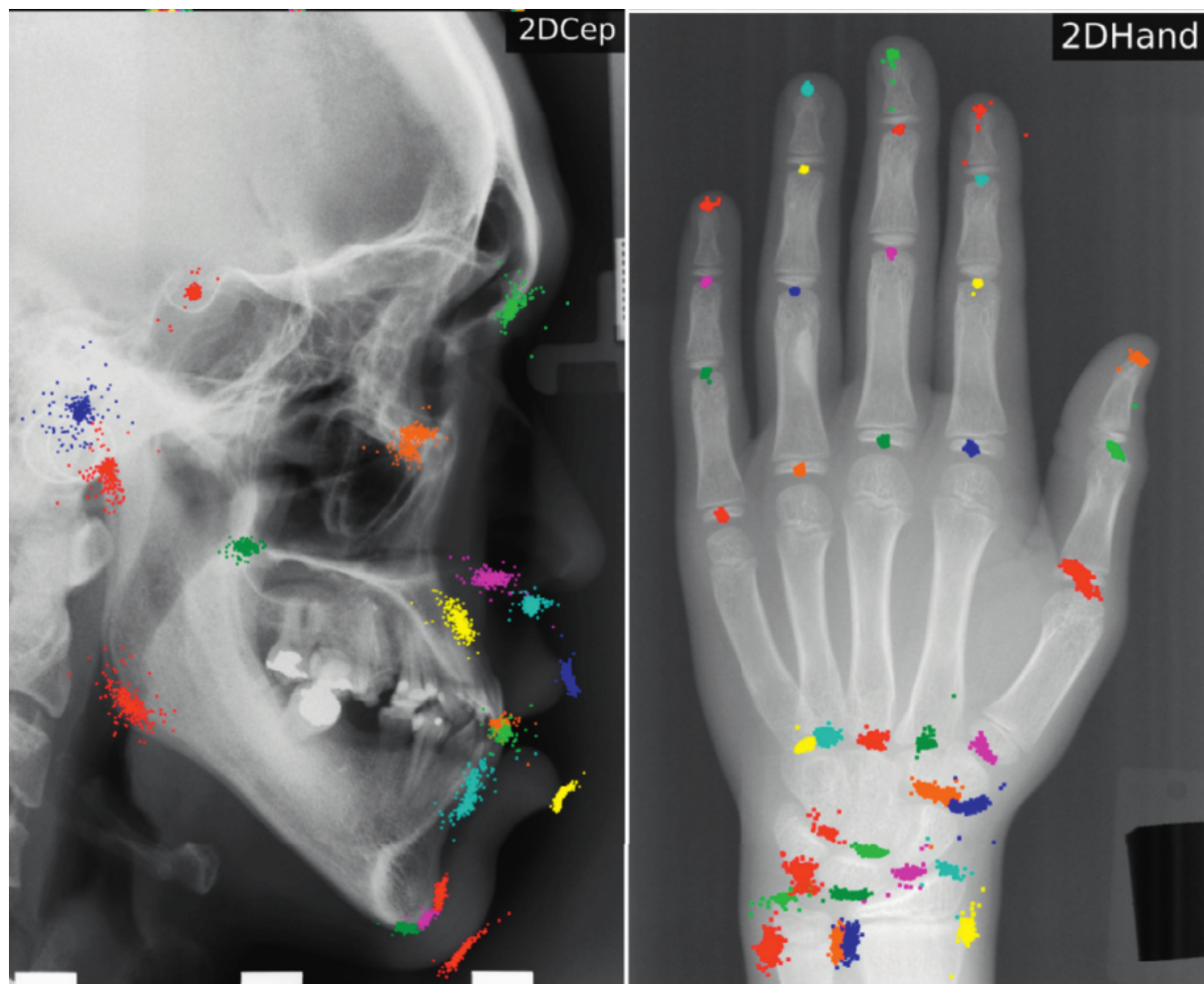


Figura 2.6. De lado izquierdo se muestra una imagen cefálica lateral y en la parte derecha una radiografía con puntos localizados en [10] (Reproducida de [10]).

Cambiando a otros métodos de machine learning en [11] hicieron uso de las redes neuronales convolucionales profundas CNNs en imágenes 2D y 3D. Para las primeras ocuparon un total de 895 radiografías (en las cuales se abordará). Los autores propusieron 4 arquitecturas de CNNs en conjunto con mapas de calor o heatmaps y funciones gaussianas que indican la posición de la *landmark*. La primera llamada **Downsampling-Net** utiliza capas de convolución y agrupación, teniendo una baja precisión en la localización. El segundo **ConvOnly-Net** no hace uso de capas de convolución y agrupación pero se deben aumentar los pesos en la red. El tercero **U-Net** utiliza un área mayor de búsqueda y con una mejor precisión, teniendo una arquitectura simétrica. La última arquitectura **SpatialConfiguration-Net** (SCN) combina la apariencia local de los puntos de interés con la configuración global de las otras *landmarks*.

En la Figura 2.7 se observa un diagrama general de su última arquitectura propuesta.

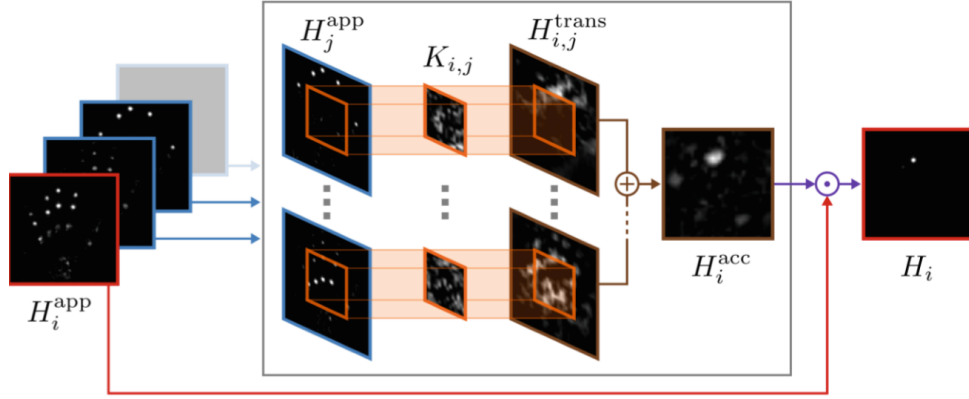


Figura 2.7. Diagrama general de [11], que muestra su arquitectura llamada **Spatial Configuration-Net** (Reproducida de [11]).

Finalmente en el trabajo [12], se retomó la arquitectura **Spatial Configuration-Net** (SCN) propuesta en [11] y también retomaron parte de su trabajo de [11]. Por lo que para ello tomaron en consideración la configuración espacial de las *landmarks* predichas. Las CNNs de las que hacen uso se entrenaron con los mapas de calor que van obteniendo al analizar nuevas imágenes. Con respecto a su modelo SCN lo entrenaron con la información local alrededor de cada *landmark* eliminando las regiones ambiguas. En total marcaron 37 *landmarks* alrededor de los dedos y en uniones de los huesos. Para su evaluación hicieron un pre-procesamiento con el método de *Otsu*. Emplearon diversos tamaños de las imágenes ya que buscaron comparar el error de la localización automática con [11]. Tomaron un total de 600 imágenes de entrenamiento y de evaluación un conjunto de 300 radiografías. Al igual que en sus anteriores trabajos comprobaron su algoritmo en otro tipo de imágenes médicas como es MRI de manos, radiografías de espina dorsal y cefálicas laterales, ver la Figura 2.8.

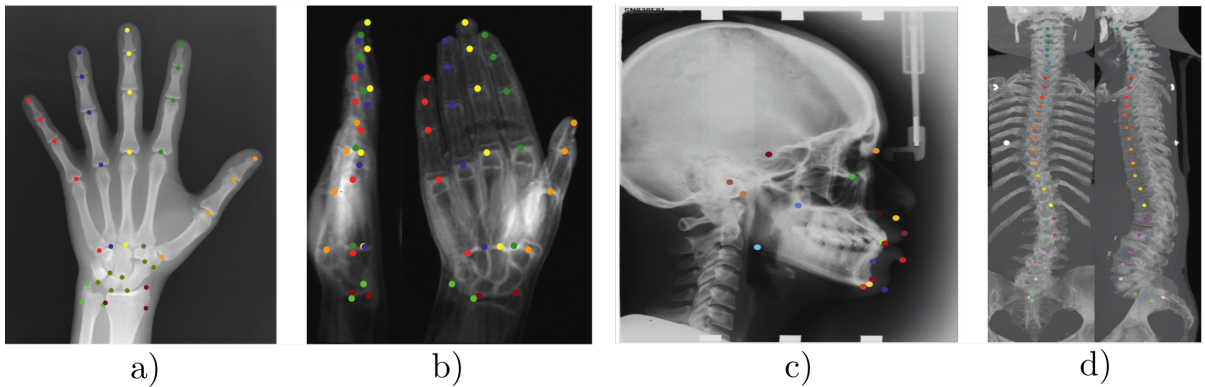


Figura 2.8. Ejemplos de imágenes marcadas en [12], donde a) es una radiografía de mano, b) MRI de mano, c) imagen cefálica lateral y d) espina dorsal (Reproducida de [12]).

Se muestra la Tabla 2.1 con información del tipo de procedimiento utilizado por cada

autor, la imagen analizada y el banco con la cantidad de radiografías que usaron. Un aspecto importante a mencionar es que algunos autores no mencionan el lugar donde obtuvieron su conjunto de imágenes. También la variedad de imágenes que se utilizan para ubicar los puntos.

Así, los autores mostrados en la Tabla 2.1 como son [28, 8, 11, 12, 10] no mencionan si con su banco de imágenes realizaron algún tipo pre-procesamiento o si se presentara algún cambio en la postura angular de la mano, escala o contraste, se seguirían localizando con éxito las *landmarks* con su respectivo método. Por lo tanto en esta tesis se propone un método que ha diferencia de los mencionados y comparados en la Tabla 2.1 se contempla lo mencionado. Puesto que el algoritmo propuesto es totalmente automático, robusto a las siguientes características: postura angular, traslación, escala, contraste y objetos alrededor de la región de la mano. Por lo que si la mano esta rotada se corrige su posición. Si la mano en la imagen médica es chica o grande, contiene objetos alrededor de ella, está ubicada en la parte superior, inferior, a la derecha o izquierda o si su contraste es oscuro o brillante las *landmarks* son localizadas.

Un aspecto a notar es que en la actualidad los algoritmos de machine learning son los más utilizados para el análisis de imágenes médicas. Lo que permite el crecimiento de ésta área, reducir el tiempo en el análisis de imágenes y tomar mejores decisiones para que el médico pueda concentrarse en otros aspectos en la salud de sus pacientes. Gracias a que también se han podido crear tecnologías que permiten digitalizar radiografías para su posterior análisis. Para finalizar como menciona [32] “el terreno de la detección, segmentación, clasificación y cuantificación” de imágenes 2D y 3D serán las de más oportunidad para la inteligencia artificial.

Autores	Método utilizado	Tipo de imagen utilizada	Banco de imágenes
Ebner et al [28]	Regression Random Forest	Resonancia Magnética	Propio (60 imágenes)
Grzejszczak et al [33]	Skin Presence Masks (SPMs)	Imágenes RGB 2D y videos secuencias	HGR (1558 imágenes)
Banda et al [7]	PCA	Radiografías	University of Cambridge (362 imágenes).Stanford University School of Medicine (14306 imágenes).
Moreno et al [6]	Aplicación de técnicas de procesamiento de imágenes	Radiografías	No mencionado (40 imágenes)
Mahmoodi et al [34]	Bayesian and Regression Technique	Radiografías	No mencionado (1000 imágenes)
Niemeijer et al [35]	ASM (Active Shape Models)	Radiografías	Hospital of the Katholieke Universiteit Leuven, Belgium (243 imágenes)
Fernández et al [36]	Registration	Radiografías	Hospital Río Carrión, Palencia, España (Cantidad de imágenes no mencionada)
Doner et al [8]	PCA	Radiografías	No mencionado (20 imágenes)
Payer et al [11, 12],	CNN's, heatmaps	MRI, radiografías	Digital Hand Atlas Database System (895 imágenes)
Urschler et al [10]	RF	MRI, radiografías	Digital Hand Atlas Database System (895 imágenes)

Tabla 2.1. Comparación de métodos para ubicar *landmarks* en distintas imágenes de manos.

Capítulo 3

Descripción del método

En esta sección se describe el método desarrollado en la búsqueda automática de los puntos de interés.

El conjunto de imágenes utilizados se tomo de [37]. Este es de acceso público por lo que no se tuvieron dificultades o restricción alguna.

3.1. Descripción general del sistema

El método propuesto se basa en dos etapas principales: entrenamiento y evaluación. Para las dos etapas se realizó una alineación de la imagen que consiste en segmentar la región de la mano. Para ello se deben de seguir los siguientes pasos. Para el primero se determinó la ubicación del baricentro de la mano. Como se sabe el baricentro es el centro de gravedad de un cuerpo, pero en Computer Vision su significado corresponde a un grupo de pixeles conectados en una imagen que comparten alguna propiedad en común. Es decir, cada objeto dentro de la imagen esta conformado por pixeles, por lo tanto el baricentro es el promedio ponderado de todos los pixeles que constituyen dicha forma. Por consiguiente el baricentro permite ubicar la región de la mano para eliminar objetos o artefactos innecesarios que puedan aparecer alrededor de la mano. El siguiente paso fue encontrar automáticamente la región de interés donde se encuentre la mano y segmentarla para posteriormente aplicar una corrección de contraste en la región central de la mano que permita tener un contraste parecido en dicha región de la mano tanto en las imágenes del conjunto de entrenamiento como en la imagen de evaluación en la que se desea realizar el proceso automático de la localización de las *landmarks*. Posteriormente, se aplicó una corrección automática de la postura de la mano (ángulo) utilizando para ello un algoritmo de reconocimiento automático del ángulo dominante de los gradientes de intensidad de la imagen, dicho algoritmo es propuesto en esta tesis. En la etapa del entrenamiento cada imagen se marcó manualmente con dos *landmarks*, una en la parte superior y la otra en la parte inferior del tercer hueso metacarpiano. Dichas *landmarks* sirvieron como puntos de referencia centrales para extraer dos regiones de interés alrededor de cada una de estas dos *landmarks*. El algoritmo de detección que se propone en esta tesis es multiresolución,

con el objetivo de acelerar el proceso de búsqueda de cada una de las dos *landmarks*. Primero, se realiza la búsqueda de una de las *landmarks* en una resolución muy baja y posteriormente en resoluciones cada vez más altas. En la práctica se utilizó una pirámide de tres resoluciones distintas, con el objetivo de agilizar la búsqueda haciéndolo primero en una resolución muy baja, y posteriormente buscando solamente en una pequeña región del siguiente nivel de resolución mayor al anterior. Por tanto, se obtuvo una región de interés alrededor de cada *landmark* por cada una de las tres resoluciones. Se utilizó el método de las *eigenfaces* [24] para representar cada una de las tres regiones de búsqueda para cada una de las dos *landmarks*. Con dicho método se pudo comparar o proyectar en el subespacio lineal generado por el método de las *eigenfaces* alguna región candidata al momento de hacer la búsqueda y comprobar si el subespacio de las *eigenfaces* es capaz de generar una reconstrucción adecuada de la región. La búsqueda es comenzada en la menor resolución y termina en la de mayor resolución, y las regiones de búsqueda alrededor de la *landmark* en cuestión abarcaran una región más pequeña de la mano conforme se cambie a una resolución mayor. Esto se observa en la Figura 3.1.

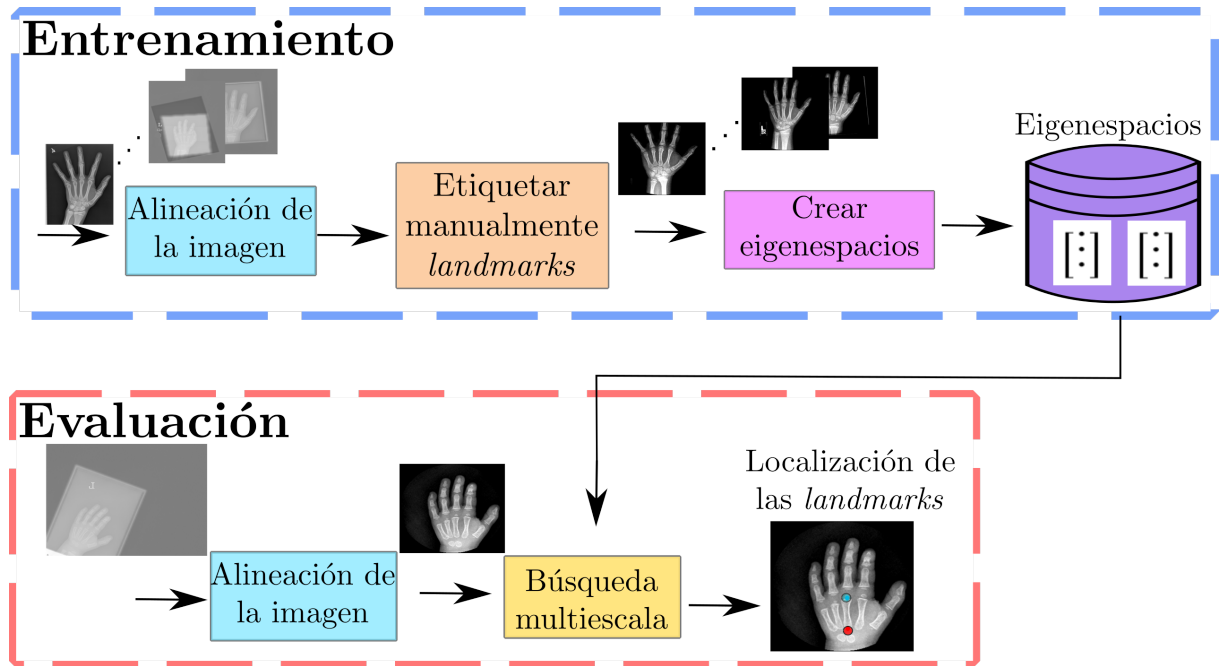


Figura 3.1. Descripción general de la localización automática de *landmarks* en el tercer hueso metacarpiano

La localización automática de *landmarks* se realizó en la zona superior e inferior del tercer hueso metacarpiano, se decidió así ya que en la parte superior de dicho hueso se encuentra el desarrollo de la epífisis de los huesos metacarpianos y las falanges proximales y en la parte inferior están ubicados los huesos carpianos, todos ellos son de utilidad para la estimación de la edad ósea.

3.2. Descripción General del proceso de Alineación de Imagen

Debido a que el método de las *eigenfaces* funciona mejor cuando las imágenes de entrenamiento son parecidas entre sí, se requiere realizar un pre-acondicionamiento de las imágenes de entrenamiento con el objetivo de que todas ellas tengan el mismo tamaño de imagen, y además, la región de la mano aparezca con la misma postura (tanto en posición, como en ángulo) y con una escala uniforme en todas. Así mismo, también se buscó uniformizar el contraste de la región de interés dentro de la mano. En pocas palabras, se realizó una alineación de las imágenes de entrenamiento en cuatro aspectos diferentes: traslación, rotación, escala y contraste.

De esta manera, se nombra $I(m, n)$ a la imagen original, y $I^g(m, n)$ a una imagen intermedia alineada en posición, escala y contraste que se obtiene mediante el Algoritmo 1. La imagen de salida de dicho algoritmo es una imagen cuadrada de 256×256 píxeles donde la mano se encuentra centrada y corregida en contraste. El último paso de la alineación de la imagen es realizar una corrección de la postura angular de la mano, para ello se aplicó el Algoritmo 5 obteniendo la imagen $I^s(m, n)$.

3.3. Alineación de la imagen

3.3.1. Segmentación de la mano

En esta sección se buscó obtener una imagen segmentada que esté corregida en contraste, posición angular, además de que solo se encuentra la región de la mano eliminando los objetos innecesarios. Con la finalidad de homogeneizar el tamaño de las imágenes y poder aplicar el mismo método para encontrar correctamente la región de la mano en todas éstas, se procedió a modificar el tamaño de la imagen de entrada $I(m, n)$ con el fin de que las demás imágenes puedan ser similares en tamaño, por lo que se determina su dimensión más pequeña y se cambia a 256 píxeles y la otra dimensión se modificó a un valor proporcional p que conserve la relación de aspecto de la imagen original, obteniendo una nueva imagen $I^p(m, n)$. Lo siguiente fue localizar el baricentro de la mano b_n . Si se considera la posición de un píxel como $\mathbf{x} = (m, n)$, entonces el baricentro representa un promedio de las posiciones de los píxeles ponderadas por su nivel de gris, esto se observa en la Ecuación 3.1.

$$baricentro = \frac{\sum_{\forall \mathbf{x}} x \bar{x}}{\sum_{\forall \mathbf{x}} \bar{x}} \quad (3.1)$$

Este cálculo puede verse en detalle en el Algoritmo 2.

Después se obtuvo solamente la región de la mano con el fin de eliminar los objetos

alrededor de ella. Con lo que se procedió a realizar el cálculo de densidad de área en la región de la mano al ocupar coronas circulares, haciendo uso de un círculo u y radio R y otro círculo v con radio r , donde $R > r$ y seleccionar la circunferencia que contendrá en su interior solamente la región de la mano. Para ello se utilizó la Ecuación 3.2.

$$densidad_de_área = \frac{u - v}{\pi(R^2 - r^2)} \quad (3.2)$$

Este cálculo y el proceso de selección de la circunferencia puede verse con más detalle en el Algoritmo 3. Una vez encontrado el baricentro y la región de la mano se continuó por ubicarlos en la imagen de entrada $I(m, n)$ con el propósito de mantener la información original de los huesos, ya que en pasos posteriores se realiza un recorte de toda la imagen que solo contenga la mano. Esto se consiguió al multiplicar el baricentro b_n por el valor proporcional p y multiplicar dicho valor por el radio r_c creando b_f y r_f los cuáles se utilizan para aplicar una máscara binaria $M(m, n)$ circular a la imagen de entrada que contenga en su interior solo la región de la mano. Se procedió a realizar un corte sobre $I(m, n)$ y obtener la región $C(m, n)$ con dimensiones $2r_f \times 2r_f$ siendo el centro geométrico de la imagen el baricentro b_f . Se cambiaron las dimensiones de la región a 256×256 píxeles para obtener la imagen cuadrada $F(m, n)$. Se continuó por corregir el contraste de la imagen de manera que se pueda tener un mejor aspecto de los huesos de la mano. Con este fin se obtuvieron dos valores promedios relativos al histograma, un máximo V_{max} y un mínimo V_{min} , los cuales se calculan al tomar un segmento cuadrado $W(j, k)$ dentro de la imagen $F(m, n)$. Después se aplicó la Ecuación 3.3 dando como resultado la imagen de la mano corregida en contraste $I^g(m, n)$.

$$corrección_de_contraste = \frac{255(F(m, n) - V_{min})}{V_{max} - V_{min}} \quad (3.3)$$

De la misma manera este cálculo se puede observar con más detalle en el Algoritmo 4. Finalmente se cambió la postura angular de la región de la mano al calcular el histograma de direcciones ponderado y la densidad en la imagen $I^g(m, n)$, tomando un segmento cuadrado $P(m, n)$ dentro de ella y realizando la sumatoria de los píxeles de dicho segmento. El cálculo se llevó a cabo como se indica en la Ecuación 3.4.

$$densidad = \sum_{m, n} P(m, n) \quad (3.4)$$

Se describe de manera más detalla éste proceso en el Algoritmo 5. Consiguiendo la imagen con su postura angular corregida $I^s(m, n)$. Todo lo anteriormente mencionado se muestra en la Figura 3.2 y su respectivo Algoritmo 1 con los pasos para obtener la imagen con la región segmentada de la mano.

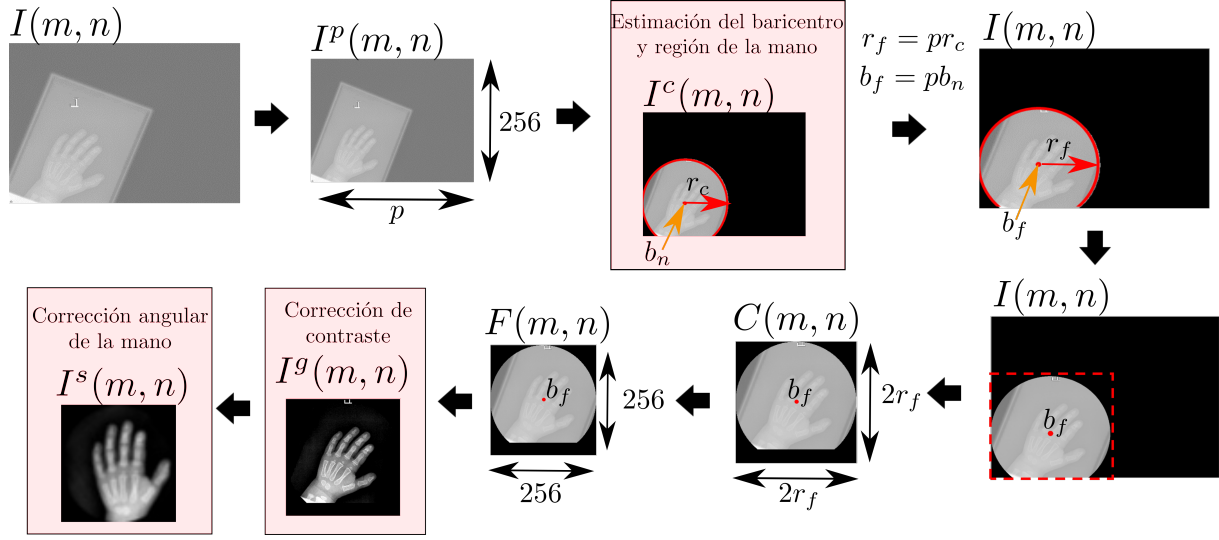


Figura 3.2. A partir de $I(m, n)$ se modificaron sus dimensiones, donde la dimensión menor es de 256 píxeles y la segunda a un valor p . Después se localizó el baricentro b_n y la región de la mano para obtener el radio r_c . Lo anterior se reubicó en la imagen $I(m, n)$. Se continuó aplicando un corte dentro de $I(m, n)$ creando $C(m, n)$ y cambiando sus dimensiones a 256×256 píxeles obteniendo $F(m, n)$. Por último se realizó una corrección de contraste $I^g(m, n)$ y postura angular obteniendo $I^s(m, n)$.

Algoritmo 1. Región segmentada de la mano

Entrada: Imagen $I(m, n)$ a evaluar.

Salidas: Imagen de la mano $I^s(m, n)$ con dimensiones de 256×256 píxeles ,
baricentro b_f de la imagen.

1 **inicio**

2 Modificar las dimensiones de $I(m, n)$ cambiando la más pequeña a 256 píxeles
y la otra un valor proporcional p creando $I^p(m, n)$.

3 Aplicar los siguientes algoritmos a $I^p(m, n)$:

4 • Algoritmo 2 para obtener la ubicación del baricentro b_n .

5 • Algoritmo 3 para obtener el radio de la mano r_c y la imagen $I^c(m, n)$.

6 Realizar $r_f = p \cdot r_c$ y $b_f = p \cdot b_n$.

7 Aplicar una máscara circular binaria $M(m, n)$ con centro en b_f y radio r_f a
 $I(m, n)$.

8 A partir de b_f como centro tomar una región cuadrada $C(m, n)$ dentro de
 $I(m, n)$ de dimensiones $2r_f \times 2r_f$.

9 Cambiar las dimensiones de $C(m, n)$ a 256×256 píxeles creando $F(m, n)$.

10 Aplicar el Algoritmo 4 a $F(m, n)$ obteniendo la imagen $I^g(m, n)$.

11 Aplicar el Algoritmo 5 a $I^g(m, n)$ obteniendo la imagen $I^s(m, n)$.

12 **fin**

3.3.1.1. Detección del centro geométrico de la mano

Se propuso localizar el baricentro de la imagen completa porque aunque suelen aparecer artefactos ajenos a la mano alrededor de ésta, se observó experimentalmente que el baricentro así encontrado de todas formas se ubicaba dentro de la región de la mano. Sin embargo, para encontrar la ubicación real del centro geométrico de la mano, se observó que el baricentro mencionado podría ser un buen punto inicial para encontrar dicho centro. Se propuso entonces un algoritmo basado en ir reduciendo una región circular alrededor del baricentro recalculando iterativamente en base a imágenes de bordes obtenidas luego de ajuste de contraste. Así se comenzó por definir un radio inicial r de 100 píxeles el cual se determinó al considerar las manos más grandes, con el fin de ir reduciéndolo y limitar la zona de búsqueda del baricentro. Lo siguiente fue obtener la imagen rectangular de bordes $E_1(m, n)$ de la imagen de entrada $I(m, n)$ al aplicar el algoritmo de detección de bordes de Canny con $U_{max} = 20$ y $U_{min} = 5$ como valores umbrales. Se continuó por ubicar por primera vez el baricentro (m_{b1}, n_{b1}) de $E_1(m, n)$ y trasladarlo con la misma posición en la imagen $I(m, n)$. Se prosiguió aplicando el Algoritmo 4 que permite corregir el contraste de la imagen, además de asignar el valor a la variable $a = 2.5$ y de tomar una región cuadrada $W(j, k)$ de tamaño 70×70 obteniendo $I^c(m, n)$. A continuación se repitió la detección de bordes con la imagen $I^c(m, n)$, pero modificando los valores umbrales con $U_{max} = 100$ y $U_{min} = 70$ resultando una nueva imagen $E_2(m, n)$ que permite distinguir de mejor manera los huesos de la mano. Ver Figura 3.3. Se realizó una nueva ubicación del baricentro (m_{b2}, n_{b2}) en la imagen $E_2(m, n)$. A partir del punto $b_a = (m_{b2}, n_{b2})$ en la imagen $E_2(m, n)$ y siendo éste el centro, se dibujó una circunferencia de radio r para comenzar el ajuste en la posición del baricentro sólo en la circunferencia. Así se inició por localizar dentro de la circunferencia y guardar la posición del baricentro en la variable b_a , reduciendo el radio en 10 píxeles, repitiendo la localización del baricentro y guardarlo en $b_n = (m_{b3}, n_{b3})$, se continuó realizando la localización y reducción de la circunferencia, pero guardando en b_n la nueva localización y la anterior en b_a como se muestra en la Figura 3.4, además de tener dos condiciones de paro al ubicar el punto. Para la primera se notó que en algunas ocasiones el baricentro ya no cambiaba su posición, así se midió la distancia euclidiana entre b_a y b_n y si es menor o igual a 1 píxel, se detiene la localización. Para la segunda condición se tomó en consideración que a partir de que el radio llegara al valor de 30 píxeles se detenga la localización, debido a que la circunferencia es pequeña en dimensiones y el baricentro ya no presentaba una mayor diferencia de distancia con respecto al anterior. Al cumplirse alguna de las dos condiciones se mantiene la ubicación final de b_n , el cuál se encuentra en el centro de la región de la mano. En el Algoritmo 2 se muestran los pasos para obtener el baricentro como se ha mencionado anteriormente.

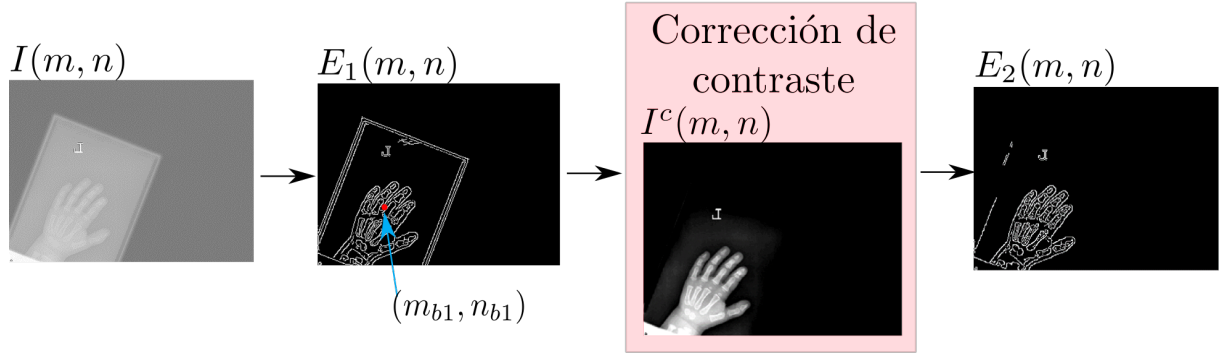


Figura 3.3. Se detectaron los bordes de la imagen de entrada $I(m, n)$ obteniendo la imagen $E_1(m, n)$ y localizar el baricentro (m_{b1}, n_{b1}) . Después se corrigió el contraste de $I(m, n)$. Finalmente se detectaron los bordes de la imagen anterior con el Algoritmo de *Canny* y obtener una nueva imagen $E_2(m, n)$ que corresponde a los bordes de los huesos de la mano.

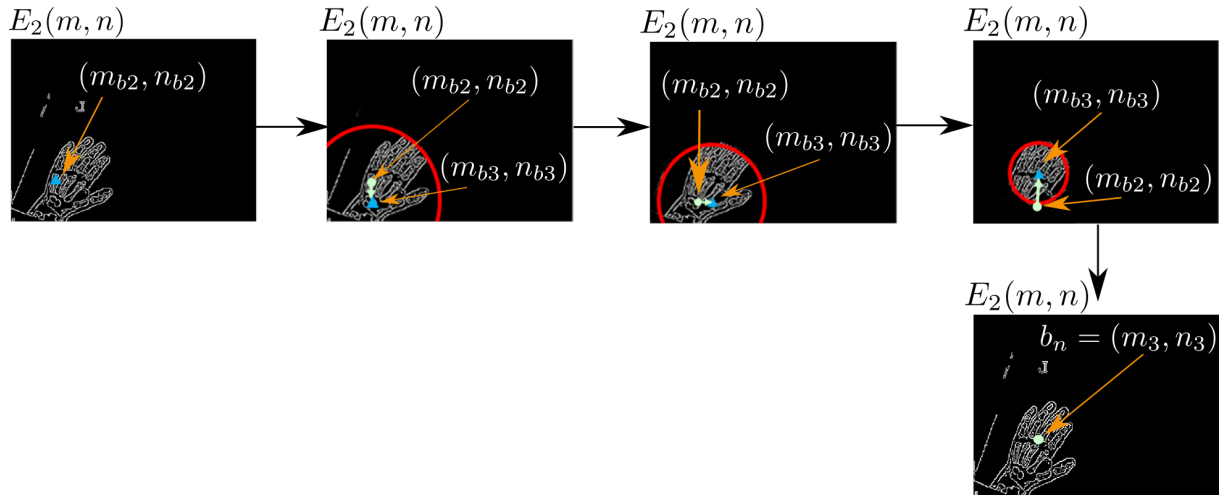


Figura 3.4. Localización del baricentro en la región de la mano. Se inició localizando el punto en la imagen $E_2(m, n)$ y siendo este el centro se dibujó una circunferencia para ajustar la ubicación del baricentro. Así el punto (m_{b2}, n_{b2}) representa el baricentro anteriormente ubicado y (m_{b3}, n_{b3}) el nuevo, notándose como se va reduciendo el área de la circunferencia.

Algoritmo 2. Estimación del centro geométrico de la mano**Entrada:** Imagen $I(m, n)$.**Salidas:** Baricentro $b_n = (m_{b3}, n_{b3})$, Imagen de bordes $E_2(m, n)$.

```

1 inicio
2   Radio inicial  $r = 100$ .
3   Obtener la imagen de bordes  $E_1(m, n)$  de  $I(m, n)$  con el algoritmo de Canny
   y valores umbrales  $U_{max} = 20$  y  $U_{min} = 5$ .
4   Localizar el baricentro  $(m_{b1}, n_{b1})$  de  $E_1(m, n)$  y trasladarlo a  $I(m, n)$ .
5   Aplicar el Algoritmo 4 para obtener  $I^c(m, n)$ .
6   Obtener la imagen binaria de bordes  $E_2(m, n)$  de  $I^c(m, n)$  con valores
   umbrales de  $U_{max} = 100$  y  $U_{min} = 70$ .
7   Localizar el baricentro  $(m_{b2}, n_{b2})$  en la imagen  $E_2(m, n)$ .
8   Sea  $b_a = (m_{b2}, n_{b2})$ .
9   mientras condición1 = Verdadero o condición2 = Verdadero hacer
10     $I_M(m, n) = M(m, n)E_2(m, n)$  con centro en  $b_a$  y radio  $r$ .
11    Calcular el nuevo baricentro  $(m_{b3}, n_{b3})$  de  $I_M(m, n)$ , siendo  $b_n = (m_{b3}, n_{b3})$ .
12    Realizar  $r = r - 10$ .
13    si la distancia euclidiana entre los baricentros es  $d \leq 1$  entonces
14      | asignar a la variable condición1  $\leftarrow$  Falso.
15    sinó
16      | asignar a la variable condición1  $\leftarrow$  Verdadero
17    fin
18    si el valor del radio es  $r \leq 30$  entonces
19      |
20      |   ■ condición1  $\leftarrow$  Falso.
21      |   ■ condición2  $\leftarrow$  Falso.
22      |
23    fin
24  fin

```

3.3.1.2. Obtención de la región de la mano

Una vez determinado el centro geométrico de la mano, se propuso un método para obtener una región circular que contenga la mano completa, de tal forma que el tamaño de la mano dentro de dicha región mantenga la misma proporción en cualquier imagen a la que se aplique el método. Esto será de utilidad tanto en la etapa de entrenamiento, como en la etapa de evaluación, pues en el caso del entrenamiento se requiere que las imágenes de ejemplo estén normalizadas en cuanto al tamaño o escala de las manos, y en el caso de la etapa de evaluación, se requiere normalizar en tamaño la imagen a evaluar. Por lo que se propuso un algoritmo basado en que la densidad de píxeles de borde de la mano es mayor a la palma de ésta que en la región periférica como los dedos. El algoritmo propuesto detecta el cambio en la densidad, cuando se mide dicha densidad en una región anular (centrada en el centro de la mano) que comienza siendo una pequeña región circular (casi siempre dentro de la palma de la mano) y que va creciendo en diámetro hasta salir de la región más densa de la mano. De esta manera se comenzó por inicializar la constante c con un valor de 1.32, la cual se obtuvo de manera empírica, dado que permitirá crear los radios para las circunferencias, además dicho valor permite que exista una proporción con respecto a las circunferencias anteriores. Se procedió creando un conjunto de radios $\mathcal{R}$ con 18 elementos, los cuales permiten realizar el cálculo de la densidad de área y mantener una proporción con respecto a la circunferencia anterior. Esto fue importante ya que así se conoció hasta que circunferencia se terminaba la mano. Por lo que para cada radio $r \in \mathcal{R}$ se realizó la siguiente operación $r_i = 30c^n$. También se calculó el área para cada elemento del conjunto de radios, creando un nuevo conjunto de áreas $\mathcal{A}$, donde para cada elemento del conjunto anterior se calculó $A_i = \pi r_i^2$. Se continuó con un tercer conjunto $\mathcal{M}$ que contiene las 18 imágenes rectangulares conformadas por máscaras binarias por lo que en el interior de la circunferencia se encuentran únicamente valores de 1 y en el exterior valores de 0, asimismo, cada circunferencia de M_i esta formada por el radio $r_i \in \mathcal{R}$ correspondiente. Continuando se multiplicó la imagen de bordes $E(m, n)$ por cada elemento del conjunto $\mathcal{M}$ para crear el conjunto $\mathcal{M}_E$. Así se da paso a el conjunto $\mathcal{H}$ que contiene coronas circulares formadas por la resta absoluta de la imagen binaria rectangular de bordes M_{Ei} por su elemento posterior M_{Ei+1} (Figura 3.5) a excepción del primero elemento $\mathcal{H}_1$ que esta conformado por la primera imagen de bordes binaria con la máscara circular $\mathcal{M}_{E1}$, siendo un total de 18 elementos para el conjunto $\mathcal{H}$. De manera similar se crearon coronas circulares pero en el conjunto de áreas $\mathcal{A}$, teniendo las mismas consideraciones donde el primer elemento del nuevo conjunto $H_{A1} = A_1$ corresponde al primer elemento del conjunto de áreas y realizando la resta absoluta del elemento A_i con el posterior A_{i+1} , por lo que cada elemento H_{Ai} esta conformado de la siguiente manera: $H_{A2} = A_2 - A_1$, $H_{A3} = A_3 - A_2$, $\dots$, $H_{A18} = A_{18} - A_{17}$. Así tanto el conjunto de coronas circulares de áreas $\mathcal{H}_A$ y el de coronas circulares $\mathcal{H}$ se puede calcular la densidad de área. La densidad permitió conocer hasta que circunferencia se han finalizado los píxeles que corresponden a la mano, por lo tanto cuando la densidad empezó a descender significó que se ha finalizado en determinar la región de la mano. Así para cada elemento $H_i \in \mathcal{H}$

y $H_{Ai} \in \mathcal{H}_A$ se calcula $D_i = \frac{H_i}{H_{Ai}}$, para $D_i \in \mathcal{D}$ con $i = 1, 2, 3, \dots, 18$ que corresponde a las densidades para cada corona circular. Formadas las coronas se debió conocer cual se seleccionará para aplicarla en la imagen quedando en su interior solamente los pixeles que corresponden a los bordes de la mano, por lo tanto se tuvo que elegir un valor umbral de tal forma que el valor de la densidad ya no correspondiera a la región de la mano. Para facilitar la elección del umbral se buscó normalizar todos los valores del conjunto $\mathcal{D}$. Para dicho objetivo se ordenó de mayor a menor todos los valores de $\mathcal{D}$ y se encontró que al considerar los dos primeros valores máximos del conjunto de densidades $D_1, D_2 \in \mathcal{D}$ se obtuvo una mejor consideración del umbral, así se obtiene el valor promedio $\bar{m}$ de los valores antes mencionados y se normalizó todo el conjunto $\mathcal{D}$ al dividir cada elemento D_i entre $\bar{m}$, creando un nuevo conjunto $\mathcal{E}$. Así se pudo determinar un valor umbral de 0.2 al graficar la mano más pequeña y la mano más grande, ver Figura 3.6. Con esta información se observó que las densidades menores al umbral seleccionados comenzaban a descender drásticamente significando que los pixeles de la mano ya no se encontraban en las siguientes circunferencias. Todo lo anteriormente mencionado se aplicó para el análisis de las demás imágenes, por lo que al analizar una nueva imagen se obtuvo su conjunto $\mathcal{E}$. Finalmente para la elección automática de la circunferencia se realizó lo siguiente: se iniciaron las variables *contador* y r_c iguales a cero, para determinar el número de circunferencia superior cercano o igual al umbral, donde el valor final del *contador* se asignó a la variable r_c e indicó el número del elemento en el conjunto $\mathcal{R}$ a seleccionar, aplicándolo a una máscara circular binaria $M(m, n)$ a $I(m, n)$, siendo el centro de la circunferencia b_n y de radio $r_c \in \mathcal{R}$, obteniendo la imagen final $I_E(m, n)$ formada por la circunferencia en su interior con los valores de gris de los pixeles que conforman la mano y en el exterior únicamente valores de 0 siendo un fondo negro. Se presenta el Algoritmo 3 con los pasos para obtener la imagen $I_E(m, n)$ antes descrita.

Algoritmo 3. Estimación de la región de la mano basado en cambios de densidad usando una región circular creciente.

Entradas: Imagen $I(m, n)$, Imagen de bordes $E(m, n)$ y baricentro b_n .

Salidas: Imagen de la mano con su circunferencia contenida $I_E(m, n)$, radio de la circunferencia r_c .

1 **inicio**

2 Inicializar $c = 1.32$.

3 Crear el conjunto de radios $\mathcal{R} = \{r_1, r_2, r_3, \dots, r_{18} | r_i = 30c^n\}$.

4 Crear el conjunto de áreas $\mathcal{A} = \{A_1, A_2, A_3, \dots, A_{18} | A_i = \pi r_i^2\}$.

5 Crear el conjunto de imágenes binarias
 $\mathcal{M} = \{M_1(m, n), M_2(m, n), \dots, M_{18}(m, n)\}$.

6 Realizar $\mathcal{M}_E = \{E(m, n)M_i(m, n) | M_i(m, n) \in \mathcal{M}\}$.

7 Crear el conjunto de coronas circulares $\mathcal{H} = \{H_1, H_2, \dots, H_{18}\}$.

8 Crea el conjunto de áreas de las coronas circulares $\mathcal{H}_A = \{H_{A1}, H_{A2}, \dots, H_{A18}\}$.

```

9
10  Calcular la densidad de área de cada corona circular  $\mathcal{D} = \{D_1, D_2, \dots, D_{18}\}$ 
    mediante:  $D_i = \frac{H_i}{H_{Ai}}$ .
11  Ordenar de mayor a menor las densidades de  $\mathcal{D}$ .
12  Obtener el primer  $D_1$  y segundo  $D_2$  máximos de  $\mathcal{D}$ .
13  Calcular el valor promedio  $\bar{m}$  al realizar:  $\bar{m} = \frac{D_1 + D_2}{2}$ .
14  Normalizar todas las densidades de  $\mathcal{D}$  con  $\mathcal{E} = \{E_1, E_2, \dots, E_{18} | E_i = \frac{D_i}{\bar{m}}\}$ .
15  Sea contador = 0 y  $r_c = 0$  .
16  para cada elemento de  $\mathcal{E}$  hacer
17      si para cada elemento  $E_i = < 0,2$  entonces
18           $r_c \leftarrow \text{contador}$ .
19          break
20      sinó
21           $\text{contador} \leftarrow \text{contador} + 1$ .
22      fin
23  fin
24  Seleccionar el número de circunferencia  $r_c$  final, en  $\mathcal{R}$ .
25  Realizar  $I_E(m, n) = M(m, n)I(m, n)$ , donde  $M(m, n)$  tiene radio  $r_c$ .
26 fin

```

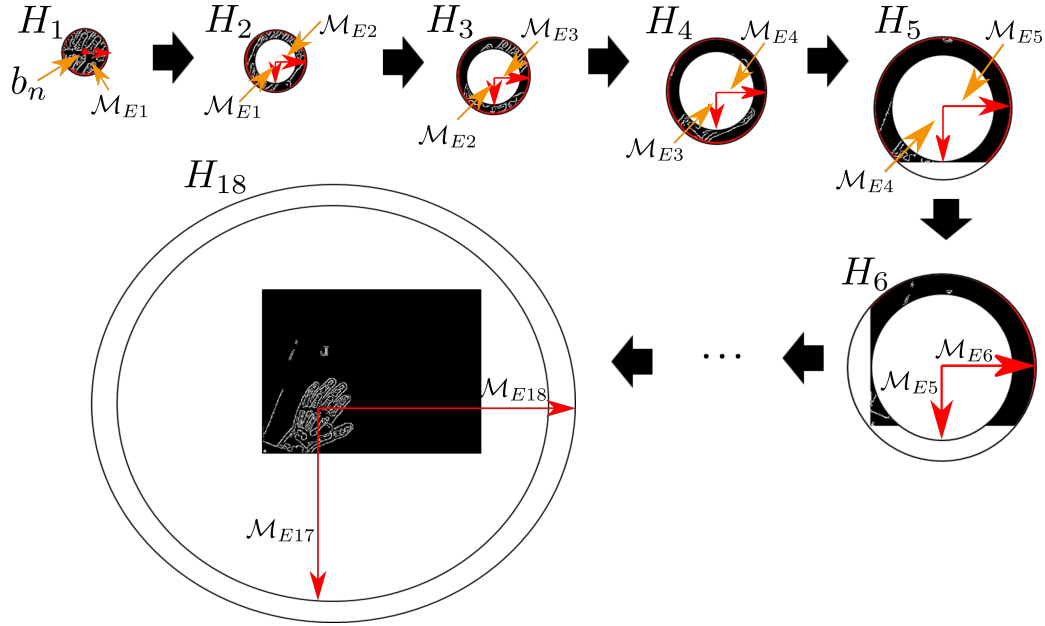


Figura 3.5. Creación de las coronas circulares con la imagen de bordes $E(m,n)$, la cual se inició con el baricentro b_n siendo el centro y comenzar con el radio $\mathcal{M}_{E1}$ para formar el primer elemento H_1 , se continuó con las coronas formadas por $H_2 \dots H_{18}$.

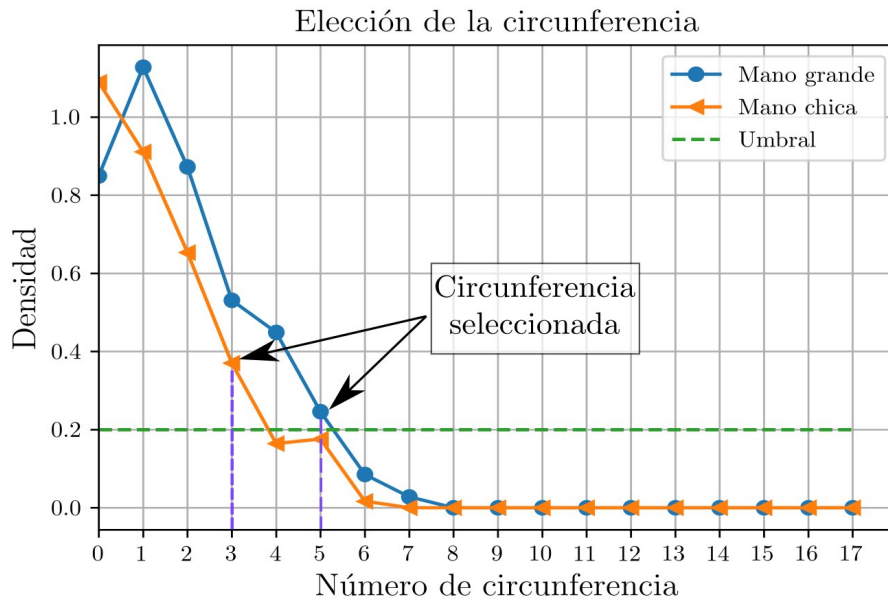


Figura 3.6. Gráfica con dos densidades que corresponden a la mano más chica y más grande del conjunto utilizado. Por lo que se pudo determinar el valor del umbral igual a 0.2 para la selección de las circunferencias en las demás imágenes.

3.3.1.3. Corrección de contraste

Se observó que las imágenes radiográficas de las manos, en distintas ocasiones contienen diferentes objetos que no son de interés para el método propuesto, pero que están situados alrededor de la mano. Al realizar la corrección de contraste en toda la imagen, los objetos influían en dicha modificación, una vez que se aplicó la corrección en todas las imágenes estas quedaban ajustadas en diferente contraste, por lo tanto se buscó que el ajuste de contraste sea exactamente igual para todas las regiones de todas las manos de todas las imágenes.

Para esto a partir del baricentro ubicado con el Algoritmo 2 como centro dentro de la mano, se tomó una región cuadrada $W(j, k)$ dentro la imagen y se calculó el valor promedio $\bar{\omega}$ de los valores de gris de dicha región. Lo siguiente fue calcular la desviación estándar σ de los valores de gris de $W(j, k)$. Los valores $\bar{\omega}$, σ y a nos permiten calcular un mínimo y máximo promedio relativos al histograma, además el valor que se asigne a la constante a permite ajustar los valores de min y max y obtener la imagen $I^c(m, n)$ que corresponde a la imagen corregida en contraste y de dimensiones iguales a la imagen $I(m, n)$. Los pasos para la corrección de contraste se muestran en el Algoritmo 4 y en la Figura 3.7.

Algoritmo 4. Corrección de contraste

Entradas: Imagen $I(m, n)$, segmento cuadrado de la imagen $W(j, k)$, valor constante a .

Salidas: Imagen corregida en contraste $I^c(m, n)$.

1 **inicio**

2 Sea N el número de píxeles en la imagen $I(m, n)$ y $W(j, k)$ que corresponde a una región dentro de $I(m, n)$.

3 Obtener el promedio de los valores de gris dentro de la región W : $\bar{\omega} = \sum_{m,n \in W} I(m, n)$.

4 Obtener la desviación estándar de los valores de nivel de gris de los píxeles

$$\text{dentro de la región } W: \sigma = \sqrt{\frac{\sum_{m,n \in W} (I(m, n) - \bar{\omega})^2}{N - 1}}.$$

5 Calcular un mínimo promedio de los valores de gris de los píxeles de la región W en base a la desviación estándar y al promedio W : $min = \bar{\omega} - a\sigma$.

6 Calcular un máximo promedio de los valores de gris de los píxeles de la región W en base a la desviación estándar y al promedio: $max = \bar{\omega} + a\sigma$.

7 De acuerdo a los valores obtenidos anteriormente se calcula $I^c(m, n) = \frac{255(I(m, n) - min)}{max - min}$, donde $I^c(m, n)$ corresponde a una imagen corregida en contraste que tiene el mismo tamaño de $I(m, n)$.

8 **fin**

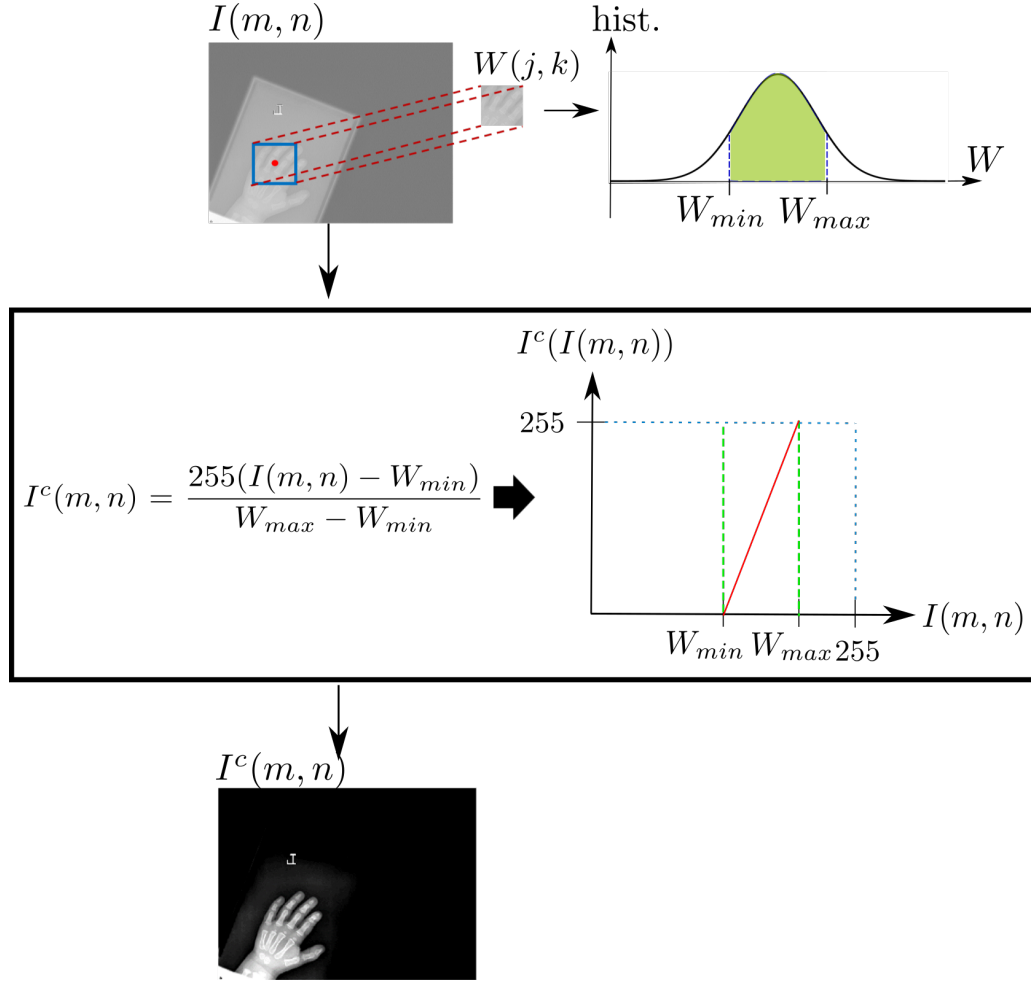


Figura 3.7. Corrección de contraste de la imagen $I(m, n)$ al tomar una región $W(j, k)$ y obtener los valores W_{min} y W_{mas} respectivos a su histograma, obteniendo la imagen $I^c(m, n)$.

3.3.1.4. Corrección angular de la imagen de la mano

Se propuso un método que permita corregir el ángulo de la imagen de la mano. Para realizarlo se debió de obtener el histograma de direcciones ponderado, para ello se inició con obtener los gradientes verticales y horizontales por medio del filtro *Sobel* de la imagen $I^g(m, n)$ también se aplicó una máscara circular binaria a la imagen y obtener $I^m(m, n)$ la cuál contiene solamente los pixeles con los niveles de gris de los huesos de la mano en su interior. Lo siguiente fue calcular su histograma polar con un total de 360 vectores, después se crearon los vectores $\mathbf{h}_i$ para $i = 1, 2, \dots, 360$, donde cada uno corresponde a la suma ponderada de una ventana que comprende a los vectores de i a $i + 90^\circ$ del histograma polar, como se observa en la Figura 3.8. A partir del histograma de direcciones ponderado $\mathbf{h}_i$ se buscó el vector con magnitud máxima $\mathbf{h}_{max1}$, se prosiguió por dibujar una línea

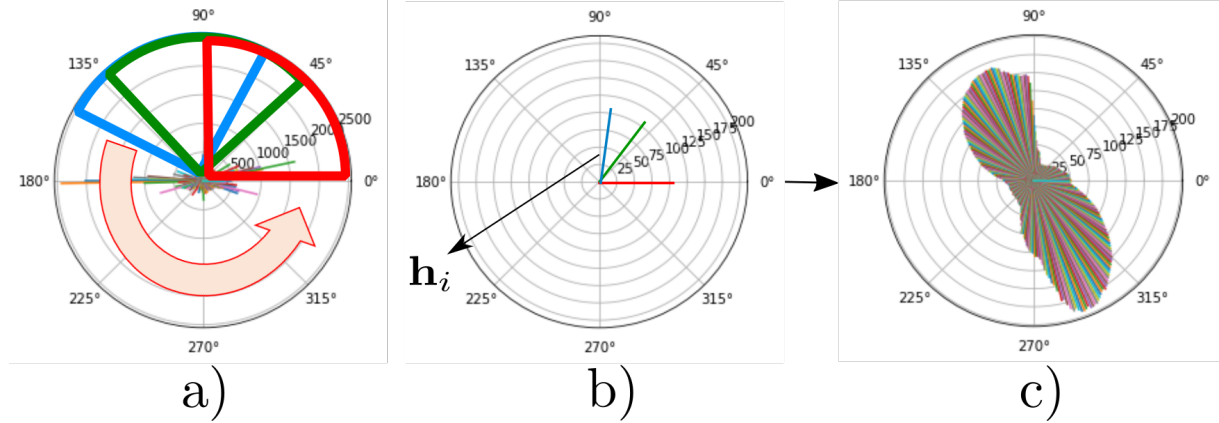


Figura 3.8. Se realizó la sumatoria ponderada de los vectores comprendidos entre una ventana de i a $i + 90^\circ$ a), obteniendo los vectores ponderados b). Este proceso se realiza por todo el plano polar (para $i = 1, 2, \dots, 360$), obteniendo así a c).

perpendicular y localizar un segundo vector máximo $\mathbf{h}_{max2}$ en la región opuesta al primer vector máximo. Se continuó por hallar el vector máximo $\mathbf{h}_M$ de los dos vectores $\mathbf{h}_{max1}$ y $\mathbf{h}_{max2}$. En algunas ocasiones al hallar el vector $\mathbf{h}_M$ y corregir la posición angular de la mano no se tenía el resultado esperado. Así se determinó crear un método que permita medir la densidad en la región de la mano y poder corregir el ángulo de manera correcta. Por lo que se percató que la parte correspondiente a los huesos de los dedos siempre presentan una mayor separación que la parte inferior de la imagen donde están los huesos del carpo. De este modo se inició por aplicar una máscara circular binaria a la imagen $I^g(m, n)$, con diámetro d_c de longitud un cuarto del tamaño de dicha imagen y centrado en su centro geométrico para obtener $I^d(m, n)$. Se dividió en dos mitades al dibujar una línea perpendicular al vector $\mathbf{h}_M$ el cuál se ubica en el centro geométrico de $I^d(m, n)$, donde la primera mitad es $C_1(m, n)$ donde están localizados los pixeles correspondientes a los huesos de los dedos y la segunda mitad $C_2(m, n)$ que contiene los huesos metacarpianos y los carpos. Así mismo se seleccionó un parche $P(m, n)$ de dimensiones un cuarto del tamaño de la imagen $I^d(m, n)$ con el fin de que determinar la densidad en las dos mitades. Por lo que para el proceso antes mencionado se realizó lo siguiente: para todos los pixeles (m, n) de $C_1(m, n)$, se centró el parche $P(m, n)$ en (m, n) y se calculó el valor $p = \sum_{m,n} P(m, n)$

(ver Figura 3.9) solo si el píxel cumple que $(m, n) > \bar{g}$ donde $\bar{g}$ corresponde al valor promedio de los valores de gris de $I^g(m, n)$, e ir guardando cada valor p en el conjunto $\mathcal{P}_1$. Para $C_2(m, n)$ también se realiza el mismo proceso anteriormente mencionado para formar un segundo vector $\mathcal{P}_2$. Por otra parte se obtuvo el valor promedio $\bar{r}_1$ del vector $\mathbf{r}_1$ y el valor promedio $\bar{r}_2$ del vector $\mathbf{r}_2$. Finalmente si se cumple la condición $\bar{r}_1 > \bar{r}_2$ y además el vector $\mathbf{h}_M$ se encuentra en la mitad $C_1(m, n)$ se rota la imagen 180° obteniendo la imagen $I_1^s(m, n)$, en otro caso si $\mathbf{h}_M$ se encuentra en la mitad $C_2(m, n)$ se rota la imagen 180° y se obtiene la imagen $I_1^s(m, n)$. Lo anterior se muestra en el Algoritmo 5.

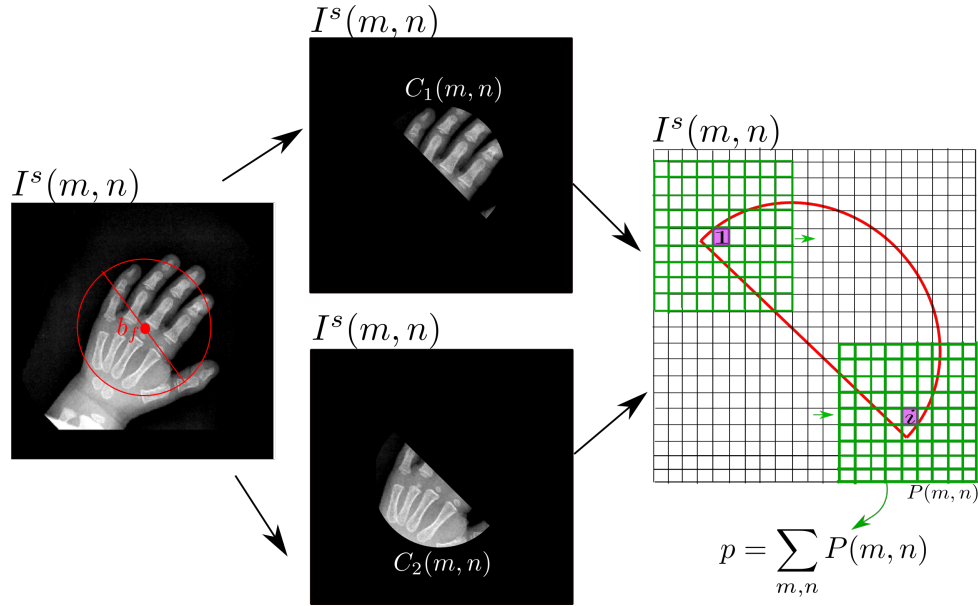


Figura 3.9. A partir de b_f se dividió en dos mitades la región de la mano siendo $C_1(m, n)$ y $C_2(m, n)$ respectivamente. Para cada mitad $C_{1,2}(m, n)$, en el primer píxel (m, n) se tomó una ventana $P(m, n)$ y si el píxel (m, n) es mayor al valor $\bar{g}$, entonces se obtiene el valor p que corresponde a la sumatoria de los valores de gris contenidos en $P(m, n)$.

Algoritmo 5. Corrección angular de la imagen de la mano

Entrada: Imagen $I^g(m, n)$.

Salida: Imagen corregida en ángulo $I_1^s(m, n)$.

1 inicio

- 2** | Aplicar una máscara circular binaria $M(m, n)$ de radio un cuarto de $I^g(m, n)$ obteniendo $I^m(m, n)$.
 - 3** | Obtener el histograma polar de los vectores de gradiente de cada uno de los píxeles de $I^m(m, n)$.
 - 4** | Obtener el histograma polar del conjunto $\mathcal{H}$ donde cada elemento $h_i \in \mathcal{H}$ corresponde a la suma ponderada de los vectores del histograma de paso 2, al tomar una ventana con los vectores de i hasta $i + 90^\circ$.
 - 5** | Obtener el vector con mayor magnitud $\mathbf{h}_{max1}$ del histograma del paso 4.
 - 6** | A partir de $\mathbf{h}_{max1}$ dibujar una línea perpendicular y localizar un segundo vector máximo $\mathbf{h}_{max2}$ en la región opuesta al primero vector máximo.
-

```
7
8   Hallar el vector  $\mathbf{h}_M$  con máxima magnitud de los vectores  $\mathbf{h}_{max1}$  y  $\mathbf{h}_{max2}$ .
9   Aplicar una máscara circular binaria de radio  $d_c$  y dimensiones de un cuarto
    de sus dimensiones a  $I^m(m, n)$  para obtener  $I^d(m, n)$ .
10  Ubicar en el centro geométrico de la imagen  $I^d(m, n)$  el vector  $\mathbf{h}_M$  para
    dibujar una línea perpendicular a dicho vector y dividir en dos mitades
     $C_1(m, n)$  y  $C_2(m, n)$  la máscara circular del paso 9.
11  Obtener el valor promedio  $\bar{g}$  de todos los niveles de gris de  $I^d(m, n)$ .
12  Sea  $j = 1$  que corresponde a la primera mitad de la circunferencia del paso 10.
13  repetir
14      para cada  $(m, n)$  de  $C_j(m, n)$  hacer
15          si  $(m, n) \geq \bar{g}$  entonces
16              Tomar una ventana cuadrada  $P(m, n)$  de tamaño un cuarto de las
                dimensiones de  $I^d(m, n)$  y centrada en  $(m, n)$ .
17              Hacer  $s = \sum_{m,n} P(m, n)$ .
18              Almacenar  $s$  en  $\mathbf{S}_j$ .
19          fin
20      fin
21       $j = j + 1$ 
22  hasta que  $j = 2$ ;
23  Obtener el valor promedio  $\bar{p}_1$  de  $\mathbf{S}_1$  y  $\bar{p}_2$  de  $\mathbf{S}_2$ .
24  si  $\bar{p}_1 > \bar{p}_2$  entonces
25      si  $\mathbf{h}_M$  esta ubicado dentro de  $C_1(m, n)$  entonces
26          Rotar  $180^\circ$  la imagen  $I^g(m, n)$  y obtener  $I_1^s(m, n)$ .
27      fin
28  sinó
29      si  $\mathbf{h}_M$  esta ubicado dentro de  $C_2(m, n)$  entonces
30          Rotar  $180^\circ$  la imagen  $I^g(m, n)$  y obtener  $I_1^s(m, n)$ .
31      fin
32  fin
33 fin
```

3.4. Etapa de entrenamiento

Por cada *landmark*, se realizó la construcción de 3 subespacios lineales usando el método de las *eigenfaces* [24]. Se etiquetaron manualmente un conjunto de 1027 imágenes con las 2 *landmarks* propuestas. Este etiquetado se realizó en las imágenes de mayor resolución para mayor precisión. Cada subespacio ó *eigenespacio* está formado por *eigenimagenes* de 32×32 píxeles para cada uno de los tres niveles de búsqueda. Las imágenes en cada nivel de búsqueda son de 64×64 , 128×128 y 256×256 píxeles. Esto puede observarse en la Figura 3.10.

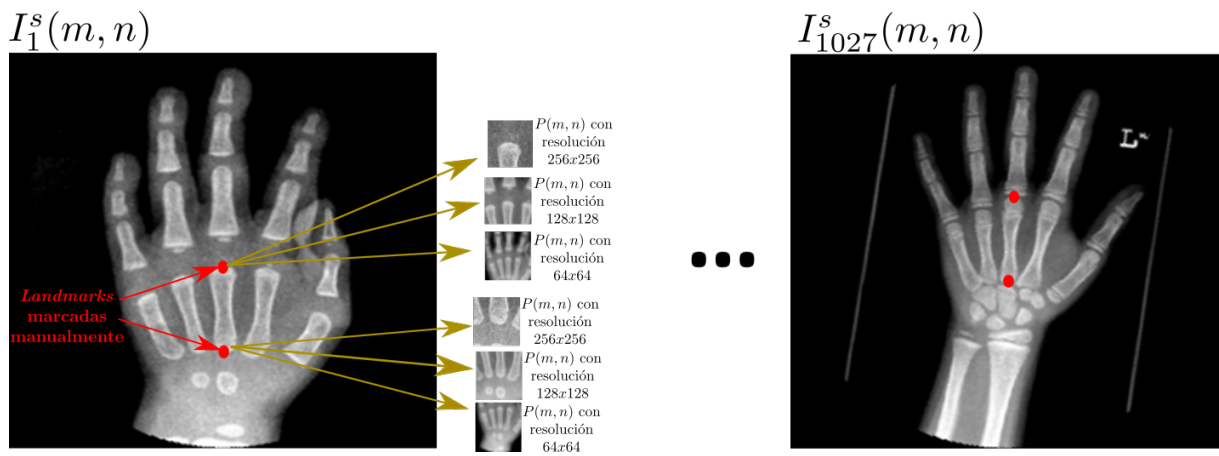


Figura 3.10. Se marcaron de manera manual las dos *landmarks* en la parte superior e inferior del tercer hueso metacarpiano para crear parches en los 3 niveles de resolución de búsqueda y crear con ellos los *eigenespacios*.

Por las imágenes de entrenamiento, se sabe que la posición probable de la *landmark* está dentro de una región rectangular cuyos límites se pueden determinar (límite izquierdo, límite derecho, límite superior, límite inferior). El límite izquierdo corresponde a la coordenada horizontal de la *landmark* posicionada más hacia la izquierda en las imágenes de ejemplo. El límite derecho corresponde a la coordenada horizontal de la *landmark* posicionada más hacia la derecha en las imágenes de ejemplo. El límite superior corresponde a la coordenada vertical de la *landmark* posicionada más hacia arriba en las imágenes de ejemplo. El límite inferior corresponde a la coordenada vertical de la *landmark* posicionada más hacia abajo en las imágenes de ejemplo. Esta delimitación permitió crear una región de búsqueda rectangular $R_3(m, n)$ dentro de la imagen $I_3^s(m, n)$ tomando los siguientes píxeles como vértices para construir la región: superior izquierdo (29, 12), superior derecho (38, 12), inferior izquierdo (29, 36), inferior derecho (38, 36), teniendo dimensiones de 24×9 píxeles, para la *landmark* superior. Para la inferior, se seleccionó una región de búsqueda rectangular $R_3(m, n)$ dentro de la imagen $I_3^s(m, n)$ tomando los siguientes píxeles como vértices para construir la región: superior izquierdo (27, 29), superior derecho (38, 29), inferior izquierdo (27, 49), inferior derecho (38, 49), teniendo dimensiones de 20×11 píxeles.

CAPÍTULO 3. DESCRIPCIÓN DEL MÉTODO

Lo anterior se observa en la Figura 3.11.

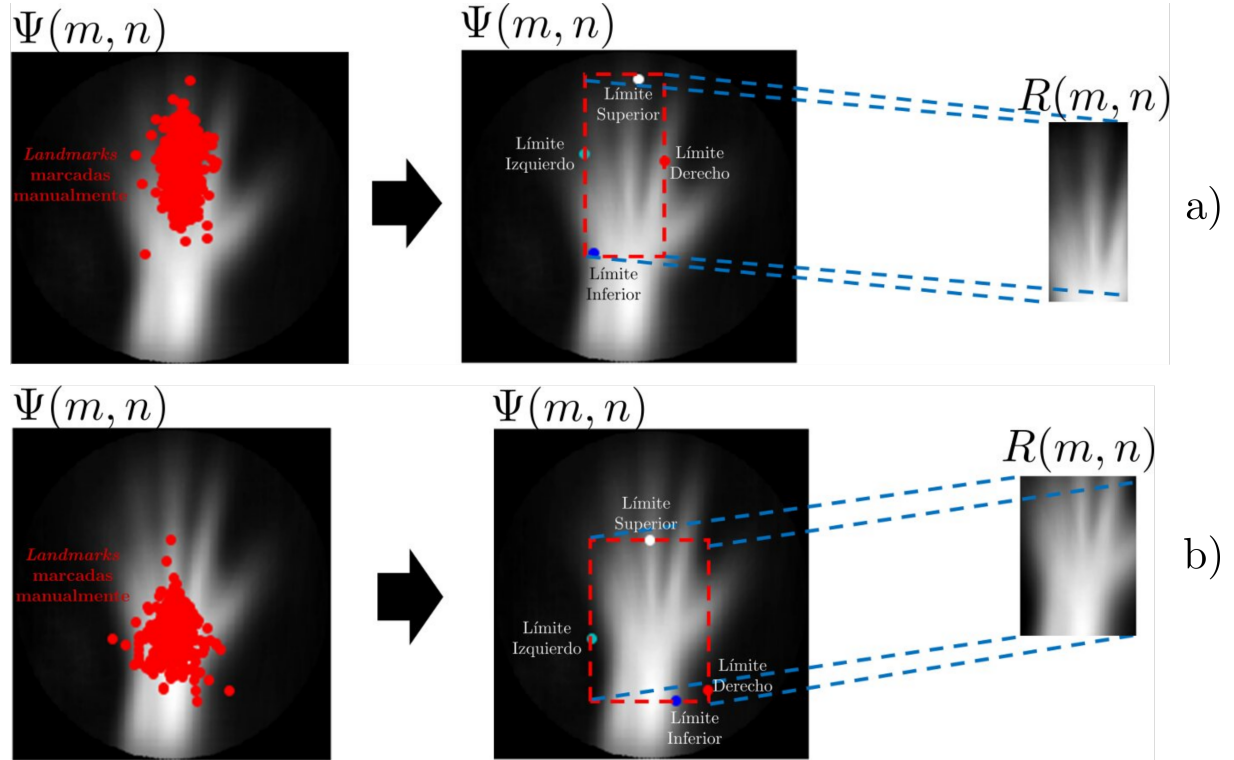


Figura 3.11. La imagen $\Psi(m, n)$ corresponde a la mano promedio obtenida al aplicar PCA y ubicar los puntos marcados manualmente, para obtener los límites: derecho, izquierdo, superior e inferior, y crear las regiones de búsqueda para la parte superior a) y para la parte inferior b).

3.5. Etapa de evaluación

3.5.1. Búsqueda de la *landmark* en diferentes niveles de resolución

En esta etapa se tuvo como objetivo la búsqueda de la *landmark* en las 3 diferentes escalas: $I_1^s(m, n)$, $I_2^s(m, n)$ y $I_3^s(m, n)$. Con tal fin primero se inició con la imagen $I_3^s(m, n)$ de 64×64 píxeles, también se tiene definida la región de búsqueda $R_3(m, n)$ dentro de dicha imagen. Se realizó la primera localización de la *landmark* l_3 , lo cual consiste en tomar distintos parches de tamaño 32×32 píxeles, siendo el píxel de dicha región el centro del parche, éste se reconstruyó en su correspondiente espacio lineal para obtener un error de reconstrucción del parche original y el reconstruido. El error se almacenó en una matriz donde, una vez finalizado toda la reconstrucción de los píxeles de la región R_3 se obtuvo el error mínimo de la matriz que corresponde a la *landmark* l_3 , estos pasos se muestran en el Algoritmo 7. Lo siguiente es redimensionar l_3 en la imagen $I_2^s(m, n)$ de 128×128 píxeles, al realizar esto la *landmark* equivale a una región cuadrada de 2×2 píxeles, que se utilizará como región de búsqueda $R_2(m, n)$ para localizar l_2 con el Algoritmo 7, al realizarlo se notó que esta localización en algunas ocasiones quedaba un poco alejado del hueso metacarpiano, así para ajustar la ubicación se agregaron 2 píxeles a cada lado de la región $R_2(m, n)$ siendo de dimensiones 6×6 píxeles, ya que agregar menos píxeles producía el mismo resultado y añadir más píxeles abarcaba zonas que ya no son de importancia al realizar la búsqueda. Se repitió la localización teniendo una mejor posición de la *landmark*. A continuación l_2 se redimensiona en la imagen $I_1^s(m, n)$, por lo tanto se obtiene una región de búsqueda también de 2×2 píxeles, agregando 2 píxeles a cada lado para formar $R_1(m, n)$ de dimensiones 6×6 píxeles. Finalmente se volvió a repetir la localización de la *landmark* al ingresar la imagen $I_1^s(m, n)$ y la región de búsqueda $R_1(m, n)$ en el Algoritmo 7, localizando la *landmark* l_1 dentro de la imagen con mayores dimensiones $I_1^s(m, n)$ y ubicada en el tercer hueso metacarpiano correspondiente. Se muestra de manera gráfica la búsqueda de la *landmark* en diferentes escalas en la Figura 3.12, en conjunto con su respectivo Algoritmo 6.

Algoritmo 6. Localización de *landmarks* a partir de una imagen en diferentes niveles de resolución

Entradas: Imágenes $I_1^s(m, n)$, $I_2^s(m, n)$, $I_3^s(m, n)$ y la región rectangular $R_3(m, n)$.

Salida: Landmark l_1 ubicada dentro de la imagen $I_1^s(m, n)$

1 **inicio**

- 2 Realizar la búsqueda de la *landmark* l_3 dentro la imagen $I_3^s(m, n)$, solo en la región de búsqueda $R_3(m, n)$ aplicando el Algoritmo 7.
- 3 Redimensionar el punto l_3 en $I_2^s(m, n)$ y aumentar dos pixeles a cada lado de la ventana para formar $R_2(m, n)$ de dimensiones 6×6 pixeles.
- 4 Localizar la *landmark* l_2 dentro de la imagen $I_2^s(m, n)$ al seguir los pasos del Algoritmo 7 e ingresar la imagen $I_2^s(m, n)$ y la región $R_2(m, n)$ en dicho algoritmo.
- 5 Redimensionar el punto l_2 en $I_1^s(m, n)$, aumentando dos pixeles a cada lado de la ventana para formar la región $R_1(m, n)$ de dimensiones 6×6 .
- 6 Localizar la *landmark* l_1 dentro de la imagen $I_1^s(m, n)$ al seguir los pasos del Algoritmo 7 e ingresar la imagen $I_1^s(m, n)$ y la región $R_1(m, n)$ en dicho algoritmo.

7 **fin**

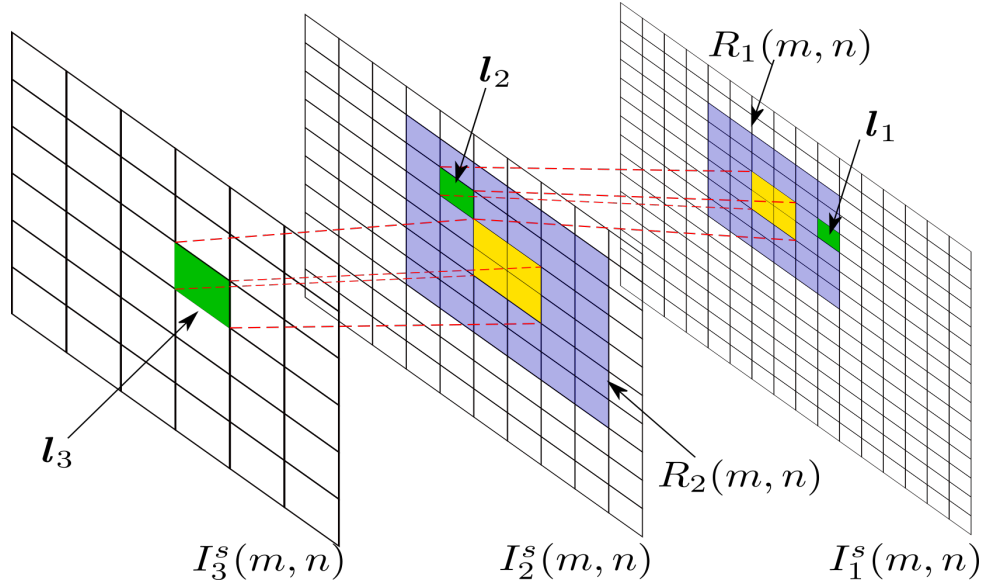


Figura 3.12. Al localizar la posición de la *landmark* l_3 en la imagen $I_3^s(m, n)$ se redimensionó el punto en la imagen $I_2^s(m, n)$ obteniendo una región de búsqueda $R_2(m, n)$, agregando dos pixeles a la región siendo de dimensiones 6×6 pixeles repitiendo la búsqueda y localizar el punto l_2 , el cual se redimensionó en la imagen $I_1^s(m, n)$ obteniendo una nueva región de búsqueda $R_1(m, n)$, también agregando dos pixeles y repitiendo la búsqueda del punto en $R_1(m, n)$, para localizar la *landmark* final l_1 .

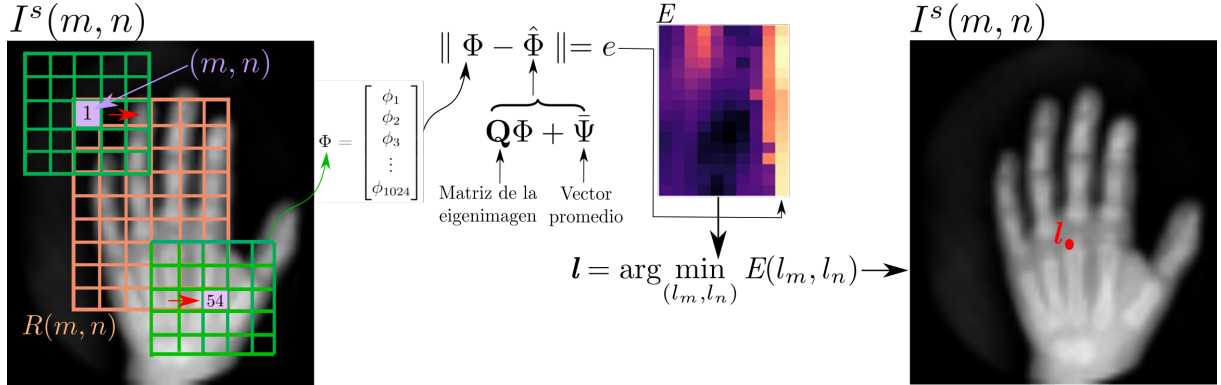


Figura 3.13. A partir de $R(m, n)$ (como ejemplo se tienen 54 píxeles), se inicia en el píxel 1 y siendo éste el centro se toma un parche que se convierte en un vector columna Φ proyectándolo en su *eigenespacio* $\mathbf{Q}$ correspondiente, obteniendo el parche reconstruido $\hat{\Phi}$ y calcular la norma e entre los dos vectores la cuál se guarda en el mapa de errores E para cada píxel de $R(m, n)$. Una vez reconstruida $R(m, n)$ se busca la posición mínima en E que corresponderá a la *landmark* l .

3.5.1.1. Proyección de parches en el subespacio de las *eigenfaces*

La búsqueda de *landmark* se realizó al tener definida una región $R(m, n)$ ubicada dentro de la imagen $I^s(m, n)$ lo que facilitó la localización, ya que buscar el punto en toda la imagen es un proceso tardado además de que hay zonas que no son de interés y se conoce que la *landmark* no se encontraría en dichas regiones. Para iniciar, lo primero fue a partir del primer píxel de $R(m, n)$ y siendo éste el centro se tomó un parche $P(m, n)$ de tamaño 32×32 píxeles. Lo siguiente fue aplicar el Algoritmo PCA al convertir $P(m, n)$ en un vector columna Φ y restarle el vector promedio $\bar{\Psi}$. Se prosiguió proyectarlo en el *eigenespacio* correspondiente $\mathbf{Q}$ de acuerdo al nivel de resolución de $I^s(m, n)$ y sumarle el vector $\bar{\Psi}$. Una vez reconstruido resultó un vector columna $\hat{\Phi}$. Posteriormente se obtuvo un error de reconstrucción que permitirá saber que tan semejante es $\hat{\Phi}$ de Φ , así entre más parecido sea el vector reconstruido del original más cerca se estará de obtener la *landmark* en la posición deseada. Esto se realiza con todos los píxeles de $R(m, n)$. Para cada uno se obtiene un error de reconstrucción y se guarda en un mapa de errores E , donde la posición (m, n) de dicha mapa corresponde al píxel utilizado en $R(m, n)$. Una vez finalizada la búsqueda en todos los píxeles de $R(m, n)$ y obtenido todos los valores del mapa se determina la posición mínima al calcular el argumento mínimo. Por lo tanto l es la posición donde se ubica la *landmark* dentro de la imagen $I^s(m, n)$. Estos pasos se muestran en la Figura 3.13 y en el Algoritmo 7.

Algoritmo 7. Búsqueda de la *landmark* al reconstruir parches de un segmento definido dentro de una imagen aplicando el Algoritmo PCA

Entradas: Imagen $I^s(m, n)$, región $R(m, n)$ dentro de la imagen $I^s(m, n)$.

Salidas: *Landmark* l dentro de la imagen $I^s(m, n)$.

```

1 inicio
2   para cada píxel  $(m, n)$  de la región  $R(m, n)$  hacer
      1. Extraer un parche  $P(m, n)$  con centro en  $(m, n)$  y tamaño  $32 \times 32$  píxeles.
      2. Convertir  $P(m, n)$  en vector columna  $\mathbf{P}$  y aplicarle PCA para
         reconstruirlo en un subespacio lineal.
      3. Calcular  $\Phi = \mathbf{P} - \bar{\Psi}$ .
      4. Calcular  $\hat{\Phi} = \mathbf{Q}\Phi + \bar{\Psi}$ .
      5. Calcular  $e = \|\Phi - \hat{\Phi}\|$ .
      6. Almacenar  $e$  en el mapa de errores  $E$ .
3   fin
4   Ubicar la landmark  $l$  de acuerdo a  $l = \underset{(l_m, l_n)}{\operatorname{argmin}} E(l_m, l_n)$ .
5 fin

```

3.5.2. Métricas de evaluación

Para poder comparar el método propuesto con los del estado del arte se tuvo que determinar el error en milímetros, pero esta información no es proporcionada por la página en donde se obtuvo el conjunto de imágenes. En consecuencia se buscó información que proporcionara la longitud en milímetros del tercer hueso metacarpiano en diferentes rangos de edad y por sexo, la cuál se obtuvo de [13]. Un problema que se percató es que el rango de edad en el que se trabajó para el algoritmo fue de 1-19 años, pero la información proporcionada por [13] es de 3-19 años, así se tuvo que realizar un ajuste por mínimos cuadrados, Figura 3.14. De modo que se obtuvo una función que permite determinar la longitud del tercer hueso metacarpiano en el rango de edad mencionado, siendo de la forma:

$$l = a - be^{-ce_o} \quad (3.5)$$

donde e_o corresponde a la edad ósea.

Así se obtuvo una función de longitud del hueso para mujeres $L3M_{mm}^f$ en milímetros donde las constantes a, b, c tienen los valores de : $a = 66.77, b = 44.15, c = 0.11$. Para la longitud de hombres $L3M_{mm}^m$ las constantes tienen los siguientes valores: $a = 86.85, b =$

66.94, $c = 0.06$. Quedando de la siguiente manera:

$$L3M_{mm}^m = 86.85 - 66.94e^{-0.06e_o} \quad (3.6)$$

$$L3M_{mm}^f = 66.77 - 44.15e^{-0.11e_o} \quad (3.7)$$

Al conocer la longitud del hueso de interés en milímetros con las funciones 3.7, 3.6 se buscó el número de milímetros en un píxel. En consecuencia se tuvo que calcular la longitud en pixeles del hueso en cuestión, esto a partir de las *landmarks* marcadas manualmente. De esta manera la longitud en pixeles para hombres es $L3M_{px}^m$, para mujeres $L3M_{px}^f$. Con lo anterior los milímetros por píxel para los dos sexos quedó de la siguiente manera:

$$mm_per_px^m = \frac{L3M_{mm}^m}{L3M_{px}^m} \quad (3.8)$$

$$mm_per_px^f = \frac{L3M_{mm}^f}{L3M_{px}^f} \quad (3.9)$$

De tal forma que para cada imagen evaluada se calculó su respectivo valor mm_per_px de acuerdo a cada sexo y se multiplica por su error en pixeles e_{px} , de esta manera se obtiene el error en milímetros e_{mm} :

$$e_{mm} = (mm_per_px)e_{px} \quad (3.10)$$

Finalmente se obtuvo el error promedio e_{mm}^p en milímetros de acuerdo a $e^p = \frac{1}{n} \sum_{i=1}^n e_{imm}$, donde $n = 250$.

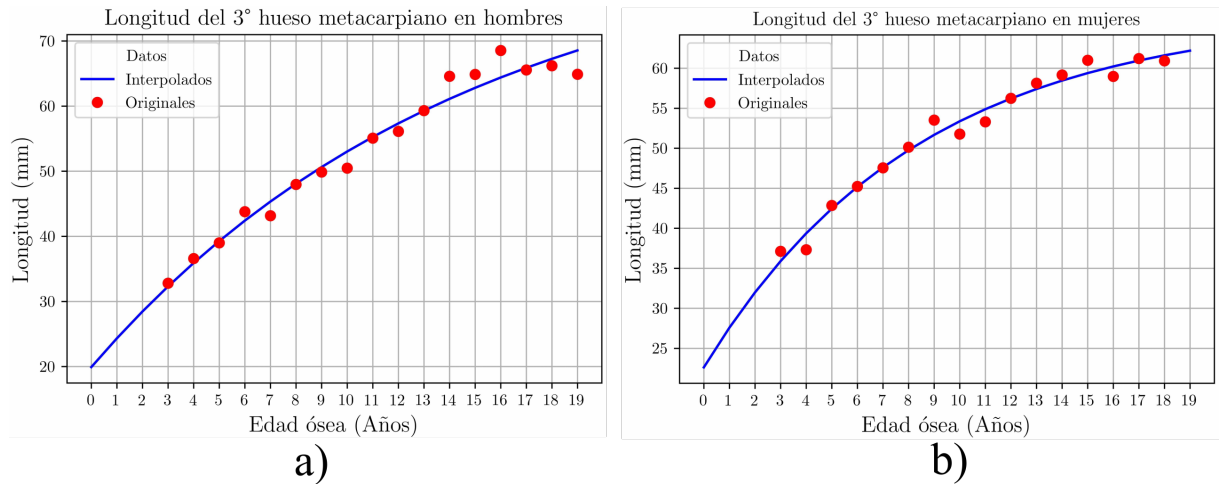


Figura 3.14. Interpolación de la longitud del tercer hueso metacarpiano en hombres a) y mujeres b), a partir de los datos proporcionados en [13].

Capítulo 4

Resultados experimentales

En total se evaluaron 250 imágenes, que no pertenecen al conjunto de entrenamiento. Al seleccionarlas se trató que el conjunto estuviera nivelado con respecto a su edad ósea, pero un problema fue que en los primeros y últimos años ya no se encontraron imágenes con respecto a las otras edades, esto se observa en la Figura 4.1.

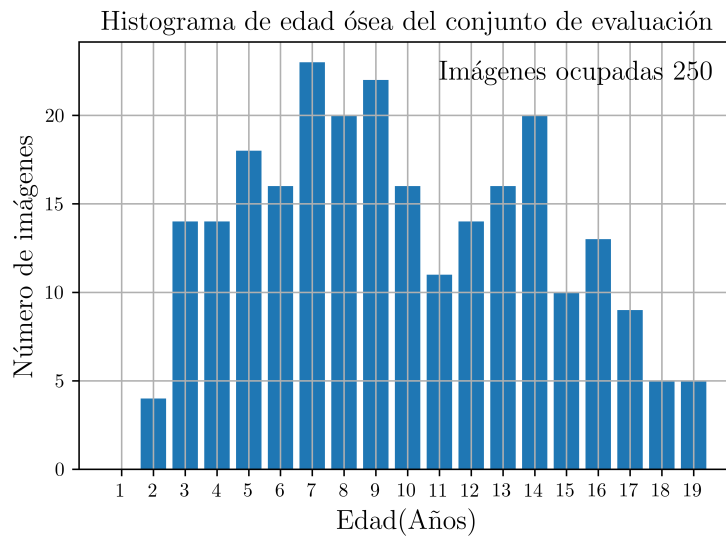


Figura 4.1. Histograma de las imágenes utilizadas en el conjunto de evaluación, mostrando la cantidad y los rangos de edad utilizados.

En la Tabla 4.1 se muestra el error estimado al marcar las *landmarks* automáticamente por el método propuesto que es de 1.15 mm promedio y 1.07 mm desviación estándar, en conjunto con otros errores obtenidos por autores con métodos en localización de *landmarks* en imágenes radiográficas de manos. Para obtener el error se obtuvo la distancia residual que consiste en comparar las imágenes marcadas automáticamente con imágenes

marcadas manualmente por un humano.

Como se observa en la Tabla 4.1 aunque el método propuesto presenta un error mayor a los mostrados, se puede notar que no son robustos a las diferentes características de las imágenes radiográficas como es a la traslación, escala, postura angular y contraste de la mano. Con lo que se obtiene un diferenciador importante, ya que además de localizar las *landmarks* automáticamente, se puede ingresar una imagen radiográfica que tenga variaciones importantes a las características antes mencionadas. Esto es que si la imagen a analizar presenta un contraste demasiado oscuro o brillante, el método es capaz de corregir esta característica de manera que se puedan observar correctamente los hueso de la mano a diferencia de [8, 9, 30, 12] que necesitan imágenes similares en contraste para aplicar su método. De igual manera si la región de la mano es bastante grande o pequeña el algoritmo es capaz de normalizar la escala de la mano y hacer que las imágenes sean similares a diferencia de trabajos como [8, 11, 30, 10], lo que significa que se tiene robustez a la escala. Asimismo el método es robusto a la traslación, ya que se encuentra de manera automática la región de la mano además de eliminar objetos que se encuentren alrededor de ella, esto sin afectar la localización de las *landmarks* en comparación de [8, 11, 30, 10, 12]. Robusto a la posición angular de la mano ya que el algoritmo es capaz de cambiar el ángulo en el que se encuentra la mano a una posición que sea adecuada para su análisis, lo que métodos como [8, 11, 10, 12] no consideran para localizar las *landmarks* en las imágenes radiográficas.

Aunque Lindner et al [9] reportan un error menor al del sistema desarrollado en esta tesis, hay que hacer resaltar que para obtener dicho error, los mencionados autores consideraron que todas las manos de las imágenes tienen un ancho de muñeca de 50 mm, lo cual puede ser una mala aproximación cuando se tiene una gran variedad en las edades de las imágenes, tal como en nuestro caso (1 a 19 años). Por otra parte, el rango de edad (5 a 18 años) de las imágenes utilizadas por Lindner et al [9] es un poco menor pero bastante comparable al rango utilizado en esta tesis. Por tanto, consideramos que aunque el error obtenido por el sistema que proponemos sea un poco mayor, es calculado de una forma más apropiada, pues se consideró la variabilidad del tamaño promedio de los huesos de acuerdo a la edad.

También en comparación con los métodos [8, 9, 11, 30, 10, 12] de la Tabla 4.1 se resalta que el método que se diseñó en esta tesis es capaz de encontrar las *landmarks* en su totalidad sin importar que alrededor de la región de la mano se encuentren artefactos que pudieran afectar dicha localización, debido a que al realizar una corrección de contraste local y aplicándola a todos los valores de gris de todos los píxeles en toda la imagen se minimiza el contraste de los artefactos que se encuentran fuera de la región de la mano a manera que no influyan en la localización de las *landmarks*. De tal manera que, las imágenes que se evaluaron habían un total de 50 radiografías con artefactos alrededor de la región de la mano y a pesar de ello las *landmarks* fueron localizadas de manera correcta.

CAPÍTULO 4. RESULTADOS EXPERIMENTALES

Tabla 4.1. Comparación del desempeño de los métodos para localización de *landmarks* investigados en el estado del arte. Aunque el método desarrollado en esta tesis tiene un error de localización ligeramente mayor al reportado por Lindner et al [9], se tiene que hacer notar que nuestro método para obtener dicho error es mucho más preciso y confiable, puesto que sí se tomó en cuenta la variación en el tamaño de los huesos en función de la edad, mientras que en [9] se supuso un tamaño fijo de ancho de muñeca (50mm) para todas las imágenes, a pesar de haber utilizado un rango amplio de edades.

Autores	Error(Pro-medio $\pm$ Desviación Estándar)	Robusto a:				
		Diversas Escalas	Traslación	Diversas posturas angulares	Mal contraste	Artefactos circundantes
Doner et al. [8]	0.77 ± 0.64	No	No	No	No	No
Lindner et al. [9]	0.61	Si	Si	Si	No	No
Payer et al. [11]	1.13 ± 0.98	No	No	No	Si	No
Štern et al. [30]	0.80 ± 0.91	No	No	Si	No	No
Urschler et al. [10]	0.80 ± 0.93	No	No	No	Si	No
Payer et al. [12]	0.66 ± 0.74	No	No	No	No	No
Propuesto	1.15 ± 1.07	Si	Si	Si	Si	Si

A continuación se muestra en la Figura 4.2 tres radiografías donde se aplicó la alineación de la imagen y la localización de las *landmarks*; el punto rojo es la *landmark* marcada en la parte superior del hueso medio metacarpiano y la azul para la parte inferior respectivamente. Después en la Figura 4.3 se muestra la comparación de las *landmarks* manuales y detectadas automáticamente. Se puede notar que la localización automática se asemeja mucho a la que un humano puede realizar.

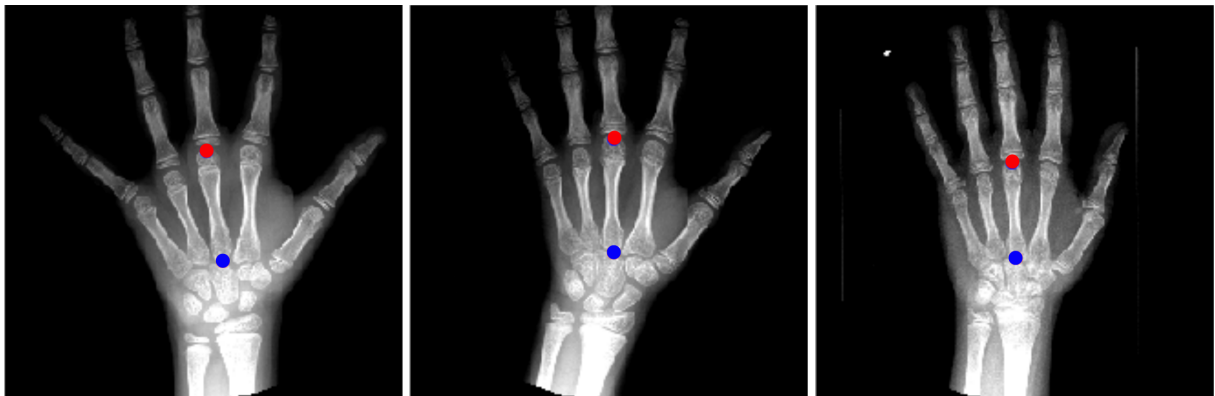


Figura 4.2. Ejemplos de la alineación de la imagen y localización automática de *landmarks* en diferentes radiografías de manos.



Figura 4.3. Comparación de *landmarks* marcadas manual y automáticamente.

4.1. Interfaz de usuario desarrollada

El siguiente paso que se tuvo que realizar fue desarrollar una interfaz que sea utilizada por un usuario. Esta debe de ser apropiada para que se pueda visualizar las *landmarks* encontradas. También que el médico sea capaz de observar con claridad las estructuras y desarrollo de los huesos de la mano. Para ello se tuvo que seguir una metodología que permita la comunicación de usuario-interfaz, así se siguió el método de casos de uso [38].

Lo que busca los casos de uso es describir el comportamiento del sistema desde el punto de vista del usuario. Por lo que permite a los actores (usuarios) saber que está realizando el software.

La metodología cuenta con el diagrama que contiene los casos de uso y una tabla por cada caso de uso que describe sus características. A continuación se muestran los diagramas 4.4, 4.5 con los casos de uso para la interfaz desarrollada y su descripción correspondiente.

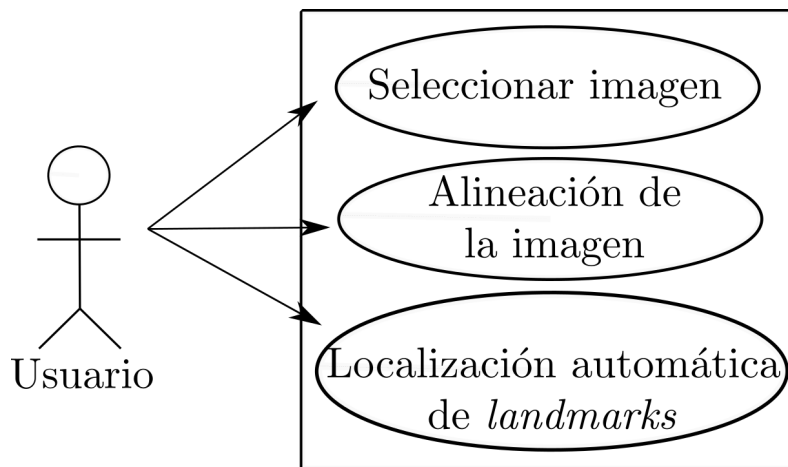


Figura 4.4. Diagrama con los casos de uso para la localización automática de *landmarks*.

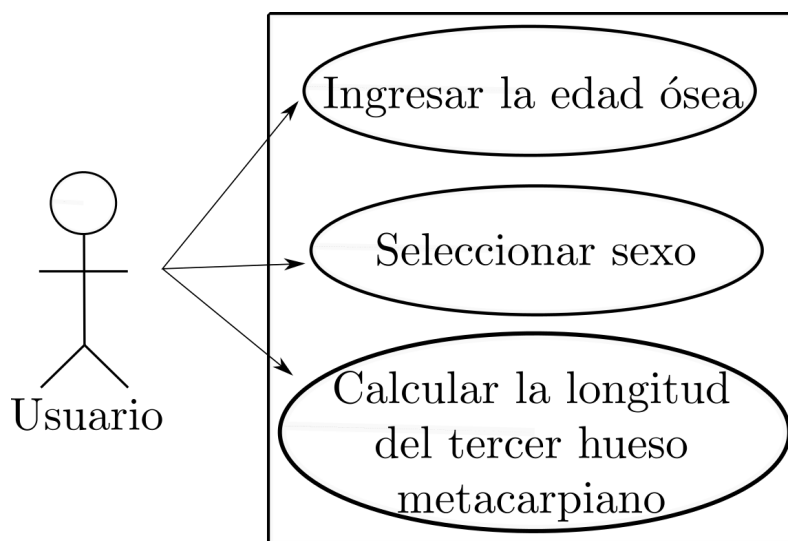


Figura 4.5. Diagrama con los casos de uso para el cálculo de la longitud del tercer hueso metacarpiano.

Tabla 4.2. Caso de uso de Seleccionar imagen.

CU-01	Seleccionar imagen
Secuencia normal	1. Al dar clic en el botón de “una imagen el sistema abre una ventana para seleccionar una radiografía que se encuentre almacenada en la computadora. 2. El usuario debe seleccionar la radiografía que deseé y dar clic en Abrir para que se inicie el análisis de ella.

Tabla 4.3. Caso de uso de Alineación de la imagen.

CU-02	Alineación de la imagen
Secuencia normal	1. El sistema obtiene la ruta donde se encuentra la imagen seleccionada para crear una carpeta llamada “Aligned_Images” y una subcarpeta nombrada “Estimated_landmarks_images”. 2. Se guarda la imagen corregida en contraste, postura angular, escala y traslación en la carpeta “Aligned_Images”.

Tabla 4.4. Caso de uso de Localización automática de *landmarks*.

CU-03	Localización automática de <i>landmarks</i>
Secuencia normal	1. A partir de la imagen alineada se localizan sus dos <i>landmarks</i> en la parte superior e inferior del tercer hueso metacarpiano. 2. Se guarda la imagen con las <i>landmarks</i> localizadas en la carpeta “Estimated_landmarks_images”. 3. Se guardan las coordenadas de las dos <i>landmarks</i> en un archivo llamado “Landmarks.csv” en la carpeta “Aligned_Images”. 4. Se muestra en la interfaz en la parte izquierda la imagen seleccionada y a la derecha la imagen alineada y con sus <i>landmarks</i> localizadas.

CAPÍTULO 4. RESULTADOS EXPERIMENTALES

Tabla 4.5. Caso de uso de Ingresar la edad ósea.

CU-04	Ingresar la edad ósea
Secuencia normal	1. El usuario ingresa la edad ósea la cual debe estar en años.

Tabla 4.6. Caso de uso de Seleccionar sexo.

CU-05	Seleccionar sexo
Secuencia normal	1. El usuario debe seleccionar de que sexo es el paciente al cual desea calcular la longitud del tercer hueso metacarpiano.

Tabla 4.7. Caso de uso de Calcular la longitud del tercer hueso metacarpiano.

CU-06	Calcular la longitud del tercer hueso metacarpiano
Secuencia normal	1. Se calcula la longitud del hueso de acuerdo al sexo seleccionado donde para mujeres se realiza con la ecuación 3.7 y para hombres con 3.6. 2. Se muestra en la interfaz la longitud aproximada del hueso en milímetros.

En la Figura 4.6 se muestra la interfaz que se desarrolló en *Qt Designer*. En ella se muestra la radiografía que el usuario seleccionó y la imagen alineada con las *landmarks* localizadas aplicando el algoritmo creado. Adicionalmente se agregó una sección que permite el cálculo de la longitud del tercer hueso metacarpiano. Para ello se le pide al usuario que ingrese la edad ósea en años del paciente y el sexo del mismo. Lo cual se logró al obtener las Ecuaciones 3.7 y 3.6 de la sección 5 del capítulo anterior.

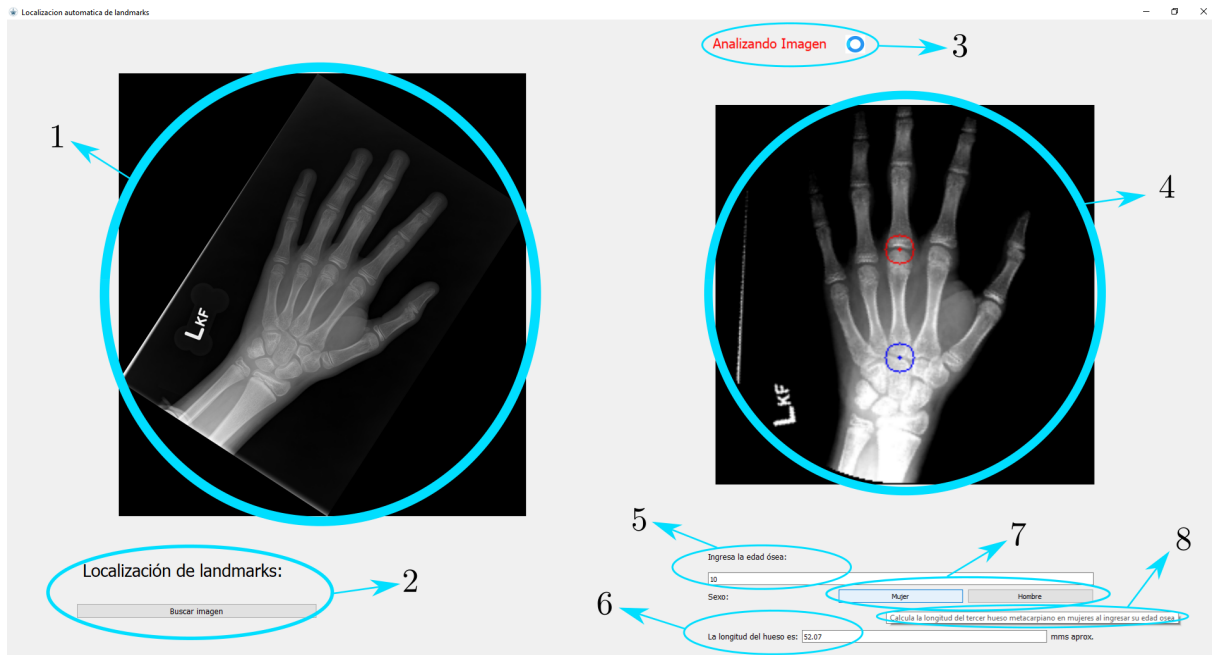


Figura 4.6. Interfaz que se desarrolló para mostrar la localización automática de *landmarks* y calcular la longitud del tercer hueso metacarpiano.

4.1.1. Descripción de cada elemento de la interfaz

1. Imagen radiográfica que el usuario selecciona.
2. Botón que permite abrir una ventana para seleccionar una imagen radiográfica almacenada en la computadora del usuario.
3. Mensaje y cursor que aparecen cuando se esta alineando la imagen y localizando sus *landmarks*. Una vez finalizado el proceso desaparece el mensaje y el cursor vuelve a su estado original.
4. Imagen alineada y con sus respectivas *landmarks* marcadas automáticamente.
5. Sección que permite al usuario ingresar la edad ósea en años de quien deseé calcular la longitud del hueso en cuestión.
6. Se muestra el cálculo aproximado de la longitud del tercer hueso metacarpiano en milímetros.
7. Botones que permiten seleccionar el sexo para el cálculo de la longitud del hueso.
8. Mensaje que describe la funcionalidad de los botones al posicionar el cursor en ellos. Estos mensajes aparecen en los tres botones que se encuentran en la interfaz.

El desarrollo del algoritmo así como las pruebas realizadas se ejecutaron en una computadora portátil con las siguientes características: procesador Intel i7 de 7° generación con 16 Gb de memoria RAM.

4.2. Discusión de resultados

La alineación de las imágenes permitió que las *landmarks* se pudieran localizar de manera adecuada. En algunos casos como es las manos de bebés o niños con edades de 2 a 3 años la alineación recortaba los huesos de los dedos, debido a que entre las falanges proximales de cada dedo existe una separación, así al seleccionar la circunferencia estos huesos quedaban por debajo del umbral que se seleccionó quedando fuera de la circunferencia. También mencionar que se tuvo que agregar una opción más al algoritmo de postura angular, ya que al tomar la densidad en la región de la mano se logró que fuera robusto a dicha propiedad.

Respecto a la aplicación del algoritmo de *eigenfaces* se pudo utilizar en las radiografías de manos de manera adecuada. Con lo que se puede notar que no solo se puede aplicar a rostros de personas como en un principio se había planteado, sino que el algoritmo es capaz de reconocer la forma que se deseó una vez que las imágenes puedan ser semejantes en sus características, que en este caso particular fue la parte superior e inferior del tercer hueso metacarpiano.

Al realizar las comparaciones del error obtenido con los del estado del arte, se notó que el algoritmo desarrollado se encuentra cerca de ellos. Una razón del por qué es un poco mayor es que se tuvieron muy pocas imágenes en los rangos de 1 a 3 años y de 17 a 19 años, que es donde se tiene un error mayor en las *landmarks* marcadas automáticamente con las marcadas manualmente.

Capítulo 5

Conclusiones y trabajo futuro

5.1. Conclusiones

Se creo un nuevo método para localizar automáticamente *landmarks*, además de algoritmos que permiten corregir las características de todas las imágenes de forma que se asemejen una de la otra. También se pudo aplicar el método de las *eigenfaces* que es un algoritmo ocupado generalmente en reconocimiento facial a las radiografías de manos. Al desarrollar y concluir todo el algoritmo propuesto se considera que las aportaciones que se hicieron en esta tesis son:

- Un método que es robusto a la escala.
- Un método que es robusto a la traslación de la región de la mano.
- Un método que es robusto a la postura angular de la mano.
- Un método para la corrección de contraste en radiografías de mano.
- Un método que es inmune a los artefactos circundantes en la región de la mano.

Con lo anteriormente mencionado se obtuvieron varios diferenciadores con los métodos revisados en el trabajo relacionado [28, 8, 11, 12, 10, 9], debido a que no nombran algunas de las características anteriormente mencionadas y solo presentan imágenes bastante similares que no se verían afectadas por su algoritmo creado.

Por lo tanto se concluye que el método es robusto a la escala ya que no importa el tamaño de la región de la mano en la imagen radiográfica, se modifica de manera que todas imágenes sean similares en tamaño. Es robusto a la traslación debido a que si la mano se encuentre a la izquierda, derecha, en la parte superior o inferior no importa, se localiza con éxito la región de la mano, además se eliminan los objetos que se encuentren alrededor de dicha región y parte de la imagen que no contiene información útil para la búsqueda de los puntos de interés. Robusto a la posición angular de la mano, ya que no interesa si la mano este al revés, ligeramente rotada, acostada o cualquiera que sea su

ángulo de inclinación las *landmarks* son localizadas. De igual forma, el método es capaz de corregir el contraste en toda la radiografía, de manera que se pueda tener una mejor segmentación y ubicación de la región de la mano. Finalmente se obtuvo un método que localiza en su totalidad las *landmarks* en imágenes que contengan objetos circundantes en la región de la mano.

5.2. Trabajo futuro

Como trabajo futuro se pretende aplicar todo el sistema desarrollado en la estimación de la edad ósea, además de crear un sistema totalmente automático capaz de ser útil a médicos radiólogos que puedan hacer uso de él. Como se ha mencionado esto reducirá el tiempo en que el médico haga una valoración en problemas de desarrollo de niños o jóvenes adultos.

Determinar la complejidad del algoritmo desarrollado, ya que esto permitirá conocer la cantidad de recursos tanto en tiempo de ejecución como de memoria utilizada. Y con esta información poder realizar una optimización de todo el código escrito.

Aplicar el algoritmo al conjunto de radiografías de manos que utilizaron los autores en el trabajo relacionado. Con esto se podrá observar como se comporta el método creado en un conjunto de imágenes diferente.

Por último se buscará tener un conjunto de imágenes con una cantidad igual entre el rango de edades utilizados. Ya que este fue uno de los problemas que no se pudo solucionar debido al tiempo y al restringido acceso que se tienen a los datos de pacientes en los hospitales.

Apéndice A

Artículos publicados

A continuación se muestra el primer artículo publicado el día 30 de septiembre del 2019 con su correspondiente constancia.



Determinación del tamaño de una mano en imágenes radiográficas utilizando landmarks

Jamid Israel Saenz Giron, Jonathan Gabriel Salazar Roldan, Salvador Eugenio Ayala Raggi y Aldrin Barreto Flores
Facultad de Ciencias de la Electrónica,
Benemérita Universidad Autónoma de Puebla, Av. San Claudio y 18 sur,
Col. Jardines de San Manuel, C.P. 72570 Puebla, Puebla, México
israel.93.ig@outlook.com,
{jonasala117, saraggi, abfver}@gmail.com

Resumen. Al estudiar y analizar imágenes radiográficas de la mano, se encuentra que no suelen estar alineadas ni comparten una buena ecualización. Debido a que las manos tienen relaciones de aspecto y ángulos diversos, encontrar de manera automática características, nos permitiría alinearlas y mejorar la relación de contraste, de este modo se facilitará el posterior trabajo con las radiografías. Para lograr este objetivo se propone un método basado en puntos de interés o *landmarks*.

Palabras Clave: Procesamiento de Imágenes, Landmarks, Set de Datos, Radiografías, Análisis por Componentes Principales, Pre-procesamiento, Eigenfaces.

1 Introducción

La localización de *landmarks* o puntos de interés en imágenes radiográficas y en este caso en particular radiografías de manos, es un paso importante para localizar regiones de interés que proporcionen información determinada, donde la más común es determinar la edad ósea [1],[2], útil para médicos radiólogos, para diagnosticar enfermedades. Desafortunadamente las zonas colindantes de los huesos de las manos resultan ser similares, por lo que se vuelve un problema desafiante, pero además se encuentran otros obstáculos como son: el contraste de las imágenes y así como diversas anomalías que pueden surgir al tomar y digitalizar la radiografía. Dichos inconvenientes ocasionan que se busquen diversos métodos para localizar *landmarks*, aplicando procesamiento de imágenes como se propone en [3]. Los autores aplican un filtro gaussiano, para extraer bordes de los huesos de la mano con la detección de bordes de Canny, a continuación, segmentan los huesos con el algoritmo Watershed y pasan la imagen a formato binario. Enseguida extraen ejes de los dedos, formando ramas que corresponden a ejes longitudinales aproximados. Se refinan con el fin de posicionarlos en el centro de cada dedo analizando secciones transversales sucesivas. En el trabajo [4], se creó un sistema

automático con el propósito de estimar la edad ósea, aplicando un método procesamiento de imágenes y visión por computadora. El método para la detección de *landmarks*, consiste en tres pasos: entrenamiento, detección inicial y refinamiento; para el primero se generaron conjuntos de parches, los cuales, se utilizarán para determinar la ubicación de los puntos en la etapa de evaluación, conjuntamente se crearon formas para considerar la zona de los puntos de acuerdo con su posición relativa y todo el conjunto de landmarks. En el área de machine learning también se han realizado métodos como en [5], donde la idea básica es usar un bosque aleatorio para determinar el vector de desplazamiento de un parche de imagen para obtener *landmarks* y buscar una zona específica del hueso de la mano como es la epífisis, la cual va determinando el crecimiento de la persona. También se han aplicado lo que son redes neuronales convolucionales como en [6], a partir de regresión de heatmap y configuración espacial, la red aprende a estimar la ubicación de los puntos y a eliminar información que ubicara de forma errónea las landmarks. Los autores además de usar radiografías de manos utilizan resonancias magnéticas de manos, encefalogramas laterales y tomografías de la medula espinal. Finalmente, en [7], se aplican métodos como es el de esquemas de registro híbrido, para modelar la deformación de la imagen como un proceso de regresión gaussiano y consideran landmarks como información adicional sobre deformaciones admisibles.

En las radiografías utilizadas, hay demasiada información, y parte de ésta no es útil, si se considera un banco grande de radiografías, es necesario obtener las zonas que aportan la información necesaria. El detectar los puntos de interés permitirá extraer solo la información que lo rodea, minimizando el poder computacional requerido. Debido a las características mencionadas se basó en el método de [8],[9], el cual es mayormente utilizado para el reconocimiento de caras, llamado Eigenfaces, compuesto por el análisis por componentes principales. Este documento se muestra como localizar landmarks de acuerdo con la obtención de información previa que es el entrenamiento y a partir de esta, localizar de manera automática puntos en las zonas del entrenamiento con radiografías nuevas. Para la primera, se extrajeron de un conjunto de imágenes sus valores característicos para crear dos eigenespacios de la parte superior e inferior del hueso medio del metacarpo y continuar con la segunda etapa llamada evaluación, donde se buscará encontrar dos landmarks en cada radiografía reduciendo la resolución y ubicar los puntos en dicha escala, y finalmente regresar a una resolución estándar la cual ubica los puntos de acuerdo con la primera localización.

2 Entrenamiento

De [10] se tomó un grupo de radiografías, para que sirvan de entrenamiento, en el grupo de entrenamiento se marcaron manualmente las *landmarks* mostrados en la Fig. 1, los puntos que nos interesa calcular automáticamente, estas marcas nos permitirán, recortar estas regiones más adelante (Fig. 3).



Fig. 1. Colocación Manual de Landmarks

Al analizar el banco de radiografías con el que se va a trabajar, es notorio una variación en el contraste de las radiografías, Fig. 2 por lo cual es necesario que se le aplique un pre-procesamiento a las imágenes, siendo un ajuste de contraste basado en los valores del histograma. Este ajuste se realiza calculando de acuerdo con las intensidades de los píxeles la media μ y desviación estándar σ . Para la media máxima y mínima se tiene que $Medmax = \mu + 1.5\sigma$ y $Medmin = \mu - 1.5\sigma$, este factor se determinó de forma empírica, de manera, por lo cual todas las imágenes ahora tienen un contraste más similar.

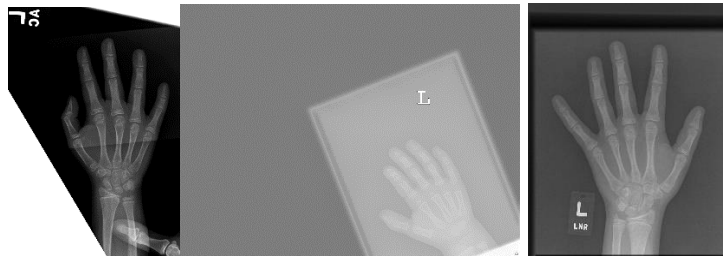


Fig. 2. Distintas radiografías del banco de imágenes, donde se aprecian los distintos contrastes de ellas.

Ahora se puede realizar un mapeo lineal de cada valor de gris en la imagen a un nuevo rango de 0 a 255. Lo siguiente fue realizar recortes de tamaño 13x13 píxeles de los puntos marcados para generar los dos espacios, Fig. 3.

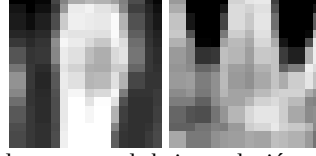


Fig. 3. Recortes del hueso del metacarpo de baja resolución, para generar los dos eigenespacios, del lado izquierdo la parte superior del metacarpo, del lado derecho la inferior.

Puesto que la posición de la mano en cada radiografía es diferente, se debe de rotar cada recorte realizado de acuerdo con su ángulo dominante θ para posicionarlo hasta que el nuevo ángulo θ sea cero. Lo siguiente es aplicar una máscara circular de forma que en pasos posteriores solo se procese lo que se encuentre en el círculo sin que ésta afecte con valores o ruido. En la Fig. 4 se muestra cómo se rotó y aplicó lo mencionado anteriormente, para cada espacio.

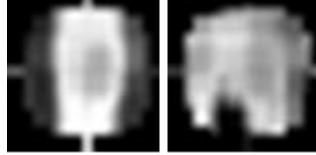


Fig. 4. Recortes del hueso del metacarpo de baja resolución para rotados con de acuerdo con su ángulo dominante θ y aplicada la máscara circular.

En este punto se tienen en total 1200 imágenes, con 600 para el primer espacio y 600 para el segundo, con lo que se procedió a llamar los espacios de Eigenmanos.

2.1 Método de Análisis por Componentes Principales (PCA)

Se busca representar datos de D dimensiones en un espacio de dimensión inferior, que mejor describan la variación, para reducir complejidad de operaciones, tiempo y espacio. Para esto el análisis de componentes principales cumple con estas condiciones, ya que se generan un conjunto ortonormal de vectores [11]. Para el entrenamiento se tienen los siguientes pasos:

1. Se tiene un conjunto de imágenes $I_1, I_2, \dots, I_M$, cada una debe convertirse a un vector $\Gamma_1, \Gamma_2, \dots, \Gamma_M$ de dimensión N^2 , concatenar cada vector formando una matriz de $N^2 \times M$.
2. Obtener el vector promedio Ψ :

$$\Psi = \frac{1}{M} \sum_{i=1}^M \Gamma_i \quad (1)$$

3. El vector promedio se debe restar a cada imagen para obtener un nuevo conjunto de vectores.

$$\Phi_i = \Gamma_i - \Psi \quad (2)$$

donde conforma la matriz $A = [\Phi_1 \Phi_2 \dots \Phi_M]$ de $N^2 \times M$.

4. Calcular la matriz de covarianza C :

$$C = \frac{1}{M} \sum_{i=1}^M \Phi_i \Phi_i^T = AA^T \quad (3)$$

de $N^2 \times N^2$

5. Calcular los eigenvectores u_i de AA^T :

Debido a que AA^T es demasiado grande y eso no es práctico, se debe buscar una aproximación de estos vectores.

- a. Calcular los eigenvectores v_i de $A^T A$:

$$A^T A v_i = \mu_i v_i \quad (4)$$

Buscar una relación entre μ_i y v_i . Se obtiene que:

$$A v_i = \mu_i u_i \quad (5)$$

- b. Normalizar μ_i de manera que $||\mu_i|| = 1$

6. Conservar solamente los K eigenvectores con más aportación.

7. Obtener la matriz

$$\Omega^T = [w_1, w_2, \dots, w_M] \quad (6)$$

y

$$w_i = u_j^T \phi_i \quad (7)$$

donde $j = 1, 2, \dots, K$; $i = 1, 2, \dots, M$.

8. Proyección

- a. Dada una imagen desconocida Γ_D

- b. Realizar

$$\Phi = \Gamma_D - \Psi \quad (8)$$

- c. Calcular la proyección de acuerdo con:

$$w_i = u_j^T \phi_i \quad (9)$$

3 Evaluación

3.1 Localización de Landmarks

Ya que se tienen las matrices de las ecuaciones 5 y 6 para la parte superior e inferior del hueso medio del metacarpo, se pasa a la evaluación de imágenes radiográficas.

Para realizar la localización, se parte de reducir su escala a la nueva imagen, delimitar la zona y localizar los puntos para que después se aumente la escala y el proceso sea más rápido, al buscar en la imagen original. Ya que se busca en imágenes sucesivas al reducir A continuación se plantean los pasos para realizar dicha búsqueda:

1. Se carga la imagen radiográfica a evaluar $M(x)$.
2. Cambiar el tamaño de la imagen, reduciendo a 255 píxeles el valor más pequeño y el otro reduciéndolo para mantener la relación de aspecto.
3. Reducir a la mitad de su escala, la cual servirá para buscar y localizar los puntos característicos.
4. Aplicar la corrección de contraste ocupada en la sección de entrenamiento.
5. Se emplea un filtro pasa-bajas para reducir aportaciones de objetos alrededor.
6. Localizar los bordes de la imagen, mediante la detección de bordes de Canny.
7. Obtener el baricentro, que deberá ubicarse dentro de la mano.
8. A partir del baricentro realizar un recorte $R(x)$ de 60 píxeles arriba, abajo, derecha e izquierda.
9. Crear una ventana que servirá para realizar la proyección de cada eigenespacio en la imagen del paso 3.
10. De acuerdo con $R(x)$ realizar lo siguiente:
 - a. Tomar una ventana del tamaño del paso 9 en la radiografía.
 - b. Rotar la ventana de acuerdo con su ángulo dominante.
 - c. Aplicar una máscara circular.
 - d. Proyectar la ventana en su eigenespacio correspondiente.
 - e. Obtener w_p y calcular el error de reconstrucción (10) para almacenarlo en una matriz de errores

$$e = ||w_p - w_j|| \quad (10)$$

11. Localizar en la matriz de errores el valor mínimo para cada eigenespacio, para

marcarlo como landmark.

4 Resultados

De la base de datos de la cual se extrajeron las imágenes de entrenamiento [9] se tomó un total de 200 imágenes para evaluar el método propuesto. A las radiografías entrantes, se reduce de escala y se les ajusta el contraste, esto es para que a continuación se pueda localizar el baricentro, el cual nos permite acotar la zona de búsqueda ya que este siempre está en dentro de la mano, Fig. 5. Y posteriormente se recorta la zona donde se buscarán las landmarks, Fig. 6.

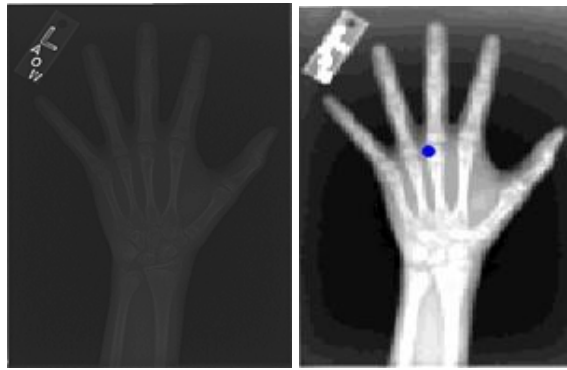


Fig. 5. Pasos iniciales para la localización de landmarks: a la izquierda, la entra original, a la derecha la imagen reducida en escala, con ajuste de constarte y localización del baricentro.

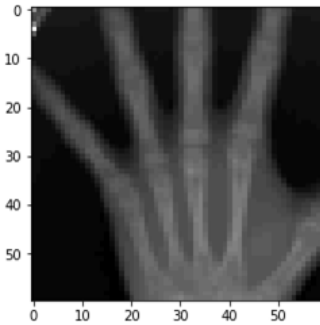


Fig. 6. Recorte de la radiografía a partir del baricentro localizado

Sobre esta sección de la radiografía de escala reducida, se fue extrayendo una ventana para ser procesada individualmente, la cual en este primer paso es de menor resolución que la imagen original, para proyectarla en los dos eigenspacios, esta reconstrucción nos dará un error , este nos indicara que tan diferentes son las ventanas de nuestro espacio, el error lo podemos ver en la Fig. 7, la posición donde el

error sea menor, implicara que es el punto que se busca, en este se tomara una ventana con mayor resolución y se repetirá el procedimiento hasta tener toda la resolución como se puede ver en Fig. 8. La cual nos muestra los puntos de interés encontrados.

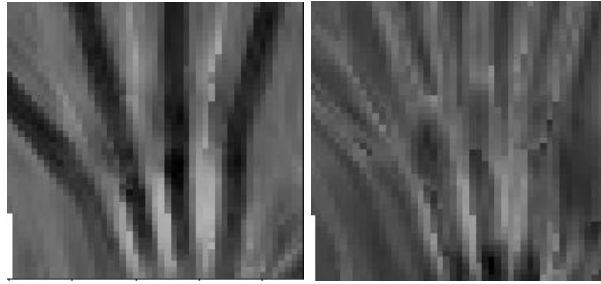


Fig. 7. Matrices de errores en los distintos eigenespacios, se busca el valor mínimo del error para ubicar la landmark, del lado izquierdo la parte superior del hueso del metacarpo y del lado derecho la parte inferior del hueso del metacarpo.

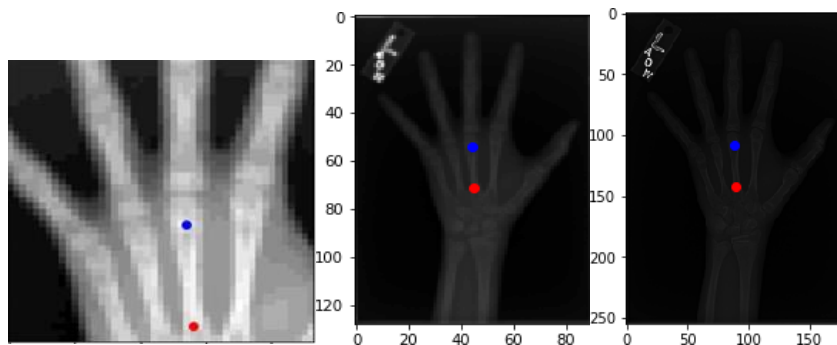


Fig. 8. Localización de landmarks en distintas escalas, se fueron localizando de izquierda a derecha.

El obtener estos puntos nos permitirá acotar de una mejor manera la mano para trabajar con ella, ya que muchas imágenes radiográficas incluyen bordes u otros objetos que no aportan información como las mostradas en Fig. 9.

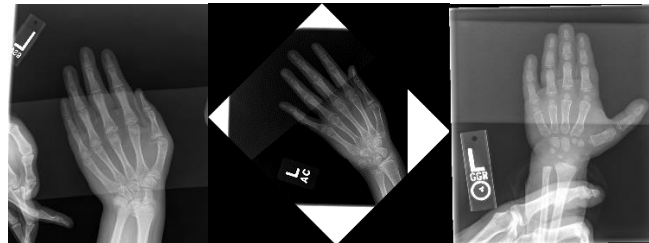


Fig. 9. Ejemplos de radiografías que contienen objetos extraños.

4 Conclusiones y trabajos futuros

En este trabajo se desarrolló un método para obtener de manera automática landmarks, estos puntos que se localizaron permitirán acotar un área en las radiografías la cual es la que trae la mayor aportación de la edad ósea, esta se utilizara para seguir desarrollando un trabajo para detectar la edad ósea de manera automática. En un futuro se aplicara este mismo método para marcar más áreas en la mano, con el fin de refinar las zonas de donde se extraerá la información, siendo estas la parte intermedia de los huesos tanto de las falanges proximales como de los huesos del metacarpo, el poder calcular la edad ósea de manera automática y eficiente, es importante ya que en la radiología es necesaria para detectar síntomas o enfermedades en menores de edad y atender enfermedades que en el futuro dañen la salud del menor siendo las más comunes: trastornos de crecimiento, problemas ortopédicos, enfermedades a nivel hormonal, entre otras.

Agradecimientos

Agradecer a Mauricio Longinos Garrido y a nuestro asesor Salvador Eugenio Ayala Raggi por su apoyo y asesoramiento en la realización de este proyecto.

Referencias

1. Tanner, J.; Whitehouse, R.; Cameron, N.; Marshall, W.; Healy, M.; Goldstein, H.: Skeletal maturity and prediction of adult height (tw2 method). *Archives of Disease in Childhood*, Vol.51, No. 6, pp. 488 (1976).
2. Greulich, Pyle, S., Radiographic atlas of skeletal development of hand wrist. *Quarterly Review of Biology*, Vol.26, Standford University Press. (1959).
3. Moreno, E. M.; Lopez, C.A.; Angel, M.; Andez, I.: Automatic detection of landmarks for image registration applied to bone age assessment. *5th WSEAS Int. Conf. on Signal Processing, Computational Geometry & Artificial Vision*, pp. 117-122, (2005).
4. Banda, J.L.T.: Sistema automático de aprendizaje incremental supervisado para reconocimiento de la edad ósea. Benemérita Universidad Autónoma de Puebla, 2018.
5. Hanhmann, F.; Böer, G.; Gabriel, E.; Deserno, T.; Meyer, C.; Scramm, H.: Classification of Voting Patterns to Improve the Generalized Hough Transform for Epiphyses Localization. . *Medical Imaging 2016: Computer-Aided Diagnosis*, Vol. 9785, No. 0, pp. 1–11, (2016).
6. Payer, C.; Štern, D.; Bischof, H.; Urschler, M.: Integrating spatial configuration into heatmap regression based CNNs for landmark localization, *Medical Image Analysis*, Vol. 54, pp. 207–219, (2019).
7. Lüthi, M.; Jud, C.; Vetter, T.: Using Landmarks as a Deformation Prior for Hybrid Image Registration. *Pattern Recognition*, pp. 196-205, (2011).
8. Turk, M.; Pentland, A.: Eigenfaces for Recognition. *Journal of Cognitive Neuroscience*, Vol. 3, No. 1, pp. 71-86, (1991).
9. Turk, M.; Pentland, A.: Face recognition using eigenfaces. *International Conference on Pattern Recognition*, pp. 586-591, (1991).
10. Mader, K.: RSNA Bone Age Predict Age from X-Rays. Kaggle.

<https://www.kaggle.com/kmader/rsna-bone-age> (2017). Accedido el 28 de Mayo de 2019.

11. Joliffe, I.T.: Introduction: Principal Component Analysis. Springer, pp. 1-9 (2002).

APÉNDICE A. ARTÍCULOS PUBLICADOS

Segundo artículo publicado el día 13 de noviembre del 2020 con su correspondiente constancia.



La Benemérita Universidad Autónoma de Puebla

A través de la Facultad de Ciencias de la Computación
otorga la presente

CONSTANCIA

A: Jamid Israel Saenz Giron, Salvador Eugenio Ayala Raggi, Mauricio Longinos Garrido, Aldrin Barreto Flores, José Francisco Portillo Robledo

Por presentar el artículo, "*Localización automática de landmarks con robustez a la traslación, postura angular y escala, y su aplicación en la región del tercer hueso metacarpiano en imágenes radiográficas de la mano*" en el II Congreso Internacional y X Congreso Nacional de Ciencias de la Computación CONACIC2020, del 9 al 13 de noviembre de 2020.

"Pensar bien, para vivir mejor"

Puebla de Z., a 13 de noviembre de 2020.


M.I. María del Consuelo Molina García
Directora.

Localización automática de *landmarks* con robustez a la traslación, postura angular y escala, y su aplicación en la región del tercer hueso metacarpiano en imágenes radiográficas de la mano

Jamid Israel Saenz Giron, Salvador E. Ayala-Raggi, Mauricio Longinos Garrido,
Aldrin Barreto-Flores, José Francisco Portillo-Robledo

Benemérita Universidad Autónoma de Puebla,
Facultad de Ciencias de la Electrónica, Av. San Claudio y 18
Sur, Jardines de San Manuel, C.P. 72570 Puebla, Puebla, México
jamid.saenz@alumno.buap.mx
salvador.raggi@correo.buap.mx
longinos.200918526@gmail.com
aldrin.barreto@correo.buap.mx
francisco.portillo@correo.buap.mx

Resumen En este artículo se propone un método nuevo de localización automática de *landmarks* en imágenes radiográficas que es robusto a cambios en posición, postura angular, tamaño y contraste. Proponemos un método de estimación de la posición de la mano en una imagen radiográfica, basado en una mejora iterativa del cálculo del centro de masa. Se propone, además, un algoritmo basado en cambio de densidad para estimar el tamaño de la mano en la imagen. Así mismo, presentamos un método fundado en la dirección de los gradientes de la imagen para realizar la estimación de la postura angular de la mano. Finalmente, presentamos un enfoque multiescala, que utiliza eigenespacios, para realizar la localización automática de las *landmarks*. Los resultados experimentales demuestran que el método propuesto consigue un error similar al de los métodos actuales, pero con la ventaja de la robustez a las variaciones en posición, postura angular, escala, y contraste de la región de interés.

Keywords: *Landmark* localisation, eigenfaces, radiographic image registration

1. Introducción

La variación en apariencia existente en las imágenes radiográficas representa un problema desafiante. Esto debido a diversos factores como el contraste, el cual no permite distinguir correctamente los huesos, el tamaño relativo de la mano en la imagen, la posición y postura angular de la mano, la variación particular de cada individuo en la forma de las estructuras óseas, y la intromisión de artefactos y otros objetos ajenos a la mano alrededor de los huesos y de la región de interés. La localización de *landmarks* es utilizada para buscar regiones de interés en distintas imágenes médicas tanto en 2D como son las radiografías y en 3D que corresponden las imágenes de resonancia magnética. En lo que respecta a la detección automática de *landmarks* en imágenes de manos, se han propuesto métodos como en [1] donde proponen la localización de puntos en radiografías al realizar de una búsqueda a diferentes escalas de la misma imagen, por medio de parches y al proyectarlos en un subespacio lineal para determinar si es semejante a las imágenes utilizadas en el entrenamiento realizado, y así ubicar las

landmarks, este método lo utilizan dichos autores en imágenes 2D y 3D. En [2] se utiliza el método de machine learning de bosques aleatorios en conjunto con el método de modelo de forma precisa y ampliamente utilizado en imágenes de rostros, para ello crean distintos modelos de apariencia de la mano y por medio de los bosques aleatorios se va seleccionando la región donde se encuentran las *landmarks*. De manera similar, en [3] utilizan los bosques aleatorios para localizar *landmarks* e identificar las zonas similares en las cuales los puntos podrían localizarse, y hacen uso de imágenes tanto en 2D como en 3D. Siguiendo el uso de bosques aleatorios, en [4], lo utilizan para modelar su regresor local basado en apariencia, combinar la información de apariencia de la imagen con la configuración geométrica de *landmarks* para ir localizando distintos puntos, también en imágenes 2D y 3D tanto en manos como en radiografías cefálicas laterales e imágenes de cuerpo completo. En [5] se utiliza otro método de machine learning, las redes neuronales convolucionales (Convolutional Neural Networks) en conjunto con el método de mapas de calor por regresión (Regressing Heatmaps) para imágenes de manos en 2D y 3D, en el cual proponen una nueva arquitectura con las redes neuronales convolucionales llamada Spatial Configuration-Net (SCN). En años recientes los autores en [6] utilizan su arquitectura SCN al localizar *landmarks* con un error bajo y además con un conjunto limitado de entrenamiento para dicha localización. Los métodos mencionados, realizan la localización de las *landmarks* y reportan errores promedio de unos pocos milímetros, sin embargo, la mayoría de ellos resuelven el problema de la localización cómo una optimización local [2], y no global, en donde la búsqueda debe realizarse muy cerca del lugar donde yacen las *landmarks*. Algunos otros, aunque con localización global [1], no reportan robustez a la escala y la rotación. Por esta razón el método aquí propuesto busca primero encontrar la mano completa en la imagen, normalizarla en tamaño y en postura angular para finalmente realizar la búsqueda de las *landmarks*. Por lo tanto proponemos un método nuevo multiresolución para la localización automática de *landmarks* que es robusto a la posición, postura angular, y escala de la mano. Hemos evaluado el nuestro método localizando automáticamente dos *landmarks* en el tercer hueso metacarpiano, uno ubicado en la parte superior y otro en la parte inferior del hueso mencionado.

2. Descripción general del sistema

El método propuesto consiste en dos etapas principales: entrenamiento y evaluación. Para las dos etapas se determina la ubicación del baricentro de la mano, el cual nos permite estimar la región de la mano para eliminar objetos innecesarios que puedan aparecer en la región de la mano. El siguiente paso es encontrar automáticamente la región de interés donde se encuentre la mano para posteriormente aplicar una corrección de contraste en la región central de la mano que permita tener un contraste parecido en dicha región de la mano tanto en las imágenes del conjunto de entrenamiento como en la imagen de evaluación en la que se desea realizar el proceso automático de la localización de las *landmarks*. El paso siguiente será corregir de forma automática la postura de la mano (ángulo) utilizando para ello un algoritmo de reconocimiento automático del ángulo dominante de los gradientes de intensidad de la imagen dicho algoritmo es propuesto en este artículo. En la etapa del entrenamiento cada imagen se marca manualmente con dos *landmarks*, una en la parte superior y la otra en la parte inferior del tercer hueso metacarpiano. Dichas *landmarks* servirán como puntos de referencia centrales para extraer dos regiones de interés alrededor de cada una de estas dos *landmarks*. El algoritmo de detección que se propone en este artículo es multiresolución, con el objetivo de agilizar el proceso de búsqueda de cada una de las dos *landmarks*. Primero, se realiza la búsqueda de una de las *landmarks* en una resolución muy baja y posteriormente en resoluciones cada vez más altas. En la práctica utilizaremos una pirámide de tres resoluciones distintas. Por tanto, tendremos una región de interés alrededor cada la *landmark* por cada una de las tres resoluciones. Se utilizará el método de las eigenfaces [7] para representar cada una de las tres regiones

de búsqueda para cada una de las dos *landmarks*. Con dicho método podemos comparar o proyectar en el subespacio lineal generado por el método de las *eigenfaces* alguna región candidata al momento de hacer la búsqueda y comprobar si el subespacio de las *eigenfaces* es capaz de generar una reconstrucción adecuada de la región. La búsqueda será comenzada en la menor resolución y terminara en la de mayor resolución, y las regiones de búsqueda alrededor de la *landmark* en cuestión abarcaran una región más pequeña de la mano conforme cambiemos a una resolución mayor. Esto se observa en la figura 1.

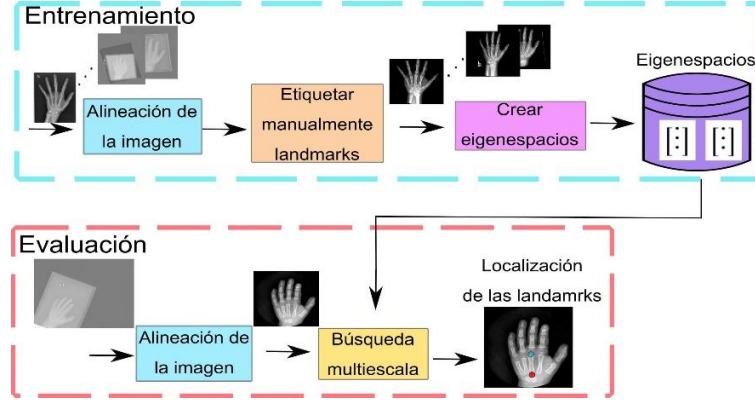


Figura 1. Diagrama a bloques del método propuesto para la localización automática de *landmarks*.

3. Método propuesto

3.1. Descripción General del proceso de Alineación de Imagen

Debido a que el método de las *eigenfaces* funciona mejor cuando las imágenes de entrenamiento son parecidas entre sí, se requiere realizar un pre acondicionamiento de las imágenes de entrenamiento con el objetivo de que todas ellas tengan el mismo tamaño de imagen, y además, la región de la mano aparezca con la misma postura (tanto en posición, cómo en ángulo) y con una escala uniforme en todas. Así mismo, también se busca uniformizar el contraste de la región de interés dentro de la mano. En pocas palabras, se realizará una alineación de las imágenes de entrenamiento en cuatro aspectos diferentes: traslación, rotación, escala y contraste. De esta manera, llamaremos $I(m, n)$ a la imagen original, y $I^q(m, n)$ a una imagen intermedia alineada en posición, escala y contraste. La imagen de salida de dicho algoritmo es una imagen cuadrada de 256x256 píxeles donde la mano se encuentra centrada y corregida en contraste. El último paso de la alineación de la imagen es realizar una corrección de la postura angular de la mano y tener la imagen $I^s(m, n)$.

3.2. Segmentación de la mano

En esta sección se busca obtener imágenes segmentadas que nos permitan disminuir las diferencias entre cada una de ellas. Primero se modifica el tamaño de la imagen de entrada $I(m, n)$, al determinar su dimensión más pequeña y cambiarla a 256 píxeles. La otra dimensión se modifica a un valor proporcional p , para obtener una nueva imagen $I^p(m, n)$. Lo siguiente es localizar el baricentro de la mano b_n , (sección 3.3). Después se eliminan los objetos alrededor de la mano (sección 3.4) para obtener la imagen $I^c(m, n)$ y el radio de la mano r_c . Se continua al ubicar el baricentro en la imagen de entrada $I(m, n)$ al multiplicar b_n por el valor proporcional p y multiplicar dicho valor por el radio r_c para aplicarle una máscara binaria $M(m, n)$ circular a la imagen. Se procede a realizar un corte sobre $I(m, n)$ y obtener la región $C(m, n)$ siendo el centro geométrico de la

imagen el baricentro (m_f, n_f) . Se cambian las dimensiones de la región a 256x256 píxeles para obtener la imagen $F(m, n)$. Finalmente la imagen se corrige en contraste para obtener $I^g(m, n)$. Lo anterior se muestra en la figura 2.

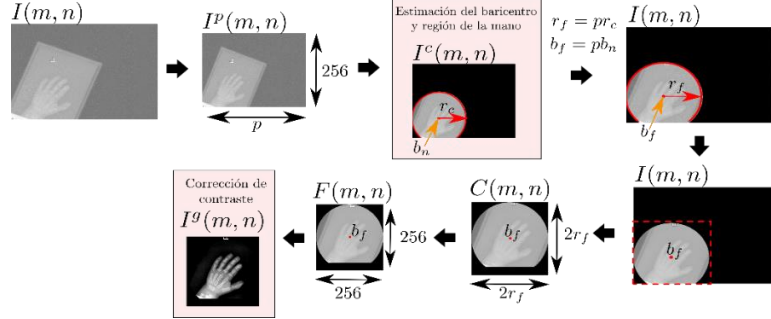


Figura 2. A partir de $I(m, n)$ se modifican las dimensiones para localizar el baricentro b_n , la región de la mano en conjunto con r_c y trasladarlos en $I(m, n)$. Se continúa aplicando un corte dentro de la imagen $I(m, n)$ creando $C(m, n)$, además se cambian sus dimensiones para tener $F(m, n)$. Por último se realiza una corrección de contraste y obtener la imagen $I^g(m, n)$.

3.3. Estimación del baricentro de la región de la mano

Se propuso localizar el baricentro, ya que a partir de éste se podrá dibujar una circunferencia que nos permitirá obtener solo la región de la mano. Para ello se seleccionó un radio inicial al considerar las manos más grandes, con el fin de ir reduciéndolo. Lo siguiente es obtener los bordes y localizar el baricentro por primera vez para tener una aproximación y trasladarlo con la misma posición en la imagen inicial $I(m, n)$. Se continúa con corregir el contraste de la imagen. A continuación se repite la detección de bordes con la imagen resultando una nueva imagen $E_2(m, n)$. Además de volver a localizar por segunda vez el baricentro en $E_2(m, n)$. A partir del punto $b_a = (m_{b2}, n_{b2})$ y siendo éste el centro, se dibuja una circunferencia de radio r para comenzar el ajuste en la posición del baricentro. Así se inicia por localizar dentro de la circunferencia y guardar la posición del baricentro en la variable b_a , e ir reduciendo el radio en 10 píxeles, repitiendo la localización del nuevo baricentro y guardarlo en $b_n = (m_{b3}, n_{b3})$, como se muestra en la figura 3. Además de tener dos condiciones de paro al ubicar el punto. Para la primera se mide la distancia euclidiana entre b_a y b_n y si es menor o igual a 1 píxel, se detiene la localización. Para la segunda condición se tomó en consideración que a partir de que el radio llegara al valor de 30 píxeles se detenga la localización. Al cumplirse una de las dos condiciones se mantiene la ubicación final de b_n .

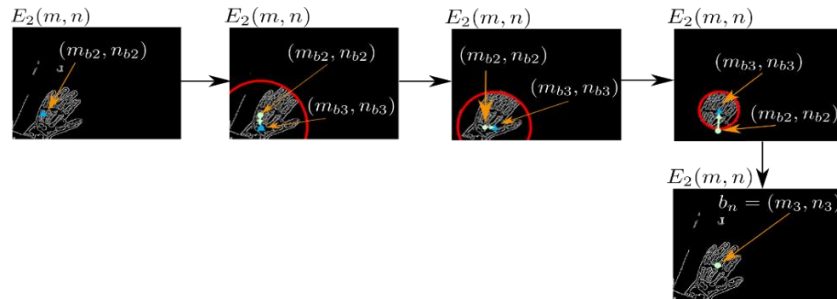


Figura 3. Localización del baricentro en la región de la mano en la imagen $E_2(m, n)$. El punto (m_{b2}, n_{b2}) representa el punto anteriormente ubicado y (m_{b3}, n_{b3}) el nuevo baricentro localizado, además de la reducción de la circunferencia para la búsqueda del baricentro final b_n .

3.4. Estimación de la región de la mano

Se buscó eliminar objetos alrededor de la mano ya que no aportan información en la búsqueda de *landmarks*. Se comienza al asignar a la constante c un valor de 1.32. Continuamos creando dos conjuntos, uno de radios R con 18 elementos, por lo que para $r_i \in R$ se calcula $r_i = 30c^n$ y el segundo para áreas de cada elemento del conjunto de radios. Se prosigue con un tercer conjunto M que contiene las 18 imágenes rectangulares conformadas por máscaras binarias donde el interior de la circunferencia se conforma por valores de 1 y en el exterior valores de 0. Continuando se multiplica la imagen de bordes $E(m, n)$ por cada elemento del conjunto M para crear el conjunto M_E . Para formar el conjunto H que contiene coronas circulares formadas por la resta absoluta de la imagen binaria rectangular de bordes M_{Ei} por su elemento posterior M_{Ei+1} (figura 4) a excepción del primero elemento H_1 conformado por la primera imagen de bordes binaria con la máscara circular M_{E1} . De manera similar se crean coronas circulares pero en el conjunto de áreas A , donde el primer elemento del nuevo conjunto corresponde a $H_{A1} = A_1$ y realizando $|A_i - A_{i+1}|$. Por lo tanto para cada $H_i \in H$ y $H_{Ai} \in H_A$ se calcula $D_i = H_i/H_{Ai}$ para $D_i \in D$. Por consiguiente se seleccionó un valor umbral que permita conocer en que circunferencia se encuentra la mano. Para facilitararlo se normalizaron todos los valores del conjunto D al ordenarlos de mayor a menor y se encontró que al considerar los dos primeros máximos del conjunto siendo D_1, D_2 , se tiene una mejor consideración del umbral, así se obtiene el valor promedio $\bar{m}$ de los valores antes mencionados y se normaliza todo el conjunto D al dividir cada elemento D_i entre $\bar{m}$, creando el conjunto E . Para poder concluir con un valor umbral de 0.2. Así para cada nueva imagen, se obtiene su conjunto E y conocer que circunferencia debe aplicarse. Esto se realiza con las variables *contador* y r_c . Los pasos posteriores se muestran en el algoritmo 1.

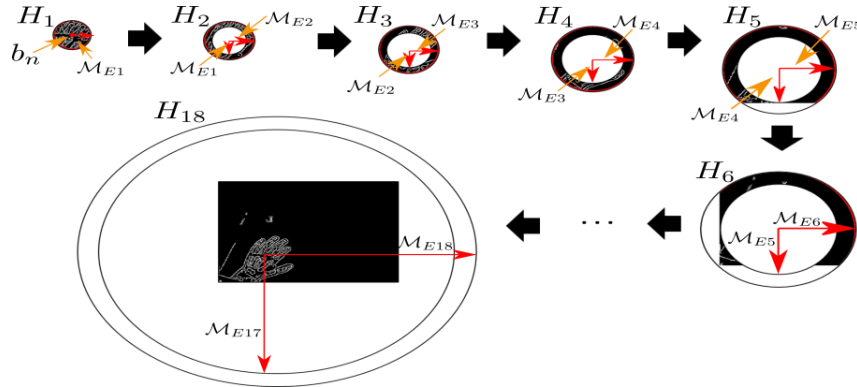


Figura 4. Creación de las coronas circulares con la imagen de bordes $E(m, n)$, la cual se inicia con el baricentro b_n siendo el centro y comenzar con el radio M_{E1} para formar el primer elemento H_1 , se continua con la corona formada por H_2 hasta formar la corona H_{18} .

Algoritmo 1: Estimación de la región de la mano basada en cambios de densidad usando una región circular creciente.

Entradas: Imagen $I(m, n)$, Imagen de bordes $E(m, n)$ y baricentro (m_b, n_b) .

Salidas: Imagen de la mano con su circunferencia contenida $I_E(m, n)$, radio de la circunferencia r_c .

1 inicio

2 Inicializar $c = 1.32$.

3 Crear el conjunto de radios $R = \{r_1, r_2, r_3, \dots, r_{18} | r_i = 30c^n\}$.

4 Crear el conjunto de áreas $A = \{A_1, A_2, A_3, \dots, A_{18} | A_i = \pi r_i^2\}$.

5 Crear el conjunto de imágenes binarias
 $M = \{M_1(m, n), M_2(m, n), \dots, M_{18}(m, n)\}$.

6 Realizar $M_E = \{E(m, n)M_i(m, n) | M_i(m, n) \in M\}$.

7 Crear el conjunto de coronas circulares $H = \{H_1, H_2, \dots, H_{18}\}$.

8 Crea el conjunto de áreas de las coronas circulares $H_A = \{H_{A1}, H_{A2}, \dots, H_{A18}\}$.

9 Calcular la densidad de área de cada corona circular $D = \{D_1, D_2, \dots, D_{18}\}$
mediante: $D_i = H_i/H_{Ai}$

```

10  Ordenar de mayor a menor las densidades  $D$ .
11  Obtener el primer  $D_1$  y segundo  $D_2$  máximos de  $D$ .
12  Calcular el valor promedio  $\bar{m}$  al realiza:  $\bar{m}=(D_1+D_2)/2$  .
13  Normalizar todas las densidades de  $D$  con  $E = \{E_1, E_2, ..., E_{18}|E_i = D_i/\bar{m}\}$ .
14  Sea  $contador = 0$  y  $r_c = 0$  .
15  para cada elemento de  $E$  hacer
16      si para cada elemento  $E_i = < 0.2$  entonces
17           $r_c \leftarrow contador$ .
18          break
19      sino
20           $contador \leftarrow contador + 1$ 
21  fin
22  fin
23  Seleccionar el número de circunferencia  $r_c$  final, en  $R$ .
24  Realizar  $I_E(m, n) = M(m, n)I(m, n)$ , donde  $M(m, n)$  tiene radio  $r_c$ .
25 fin

```

3.5. Corrección angular de la imagen de la mano

Se propuso un método que nos permita corregir el ángulo de la imagen de la mano. Se inicia con obtener los gradientes verticales y horizontales de la imagen $I^g(m, n)$ también se aplica una máscara circular binaria a la imagen y obtener $I^m(m, n)$. Lo siguiente es calcular su histograma polar con un total de 360 vectores, después se crean los vectores $\mathbf{h}_i$ para $i = 1, 2, ..., 360$, donde cada uno corresponde a la suma ponderada de una ventana que comprende a los vectores de i a $i + 90^\circ$ del histograma polar. A partir de los vectores $\mathbf{h}_i$ se busca el vector con magnitud máxima $\mathbf{h}_{max1}$. Se prosigue por dibujar una línea perpendicular y localizar un segundo vector máximo $\mathbf{h}_{max2}$ en la región opuesta al primer vector máximo. Se continúa por hallar el vector $\mathbf{h}_M$ de los vectores $\mathbf{h}_{max1}$ y $\mathbf{h}_{max2}$. En algunas ocasiones al hallar el vector máximo y corregir su ángulo no se tenía el resultado esperado. Por lo tanto se decidió crear un método que permita medir la densidad en la región de la mano y poder corregir el ángulo de manera correcta. Así se inicia por aplicar una máscara circular binaria a la imagen $I^g(m, n)$, con diámetro d_c de longitud un cuarto de dicha imagen y centrado en el centro geométrico para obtener $I^d(m, n)$ y se divide en dos mitades al dibujar una línea perpendicular al vector $\mathbf{h}_M$ el cual se ubica en el centro geométrico de $I^d(m, n)$, donde la primera mitad es $C_1(m, n)$ y la segunda $C_2(m, n)$, así mismo se selecciona un parche $P(m, n)$ de tamaño un cuarto del tamaño de la imagen $I^d(m, n)$. Para todos los píxeles (m, n) de $C_1(m, n)$, se centra el parche $P(m, n)$ en (m, n) y se calcula el valor p solo si el píxel cumple que $(m, n) > \bar{g}$ donde $\bar{g}$ e ir guardando cada valor p en el conjunto P_1 . Para $C_2(m, n)$ también se realiza el mismo proceso anteriormente mencionado para formar el segundo vector $\mathbf{r}_2$. Por otra parte se obtiene el valor promedio $\bar{r}_1$ del vector $\mathbf{r}_1$ y el valor promedio $\bar{r}_2$ del vector $\mathbf{r}_2$. Finalmente si $\bar{r}_1 > \bar{r}_2$ y además el vector $\mathbf{h}_M$ se encuentra en la mitad $C_1(m, n)$ se rota la imagen 180° y se tiene la imagen $I^{r_1}(m, n)$, en otro caso si $\mathbf{h}_M$ se encuentra en la mitad $C_2(m, n)$ se rota la imagen 180° y se tiene la imagen $I^{r_2}(m, n)$. Se muestra en el algoritmo 2 con lo anteriormente mencionado.

Algoritmo 2: Corrección angular de la imagen de la mano

Entrada: Imagen $I^g(m, n)$.
Salida: Imagen corregida en ángulo $I^s(m, n)$.

```

1 inicio
2  Aplicar una máscara circular binaria  $I^m(m, n)$  a  $I^g(m, n)$ .
3  Obtener el histograma polar de los vectores de gradiente de cada uno de los
    píxeles de  $I^m(m, n)$ .
4  Crear y obtener su histograma polar del conjunto  $H$  donde  $h_i$  corresponde a la suma
    ponderada de los vectores del histograma de paso 2, al tomar una ventana de
    vectores  $i$  a  $i + 90^\circ$ .
5  Obtener el vector con mayor magnitud  $\mathbf{h}_{max1}$  del histograma del paso 4.
6  A partir de  $\mathbf{h}_{max1}$  dibujar una línea perpendicular y localizar  $\mathbf{h}_{max2}$  en la región

```

opuesta al primero vector máximo.

7 Hallar el vector $\mathbf{h}_M$ con máxima magnitud de los vectores $\mathbf{h}_{max1}$ y $\mathbf{h}_{max2}$.

8 Aplicar una máscara circular binaria de radio d_c y dimensiones de un cuarto de sus dimensiones a $I^m(m, n)$ para obtener $I^d(m, n)$.

9 Ubicar en el centro geométrico de la imagen $I^d(m, n)$ a $\mathbf{h}_M$ para dibujar una línea perpendicular a dicho vector y dividir en dos mitades $C_1(m, n)$ y $C_2(m, n)$ la máscara circular del paso 8.

10 Obtener el valor promedio $\bar{g}$ de todos los niveles de gris de $I^d(m, n)$.

11 Sea $j = 1$ que corresponde a la primera mitad de la circunferencia del paso 9.

12 **repetir**

13 **para cada** (m, n) de $C_j(m, n)$ **hacer**

14 **si** $(m, n) \geq \bar{g}$ **entonces**

15 Tomar $P(m, n)$ de tamaño 1/4 de las dimensiones de $I^d(m, n)$ y centrada en (m, n) .

16 Hacer $s = \sum_{m,n} P(m, n)$.

17 Almacenar s en S_j .

18 **fin**

19 **fin**

20 $j = j + 1$

21 **hasta que** $j = 2$;

22 Obtener el valor promedio $\bar{p}_1$ de S_1 y $\bar{p}_2$ de S_2 .

23 **si** $\bar{p}_1 > \bar{p}_2$ **entonces**

24 **si** $\mathbf{h}_M$ está ubicado dentro de $C_1(m, n)$ **entonces**

25 Rotar 180° la imagen $I^g(m, n)$ y obtener $F(m, n)$.

26 **fin**

27 **sinó**

28 **si** $\mathbf{h}_M$ está ubicado dentro de $C_2(m, n)$ **entonces**

29 Rotar 180° la imagen $I^g(m, n)$ y obtener $F(m, n)$

30 **fin**

31 **fin**

32 **fin**

3.6. Etapa de entrenamiento

Por cada *landmark*, se realiza la construcción de 3 subespacios lineales usando el método de las *eigenfaces*. Se deben etiquetar manualmente un conjunto de imágenes con las 2 *landmarks* propuestas. Este etiquetado se realiza en las imágenes de mayor resolución. Cada subespacio o *eigenespacio* está formado por *eigenimagenes* de 32x32 píxeles para cada uno de los tres niveles de búsqueda. Las imágenes en cada nivel de búsqueda son de 64x64, 128x128 y 256x256 píxeles. Por las imágenes de entrenamiento, sabemos que la posición probable de la *landmark* está dentro de una región rectangular cuyos límites podemos determinar. El límite izquierdo corresponde a la coordenada horizontal de la *landmark* posicionada más hacia la izquierda. El límite derecho corresponde a la coordenada horizontal de la *landmark* posicionada más hacia la derecha. El límite superior corresponde a la coordenada vertical de la *landmark* posicionada más hacia arriba. El límite inferior corresponde a la coordenada vertical de la *landmark* posicionada más hacia abajo en las imágenes de ejemplo. Esta delimitación nos permite crear una región de búsqueda rectangular $R_3(m, n)$ dentro de la imagen $I^s(m, n)$ tomando los siguientes píxeles como vértices para construir la región: superior izquierdo (29, 12), superior derecho (38, 12), inferior izquierdo (29, 36), inferior derecho (38, 36), teniendo dimensiones de 24x9 píxeles, para la *landmark* superior. Para la inferior, selecciona una región de búsqueda rectangular $D_3(m, n)$ dentro de la imagen $I^s(m, n)$ tomando los siguientes píxeles como vértices para construir la región: superior izquierdo (27, 29), superior derecho (38, 29), inferior izquierdo (27, 49), inferior derecho (38, 49), teniendo dimensiones de 20x11 píxeles.

3.7. Búsqueda de la *landmark* en diferentes niveles de resolución

En esta etapa se tiene como objetivo la búsqueda de la *landmark* en diferentes escalas, ya que se cuenta con la imagen de la región de la mano en 3 diferentes tamaños. Primero se inicia con la imagen $I^s(m, n)$ de 64x64 píxeles, también se tiene definida la región de búsqueda $R_3(m, n)$ dentro de dicha imagen. Se realiza la primera localización de la *landmark* I_3 (sección 3.7.1). Lo siguiente es redimensionar I_3 en la imagen $I_2^s(m, n)$ de 128x128 píxeles, al realizar esto la *landmark* equivale a una región cuadrada de 2x2 píxeles, pero agregando 2 píxeles a cada lado de la región siendo de dimensiones 6x6 píxeles que se utilizará como región de búsqueda $R_2(m, n)$ para localizar I_2 . Repetimos la búsqueda teniendo una mejor posición de la *landmark*. A continuación I_2 se redimensiona en la imagen $I_1^s(m, n)$, por lo tanto se obtiene una región de búsqueda también de 2x2 píxeles, agregando 2 píxeles a cada lado para formar $R_1(m, n)$. Finalmente se repite la localización de la *landmark* en la imagen $I_1^s(m, n)$ y la región de búsqueda $R_1(m, n)$, localizando la *landmark* I_1 dentro de la imagen con mayores dimensiones $I_1^s(m, n)$ y ubicada en el hueso medio metacarpiano correspondiente. Se muestra de manera gráfica la búsqueda de la *landmark* en diferentes escalas en la figura 5.

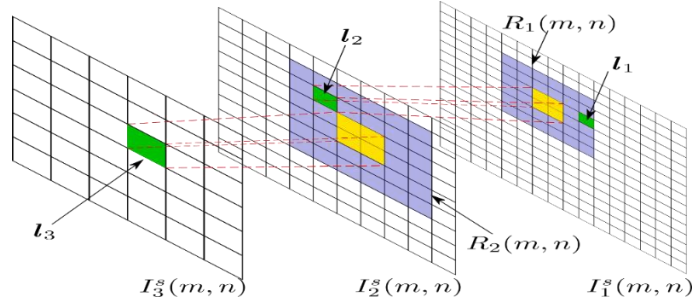


Figura 5. Al localizar la posición de la *landmark* I_3 en la imagen $I_3^s(m, n)$ se redimensiona el punto en la imagen I_2^s siendo una región de búsqueda $R_2(m, n)$, y se repite la búsqueda, para localizar el punto I_2 , el cual se redimensiona en la imagen $I_1^s(m, n)$ obteniendo una nueva región de búsqueda $R_1(m, n)$, repitiendo la búsqueda del punto en $R_1(m, n)$, para localizar la *landmark* final I_1 .

3.7.1 Proyección de parches en el subespacio de las eigenfaces La búsqueda de la *landmark* solo se realiza en $R(m, n)$. Lo primero es a partir del primer píxel de $R(m, n)$ y siendo éste el centro se toma un parche $P(m, n)$ de tamaño de 32x32 píxeles, lo siguiente es aplicar el algoritmo PCA al convertir $P(m, n)$ en un vector columna Φ y restarle el vector promedio $\bar{\Psi}$, después se proyecta en el subespacio lineal correspondiente Q de acuerdo al nivel de resolución de $I^s(m, n)$ y sumarle el vector $\bar{\Psi}$ para obtener el vector columna $\hat{\Phi}$. Posteriormente se obtiene un error de reconstrucción $e = \|\Phi - \hat{\Phi}\|$. Esto se realiza con todos los píxeles de $R(m, n)$ para cada uno se obtiene un error de reconstrucción y se guarda en un mapa de errores E , donde la posición (m, n) de dicha mapa corresponde al píxel utilizado en $R(m, n)$. Una vez finalizada la búsqueda se determina la posición mínima al calcular $l = \arg \min_{(l_m, l_n)} E(l_m, l_n)$. por lo cual l es la posición donde se ubica la *landmark* dentro de la imagen $I^s(m, n)$. Estos pasos se muestran en la figura 6.

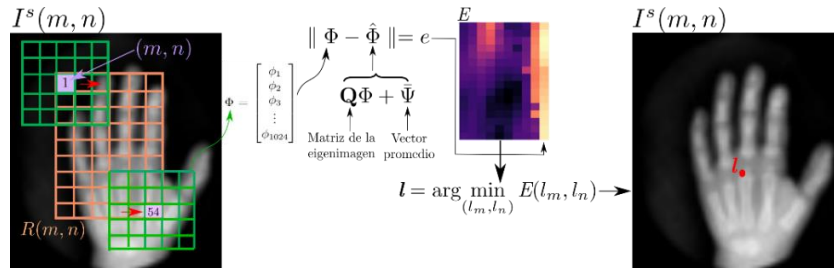


Figura 6. A partir de la región $R(m, n)$, se inicia en el píxel 1 y siendo éste el centro se toma un parche que se convierte en un vector columna Φ el cual se proyecta en Q , para obtener $\hat{\Phi}$ y calcular la norma e de los dos vectores y guardarlos en E . Una vez finalizada toda la reconstrucción de la región $R(m, n)$ se busca la posición mínima en E que corresponderá a la *landmark* l en el hueso medio metacarpiano.

3.8. Métricas de evaluación

Para determinar el error se tuvo que buscar información que nos pudiera dar la longitud en milímetros del tercer hueso metacarpiano en diferentes rangos de edad y por sexo, el cual se obtuvo de [8]. Un problema es que el rango de edad en el que se trabajó para el algoritmo fue de 1-19 años, pero la información proporcionada es de 3-19 años, así se tuvo que realizar un ajuste por mínimos cuadrados. La función obtenida para determinar la longitud del tercer hueso metacarpiano es de la forma:

$$l = a - be^{-ce_o} \quad (1)$$

donde e_o corresponde a la edad ósea. Así se tiene una función de longitud del hueso para mujeres $L3M_{mm}^f$ y de hombres $L3M_{mm}^m$ Quedando de la siguiente manera:

$$L3M_{mm}^m = 86.85 - 66.94e^{-0.06e_o} \quad (2)$$

$$L3M_{mm}^f = 66.77 - 44.15e^{-0.11e_o} \quad (3)$$

Al conocer la longitud del hueso en milímetros con (2) y (3) se buscó el número de milímetros en un píxel con su respectiva longitud en pixeles obtenida por las *landmarks* marcadas manualmente en su correspondiente sexo donde para hombres es $L3M_{px}^m$ y para mujeres $L3M_{px}^f$, quedando de la siguiente manera:

$$mm_per_px_m = L3M_{mm}^m / L3M_{px}^m \quad (4)$$

$$mm_per_px_m = L3M_{mm}^f / L3M_{px}^f \quad (5)$$

Por lo tanto para cada imagen evaluada se calculó su respectivo valor *mm per px* dependiendo del sexo y se multiplica por su error en pixeles e_{px} por lo que de esta manera obtenemos el error en milímetros e_{mm} como:

$$e_{mm} = (mm_per_px)e_{px} \quad (6)$$

Finalmente se obtiene el error promedio e^p en milímetros de acuerdo a $e^p = (1/n) \sum_{i=1}^n e_{imm}$, donde $n = 250$.

4. Resultados experimentales

En la tabla 1 se muestra el error estimado al marcar las *landmarks* automáticamente por nuestro algoritmo en conjunto con otros errores obtenidos por autores con métodos en localización de *landmarks* en imágenes radiográficas de manos. Para obtener el error se obtuvo la distancia residual que consiste en comparar las imágenes marcadas automáticamente con imágenes marcadas manualmente por un humano. Para esto se ocuparon 250 radiografías que no pertenecen al conjunto de entrenamiento.

Tabla 1. Error obtenido y los que se presentan en el estado del arte.

Autores	Error (Promedio $\pm$ Desviación estándar) mm
Doner et al. [1]	0.77 $\pm$ 0.64
Lindner et al. [2]	0.61
Payer et al. [9]	1.13 $\pm$ 0.98
Štern et al. [5]	0.80 $\pm$ 0.91
Urschler et al. [4]	0.80 $\pm$ 0.93
Payer et al. [6]	0.66 $\pm$ 0.74
Propuesto	1.15 $\pm$ 1.07

4. Conclusiones

En este artículo se ha presentado un método para la localización automática de *landmarks* en radiografías de manos, el cual ubica dos *landmarks* en el tercer hueso metacarpiano, lo cual es posible al tener las imágenes lo más parecidas posibles entre sí. Para esto se tomaron en consideración características como el contraste, la rotación, traslación y escala de cada imagen. Nuestros resultados muestran que nuestro método tiene un error de 1.15 1.07 mm similar a los del estado del arte como son: [1], [2], [3], [4], [6]. Así consideramos que nuestras contribuciones son: 1.Un método para la ubicación de la región de la mano en las radiografías, 2.Un método para corregir la posición angular de la mano, 3.Un método para la corrección de contraste. Como trabajo futuro se pretende ocupar este método para obtener regiones de interés que contengan las zonas de las epífisis en los huesos metacarpianos y los huesos carpa- les, los cuales son útiles para la estimación de la edad ósea. Donde se ha observado que un error mínimo en la localización de *landmarks* permite tener una mejor estimación de la edad ósea.

Referencias

1. Donner, R.; Menze, B.H.; Bischof, H.; Langs, G.: Fast anatomical structure localization using top-down image patch regression. In: Menze, B.H., Langs, G., Lu, L., Montillo, A., Tu, Z., Criminisi, A. (Eds.) *Medical Computer Vision. Recognition Techniques and Applications in Medical Imaging*. pp. 133–141. Springer Berlin Heidelberg, Berlin, Heidelberg (2013)
2. Lindner, C.; Bromiley, P.A.; Ionita, M.C.; Tes, T.F.: Robust and accurate shape model matching using random forest regression-voting. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 37(9), 1862–1874 (2015)
3. Štern, D.; Ebner, T.; Urschler, M.: From local to global random regression forests: Exploring anatomical landmark localization. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (Eds.) *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*. pp. 221–229. Springer International Publishing, Cham (2016)
4. Urschler, M.; Ebner, T.; Štern, D.: Integrating geometric configuration and appearance information into a unified framework for anatomical landmark localization. *Medical image analysis* 43, (January 2018), <https://doi.org/10.1016/j.media.2017.09.003>
5. Payer, C.; Štern, D.; Bischof, H.; Urschler, M.: Regressing heatmaps for multiple landmark localization using CNNs. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*. pp. 230–238 (2016)
6. Payer, C.; Štern, D.; Bischof, H.; Urschler, M.: Integrating spatial configuration into heatmap regression based CNNs for landmark localization. *Medical Image Analysis* 54, 207–219 (may 2019)
7. Turk, M.; Pentland, A.: Eigenfaces for recognition. *Journal of Cognitive Neuroscience* 3(1), 71–86 (Jan 1991)
8. Haghnegahdar, A.; Pakshir, H.; Ghanbari, I.: Correlation between skeletal age and metacarpal bones and metacarpophalangeal joints dimensions. *Journal of dentistry (Shiraz, Iran)* 20, 159–164 (09 2019)
9. Payer, C.; Štern, D.; Bischof, H.; Urschler, M.: Regressing heatmaps for multiple landmark localization using cnns. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (Eds.) *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*. pp. 230–238. Springer International Publishing, Cham (2016)

Bibliografía

- [1] A. L. Galán, “Morfometría geométrica: el estudio de la forma y su aplicación en biología,” *TEMAS DE CIENCIA Y TECNOLOGÍA*, vol. 19, no. 55, pp. 53–59, 2015.
- [2] R. Vandaele, J. Aceto, M. Muller, F. Péronnet, V. Debat, C.-W. Wang, C.-T. Huang, S. Jodogne, P. Martinive, P. Geurts, and R. Marée, “Landmark detection in 2d bioimages for geometric morphometrics: a multi-resolution tree-based approach,” *Scientific Reports*, vol. 8, p. 538, Jan 2018.
- [3] D. Zhu, J. Lv, H. Chen, and T. Liu, “Group-wise optimization of common brain landmarks with joint structural and functional regulations,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2014* (P. Golland, N. Hata, C. Barillot, J. Hornegger, and R. Howe, eds.), (Cham), pp. 716–723, Springer International Publishing, 2014.
- [4] N. Xue, M. Doellinger, C. P. Ho, R. K. Surowiec, and R. Schwarz, “Automatic detection of anatomical landmarks on the knee joint using mri data,” *J Magn Reson Imaging*, vol. 41, pp. 183–192, Jan 2015.
- [5] C. Chen, W. Xie, J. Franke, P. Grutzner, L.-P. Nolte, and G. Zheng, “Automatic x-ray landmark detection and shape segmentation via data-driven joint estimation of image displacements,” *Medical Image Analysis*, vol. 18, no. 3, pp. 487 – 499, 2014.
- [6] E. Muñoz Moreno, R. Cárdenes, R. De Luis-García, M. A. Martín-Fernández, and C. Alberola-López, “Automatic detection of landmarks for image registration applied to bone age assessment,” in *Proceedings of the 5th WSEAS International Conference on Signal Processing, Computational Geometry & Artificial Vision*, ISCGAV’05, (Stevens Point, Wisconsin, USA), p. 117–122, World Scientific and Engineering Academy and Society (WSEAS), 2005.
- [7] J. Banda, *Sistema automático de aprendizaje incremental supervisado para reconocimiento de la edad ósea*. tesis de Maestría, Facultad de Ciencias de la Electrónica, BUAP, 2018.
- [8] R. Donner, B. H. Menze, H. Bischof, and G. Langs, “Fast anatomical structure localization using top-down image patch regression,” in *Medical Computer Vision. Recognition Techniques and Applications in Medical Imaging* (B. H. Menze, G. Langs,

- L. Lu, A. Montillo, Z. Tu, and A. Criminisi, eds.), (Berlin, Heidelberg), pp. 133–141, Springer Berlin Heidelberg, 2013.
- [9] C. Lindner, P. A. Bromiley, M. C. Ionita, and T. F. Cootes, “Robust and accurate shape model matching using random forest regression-voting,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 9, pp. 1862–1874, 2015.
- [10] M. Urschler, T. Ebner, and D. Štern, “Integrating geometric configuration and appearance information into a unified framework for anatomical landmark localization,” *Medical image analysis*, vol. 43, p. 23–36, January 2018.
- [11] C. Payer, D. Štern, H. Bischof, and M. Urschler, “Regressing heatmaps for multiple landmark localization using cnns,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016* (S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal, and W. Wells, eds.), (Cham), pp. 230–238, Springer International Publishing, 2016.
- [12] C. Payer, D. Štern, H. Bischof, and M. Urschler, “Integrating spatial configuration into heatmap regression based CNNs for landmark localization,” *Medical Image Analysis*, vol. 54, pp. 207–219, may 2019.
- [13] A. Haghnegahdar, H. Pakshir, and I. Ghanbari, “Correlation between skeletal age and metacarpal bones and metacarpophalangeal joints dimensions,” *Journal of dentistry (Shiraz, Iran)*, vol. 20, pp. 159–164, 09 2019.
- [14] T. M. Deserno, *Fundamentals of Medical Image Processing*, pp. 1139–1165. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011.
- [15] F. L. Bookstein, “Morphometric tools for landmark data,” 1992. Title from publisher’s bibliographic system (viewed on 05 Oct 2015).
- [16] D. G. Kendall, “The diffusion of shape,” *Advances in applied probability*, vol. 9, no. 3, pp. 428–430, 1977.
- [17] C. P. Klingenberg and L. R. Monteiro, “Distances and Directions in Multidimensional Shape Spaces: Implications for Morphometric Applications,” *Systematic Biology*, vol. 54, pp. 678–688, 08 2005.
- [18] C. Cintas, C. Delrieux, and R. González José, “Posicionamiento automático de landmarks anatómicos en ojos,” in *XVIII Congreso Argentino de Ciencias de la Computación*, 2012.
- [19] W. Greulich and S. Pyle, “Radiographic atlas of skeletal development of the hand and wrist,” *The American Journal of the Medical Sciences*, vol. 238, no. 3, 1959.

- [20] J. Tanner, R. Whitehouse, W. Marshall, M. Healy, and H. Goldstein, “Skeletal maturity and prediction of adult height (tw2 method),” *New York, NY: Academic*, 1975.
- [21] P. Thangam, T. Mahendiran, and K. Thanushkodi, “Skeletal bone age assessment – research directions,” *Journal of Engineering Science and Technology Review*, vol. 5, pp. 90–96, 03 2012.
- [22] W. Unterpinker, *Anatomical Landmark Localization for an Automatic Multi-Factorial Age Assessment System*. M.S. thesis, Graz University of Technology, Graz, Austria, 2015.
- [23] M. Kirby and L. Sirovich, “Application of the karhunen-loeve procedure for the characterization of human faces,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 103–108, 1990.
- [24] M. Turk and A. Pentland, “Eigenfaces for recognition,” *Journal of Cognitive Neuroscience*, vol. 3, pp. 71–86, jan 1991.
- [25] I. Jolliffe, *Principal Component Analysis*. Springer Series in Statistics, Springer, 2002.
- [26] H. Ouanan, M. Ouanan, and B. Aksasse, “Facial landmark localization: Past, present and future,” in *2016 4th IEEE International Colloquium on Information Science and Technology (CiSt)*, pp. 487–493, Oct 2016.
- [27] K. Doi, “Computer-aided diagnosis in medical imaging: historical review, current status and future potential,” *Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society*, vol. 31, no. 4-5, pp. 198–211, 2007.
- [28] T. Ebner, D. Stern, R. Donner, H. Bischof, and M. Urschler, “Towards automatic bone age estimation from mri: Localization of 3d anatomical landmarks,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2014* (P. Golland, N. Hata, C. Barillot, J. Hornegger, and R. Howe, eds.), (Cham), pp. 421–428, Springer International Publishing, 2014.
- [29] J. Gall and V. Lempitsky, *Class-Specific Hough Forests for Object Detection*, pp. 143–157. London: Springer London, 2013.
- [30] D. Štern, T. Ebner, and M. Urschler, “From local to global random regression forests: Exploring anatomical landmark localization,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016* (S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal, and W. Wells, eds.), (Cham), pp. 221–229, Springer International Publishing, 2016.

- [31] L. Peter, O. Pauly, P. Chatelain, D. Mateus, and N. Navab, “Scale-adaptive forest training via an efficient feature sampling scheme,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (N. Navab, J. Hornegger, W. M. Wells, and A. Frangi, eds.), (Cham), pp. 637–644, Springer International Publishing, 2015.
- [32] G. Febles, “Inteligencia artificial en imagenología. revisión de conceptos, aplicaciones y consecuencias.” Accedido el 28 de mayo del 2020 de <http://sriuy.org.uy/ojs/index.php/Rdi/article/download/53/70?inline=1>.
- [33] T. Grzejszczak, M. Kawulok, and A. Galuszka, “Hand landmarks detection and localization in color images,” *Multimedia Tools and Applications*, vol. 75, no. 23, pp. 16363–16387, 2016.
- [34] S. Mahmoodi, B. S. Sharif, E. G. Chester, J. P. Owen, and R. E. J. Lee, “Automated vision system for skeletal age assessment using knowledge based techniques,” in *Proc. Sixth Int. Conf. Image Processing and Its Applications*, vol. 2, pp. 809–813 vol.2, July 1997.
- [35] M. Niemeijer, B. van Ginneken, C. A. Maas, F. J. A. Beek, and M. A. Viergever, “Assessing the skeletal age from a hand radiograph: automating the tanner-whitehouse method,” in *Medical Imaging 2003: Image Processing, San Diego, California, United States, 15-20 February 2003* (M. Sonka and J. M. Fitzpatrick, eds.), vol. 5032 of *SPIE Proceedings*, SPIE, 2003.
- [36] M. Á. Martín-Fernández, M. Martín-Fernández, and C. Alberola-López, “Automatic bone age assessment: a registration approach,” in *Medical Imaging 2003: Image Processing, San Diego, California, United States, 15-20 February 2003* (M. Sonka and J. M. Fitzpatrick, eds.), vol. 5032 of *SPIE Proceedings*, SPIE, 2003.
- [37] K. Mader, “Rsna bone age predict age from x-rays,” 05 2017. Accedido el 28 de mayo del 2019 de <https://www.kaggle.com/kmader/rsna-bone-age>.
- [38] J. Arlow and I. Neustadt, *UML 2.0 and the Unified Process: Practical Object-Oriented Analysis and Design (2nd Edition)*. Addison-Wesley Professional, 2005.