

REVISIÓN SISTEMÁTICA EXPLORATORIA: DIAGNÓSTICO ASISTIDO POR COMPUTADORA CON USO DE INTELIGENCIA ARTIFICIAL EN IMAGENOLOGÍA MAXILOFACIAL



Meléndez Rojas, Patricio; Araneda Fuentes, Cristoffer; Tapia Vilches, Ignacio.
Universidad Nacional Andrés Bello, Facultad de Odontología
Viña del Mar, Chile, 2020

Resumen:

Objetivo: Explorar el uso y viabilidad de la Inteligencia Artificial en el campo de la Imagenología Maxilofacial (u odontología) ante la incorporación de la era digital a la práctica clínica. Demostrar si la literatura disponible la califica como un método realmente efectivo. **Materiales y métodos:** Revisión Sistemática Exploratoria. Estudios en inglés publicados en Bases de datos EBSCO, PubMed, y World of Science, con un algoritmo de búsqueda para identificar artículos que relacionen el uso de Inteligencia e Imagenología Maxilofacial de los últimos 10 años, en detección de patologías relacionadas al área maxilofacial (caries y quistes). La búsqueda fue realizada en julio de 2020 y los trabajos fueron seleccionados valorando su capacidad de responder a la pregunta de investigación planteada. **Resultados:** Se espera encontrar un avance en detección de lesiones del territorio maxilofacial con sistemas de Diagnostico Asistido por Computadora basadas en algún tipo de Inteligencia Artificial. **Conclusiones:** Los resultados obtenidos en el presente estudio han permitido revisar que la evidencia disponible sobre diagnostico asistido por computadora en sistemas basados en inteligencia artificial presentan sensibilidad y especificidad similar a la de clínicos experimentados en algunos casos, sin embargo, tampoco se puede depender solo de estas aplicaciones inteligentes ya que se detectó que existen factores que afectan este desempeño.

Palabras clave:

Inteligencia Artificial, Machine learning, Deep-learning, Computer Aided Diagnosis (CAD), Big Data

Abstract:

Objective: To explore the use and viability of Artificial Intelligence in the field of Maxillofacial Imaging (or dentistry) before the incorporation of digital age into clinical. Demonstrate if the available literature qualifies it as an effective method. **Materials and methods:** Scoping Review. Studies in English published in EBSCO, PubMed, and World of Science databases, with a search algorithm to identify articles that relate the use of Intelligence and Maxillofacial Imaging of the last 10 years, in the detection of pathologies related to the maxillofacial area (caries and cysts). The search was carried out in July 2020 and the papers were selected evaluating their ability to answer the research question posed. **Results:** It is expected to find an advance in the detection of lesions of the maxillofacial territory with Computer Aided Diagnosis systems based on some type of Artificial Intelligence. **Conclusions:** The results obtained in the present study have allowed us to review that the available evidence on computer-assisted diagnosis in systems based on artificial intelligence presents sensitivity and specificity like experienced clinicians in some cases, however, it is not possible to depend only on these intelligent applications since it was found that there are methods that affect this performance.

Keywords:

Artificial Intelligence, Machine learning, Deep-learning, Computer Aided Diagnosis (CAD), Big Data.

INTRODUCCIÓN

En los últimos años, la era digital ha significado una reestructuración de la vida moderna. Nuevas formas de comunicación e información han revolucionado a la sociedad desde sus cimientos. Bajo este ámbito, la implementación y uso de inteligencia artificial (IA) ha adquirido un rol importante en la vida diaria como también en lo laboral.

Ante la posibilidad de que ciertos sistemas computacionales logran “aprender” a clasificar y detectar patrones en imágenes mediante la tecnología de Aprendizaje Automático (“*Machine Learning*” en inglés o “ML” por sus iniciales), ha logrado generar profundos cambios en áreas relacionadas a la imagenología médica en los últimos años. Uno de los algoritmos más usados para lograr esta tarea corresponde a las Redes Neuronales Artificiales (“RNA” por sus siglas), cuya arquitectura utiliza un gran número de elementos interconectados llamados neuronas (también denominados nodos o células) para simular en cierta manera, el fenómeno de transmisión y procesado de información en una sinapsis cerebral.(1)

Una RNA está constituida por el sistema de procesamiento computacional caracterizándose por la presencia de una o más capas intermedias ocultas (“intermediate hidden layers” en inglés). Al igual que en la naturaleza, cada capa intermedia oculta de neuronas está conectada con cada una de las neuronas en la capa anterior a modo que exista “sinapsis” entre ellas (2)

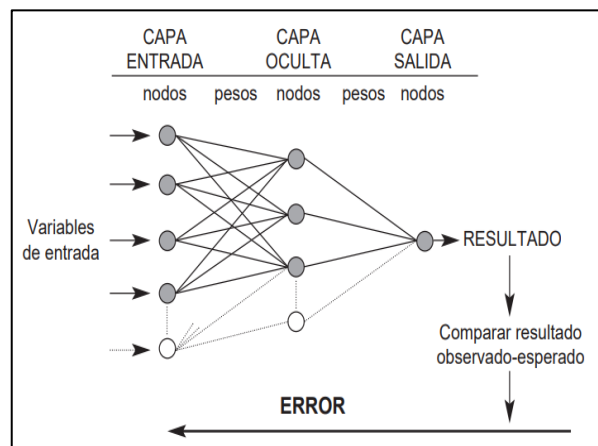


Imagen 1 Esquema básico de Red Neuronal Artificial (ARN) “Arquitectura de una red neuronal artificial (perceptrón de tres capas). Los pesos son los parámetros que deben ajustarse durante el proceso de entrenamiento para conseguir que la capa de salida de los resultados coincida con los observados. El error entre lo observado y el resultado de la red se propaga hacia atrás (backpropagation) y debe ir disminuyendo en las sucesivas iteraciones en las que se presentan los valores de las variables predictoras. La complejidad de una red depende del número de nodos de su capa oculta. (3)

Mediante los nodos que conforman la capa de entrada (*input*), diversas variables en forma de valores ingresaran a la red artificial, para luego ser sometidos a procesos de ajuste en las capas ocultas

para finalmente arrojar un resultado (*output*). (4).

Más recientemente, una subrama de “ML”, el Aprendizaje Profundo (“*Deep Learning*” en inglés o “DL” por sus iniciales), ha logrado entrenar modelos de mayor complejidad mediante el empleo de modelos con una gran cantidad de capas intermedias, otorgando mejor interpretación de las características a detectar o diagnosticar en donde las técnicas clásicas de ML fallan. Con el surgimiento de macro datos (“*Big data*” en inglés), DL logra beneficiarse de esta masiva cantidad de información debido a que mientras más datos maneja, mejor se comporta ante análisis complejos (2)(5).

Dentro de DL hasta la fecha, se debe destacar el tipo de arquitectura basada en Redes Neuronales Convolucionales (“*Deep Neural Networks*” en inglés o “DNN’s” por sus siglas). En este caso, como no todas sus neuronas están conectadas en todos los niveles como en una arquitectura clásica de RNA, estas “modelan la información en partes más pequeñas, para combinar después esta información en las capas más profundas de la red y dividir el problema en partes

para conseguir predicciones más sencillas y precisas” (6).

En Odontología, la creación de equipos y sistemas computacionales cada vez más poderosos ha permitido incursionar en el uso de inteligencia artificial mediante arquitecturas DNN’s, especializadas para búsqueda de alteraciones o detección de patrones en imagenología maxilofacial o patología, optimizando los diagnósticos realizados, agilizando los procedimientos y ofreciendo al clínico cierta guía o una segunda opinión. En ciertos casos, estos sistemas pueden llegar a entregar información de vital importancia que a su vez no fue pesquisada por el usuario, logrando convertirse así en máquinas de Soporte de Decisión Diagnostica (o CAD por su sigla en inglés. (7)

El propósito de este artículo es revisar el contexto de la inteligencia artificial y sus conceptos asociados para las entender las aplicaciones en medicina y perspectivas futuras que tendrá en la odontología.

MATERIALES Y MÉTODOS

Para este *Scoping Review* se planteó la siguiente pregunta de investigación: ¿Cuál es la evidencia disponible sobre el diagnóstico imagenológico elemental (diagnóstico de caries, restauraciones, fracturas radiculares, nivel óseo, lesiones periapicales, y quistes) usando sistemas informáticos basados en inteligencia artificial?

Búsqueda

La búsqueda incluyó estudios publicados entre el año 2010 y el 31 de Julio de 2020. La búsqueda fue realizada en las bases de datos PubMed, Medline (vía EBSCO), Scopus y Web of Science. Se utilizaron los siguientes términos MeSH: *“Artificial Intelligence”*, *“Diagnosis, Computer-Assisted”* *“Machine Learning”*, *“Deep Learning”*, *“Dental Radiology”*, con las respectivas siglas para inteligencia artificial: “AI”, “ML”, “DL” y “CNN”. Para la obtención de resultados se siguió la siguiente estrategia: *“Artificial Intelligence”* OR *“Diagnosis, Computer-Assisted”* OR *“Machine Learning”* OR *“Deep Learning”* OR “AI” OR “ML” OR “DL” OR “CNN” AND *“Dental Radiograph**”*.

Selección de Estudio

Publicaciones académicas y/o artículos científicos que estén disponibles en idioma inglés, con una antigüedad máxima de 10 años, indexados en alguna de las siguientes bases de datos: “PubMed”, “Scopus”, “MEDLINE” (vía EBSCO) y “Web of Science”. Los artículos seleccionados deben contar con alguna herramienta que involucre algún modelo de inteligencia artificial, con elementos del diagnóstico imagenológico general, (para esta revisión se considerarán identificación dentaria, diagnóstico de caries, patología periapical, fracturas dentarias y/o radiculares, quistes, tumores, nivel óseo). Se excluyeron: artículos de opinión o editoriales de revistas científicas afines al tema, publicaciones que utilicen inteligencia artificial para cualquier otra aplicación en odontología (implantología, diagnóstico de osteoporosis, etc.), publicaciones a las que no se pudo acceder al texto completo, o aquellas que no se encontraron disponibles en idioma inglés.

Procesamiento y selección de la información

Los trabajos seleccionados, luego de una lectura crítica, serán organizados en una

tabla de datos que incluirá los siguientes apartados: año y país de publicación, autor (es), tipo de imagen, campo de diagnóstico, tipo de inteligencia artificial, número de imágenes procesadas, presencia de revisores de las imágenes y principales resultados de cada investigación.

La selección fue realizada por dos revisores de manera independiente y luego fueron discutidos las selecciones no coincidentes, siguiendo el flujograma PRISMA.

RESULTADOS

La búsqueda identificó 1391 registros, distribuidos de la siguiente manera. 690 fueron obtenidos desde “PubMed”, 369 desde “Scopus”, 169 desde “Medline” (vía EBSCO) y 153 desde “Web of Science”. Para mayor referencia revisar el flujograma prisma (Imagen 2). Un listado de los trabajos incluidos y las principales características de cada uno puede ser encontrado en la Tabla 1. A su vez los trabajos excluidos en la fase de revisión de lectura completa y los motivos de exclusión pueden ser encontrados en la Tabla 2

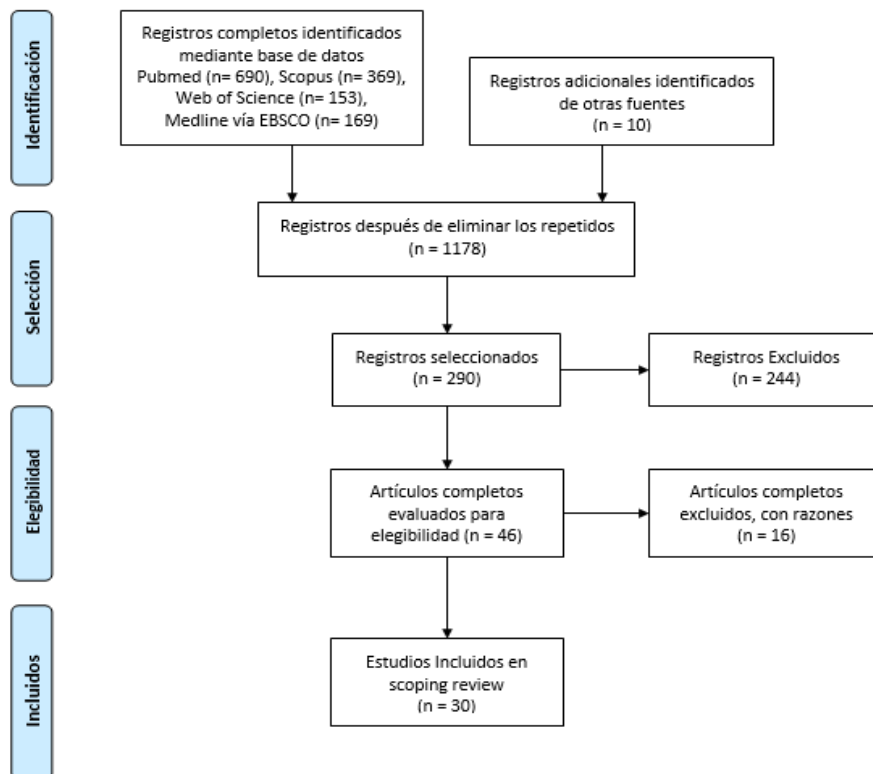


Imagen 2 Flujograma adaptado de PRISMA

Tabla 1 Estudios Incluidos y sus principales hallazgos

Estudio	País y Año	Campo Diagnóstico	Imagen utilizada	Imágenes Totales	Revisores	Tipo de Inteligencia artificial	Principal Resultado
Chang HJ, et al.	Corea del Sur, 2020	Nivel óseo radiográfico	Radiografías Panorámicas	3 set con 330,115 y 73 imágenes en cada set	3 radiólogos Dentales	<i>Hybrid Framework Deep Learning</i>	El coeficiente de correlación de Pearson del método automático con los diagnósticos de los radiólogos fue de 0,73 en general para toda la mandíbula ($p < 0,01$), y el valor de correlación de intracase 0,91 global para toda la mandíbula ($p < 0,01$).
Khan y Mukati	Pakistán, 2020	Identificación dentaria	Radiografías panorámicas	116 radiografías (2910 imágenes) para entrenamiento. 80000 imágenes para transfer learning	Los autores	<i>Convolutional neural network (AlexNet)</i>	Efectividad de la clasificación dentaria inicial de 88.31%. Efectividad de la clasificación dentaria después de Transfer Learning de 98.88%
Geetha V, et al.	India, 2020	Diagnóstico de Caries	Radiografías retroalveolares	105 radiografías intraorales	Radiólogo dental (identificó caries en 49 imágenes)	<i>Back-propagation neural network</i>	Efectividad del diagnóstico de caries de 97.1%, tasa de falso positivo de 2.8%
Lee JH, et al.	Corea del Sur, 2019	Clasificación de quistes	CBCT Radiografías Panorámicas	2126 imágenes (1140 panorámicas y 986 CBCT)	Comparación con estudios histológicos hechos por histopatología	<i>Deep Convolutional neural network (GoogLeNet Inception V3)</i>	El diagnóstico de quistes presenta sensibilidad de 96.1%, y especificidad 77.1%
Ekert T, et al.	Alemania, 2019	Diagnóstico de lesiones periapicales	Radiografías panorámicas	2001 imágenes a partir de 85 radiografías panorámicas	6 dentistas con experiencia variable entre 3 a 10 años	<i>Deep Convolutional neural network</i>	El diagnóstico de quistes presenta sensibilidad de 0.65, y especificidad 0.87
Krois J, et al.	Alemania, 2019	Nivel óseo radiográfico	Radiografías panorámicas	2001 imágenes a partir de 85 radiografías panorámicas	3 especialistas independientes	<i>Deep Convolutional neural network</i>	La precisión de clasificación de la CNN fue de 0,81 (0,02). La sensibilidad y la especificidad medias fueron 0,81 (0,04), 0,81 (0,05), respectivamente.
Lee JH, et al.	Corea del Sur, 2019	Diagnóstico de caries	Radiografías retroalveolares	3000 radiografías (2400 para entrenamiento, 600 test)	4 examinadores	<i>Deep Convolutional neural network (GoogLeNet Inception V3)</i>	Las precisiones diagnósticas de los modelos premolar, molar, y premolar y molar en conjunto fueron del 89,0%. (80,4–93,3), 88,0% (79,2–93,1) y 82,0% (75,5–87,1), respectivamente
Abdalla-Aslam, et al.	Israel, 2020	Diagnóstico de restauraciones	Radiografía Panorámica	83 radiografías panorámicas (738 restauraciones)	Experto en medicina oral con conocimientos de imagenología dental	<i>Machine Learning, Cube Vector Support Machine,</i>	El algoritmo detectó un 94.6% de restauraciones. La clasificación eliminó los erróneos y finalmente quedó un 90.5% de las restauraciones. La eficacia promedio de clasificar y discriminar de verdaderas restauraciones fue de 93.6%.
Endres MG, et al.	EE. UU., 2020	Clasificación de quistes dentarios	Radiografías panorámicas	3240 imágenes en total	102 imágenes por cada cirujano maxilofacial (experiencia variable entre 5 a 20 años)	<i>Supervised Deep Learning</i>	El modelo de Deep Learning tiene mejor desempeño en 14 de los 24 cirujanos maxilofaciales con los que se comparó el modelo
Osmanovic A, et al.	EE. UU., 2020	Diagnóstico de lesiones periapicales	No especifica	112 imágenes	Diagnosticado mediante histopatología	<i>Machine Learning (Random Forest)</i>	Random Forest algorithm puede incrementar el diagnostico verdadero a 85.71%.

García Cantú A, et al.	Alemania, 2020	Diagnóstico de caries	Radiografías retroalveolares	3686 radiografías bitewing divididas en 3 grupos (3293 entrenamiento, 252 validación y 141 prueba)	4 dentistas especialistas en diagnóstico odontológico preventiva y operatoria, cariología	<i>Convolutional neural network (U-Net)</i>	Presenta un 0.80 de efectividad diagnóstica versus 0.71 de los dentistas que evaluaron. Sensibilidad de 0.75 versus 0.36 Especificidad de 0.83 versus 0.91
Thanathornwong y Suebnukarn	Tailandia, 2020	Nivel óseo radiográfico	Radiografías panorámicas	100 radiografías panorámicas de pacientes comprometidos periodontalmente	Evaluación de periodoncistas	<i>R-Convolutional neural network</i>	Promedio de precisión de 0.81, sensibilidad de 0.84 y especificidad de 0.88
Orhan K, et al.	Turquía, 2020	Diagnóstico de lesiones periapicales	CBCT	153 imágenes de lesiones periapicales	2 especialistas en imagenología	<i>Deep Convolutional neural network (U-Net)</i>	El AI fue capaz de detectar 142 de 153 lesiones periapicales con una fiabilidad de 92.8%
Ariji Y, et al.	Japón, 2019	Clasificación de quistes dentarios	Radiografías panorámicas	210 radiografías	1 especialistas en imagenología	<i>Deep neural network (DetectNet)</i>	La sensibilidad fue de 0,88 para las pruebas 1 y 2. La tasa de falsos positivos por imagen fue de 0,00 para la prueba 1 y de 0,04 para la prueba 2
Singh y Sehgalb	India, 2020	Identificación dentaria	Radiografías panorámicas	400 radiografías (240 entrenamiento, 160 de prueba)	No presenta	<i>Deep Convolutional neural network</i>	Se ha logrado una precisión del 95% para la base de datos de prueba y del 92% para el conjunto de datos original con el algoritmo propuesto
Lee JH, et al.	Corea del Sur, 2018	Nivel óseo radiográfico	Radiografías retroalveolares	3 set de imágenes: 1,044 para entrenamiento, 348 para validación y 348 para prueba.	3 periodoncistas experimentados y calibrados entre sí.	<i>Deep Convolutional neural network (VGG-19 network)</i>	La precisión diagnóstica de la PCT fue del 81,0% para los premolares y del 76,7% para los molares. Utilizando 64 premolares y 64 molares que fueron diagnosticados clínicamente como PCT grave, la precisión de predicción de la extracción fue del 82,8% (IC del 95%, 70,1% -91,2%) para los premolares y del 73,4% (IC del 95%, 59,9% -84,0%) para molares.
Fukuda M	Japón, 2019	Diagnóstico de fractura radicular	Radiografía Panorámica	300 radiografías, con 330 dientes fracturados en ellas	2 radiólogos y 1 endodoncista	<i>Deep neural network (DetectNet)</i>	El diagnostico presenta una precisión de 0.93 y sensibilidad de 0.75
Son LH	Vietnam, 2017	Identificación dentaria. Diagnóstico de caries, fracturas radiculares, y nivel óseo radiográfico.	Radiografías retroalveolares y panorámicas	87 imágenes de lesiones orales	Evaluación por radiólogos dentales	<i>Semi-Supervised fuzzy clustering - fuzzy satisficing (SSFC-FS)</i>	La precisión diagnostica de las cinco enfermedades dentales evaluadas fue de 92,74%
Nurtanio I, et al	Rumania, 2013	Clasificación de quistes dentarios	Radiografías Panorámicas digital	133 radiografías panorámicas (53 imágenes de quistes y 80 imágenes de lesiones tumorales)	1 radiólogo	<i>Support Machine Vector</i>	Los resultados muestran que la diferenciar quistes de lesiones tumorales pueden alcanzar una precisión de hasta 87.18% y AUC hasta 0.9444
Wang L, et al	China, 2016	Identificación dentaria y de estructuras dentales	Imágenes Micro-CT	280 imágenes	Dentistas con experiencia clínica que manualmente etiquetaron diferentes estructuras dentarias	<i>Pulse Coupled Neural Network</i>	Segmentación y clasificación individual de dientes mediante PCNM. Este resultado obtiene mejor desempeño para identificar estructuras comparado con FCM, HCA, DBSCAN y GMMs

Yilmaz E, et al.	Turquía, 2017	Clasificación de quistes dentarios	CBCT	50 archivos de imagen CBCT	No presenta	<i>K-nearest neighbors (k-NN), Naive Bayes, Decision trees, Random forest, Neural Network (NN) Support Vector Machine (SVM)</i>	De todos los modelos, SVM obtuvo el mejor desempeño con un 94% de precisión. Resultado obtenido mediante ANOVA.
Miki Y, et al	Japón, 2017	Identificación dentaria	CBCT	52 archivos de imágenes CBCT obtenido por 2 sistemas (Veraviewepocs3D y Alphard VEGA), de las cuales 5 imágenes de cada sistema fueron elegidas al azar como información de prueba y los 42 restantes como información de entrenamiento.	No presenta	<i>Convolutional neural network (Fully Convolutional Network – AlexNet)</i>	El porcentaje de detección promedio corresponde a 77.4% precisión con 5.8 falsos positivos por imagen.
Wang C, et al	China, 2016	Diagnóstico de caries	Telerradiografía lateral de cráneo, Radiografías Bite-wing	400 cefalometrías 120 radiografías bitewing obtenidas de 120 pacientes	2 experimentados doctores	Al ser un desafío, 9 equipos se registraron, pero solo 2 equipos lo implementaron correctamente. (1) <i>Ronneberger et al: "U-shaped" deep convolutional neural network</i> (2) <i>Lee et al: Random forest machine learning system</i>	El promedio del valor-F de los equipos (Ronneberger y et al, y Lee et al.) son de 0.560 y 0.268, respectivamente, y la red profunda convolucional "u-shaped" por Ronneberger et al. se desempeña mejor para las 3 estructuras dentales (esmalte, dentina y pulpa)
Johari M, et al	Irán, 2017	Detección de fractura vertical radicular	Radiografías retroalveolares y CBCT	240 radiografías de dientes (120 sin fractura vertical y 120 con fractura vertical). Separadas en 3 grupos para entrenamiento y prueba de la red neuronal.	No presenta	<i>Probabilistic Neural Network (Feed Forward Neural Network) con Multilayer Perceptron Network (MLP)</i>	En las radiografías periapicales, los valores máximos de precisión, sensibilidad y especificidad en los 3 grupos fueron 70.00%, 97.78% y 67.7%, respectivamente. Estos valores en las imágenes CBCT fueron 96.6%, 93.3% y 100%, respectivamente, con un cambio en el rango de varianza de 0.1-0.65.
Yang H, et al	Corea del sur, 2020	Detección de quistes y tumores	Radiografías panorámicas	1603 radiografías panorámicas. Imágenes fueron clasificadas en 4 categorías.	Se realizó un análisis comparativo entre 3 grupos (YOLO, cirujanos orales maxilofaciales y practicantes generales).	<i>Deep Convolutional Neural Network YOLO (You Only Look Once)</i>	YOLO clasifica la mayor puntuación entre los 3 grupos (precisión = 0,707, exhaustividad =0,680). Las diferencias de desempeño entre la maquina y los clínicos fue estadísticamente insignificante. Los resultados indican la utilidad de la autodetección de redes convolucionales.

Mahmoud, et al	Egipto, 2016	Diagnóstico de lesiones periapicales.	Radiografías retroalveolares periapicales digitales	201 radiografías periapicales	No presenta	<i>Feed Forward Neural Network y K-nearest Neighbor</i>	De acuerdo con los resultados, K- nearest Neighbor es mejor que la red neural Feed Forward
Tracy KD, et al	EE. UU, 2011	Diagnóstico de caries y de restauraciones ("Logitech Caries Detector")	Radiografía Bitewing digital e imágenes clínicas	17 radiografías más fotografías clínicas para documentar.	12 dentistas (evaluadores) y 1 dentista que restauró 28 superficies proximales	<i>Feed Forward Neural Network</i>	La sensibilidad entre todos los dentistas fue del 30% con la imagen; 35% con el brillo y contraste ajustado; 39% cuando la imagen se hizo más nítida; y 69% cuando la herramienta de análisis de densidad fue usada. La especificidad corresponde a un 97% en la imagen inicial, 95% con el brillo y contraste ajustado a la imagen; 93% con la imagen más nítida; y 94% cuando la herramienta de análisis de densidad fue utilizada.
Kositbowornchai S, et al	Tailandia, 2013	Detección de fractura vertical radicular	Radiografías retroalveolares periapicales digitales	200 imágenes (50 sanos y 150 con fractura vertical radicular) para entrenar y probar la red neuronal artificial, dividido en 3 grupos de acuerdo con el número de dataset de entrenamiento y prueba: 80/120, 105/95, 130/70 respectivamente.	No presenta	<i>Probabilistic Neural Network</i>	La mayor sensibilidad (98%), especificidad (90,5%) y precisión (95.7%) ocurrió en el grupo 3.
Behere RR, et al	India, 2011	Diagnóstico de caries dental ("Logitech Caries Detector")	Radiografía Digital estandarizada e imágenes digitales.	100 superficies proximales (de 50 dientes extraídos) de los que se obtuvieron radiografías digitales. también se obtuvieron imágenes digitales de las superficies	3 observadores independientes	<i>Feed Forward Neural Network</i>	La sensibilidad, especificidad, valor predictivo positivo, valor predictivo negativo y precisión general de LCD para los distintos grados son: -Sin caries: 33,96,73,82 y 81% respectivamente -Caries de esmalte: 5,97,33,80 y 79% respectivamente -Caries dentinaria: 100,96,50, 100 y 96% respectivamente Bajo estos resultados, LCD parece más confiable en descartar caries que en detectarlas
Kwon O, et al	Corea del sur, 2020	Diagnóstico de quistes y tumores	Radiografías Panorámicas digitales	1282 radiografías panorámicas. Estas incluyen 350 imágenes con quiste dentígero, 302 quistes periapicales, 300 queratoquistes odontogénicos, 230 ameloblastomas y 100 radiografías panorámicas normales como grupo control	Los quistes y tumores fueron clasificados basados en diagnostico histopatológico por biopsia y diagnóstico clínico. 2 radiólogos con más de 15 años de experiencia diagnosticaron las radiografías panorámicas seleccionadas.	<i>Convolutional Neural Network modificada</i>	El rendimiento general de la clasificación para las enfermedades mejoró desde un 78,2% de sensibilidad, 93,9% de especificidad, 91,3% de precisión y 0,86 AUC utilizando la CNN con datos no aumentados establecidos en 88,9% de sensibilidad, 97,2% de especificidad, 95,6% de precisión y 0,94 AUC utilizando la CNN con conjunto de datos aumentado. CNN que usaba un conjunto de datos aumentados tenía las siguientes sensibilidades, especificidades, precisiones y AUC: 91,4%, 99,2%, 97,8% y 0,96 para países en desarrollo, 82,8%, 99,2%, 96,2% y 0,92 para PC, 98,4%, 92,3%, 94,0% y 0,97 para OKC, 71,7%, 100%, 94,3% y 0,86 para AB, y 100,0%, 95,1%, 96,0% y 0,97 para mandíbulas normales, respectivamente.

Tabla 2 Estudios excluidos y motivo de exclusión

Estudio	País y Año	Motivo de exclusión
Birbal R, et al.	Turquía, 2015	No especifica qué tipo de sistema de Inteligencia Artificial utiliza para lograr el resultado
Kim, et al	Corea del Sur, 2019	Sólo disponible en idioma original (coreano)
Marín I, et al	Hungría, 2017	No hay diagnóstico asistido por Inteligencia artificial, sólo hay desarrollo de algoritmo.
Raith S, et al	Alemania, 2017	Uso de Inteligencia Artificial para detectar maloclusión dentaria basado en modelos.
Son L.H, et al	2016, Vietnam	Emplea métodos de segmentación. No hay diagnóstico asistido por Inteligencia artificial.
Jung SK, et al	Corea del sur, 2016	Uso de Inteligencia Artificial para evaluar exodoncia por indicación ortodóncico.
Bouchahma M, et al	Emiratos Árabes, 2019	Inteligencia artificial utilizada para la planificación del tratamiento de caries
Vaccaro G, et al	España, 2014	Aplica detección y diagnóstico mediante inteligencia artificial aplicado en fotografías clínicas. No utiliza radiografías digitales en el estudio
Veena DK, et al	India 2017	Técnica mediante ajuste de contraste y negativo de imagen para mejorar la detección de patologías. No realiza diagnóstico asistido por inteligencia artificial
Lin PL, et al	China, 2017	Emplea detección mediante diferentes algoritmos y escala de grises para detectar nivel óseo. No realiza diagnóstico asistido por inteligencia artificial
Zhang ZL, et al	China, 2017	Sólo disponible en idioma original (chino)
Sakamoto J, et al	Japón, 2016	Evaluación de técnica basada en difusión curtosis para la detección de quistes y tumores en imágenes de resonancia magnética. No realiza diagnóstico asistido por inteligencia artificial
Elsatani MH, et al	Egipto, 2016	Comparación de distintos sistemas CBCT para la detección de fracturas verticales radiculares. No realiza diagnóstico asistido por inteligencia artificial
Nascimento H, et al	Brasil, 2015	Detección de fracturas verticales radiculares mediante la aplicación de distintos filtros de enfoque y suavizado de imagen. No realiza diagnóstico asistido por inteligencia artificial
Siotia J, et al	Alemania, 2011	Reporte de caso del uso de "DentaScan". No realiza diagnóstico asistido por inteligencia artificial
Taylor AM, et al	Inglaterra, 2010	Utiliza algoritmo de "detección de bordes" en imágenes micro-CT. No realiza diagnóstico asistido por inteligencia artificial

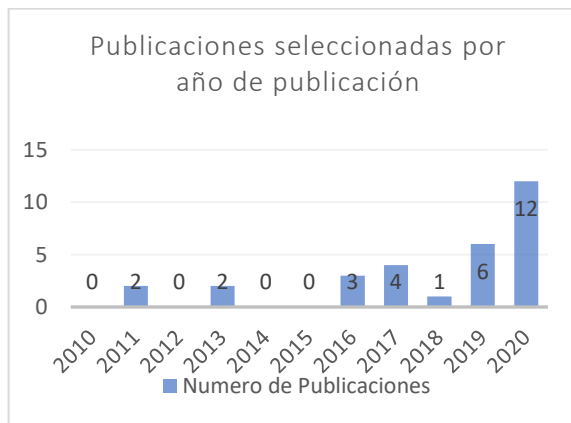


Imagen 3 Distribución de trabajos seleccionados por año de Publicación

La temática de los estudios abarcó las siguientes disciplinas diagnósticas: diagnóstico de caries, identificación dentaria, fracturas radiculares, lesiones periapicales, diagnóstico y clasificación

de quistes de los maxilares y evaluación de la pérdida de hueso alveolar. La distribución los trabajos por disciplina diagnóstica se resume en Imagen 16.

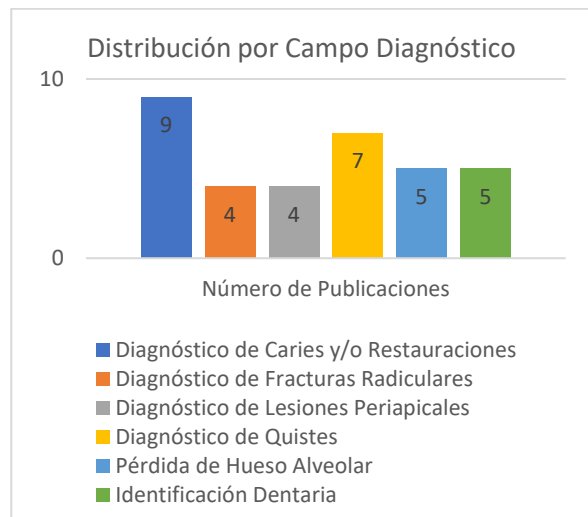


Imagen 4 Distribución por Campo Diagnóstico

Cabe destacar que dentro de los trabajos seleccionados hubieron 18 que utilizaron menos de 500 imágenes, 2 que utilizaron entre 500 y 1.000 imágenes, y 10 que usaron más de 1.000 imágenes para lograr su finalidad diagnóstica.

Dentro del tipo de imagen utilizada para el proceso diagnóstico, las que más se utilizaron corresponden a radiografías panorámicas, seguido de radiografías retroalveolares seguido por radiografías periapicales, en más de un trabajo se utiliza más de un tipo de radiografía para investigar. Un desglose por tipo de imagen se encuentra en la Imagen 17.

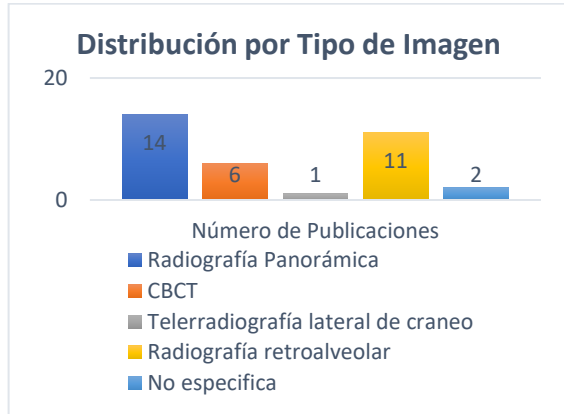


Imagen 5 Distribución por tipo de imagen

Respecto a los revisores, en 6 trabajos no se menciona presencia de revisores que realicen una revisión previa o un medio de comparación para el banco de datos que presenta la publicación. Principalmente los revisores que se mencionan son

personas con conocimientos y experiencia variable en imagenología maxilofacial. En estudios que evaluaron nivel óseo fueron evaluados principalmente por periodoncistas, en el caso de fracturas radicales fueron evaluados tanto por radiólogos y endodoncistas.

Respecto al tipo de inteligencia artificial utilizada priman diferentes formas de Deep Learning (26 publicaciones), donde las diversas arquitecturas de *convolutional neural network* (CNN) priman sobre otros tipos de arquitecturas de inteligencia artificial. Se mencionan diversos tipos de CNN's entre ellos: *back propagation neural network*, *AlexNet*, *U-Net*, *DetectNet*, entre otros. Se presentan 4 Estudios que utilizan *Machine Learning* dentro de esta revisión.

El rango de efectividad de los estudios seleccionados fue el siguiente: para la identificación dentaria, el valor fue de entre 0.77 y 0.98; para el diagnóstico de caries y restauraciones, el valor fue entre 0.80 y 0.97; para la identificación de fracturas radicales el valor fue entre 0.7 y 0.92; para la detección del nivel de hueso alveolar el valor fue 0.81 a 0.92; para el diagnóstico de lesiones

periapicales el valor fue de y para los quistes 0.85 y 0.92 y tumores el valor fue de 0.70 y 0.94.

DISCUSIÓN

En la última década se ha evidenciado un crecimiento exponencial tanto en búsquedas como estudios sobre inteligencia artificial. La aparición de modelos capaces de aplicar inteligencia artificial en el diagnóstico imagenológico maxilofacial se ha desarrollado con un retraso bastante considerable con respecto al área de la imagenología médica (9)

Este estudio analizó un total de 30 trabajos realizados en los últimos 10 años, lo que permitió evidenciar en cierta medida el progreso de la implementación de estas tecnologías, la que está enfocada principalmente en detección y diagnóstico de lesiones en el territorio maxilofacial. El número de publicaciones que involucraban inteligencia artificial y sus diversos tipos de uso en la imagenología maxilofacial tuvo un aumento considerable, año a año.

Identificación y clasificación dentaria

Khan *et al.*, obtuvo en un desempeño relativamente bajo al poseer solo 2910 imágenes, entregando una precisión de

88.31%. El reentrenamiento de *AlexNet* permitió mejorar la precisión de la clasificación hasta un 98.88%. (10) Singh *et al.*, clasificaron incisivos, caninos, premolares y molares comparando un dataset original de 400 imágenes panorámicas segmentadas en 12000 imágenes de dientes y un dataset aumentado de 48000 imágenes de dientes obteniendo una mejora en la precisión de clasificación de 92 a 95% con la data aumentada probando la eficiencia del sistema propuesto en base a 6 capas de *CNN* vs la arquitectura de *AlexNet*. (11)

Miki *et al.*, encontraron resultados beneficiosos para la aplicación de *CNN* en cortes de CT con una detección de 77.4% con 5.8 falsos positivos por imagen. La precisión de la clasificación fue mucho menor que estudios previos donde se obtuvo un 91% de precisión, pudiendo deberse al pequeño número de muestras. (12) Wang L. *et al.*, propusieron un método GFRLS basado en información 3D y P-CNN. Se evaluó la efectividad de este método en 3 datasets de Micro CT en 3 molares mandibulares distintos mediante la segmentación de dientes y la conversión de imágenes binarias en escala de grises junto a la reducción de

ruido de las imágenes. Los valores de sensibilidad para dentina son en general mayores que los de especificidad, siendo la dentina mejor clasificada y la pulpa siendo confundida con parte de la dentina por lo pequeña en comparación a esta. Sin embargo, pese a las adversidades, los resultados fueron satisfactorios. (13)

Evaluación del nivel óseo radiográfico

Thanathornwong *et al.* propusieron que la cantidad de datos necesarios para el entrenamiento de una red depende de la variabilidad de lo que se desea estudiar y que puede ser simplificada mediante transfer learning, usando la base de un modelo anterior para programar nuevos. La cantidad de datos obtenidos de radiografías periapicales y panorámicas fue reducido a través de la alineación simétrica de los dientes. Se utilizó un aprendizaje profundo de *R-CNN* para detección de objetos. Sensibilidad y especificidad demostraron tener rendimiento aceptable para clasificación de dientes, pero aquellos dientes con compromiso periodontal influyeron en la comparación con el rendimiento de los expertos, obteniendo de igual forma una detección satisfactoria. Sugieren la necesidad de incluir técnicas adicionales

para mejorar el sistema ante investigaciones futuras.(14)

Krois *et al.*, evaluaron el rendimiento de una *CNN* para detectar pérdida ósea periodontal, a la vez que se comparaba con seis dentistas experimentados. Si bien el estudio asumió que una *CNN* es útil al ofrecer mayor precisión y con menor tiempo de diagnóstico que los dentistas, diversos análisis que afectaban el desempeño de ambos grupos conllevaron a rechazar la hipótesis, debido a que no se pudo confirmar un desempeño significativamente mejor. También se destacó el poco acuerdo existente entre los mismos dentistas revisores. Al igual que Thanathornwong *et al.*, estableció que la aplicabilidad y precisión de *CNN* puede mejorar cuando existen más datos e imágenes adicionales. (15)

Detección de caries y restauraciones

Estudios más antiguos incorporados presentaron resultados no tan alentadores. Behere y Lele, mediante el software *Logicon Caries Detection* presentaron valores de sensibilidad bajos a la hora de descartar la presencia de caries (33%), un valor similar al observado por evaluadores. Al diagnosticar caries de esmalte la

sensibilidad fue de 5%.(16) Con una metodología diferente, usando el mismo software, Tracy *et al.*, lograron un valor de sensibilidad de 69%, la diferencia de los resultados radica en que se aplicaron filtros de contraste y brillo para procesar las imágenes.(17) En ambos estudios se clasificaron las lesiones diferenciando los dientes entre los que no tienen presencia de caries, caries en esmalte y caries en dentina, además en ambos trabajos el valor de la especificidad fue alto (97% en Behere y 96% en Tracy), para diagnosticar caries de esmalte.

Wang C *et al.*, resumieron y compararon los resultados de Ronneberger y Lee cuyos valores de sensibilidad al diagnóstico de caries fueron de 0.12 y 0.06 respectivamente mejorando su precisión al momento de diagnosticar caries dentinarias. (18) García Cantú *et al.*, demostraron que su modelo de red neural tenía una sensibilidad mayor (0.75) comparado con los examinadores que revisaron las imágenes (0.36). La especificidad fue más baja en comparación a los examinadores (0.83 versus 0.91) (19).

Lee *et al.*, utilizando una *Deep – CNN*, lograron una sensibilidad de 81% al

diagnosticar lesiones de caries premolares y molares. Afirmaron que la sensibilidad es mayor en molares que en premolares (92.3% versus 84%). La especificidad es mayor en premolares que en molares (94% versus 84%). (20)

Geetha *et al.*, utilizando una *back propagation neural network* lograron una eficacia de 97.1% para el diagnóstico y un falso positivo de 2.8%. El sistema estaba compuesto por una red de 2 capas. Estos autores destacaban que su resultado era superior a los logrados por 4 estudios diferentes cuyas eficacias variaban entre 65% y 94%. (21)

Abdalla-Aslan *et al.*, lograron una precisión de un 93.6% con un modelo de *Machine Vector Support*, demostrando un buen desempeño al momento de detectar y clasificar restauraciones. (22).

Detección de fracturas.

Fukuda *et al.*, basándose en estudios anteriores, seleccionaron radiografías panorámicas que incluían dientes con líneas de fractura radicular vertical (VRF por su sigla en inglés). La creación de un sistema de diagnóstico totalmente automatizado demostró que, aunque detectara aproximadamente 80% de 300 imágenes, el desempeño podría ser

insuficiente para la práctica clínica debido a que más de la mitad de los dientes endodónticamente no tratados con *VRF* no fueron detectados. Esto pudo deberse al bajo número de dientes usado para el proceso de entrenamiento (23).

Johari *et al.*, programaron una red neural probabilística para el diagnóstico de *VRF* en imágenes periapicales y de CBCT para premolares tratados endodónticamente e intactos sin caries. Ambas radiografías se dividieron en diferentes grupos en los cuales se asignaba más información y datos de entrenamiento. Se logró confirmar que el desempeño de una red neural en imágenes CBCT fue más efectivo que en aquel empleado en radiografías periapicales, pero es necesario determinar a futuro el éxito de esta red mediante estudios clínicos (24).

Kositbowornchai *et al.*, desarrollaron el modelo de una red neuronal artificial para diagnosticar *VRF*. La sensibilidad, especificidad y precisión de la red varió en función de la data de reconocimiento. Definieron dos aspectos clave: 1) Mayor número de especímenes para entrenamiento otorga una mejor precisión para reconocer un objeto desconocido y 2) Mientras más cercana sea la varianza

de la función en la data de reconocimiento, mayor el número de correctos diagnósticos realizados.

Al comparar métodos para detección de *VRF*'s, CBCT corresponde a la técnica con mayor sensibilidad y especificidad para detección en estudios clínicos con un 88% y 75% respectivamente. Además, los escaneos de CBCT obtuvieron mayor precisión en detectar *VRF* de 0.2 mm y *VRF* 0.4 mm (70% y 90%, respectivamente) que comparación con radiografía digital (43.3% y 60%, respectivamente). La sensibilidad y especificidad para radiografía intraoral convencional en detección de *VRF* fue de 38% y 87% y en radiografía digital fue de 48 y 89%. Es en este panorama donde la aplicación de redes neuronales artificiales otorga las mayores sensibilidades (97,2-98%) y especificidades comparables a las anteriores (60.0-90.5%). Para la práctica clínica, la aplicación de estas técnicas ofrece una ventaja debido a que guiarse solo por la vista es insuficiente y muchas veces resulta caer en la subjetividad. (25)

Detección y clasificación de quistes

Nurtanio *et al.*, desarrollaron un sistema basado en un *Support Vector Machine* para clasificar lesiones entre quistes o

tumores. Los mejores resultados los obtuvieron al utilizar una combinación de una textura de estadísticas de primer orden (FO), nivel de coocurrencia de la matriz Gray (GLCM, por su sigla en inglés) y Nivel de comprensión de datos de la matriz Gray, (GLRLM, por su sigla en inglés), cuyo valor de eficacia fue de un 87.18%. También se logró la misma eficacia utilizando una combinación de FO y GLRLM. Esta última combinación también obtuvo el resultado de área bajo la curva más alto (0.944) y calificado según una tabla criterios propuesta por Mohanty *et al.*, como un resultado excelente. (26)

Yilmaz *et al.*, probaron diferentes tipos de arquitectura de *CNN's* para clasificar entre lesiones periapicales y queratoquiste odontogénico, siendo *Support Vector Machine* la de mejor comportamiento al diagnosticar con una eficacia de un 94%. (27)

Ariji *et al.*, mediante el software *DIGITS* clasificaron las lesiones en: ameloblastomas, quiste dentígero, quiste radicular, queratoquiste odontogénico y quiste óseo simple. La sensibilidad de la prueba realizada para identificar las diversas lesiones fue de 0.88. La

sensibilidad para diagnosticar ameloblastomas fue de 0.71; para el queratoquiste odontogénico fue de 1 pero tiene un pobre rango de clasificación del modelo de aprendizaje, este valor es de 0.13; para el quiste dentígero fue de 0.88 y el del quiste radicular fue de 0.81. Debido a que la cantidad de casos de quiste óseo simple era muy baja, se excluyeron del proceso. (28)

Lee JH *et al.*, usando *GoogLeNet Inception v3*, logró una sensibilidad de un 88.2 y especificidad de 77% para radiografías panorámicas. En CBCT su sensibilidad fue de 96.1% y especificidad de 77.1%. La eficacia diagnóstica total fue de un 84.6% utilizando radiografías panorámicas, y 91.4% utilizando CBCT. En ambos casos los quistes periapicales quienes tuvieron la mayor eficacia diagnóstica y la peor eficacia se observó en el diagnóstico de queratoquiste odontogénico. (29) Yang *et al.*, propusieron un modelo de *Deep CNN*, basado en *YOLO*, el cual está caracterizado por un procesamiento simple de los datos, donde se detectó y clasificó al mismo tiempo en una imagen. Mediante imágenes de radiografías panorámicas, diferenció entre quiste dentígero, queratoquiste odontogénico y

ameloblastoma, la eficacia diagnóstica fue de 0.663 versus los examinadores que tuvieron eficacias inferiores, en conclusión, el modelo fue más efectivo que los profesionales. (30)

Kwon *et al.*, utilizaron una *Deep CNN*, basada también en el modelo *YOLO*. Se logró para la clasificación general, una sensibilidad de 78.2%, especificidad de 93.9%, eficacia de 95.6% y un área bajo la curva de 0.94, utilizando la *CNN* con el conjunto de datos aumentado. Este estudio concluyó que, a pesar del número limitado de imágenes, el utilizar una *CNN* con el conjunto de datos aumentado permitía tener resultados elevados de sensibilidad, especificidad, eficacia y área bajo la curva. (31)

Diagnóstico de lesiones periapicales

Mahmoud *et al.*, probaron una *Feed Forward Neural Network* y una *K-Nearest Neighbor Classifier*, siendo esta última la que obtuvo el mejor desempeño al momento de identificar lesiones periapicales, la eficacia en este caso fue de un 82.61%. (32)

Osmanovic *et al.*, realizaron un modelo de *Random Forest*, el modelo más prometedor de los aplicados tuvo una eficacia de 85.71% en promedio, siendo

más efectiva para el diagnóstico de quistes con una eficacia de 87.7% comparado con el de granulomas, cuya eficacia fue de 82.1%. Este modelo incluía un total de 7 características. Otros modelos incluían menos características y presentaban menor efectividad general, mayor a la hora de diagnosticar quistes, pero muy menor a la hora de diagnosticar granulomas.(33)

Orhan *et al.*, utilizando imágenes de CBCT y a través una *Deep CNN*, lograron tener una detección correcta de lesiones periapicales de un 92.8%, un resultado similar al examinador.(34)

Ekert *et al.*, desarrollaron una red neural de 7 capas, cuya sensibilidad y especificidad fueron de 0.65 and 0.87. Concluyeron que si bien su modelo podría diagnosticar desde radiografías panorámicas, debía mejorar en su sensibilidad para ser empleado como una herramienta de apoyo al clínico.(35)

Dentro de las limitaciones de este estudio se puede indicar que la estrategia de búsqueda realizada incorporó siglas de los diversos tipos de Inteligencia Artificial, entre esas siglas se encuentra la sigla de "*Machine Learning*" (ML). Esta sigla incorporó en la búsqueda publicaciones

relacionadas con “Medicina Legal” (ML), lo que hizo que el número de publicaciones en la búsqueda inicial fuese aumentado en comparación a la cantidad de publicaciones finalmente seleccionadas para las distintas fases de revisión.

Muchas publicaciones en su título no mencionaban la presencia de Inteligencia Artificial en su investigación, lo que pudo eliminar algunas publicaciones en la selección final.

Se presentó en algunas publicaciones, resultados de desempeño general, cuando analizaban más de una variable, por ende, no se puede saber si este sistema tuvo mejores resultados al analizar una de las variables en particular.

CONCLUSIÓN

Los resultados obtenidos en el presente estudio han permitido revisar que la evidencia disponible sobre diagnóstico asistido por computadora en sistemas basados en inteligencia artificial, pese a lo discutido, tiene sensibilidad y especificidad similar a la de clínicos, sin embargo, tampoco se puede depender solo de estas aplicaciones inteligentes ya que se detectó que existen métodos que afectan este desempeño. Tipo de

entrenamiento de los modelos neuronales, cantidad de información para entrenar los modelos y segmentación de imágenes resultarían claves en la efectividad de clasificación y predicción de lesiones.

De acuerdo con resultados explorados, se puede inferir que los sistemas basados en redes neurales convolucionales (o *CNN* por su sigla en inglés), presentan un mayor desempeño en comparación a otros sistemas basados en Inteligencia Artificial. Las *CNN* han presentado una mayor precisión al momento de clasificar y diagnosticar lesiones en el territorio maxilofacial.

La mejora en sistemas computacionales ha permitido acelerar los tiempos tanto de desarrollo de modelos como de análisis de grandes conjuntos de datos, obteniendo resultados comparables a las de un experto. Aun así, se necesita mayores investigaciones en los algoritmos y sobre su aplicación clínica en la práctica para demostrar el impacto real en la práctica diagnóstica.

Adicionalmente se requiere crear una terminología básica en Inteligencia Artificial, para estandarizar y comprender los términos asociados.

BIBLIOGRAFÍA

1. Mateo F. Redes neuronales y preprocesado de variables para modelos y sensores en bioingeniería. Thesis. 2012;322.
2. Choy G, Khalilzadeh O, Michalski M, Do S, Samir AE, Pinykh OS, et al. Current applications and future impact of machine learning in radiology. *Radiology*. 2018;288(2):318–28.
3. Trujillano Cabello J, Badía Castelló M, March Llanes J, Rodríguez Pozo À, Serviá Goixart L, Sorribas Tello A. Redes neuronales artificiales en medicina intensiva. Ejemplo de aplicación con las variables del MPM II. *Med Intensiva [Internet]*. 2005;29(1):13–20. Available from: [http://dx.doi.org/10.1016/S0210-5691\(05\)74198-X](http://dx.doi.org/10.1016/S0210-5691(05)74198-X)
4. Izeboudjen N, Larbes C, Farah A. A new classification approach for neural networks hardware: From standards chips to embedded systems on chip. *Artif Intell Rev*. 2014;41(4):491–534.
5. López BR. ¿Cuáles son las diferencias entre Data Science, Machine Learning, Deep Learning y Artificial Intelligence? [Internet]. Available from: <https://www.cursosgis.com/cuales-son-las-diferencias-entre-data-science-machine-learning-deep-learning-y-artificial-inteligence/>
6. González Muñoz A, Díaz Blanco I. Aplicaciones de Técnicas de Inteligencia Artificial Basadas en Aprendizaje Profundo (deep learning) al Análisis y Mejora de la Eficiencia de Procesos Industriales. 2018;(February 2018):75.
7. Giger ML. Machine Learning in Medical Imaging. *J Am Coll Radiol [Internet]*. 2018;15(3):512–20. Available from: <https://doi.org/10.1016/j.jacr.2017.12.028>
8. Lakhani P, Gray DL, Pett CR, Nagy P, Shih G. Hello World Deep Learning in Medical Imaging. *J Digit Imaging*. 2018;31(3):283–9.
9. Schwendicke F, Golla T, Dreher M, Krois J. Convolutional neural networks for dental image diagnostics: A scoping review. *J Dent*. 2019 Dec;91:103226.
10. Khan S, Mukati A. Dataset augmentation for machine learning applications of dental radiography. *Int J Adv Comput Sci Appl*. 2020;11(2):453–6.
11. Singh P, Sehgal P. Numbering and Classification of Panoramic Dental Images Using 6-Layer Convolutional Neural Network. *Pattern Recognit Image Anal*. 2020;30(1):125–33.
12. Miki Y, Muramatsu C, Hayashi T, Zhou X, Hara T, Katsumata A, et al. Tooth labeling in cone-beam CT using deep convolutional neural network for forensic identification. *Med Imaging 2017 Comput Diagnosis*. 2017;10134:101343E.
13. Wang L, Li S, Chen R, Liu SY, Chen JC. An automatic

- segmentation and classification framework based on PCNN model for single tooth in MicroCT images. *PLoS One*. 2016;11(6):1–19.
14. Chen Y, Stanley K, Att W. Artificial intelligence in dentistry: current applications and future perspectives. Vol. 51, Quintessence International. 2020. p. 248–57.
 15. Krois J, Ekert T, Meinhold L, Golla T, Kharbot B, Wittemeier A, et al. Deep Learning for the Radiographic Detection of Periodontal Bone Loss. *Sci Rep*. 2019;9(1).
 16. Behere RR, Lele SM. Reliability of Logicon caries detector in the detection and depth assessment of dental caries: an in-vitro study. *Indian J Dent Res Off Publ Indian Soc Dent Res*. 2011;22(2):362.
 17. Tracy KD, Dykstra BA, Gakenheimer DC, Scheetz JP, Lacina S, Scarfe WC, et al. Utility and effectiveness of computer-aided diagnosis of dental caries. *Gen Dent*. 2011;59(2):136–44.
 18. Wang C-W, Huang C-T, Lee J-H, Li C-H, Chang S-W, Siao M-J, et al. A benchmark for comparison of dental radiography analysis algorithms. *Med Image Anal*. 2016 Jul;31:63–76.
 19. Cantu AG, Gehrung S, Krois J, Chaurasia A, Rossi JG, Gaudin R, et al. Detecting caries lesions of different radiographic extension on bitewings using deep learning. *J Dent*. 2020 Sep;100:103425.
 20. Lee J-H, Kim D-H, Jeong S-N, Choi S-H. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. *J Dent*. 2018 Oct;77:106–11.
 21. Geetha V, Aprameya KS, Hinduja DM. Dental caries diagnosis in digital radiographs using back-propagation neural network. *Heal Inf Sci Syst*. 2020 Dec;8(1):8.
 22. Abdalla-Aslan R, Yeshua T, Kabla D, Leichter I, Nadler C. An artificial intelligence system using machine-learning for automatic detection and classification of dental restorations in panoramic radiography. *Oral Surg Oral Med Oral Pathol Oral Radiol*. 2020 Jun;
 23. Fukuda M, Inamoto K, Shibata N, Arijii Y, Yanashita Y, Kutsuna S, et al. Evaluation of an artificial intelligence system for detecting vertical root fracture on panoramic radiography. *Oral Radiol [Internet]*. 2020;36(4):337–43. Available from: <https://doi.org/10.1007/s11282-019-00409-x>
 24. Type CT, Prof TK, Conflicts P, Morawska-kochman M, Morawska-kochman M. Detection of Vertical Root Fractures in Intact and Endodontically Treated Premolar Teeth by Designing a Probabilistic Neural Network: An ex Vivo Study. Type, CT, Prof, TK, Conflicts, P, Morawska-kochman, M, Morawska-kochman, M, 2016 © 2016 Authors Publ by Br Inst Radiol. 2016;(C).
 25. Kositbowornchai S, Plermkamon S,

- Tangkosal T. Performance of an artificial neural network for vertical root fracture detection: An ex vivo study. *Dent Traumatol.* 2013;29(2):151–5.
26. Nurtanio I, Purnama IE, Hariadi M, Purnomo MH. Classifying Cyst and Tumor Lesion Using Support Vector Machine Based on Dental Panoramic Images Texture Features. In 2013.
27. Yilmaz E, Kayikcioglu T, Kayipmaz S. Computer-aided diagnosis of periapical cyst and keratocystic odontogenic tumor on cone beam computed tomography. *Comput Methods Programs Biomed [Internet].* 2017;146:91–100. Available from: <http://dx.doi.org/10.1016/j.cmpb.2017.05.012>
28. Arijji Y, Yanashita Y, Kutsuna S, Muramatsu C, Fukuda M, Kise Y, et al. Automatic detection and classification of radiolucent lesions in the mandible on panoramic radiographs using a deep learning object detection technique. *Oral Surg Oral Med Oral Pathol Oral Radiol.* 2019;128(4):424–30.
29. Lee J-H, Kim D-H, Jeong S-N. Diagnosis of cystic lesions using panoramic and cone beam computed tomographic images based on deep learning neural network. *Oral Dis.* 2020 Jan;26(1):152–8.
30. Yang H, Jo E, Kim HJ, Cha I, Jung Y-S, Nam W, et al. Deep Learning for Automated Detection of Cyst and Tumors of the Jaw in Panoramic Radiographs. *J Clin Med.* 2020;9(6):1839.
31. Kwon O, Yong T-H, Kang S-R, Kim J-E, Huh K-H, Heo M-S, et al. Automatic diagnosis for cysts and tumors of both jaws on panoramic radiographs using a deep convolution neural network. *Dentomaxillofac Radiol.* 2020 Jul;20200185.
32. Mahmoud YE, Labib SS, Mokhtar HMO. Teeth periapical lesion prediction using machine learning techniques. *Proc 2016 SAI Comput Conf SAI 2016.* 2016;129–34.
33. Osmanovic A, Halilovic S, Jukic S, Kevric J, Hadziabdic N. Application of a Computer-Aided Diagnostic System for Early Identification of Periapical Lesions---A Pilot Study. In: Badnjevic A, Škrbić R, Gurbeta Pokvić L, editors. *CMBEBIH 2019.* Cham: Springer International Publishing; 2020. p. 75–9.
34. Orhan K, Bayrakdar IS, Ezhov M, Kravtsov A, Özyürek T. Evaluation of artificial intelligence for detecting periapical pathosis on cone-beam computed tomography scans. *Int Endod J.* 2020;53(5):680–9.
35. Ekert T, Krois J, Meinhold L, Elhennawy K, Emara R, Golla T, et al. Deep Learning for the Radiographic Detection of Apical Lesions. *J Endod.* 2019 Jul;45(7):917-922.e5.