

Evaluación de la fiabilidad del programa basado en Inteligencia Artificial
Denti.Ai para la identificación de estructuras y tratamientos dentarios
presentes en radiografías panorámicas.



TRABAJO FIN DE MÁSTER

Alumna: Mónica Bonfanti Gris
Tutor: Prof. Dr. Guillermo Pradíes Ramiro

MÁSTER EN CIENCIAS ODONTOLÓGICAS

Facultad de Odontología, Universidad Complutense de Madrid
Curso Académico 2020-2021

3. Normativa

Los firmantes del presente compromiso declaran conocer la normativa vigente reguladora para la realización y defensa de los TFM y aceptan las disposiciones contenidas en la misma.

4. Obligaciones del estudiante de Máster

- Elaborar, consensuado con el/los Tutor/es del TFM un cronograma detallado de trabajo que abarque el tiempo total de realización del mismo hasta su lectura.
- Informar regularmente al Tutor/es del TFM de la evolución de su trabajo, los problemas que se le planteen durante su desarrollo y los resultados obtenidos.
- Seguir las indicaciones que, sobre la realización y seguimiento de las actividades formativas y la labor de investigación, le hagan su tutor/es del TFM.
- Velar por el correcto uso de las instalaciones y materiales que se le faciliten por parte de la Universidad Complutense con el objeto de llevar a cabo su actividad de trabajo, estudio e investigación.

5. Obligaciones del tutor/es del TFM

- Supervisar las actividades formativas que desarrolle el estudiante; así como desempeñar todas las funciones que le sean propias, desde el momento de la aceptación de la tutorización hasta su defensa pública.
- Facilitar al estudiante la orientación y el asesoramiento que necesite.

6. Buenas prácticas

El estudiante y el tutor/es del TFM se comprometen a seguir, en todo momento, prácticas de trabajo seguras, conforme a la legislación actual, incluida la adopción de medidas necesarias en materia de salud, seguridad y prevención de riesgos laborales.

También se comprometen a evitar la copia total o parcial no autorizada de una obra ajena presentándola como propia tanto en el TFM como en las obras o los documentos literarios, científicos o artísticos que se generen como resultado del mismo. Para tal, el estudiante firmará la Declaración de No Plagio del ANEXO I, que será incluido como primera página de su TFM.

7. Procedimiento de resolución de conflictos académicos

En el caso de producirse algún conflicto derivado del incumplimiento de alguno de los extremos a los que se extiende el presente compromiso a lo largo del desarrollo de su TFM, incluyéndose la posibilidad de modificación del nombramiento del

tutor/es, la coordinación del máster buscará una solución consensuada que pueda ser aceptada por las partes en conflicto. En ningún caso el estudiante podrá cambiar de Tutor directamente sin informar a su antiguo Tutor y sin solicitarlo oficialmente a la Coordinación del Máster.

En el caso de que el conflicto persista se gestionará según lo previsto en el SGIC de la memoria verificada.

8. Confidencialidad

El estudiante que desarrolla un TFM dentro de un Grupo de Investigación de la Universidad Complutense, o en una investigación propia del Tutor, que tenga ya una trayectoria demostrada, o utilizando datos de una empresa/organismo o entidad ajenos a la Universidad Complutense de Madrid, se compromete a mantener en secreto todos los datos e informaciones de carácter confidencial que el Tutor/es del TFM o de cualquier otro miembro del equipo investigador en que esté integrado le proporcionen así como a emplear la información obtenida, exclusivamente, en la realización de su TFM.

Asimismo, el estudiante no revelará ni transferirá a terceros, ni siquiera en los casos de cambio en la tutela del TFM, información del trabajo, ni materiales producto de la investigación, propia o del grupo, en que haya participado sin haber obtenido, de forma expresa y por escrito, la autorización correspondiente del anterior Tutor del TFM.

9. Propiedad intelectual e industrial

Cuando la aportación pueda ser considerada original o sustancial el estudiante que ha elaborado el TFM será reconocido como cotitular de los derechos de propiedad intelectual o industrial que le pudieran corresponder de acuerdo con la legislación vigente.

10. Periodo de Vigencia

Este compromiso entrará en vigor en el momento de su firma y finalizará por alguno de los siguientes supuestos:

- Cuando el estudiante haya defendido su TFM.
- Cuando el estudiante sea dado de baja en el Máster en el que fue admitido.
- Cuando el estudiante haya presentado renuncia escrita a continuar su TFM.
- En caso de incumplimiento de alguna de las cláusulas previstas en el presente documento o en la normativa reguladora de los Estudios de Posgrado de la Universidad Complutense.

La superación académica por parte del estudiante no supone la pérdida de los derechos y obligaciones intelectuales que marque la Ley de Propiedad Intelectual para ambas partes, por lo que mantendrá los derechos de propiedad intelectual sobre

su trabajo, pero seguirá obligado por el compromiso de confidencialidad respecto a los proyectos e información inédita del tutor.

Firmado en Madrid, a 2 de SEPTIEMBRE de 20 21

El estudiante de Máster Fdo.:	El Tutor/es Fdo.: 
--------------------------------------	--

SR. COORDINADOR DEL MÁSTER EN CIENCIAS ODONTOLÓGICAS

ANEXO I: DECLARACIÓN DE NO PLAGIO

D./Dña. MÓNICA BONFANTI GRIS
con NIF _____, estudiante de Máster en la Facultad de
ODONTOLOGÍA de la Universidad Complutense de Madrid en el
curso 20 -20 , como autor/a del trabajo de fin de máster titulado
EVALUACIÓN DE LA FIABILIDAD DEL PROGRAMA BASADO EN INTELIGENCIA ARTIFICIAL
DENTI.AI PARA LA IDENTIFICACIÓN DE ESTRUCTURAS Y TRATAMIENTOS DENTARIOS EN
RADIOGRAFIAS PANORÁMICAS
y presentado para la obtención del título correspondiente, cuyo/s tutor/ es/son:

DR. PROF. GUILLERMO JESÚS PRADIES RAMIRO

DECLARO QUE:

El trabajo de fin de máster que presento está elaborado por mí y es original. No copio, ni utilizo ideas, formulaciones, citas integrales e ilustraciones de cualquier obra, artículo, memoria, o documento (en versión impresa o electrónica), sin mencionar de forma clara y estricta su origen, tanto en el cuerpo del texto como en la bibliografía. Así mismo declaro que los datos son veraces y que no he hecho uso de información no autorizada de cualquier fuente escrita de otra persona o de cualquier otra fuente.

De igual manera, soy plenamente consciente de que el hecho de no respetar estos extremos es objeto de sanciones universitarias y/o de otro orden.

En Madrid, a 2 de SEPTIEMBRE de 20 21

Fdo.:

Esta DECLARACIÓN debe ser insertada en primera página de todos los trabajos fin de máster conducentes a la obtención del Título.



UNIVERSIDAD COMPLUTENSE DE MADRID. Facultad de Odontología
TRABAJO DE FIN DE MÁSTER
VISTO BUENO DEL TUTOR
MASTER OFICIAL EN CIENCIAS ODONTOLÓGICAS

El profesor/a tutor

Nombre y apellidos:	GUILLERMO JESÚS PRADÍES RAMIRO
---------------------	--------------------------------

del alumno/a

Nombre y apellidos	MÓNICA BONFANTI GRIS
--------------------	----------------------

encuadrado en la línea de investigación

LÍNEAS Y MÉTODOS DE INVESTIGACIÓN EN MATERIALES Y TECNOLOGÍA DE APLICACIÓN PROTÉSICA.

DA EL VISTO BUENO
para que el Trabajo de Fin de Máster titulado

Evaluación de la fiabilidad del programa basado en Inteligencia Artificial Denti.Ai para la identificación de estructuras y tratamientos dentarios presentes en radiografías panorámicas
--

sea admitido para su defensa ante Tribunal.

En MADRID , a 1 de SEPTIEMBRE de 2021 .

Fdo: el profesor/a

El presente Visto Bueno se debe acompañar del Trabajo de Investigación en formato electrónico y tres copias en papel

ÍNDICE

1. Resumen.....	3
2. Abstract.....	4
3. Introducción.....	5
4. Justificación.....	11
5. Objetivos.....	12
6. Hipótesis de trabajo.....	13
7. Material y Método.....	14
7.1. Selección de Muestra.....	14
7.2. Análisis de Muestra.....	14
7.3. Diseño de Estudio.....	15
7.4. Recogida de datos.....	18
7.5. Análisis Estadístico.....	18
8. Resultados.....	20
9. Discusión.....	36
9.1. Discusión sobre la metodología utilizada.....	36
9.2. Discusión sobre los resultados obtenidos.....	40
10. Conclusiones.....	45
11. Bibliografía.....	46
12. Anexos.....	50

Resumen

Introducción. La implantación de protocolos digitales, incluidos herramientas de Inteligencia artificial en el ámbito odontológico ha supuesto una completa revolución en la rutina diaria de los clínicos dedicados a la profesión. Actualmente, el uso de softwares que complementan el diagnóstico clínico realizado por el profesional están siendo cada vez más desarrollados e implementados en la rutina diaria.

Objetivo. Evaluar la capacidad de diagnóstico del programa basado en inteligencia artificial “Denti.Ai” en relación a las estructuras y tratamientos dentarios presentes en ortopantomografías (OPGs).

Material y Método. Un total de 300 radiografías panorámicas (OPG) de pacientes adultos fueron aleatoriamente seleccionadas y utilizadas para la consecución de este estudio. Las imágenes fueron analizadas manualmente por dos operadores con experiencia en radiodiagnóstico, indicando y clasificando las piezas dentarias (Nomenclatura FDI) y tratamientos presentes en ellas (obturaciones metálicas, RM; obturaciones plásticas, RP; Tratamiento de Conductos, TC; Coronas, C; Implantes, IOI). En el caso de discrepancia en el diagnóstico un 3 evaluador tomó parte en la decisión final. Posteriormente las imágenes fueron cargadas y analizadas por el software Denti.Ai. Los resultados fueron registrados y analizados mediante la realización de un estudio estadístico basado en la utilización de estadística descriptiva, realización de pruebas diagnósticas: sensibilidad (S), especificidad (E), valor predictivo positivo (VPP), valor predictivo negativo (VPN) del programa Denti.Ai, y su posterior representación en Curvas de Roc; y pruebas de concordancia.

Resultados: El programa Denti.Ai es capaz de identificar correctamente las estructuras dentarias en el 69.3% de los casos, obteniendo errores de diversas categorías en 30.7% restante. Su precisión diagnóstica para la detección de tratamientos dentarios es de 41.11% para restauraciones plásticas, 85.48% para restauraciones metálicas, 91.9% para tratamientos de conductos, 89.53% para coronas y 100% para implantes osteointegrados. La concordancia obtenida entre el Gold Standard y el programa fue siempre superior al 0.90.

Conclusiones: Denti.Ai puede considerarse un software cuyas tareas permiten dar soporte y asistencia al profesional odontólogo en procesos de diagnóstico, prevención y pronóstico dentarios. Mayor número de estudios implicados en la capacidad diagnóstica del programa deberían ser desarrollados con el fin de mejorar el desempeño del mismo.

Palabras clave: *Inteligencia Artificial; Diagnóstico asistido; Redes Neuronales Convolucionales; Radiografía Panorámica; Denti.Ai.*

Abstract

Introduction. The implementation of digital protocols, including Artificial Intelligence tools in the dental field, has meant a complete revolution in the daily routine of clinicians dedicated to the profession. Currently, the use of software that complements the clinical diagnosis carried out by the professional is being increasingly developed and implemented in the daily routine.

Objectives. To evaluate the diagnostic capacity of the program based on artificial intelligence "Denti.Ai" in relation to the dental structures and treatments present in panoramic radiographs.

Material and methods. 300 OPG radiographs of adult patients were randomly selected and used to carry out this study. The images were analyzed manually by a total of two operators with experience in radiodiagnosis, indicating and classifying the dental pieces (FDI Nomenclature) and treatments present in them (metal fillings, RM; plastic fillings, RP; Canal Treatment, CT; Crowns, C; Implants, IOI). Later the images were added and analyzed by the Denti.Ai software. The results were stored and finally analyzed by conducting a statistical study based on the measurement of sensitivity (S), specificity (E), positive predictive value (PPV), negative predictive value (NPV) of the Denti.Ai program, and its later representation in Roc Curves.

Results: Artificial Intelligence-based program (Denti.Ai) is capable of identifying dental structures correctly 69.3% of cases, obtaining errors from various categories in the remaining 30.7%. Its diagnostic accuracy for the detection of different dental treatments is 41.11% for plastic restorations, 85.48% for metal restorations, 91.9% for root canal treatments, 89.53% for crowns and 100% for osseointegrated implants. The ICC obtained between the Gold Standard and the program was always higher than 0.90.

Conclusions: Denti.Ai can be considered a software that allows to give support and assistance to dental clinicians in dental diagnosis, prevention, and prognosis processes. However, a greater number of studies involved in the diagnostic capacity of the program should be developed in order to improve its performance.

Keywords: *Artificial Intelligence; Computer-aided Diagnostics; Convolutional Neural Networks; Panoramic Radiograph; Denti.Ai.*

Introducción.

El **diagnóstico** es, uno de los procedimientos de mayor importancia en el ámbito clínico odontológico. Se trata de un proceso de maduración del caso que no sólo se compone de la experiencia y conocimiento del profesional odontólogo, sino que también se apoya en técnicas complementarias que aportan información esencial que, de otra forma, el clínico sería incapaz de obtener, como el radiodiagnóstico.

Tradicionalmente, y dependiendo de la patología que presentase el paciente, el **diagnóstico clínico se ha realizado por medio de la anamnesis y las exploraciones extraoral e intraoral**, incluyendo pruebas de percusión para la valoración de lesiones periapicales o periodontales, pruebas térmicas para la determinación de la vitalidad pulpar, sondaje periodontal o evaluación oclusal, entre otras.

El **diagnóstico radiográfico siempre ha estado presente en odontología** desde el descubrimiento de los rayos X por *Wilhelm Röntgen* en 1896. Este tipo de pruebas aportan una serie de imágenes que son de gran utilidad para el profesional, dado que **permiten extraer una información muy valiosa que determina el tratamiento y pronóstico finales** de las piezas dentarias o estructuras bucales afectadas.

Aunque este tipo de técnicas sean de indudable ayuda en la rutina diaria de cualquier clínico, se ha de tener en cuenta que requieren una gran curva de aprendizaje para ser realmente útiles. Así, es de vital importancia saber identificar y analizar todos los signos radiográficos que indiquen la presencia de posibles patologías o tratamientos dentarios.

Debido a la fatiga ocular y mental general a lo largo del tiempo de consulta a la que se ven expuestos los profesionales odontólogos, se puede decir que el criterio de un operador no siempre es el mismo y que, muchas veces, se pasan por alto determinados signos radiográficos de interés que podrían marcar un punto de inflexión en el tratamiento de cualquier paciente. Además, la frecuente aparición de artefactos en las imágenes radiográficas puede dificultar el diagnóstico o provocar una mala interpretación de las mismas.

La introducción en el mercado biosanitario de aplicaciones de los recursos matemáticos e informáticos basados en la **Inteligencia Artificial (IA)**, tiene por objetivo facilitar y mejorar la capacidad diagnóstica y terapéutica de los profesionales. Se trata de un concepto establecido por *John MacCarthy* en 1955 que se refiere a la **capacidad de ciertas maquinarias de imitar el conocimiento y el comportamiento humanos** implementados a base de secuencias de algoritmos. (1)

En 1978, el matemático *Richard Bellman* definió la IA como la automatización de actividades relacionadas con habilidades propias del pensamiento humano, incluyendo la capacidad de aprendizaje, toma de decisiones y resolución de problemas. (2) El objetivo de este tipo de tecnología tan revolucionaria es desarrollar “**cerebros inteligentes**” integrados en máquinas que aprendan a base de “nutrirlos” con datos (data).

Se trata de cerebros que han sido creados en imitación a los humanos y cuyos principales constituyentes son las **neuronas artificiales**.

A base de la concatenación y conexión de estas, se consiguen crear una serie de capas que, junto con la aplicación de operaciones matemáticas específicas, desarrollan una red que permite la resolución de una tarea en concreto como la clasificación de estructuras en imágenes radiográficas, o de determinados tratamientos visibles en las mismas, por ejemplo. (3,4)

Dentro de la IA se pueden diferenciar los siguientes campos de estudio: (2) **Fig. 1.**

- **Machine Learning (ML):** se trata de un sistema creado con algoritmos y diseñado para predecir resultados basándose en un conjunto de datos obtenidos de forma previa. Dentro de este, las formas de inteligencia artificial más comúnmente utilizados en Odontología son los denominados *Genetic Algorithm* (Algoritmo Genético), *Artificial Neural Network* (ANN) y *Fuzzy Logic* (lógica difusa). (4)
 - **Genetic Algorithm:** Este subtipo de ML ha ganado cierta fama a lo largo de la última década por su progresiva optimización en la resolución de problemas complejos. (4)
 - **Artificial Neural Network (ANN):** *Minsky y Dean Edmonds*, en 1951, fueron los primeros en desarrollar lo que se considera actualmente como la primera ANN. Estos conjuntos neuronales fueron inventados manteniendo la estructura del cerebro artificial en su forma más básica, aunque permitiéndole desarrollar ciertas habilidades en imitación a su semejante humano. La mayor ventaja de este tipo de sistemas es **su capacidad de resolución de problemas que son considerados como imposibles por métodos convencionales**. Son especialmente útiles en el **diagnostico de patologías, en la identificación biomédica e incluso en el análisis de imágenes**. Se trata del tipo de ML más empleado en Odontología. (4)
 - **Fuzzy Logic:** El término *fuzzy* (en inglés) hace alusión a todo aquello que no está claro, o que parece vago de entender. En aquellas situaciones en las que el clínico no sepa tomar una decisión es cuando este subgrupo de ML proporciona un posible razonamiento alternativo y, a su vez, una solución al problema. Este término fue acuñado por *Lotfi Zadeh* y se compone de cuatro importantes partes: **Rule Base** (conocimiento base compuesto de una base de reglas y datos iniciales), **Fuzzification** (proceso de transformación de la información entrada en lenguaje apto para el cerebro artificial), **Inference Engine** (realiza los procesos de inferencia sobre sus propias reglas y datos) y **Defuzzification** (generación de una inferencia específica). (4)
- **Neural Networks (NN):** se trata de un sistema de algoritmos cuya finalidad es el cálculo de señales gracias a una red neuronal formada por neuronas artificiales. (4)
- **Deep Learning (DL) o Convolutional Neural Network (CNN):** red neuronal conformada por distintas capas computacionales con el fin de identificar patrones o características de forma automática. (4)

No obstante, desde la creación de esta tecnología, las aplicaciones de la IA en la vida real han sido percibidas tanto positiva como negativamente, dado que puede suponer un conflicto ético si es aplicado al diagnóstico en Medicina sin la posterior supervisión humana. Sin embargo, con el paso de los años y el desarrollo de las ingenierías, la situación actual es muy prometedora: la última década ha marcado una extraordinaria evolución en cuanto a lo que ML se refiere. (3)

En el ámbito biosanitario, el papel de la IA en el diagnóstico en imágenes de patologías es fundamental, dado que ayuda a eliminar cualquier tipo discrepancia de opinión como parte del sesgo producido por la subjetividad del operador. Además, permite analizar de forma inteligente la evolución del paciente a lo largo de los años, estableciendo un posible pronóstico. (3)

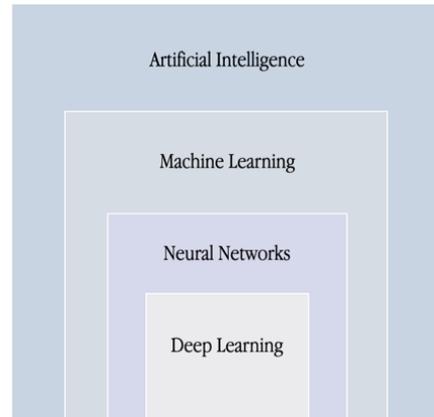


Figura 1. Esquema ilustrativo de los componentes de que engloba la IA.

Hasta ahora se ha descrito su empleo en la detección de patologías cardiovasculares, en estudios sobre fármaco-terapia relacionados con el manejo de diversas patologías, así como en la mejora del diagnóstico de patologías retino-oculares. (1) Además, su uso en dermatología está siendo cada vez más protagonista: se han publicado estudios sobre la detección de nevus y la clasificación de los mismos (en benignos o malignos), así como para la determinación de gravedad de enfermedades inflamatorias (como psoriasis), a través de procedimientos de reconocimiento por imagen. (5–8)

De igual forma, también ha sido aplicada en modelos de análisis pronósticos y predictivos, en la identificación de signos desconocidos relacionados a diferentes patologías en imágenes radiográficas, y como soporte para la toma de decisiones en tratamientos de nueva introducción. (9)

¿Cuáles pueden ser las aplicaciones de la IA en Odontología? Existen muchas posibles alternativas a la aplicación de este tipo de tecnología en nuestro campo. Algunos ejemplos son: (10)

- **Evaluación de la limpieza oral del paciente**, por medio del desarrollo de cepillos de dientes inteligentes que evalúan la técnica de higiene del paciente. (11)
- **Destrucción de biofilm intraconducto en endodoncia**, por medio de micro-robots con habilidades catalíticas. Además, también se ha estudiado su uso en la prevención de infecciones peri-implantarias. (12)
- **Contribuir en la preparación dentaria para prótesis fijas**, por medio de sistemas que ayudan al clínico a la hora de hacer la preparación de la pieza dentaria (13), el uso de sistemas CAD/CAM para conseguir restauraciones protésicas finales precisas, etc. (4)
- **Registrar movimientos mandibulares**, por medio de sistemas específicos para ello. También se ha descrito la existencia de articuladores virtuales. (14)
- **Incorporación de tratamientos ortodóncicos guiados por IA** (ortodoncia de alineadores) y planificación de cirugías ortognáticas guiadas.

- **Incorporación de sistemas de IA en escáneres intraorales** para la eliminación de estructuras no relacionadas con el escaneado, como tejidos blandos (mejillas, lengua, etc.)

Además, como se ha mencionado anteriormente, los profesionales dedicados a esta rama de la salud deben aplicar todo su conocimiento en la detección de patologías, tanto de forma clínica como por medio de pruebas complementarias.

Aunque las imágenes radiodiagnósticas son de indudable apoyo, se debe tener en consideración que no todos los odontólogos están igualmente formados ni tienen la misma experiencia clínica, lo cual puede interferir en su capacidad diagnóstica y de análisis. Por ello, la IA puede ser de gran utilidad en estos casos: puede automatizar el diagnóstico y guiar al clínico en la toma de una mejor decisión para asegurar un buen pronóstico al paciente. (2)

El desarrollo de estos sistemas inteligentes se ha visto aplicado tanto en el análisis de imágenes de odontología general (radiografías panorámicas, radiografías periapicales, etc.) como en otras especialidades como cirugía (imágenes en 3D) u ortodoncia (cefalometrías). (15–17)

En concreto, en lo que se refiere a las **radiografías panorámicas u ortopantomografías (OPG)**, la visualización de todas las estructuras de interés oro-faciales en una única imagen, así como la comodidad con la que se lleva a cabo y la baja radiación que es absorbida por el paciente (14,15), hacen de ella una imagen esencial con la que los profesionales han de partir en su proceso de maduración del caso. (20)

Se han descrito numerosos estudios realizados sobre OPG cuyo principal objetivo es la evaluación de diversas redes neuronales que permiten la detección de estructuras anatómicas, posibles patologías, tanto óseas (21,22) como dentarias y tratamientos ya realizados en la cavidad oral del paciente. (20)

La aplicación de tecnologías basadas en IA en nuestro campo puede ser de interés porque, según lo comentado en el artículo publicado por *Rushton et al.*, **existe cierta tendencia a que el profesional odontólogo no diagnostique todas las patologías presentes en la imagen panorámica**. Específicamente, se afirma que en lo referente a las patologías no dentarias (como lesiones quísticas, patologías del seno maxilar, etc.), éstas suelen ser mayormente diagnosticadas por radiólogos dentales especializados. (23)

No obstante, y como todo en investigación, el uso de cerebros artificiales en el diagnóstico odontológico ha sufrido un largo proceso de evolución en los últimos años:

- En un primer momento se desarrollaron programas inteligentes basados en algoritmos con el fin de realizar una **identificación y clasificación de las piezas dentarias y maxilo-faciales**. *Patanachai et al.* propusieron un sistema de segmentado dentario basado en onículas transformadas a la vez que *Poonsri et al.* introdujeron un algoritmo completamente automático que permitía el mismo propósito. (24,25)
- Posteriormente, *Oliveira et al.* desarrollaron un algoritmo que permitía la segmentación y la caracterización de las piezas dentarias con y sin lesiones cariosas. Es decir, **se comenzó a introducir una nueva variable que permitiese identificar patologías**. No

obstante, este algoritmo no detectaba con precisión la localización exacta de la caries. (26).

- La utilización de IA en el ámbito odontológico no puede únicamente basarse en la identificación, segmentación y clasificación de piezas dentarias y sus patologías. Poco a poco se han introducido otras posibles formas de aplicar estos sistemas tan novedosos, como son la predicción eruptiva del tercer molar en función de su angulación (27) o la detección de estadios tempranos de patologías óseas como la osteoporosis (22).

A día de hoy, se podría decir que la IA está cada vez más cerca de ser utilizada rutinariamente por los clínicos dedicados a la salud de sus pacientes. En Odontología, los avances en los últimos años han sido abrumadores: además de lo mencionado anteriormente, se están entrenando softwares para la detección de tumores mandibulares, fracturas radiculares, posibles signos radiológicos asociados al Síndrome de Sjögren, etc. (28)

Se han descrito diferentes programas diseñados para automatizar la clasificación de estructuras y la detección de patologías y tratamientos presentes en la cavidad oral del paciente. Entre ellos merece la pena destacar Resnet-101 (27,28), GoogLeNet Inception v3 (29,30) y Denti.Ai. (31).

ResNet-101 es un programa basado en inteligencia artificial que consta de 152 capas cuyo método de aprendizaje varía del considerado como “normal” en estos casos. Se trata de un sistema que se nutre a partir de representaciones residuales y que introduce la denominada conexión de acceso directo (ajustar la unión de una capa con la siguiente sin recurrir a ninguna modificación). Esto permite crear una red neuronal mucho más extensa y profunda.

GoogLeNet Inception v3 es un sistema neuronal que se dio a conocer en el año 2014 por su gran papel en el *ImageNet Large Scale Visual Recognition Challenge*. Se trata de una red convolucional formada por 22 capas que ya ha sido entrenada con alrededor de 1.28 millones de imágenes, sabiendo reconocer hasta 1000 objetos presentes en ellas. (29)

Denti.Ai es un servicio automatizado basado en IA fundado en 2017 por *Dmitry Tuzoff* que genera un diagnóstico automático de imágenes radiológicas por medio de tecnologías de Deep Learning (DL). Este programa promete ser de gran utilidad gracias a su versatilidad de análisis: es compatible tanto con OPG como con radiografías intraorales (periapicales o aletas de mordida).

Se trata de un sistema en el que, a través de una serie de algoritmos matemáticos, **se consigue la extracción y el aprendizaje del cerebro artificial de determinados eventos de interés para mejorar las interpretaciones de imágenes posteriores.** (31)

Para llevar a cabo la tarea de detectar y clasificar las estructuras dentarias encontradas en las OPG, este programa emplea diferentes módulos de detección: en primer lugar, se **procesa la imagen para establecer los límites de cada diente** presente en la misma para, a continuación, **marcar con recuadros delimitadores las estructuras encontradas.**

A continuación, el módulo de numeración de dientes clasifica cada región dentaria delimitada según la nomenclatura dentaria FDI (*Federation Dentaire Internationale*). De esta forma, el sistema genera unas coordenadas en los cuadros delimitadores y los asocia a números de la dentición correspondientes a todas las estructuras detectadas en la imagen. (31)

Denti.Ai utiliza la última generación del modelo **Faster R-CNN (Region based Convolutional Neural Network)** para la **detección radiográfica de piezas dentarias**. Específicamente, este modelo consta de dos sistemas: (31)

- **Sistema de detección de regiones (RPN o Regional Proposal Network):** su principal función es proponer regiones en las que pueden encontrarse los objetos de interés, en este caso las estructuras dentarias.
- **Sistema de detección de objetos (Object Detector):** trata de usar las regiones propuestas por el sistema anterior para ahondar en la localización y en la clasificación de los objetos de interés (estructuras dentarias).

Sin embargo, la numeración de las estructuras dentarias identificadas por el Programa Denti.Ai se realiza a través de la arquitectura informática **VGG-16**. Esta permite clasificar los recuadros de delimitación mostrados gracias al modelo **Faster R-CNN** para asignarlos a un número de acuerdo con la **numeración de la FDI**. (31)

Toda la información generada es posteriormente procesada y revisada mediante un método heurístico para mejorar los resultados de predicción alcanzados por el sistema inteligente. Este proceso se basa en la suposición de que cada diente puede aparecer como máximo una vez en la imagen (y en el orden específico de la dentición). (31)

De forma resumida, pues, el algoritmo que compone el programa Denti.Ai funciona de la siguiente manera:

1. **Delimitación de todas las estructuras dentarias identificadas en cada arcada**
2. **Asignación de una numeración FDI a cada pieza identificada**
3. **Repetición del proceso para el cálculo de la mayor puntuación de confianza** (extraída de todas las posibles combinaciones válidas en la asignación de dientes y numeraciones)
4. **Elección de la combinación con la mayor puntuación de confianza.**

En lo que se refiere al diagnóstico de patologías (como lesiones cariosas, lesiones periapicales, etc.) y la detección y clasificación de tratamientos dentarios (como obturaciones de composite o de amalgama, tratamientos de conductos, prótesis fijas, implantes, etc.), el programa se rige por una **codificación basada en una escala de grises**, a la vez que lo mencionado anteriormente para la **delimitación de la estructura patológica/terapéutica identificada**.

Resumidamente, su compleja red neuronal permite examinar la imagen radiográfica en un tiempo máximo de 4 segundos, obteniendo como resultado una gráfica en la que se muestran todos los hallazgos encontrados en la imagen (marcados también en ella). El odontólogo puede auditar y verificar el diagnóstico creado por el programa, pudiendo modificarlo en todo momento de acuerdo con su experiencia y opinión profesional permitiendo, además, “nutrir” la red neuronal con la corrección del fallo que haya podido tener el sistema inteligente.

Además, en caso de duda, Denti.Ai permite al usuario tener acceso a una segunda opinión generada en base a la experiencia de radiólogos expertos que participaron en el proceso de entrenamiento del software.

Justificación.

La introducción de la reciente tecnología en el ámbito odontológico ha supuesto una completa revolución en la rutina diaria de los clínicos dedicados a la profesión.

Es indudable que, aunque existen numerosas ventajas relacionadas con la incorporación de flujo digital en las clínicas dentales, la rapidez y la comodidad de trabajo son las más destacables. Este nuevo concepto de odontología ha permitido acortar el número de citas necesarias para terminar un determinado tratamiento, así como el desarrollo de otras nuevas fuentes de ayuda y soporte al profesional.

En los últimos años, la Inteligencia Artificial ha evolucionado en varios campos dentro del área de la salud. En odontología, la creación de programas de soporte diagnóstico ha supuesto un punto de inicio en una trayectoria de investigación que, indudablemente, seguirá ampliándose de manera muy importante durante los próximos años.

Actualmente, la forma de diagnosticar las posibles patologías presentes en la cavidad oral de los pacientes consiste en un análisis clínico y radiológico. En el presente estudio se utilizará el programa basado en IA denominado Denti.Ai para cumplir con el propósito que tradicionalmente ha sido llevado a cabo por el clínico.

Este cerebro artificial permite eliminar ciertos factores a los que todos los clínicos están expuestos: fatiga ocular, cansancio e incluso falta de experiencia diagnóstica (siendo estos motivos una razón por los que se puede pasar por alto la detección de patologías normalmente incipientes).

Además, también se ha de tener en consideración la inequidad de diagnóstico y toma de decisiones terapéuticas de los odontólogos: puede ser que algunos tiendan a infradiagnosticar y a proceder con tratamientos menos invasivos en determinados casos y otros, sin embargo, todo lo contrario.

Al tratarse de un programa de IA de reciente introducción, apenas se dispone de literatura científica que describa las ventajas e inconvenientes que puede mostrar. Así pues, es necesaria la realización de estudios de esta índole para su validación clínica.

Por este motivo y por los mencionados anteriormente, se justifica la realización de este trabajo de investigación, cuyo fin es evaluar la capacidad diagnóstica del programa Denti.Ai en la detección de patologías y tratamientos dentarios presentes en la cavidad oral del paciente.

Objetivos

El **objetivo principal** de este estudio es **evaluar la fiabilidad diagnóstica del programa Denti.Ai en la detección de estructuras y tratamientos dentarios presentes en radiografías panorámicas.**

En relación con este objetivo general, los **objetivos específicos** son:

1. **Evaluar la capacidad de detección de estructuras dentarias** presentes en radiografías panorámicas del programa Denti.Ai.
2. **Evaluar la capacidad de detección de restauraciones** presentes en radiografías panorámicas (en concreto, restauraciones plásticas y metálicas) presentes en radiografías panorámicas del programa Denti.Ai.
3. **Evaluar la capacidad de diferenciación entre restauraciones metálicas y plásticas** presentes en radiografías panorámicas del programa Denti.Ai.
4. **Evaluar la capacidad de detección de tratamientos de conductos** presentes en radiografías panorámicas del programa Denti.Ai.
5. **Evaluar la capacidad de detección de coronas dentarias, pónicos y coronas sobre implantes** presentes en radiografías panorámicas del programa Denti.Ai.
6. **Evaluar la capacidad de detección de implantes dentales** presentes en radiografías panorámicas del programa Denti.Ai.

Hipótesis

Como se ha mencionado anteriormente, los programas de aplicación de inteligencia artificial en odontología se encuentran todavía en claro proceso de desarrollo. Actualmente se está poniendo a prueba su uso estudiando su capacidad de análisis en las imágenes bidimensionales utilizadas en Odontología, aunque también se están empleado de muchas otras formas y en muchos otros campos.

En este estudio se pretende evaluar la capacidad de detección de estructuras y tratamientos dentarios del programa Denti.Ai en radiografías panorámicas. Todo ello tiene por objetivo aumentar la sensibilidad y la especificidad del programa en su detección de patologías y tratamientos presentes en la cavidad oral del paciente.

Por tanto, en nuestro estudio se formulan las siguientes **hipótesis nulas** de aplicación estadística:

1. **No existen diferencias estadísticamente significativas en la detección de estructuras dentarias** entre lo diagnosticado por el programa y lo diagnosticado por la evaluación clínica.
2. **No existen diferencias estadísticamente significativas en la detección de restauraciones dentarias** entre lo diagnosticado por el programa y lo diagnosticado por la evaluación clínica.
3. **No existen diferencias estadísticamente significativas en la clasificación de restauraciones (metálicas o plásticas)** entre lo diagnosticado por el programa y lo diagnosticado por la evaluación clínica.
4. **No existen diferencias estadísticamente significativas en la detección de tratamientos de conductos** entre lo diagnosticado por el programa y lo diagnosticado por la evaluación clínica.
5. **No existen diferencias estadísticamente significativas en la detección de coronas de recubrimiento total** entre lo diagnosticado por el programa y lo diagnosticado por la evaluación clínica.
6. **No existen diferencias estadísticamente significativas en la detección de implantes osteointegrados** entre lo diagnosticado por el programa y lo diagnosticado por la evaluación clínica.
7. **No existen diferencias estadísticamente significativas en la detección de implantes osteointegrados** entre lo diagnosticado por el programa y lo diagnosticado por la evaluación clínica.

Material y Método

Selección de muestras

Todas las pruebas de imagen utilizadas fueron realizadas con el equipo **Kodak 9000 3D Extraoral Imaging System** (Carestream Health, Inc.), siendo exportadas con un **tamaño de imagen de 2706 x 1536 píxeles** y una **resolución de 323 x 323 píxeles**.

Para la consecución de este estudio se seleccionaron **300 radiografías panorámicas en dentición permanente** de forma **completamente aleatoria**, sin tener en cuenta sexo ni raza de los pacientes. Todas ellas fueron exportadas en formato *.jpeg*, analizadas por dos operadores y posteriormente añadidas al programa Denti.Ai. **Fig. 2.**



Figura 2. Esquema del proceso de selección de las muestras utilizadas en el estudio.

Para la selección de muestras del estudio, se siguieron los siguientes criterios de inclusión y de exclusión:

- **Criterios de inclusión.**
 - **Pacientes mayores de edad**
 - **Pacientes de ambos sexos**
 - **Pacientes de cualquier nacionalidad**
 - **Radiografías panorámicas**
 - **Pacientes con dentición permanente**
 - **Paciente con dentición permanente con ausencias dentarias**
 - **Paciente con dentición permanente sin ausencias dentarias**

- **Criterios de exclusión.**
 - **Pacientes con dentición mixta**
 - **Pacientes con dentición temporal**

Análisis de muestras

Una vez seleccionadas todas las imágenes para la consecución de este estudio, se realizó un análisis exhaustivo de cada una de ellas por dos operadores con experiencia diagnóstica de forma independiente.

Los operadores fueron entrenados y calibrados previamente al análisis de las radiografías. Para medir la concordancia entre las valoraciones cuantitativas obtenidas por los distintos operadores, se calculó el **Índice de Correlación Interclase**, obteniendo un resultado de entre **0,993 y 1,000** (indicando niveles de concordancia muy buenos). El **consenso diagnóstico establecido entre los dos operadores** conformó la **base** sobre la cual se llevó a cabo la comparación de resultados **obtenidos** por el programa **Denti.Ai**.

En caso de que no existiera un acuerdo entre los dos operadores, un tercero actuó evaluando el caso y logrando el consenso entre los tres evaluadores. De no lograrse, ese caso fue descartado. Por lo tanto, en función del diagnóstico acordado, encontramos las siguientes situaciones:

- **El diagnóstico base y el sugerido por Denti.Ai son el mismo:** Se consideran correctas las dos fuentes.
- **El diagnóstico base difiere del sugerido por Denti.Ai:** en estos casos se necesitará de un tercer operador que realice un análisis de los casos para determinar el resultado real, además de identificar las posibles vías por las que se ha podido llegar a la malinterpretación (por parte del operador o del programa).

Diseño del estudio

Este proyecto de investigación es un **estudio observacional** (los investigadores se limitan a medir y a observar lo detectado por el programa Denti.Ai) y **transversal** (los datos obtenidos representan un punto aislado en el tiempo).

Al tratarse de un sistema en el que **se anonimizan de forma automática todas las imágenes radiológicas** aportadas por el clínico, no fue necesaria la firma del consentimiento informado por parte de los pacientes, aunque sí fue requerida la aprobación del Comité Ético de Investigaciones del Hospital Clínico San Carlos para la consecución del estudio (Código de aceptación N° 21/297E). **Anexo 1.**

Para poder llevar a cabo esta investigación, se dividió la metodología de trabajo en dos partes diferenciadas: primeramente, se llevó a cabo el análisis de las 300 radiografías por ambos operadores de forma individual. En cada una de las pruebas radiológicas se asignaron, de forma manual, las siguientes categorías según los hallazgos encontrados en ellas:

- **Estructuras dentarias:** El operador marcó en **nomenclatura FDI** tanto las piezas presentes como aquellas ausentes en la imagen panorámica.
- **Restauraciones dentarias:** El operador marcó con las siglas **RM** (Restauración Metálica) o **RP** (Restauración Plástica).
- **Tratamientos de conductos:** El operador marcó con las siglas **TC**.
- **Coronas dentarias y sobre implantes:** El operador marcó con las siglas **C**.
- **Implantes osteointegrados:** El operador marcó con las siglas **IOI**.

Una vez recopilados todos los datos almacenados por los mismos, se procedió al **estudio de la concordancia entre operadores** por medio de una revisión general de los resultados obtenidos por cada uno de los mismos. De esta manera, se obtuvieron los resultados finales modificando aquellos casos en los que los operadores discernían. Esta información fue considerada la base sobre la cual se comparó posteriormente los resultados obtenidos por Denti.Ai.

En segundo lugar, se procedió al análisis de las radiografías por el programa basado en inteligencia artificial. Para ello, se añadieron todas las radiografías introduciendo únicamente el número de muestra que había sido asignada por los operadores anteriormente, en el campo “Patient’s ID”. **Fig 3.**



Figura 3. Proceso de anonimización de las imágenes en Denti.Ai.

Se categorizaron las radiografías en las siguientes carpetas con el fin de mejorar la accesibilidad de las radiografías en caso de que tuvieran que ser revisadas:

- **Carpeta 1:** OPGs 1-25.
- **Carpeta 2:** OPGs 26-50.
- **Carpeta 3:** OPGs 51-75.
- **Carpeta 4:** OPGs 76-100.
- **Carpeta 5:** OPGs 101-125.
- **Carpeta 6:** OPGs 126-150.
- **Carpeta 7:** OPGs 151-175.
- **Carpeta 8:** OPGs 176-200.
- **Carpeta 9:** OPGs 201-225.
- **Carpeta 10:** OPGs 226-250.
- **Carpeta 11:** OPGs 251-275.
- **Carpeta 12:** OPGs 276-300.

Aunque el programa diera la posibilidad de modificar las características de brillo y contraste de las imágenes, **en este estudio no se realizó ninguna modificación con el fin de asegurar cierta consistencia e igualdad de términos entre la evaluación de los operadores y el del cerebro artificial (Denti.Ai).**

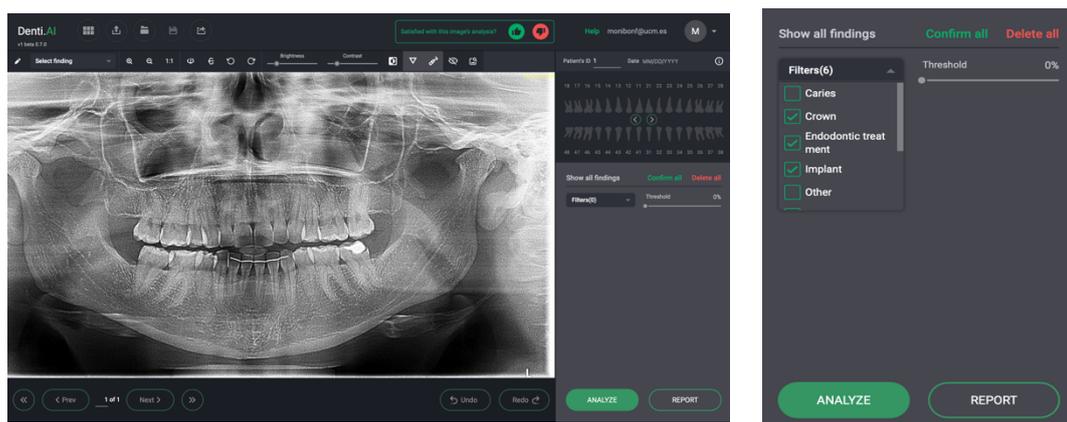


Figura 4. Proceso de selección de hallazgos a diagnosticar en el programa Denti.Ai.

Una vez cargada la radiografía, Denti.Ai permite marcar el tipo de tratamientos que se desea que sean diagnosticados. En el caso de este estudio, **se marcaron las siguientes opciones** (por orden de aparición en el programa): **crown, endodontic treatment, implant, pontic y filling.** **Fig.4.**

Evaluación de la fiabilidad del programa basado en Inteligencia Artificial Denti.Ai para la identificación de estructuras y tratamientos dentarios presentes en radiografías panorámicas.

Además, se ajustó el umbral de detección del programa al 0% para poder evaluar todo aquello que era detectado por el mismo (independientemente del porcentaje de fiabilidad que se obtuviese). Una vez realizado todo lo anterior, se procedió al análisis de la imagen.

De forma automática, el programa detectó los dientes tanto presentes como ausentes en boca (adjudicándoles su correspondiente numeración de acuerdo con el sistema FDI) así como la posible presencia de caries o de pérdida de hueso alveolar. Denti.Ai ofrece un esquema simplificado y codificado por colores en el que distinguimos: **Fig. 5.**

- **Colores verdes:** dientes sanos, presentes en boca.
- **Colores amarillos:** dientes sanos, con pérdida de hueso alveolar.
- **Colores rojos:** dientes con patología cariosa.
- **Colores grises:** dientes ausentes en boca.



Figura 5. Sistema de codificación de colores utilizado por Denti.Ai y esquema de detección de estructuras dentarias.

Los tratamientos encontrados serán aportados por el programa en la parte inferior derecha, indicando la pieza dentaria de la que se trata, el tipo de hallazgo encontrado y, en caso de las obturaciones, una etiqueta que especifica si se ha considerado de composite o de amalgama. **Fig. 6.**

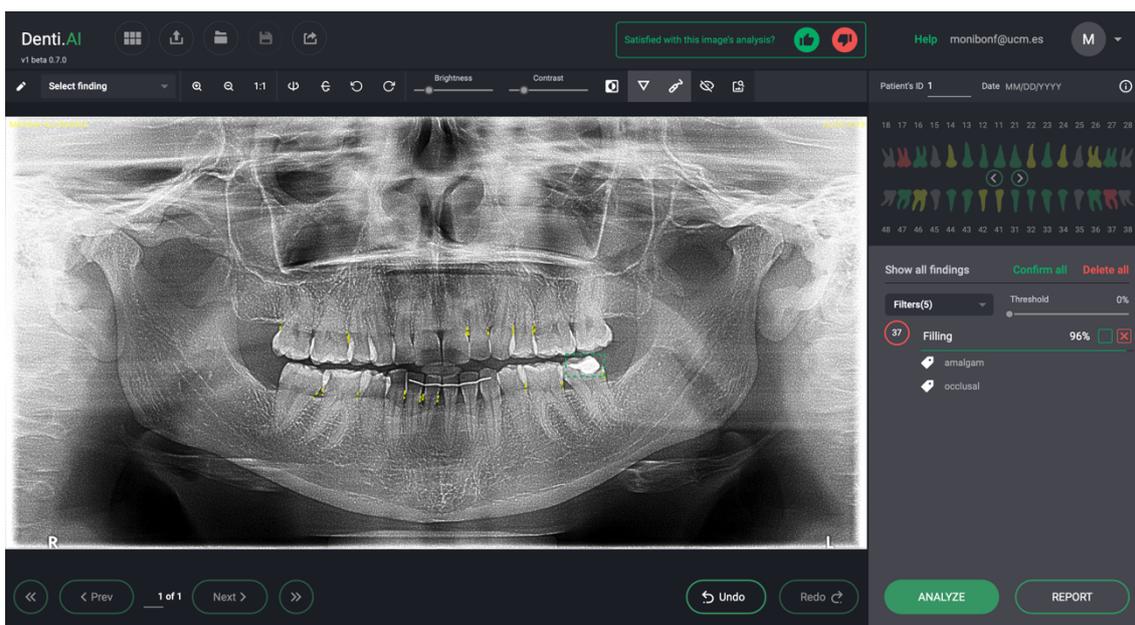


Figura 6. Hallazgos encontrados por Denti.ai tras el análisis de la radiografía.

Finalmente, se procedió a la recogida y almacenamiento de los datos obtenidos en hojas de cálculo del programa Microsoft Excel.

Recogida de datos

La recogida de datos analizados por los operadores se realizó en dos pasos: una primera vez de forma manual y una segunda de forma digital a través de una hoja de cálculo del programa Microsoft Excel. Para la evaluación manual de cada uno de los operadores, se llevó a cabo una plantilla que facilitara la agrupación de la mayor cantidad de información posible por radiografía, por hoja. **Anexo 2.**

Sin embargo, los datos obtenidos por el programa fueron recopilados directamente por el investigador principal de forma digital (también con el programa Microsoft Excel).

En la hoja de cálculo creada se especificaron: el **número de piezas dentarias** detectadas, el **número de restauraciones totales** detectadas, el **número de restauraciones plásticas y metálicas** detectadas, el **número de tratamientos de conductos** detectados, el **número total de coronas** (sobre diente natural, sobre implante y pónicos) detectados y el **número de implantes** detectados; por cada operador y por Denti.Ai, respectivamente.

Además, a la hora de la detección de las estructuras dentarias presentes en boca por parte de Denti.Ai, se observaron algunos errores que fueron clasificados en los siguientes valores cuantitativos:

- **0:** Ausencia de error en la descripción de piezas dentarias presentes en la OPG.
- **1:** Denti.Ai confunde diente natural con corona sobre implante.
- **2:** Denti.Ai confunde diente natural con pónico.
- **3:** Denti.Ai confunde piezas dentarias (asigna nomenclaturas FDI incorrectas).
- **4:** Denti.Ai no detecta supernumerarios o dientes incluidos.
- **5:** Denti.Ai no detecta piezas presentes en boca.

Análisis estadístico

En este apartado se evaluó la precisión con la que el programa identificó las variables de estudio en las imágenes diagnósticas. Para ello fue necesario calcular los valores correspondientes a Verdadero Positivo (VP), Falso Positivo (FP), Verdadero Negativo (VN) y Falso Negativo (FN):

- **Verdadero positivo:** Diagnóstico positivo (pieza dentaria, implantes o diversos tipos de restauración presentes).
- **Falso positivo:** Diagnóstico positivo, pero pieza dentaria, implantes o diversos tipos de restauración ausentes.
- **Verdadero negativo:** Diagnóstico negativo (pieza dentaria, implantes o diversos tipos de restauración ausentes).

Evaluación de la fiabilidad del programa basado en Inteligencia Artificial Denti.Ai para la identificación de estructuras y tratamientos dentarios presentes en radiografías panorámicas.

- **Falso negativo:** Diagnóstico negativo, pero pieza dentaria, implantes o diversos tipos de restauración presentes.

Con los datos anteriormente recolectados, se llevará a cabo el **cálculo de las siguientes pruebas diagnósticas:**

- **Sensibilidad (S):** Capacidad de Denti.Ai de detectar la presencia de estructuras y de restauraciones dentarias.

$$S = \frac{VP}{VP + FN}$$

- **Especificidad (E):** Probabilidad de que, no habiendo estructuras o restauraciones dentarias, el resultado obtenido por el programa sea negativo.

$$E = \frac{VN}{VN + FP}$$

- **Valor predictivo para positivos (VPP):** Capacidad del programa de **detectar los verdaderos positivos** entre el total de imágenes radiográficas con presencia de piezas y restauraciones dentarias. También puede entenderse como la probabilidad de que, obteniendo un resultado positivo por parte del programa, la presencia de piezas y restauraciones dentarias en la imagen radiográfica sean realmente ciertas.

$$VPP = \frac{VP}{VP + FP}$$

- **Valor predictivo para negativos (VPN):** Capacidad del programa de **detectar los verdaderos negativos** entre el total de imágenes radiográficas con ausencia de restauraciones o piezas dentarias. Probabilidad de que, obteniendo un resultado negativo por parte del programa, la ausencia de piezas o restauraciones dentarias sean realmente ciertas.

$$VPN = \frac{VN}{VN + FN}$$

- **Curvas ROC:** Se representarán la **sensibilidad (S)** y la **especificad (E)** de la prueba diagnóstica (Denti.Ai), para todos los valores de los puntos de corte, dibujándose, de esta forma, la curva de ROC (que **une pares de valores**). El área encontrada bajo la curva **permitirá medir la probabilidad de clasificar correctamente dos sujetos del estudio** (VP y VN). De esta forma, teniendo en cuenta el área podremos decir que:

- La curva de una prueba con **validez diagnóstica perfecta**, tendrá un **área bajo la curva de 1 (S y E = 1)**.
- La curva de una **prueba sin validez diagnóstica**, tendrá un **área bajo la curva de 0,5 (S y E = 0,5)**.

Resultados

Análisis descriptivo de las variables

Detección de piezas dentarias

La media de piezas dentarias detectadas por Denti.Ai fue de 28,12 dientes por radiografía panorámica, cifra muy cercana a lo establecido como gold estándar de acuerdo a los resultados obtenidos por parte de los operadores (28,11). En la realización de las pruebas de normalidad (*Kolmogorov-Smirnov*) pudimos observar que los **resultados no cumplían criterios de normalidad ($p < 0.05$)**. Esto puede estar relacionado con el hecho de que **las radiografías panorámicas escogidas para la consecución de este estudio no eran iguales ni contenían la misma información** (había número variable de piezas dentarias en cada una, con tratamientos diversos, etc).

En el análisis descriptivo del tipo de errores cometidos por el programa Denti.Ai en referencia a las piezas dentarias detectadas en las radiografías panorámicas, se encontraron los siguientes resultados: en el 50% de los casos, el programa supo detectar y clasificar las piezas dentarias sin ningún problema. Sin embargo, en el 20,7% de los casos de encontró que el Denti.Ai. cometía errores en la clasificación de piezas dentarias existentes en la radiografía panorámica (es decir, daba una nomenclatura FDI incorrecta a las piezas detectadas). El segundo error más frecuente se relacionó a la confusión del programa entre pónicos y dientes naturales (asignaba pónicos como piezas naturales presentes en boca, cuando evidentemente eran inexistentes). De igual manera, con una probabilidad del 8,8% el programa detectó coronas sobre implantes como dientes naturales. Por último, en un 7,3% de los casos el programa no supo reconocer todas las piezas presentes en la radiografía o la presencia de dientes supernumerarios (1,8%). (*Tabla 8, Figuras 7 y 8*).

		Respuestas		Porcentaje de casos
		N	Porcentaje	
TIPO ERROR	ET0	164	50,0%	54,7%
	ET1	29	8,8%	9,7%
	ET2	37	11,3%	12,3%
	ET3	68	20,7%	22,7%
	ET4	6	1,8%	2,0%
	ET5	24	7,3%	8,0%
Total		328	100,0%	109,3%

Tabla 8. Relación de proporciones de los tipos de errores cometidos por el programa Denti.Ai. en la detección y clasificación de piezas dentarias.

Evaluación de la fiabilidad del programa basado en Inteligencia Artificial Denti.Ai para la identificación de estructuras y tratamientos dentarios presentes en radiografías panorámicas.

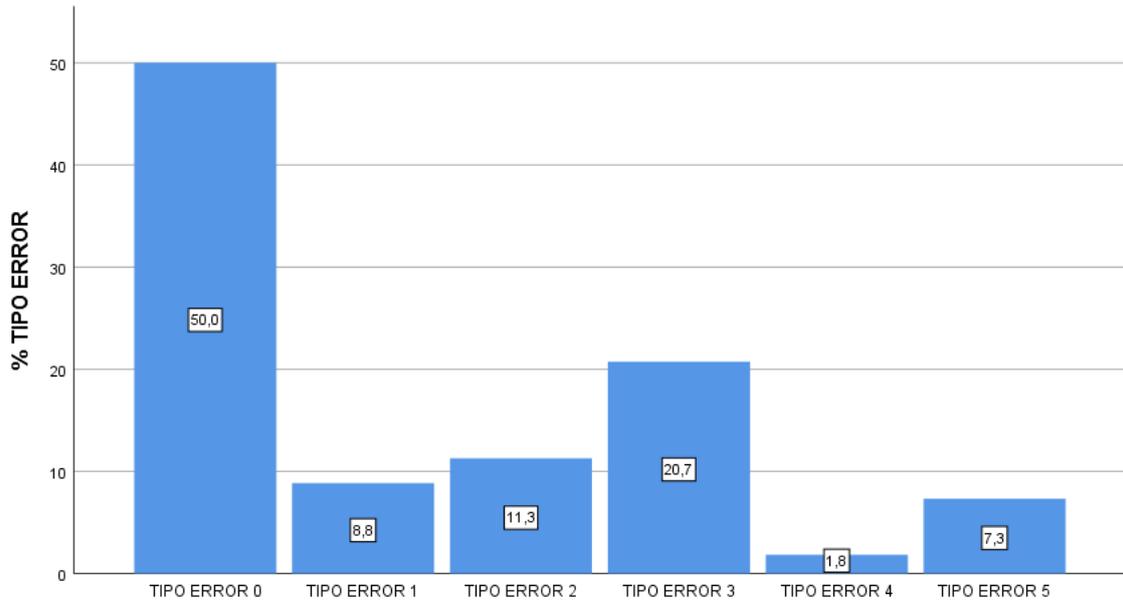


Figura 7. Representación gráfica de los tipos de errores cometidos por el programa Denti.Ai en la detección y clasificación de piezas dentarias.

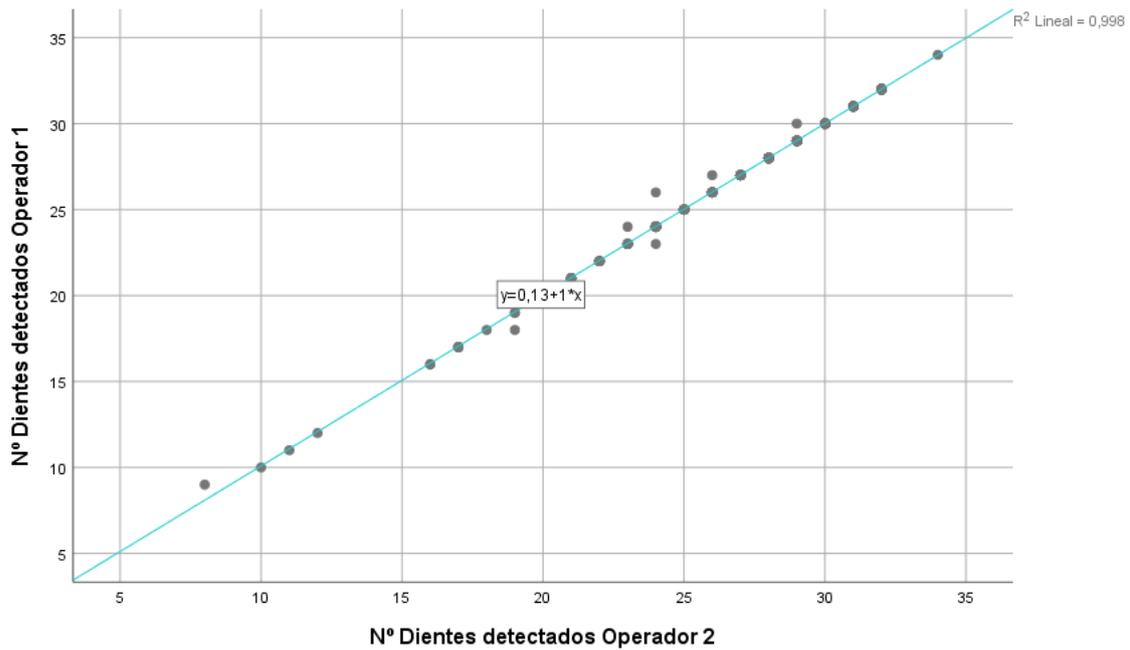


Figura 8. Análisis exploratorio de la concordancia entre operadores para el número total de dientes detectados (siendo Operador 1 Denti.Ai y Operador 2 el gold estándar)

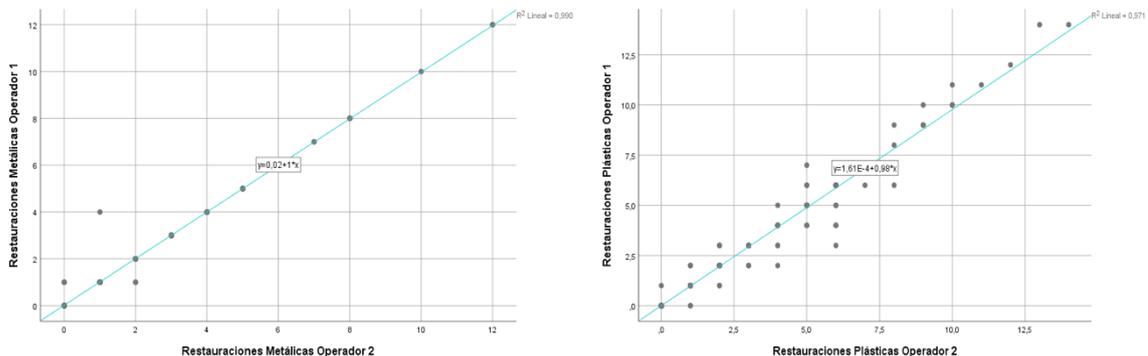
La gráfica representada en la Figura 8 muestra el análisis de concordancia entre los resultados obtenidos por Denti.Ai y el gold estándar (siendo este el consenso entre operadores) por medio de la utilización de la recta de regresión (no existiendo variables independientes y dependientes). Como se puede observar, en aquellos casos en los que las radiografías presentaban mayor número de piezas dentarias, mayor fue la probabilidad de error del sistema.

Detección de obturaciones

La media de obturaciones totales detectadas por Denti.Ai fue de 3,42 dientes por radiografía panorámica, ligeramente inferior a lo establecido como gold estándar de acuerdo a los resultados obtenidos por parte de los operadores (3,45). **Figuras 9, 10.** En la realización de las pruebas de normalidad (*Kolmogorov-Smirnov*) pudimos observar que **los resultados no cumplían criterios de normalidad ($p < 0.05$)**. Esto puede estar relacionado con el hecho de que **las radiografías panorámicas escogidas para la consecución de este estudio no eran iguales ni contenían la misma información** (había número variable de obturaciones en cada una de ellas, tanto plásticas como metálicas).

La gráfica representada en la Figura 9 muestra el análisis de concordancia entre los resultados obtenidos por Denti.Ai y el gold estándar (siendo este el consenso entre operadores) por medio de la utilización de la recta de regresión (no existiendo variables independientes y dependientes). Como se puede observar, en aquellos casos en los que las radiografías presentaban mayor número de piezas dentarias, mayor fue la probabilidad de error del sistema.

- **Detección de obturaciones plásticas:** De acuerdo a la estadística descriptiva realizada, se pudo observar que el programa Denti.Ai tuvo cierta dificultad en la detección de obturaciones plásticas debido a la presencia, muchas veces, de radio densidades parecidas a las del tejido dentario por la utilización de composites fluidos tipo Bulk-Fill. Se estimó una media detectada por el programa de 2,16 frente a 2,21 (establecido como gold estándar de acuerdo a los resultados obtenidos por parte de los operadores).
- **Detección de obturaciones metálicas:** De acuerdo a la estadística descriptiva realizada, se pudo observar que el programa Denti.Ai tuvo resultados muy favorables en la detección de obturaciones metálicas (1,26 frente a 1,25 establecido como gold estándar de acuerdo a los resultados obtenidos por parte de los operadores).



Figuras 9, 10. Análisis exploratorio de la concordancia entre operadores para las variables restauraciones metálicas y plásticas (siendo Operador 1 Denti.Ai y Operador 2 el gold estándar)

La gráfica representada en la Figura 9 muestra el análisis de concordancia entre los resultados obtenidos por Denti.Ai y el gold estándar (siendo este el consenso entre operadores) por medio de la utilización de la recta de regresión (no existiendo variables independientes y dependientes).

Como se puede observar, en aquellos casos en los que las radiografías presentaban obturaciones metálicas, la capacidad diagnóstica fue prácticamente perfecta. Sin embargo, existió mayor discordancia en la detección de restauraciones plásticas, lo cual puede hacer poco fiable la detección de las mismas.

Detección de tratamientos de conductos

La media de tratamientos de conductos detectados por Denti.Ai fue exactamente igual que lo establecido como gold estándar de acuerdo a los resultados obtenidos por parte de los operadores (0,88), por lo que encontramos una elevada capacidad diagnóstica en este caso.

No obstante, en la realización de las pruebas de normalidad (*Kolmogorov-Smirnov*) pudimos observar que los **resultados no cumplían criterios de normalidad ($p < 0.05$)**. Esto puede estar relacionado con el hecho de que **las radiografías panorámicas escogidas para la consecución de este estudio no eran iguales ni contenían la misma información** (había número variable de tratamientos de conductos en cada una de ellas: algunas no tenían y otras, sin embargo, tenían uno o varios).

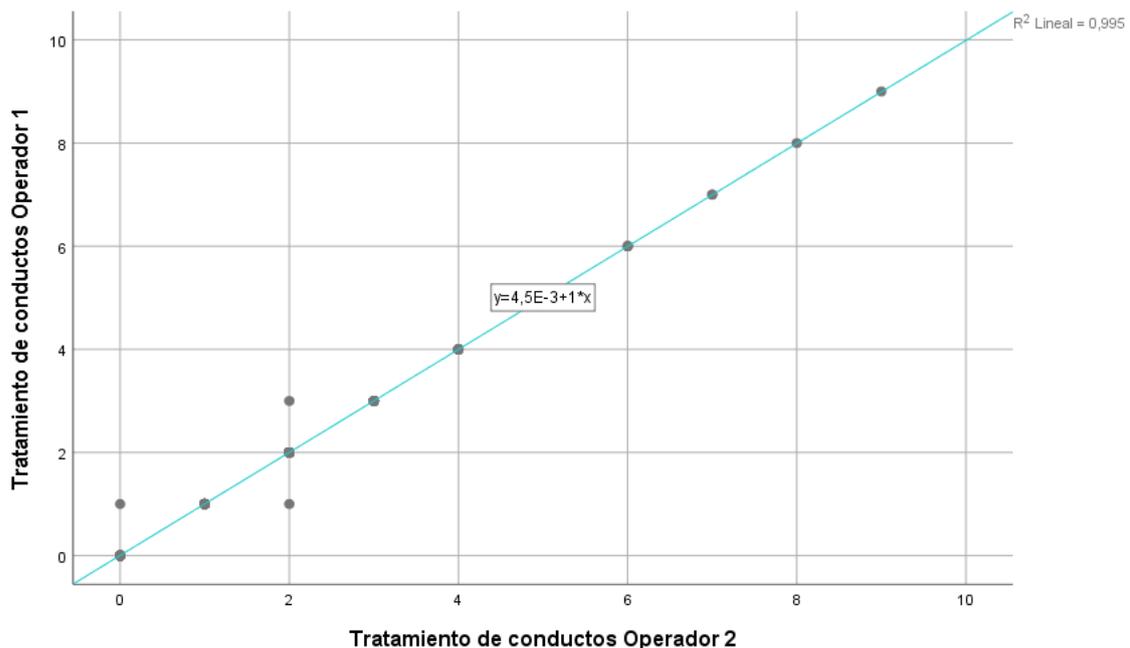


Figura 11. Análisis exploratorio de la concordancia entre operadores para la variable tratamiento de conductos (siendo Operador 1 Denti.Ai y Operador 2 el gold estándar)

Como podemos observar en la gráfica representada en la *Figura 11*, los resultados obtenidos por Denti.Ai en comparación al gold estándar mostraron una capacidad diagnóstica prácticamente perfecta en la detección de tratamientos endodónticos por parte del programa.

Detección de coronas (sobre diente natural, implantes y pónicos)

La media de coronas (sobre diente natural, implantes y pónicos) detectadas por Denti.Ai fue exactamente igual que lo establecido como gold estándar de acuerdo a los resultados obtenidos por parte de los operadores (1,32), por lo que también encontramos una elevada capacidad diagnóstica en este caso.

No obstante, en la realización de las pruebas de normalidad (*Kolmogorov-Smirnov*) pudimos observar que los **resultados no cumplían criterios de normalidad ($p < 0.05$)**. Esto puede estar relacionado con el hecho de que **las radiografías panorámicas escogidas para la consecución de este estudio no eran iguales ni contenían la misma información** (había número variable de coronas en cada una de ellas: algunas no tenían y otras, sin embargo, tenían una o varias).

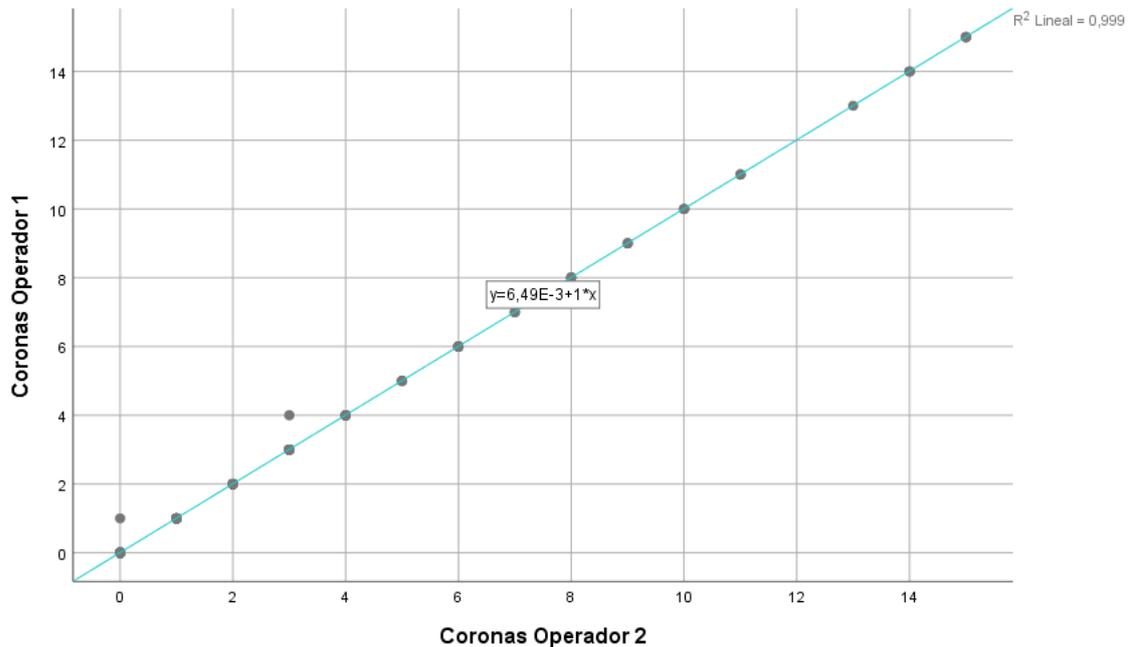


Figura 12. Análisis exploratorio de la concordancia entre operadores para la variable coronas (siendo Operador 1 Denti.Ai y Operador 2 el gold estándar).

Como podemos observar en la gráfica representada en la *Figura 12*, los resultados obtenidos por Denti.Ai en comparación al gold estándar mostraron una capacidad diagnóstica prácticamente perfecta en la detección de coronas por parte del programa.

Detección de implantes osteointegrados

La media de implantes detectados por Denti.Ai fue exactamente igual que lo establecido como gold estándar de acuerdo a los resultados obtenidos por parte de los operadores (0,30), por lo que concluimos una elevada capacidad diagnóstica en este caso.

No obstante, en la realización de las pruebas de normalidad (*Kolmogorov-Smirnov*) pudimos observar que los **resultados no cumplían criterios de normalidad ($p < 0.05$)**. Esto puede estar relacionado con el hecho de que **las radiografías panorámicas escogidas para la consecución de este estudio no eran iguales ni contenían la misma información** (había número variable de tratamientos de conductos en cada una de ellas: algunas no tenían y otras, sin embargo, tenían uno o varios).

Evaluación de la fiabilidad del programa basado en Inteligencia Artificial Denti.Ai para la identificación de estructuras y tratamientos dentarios presentes en radiografías panorámicas.

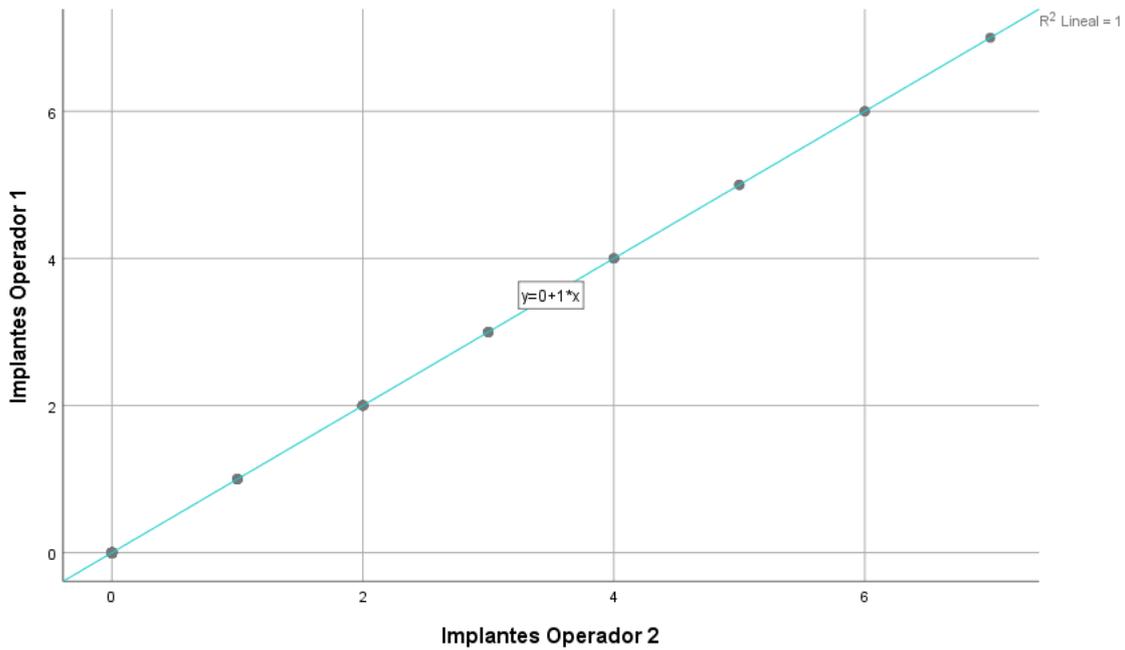


Figura 13. Análisis exploratorio de la concordancia entre operadores para la variable implantes (siendo Operador 1 Denti.Ai y Operador 2 el gold estándar).

Como podemos observar en la gráfica representada en la *Figura 12*, los resultados obtenidos por Denti.Ai en comparación al gold estándar mostraron una capacidad diagnóstica perfecta en la detección de implantes por parte del programa.

Concordancia entre Gold Standard y Denti.Ai

Una vez concluida la recogida de datos de los dos operadores principales, se llevó a cabo un estudio de la concordancia de los mismos por medio del cálculo del **coeficiente de correlación intraclase (ICC)**. Este índice permite medir la consistencia y la concordancia entre medidas cuantitativas obtenidas de diferentes instrumentos de medida o evaluadores.

Debido a que el conjunto de muestras o especímenes son obtenidas de una forma aleatoria y los evaluadores constituyen la totalidad de la población o *gold estándar*, utilizaremos el **modelo de efectos mixtos**. Además, utilizará el **acuerdo absoluto** porque de esta forma se contemplará cualquier diferencia entre medidas como una discordancia (siendo menor la concordancia cuantos mayores sean los casos discrepantes).

Como puede observarse, al realizar el cálculo del ICC obtenemos el valor del mismo para dos casos: **medidas únicas** (grado de concordancia al comparar cada observación de ambos evaluadores) y **medidas promedio** (concordancia promedio en todas las observaciones realizadas por los evaluadores). Por lo general, los **valores son siempre superiores a 0,90**, por lo que se puede concluir que la **concordancia entre operadores y Denti.Ai es prácticamente perfecta**. (*Tablas 1-7*). Además, tanto en las medidas únicas como en las medidas promedio, el **p-valor es significativo ($p < 0,05$)**, por lo que podemos aceptar (**con un 95% de confianza**) la hipótesis de una **perfecta concordancia entre ambos evaluadores**.

Evaluación de la fiabilidad del programa basado en Inteligencia Artificial Denti.Ai para la identificación de estructuras y tratamientos dentarios presentes en radiografías panorámicas.

Coeficiente de correlación intraclase							
	Correlación intraclase ^b	95% de intervalo de confianza		Prueba F con valor verdadero 0			
		Límite inferior	Límite superior	Valor	gl1	gl2	Sig
Medidas únicas	,999 ^a	,999	,999	2052,714	299	299	,000
Medidas promedio	1,000 ^c	,999	1,000	2052,714	299	299	,000

Tabla 1. Cálculo de nivel de concordancia entre operadores (ICC) en el total de número de dientes detectados.

Coeficiente de correlación intraclase							
	Correlación intraclase ^b	95% de intervalo de confianza		Prueba F con valor verdadero 0			
		Límite inferior	Límite superior	Valor	gl1	gl2	Sig
Medidas únicas	,991 ^a	,989	,993	234,324	299	299	,000
Medidas promedio	,996 ^c	,995	,997	234,324	299	299	,000

Tabla 2. Cálculo del nivel de concordancia entre operadores (ICC) en el total de obturaciones detectadas.

Coeficiente de correlación intraclase							
	Correlación intraclase ^b	95% de intervalo de confianza		Prueba F con valor verdadero 0			
		Límite inferior	Límite superior	Valor	gl1	gl2	Sig
Medidas únicas	,985 ^a	,982	,988	135,719	299	299	,000
Medidas promedio	,993 ^c	,991	,994	135,719	299	299	,000

Tabla 3. Cálculo del nivel de concordancia entre operadores (ICC) en el total de obturaciones plásticas detectadas.

Coeficiente de correlación intraclase							
	Correlación intraclase ^b	95% de intervalo de confianza		Prueba F con valor verdadero 0			
		Límite inferior	Límite superior	Valor	gl1	gl2	Sig
Medidas únicas	,995 ^a	,993	,996	382,384	299	299	,000
Medidas promedio	,997 ^c	,997	,998	382,384	299	299	,000

Tabla 4. Cálculo del nivel de concordancia entre operadores (ICC) en el total de obturaciones metálicas detectadas.

Evaluación de la fiabilidad del programa basado en Inteligencia Artificial Denti.Ai para la identificación de estructuras y tratamientos dentarios presentes en radiografías panorámicas.

Coeficiente de correlación intraclase							
	Correlación intraclase ^b	95% de intervalo de confianza		Prueba F con valor verdadero 0			
		Límite inferior	Límite superior	Valor	gl1	gl2	Sig
Medidas únicas	,998 ^a	,997	,998	883,047	299	299	,000
Medidas promedio	,999 ^c	,999	,999	883,047	299	299	,000

Tabla 5. Cálculo del nivel de concordancia entre operadores (ICC) en el total de tratamiento de conductos detectados.

Coeficiente de correlación intraclase							
	Correlación intraclase ^b	95% de intervalo de confianza		Prueba F con valor verdadero 0			
		Límite inferior	Límite superior	Valor	gl1	gl2	Sig
Medidas únicas	1,000 ^a	1,000	1,000	5524,389	299	299	,000
Medidas promedio	1,000 ^c	1,000	1,000	5524,389	299	299	,000

Tabla 6. Cálculo del nivel de concordancia entre operadores (ICC) en el total de coronas (sobre diente o implante) y pónicos detectados.

Coeficiente de correlación intraclase							
	Correlación intraclase ^b	95% de intervalo de confianza		Prueba F con valor verdadero 0			
		Límite inferior	Límite superior	Valor	gl1	gl2	Sig
Medidas únicas	1,000 ^a	.	.	.	299	.	.
Medidas promedio	1,000 ^c	.	.	.	299	.	.

Tabla 7. Cálculo del nivel de concordancia entre operadores (ICC) en el total de implantes detectados.

Sensibilidad, Especificidad, Valor Predictivo Positivo y Valor Predictivo Negativo. Curvas ROC de Prueba Diagnóstica Denti.Ai

Con el fin de determinar la validez de la prueba diagnóstica, se dicotomizaron todos los datos obtenidos en el estudio para poder llevar a cabo las mediciones paramétricas de Sensibilidad (S), Especificidad (E), Valor Predictivo Positivo (VPP) y Valor Predictivo Negativo (VPN). Los resultados obtenidos tanto por el operador como por el programa (Denti.Ai) fueron anotados de forma numérica en una base de datos por lo que, para dicotomizarlos, se emplearon las siguientes categorías: 0 (el operador/el programa sí detectó la variable de estudio) y 1 (el operador/el programa no detectó la variable de estudio).

De esta manera, en el caso de presencia de restauraciones plásticas (por ejemplo), se adjudicó el valor 0; y en caso de ausencia de dichos tratamientos en las radiografías panorámicas, el valor 1. La categorización obtenida por los resultados de los operadores fue considerada como *gold standard*. En caso de discrepancia de resultados entre lo obtenido por los operadores y lo obtenido por Denti.Ai, se describieron los siguientes escenarios:

Evaluación de la fiabilidad del programa basado en Inteligencia Artificial Denti.Ai para la identificación de estructuras y tratamientos dentarios presentes en radiografías panorámicas.

1. Si el operador no observó tratamientos realizados (1), y Denti.Ai sí los detectó (0), se categorizó el resultado obtenido por el programa con un 0. De esta manera, los tratamientos detectados por el programa serían Falsos Positivos.
2. Si el operador observó tratamientos realizados (0) y Denti. Ai no detectó los mismos (1), se categorizó el resultado obtenido por el programa con un 1. De esta manera, los tratamientos no detectados por el programa serían Falsos Negativos.
3. Si el operador observo N tratamientos realizados (0) y Denti.Ai detectó menos tratamientos de los indicados por el *gold standard*, se categorizó el resultado obtenido por el programa con un 1.
4. Si el operador observo N tratamientos realizados (0) y Denti.Ai detectó más tratamientos de los indicados por el *gold standard*, se categorizó el resultado obtenido por el programa con un 0.

Detección de piezas dentarias

El operador detectó presencia de estructuras dentarias en todas las radiografías panorámicas, por lo que todos los casos fueron dicotomizados con el valor 0 en la columna gold standard. Los valores que adquirió el programa para cada caso dependieron de la concordancia entre los resultados:

- Se adjudicó el **valor 0** en casos en los que el **número de piezas dentarias detectadas por ambos operadores** (gold standard y Denti.Ai) **fuera el mismo**.
- Se adjudicó el **valor 1** en casos en los que **el número de piezas dentarias detectadas por ambos operadores** (gold standard y Denti.Ai) **discerniesen**.

De esta manera se obtuvo que **en el 69,3% (S) de los casos, el programa detectó exactamente el mismo número de piezas dentarias que el operador**.

Detección de obturaciones

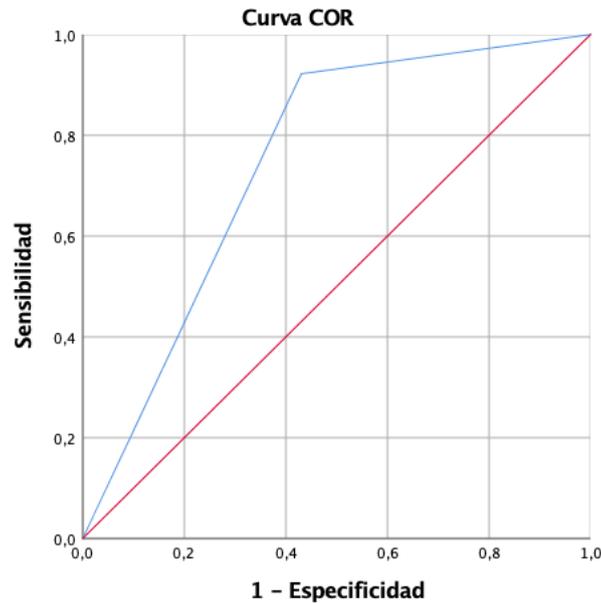
Teniendo en cuenta los criterios mencionados anteriormente, se calcularon los valores de S (56,95%), E (92,2%), VPP (95,48%) y VPN (42,51%) para la variable número total de obturaciones detectadas (tanto plásticas como metálicas). *Tabla 8*.

Tabla cruzada NRestTotOp*NRestTotDentiAi

			NRestTotDentiAi		Total
			0	1	
NRestTotOp	0	Recuento	127	96	223
		% dentro de NRestTotOp	57,0%	43,0%	100,0%
	1	Recuento	6	71	77
		% dentro de NRestTotOp	7,8%	92,2%	100,0%
Total		Recuento	133	167	300
		% dentro de NRestTotOp	44,3%	55,7%	100,0%

Tabla 8. *Tabla de contingencia obtenida por SPSS a partir de la cual se calcularon los valores S, E, VPP, VPN para la variable estudiada.*

Asimismo, se llevó a cabo la representación de los valores obtenidos en Curvas ROC. **Figura 14.** Como se puede observar, el **área bajo la curva se corresponde con 0.746**, por lo que puede concluirse que el diagnóstico del programa Denti.Ai puede ser útil para ciertos propósitos (tiene una exactitud buena, pero no perfecta).



Los segmentos de diagonal se generan mediante empates.

Figura 14. Gráfico de la Curva ROC para la variable estudiada.

Área bajo la curva

Variables de resultado de prueba: RestTotDentiAi

Área	Desv. Error ^a	Significación asintótica ^b	95% de intervalo de confianza asintótico	
			Límite inferior	Límite superior
,746	,029	,000	,688	,803

Las variables de resultado de prueba: RestTotDentiAi tienen, como mínimo, un empate entre el grupo de estado real positivo y el grupo de estado real negativo. Las estadísticas podrían estar sesgadas.

a. Bajo el supuesto no paramétrico

b. Hipótesis nula: área verdadera = 0,5

Tabla 9. Tabla informativa sobre la representación gráfica de la Curva ROC para la variable estudiada.

- **Detección de obturaciones metálicas:** Teniendo en cuenta los criterios de dicotomización de variables mencionados anteriormente, se obtuvieron resultados para los valores paramétricos de S (85,48%), E (87,5%), VPP (82,8) y VPN (89,53%). **Tabla 10.**
- **Detección de obturaciones plásticas:** Teniendo en cuenta los criterios de dicotomización de variables mencionados anteriormente, se obtuvieron resultados para los valores paramétricos de S (41,11%), E (93,3%), VPP (90,24%) y VPN (87,5%%). **Tabla 11.**

Evaluación de la fiabilidad del programa basado en Inteligencia Artificial Denti.Ai para la identificación de estructuras y tratamientos dentarios presentes en radiografías panorámicas.

		RestMetDentiAI		Total	
		0	1		
RestMetOp	0	Recuento	106	18	124
		% dentro de RestMetOp	85,5%	14,5%	100,0%
	1	Recuento	22	154	176
		% dentro de RestMetOp	12,5%	87,5%	100,0%
Total		Recuento	128	172	300
		% dentro de RestMetOp	42,7%	57,3%	100,0%

		RestPLDentiAI		Total	
		0	1		
RestPIOp	0	Recuento	74	106	180
		% dentro de RestPIOp	41,1%	58,9%	100,0%
	1	Recuento	8	112	120
		% dentro de RestPIOp	6,7%	93,3%	100,0%
Total		Recuento	82	218	300
		% dentro de RestPIOp	27,3%	72,7%	100,0%

Tablas 10 y 11. Tablas de contingencia obtenidas por SPSS a partir de la cual se calcularon los valores S, E, VPP, VPN para las variables restauraciones metálicas y plásticas, respectivamente.

A continuación se muestran las curvas ROC obtenidas como resultado de la representación de la S y la E para cada variable. Como se puede observar, la capacidad diagnóstica del programa para la detección de restauraciones metálicas (0,869) es superior que en restauraciones plásticas (0,672). Figuras 15 y 16; Tablas 12 y 13.

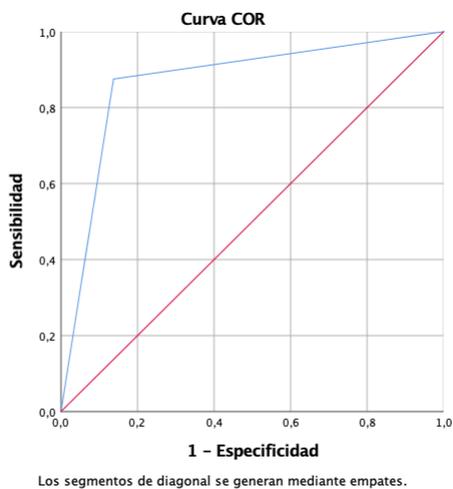


Figura 15. Gráfico de la curva ROC para la variable restauraciones metálicas

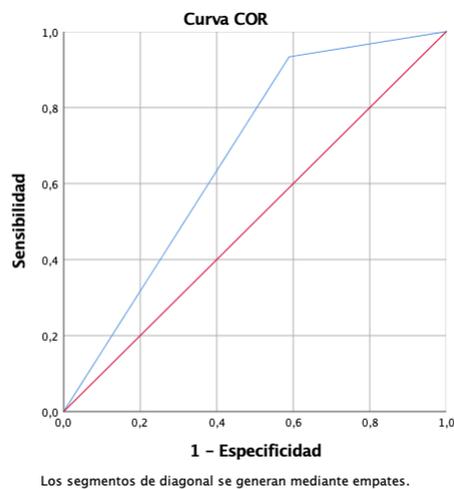


Figura 16. Gráfico de la curva ROC para la variable restauraciones plásticas

Área bajo la curva

Variables de resultado de prueba: RestMetDentiAi

Área	Desv. Error ^a	Significación asintótica ^b	95% de intervalo de confianza asintótico	
			Límite inferior	Límite superior
,869	,023	,000	,824	,914

Las variables de resultado de prueba: RestMetDentiAi tienen, como mínimo, un empate entre el grupo de estado real positivo y el grupo de estado real negativo. Las estadísticas podrían estar sesgadas.

a. Bajo el supuesto no paramétrico

b. Hipótesis nula: área verdadera = 0,5

Tabla 12. Tabla informativa sobre la representación gráfica de la Curva ROC para la variable restauraciones metálicas

Área bajo la curva

Variables de resultado de prueba: RestPIDentiAi

Área	Desv. Error ^a	Significación asintótica ^b	95% de intervalo de confianza asintótico	
			Límite inferior	Límite superior
,672	,031	,000	,612	,732

Las variables de resultado de prueba: RestPIDentiAi tienen, como mínimo, un empate entre el grupo de estado real positivo y el grupo de estado real negativo. Las estadísticas podrían estar sesgadas.

- a. Bajo el supuesto no paramétrico
- b. Hipótesis nula: área verdadera = 0,5

Tabla 13. Tabla informativa sobre la representación gráfica de la Curva ROC para la variable restauraciones plásticas.

Detección de tratamientos de conductos

Teniendo en cuenta los criterios de categorización mencionados anteriormente, se procedió a calcular los valores de S (91,9%), E (100%), VPP (100%) y VPN (94,62%) de la variable estudiada. **Tabla 14.** En este caso, el programa detecta correctamente las piezas que tienen tratamiento de conductos pero, sin embargo, existen algunas de las mismas que lo presentan pero que no es detectado por el mismo.

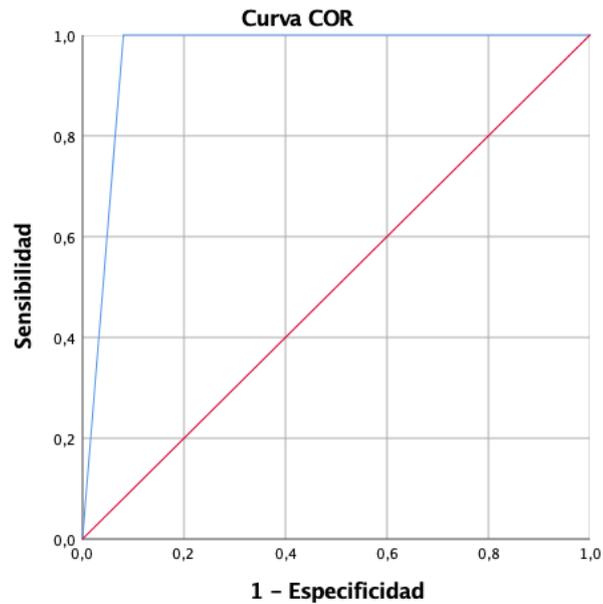
Tabla cruzada TtoCondOp*TtoCondDentiAi

			TtoCondDentiAi		Total
			0	1	
TtoCondOp	0	Recuento	114	10	124
		% dentro de TtoCondOp	91,9%	8,1%	100,0%
	1	Recuento	0	176	176
		% dentro de TtoCondOp	0,0%	100,0%	100,0%
Total		Recuento	114	186	300
		% dentro de TtoCondOp	38,0%	62,0%	100,0%

Tabla 14. Tabla de contingencia obtenidas por el programa SPSS sobre la cual se calcularon los valores de S, E, VPP y VPN para la variable tratamiento de conductos.

No obstante, la capacidad diagnóstica de Denti.Ai para esta variable puede decirse que es elevada, tal y como indica su representación en la curva de ROC (0,960). **Figura 17, Tabla 15.**

Evaluación de la fiabilidad del programa basado en Inteligencia Artificial Denti.Ai para la identificación de estructuras y tratamientos dentarios presentes en radiografías panorámicas.



Los segmentos de diagonal se generan mediante empates.

Figura 17. Gráfico de la Curva ROC para la variable tratamiento de conductos

Área bajo la curva

Variables de resultado de prueba: TtoCondDentiAi

Área	Desv. Error ^a	Significación asintótica ^b	95% de intervalo de confianza asintótico	
			Límite inferior	Límite superior
,960	,014	,000	,931	,988

Las variables de resultado de prueba: TtoCondDentiAi tienen, como mínimo, un empate entre el grupo de estado real positivo y el grupo de estado real negativo. Las estadísticas podrían estar sesgadas.

- a. Bajo el supuesto no paramétrico
- b. Hipótesis nula: área verdadera = 0,5

Tabla 15. Tabla informativa sobre la representación gráfica de la Curva ROC para la variable tratamiento de conductos.

Detección de coronas (sobre diente natural, implantes y pónicos)

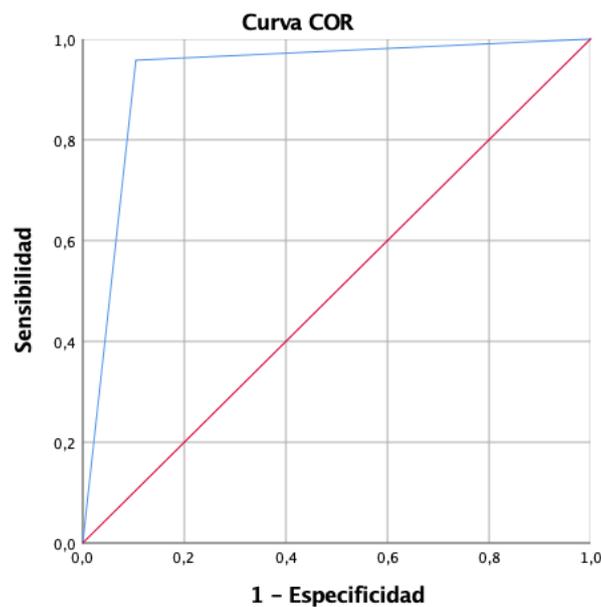
Teniendo en cuenta los criterios de categorización mencionados anteriormente, se procedió a calcular los valores de S (89,53%), E (95,79%), VPP (89,53%) y VPN (95,79%) de la variable estudiada. **Tabla 16.**

Tabla cruzada CorOp*CorDentiAi

		CorDentiAi		Total	
		0	1		
CorOp	0	Recuento	77	9	86
		% dentro de CorOp	89,5%	10,5%	100,0%
	1	Recuento	9	205	214
		% dentro de CorOp	4,2%	95,8%	100,0%
Total		Recuento	86	214	300
		% dentro de CorOp	28,7%	71,3%	100,0%

Tabla 16. Tabla de contingencia obtenidas por el programa SPSS sobre la cual se calcularon los valores de S, E, VPP y VPN para la variable coronas.

La capacidad diagnóstica del programa para la detección de coronas es elevada, dado que en la curva de ROC realizada se obtiene un área bajo la curva de 0,927. **Figura 18, Tabla 17.**



Los segmentos de diagonal se generan mediante empates.

Figura 18. Gráfico de la Curva ROC para la variable coronas.

Área bajo la curva

Variables de resultado de prueba: CorDentiAi

Área	Desv. Error ^a	Significación asintótica ^b	95% de intervalo de confianza asintótico	
			Límite inferior	Límite superior
,927	,021	,000	,886	,967

Las variables de resultado de prueba: CorDentiAi tienen, como mínimo, un empate entre el grupo de estado real positivo y el grupo de estado real negativo. Las estadísticas podrían estar sesgadas.

a. Bajo el supuesto no paramétrico

b. Hipótesis nula: área verdadera = 0,5

Tabla 17. Tabla informativa sobre la representación gráfica de la Curva ROC para la variable coronas.

Detección de implantes osteointegrados

En el caso de la detección de implantes osteointegrados, el programa obtuvo valores perfectos (100%) de diagnóstico de los mismos (S, E, VPP y VPPN). **Tabla 18.** Además, tal y como se muestra en la curva de ROC, el área bajo la misma es de 1 (lo cual indica perfección diagnóstica del programa para esta variable). **Figura 19, Tabla 19.**

Tabla cruzada IOIOp*IOIDentiAI

			IOIDentiAI		Total
			0	1	
IOIOp	0	Recuento	35	0	35
		% dentro de IOIOp	100,0%	0,0%	100,0%
	1	Recuento	0	265	265
		% dentro de IOIOp	0,0%	100,0%	100,0%
Total		Recuento	35	265	300
		% dentro de IOIOp	11,7%	88,3%	100,0%

Tabla 16. Tabla de contingencia obtenidas por el programa SPSS sobre la cual se calcularon los valores de S, E, VPP y VPN para la variable implantes osteointegrados.

Evaluación de la fiabilidad del programa basado en Inteligencia Artificial Denti.Ai para la identificación de estructuras y tratamientos dentarios presentes en radiografías panorámicas.

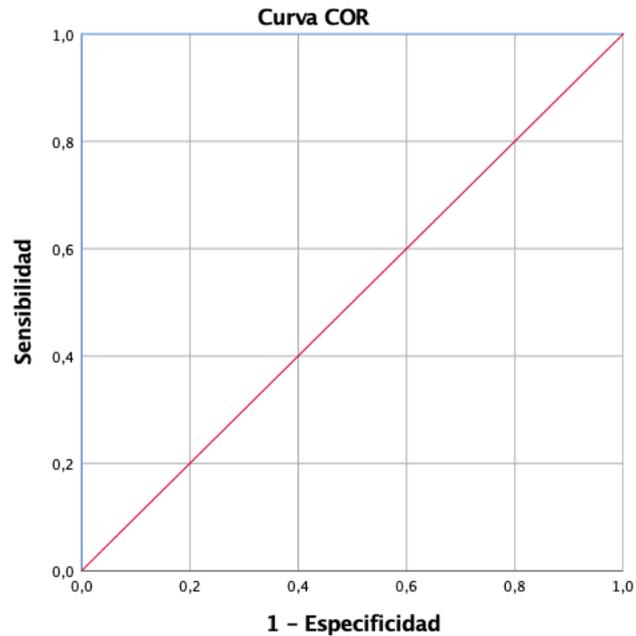


Figura 19. Gráfico de la Curva ROC para la variable implantes osteointegrados.

Área bajo la curva

Variables de resultado de prueba: ImpDentiAi

Área	Desv. Error ^a	Significación asintótica ^b	95% de intervalo de confianza asintótico	
			Límite inferior	Límite superior
1,000	,000	,000	1,000	1,000

a. Bajo el supuesto no paramétrico

b. Hipótesis nula: área verdadera = 0,5

Tabla 19. Tabla informativa sobre la representación gráfica de la Curva ROC para la variable implantes osteointegrados.

Discusión

Este estudio fue realizado con el fin de **evaluar la fiabilidad del programa basado en inteligencia artificial Denti.Ai para la detección de estructuras y tratamientos dentarios**. Como se ha podido observar con anterioridad en el apartado de Resultados, el programa Denti.Ai podría ser utilizado como soporte diagnóstico para la detección de estructuras y tratamientos dentarios presentes en radiografías panorámicas, obteniendo unos **valores cercanos al 100% de precisión diagnóstica en la detección de coronas, tratamiento de conductos e implantes; en torno al 80% en la detección de restauraciones metálicas y del 40% en restauraciones plásticas**.

A continuación, se describirá la discusión en dos apartados para facilitar su comprensión: discusión de la metodología utilizada y discusión de los resultados obtenidos.

Discusión de la metodología utilizada

La inteligencia artificial aplicada al campo de la salud es, a día de hoy, uno de los avances más destacables a los que se ha llegado gracias al desarrollo de la tecnología. Como se ha mencionado anteriormente, se trata de una nueva forma metodológica que permite la automatización del análisis y evaluación de situaciones y la aplicación de acciones o resultados directos en base a dicho análisis. Esta novedad ha sido aplicada en múltiples especialidades médicas y en algunas odontológicas (como puede ser en ortodoncia, con la digitalización de las arcadas dentarias y la implementación de técnicas de previsión de movimientos dentarios mediante softwares, para después confeccionar alineadores de plástico secuenciales; o en odontología general, por medio de la utilización de cepillos inteligentes que permiten conocer la calidad del cepillado realizado por los pacientes).

En lo referente al radiodiagnóstico dental, los sistemas de inteligencia artificial que han sido desarrollados centran su atención en las imágenes en dos dimensiones (ya sean radiografías panorámicas, periapicales o aletas de mordida). (29) No obstante, cabe destacar que existen algunos estudios en los que se utilizan estos mecanismos en pruebas diagnósticas en las tres dimensiones del espacio. (32)

En concreto, las radiografías panorámicas suponen un apoyo para el profesional a la hora de detectar patologías clínicamente no evidentes, así como para la realización de un diagnóstico inicial del paciente. No obstante, presentan algunas limitaciones que pueden entorpecer el proceso tecnológico de los sistemas inteligentes, como pueden ser la presencia de distorsiones anatómicas, variabilidad de los niveles de grises (dependiendo de las características del aparato de rayos utilizado) o artefactos. (20,33)

En el presente estudio se utilizaron radiografías panorámicas de pacientes que fueron realizadas con un único equipo de rayos. Sin embargo, este hecho no permite asegurar que todas las radiografías fueran tomadas con las mismas características de kilo voltaje (kVp), por lo que el contraste en cada una de las imágenes pudo variar y, como consecuencia, la detección automática de la red neuronal artificial pudo verse alterada. (20)

De igual manera, aunque la posición requerida del paciente para realización de la radiografía no es nada compleja, cualquier movimiento que este realice, una colocación inadecuada de los parámetros radiográficos, o diferencias entre los volúmenes a explorar, puede incorporar alteraciones a la imagen obtenida suponiendo, pues, una posible fuente de error del programa. (20)

Existe literatura relacionada en la que se describe el uso de radiografías obtenidas de varios equipos de radiodiagnóstico con el fin de asegurar cierta heterogeneidad en las características de las imágenes y permitiendo a su vez, poner a prueba el sistema artificial (21,28). Este procedimiento puede considerarse una óptima forma de asegurar un correcto entrenamiento de la red neuronal para la detección de estructuras anatómicas y tratamientos o patologías presentes en cualquier tipología de imagen. No obstante, supone una curva de aprendizaje mucho mayor y el empleo de un mayor número de pruebas para obtener resultados fiables.

El número de radiografías utilizadas en este estudio fue N=300. A diferencia de lo estipulado por otros autores (34), en este proyecto de investigación se destinó el total de imágenes a la fase experimental.

De acuerdo a la literatura publicada hasta el año 2021, lo ideal en este tipo de casos es la recopilación de un número elevado de imágenes con el fin de realizar una división aleatoria de las mismas en función del papel que ejerzan dentro del estudio: un mayor porcentaje iría destinado al “entrenamiento” del modelo artificial, y el restante a la realización de la fase experimental (a partir del cual se obtendrán los resultados). No obstante, cabe destacar que esta metodología ha sido únicamente descrita en aquellos estudios en los que se ha creado una red neuronal desde cero. (34) En nuestro caso, al tratarse de un programa de inteligencia artificial disponible desde el año 2017 y “entrenado” a priori (31), consideramos que la fase previa a la experimental ya había sido conseguida y, por tanto, omitible.

En el artículo publicado por *Schwendicke et al*, se menciona que el número medio total de radiografías panorámicas utilizadas por los autores es de aproximadamente 1053 (34). Aunque se han descrito estudios en los que se utilizaron un número similar al propuesto por los autores (31), *Kim et al.* emplearon un total 12.179 imágenes para el entrenamiento (11.189), validación (180) y realización de la fase experimental (800). (21)

En este estudio se utilizó el programa Denti.Ai para la realización del análisis de las radiografías. Sin embargo, la mayor parte de artículos publicados a día de hoy optaron por otras alternativas, como AlexNet, VGG16, ResNet, U-Net, VGG19 o GoogleNet, entre otros. (34)

En la actualidad, existen numerosos artículos científicos publicados relacionados con el empleo de sistemas de inteligencia artificial en el diagnóstico de imágenes radiográficas en odontología. Para facilitar su acceso, se ha elaborado una tabla resumen con los artículos científicos revisados que incluye la siguiente información de interés: autores, revista, año, tipo de red neuronal utilizada, tipo de imágenes utilizadas, número de imágenes totales y número de imágenes destinadas a fases de entrenamiento y experimental. (**Tabla 20**)

De los artículos incluidos en este estudio, no todos trataron de realizar un diagnóstico automático de tratamientos presentes en boca. Existen otros cuyo objetivo es utilizar los sistemas basados en IA con el fin de detectar presencia de caries (29,35) o enfermedad periodontal (pérdida ósea) (21).

De acuerdo al estudio publicado por *Abdalla-Aslan et al.* en el que se utilizaron radiografías panorámicas para la detección de restauraciones dentales, los modelos artificiales parecen nutrirse a partir de dos fases claramente diferenciadas: una primera etapa en la que se realiza la **delimitación automática de la región de interés (Region of Interest, ROI)** y una segunda en la que se realiza la **detección y la clasificación** de las restauraciones. Por lo general, para poderlo llevar a cabo, el algoritmo inteligente utiliza un método basado en la diferenciación de los niveles de grises. (20)

Aunque este procedimiento no sea del todo cuestionable, cabe destacar que la asignación de un único valor umbral de niveles de grises en la totalidad de la superficie radiográfica no es apropiado debido a la falta de homogeneidad de los mismos en la propia imagen. Por ello, los autores propusieron una alternativa metodológica que consistía en el uso de diferentes valores umbral de acuerdo a las diferentes regiones presentes en la misma. (20)

Autores	Revista	Año	Red Neuronal	Tipo de imagen	Nº total de imágenes (entrenamiento, fase experimental, validación)
<i>Abdalla-Aslan, et al.</i> (20)	Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology	2020	Matlab (Machine Learning)	Radiografías Panorámicas	83 (no específica)
<i>Aliaga, et al.</i> (22)	Artificial Intelligence in Medicine	2020	Matlab (Machine Learning)	Radiografías Panorámicas	370 (no específica)
<i>Devito, et al.</i> (35)	Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology	2008	No específica CNN (Convolutional Neural Network)	Radiografías Aletas de mordida	160 (80, 40, 40)
<i>Hiraiwa, et al.</i> (36)	Dentomaxillofacial Radiology	2019	AlexNet + GoogleNet Inception	Radiografías Panorámicas + CBCTs	760 (no específica)
<i>Jae-Hong, et al.</i> (30)	Oral Diseases	2019	GoogleNet Inception	Radiografías Panorámicas + CBCTs	2,126 1140 Rx Panorámicas: 912 entrenamiento y validación, y 228 fase experimental (684+228, 228) 986 CBCTs: 789 entrenamiento y validación, y 197 fase experimental (592+197, 197)
<i>Kim, et al.</i> (21)	Nature Research	2019	DeNTNet	Radiografías Panorámicas	12.179 (11.189, 190, 800)
<i>Lee, et al.</i> (29)	Journal of Dentistry	2018	GoogleNet Inception	Radiografías periapicales	3000 (2400, 600,2400)
<i>Lee, et al.</i> (37)	Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology	2020	R-CNN	Radiografías Panorámicas	50 (30,10,10)
<i>Leite, et al.</i> (28)	Clinical Oral Investigations	2020	DeepLab-v3 + Resnet 101	Radiografías Panorámicas	153 (70,65,18)
<i>Miki, et al.</i> (38)	Computers un Biology and Medicine	2017	AlexNet	CBCTs	52 (0,47,5)
<i>Setzer, et al.</i> (39)	Journal of Endodontics	2020	U-Net	CBCTs	20
<i>Tuzoff, et al.</i> (31)	Dentomaxillofacial Radiology	2019	Denti.Ai	Radiografías Panorámicas	1574 (1352,222,0)

Tabla 20. Descripción de los artículos revisados para la realización de este estudio.

El proceso de análisis y detección manual de estructuras y tratamientos dentarios fue realizado por profesionales odontólogos con experiencia radiodiagnóstica. Aunque en la mayoría de artículos esta primera parte metodológica es realizada por los anteriormente mencionados (29), existen algunos en los que se describe la participación de técnicos expertos en radiodiagnóstico (28,31,37), auxiliares/higienistas dentales (21), expertos en medicina oral (20,37) y maxilofacial (28), o estudiantes de grado para la realización de esta tarea (39).

En cuanto al número de operadores utilizados para la realización de la primera fase del proyecto, de acuerdo a lo reportado por *Schwendicke et al.*, en la literatura publicada existe una variación de entre 1 a 5 profesionales implicados. En el presente estudio se optó por la realización de dicha tarea por dos odontólogos con experiencia radiodiagnóstica teniendo, en caso de discordancia, un tercer operador que permitiría establecer un acuerdo final entre ambos evaluadores.

Con el fin de agrupar la información sobre la experiencia diagnóstica de los operadores y el número de ellos que participaron en los estudios, se ha elaborado una tabla en la que se especifican dichos datos para cada artículo incluido. **Tabla 21.**

Autores	Revista	Año	Tipo de Operadores	Nº de Operadores
<i>Abdalla-Aslan, et al.</i> (20)	Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology	2020	Odontólogo especialista en Medicina Oral	1
<i>Aliaga, et al.</i> (22)	Artificial Intelligence in Medicine	2020	No específica	No específica
<i>Devito, et al.</i> (35)	Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology	2008	Radiólogos dentales, odontólogos generales, periodoncistas, odontopediatras.	25
<i>Hiraiwa. Et al.</i> (36)	Dentomaxilofacial Radiology	2019	No específica	1
<i>Jae-Hong, et al.</i> (30)	Oral Diseases	2019	No específica	No específica
<i>Kim, et al.</i> (21)	Nature Research	2019	Odontólogos e higienistas dentales	6
<i>Lee, et al.</i> (29)	Journal of Dentistry	2018	No específica	No específica
<i>Lee, et al.</i> (37)	Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology	2020	Radiólogo oral y maxilofacial	1
<i>Leite, et al.</i> (28)	Clinical Oral Investigations	2020	No específica	No específica
<i>Miki, et al.</i> (38)	Computers in Biology and Medicine	2017	No específica	No específica
<i>Setzer, et al.</i> (39)	Journal of Endodontics	2020	Radiólogo oral y maxilofacial, endodoncista, estudiante de último año.	3
<i>Tuzoff, et al.</i> (31)	Dentomaxilofacial Radiology	2019	Expertos en Radiodiagnóstico	5

Tabla 21. Comparación de operadores empleados en los artículos incluidos.

Discusión de los resultados obtenidos

La detección de estructuras dentarias en radiografías panorámicas y periapicales es el primer paso para el desarrollo de softwares que permitan la automatización en el diagnóstico de tratamientos y patologías dentarias. Se trata de una nueva forma de proporcionar un soporte médico que conlleva un proceso bastante complejo.

De acuerdo al *Institute of Electrical and Electronics Engineers (IEEE)*, en el que en el año 2015 se realizó un simposio internacional sobre nuevas técnicas de diagnóstico en imágenes biomédicas, se llegó a la conclusión de que la segmentación de las estructuras dentarias en radiografías es una tarea difícil de automatizar. En este evento se describieron tan sólo dos grupos de investigadores que consiguieron crear una red convolucional de neuronas artificiales que pudieran realizar una adecuada segmentación y clasificación dentaria (aunque el diagnóstico de lesiones cariosas no demostrase resultados satisfactorios). (37, 40, 41)

Desde la introducción de la inteligencia artificial en el campo de la odontología, se han descrito numerosos estudios en los que se trata de automatizar el diagnóstico de patologías y tratamientos en radiografías intraorales (37,42–45) siendo, en menor medida, el número de artículos dirigidos a la aplicación de esta nueva tecnología en ortopantomografías. (37,46).

La detección y segmentación automáticas de estructuras dentarias es una labor mucho más desafiante que la que supone realizarlo sobre estructuras de soporte óseas. Esto se debe a múltiples motivos, dentro de los cuales destacamos la presencia de varias piezas muy próximas entre si en ambas arcadas o la multitud de densidades radiológicas que pueden existir en una misma pieza dentaria, por ejemplo. (28) Además, de forma general, se puede afirmar que la delimitación de las raíces en radiografías panorámicas es más compleja que en radiografías intraorales. Tratándose de piezas dentarias multi-radiculares, esta tarea puede verse todavía más dificultada. Es por ello, por lo que la segmentación automática de estructuras dentarias en este tipo de imágenes puede verse algo comprometida. (28,37,41)

El artículo publicado por *Hiraiwa et al.*, hace referencia al uso de programas basados en inteligencia artificial para la delimitación y detección de anomalías morfológicas radiculares en radiografías panorámicas. En este estudio se realizó una comparación de la capacidad diagnóstica de dos programas anteriormente mencionados: Alexnet y GoogleNet. Los autores finalmente concluyeron que, aunque la calidad de los resultados depende del número de imágenes con el que sea entrenado el software, GoogleNet puede ser considerado como el mejor para esta tarea por sus características informáticas (aunque sin diferencias estadísticamente significativas). (36) De acuerdo a lo descrito en la publicación, estos dos programas de inteligencia artificial mostraron resultados diagnósticos similares o superiores a los obtenidos por radiólogos con más de 20 años de experiencia (que utilizaron como *Ground Truth*). (36)

En el caso del artículo publicado por *Silva et al.*, se describió la creación de un sistema basado en inteligencia artificial que fue entrenado con una sola imagen panorámica en la que se seleccionaron, de forma automática y en función de diferentes regiones de interés (ROI), las 32 estructuras dentarias presentes en boca. Sin embargo, este estudio no obtuvo resultados satisfactorios debido a una inadecuada segmentación de las raíces dentarias, obteniendo una forma de las mismas no fieles a la realidad. (46)

Sin embargo, tanto *Wirtz et al.* como *Lee et al.*, presentaron un modelo basado en 10 radiografías panorámicas en las que se realizó una clasificación manual de las estructuras presentes en la imagen, la cual fue utilizada como “*Ground Truth*”. Los resultados obtenidos por estos equipos investigadores fueron significativamente mejores que a los que llegaron *Silva et al.* (37,41) Además, estos (junto con otros autores) afirmaron que, para obtener un mínimo de calidad en la segmentación realizada por el software, es necesario pasar por una primera etapa de selección manual realizada por un operador humano. (28,37,41) **Figura. 20.**



Figura 20. Delimitación de estructuras dentarias llevada a cabo por operadores (*Ground Truth*, verde); y lo sugerido por el programa (rojo) en el estudio realizado por *Lee et al.* (36)

Leite et al. llevaron a cabo una investigación en la que se obtuvieron resultados muy positivos en la detección, numeración y segmentación dentaria en radiografías panorámicas. En concreto, los autores señalaron que el mejor rendimiento en la ejecución de las tareas por parte de la red convolucional (DeepLab-v3 + Resnet 101) se dio, por orden decreciente, en: caninos, incisivos, premolares y molares. En este caso, se utilizaron dos redes convolucionales con el fin de aumentar la rapidez y la precisión de los resultados, de tal forma que DeepLab-v3 fue empleado para la detección y clasificación de las estructuras dentarias; y ResNet-101 para la segmentación de las mismas. De esta forma, y de acuerdo a los autores, se consiguieron resultados superiores a los de los estudios anteriormente publicados (36).

En esta investigación no se realizó una selección manual de todas las estructuras dentarias debido a que el programa objeto de estudio (Denti.Ai) delimita de forma automática la estructura de interés dentro del marco maxilo-mandibular. A diferencia de los programas mencionados anteriormente, el programa únicamente marca con un recuadro las piezas presentes en la radiografía panorámica (no atendiendo, pues, a la perfecta delimitación de la forma de la estructura en cuestión). **Figura 21.**



Figura 21. Delimitación en forma de recuadros de estructuras dentarias en el programa *Denti.Ai*.

En el presente estudio se encontró que el **programa acertó prácticamente en un 100% en la detección de las estructuras dentarias presentes en la radiografía panorámica**. Sin embargo, se observaron ciertos errores en la clasificación de los mismos (nomenclatura FDI), así como en la asignación a categoría dentaria de restauraciones como pónicos o coronas sobre implantes.

En el 50% de los casos se observó que Denti.Ai cometió errores de diversa índole (de mayor a menor frecuencia): asignación errónea de nomenclatura dentaria a piezas presentes en boca, confusión de piezas naturales con pónicos o coronas sobre implantes, no detección de piezas incluidas, supernumerarios o piezas presentes en boca (error asociado en imágenes borrosas).

Los resultados obtenidos en la **detección dentaria según** el artículo publicado por *Tuzoff et al.* fueron muy positivos, dado que se registraron una sensibilidad y una especificidad en dicha tarea de 99,41% y 99,45%, respectivamente (en comparación con la medición base del operador, la cual registró valores de sensibilidad y especificidad de 99,80% y 99,98%, respectivamente). Sin embargo, en el presente estudio se obtuvieron resultados diferentes, obteniendo una S de 56,95% y una E de 92,2%. De acuerdo al valor de VPN obtenido, del total de casos en los que el programa falló en la detección de estructuras dentarias, tan sólo el 42,51% de ellos fueron realmente acertados.

En la tarea de **numeración de las estructuras dentarias** (con nomenclatura FDI), y según el artículo de *Tuzoff et al.*, el programa obtuvo unos valores de sensibilidad y especificidad de 98% y 99,94% respectivamente (en comparación con la medición base del operador, que registró valores de sensibilidad y especificidad de 98,93% y 99,97% respectivamente).(31)

Por lo general, en su estudio se describieron buenos resultados en la detección dentaria en radiografías de buena calidad, pero con situaciones diversas: desde normoposiciones dentarias a casos con solapamiento de estructuras por apiñamiento. Los autores describieron que el programa utilizado (Denti.Ai) obtuvo también resultados satisfactorios en casos de imágenes borrosas o con peor calidad. Además, afirman que el programa supo excluir correctamente aquellas piezas reemplazadas con pónicos e implantes (contrario a lo encontrado en el presente estudio). (31)

Los casos en los que el programa no supo detectar estructuras dentarias presentes (falsos negativos), fueron asociados a pequeños restos radiculares (de difícil detección incluso para el operador), presencia de objetos ortopédicos que dificultaban la delimitación dentaria o un gran apiñamiento. Del mismo modo, el programa detectó como presentes algunas estructuras ausentes pero reemplazadas por implantes y pónicos. (31)

En cuanto a la numeración proporcionada por Denti.Ai, los autores reflejaron que los principales errores fueron asociados a la ausencia de piezas dentarias adyacentes a las estructuras presentes en la radiografía (confundía el diente presente con el ausente adyacente a él, también reportado en el presente estudio). Se observó una mayor frecuencia de este error asociada a la región molar (tanto superior como inferior). (31)

En el estudio publicado por *Leite et al.* se menciona que las posibles causas de un aumento de error (falsos negativos) en la segmentación dentaria automática pueden ser producidos por la presencia de imágenes radiopacas que aparecen en la región anterior (como línea ósea palatina, espina nasal anterior, o incluso el dispositivo intermaxilar del aparato de rayos que es utilizado para el correcto posicionamiento de los pacientes). Según lo reportado por los autores, estos

artefactos pueden inducir cierta confusión sobre los softwares de inteligencia artificial, produciendo una delimitación de estructuras algo distorsionada o completamente incorrecta. (28) Sin embargo, los autores hacen hincapié en que el diagnóstico sugerido por el sistema en casos de presencia de tratamientos como implantes o puentes no es el adecuado, por lo que debería ser posteriormente estudiado.

La calidad de las imágenes utilizadas o la edad de los pacientes incluidos en el estudio, son otros factores limitantes del estudio publicado, dado que los autores mencionan que se realizó una selección de radiografías de pacientes jóvenes (18 años) en las que la calidad era casi perfecta y no había ausencias dentarias (obteniendo, pues, menor probabilidad de fallo por parte del sistema). De hecho, se reconoce que posteriores investigaciones tendrán que ser llevadas a cabo con la inclusión de radiografías de diversas características en términos de edad de pacientes, calidad de la imagen y del balance de grises. (28)

En cuanto a la **detección y clasificación de restauraciones dentarias**, *Abdalla-Aslan et al.*, obtuvieron resultados muy satisfactorios. En su estudio se realizó una categorización inicial en la que se incluyeron los tratamientos más comúnmente encontrados en los pacientes (obturaciones de amalgama y de composite, coronas dentarias, implantes y tratamiento de conductos); y una segunda categorización en la que se combinaron las anteriores según frecuencia de aparición de forma combinada (implante + corona, tratamiento de conductos + corona). (20)

En este artículo se describe un elevado rendimiento diagnóstico por parte del sistema artificial, obteniendo una sensibilidad en la detección de tratamientos de alrededor del 94,6% (siendo mayores los resultados obtenidos en la detección de coronas, implantes y obturaciones de amalgama) y una precisión en la fase de clasificación de los mismos del 93,6%. (20)

El algoritmo utilizado por el sistema basado en inteligencia artificial del artículo basa su toma de decisiones en función de diferentes valores de umbral de grises y la posición que estos tengan dentro de la imagen. Según lo reportado por los autores, este mecanismo funciona adecuadamente en regiones en las que se diferencien claramente objetos de dos densidades distintas, pero no en aquellas en las que el balance de grises sea parecido. Este es el motivo por el cual se obtuvieron valores más bajos en la detección y clasificación de tratamientos como obturaciones de composite o tratamientos de conductos (de 83,1% y 83,2% respectivamente).(20) Sin embargo, los resultados obtenidos en nuestro estudio otorgaron una menor sensibilidad de detección a tratamientos de restauraciones plásticas (S=41,11%) y las coronas (S=89,53%), que a veces eran confundidas con restauraciones metálicas extensas.

En el presente estudio también se encontraron resultados satisfactorios en la detección y clasificación de obturaciones metálicas (S=85,58%), así como en casos de presencia de coronas (S= 89,53%), tratamiento de conductos (S= 91,9%) e implantes (100%). Sin embargo, la capacidad de detección del software Denti.Ai se vio comprometida en la clasificación de obturaciones plásticas debido a la presencia de radiopacidades similares al de la estructura dentaria. No obstante, cabe destacar que los casos en los que el software presentó mayor probabilidad de error fueron aquellos en los que se utilizaron composites fluidos que eran difíciles de detectar incluso por los operadores que participaron en el estudio.

De acuerdo a lo mencionado por *Abdala-Aslan et al.*, dentro de las limitaciones de su estudio puede destacarse la realización de todas las radiografías con un mismo aparato de rayos, lo cual favorece la obtención de resultados positivos por entrenar al sistema con unos parámetros determinados. Asimismo, se describe que la presencia pocos casos en ciertas categorías de tratamientos podría haber perjudicado el resultado por una escasa “formación” del sistema en el reconocimiento de los mismos. (20)

En definitiva, y como se ha mencionado anteriormente, el uso de programas de inteligencia artificial está siendo cada vez más utilizado en el ámbito biosanitario. En Odontología, se han descrito casos en los que este tipo de sistemas han sido empleados para la mejora del diagnóstico de patologías como caries (47), fracturas dentarias (48), enfermedades degenerativas periodontales (49), clasificación esquelética ortodóncica (cefalometrías) (50) o incluso para la diferenciación radiográfica entre tumores y quistes maxilo-mandibulares (51).

Además, los autores apuntan que el desarrollo de este tipo de tecnología puede mejorar la planificación y el uso de realidad virtual en procedimientos quirúrgicos, endodónticos o incluso estéticos. (28) De esta forma, con sugerencias automatizadas sobre la planificación y predicción de enfermedades, el trabajo de los profesionales será más eficiente y podrá asegurar resultados más predecibles. (49)

El presente estudio posee ciertas limitaciones. En primer lugar, y como se ha señalado en anteriores ocasiones, se utilizó un único aparato de radiodiagnóstico. Este hecho pudo influenciar la capacidad diagnóstica tanto del software como de los evaluadores (dado que ambos estaban condicionados por la calidad del mismo). Así mismo, incluso realizando las imágenes en un único sistema, la capacidad diagnóstica del software hubiera sido puesta a prueba realizando ciertas modificaciones en las características gráficas de cada imagen (exposición, contraste, etc). Al no haberse sometido a ningún tipo de alteración de las mismas y haber obtenido radiografías similares, puede decirse que el software (en nuestro caso) fue únicamente entrenado en la identificación de estructuras y tratamientos sobre imágenes de buena calidad. No obstante, cabe destacar que existen limitaciones intrínsecas al tipo de imágenes elegidas para la realización de este estudio y que pueden entorpecer los resultados del estudio, como son la presencia de artefactos o distorsiones anatómicas en las mismas.

Otra de las limitaciones de este estudio es la obtención de un único grupo de imágenes destinadas a la fase de prueba de Denti.Ai. La mayor parte de operadores realizaron una división de las imágenes en dos grupos: uno de entrenamiento del software, y otro para la prueba final del mismo. Con ello se pretendía corregir los fallos del sistema para que “aprendiese” a detectar correctamente en las imágenes de prueba finales. En el presente estudio, las anotaciones erróneas realizadas por Denti.Ai no fueron corregidas por los operadores. Este hecho supone que el programa entendía que lo que él detectaba era lo correcto (cuando ciertamente no lo era siempre), así como la imposibilidad de medir la capacidad de aprendizaje del mismo a lo largo del diagnóstico realizado en las 300 radiografías que se incluyeron en el estudio.

Conclusiones

Dados los resultados obtenidos en el presente estudio, se puede concluir que:

1. Los sistemas basados en inteligencia artificial podrán ser utilizados en distintas áreas del campo odontológico en un futuro cercano.
2. Denti.Ai puede considerarse un software cuyas tareas permiten dar soporte y asistencia al profesional odontólogo en procesos de diagnóstico, prevención y pronóstico dentarios.
3. **En la mitad de los casos el programa acertará en la detección de estructuras dentarias permanentes**, siendo la mitad restante ocupada por errores de diversas categorías.
4. **El programa Denti.Ai parece tener un adecuado entrenamiento para la detección de restauraciones dentarias metálicas**, siendo no tan acertada dicha tarea en el caso de restauraciones plásticas.
5. **La detección de tratamientos de conductos, coronas e implantes osteointegrados llevada a cabo por el programa Denti.Ai es prácticamente perfecta**, no habiendo apenas discrepancias diagnósticas en contraste a lo estipulado por el gold estándar.
6. **Mayor número de estudios implicados en la capacidad diagnóstica del programa deberían ser desarrollados** con el fin de mejorar el desempeño del mismo.

Bibliografía

1. Chen YW, Stanley K, Att W. Artificial intelligence in dentistry: Current applications and future perspectives. *Quintessence Int (Berl)*. 2020;51(3):248–57.
2. Khanagar SB, Al-ehaideb A, Maganur PC, Vishwanathaiah S, Patil S, Baeshen HA, et al. Developments, application, and performance of artificial intelligence in dentistry – A systematic review. *J Dent Sci*. 2020;1–15.
3. Schwendicke F, Samek W, Krois J. Artificial Intelligence in Dentistry: Chances and Challenges. *J Dent Res [Internet]*. 2020 Jul 21 [cited 2020 Oct 12];99(7):769–74. Available from: <http://journals.sagepub.com/doi/10.1177/0022034520915714>
4. Tandon D, Rajawat J. Present and Future of Artificial Intelligence in Dentistry. *J Oral Biol Craniofacial Res [Internet]*. 2020;1--20. Available from: <http://dx.doi.org/10.1016/j.jobcr.2017.01.002>
5. Gomolin A, Netchiporouk E, Gniadecki R, Litvinov I V. Artificial Intelligence Applications in Dermatology: Where Do We Stand? *Front Med*. 2020;7(March):1–7.
6. Shrivastava VK, Londhe ND, Sonawane RS, Suri JS. A novel and robust Bayesian approach for segmentation of psoriasis lesions and its risk stratification. *Comput Methods Programs Biomed*. 2017;150:9–22.
7. Xie F, Fan H, Li Y, Jiang Z, Meng R, Bovik A. Melanoma classification on dermoscopy images using a neural network ensemble model. *IEEE Trans Med Imaging*. 2017;36(3):849–58.
8. Yu C, Yang S, Kim W, Jung J, Chung KY, Lee SW, et al. Acral melanoma detection using a convolutional neural network for dermoscopy images. *PLoS One*. 2018;13(3):1–14.
9. Joda T, Waltimo T, Probst-Hensch N, Pauli-Magnus C, Zitzmann NU. Health Data in Dentistry: An Attempt to Master the Digital Challenge. *Public Health Genomics*. 2019;22(1–2):1–7.
10. Grischke J, Johannsmeier L, Eich L, Griga L, Haddadin S. Dentronics: Towards robotics and artificial intelligence in dentistry. *Dent Mater [Internet]*. 2020 Jun [cited 2020 Oct 12];36(6):765–78. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0109564120300762>
11. Lang T, Staufer S, Jennes B, Gaengler P. Clinical validation of robot simulation of toothbrushing - comparative plaque removal efficacy. *BMC Oral Health*. 2014;14(1):1–9.
12. Hwang G, Paula AJ, Hunter EE, Liu Y, Babeer A, Karabucak B, et al. Catalytic antimicrobial robots for biofilm eradication. *Sci Robot*. 2019;4(29):1–14.
13. Ortiz Simon, Jose Luis; Minor Martínez, Arturo; Lorias Espinoza, Daniel; Romero Velázquez JG. Mechatronic assistant system for dental drill handling. *Int J*. 2011;(7):22–6.
14. Rodr E, Kypson AP, Moten SC, Nifong LW, Jr WRC. Robotic mitral surgery at East Carolina University : *Int J*. 2006;(April):211–5.

15. Kunz F, Stellzig-Eisenhauer A, Zeman F, Boldt J. Artificial intelligence in orthodontics. Evaluation of a fully automated cephalometric analysis using a customized convolutional neural network. *Jornal Orofac Orthop.* 2020;(81):52–68.
16. Takahashi T, Nozaki K, Gonda T, Mameno T, Wada M, Ikebe K. Identification of dental implants using deep learning—pilot study. *Int J Implant Dent [Internet].* 2020 Dec 22 [cited 2020 Oct 12];6(1):53. Available from: <https://journalimplantdent.springeropen.com/articles/10.1186/s40729-020-00250-6>
17. Erpaçal B, Adıgüzel Ö, Cangül S. The use of micro-computed tomography in dental applications. *Int Dent Res.* 2019;9(2):78–91.
18. Boeddinghaus R, Whyte A. Dental panoramic tomography: An approach for the general radiologist. *Australas Radiol.* 2006;50(6):526–33.
19. Ludlow JB, Davies-Ludlow LE, White SC. Patient Risk Related to Common Dental Radiographic Examinations. *J Am Dent Assoc.* 2008;139(9):1237–43.
20. 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 [Internet].* 2020 Jun [cited 2020 Oct 12]; Available from: <https://linkinghub.elsevier.com/retrieve/pii/S221244032031018X>
21. Kim J, Lee HS, Song IS, Jung KH. DeNTNet: Deep Neural Transfer Network for the detection of periodontal bone loss using panoramic dental radiographs. *Sci Rep.* 2019;9(1):1–10.
22. Aliaga I, Vera V, Vera M, García E, Pedrera M, Pajares G. Automatic computation of mandibular indices in dental panoramic radiographs for early osteoporosis detection. *Artif Intell Med [Internet].* 2020 Mar [cited 2020 Oct 12];103(101816):1–17. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S093336571930421X>
23. Rushton VE, Horner K, Worthington H V. Screening panoramic radiology of adults in general dental practice: Radiological findings. *Br Dent J.* 2001;190(9):495–501.
24. Patanachai N, Covavisaruch N, Sinthanayothin C. Wavelet transformation for dental X-ray radiographs segmentation technique. *Proc - 2010 8th Int Conf ICT Knowl Eng ICT KE 2010.* 2010;103–6.
25. Poonsri A, Aimjirakul N, Charoenpong T, Sukjamsri C. Teeth segmentation from dental x-ray image by template matching. *BMEiCON 2016 - 9th Biomed Eng Int Conf.* 2017;10–3.
26. Oliveira JF, Proença H. Dental X-Ray: A Data Set of Panoramic Dental Radiographs for Stomatologic Image Processing Purposes. *Computational Methods in Applied Sciences.* 2011.
27. Vranckx M, Van Gerven A, Willems H, Vandemeulebroucke A, Leite AF, Politis C, et al. Artificial intelligence (Ai)-driven molar angulation measurements to predict third molar eruption on panoramic radiographs. *Int J Environ Res Public Health.* 2020;17(10):1–14.
28. Leite AF, Gerven A Van, Willems H, Beznik T, Lahoud P, Gaêta-Araujo H, et al. Artificial intelligence-driven novel tool for tooth detection and segmentation on panoramic radiographs. *Clin Oral Investig.* 2020;1–11.

29. Lee JH, Kim DH, Jeong SN, Choi SH. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. *J Dent*. 2018;77(June):106–11.
30. Jae-Hong L, Do-Hyung K, Seong-Nyum J. Diagnosis of cystic lesions using panoramic and cone beam computed tomographic images based on deep learning neural network. *Oral Dis*. 2019;(26):152–8.
31. Tuzoff D V., Tuzova LN, Bornstein MM, Krasnov AS, Kharchenko MA, Nikolenko SI, et al. Tooth detection and numbering in panoramic radiographs using convolutional neural networks. *Dentomaxillofacial Radiol* [Internet]. 2019 May [cited 2020 Oct 12];48(4):1–10. Available from: <https://www.birpublications.org/doi/10.1259/dmfr.20180051>
32. Xu X, Liu C, Zheng Y. 3D Tooth Segmentation and Labeling Using Deep Convolutional Neural Networks. *IEEE Trans Vis Comput Graph*. 2019;25(7):2336–48.
33. McDavid WD, Langlais RP, Welander U, Morris CR. Real, double, and ghost images in rotational panoramic radiography. *Dentomaxillofac Radiol*. 1983;12(2):122–8.
34. Schwendicke F, Golla T, Dreher M, Krois J. Convolutional Neural Networks for dental image diagnostics: A scoping review. *J Dent* [Internet]. 2019;91:1–14. Available from: <https://doi.org/10.1016/j.chieco.2019.101362>
35. Devito KL, de Souza Barbosa F, Filho WNF. An artificial multilayer perceptron neural network for diagnosis of proximal dental caries. *Oral Surgery, Oral Med Oral Pathol Oral Radiol Endodontology* [Internet]. 2008;106(6):879–84. Available from: <http://dx.doi.org/10.1016/j.tripleo.2008.03.002>
36. Hiraiwa T, Arijji Y, Fukuda M, Kise Y, Nakata K, Katsumata A, et al. A deep-learning artificial intelligence system for assessment of root morphology of the mandibular first molar on panoramic radiography. *Dentomaxillofacial Radiol*. 2019;48(3):1–7.
37. Lee JH, Han SS, Kim YH, Lee C, Kim I. Application of a fully deep convolutional neural network to the automation of tooth segmentation on panoramic radiographs. *Oral Surg Oral Med Oral Pathol Oral Radiol*. 2020;129(6):635–42.
38. Miki Y, Muramatsu C, Hayashi T, Zhou X, Hara T, Katsumata A, et al. Classification of teeth in cone-beam CT using deep convolutional neural network. *Comput Biol Med* [Internet]. 2017;80:24–9. Available from: <http://dx.doi.org/10.1016/j.compbio.2016.11.003>
39. Setzer FC, Shi KJ, Zhang Z, Yan H, Yoon H, Mupparapu M, et al. Artificial Intelligence for the Computer-aided Detection of Periapical Lesions in Cone-beam Computed Tomographic Images. *J Endod* [Internet]. 2020 Jul [cited 2020 Oct 12];46(7):987–93. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0099239920302351>
40. Lin N-H, Lin T-L, Wang X, Kao W-T, Tseng H-W, Chen S-L, et al. Teeth Detection Algorithm and Teeth Condition Classification Based on Convolutional Neural Networks for Dental Panoramic Radiographs. *J Med Imaging Heal Informatics*. 2018;8(3):509–13.
41. Wirtz A, Wambach J, Wesarg S. Automatic teeth segmentation in cephalometric X-ray images using a coupled shape model. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2018. p. 194–203.

42. Keshtkar F, Gueaieb W. Segmentation of dental radiographs using a swarm intelligence approach. *Can Conf Electr Comput Eng*. 2006;(May):328–31.
43. Rad AE, Rahim M, Kumoi R, Norouzi A. Dental x-ray image segmentation and multiple feature extraction. *2nd World Conf Innov Comput Sci - 2012 Dent*. 2012;(May 2014):188–97.
44. Lin PL, Huang PY, Huang PW. An effective teeth segmentation method for dental periapical radiographs based on local singularity. *ICSSE 2013 - IEEE Int Conf Syst Sci Eng*. 2013;407–11.
45. Dighe S, Revati S. Preprocessing , Segmentation and Matching of Dental Radiographs used in. *Int J Sci Appl Inf Technol*. 2012;1(2278):52–6.
46. Silva G, Oliveira L, Pithon M. Automatic segmenting teeth in X-ray images: Trends, a novel data set, benchmarking and future perspectives. *Expert Syst Appl*. 2018;107:15–31.
47. Moutselos K, Berdouses E, Oulis C, Maglogiannis I. Recognizing Occlusal Caries in Dental Intraoral Images Using Deep Learning. *Proc Annu Int Conf IEEE Eng Med Biol Soc EMBS*. 2019;1617–20.
48. Fukuda M, Inamoto K, Shibata N, Ariji Y, Yanashita Y, Kutsuna S, et al. Evaluation of an artificial intelligence system for detecting vertical root fracture on panoramic radiography. *Oral Radiol*. 2020;36(4):337–43.
49. Lee JH, Kim DH, Jeong SN, Choi SH. Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm. *J Periodontal Implant Sci*. 2018;48(2):114–23.
50. Yu HJ, Cho SR, Kim MJ, Kim WH, Kim JW, Choi J. Automated Skeletal Classification with Lateral Cephalometry Based on Artificial Intelligence. *J Dent Res*. 2020;99(3):249–56.
51. Ariji 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.

Anexos

Anexo 1. Evaluación favorable del Comité de Ética del Hospital Clínico San Carlos



CEIm Hospital Clínico San Carlos

Dictamen Protocolo Favorable

C.I. 21/297--E

22 de abril de 2021

Dra. Mar García Arenillas
Presidenta del CEIC Hospital Clínico San Carlos

CERTIFICA

Que el CEIm Hospital Clínico San Carlos en reunión de Comisión Permanente, acta 4.2/21, ha evaluado la propuesta del estudio:

Título: **EVALUACIÓN DE LA FIABILIDAD DEL PROGRAMA DENTI.AI EN LA IDENTIFICACIÓN DE ESTRUCTURAS Y TRATAMIENTOS DENTARIOS PRESENTES EN RADIOGRAFÍAS PANORÁMICAS**

Código Interno: **21/297-E**

Investigadores:	
GUILLERMO PRADÍES RAMIRO	TUTOR
MONICA BONFANTI GRIS	ALUMNO

Tipo documento	Versión
Protocolo	Abril 2021

Que en este estudio:

- Se cumplen los requisitos necesarios de idoneidad del protocolo en relación con los objetivos del estudio y están justificados los riesgos y molestias previsibles para el sujeto.
- La capacidad del investigador y los medios disponibles son adecuados para llevar a cabo el estudio.
- El alcance de las compensaciones económicas previstas no interfiere con el respeto de los postulados éticos.
- Se cumplen los preceptos éticos formulados en la Orden SAS 3470/2009 y la Declaración de Helsinki de la Asociación Médica mundial sobre principios éticos para las investigaciones médicas en seres humanos y en sus posteriores revisiones, así como aquellos exigidos por la normativa legal aplicable en función de las características del estudio.

Anexo 2. Plantilla de recolección de datos manual por los operadores.

N	Nº PIEZAS TOTALES	PIEZAS AUSENTES (FDI)	RESTAURACIONES PZA (FDI): RP, RM	TTO CONDUCTOS (FDI)	CORONA/ PÓNTICO (FDI)	IOI (FDI)	N	Nº PIEZAS TOTALES	PIEZAS AUSENTES (FDI)	RESTAURACIONES PZA (FDI): RP, RM	TTO CONDUCTOS (FDI)	CORONA/ PÓNTICO (FDI)	IOI (FDI)
			1C: 2C: 3C: 4C:							1C: 2C: 3C: 4C:			
			1C: 2C: 3C: 4C:							1C: 2C: 3C: 4C:			
			1C: 2C: 3C: 4C:							1C: 2C: 3C: 4C:			
			1C: 2C: 3C: 4C:							1C: 2C: 3C: 4C:			
			1C: 2C: 3C: 4C:							1C: 2C: 3C: 4C:			
			1C: 2C: 3C: 4C:							1C: 2C: 3C: 4C:			
			1C: 2C: 3C: 4C:							1C: 2C: 3C: 4C:			
			1C: 2C: 3C: 4C:							1C: 2C: 3C: 4C:			
			1C: 2C: 3C: 4C:							1C: 2C: 3C: 4C:			
			1C: 2C: 3C: 4C:							1C: 2C: 3C: 4C:			