

## Localización de referencias cefalométricas en radiografías laterales de cráneo: Una revisión sistemática

José Luis López-Ramírez, José Ruiz-Pinales,  
Juan Gabriel Aviña-Cervantes, Pedro López-Rodríguez,  
José Guadalupe Beltrán-Hernández, Mario Alberto Ibarra-Manzano

Universidad de Guanajuato, División de Ingenierías  
Campus Irapuato Salamanca, Departamento de Ingeniería Electrónica,  
México

{jl.lopezramirez,pinales}@ugto.mx

**Resumen.** Los sistemas de diagnóstico asistidos por computadora aplicados en la medicina juegan un papel fundamental en la detección de patologías (en etapa temprana y/o progreso), cuyo diagnóstico requiere la toma de muestras de "imagenología diagnóstica". Actualmente existe un número considerable de algoritmos computacionales que han obtenido resultados destacados en la estimación (automática y semiautomática) de referencias cefalométricas aplicada en radiografías laterales de cráneo. Siendo esta, una tarea que requiere de un lapso considerable de tiempo por parte de los médicos especialistas, dado que la identificación de los diferentes puntos de referencia requiere de un conocimiento amplio en el tema a la par de agudeza visual para identificar los diferentes niveles de resolución y clasificación dentro de los cefalogramas. En el área de diagnóstico clínico odontológico la planificación maxilofacial es utilizada comúnmente. Sin embargo, solo existe una base de datos pública presentada en el Simposio Internacional sobre Imágenes Biomédicas (SIIB 2014-2015). Este proyecto consiste en una revisión sistemática de los diferentes trabajos presentados hasta el momento. Adicionalmente, se presenta una comparación (cuantitativa y cualitativa) de los resultados obtenidos por cada uno de los autores en las diferentes técnicas propuestas para dar solución a esta tarea.

**Palabras clave:** Aprendizaje profundo, aprendizaje máquina, redes neuronales convolucionales (RNC), revisión bibliográfica, análisis cefalométrico.

### Cephalometric References' Localization in Lateral Skull Radiographs: A Systematic Review

**Abstract.** Computer-Aided Diagnostic (CADx) systems applied in medicine play a fundamental role in the detection of pathologies (at an early stage

or progress), which diagnosis requires taking samples of medical images. Currently, there is a considerable number of computational algorithms that have obtained outstanding results in the automatic and semi-automatic cephalometric references localization systems, applied to lateral skull radiographs. Otherwise, this is a time-consuming task for the expert physicians, since it requires knowledge and significant lapse of time for the identification of the reference points and a visual acuity to identify the different levels of resolution and classification within the cephalograms. In the orthodontics field, the maxillofacial treatment plan is used commonly by dentists. Nevertheless, there is only an available public database which was presented at the International Symposium on Biomedical Images (ISBI) 2015. This project consists of a systematic review of the different works presented to date. Additionally, the comparison (quantitative and qualitative) of the results achieved by the authors in the different techniques proposed to solve this task.

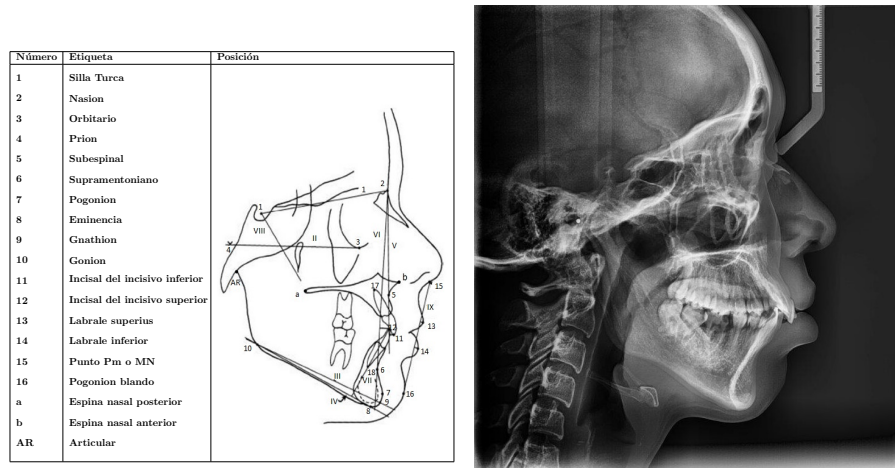
**Keywords:** Deep learning, machine learning, convolutional neural networks, literature review, cephalometric analysis.

## 1. Introducción

Los sistemas computacionales han tenido un desarrollo significativo en las últimas décadas. En especial, han mostrado gran influencia en las ciencias médicas. En este campo se desarrollan sistemas de diagnóstico asistidos por computadora cuya finalidad tiene como objetivo la predicción y/o detección temprana de patologías. Estos sistemas ayudan a los médicos a emitir un prediagnóstico con alto grado de exactitud en estudios que requieren “imagenología diagnóstica”. En el campo médico de la odontología existen estudios como la cefalometría. Esta última se encarga, de obtener mediciones del cráneo humano a partir de imágenes tomadas con rayos-X, bajo un procedimiento estricto de captura.

En la Figura 1a, se muestra un diagrama esquemático con los puntos cefalométricos propuestos a detectar. Adicionalmente la Figura 1b, proporciona un ejemplo de una radiografía lateral de cráneo, utilizada en el “Análisis cefalométrico”. Para la identificación de referencias cefalométricas se debe contar con máquinas de rayos-X que cuya salida sea una imagen digital o impresa de las condiciones médicas del sistema óseo en la vista lateral del cráneo. También se requiere del tiempo, conocimiento y dedicación de médicos especialistas en el área para la identificación de los puntos cefalométricos. En la Figura 1a, se observan los nombres y ubicación de 19 referencias cefalométricas importantes para el diagnóstico clínico del crecimiento de cráneo, planificación de cirugías maxilofaciales, entre otros.

Actualmente existe una base de datos pública presentada en el SIIB 2014-2015. Esta base de datos ha sido utilizada desde entonces por diferentes autores, donde presentan sus algoritmos computacionales para resolver la identificación automática y semiautomática de las referencias cefalométricas. Se han obtenido resultados sumamente importantes con una alta precisión y exactitud.



(a) Posiciones de los 19 puntos que conforman el análisis cefalométrico [13]. (b) Radiografía lateral de cráneo.

Fig. 1. Imágenes de rayos-X para estudio de Análisis cefalométrico.

En este trabajo se presenta una revisión sistemática de los trabajos que han sido publicados hasta el momento. Es posible obtener una comparación cuantitativa de dichos estudios, debido a que han sido entrenados y evaluados previamente con una misma base de datos, ya que de no ser así, solo es posible comparar las diferentes metodologías.

La evaluación del rendimiento computacional se realiza con una fracción de la base de datos utilizada. Esta última es elegida de manera aleatoria y destinada únicamente para las pruebas de rendimiento en todos los algoritmos propuestos. Las métricas de evaluación empleadas son: el Error Radial Medio (ERM) y la Tasa de Detección (TD) de éxito.

Este trabajo se organiza de a siguiente manera. La Sección **Materiales y Metodología**, donde se presentan las características de la única base de datos pública que utilizaron la mayoría de los autores de los trabajos presentados en esta revisión. Por otro lado se presenta una breve descripción de las diferentes técnicas utilizadas por cada uno de los autores. La Sección **Resultados**, presenta un análisis comparativo de los resultados cuantitativos y cualitativos, basados en las métricas de evaluación y obtenidos por los diferentes autores. En la Sección **Conclusiones** se presentan las deducciones de esta revisión sistemática.

## 2. Materiales y metodologías

### 2.1. Base de datos pública (SIIB 2014-2015)

La competición convocada por el Simposio Internacional sobre Imágenes Biomédicas 2015 respaldada por el IEEE (por sus siglas en inglés, “Institute of

Electrical and Electronics Engineers”), proporciona un conjunto de imágenes públicas de las competiciones para la detección de puntos de referencia cefalométricos en sus ediciones SIIB 2014 [18] y SIIB 2015 [19]. Esta base de datos se encuentra en el sitio web de la competición SIIB 2014-2015.

La base de datos está compuesta por 400 radiografías de pacientes de entre 7 y 76 años (edad promedio: 27 años; 235 mujeres, 165 hombres). Todos los cefalogramas se adquirieron en formato TIFF con una máquina “Soredex CRANEX Excel CEPH” (Tuusula, Finlandia), utilizando el software “Soredex SorCom 3.1.5, versión 2.0”. La resolución de cada imagen es de  $1935 \times 2400$  píxeles con una resolución espacial de  $0,1 \times 0,1 \text{ mm}^2$  por píxel. Un ejemplo de imagen de esta base de datos se ilustra en la Figura 2a.

El etiquetaje manual fue realizado por dos ortodoncistas clínicos expertos (con 15 años y seis años de experiencia, respectivamente) del departamento dental del hospital general de Taiwan, el etiquetaje del médico con más experiencia es llamado “Experto 1” y por consecuente el médico con 6 años es reconocido como “Experto 2”. Los 19 puntos de referencia marcados fueron elegidos por involucrar estructuras comunes utilizadas en evaluaciones cefalométricas como el análisis cefalométrico de Steiner [17] y el análisis Wits Appraisal [7]. La base de datos fue dividida de manera aleatoria en 150 imágenes de entrenamiento y dos conjuntos de prueba: el primer conjunto también llamado “Prueba 1” está conformado por 150 imágenes con la que se evaluaron los resultados del SIIB 2014-2015 [18] y el segundo conjunto de prueba llamado “Prueba 2” consta de 100 imágenes y de igual forma fueron las imágenes de evaluación del SIIB 2014-2015 [19].

## 2.2. Métricas de evaluación

Son considerados dos criterios fundamentales para la evaluación del desempeño de los algoritmos computacionales, presentados en la Sección 2.3. Ambos criterios fueron propuestos durante las competiciones SIIB 2014 [18] y SIIB 2015 [19].

**Error Radial Medio (ERM).** El primer criterio de evaluación es el Error Radial Medio (ERM) en conjunto con la Desviación Estándar (DE) asociada. El primer criterio de evaluación es el error radial medio (ERM) en conjunto con la Desviación Estándar (DE) asociada. El primer parámetro a considerar es la distancia euclidiana entre cada píxel propuesto por el algoritmo y el píxel objetivo (referencia etiquetada por el médico experto) directamente sobre el cefalograma. Ésta distancia es definida matemáticamente por  $R$  en la Ecuación (1):

$$R = \sqrt{(\Delta X)^2 + (\Delta Y)^2}, \quad (1)$$

donde,  $\Delta X$  y  $\Delta Y$  son las distancias en sus respectivos ejes (X y Y). Siendo  $\Delta X = x_p - x_o$  y  $\Delta Y = y_p - y_o$  (el subíndice “p” es de propuesto y el subíndice “o” de objetivo). Teniendo esto en cuenta, el ERM o distancia promedio es formulada como se muestra en la Ecuación (2):

$$\text{ERM} = \frac{1}{N} \sum_{i=1}^N R_i \quad [\text{mm}], \quad (2)$$

$$R_i = \frac{1}{M} \sum_{j=1}^M R_{ij} \quad [\text{mm}], \quad (3)$$

donde, la constante  $N$  es el número de cefalogramas del conjunto que se está evaluando y  $M$  es el número de referencias cefalométricas a localizar.

La Desviación Estándar (DE), representada en la Ecuación (4). La DE es una medida de la dispersión de datos (referencias cefalométricas), cuanto mayor sea la dispersión mayor es la DE, si no hubiese ninguna variación en los datos, es decir, si fueran todos iguales, la DE sería cero. El cálculo de la desviación estándar se realiza de la siguiente manera:

$$\text{DE} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (R_i - \text{ERM})^2}. \quad (4)$$

**Tasa de Detección (TD)** Para el cálculo de la tasa de detección, se considera exitosa la detección cuando un punto cefalométrico propuesto por un algoritmo que se encuentra dentro de cierto rango médico aceptable de distancias, tales como (2, 2.5, 3 y 4) mm. De esta forma es como se establece la tasa de detección (TD) dada por la Ecuación (5):

$$p_z = \frac{|\{i \in \mathbb{N} : R_i < z, 1, 0 \leq i \leq N\}|}{N} \times 100 \%, \quad (5)$$

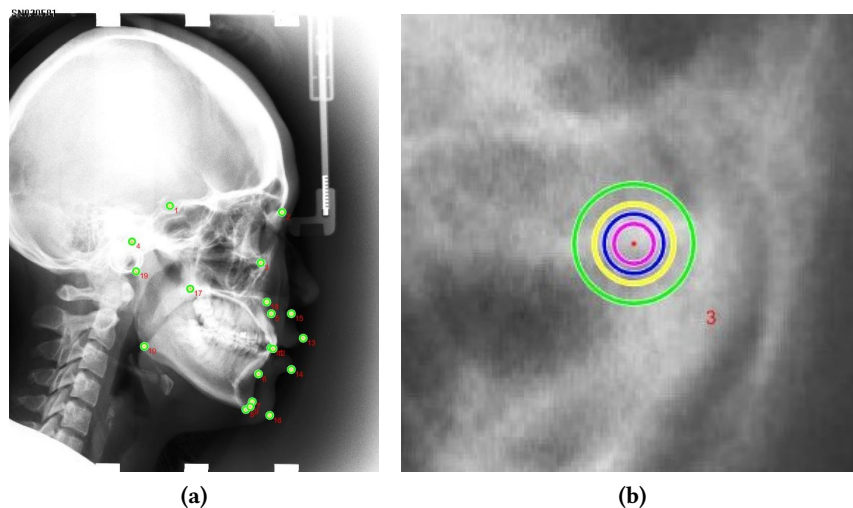
donde  $|\cdot|$  simboliza la cardinalidad,  $R_i$  representa la distancia radial de la etiqueta propuesta. La literal  $z$  expresa el rango de precisión usado en la evaluación (Rangos: {2.0, 2.5, 3.0, 4.0} mm).

La Figura 2a presenta una radiografía lateral de cráneo, perteneciente al conjunto de datos SIIB 2014-2015, con su respectivo etiquetado manual. Por otro lado la Figura 2b, representa un recorte de la etiqueta llamada "Orbitario". Siendo el punto rojo (centro), el píxel etiquetado por el experto y los anillos concéntricos (magenta, azul, amarillo y verde), los rangos de detección médicamente aceptables {2.0, 2.5, 3.0, 4.0} mm, respectivamente. Estos rangos de detección determinan la capacidad de precisión y exactitud de los sistemas computacionales para resolver este desafío.

En la siguiente Sección presenta una breve descripción de los métodos y técnicas propuestas en la literatura para dar solución a esta importante tarea.

### 2.3. Revisión literaria

Ibragimov *et al.* [5,6] presentan un marco de trabajo basado en "Teoría de juegos" (GTF) y "Random Forests" para detectar las etiquetas cefalométricas. Los puntos, detectados por GTF usando un modelo de apariencia multi-escala, son refinados por medio de "Random Forest" usando un modelo de forma. Primero es aplicado el modelo de apariencia multi-escala a la imagen-objetivo escalada y



**Fig. 2.** a) Imagen cefalométrica de la base de datos. b) Rangos de detección (rojo: Verdad de Referencia (“*Ground-Truth*”), magenta: 2 mm, azul: 2.5 mm, amarillo: 3 mm y verde: 4 mm).

cuando se requiere una posición más precisa se aplica en la imagen original sin escalar. De esta manera, se considera información global al analizar la imagen escalada, e información local al analizar la imagen original, lo que conduce a una reducción de complejidad computacional. El resultado obtenido muestra que al menos el 71.7% de las etiquetas detectadas se encuentran en el rango de 2mm.

Kaur *et al.* [8] utilizan la detección de características globales basadas en los momentos de Zernike. Este análisis, consiste en la clasificación de objetos obteniendo información sobre los diferentes rasgos geométricos descritos por los polinomios de Zernike debido a su alta capacidad de representación y la pequeña sensibilidad al ruido presentado en el fondo de las imágenes de rayos X. Cabe resaltar que la base de datos utilizada es una base de datos privada, conformada por un total de 135 imágenes y para su validación se utilizó el algoritmo *drop-one-out*, utilizando 134 imágenes para entrenamiento, y una imagen fue excluida para la prueba.

Claudia Lindner and Tim F. Cootes [10]. Proponen solucionar el etiquetaje automático de radiografías laterales de cráneo usando “*Random Forest-Regression Voting*”. Esta técnica cuenta con una alta capacidad de generalización para la solución de muchos problemas. Consiste en mejorar la generalización y aprender los patrones que se presentan en los datos de entrenamiento, compensando los errores de las predicciones de los distintos árboles de decisión mediante la votación por regresión, obteniendo resultados de precisión en el rango de los 2mm.

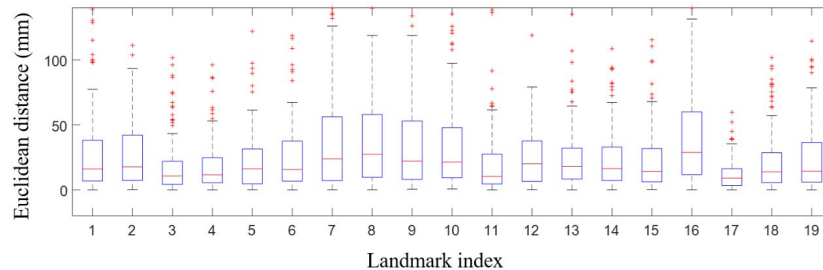
Claudia Lindner *et al.* [19]. Proponen un sistema llamado FALA (por sus siglas en inglés: “*Fully Automatic Landmark Annotation System*”). FALA es implementado

usando su técnica de “*Random Forest-Regression Voting*”, con la cual detectan la posición escala y orientación del cráneo, para posteriormente aplicar una técnica CML (por sus siglas en inglés: “*Constrained Local Model*”) que detecta las posiciones de los puntos de referencia. El primer paso hace que el sistema sea robusto a cualquier variación en la adquisición de imágenes, y el segundo paso permite la colocación precisa de los 19 puntos de referencia. La combinación de estas dos técnicas muestra el mejor desempeño para resolver esta tarea obteniendo los mejores resultados en la literatura existente. Este sistema logró un ERM de 1.2mm, y el 84.7% de los puntos de referencia se ubicaron dentro del rango de precisión clínicamente aceptado de 2.0 mm.

Sercan Ö. Arık *et al.* [1]. Propone por primera vez una arquitectura de redes neuronales “*feed-forward*” profundas para el análisis de cefalometría cuantitativa que proporciona la estimación de 19 pares de datos cefalométricos localizados en radiografías laterales de cráneo, utilizando la base de datos públicas propuesta en la competición anteriormente descrita [19]. Los resultados obtenidos muestran que al menos un 76% de los puntos propuestos de la partición de prueba son etiquetados dentro del rango de 2mm. En este caso, se utilizaron recortes (parches) de imágenes como entrada a la red neuronal.

Hansang Lee *et al.* [9] presentan un sistema basado en aprendizaje profundo, proponiendo una arquitectura de RNC para estimar por medio de regresión los 19 puntos de referencia. Se utilizó la base de datos pública SIIB 2014-2015. La metodología es descrita como el entrenamiento de 38 sistemas de regresión con la misma arquitectura RNC en variables de 19 pares coordenados  $(x, y)$ , respectivamente. Finalmente, se calculan 38 variables de coordenadas como salida de estos sistemas entrenados a partir de imágenes invisibles. Obteniendo resultados plasmados en la gráfica de caja mostrada en la Figura 3. En el eje horizontal, se enumeran las 19 referencias cefalométricas obtenidas por el algoritmo propuesto. En el eje vertical, se muestra la distancia euclidiana que hay entre los valores que propone el algoritmo computacional y el punto etiquetado por el médico experto. Estos resultados muestran un rendimiento prometedor, sin embargo, no se mencionan resultados cuantitativos, por lo que visualmente no se puede estimar un resultado que pueda ser comparado con los resultados establecidos por las otras metodologías propuestas de esa revisión. Ji-Hoon Park *et al.* [4] comparan el rendimiento computacional de dos algoritmos basados en aprendizaje profundo. La base datos es compuesta por 1311 imágenes. Los sistemas fueron entrenados con un total de 1028 imágenes que se consideraron únicamente para el entrenamiento de las arquitecturas. Para las pruebas de rendimiento se utilizaron las 283 imágenes restantes. Cada imagen es compuesta por 80 etiquetas cefalométricas. Los métodos propuestos son la arquitectura “*You-Only-Look-Once*” versión 3 (YOLOv3) propuesta por Redmon *et al.* [15] y el método “*Single Shot Multibox Detector*” (SSD) propuesto por Wei Liu *et al.* [11]. El algoritmo YOLOv3 superó a SSD en precisión para 38 de 80 puntos de referencia. Los otros 42 de 80 puntos de referencia no mostraron una diferencia estadísticamente significativa entre YOLOv3 y SSD.

Runnan Chen *et al.* [2]. utilizan como metodología la combinación de un módulo basado en una RNC simple, que es utilizada para extraer las características de la imagen de entrada y posteriormente se utiliza una técnica llamada “*Attentive*



**Fig. 3.** Diagrama de caja de distancias euclidianas entre puntos de referencia predichos y los puntos objetivo, propuestos por Hansang Lee *et al.* [9].

*Feature Pyramid Fusion* (AFPF). Esta última, permite resaltar de forma explícita las características de alta resolución. De manera conjunta, se combinaron los “*heatmaps*” (mapas de atención) y mapas de desplazamiento, estas etapas permiten que el rendimiento del sistema mejore. Y así mismo sea más eficaz la regresión por votación de los píxeles. Obteniendo resultados de etiquetado de manera automática en tiempo real, directamente sobre imágenes del dispositivo que realiza la radiografía.

Xiubin Dai *et al.* [3] proponen una arquitectura basada en aprendizaje profundo utilizando redes generativas adversariales. Esta arquitectura está basada en el manejo de mapas de distancia de cada uno de los puntos cartométricos a encontrar. Teniendo la red generadora entrenada se obtienen mapas de distancia que proporcionan información de los 19 puntos cefalométricos. Obteniendo los 19 mapas con mejor aproximación al conjunto de imágenes objetivo son los que por otro lado la red discriminadora entrega como resultado.

Jiahong Qian *et al.* [14] diseñan un método rápido basado en R-CNN “*Region Based Convolutional Neural Network*”, por sus siglas en inglés). Los autores han nombrado a su método como “CephaNET”. Esta arquitectura consiste en la construcción de un grafo 2D (no dirigido), donde se plasma la distribución de puntos de referencia en los datos de entrenamiento. Para eliminar los puntos anormales detectados, se adopta la “máxima confianza” y la transformación Laplaciana. Este resultado es comparando con otras metodologías propuestas como se puede observar en la Figura 4.

Hengameh Mirzaalian *et al.* [12] desarrollaron este algoritmo computacional que extrae un conjunto de características que incluyen: características locales basadas en patrones binarios, coordenadas espaciales (X, Y) de los puntos cefalométricos, *blobness*, *tubularness* y características de Zernike. Se emplearon estas características para calcular la verosimilitud de cada uno de los puntos de referencia usando un clasificador “*Random Forest*”. En la Figura 5, se presentan los resultados obtenidos en esta investigación.

Yu Song *et al.* [16] proponen un método semiautomático para la detección de puntos de referencia cefalométricos mediante aprendizaje profundo. Primero se extrae manualmente una región ROI (por sus siglas en inglés: “*Region Of*



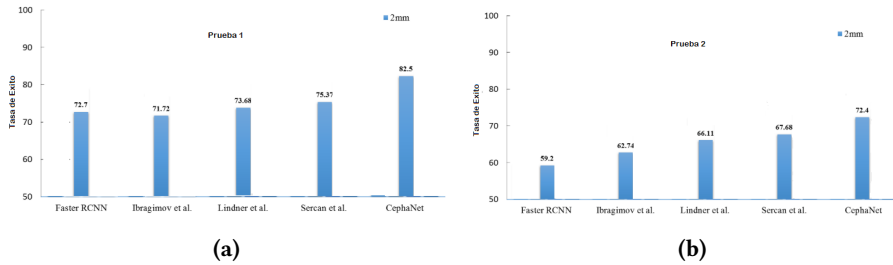


Fig. 4. Resultados de CephNET Vs Otros-autores, en el rango clínico de 2mm, para los dos subconjuntos de imágenes de prueba [14].

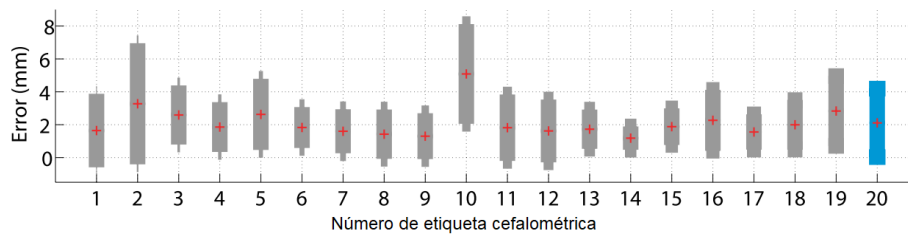


Fig. 5. Resultados obtenidos por el algoritmo computacional propuesto por Hengameh Mirzaalian *et al.* [12].

*Interest*”) para cada punto de referencia de manera manual. Posteriormente, se utiliza una arquitectura Resnet50 para obtener las coordenadas de la referencia cefalométrica que se encuentra dentro de la ROI. Obteniendo resultados que superan cuantitativamente los resultados del estado del arte.

Los diferentes trabajos presentados anteriormente concuerdan en que la cefalometría cuantitativa y precisa toma un papel importante para los diagnósticos o tratamientos odontológicos. Sin embargo, es un trabajo que conlleva mucho tiempo para etiquetar manualmente los puntos de referencia cefalométricos. Dicha implementación requiere de médicos profesionales expertos. En conclusión, el desarrollo de un sistema de diagnóstico que ofrezca una precisión y exactitud comparable a la de un médico experto en el área, agilizaría el diagnóstico para posteriormente dar el tratamiento adecuado y si es el necesaria la planificación de la cirugía.

### 3. Resultados

A continuación, se presenta la Tabla 1, donde se muestran los resultados obtenidos por los diferentes autores de acuerdo a la métrica ERM.

Por lo tanto no se calculó el ERM para este subconjunto de datos. De igual forma los autores han tomado como regla, sólo entregar sus resultados con el

primer subconjunto de datos llamado “Prueba 1”, por lo tanto no es posible tener un registro de sus resultados para esta métrica y en la Tabla 1, está marcada con SR (Sin Resultado). En el caso específico de los trabajos de Sercan *et al.* y CephaNET. Los autores no presentaron los resultados globales del ERM, sólo se limitaron a presentar la Tasa de Detección (TD). Como previamente se ha mencionado en la Sección 2.3.

**Tabla 1.** Tabla de resultados del ERM, para autores que utilizaron la base de datos propuesta por el SIIB 2014-2015.

Autor	ERM(Prueba 1)	ERM(Prueba 2)
Ibragimov <i>et al.</i> (2014)	1.919	SR
Ibragimov <i>et al.</i> (2015)	1.851	1.82
Chu <i>et al.</i>	2.679	SR
Chen y Zhang	2.847	SR
Mirzaalian y Harmarneh	2.353	SR
Vandaele <i>et al.</i>	2.198	SR
Lidner and Cootes	1.67	SR
Sercan <i>et al.</i>	SR	SR
CephaNET	SR	SR
Y. Song <i>et al.</i>	1.14746	1.22347
Runnan Chen <i>et al.</i>	1.17	1.48

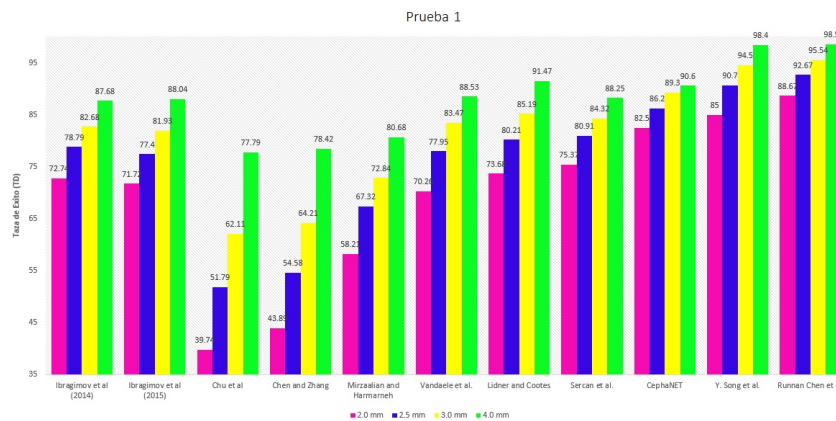
Como se puede observar en la Tabla 1, el mejor resultado es presentado por Y. Song *et al.* [20], seguido de Runnan Chen *et al.* [2]. Por un lado, Song presenta un sistema semiautomático que requiere que el médico seleccione el de manera manual, recortes de ventanas de 128 x 128 píxeles que contengan un punto cefalométrico, por lo que el punto es encontrado con mayor precisión y exactitud. Por otro lado, Chen *et al.* presentan una arquitectura donde, de manera indirecta generan un mapa 2D que contiene la distribución de los puntos cefalométricos de las imágenes etiquetadas.

De igual manera esto representa una ayuda indirecta al algoritmo computacional y posteriormente aplican transformaciones Laplacianas para encontrar el punto objetivo con mayor certidumbre.

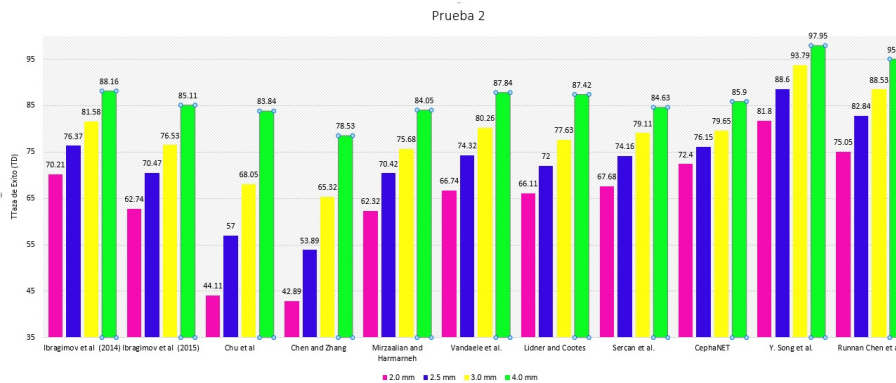
Además, se muestra que el mejor resultado para un sistema completamente automático es el presentado por Claudia Lindner [10]. Este sistema utiliza la técnica computacional *Random Forest-Regression Voting*. Un árbol de decisión representa un modelo de clasificación o regresión aplicado al procesamiento de imágenes en una estructura de árbol. Cada nodo en el árbol representa una característica del espacio de entrada, cada decisión una rama y cada hoja al final de una rama es el valor de salida correspondiente. Esto es realizado múltiples ocasiones para cada una de las 19 etiquetas a identificar. Posterior a que se generalizan y se aprenden patrones, se compensan los errores de predicción con una votación, haciendo un

algoritmo robusto, con un rendimiento alto para superar los resultados de los otros sistemas automáticos presentados anteriormente.

En la Figura 6 se presentan dos gráficas de barras donde se muestra el porcentaje de la tasa de detección para los cuatro rangos médicamente aceptables (2 mm, 2.5 mm, 3 mm, 4 mm). Para los 11 trabajos presentados que utilizaron la



(a) Resultado Experimental 1



(b) Resultado Experimental 2

Fig. 6. Comparación de resultados entre autores para la métrica TD.

base de datos de la competición del SIIB 2014-2015. La Figura 6b representan las TDs, para el subconjunto de imágenes llamado “Prueba 1”, que consta de un total de 150 imágenes. La Figura 6a muestra las TDs para el subconjunto de imágenes “Prueba 2”, de igual manera es un subconjunto de 150 imágenes de prueba.

Por otro lado, existen trabajos que utilizaron su propia base de datos y presentan metodologías importantes que dan solución a la localización de etiquetas

cefalométricas.

Kaur *et al.* obtuvieron una localización del 89.5% en la tasa de detección, dentro del rango de 2 mm. Siendo este un resultado prometedor y poder ser cuantitativamente comparables los resultados con el obtenido por Claudia Lindner *et al.*, debe ser probado este algoritmo con la base de datos del SIIB 2014-2015.

Por otro lado Ji-Hoon Park *et al.* comparan dos algoritmos computacionales, obteniendo como resultado que entre el *Single Shot Multibox Detector* y la arquitectura YOLOv3 el que obtiene el mejor desempeño es esta última, obteniendo tasas de detección de, 96.2%, 92%, 87.4% y 80.4% en los rangos de {4, 3, 2.5, 2} mm respectivamente. En su manuscrito comparan los resultados contra los obtenidos por Sercan Ö. Arik *et al.* por ser arquitecturas que utilizan el mismo fundamento científico, pero al no utilizar la misma base de datos los resultados no pueden ser sistemas competitivos cuantitativamente.

La siguiente Sección se presenta las conclusiones obtenidas respecto al diagnóstico clínico de la identificación de etiquetas cefalométricas y así mismo los aspectos que se deben cuidar al intentar resolver esta tarea.

#### 4. Conclusiones

La identificación automática de referencias cefalométricas es un procedimiento médico rutinario para la planificación de cirugías maxilofaciales en el área de la ortodoncia. Este proceso ocupa tiempo y conocimiento de un humano experto, lo cual hace que el realizar un algoritmo computacional para un sistema de diagnóstico asistido por computadora, que tenga características de precisión y exactitud dentro de los rangos médicos aceptables sea aún una tarea pendiente. Los algoritmos presentados en este trabajo reflejan que primero se tiene que implementar un mecanismo automático para la detección de las regiones de interés, donde estas regiones contengan las etiquetas cefalométricas y eliminar las variaciones que afectan a muchos algoritmos. Una de las técnicas que en los últimos años han mostrado un alto rendimiento computacional, son los modelos de aprendizaje profundo. Estos algoritmos están inspirados biológicamente en el comportamiento visual humano y pueden resultar poderosos. Además, los resultados presentados muestran que ya se tiene un TD cercano al 99% dentro del rango de 4mm, sin embargo, aún no se ha dado por concluida esta tarea.

#### Referencias

1. Arik, S.Ö., Ibragimov, B., Xing, L.: Fully automated quantitative cephalometry using convolutional neural networks. *Journal of Medical Imaging* 4(1), 014501 (2017)
2. Chen, R., Ma, Y., Chen, N., Lee, D., Wang, W.: Cephalometric landmark detection by attentive feature pyramid fusion and regression-voting. In: Shen, D., Liu, T., Peters, T.M., Staib, L.H., Essert, C., Zhou, S., Yap, P.T., Khan, A. (eds.) *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*. pp. 873–881. Springer International Publishing, Cham (2019)

3. Dai, X., Zhao, H., Liu, T., Cao, D., Xie, L.: Locating Anatomical Landmarks on 2D Lateral Cephalograms through Adversarial Encoder-Decoder Networks. *IEEE Access* 7, 132738–132747 (2019)
4. Hwang, H.W., Park, J.H., Moon, J.H., Yu, Y., Kim, H., Her, S.B., Srinivasan, G., Aljanabi, M.N.A., Donatelli, R.E., Lee, S.J.: Automated identification of cephalometric landmarks: Part 2-Might it be better than human? *Angle Orthodontist* 90(1), 69–76 (2020)
5. Ibragimov, B.: Automatic cephalometric x-ray landmark detection by applying game theory and random forests. In: *Grand Challenges in Dental X-ray Image Analysis* (2014)
6. Ibragimov, B., Likar, B., Pernus, F., Vrtovec, T.: Computerized cephalometry by game theory with shape-and appearance-based landmark refinement. In: *Grand Challenges in Dental X-ray Image Analysis* (2015)
7. Jacobson A: The “Wits” appraisal of jaw disharmony. *American Journal of Orthodontics* 67(2), 125–138 (1975)
8. Kaur, A., Singh, C.: Automatic cephalometric landmark detection using Zernike moments and template matching. *Signal, Image and Video Processing* 9(1), 117–132 (2015)
9. Lee, H., Park, M., Kim, J.: Cephalometric landmark detection in dental x-ray images using convolutional neural networks. *Medical Imaging 2017: Computer-Aided Diagnosis* 10134(March 2017), 101341W (2017)
10. Lindner, C., Cootes, T.F.: Fully automatic cephalometric evaluation using Random Forest regression-voting. *Isbi* pp. 1–8 (2015)
11. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y., Berg, A.C.: SSD: Single shot multibox detector. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 9905 LNCS, 21–37 (2016)
12. Mirzaalian, H., Hamarneh, G.: Automatic globally-optimal pictorial structures with random decision forest based likelihoods for cephalometric x-ray landmark detection. In: *Automatic Cephalometric X-Ray Landmark Detection Challenge 2014, in conjunction with IEEE International Symposium on Biomedical Imaging (IEEE ISBI)*. pp. 1–12 (2014), [http://www.cs.sfu.ca/~hamarneh/ecopy/isbi2014\\_chlg\\_ceph.pdf](http://www.cs.sfu.ca/~hamarneh/ecopy/isbi2014_chlg_ceph.pdf)
13. Ortoface: ANÁLISIS DEL TRAZADO: REFERENCIAS CEFALOMÉTRICAS. url <http://ortoface.com/wp-content/uploads/2016/12/An%CC%81alisis-del-trazado.pdf> (2016)
14. Qian, J., Cheng, M., Tao, Y., Lin, J., Lin, H.: Cephanet: An improved faster r-cnn for cephalometric landmark detection. In: *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. pp. 868–871 (2019)
15. Redmon, J., Farhadi, A.: YOLO v.3. Tech report pp. 1–6 (2018), <https://pjreddie.com/media/files/papers/YOLOv3.pdf>
16. Song, Y., Qiao, X., Iwamoto, Y., Chen, Y.W.: Semi-automatic cephalometric landmark detection on x-ray images using deep learning method. In: Liu, Y., Wang, L., Zhao, L., Yu, Z. (eds.) *Advances in Natural Computation, Fuzzy Systems and Knowledge Discovery*. pp. 585–592. Springer International Publishing, Cham (2020)
17. Steiner, C.C.: Cephalometrics for you and me. *American Journal of Orthodontics* 39(10), 729–755 (1953)
18. Wang, C.W., Huang, C.T., Hsieh, M.C., Li, C.H., Chang, S.W., Li, W.C., Vandaele, R., Marée, R., Jodogne, S., Geurts, P., Chen, C., Zheng, G., Chu, C., Mirzaalian, H., Hamarneh, G., Vrtovec, T., Ibragimov, B.: Evaluation and Comparison of Anatomical Landmark Detection Methods for Cephalometric X-Ray Images: A Grand Challenge. *IEEE Transactions on Medical Imaging* 34(9), 1890–1900 (2015)

*José Luis López-Ramírez, José Ruiz-Pinales, Juan Gabriel Aviña-Cervantes, et al.*

19. Wang, C.W., Huang, C.T., Lee, J.H., Li, C.H., Chang, S.W., Siao, M.J., Lai, T.M., Ibragimov, B., Vrtovec, T., Ronneberger, O., Fischer, P., Cootes, T.F., Lindner, C.: A benchmark for comparison of dental radiography analysis algorithms. *Medical Image Analysis* 31, 63–76 (jul 2016), <https://linkinghub.elsevier.com/retrieve/pii/S1361841516000190>
20. Yu, H.J., Cho, S.R., Kim, M.J., Kim, W.H., Kim, J.W., Choi, J.: Automated Skeletal Classification with Lateral Cephalometry Based on Artificial Intelligence. *Journal of Dental Research* (2020)