



**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL
FACULTAD DE CIENCIAS DE LA SALUD
CARRERA DE ODONTOLOGÍA**

TEMA:

Precisión diagnóstica de la inteligencia artificial comparada con la interpretación humana en radiografías panorámicas.

AUTORA:

Jarro Jaramillo, Kerly Arlen

**Trabajo de titulación previo a la obtención del título de
ODONTÓLOGA**

TUTOR:

Dra. Díaz Rojas, Dennisse Fernanda

**Guayaquil, Ecuador
7 de septiembre del 2026**



**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL
FACULTAD DE CIENCIAS DE LA SALUD
CARRERA DE ODONTOLOGÍA**

CERTIFICACIÓN

Certificamos que el presente trabajo de titulación fue realizado en su totalidad por **Jarro Jaramillo, Kerly Arlen**, como requerimiento para la obtención del título de **Odontólogo**.

TUTOR (A)

f. 

Dra. Díaz Rojas, Dennisse Fernanda

DIRECTOR DE LA CARRERA

f. _____

Dra. Bermúdez Velásquez, Andrea Cecilia

Guayaquil, a los 7 días del mes de septiembre del año 2026



**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL
FACULTAD DE CIENCIAS DE LA SALUD
CARRERA DE ODONTOLOGÍA**

DECLARACIÓN DE RESPONSABILIDAD

Yo, Jarro Jaramillo, Kerly Arlen

DECLARO QUE:

El Trabajo de Titulación, **Precisión diagnóstica de la inteligencia artificial comparada con la interpretación humana en radiografías panorámicas** previo a la obtención del título de **Odontóloga**, ha sido desarrollado respetando derechos intelectuales de terceros conforme las citas que constan en el documento, cuyas fuentes se incorporan en las referencias o bibliografías. Consecuentemente este trabajo es de mi total autoría.

En virtud de esta declaración, me responsabilizo del contenido, veracidad y alcance del Trabajo de Titulación referido.

Guayaquil, a los 7 días del mes de septiembre del año 2026

LA AUTORA

f. Kerly Jarro
Jarro Jaramillo, Kerly Arlen



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL
FACULTAD DE CIENCIAS DE LA SALUD
CARRERA DE ODONTOLOGÍA

AUTORIZACIÓN

Yo, **Jarro Jaramillo, Kerly Arlen**

Autorizo a la Universidad Católica de Santiago de Guayaquil a la **publicación** en la biblioteca de la institución del Trabajo de Titulación, **Precisión diagnóstica de la inteligencia artificial comparada con la interpretación humana en radiografías panorámicas**, cuyo contenido, ideas y criterios son de mi exclusiva responsabilidad y total autoría.

Guayaquil, a los 7 días del mes de septiembre del año 2026

LA AUTORA

f. Kerly Jarro
Jarro Jaramillo, Kerly Arlen



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL
FACULTAD DE CIENCIAS DE LA SALUD
CARRERA DE ODONTOLOGÍA

REPORTE COMPILATIO



Informe de análisis
Compilatio Magister+ | UCSG-EC- Universidad Católica de Santiago de Guayaquil

KERLY ARLEN JARRO copylatio

ID : 3959a0d2937de4ef3b1cb4d93f6f895d5f333e40

0%

Textos sospechosos

Nombre del fichero : KERLY ARLEN JARRO copylatio.txt
Tamaño del archivo original : 42,53 kB
Número de palabras : 3040

Depositante : Dennisse Fernanda Díaz Rojas
Fecha de depósito : 3 de septiembre de 2026
Tipo de carga : Interface
Fecha de fin de análisis : 3 de septiembre de 2026

Resumen (sección 1/3)

Localización de los textos sospechosos en el documento :



Incluido en el porcentaje de textos sospechosos :

Similitudes

Pasajes con similitudes a fuentes encontradas en diferentes colecciones.

0%

Detección de IA

Textos estilísticamente próximos a un texto generado por una IA.
Este índice es un indicador y no una prueba. Comprueba con el autor si domina los conocimientos mencionados en el documento.

0%

Idiomas no reconocidos

Pasajes en los que parte del vocabulario utilizado no forma parte del diccionario de la lengua.
Puede tratarse de un intento del autor de modificar el texto para evitar ser detectado.

0%

No incluido en el porcentaje de textos sospechosos :

Textos entre comillas

Pasajes entre comillas, a menudo indicativos de una cita.

0%

TUTOR (A)

f. _____

Dra. Díaz Rojas, Dennisse Fernanda

AGRADECIMIENTO

En primer lugar, le quiero dar gracias a Dios, quien es el principal autor de este sueño y a quien me encomendé siempre en cada momento durante mi formación académica. Gracias, mi Señor del Cielo, por no soltarme y por darme fuerzas para seguir luchando; Tu fe y Tu amor siempre estuvieron presentes.

A mis padres, Carlos Jarro y Betsabé Jaramillo, quienes me llevaron de la mano desde niña a mi primer día de clases y hasta el día de hoy, que ingresan conmigo a terminar un sueño que también es de ustedes. A través de mí se cumplen sus sueños y me siento orgullosa de poder dárselos. Gracias por ser mis padres, los amo con todo mi corazón.

También a mis hermanos: Nathaly, Jean, Nicol y Michael, quienes creyeron en mí y cuyo amor me acompañó siempre. Gracias por ser mis cómplices de travesuras y de amor; tenerlos como hermanos es el tesoro más maravilloso que Dios me dio.

A mis sobrinos, que son el amor de mi vida y cuya fuerza me impulsó desde que nacieron. A mi cuñada, Angie Morocho, le doy las gracias por creer en mí y por ser mi apoyo en las publicaciones para la búsqueda de pacientes muchas gracias, cuñada

A mis docentes, quienes me enseñaron el amor hacia la Odontología y de quienes aprendí a desenvolverme mejor en el área clínica: Dra. Geoconda Luzardo y Dra. Alejandra Torres, muchas gracias.

A mi tutora, Dra. Denisse Díaz, por guiarme durante mi trabajo de titulación, ya que sin su orientación este trabajo no hubiera sido posible.

Gracias a todos por haber formado parte de este capítulo en mi vida. Hoy se cierra una etapa de mi formación universitaria, pero se vienen muchas más junto a ustedes. Espero que la vida me permita seguir teniéndolos a mi lado. Muchas gracias. Cierro este capítulo con lágrimas y con el corazón lleno de gratitud hacia ustedes, porque fui bendecida al tenerlos en mi vida.

Muchas gracias por haber estado conmigo.

Con amor,

Kerly Jarro

DEDICATORIA

Esta tesis va dedicada a mi madre, Betsabé Jaramillo, quien me apoyó de principio a fin en mi carrera universitaria. Si no fuera por ella, este sueño no sería posible. A ti, mamita, gracias por llevarme en tu vientre nueve meses, por haberme criado de una manera maravillosa y por hacer que sea la mujer que soy hoy. No me alcanzará la vida para recompensarte todo lo que me has dado, pero con este trabajo que hice con amor para ti, te compenso un poco de todo lo que me has brindado. Eres mi orgullo como mujer y un ejemplo de madre.

Junto a ti la vida ha sido maravillosa, aunque muchas veces fue difícil; tú, como madre, hiciste que no sintiera tanto los golpes de la vida porque muchas veces abandonaste tus sueños por cumplir los míos, por verme feliz y dejarme soñar. Gracias por todo lo que hiciste por mí, por tus sacrificios y por todas las lágrimas que derramaste por mí cuando el corazón lo tenía cansado y la fe quebrada. Gracias por curar mi corazón, porque Dios no pudo darme mejor madre que tú, mamá.

Te amo con todo el corazón y, si hay vida después de la muerte, espero que Dios me siga permitiendo ser tu hija para volvernos a encontrar. Te amo. Hoy te regalo esta tesis a ti y hoy no solo se titula tu hija, sino también la madre de una odontóloga, porque a través de mí vive tu esencia.

Gracias por todo, te amo mamita



**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL
FACULTAD DE CIENCIAS DE LA SALUD
CARRERA DE ODONTOLOGÍA**

TRIBUNAL DE SUSTENTACIÓN

f. _____

Dra. Andrea Cecilia Bermúdez Velásquez
DECANO O DIRECTOR DE CARRERA

f. _____

Dra. Estefania Del Rocio Ocampo Poma
COORDINADOR DEL ÁREA O DOCENTE DE LA CARRERA

f. _____

Dra. Andrea Cecilia Bermúdez Velásquez
OPONENTE



**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL
FACULTAD DE CIENCIAS DE LA SALUD
CARRERA DE ODONTOLOGÍA**

CALIFICACIÓN

TUTOR (A)

f. _____
Dra. Dennisse Diaz Rojas

RESUMEN

Antecedentes: La inteligencia artificial se ha incorporado al análisis de imágenes odontológicas como herramienta de apoyo para identificar diferentes hallazgos. Sin embargo, su desempeño puede variar según el sistema utilizado y el tipo de alteración evaluada, por lo que es necesario compararla con la interpretación profesional. **Objetivo:** Determinar si la precisión diagnóstica de los sistemas de inteligencia artificial Gemini y NeuroVet Dental es comparable o superior a la interpretación humana en radiografías panorámicas odontológicas. **Materiales y métodos:** Estudio cuantitativo, observacional, retrospectivo y de exactitud diagnóstica. Se analizaron 140 radiografías panorámicas digitales mediante Gemini, NeuroVet Dental y una especialista en odontología. Se evaluó la presencia de caries, fracturas coronarias, restauraciones, implantes y prótesis fija, utilizando como patrón de referencia la interpretación de una especialista en Radiología Oral. Se calcularon sensibilidad, especificidad, exactitud, valores predictivos y concordancia. **Resultados:** La evaluación humana presentó la mayor exactitud y concordancia en la mayoría de los hallazgos. Gemini obtuvo resultados comparables para caries, con una exactitud del 94,3% frente al 96,4% de la evaluadora humana. En implantes no se encontraron diferencias significativas entre los métodos ($p=0,135$). En fracturas, restauraciones y prótesis fija, la evaluación humana fue significativamente más exacta que ambos sistemas de inteligencia artificial ($p<0,001$). **Conclusión:** Gemini y NeuroVet Dental no alcanzaron una precisión diagnóstica comparable o superior a la interpretación humana. Su desempeño dependió del hallazgo evaluado, por lo que pueden utilizarse como herramientas de apoyo, pero requieren supervisión profesional.

Palabras Clave: Inteligencia Artificial; Radiografía Panorámica; Diagnóstico por Imagen; Sensibilidad y Especificidad; Odontología.

ABSTRACT

Background: Artificial intelligence has been incorporated into dental image analysis as a supportive tool for identifying various findings. However, its performance can vary depending on the system used and the type of abnormality being assessed; therefore, comparison with professional interpretation is necessary. **Objective:** To determine whether the diagnostic accuracy of the Gemini and NeuroVet Dental artificial intelligence systems is comparable or superior to human interpretation in dental panoramic radiographs. **Materials and Methods:** A quantitative, observational, retrospective study on diagnostic accuracy. A total of 140 digital panoramic radiographs were analyzed using Gemini, NeuroVet Dental, and an rehabilitation specialist. The presence of caries, crown fractures, restorations, implants, and fixed prostheses was evaluated, using the interpretation of an oral radiology specialist as the reference standard. Sensitivity, specificity, accuracy, predictive values, and agreement were calculated. **Results:** Human evaluation demonstrated the highest accuracy and agreement for the majority of findings. Gemini yielded comparable results for caries, with an accuracy of 94.3% compared to 96.4% for the human evaluator. No significant differences were found between the methods regarding implants ($p=0.135$). For fractures, restorations, and fixed prostheses, human evaluation was significantly more accurate than both artificial intelligence systems ($p<0.001$). **Conclusion:** Gemini and NeuroVet Dental did not achieve diagnostic accuracy comparable or superior to human interpretation. Their performance depended on the specific finding being evaluated; thus, while they can serve as supportive tools, they require professional supervision.

Keywords: Artificial Intelligence; Radiography, Panoramic; Diagnostic Imaging; Sensitivity and Specificity; Dentistry.

Introducción

La radiografía panorámica es esencial para los diagnósticos odontológicos porque nos da una vista completa de la región dentomaxilofacial en una imagen, lo que permite realizar una planificación de tratamientos restaurativos, quirúrgicos, protésicos y de implantología, ya que facilita la identificación de alteraciones que puedan complementar la evaluación clínica del paciente.^{1,2} No obstante, su interpretación tiene retos técnicos como la distorsión geométrica, la superposición anatómica de las estructuras y la variabilidad en la calidad de adquisición. Estos elementos alteran la exactitud del diagnóstico y crean diferencias entre los observadores.³

La precisión diagnóstica no solo depende de la calidad de la imagen radiográfica adquirida, sino también por el criterio clínico, la experiencia y el conocimiento del profesional encargado de evaluar e interpretar.⁴ Esta variabilidad puede

adquirir mayor importancia cuando se analizan varios hallazgos al mismo tiempo (implantes, caries, prótesis o restauraciones), cuya complejidad visual dificulta la detección e interpretación de las mismas. Por ello, se ha incrementado el interés por crear instrumentos tecnológicos que ayuden al médico odontólogo durante el proceso de diagnóstico.^{3,5-7}

En este contexto, la inteligencia artificial (IA) progresa en el estudio automatizado de imágenes odontológicas a través del aprendizaje de algoritmos (DeepLearning), estos sistemas tienen la capacidad de identificar patrones complejos con rapidez y precisión, utilizando bases de datos radiográficas amplias.^{8,9} En odontología, se han mostrado resultados favorables al reconocer caries, dientes faltantes, restauraciones e implantes en radiografías panorámicas.^{10,11}

A pesar de estos avances, siguen existiendo limitaciones para garantizar la utilidad clínica definitiva. Estudios previos han señalado que, pese a los altos valores de

precisión reportados por diferentes modelos de IA, existe una amplia variabilidad entre los estudios debido a diferencias en las bases de datos utilizadas, algoritmos empleados, criterios diagnósticos y métodos de validación. Esta heterogeneidad dificulta comparar directamente los resultados y establecer si estos sistemas mantienen un desempeño confiable fuera de los escenarios específicos en los que fueron entrenados y evaluados.^{3,12}

Aparte de este ámbito, existen muchas investigaciones que evalúan algoritmos aislados bajo metodologías distintas, existiendo escasa evidencia sobre la comparación directa entre plataformas de IA usando una misma muestra y un patrón de referencia común. Esta brecha demanda una mayor validación independiente que permita conocer su verdadero rendimiento real de diagnóstico antes de incorporarla como apoyo en la práctica clínica.^{13,14}

Por lo cual realizar evaluaciones con precisión en los diagnósticos mediante indicadores como la sensibilidad,

especificidad, exactitud, valores predictivos y concordancia resulta muy necesario para medir la capacidad de estos sistemas. La comparación entre herramientas de inteligencia artificial y evaluadores humanos permite no solo identificar el grado de acierto de la tecnología, sino también sus principales limitaciones y posibles errores en los diagnósticos.

El propósito de esta investigación consiste en determinar si la precisión diagnóstica de los sistemas de inteligencia artificial Gemini y NeuroVet Dental es comparable o superior a la interpretación humana en la evaluación de radiografías panorámicas odontológicas.

MATERIALES Y MÉTODOS

El presente estudio tuvo un enfoque cuantitativo, con un diseño analítico, observacional y retrospectivo, de exactitud diagnóstica. La investigación fue desarrollada en las instalaciones de la Clínica Odontológica de la Universidad

Católica de Santiago de Guayaquil (UCSG), mediante la revisión de una base de radiografías panorámicas digitales del semestre B-2025 pertenecientes al laboratorio radiológico de la institución.

Durante el desarrollo de este estudio se evaluó la capacidad diagnóstica de Gemini, NeuroVet Dental y una especialista en dentística para identificar hallazgos específicos en las radiografías seleccionadas (lesiones cariosas, fracturas coronarias, restauraciones, implantes dentales y prótesis fija). Los resultados obtenidos fueron comparados con la interpretación realizada por una Odontóloga Especialista en Radiología Oral, la cual fue considerada como Gold Standard.

Población y muestra

La base de datos sobre la que se trabajó estuvo conformada por 1.539 radiografías panorámicas digitales. De forma inicial, se realizó una revisión de todas las radiografías disponibles, durante este

proceso se excluyeron imágenes radiográficas de pacientes pediátricos, radiografías digitales de baja calidad, así como aquellas que presentaron alguna distorsión o presentaban algún artefacto que impidiera realizar una correcta revisión. Finalmente, el universo estuvo constituido por 219 radiografías panorámicas digitales de pacientes adultos.

Se utilizó SurveyMonkey para realizar un cálculo muestral, para esto se consideró el universo ($n = 219$), un nivel de confianza del 95%, margen de error del 5% y una proporción esperada del 50%. El tamaño de muestra final fue de 140 radiografías digitales. La unidad de análisis correspondió a cada radiografía panorámica incluida en el estudio, donde se evaluó la presencia o ausencia de los hallazgos ya establecidos en cada imagen.

Criterios de selección del estudio

Se incluyeron: radiografías panorámicas digitales de pacientes adultos, que presenten una calidad diagnóstica aceptable

(alta o media) y que cuenten con la información completa: edad y sexo.

Se excluyeron: radiografías de pacientes pediátricos, con distorsiones o artefactos que impidan realizar una correcta interpretación radiográfica, así como aquellas con datos incompletos o provenientes de otras bases de datos o período académico.

Variables

Variable dependiente: Precisión diagnóstica.

Se evaluó mediante los siguientes indicadores: sensibilidad (Se), especificidad (Sp), exactitud diagnóstica (Acc), valor predictivo positivo (VPP), valor predictivo negativo (VPN), falsos negativos y positivos, concordancia diagnóstica con el gold standard.

Variable independiente: Método de interpretación radiográfica.

Se categorizó en:

- Gemini

- NeuroVet Dental
- Evaluador humano (Especialista en Dentística restauradora).

La variable de control fue la calidad de la imagen radiográfica, clasificada en alta o media, y como variables de caracterización la edad (expresada en años y categorizada en grupos etarios) y el sexo (masculino o femenino).

Procedimientos

Tras obtener la autorización institucional para acceder a la base de datos de radiografías panorámicas digitales, se inició con el proceso de revisión y selección de las imágenes radiográficas disponibles teniendo en cuenta los criterios de inclusión y exclusión.

Con las imágenes seleccionadas, se elaboró una ficha de recolección de datos digital en Microsoft Excel. Antes de iniciar el proceso de evaluación radiográfica por el evaluador humano y sistemas IA, se asignó un código alfanumérico con el objetivo de garantizar la confidencialidad de los pacientes. La

relación entre los códigos asignados y la identidad de los pacientes fue conocida únicamente por la investigadora principal.

Cada una de las radiografías que fueron incluidas se codificaron con códigos alfanuméricos con la finalidad de garantizar la confidencialidad de los participantes. Los datos obtenidos fueron registrados en aquella base.

Los datos fueron también codificados y registrados de manera dicotómica: 1 = presencia y 0 = ausencia. Se consideró “presencia” cuando se detectó al menos un hallazgo de la categoría evaluada en la radiografía, independientemente de su cantidad o ubicación en la cavidad bucal. Con esta información, un especialista en imagenología oral se encargó de realizar una evaluación independiente de todas las radiografías ($n = 140$) y se consideraron como el patrón de referencia para el estudio (gold standard).

El evaluador humano se encargó de interpretar las mismas radiografías de

forma independiente y ciega a los resultados establecidos como gold standard. Esto fue repetido con ambos sistemas IA (Gemini y NeuroVet Dental). Finalmente, todos los datos obtenidos se unificaron en una misma hoja de cálculo y posteriormente procesados en un programa estadístico.

Plan de tabulación y análisis estadístico

Los datos fueron tabulados y codificados en una hoja de cálculo de Microsoft Excel. Su análisis posterior fue realizado mediante el software estadístico IBM SPSS Statistics. Para describir las variables de caracterización de la muestra se realizó un análisis descriptivo, utilizando frecuencias, porcentajes, medidas de tendencia central y dispersión para presentar la edad y sexo.

Para comparar los hallazgos entre los tres métodos de evaluación establecidos se utilizó tablas de contingencia. Se determinó los verdaderos positivos y negativos, falsos positivos y negativos. Se calculó la Se, Sp, Acc, VPP y VPN. La Se y Sp fueron presentados con IC del 95%.

Para evaluar la concordancia entre los 3 métodos vs. El Gold standard se utilizó el coeficiente Kappa de Cohen. Para evaluar la exactitud de los tres métodos se utilizó la prueba Q de Cochran. Cuando la comparación global resultó ser significativa, se realizó comparaciones pareadas utilizando la prueba exacta de McNemar con corrección de Bonferroni. Se consideró estadísticamente significativo un valor de $p < 0,05$.

RESULTADOS

En la **Tabla 1** se observó que el evaluador humano fue el que alcanzó un 100% de sensibilidad al detectar lesiones cariosas, restauraciones y los implantes, mientras que para fracturas coronarias obtuvo un 96,9%. Por su parte, Gemini presentó una sensibilidad 98,5% para caries y 88,1% para restauraciones. Su sensibilidad fue menor para prótesis fija, implantes y fracturas coronarias. NeuroVet Dental alcanzó su mayor sensibilidad en restauraciones, con un 91%. Su desempeño fue menor en caries, con un 50%, y

principalmente en fracturas coronarias, con un 23,1%.

Respecto al valor predictivo positivo, el evaluador humano alcanzó un VPP del 100% para implantes. NeuroVet Dental también obtuvo un VPP del 100% para caries y fracturas coronarias (Tabla 1).

Tabla 1. Sensibilidad de los métodos evaluados según el hallazgo radiográfico

Halla zgo	GS positi vos n	Evalua dor % (IC95 %)	VPP %	Gemi ni % (IC95 %)	VPP %	Neur oVet % (IC95 %)	VPP %
Caries	134	100,0 (97,3 – 100,0)	96,4	98,5 (94,7 – 99,8)	95,7	50,0 (41,2 – 58,8)	100,0
Fracturas coronarias	65	96,9 (89,3 – 99,6)	96,9	61,5 (48,6 – 73,3)	64,5	23,1 (13,5 – 35,2)	100,0
Restauraciones	134	100,0 (97,3 – 100,0)	98,5	88,1 (81,3 – 93,0)	95,9	91,0 (84,9 – 95,3)	96,1
Implantes	3	100,0 (29,2 – 100,0)	100,0	66,7 (9,4– 99,2)	66,7	66,7 (9,4– 99,2)	66,7
Prótesis fija	15	86,7 (59,5 – 98,3)	92,9	73,3 (44,9 – 92,2)	40,7	60,0 (32,3 – 83,7)	45,0

En la **Tabla 2** se encontró que el evaluador humano obtuvo una especificidad del 100% en implantes. Gemini de igual forma

presentó un porcentaje alto para implantes, con un 99,3% al igual que NeuroVet Dental. En prótesis fija se encontró una especificidad alta para el evaluador humano y NeuroVet Dental (91,2%), mientras Gemini presentó el porcentaje más bajo (87,2%). Para fracturas coronarias, NeuroVet Dental obtuvo un 100%, el evaluador humano obtuvo un 97,3% y Gemini un 70,7%. Para las restauraciones se observaron porcentajes más bajos en los 3 métodos, el evaluador humano con un 66,7%, Gemini y NeuroVet Dental con un 16,7% cada uno. Se observó que, para caries, NeuroVet Dental alcanzó una especificidad alta (100%), mientras que el evaluador humano y Gemini presentaron porcentajes bajos.

En implantes, el evaluador humano alcanzó un VPN del 100%, mientras que Gemini y NeuroVet Dental obtuvieron un 99,3%. Sin embargo, NeuroVet presentó un VPN de apenas 8,2% para caries y del 60% para fracturas coronarias (Tabla 2).

Tabla 2. Especificidad de los métodos evaluados según el hallazgo radiográfico

Hallazgo	GS ne gat ivo s n	Eval uado r % (IC9 5 %)	V P N %	Ge min i % (IC 95 %)	V P N %	Neur oVet % (IC9 5 %)	V P N %
Caries	6	16,7 (0,4– 64,1)	10 0, 0	0,0 (0– 45, 9)	0	100, 0 (54,1 – 100, 0)	8, 2
Fracturas coronaria s	75	97,3 (90,7 – 99,7)	97 ,3	70, 7 (59, 0– 80, 6)	67 ,9	100, 0 (95,2 – 100, 0)	60 ,0
Restaura ciones	6	66,7 (22,3 – 95,7)	10 0, 0	16, 7 (0,4 – 64, 1)	5, 9	16,7 (0,4– 64,1)	7, 7
Implante s	13 7	100,0 (97,3 – 100,0)	10 0, 0	99, 3 (96, 0– 100 ,0)	99 ,3	99,3 (96,0 – 100, 0)	99 ,3
Prótesis fija	12 5	99,2 (95,6 – 100,0)	98 ,4	87, 2 (80, 0– 92, 5)	96 ,5	91,2 (84,8 – 95,5)	95 ,0

En la **Tabla 3** se observó una concordancia perfecta entre el evaluador humano y el gold standard para implantes ($\kappa = 1,000$; $p < 0,001$). Por su parte, Gemini alcanzó una

concordancia considerable con el gold standard ($\kappa = 0,659$; $p < 0,001$) para implantes, al igual que NeuroVet Dental ($\kappa = 0,659$; $p < 0,001$).

Tabla 3. Concordancia de los métodos evaluados con el especialista en radiología oral

Método	Hallazgo	Acuerdo %	Kappa	Interpretación	z	p
Evaluador humano	Caries	96,4	0,277	Aceptable	4,74	<0,001
	Fracturas coronarias	97,1	0,943	Casi perfecta	11,2	<0,001
	Restauraciones	98,6	0,793	Considerable	9,59	<0,001
	Implantes	100,0	1,000	Perfecta	11,8	<0,001
	Prótesis fija	97,9	0,885	Casi perfecta	10,5	<0,001
Gemini	Caries	94,3	-0,022	Sin concordancia	-0,301	0,763
	Fracturas coronarias	66,4	0,323	Aceptable	3,83	<0,001
	Restauraciones	85,0	0,025	Leve	0,347	0,729
	Implantes	98,6	0,659	Considerable	7,80	<0,001
	Prótesis fija	85,7	0,448	Modera da	5,61	<0,001

NeuroVet Dental	Caries	52,1	0,079	Leve	2,40	0,016
	Fracturas coronarias	64,3	0,243	Aceptable	4,40	<0,001
	Restauraciones	87,9	0,050	Leve	0,637	0,524
	Implantes	98,6	0,659	Considerable	7,80	<0,001
	Prótesis fija	87,9	0,447	Modera da	5,35	<0,001

En la **Tabla 4** se observó que el evaluador humano obtuvo una exactitud del 96,4% para caries, seguido de Gemini con un 94,3%, sin diferencias estadísticamente significativas. Para fracturas coronarias, se encontró que el evaluador humano obtuvo una exactitud del 97,1%, seguido de Gemini con un 66,4% y NeuroVet Dental con un 64,3%. En restauraciones el evaluador humano alcanzó un 98,6%, Gemini (85%) y NeuroVet Dental (87,9%). En prótesis fija el evaluador humano alcanzó 97,9 %, frente al 85,7% de Gemini y al 87,9% de NeuroVet Dental.

Tabla 4. Comparación de la exactitud diagnóstica entre los métodos evaluados

Hallazgo	Evalua dor human o %	Gem ini %	Neuro Vet %	Q de Cochr an	p
Caries	96,4	94,3	52,1	99,10 8	<0,0 01
Fracturas coronarias	97,1	66,4	64,3	54,43 8	<0,0 01
Restauraci ones	98,6	85,0	87,9	22,29 6	<0,0 01
Implantes	100,0	98,6	98,6	4,000	0,13 5
Prótesis fija	97,9	85,7	87,9	19,00 0	<0,0 01

En la **Tabla 5** se observa que el evaluador humano presentó un total de 14 clasificaciones incorrectas: 10 correspondieron a falsos positivos (71,4%) y cuatro a falsos negativos (28,6%). Gemini por su parte, registró 98 errores y diagnósticos: 51% falsos positivos y 49% falsos negativos, siendo las fracturas coronarias el hallazgo con más cantidad de errores. NeuroVet Dental, por su parte, fue el método con la mayor cantidad de errores registrados, siendo 153: donde el 88,9%

fueron falsos negativos y el 11,1% falsos positivos.

Tabla 5. Falsos positivos y falsos negativos según método y hallazgo radiográfico

Método	Hallazgo	F P n	F N n	Total de error es	Error predomina nte
Evaluad or humano	Caries	5	0	5	Falso positivo
	Fracturas coronarias	2	2	4	Ambos por igual
	Restauracio nes	2	0	2	Falso positivo
	Implantes	0	0	0	Ninguno
	Prótesis fija	1	2	3	Falso negativo
Subtot al		1 0	4	14	FP: 71,4 %
Gemini	Caries	6	2	8	Falso positivo
	Fracturas coronarias	2 2	25	47	Falso negativo
	Restauracio nes	5	16	21	Falso negativo
	Implantes	1	1	2	Ambos por igual
	Prótesis fija	1 6	4	20	Falso positivo
Subtot al		5 0	48	98	FP: 51,0 %
	Caries	0	67	67	Falso negativo

NeuroVet Dental	Fracturas coronarias	0	50	50	Falso negativo
	Restauraciones	5	12	17	Falso negativo
	Implantes	1	1	2	Ambos por igual
	Prótesis fija	11	6	17	Falso positivo
Subtotal		17	136	153	FN: 88,9 %

DISCUSIÓN

Se evidenciaron diferencias significativas en el desempeño diagnóstico según el tipo de método de evaluación (Gemini, NeuroVet Dental y Evaluador Humano) y el tipo de hallazgos a evaluar en la imagen radiográfica. Los resultados obtenidos de los tres métodos fueron comparados con el gold standard establecido por una especialista en imagenología oral. Se encontró que el evaluador humano fue el que presentó una mayor concordancia con el estándar de referencia que fue la especialista en imagenología oral, especialmente en el hallazgo “fracturas coronarias” con un $k = 0,943$, a este le

siguió prótesis fija con un $k = 0,885$ e implantes con un $k = 1.000$. A diferencia de estos resultados, la respuesta y comportamiento de Gemini y NeuroVet Dental fue variable.

En cuanto a las lesiones cariosas, se encontraron diferencias significativas entre el evaluador humano y ambas IA. El evaluador humano alcanzó una sensibilidad del 100% y una exactitud del 96,4%. Gemini presentó una mayor capacidad en la detección de lesiones cariosas (Se = 98,5%; Acc = 94,3%), sin embargo, la especificidad obtenida fue del 0,0% debido a que clasificó como positivas seis radiografías consideradas negativas por el gold standard. En cambio, NeuroVet Dental presentó un comportamiento opuesto (Se = 50%; Sp = 50%), la exactitud obtenida fue del 52,1% producto de la presencia de 67 falsos positivos. Zhu, et al.¹⁵, evaluó un sistema IA en la detección de caries, obteniendo sensibilidad del 55,4%, una especificidad del 99% y un área bajo la curva de 0,772. Se reportó un mejor

desempeño para hallazgos radiográficamente definidos, tales como coronas completas, restos radiculares y molares impactados, así como piezas ausentes. Otros autores como Batgerel, et al.⁴ evaluaron y compararon la capacidad de un grupo de estudiantes de odontología y de un sistema IA para detectar lesiones cariosas en radiografías panorámicas. El sistema IA alcanzó una sensibilidad del 88,3%, especificidad del 91,7% y una exactitud del 90,1%, superando de forma significativa a los estudiantes evaluados ($p < 0,001$).

Las diferencias entre ambos estudios pueden estar relacionados con el algoritmo empleado, severidad de las lesiones y la unidad de análisis, demostrando que los hallazgos obtenidos con un sistema IA no debe ser generalizado a otras herramientas. Asimismo, pone en evidencia que la detección de lesiones cariosas en radiografías panorámicas continúa siendo una tarea compleja para los sistemas automatizados.

En cuanto a las restauraciones, se encontraron buenos resultados en la detección de estas para Gemini y NeuroVet Dental, sin embargo, la capacidad para reconocer los casos negativos de forma correcta fue baja. Otros autores como Bonfanti-Gris et al.¹⁷, evaluó un sistema IA (Denti.Ai) en 300 radiografías panorámicas, alcanzando una sensibilidad de 85,48% y especificidad del 87,50% en la detección de restauraciones metálicas, mientras que en la detección de restauraciones de resina se presentó una disminución porcentual de la sensibilidad a 41,11%, aunque se mantuvo una especificidad de 93,30%. Estos resultados respaldan que la capacidad de la IA puede verse influenciada según el tipo de factores evaluado, material y características radiográficas.

En cuanto a la detección de implantes dentales, la situación observada fue diferente. La exactitud obtenida por Gemini y NeuroVet Dental fue de 98,6%, sin embargo, su sensibilidad fue de 66,7%,

resultado que debe ser interpretado considerando que en toda la muestra solo existieron tres casos de implantes positivos. Arijji et al.¹⁸, utilizó un modelo YOLOv7, el cual estuvo entrenado específicamente para este propósito, obteniendo una sensibilidad del 100%, precisión del 97,9% y F1-score del 98,9%. En otro artículo realizado por Bonfanti-Gris et al., reportaron un rendimiento del 100% para la detección de implantes con el sistema IA utilizado. Las diferencias descritas pueden estar relacionadas tanto con el pequeño grupo implantes presentes en la muestra como con el entrenamiento específico de cada sistema IA, por lo cual, la elevada exactitud reportada en el presente estudio deberá ser interpretada con cautela.

Para las prótesis fijas, ambos sistemas IA presentaron un menor desempeño en comparación con el evaluador humano. Gemini presentó una sensibilidad de 73,3% y NeuroVet Dental de 60%, con valores predictivos positivos de 40,7% y 45% cada uno. De forma similar, Turosz et al.¹⁹,

reportaron dificultades con las prótesis fijas; en la detección de coronas la precisión fue de 65,3% a pesar de obtener indicadores superiores al 90% para otros hallazgos presentes en la radiografía. Estos hallazgos respaldan que el rendimiento de estos sistemas IA no es uniforme, pudiendo cambiar dependiendo de la estructura o tratamiento que se intenta detectar.

Las diferencias encontradas entre los métodos de evaluación fueron más evidentes en el tipo y cantidad de errores. En cuanto al evaluador humano los errores acumulados fueron 14, Gemini con 98 y NeuroVet Dental con 153, dentro de los cuales 136 correspondieron a falsos negativos, lo cual pone en evidencia un patrón de omisión de hallazgos presentes, especialmente de lesiones cariosas y fracturas. Por su parte, Gemini mostró un comportamiento más equilibrado entre falsos positivos (n = 50) y falsos negativos (n = 48). Estos hallazgos explican la importancia de no considerar de forma independiente la exactitud y especificidad

al momento de establecer la utilidad diagnóstica de una herramienta.

Finalmente, los resultados obtenidos con Gemini son consistentes con lo señalado por Erden et al.²⁰, quienes realizaron una comparación entre Gemini, ChatGPT-4°, Claude Sonnet 4.0 y Copilot en 93 terceros molares inferiores impactados. A pesar de que Gemini presentó una concordancia moderada en ciertos parámetros, ninguno de los sistemas evaluados logró obtener un desempeño comparable con el gold standard (expertos) en las clasificaciones/lesiones/tratamientos evaluados. De forma conjunta, estos resultados demuestran que Gemini y NeuroVet Dental pueden ser utilizados como una herramienta de apoyo en la interpretación radiográfica, teniendo en cuenta que su desempeño depende del hallazgo evaluado y todavía presenta errores que vuelven necesaria una segunda revisión por parte de los profesionales a cargo.

CONCLUSIÓN

Se encontró que ambos sistemas IA evaluados en el presente estudio no obtuvieron una precisión diagnóstica comparable o superior a la interpretación realizada por el evaluador humano. Se hizo evidente que la interpretación realizada por el evaluador humano tuvo mayor concordancia con el estándar de referencia, especialmente al detectar fracturas coronarias, prótesis fija y restauraciones, esto evidencia a su vez un mejor desempeño por parte del profesional en odontología.

A diferencia de esto, se encontró que el desempeño de la IA fue variable según la lesión a detectar o tratamiento a evaluar, como las prótesis fijas o implantes. De ambas, Gemini presentó mejores resultados, comparables con el evaluador humano específicamente en la detección de lesiones cariosas. Se concluye que las IA pueden ser utilizadas como herramientas complementarias en el proceso de análisis e interpretación de radiografías panorámicas. Sin embargo, es importante recalcar que los

resultados arrojados por estos sistemas exigen una valoración profesional previa, ya que no sustituyen el criterio diagnóstico humano.

RECOMENDACIÓN

Para futuras investigaciones se hace la recomendación de incluir una muestra mucho más amplia, además de equilibrar los grupos de casos positivos y negativos, es decir, con y sin los hallazgos establecidos, de esta forma se podrá medir con una precisión mucho más alta la sensibilidad y especificidad de los métodos de evaluación. También se recomienda incluir otros sistemas de inteligencia artificial mencionados en estudios previos.

REFERENCIAS

1. Jagtap R, Samata Y, Parekh A, Tretto P, Vujanovic T, Naik P, et al. Automatic feature segmentation in dental panoramic radiographs. *Sci Prog.* 2024;107(4):368504241286659. doi:10.1177/00368504241286659 PubMed PMID: 39415666; PubMed Central PMCID: PMC11489955.
2. Ibraheem WI. Accuracy of Artificial Intelligence Models in Dental Implant Fixture Identification and Classification from Radiographs: A Systematic Review. *Diagnostics.* enero de 2024;14(8):806. doi:10.3390/diagnostics14080806
3. Turosz N, Chęcińska K, Chęciński M, Brzozowska A, Nowak Z, Sikora M. Applications of artificial intelligence in the analysis of dental panoramic radiographs: an overview of systematic reviews. *Dento Maxillo Facial Radiol.* octubre de 2023;52(7):20230284. doi:10.1259/dmfr.20230284 PubMed PMID: 37665008; PubMed Central PMCID: PMC10552133.
4. Batgerel OE, Akkitap MP, Sasany R. Diagnostic competence of senior dental students in detecting caries on panoramic radiographs with and

- without artificial intelligence assistance: a cross-sectional study on caries detection on panoramic radiographs. *BMC Med Educ.* 28 de noviembre de 2025;25(1):1662. doi:10.1186/s12909-025-08269-2 PubMed PMID: 41316237; PubMed Central PMCID: PMC12661667.
5. Umer F, Habib S, Adnan N. Application of deep learning in teeth identification tasks on panoramic radiographs. *Dento Maxillo Facial Radiol.* 1 de julio de 2022;51(5):20210504. doi:10.1259/dmfr.20210504 PubMed PMID: 35143260; PubMed Central PMCID: PMC10043607.
 6. Mohammad-Rahimi H, Motamedian SR, Rohban MH, Krois J, Uribe SE, Mahmoudinia E, et al. Deep learning for caries detection: A systematic review. *J Dent.* julio de 2022;122:104115. doi:10.1016/j.jdent.2022.104115 PubMed PMID: 35367318.
 7. Khurshid Z, Faridoon F, Silkosessak OCUD, Trachoo V, Waqas M, Hasan S, et al. Multi-Regional deep learning models for identifying dental restorations and prosthesis in panoramic radiographs. *BMC Oral Health.* 5 de noviembre de 2025;25(1):1750. doi:10.1186/s12903-025-07138-0
 8. Asci E, Kilic M, Celik O, Cantekin K, Bircan HB, Bayrakdar İS, et al. A Deep Learning Approach to Automatic Tooth Caries Segmentation in Panoramic Radiographs of Children in Primary Dentition, Mixed Dentition, and Permanent Dentition. *Children.* 5 de junio de 2024;11(6):690. doi:10.3390/children11060690 PubMed PMID: 38929269; PubMed Central PMCID: PMC11202197.

9. Umer F, Habib S, Adnan N. Application of deep learning in teeth identification tasks on panoramic radiographs. *Dentomaxillofacial Radiol.* 1 de julio de 2022;51(5):20210504. doi:10.1259/dmfr.20210504
10. Umer F, Habib S, Adnan N. Application of deep learning in teeth identification tasks on panoramic radiographs. *Dento Maxillo Facial Radiol.* 1 de julio de 2022;51(5):20210504. doi:10.1259/dmfr.20210504 PubMed PMID: 35143260; PubMed Central PMCID: PMC10043607.
11. Chen YW, Stanley K, Att W. Artificial intelligence in dentistry: current applications and future perspectives. *Quintessence Int.* 2020;51(3):248-57. doi:10.3290/j.qi.a43952 PubMed PMID: 32020135.
12. Liu Z, Ai QYH, Yeung AWK, Tanaka R, Nalley A, Hung KF. Performance of a Vision-Language Model in Detecting Common Dental Conditions on Panoramic Radiographs Using Different Tooth Numbering Systems. *Diagnostics.* enero de 2025;15(18):2315. doi:10.3390/diagnostics15182315
13. Turosz N, Chęcińska K, Chęciński M, Brzozowska A, Nowak Z, Sikora M. Applications of artificial intelligence in the analysis of dental panoramic radiographs: an overview of systematic reviews. *Dento Maxillo Facial Radiol.* octubre de 2023;52(7):20230284. doi:10.1259/dmfr.20230284 PubMed PMID: 37665008; PubMed Central PMCID: PMC10552133.
14. Ayhan B, Ayan E, Atsü S. Detection of dental caries under fixed dental prostheses by analyzing digital panoramic radiographs with artificial

- intelligence algorithms based on deep learning methods. BMC Oral Health. 10 de febrero de 2025;25(1):216. doi:10.1186/s12903-025-05577-3 PubMed PMID: 39930440; PubMed Central PMCID: PMC11809006.
15. Zhu J, Chen Z, Zhao J, Yu Y, Li X, Shi K, et al. Artificial intelligence in the diagnosis of dental diseases on panoramic radiographs: a preliminary study. BMC Oral Health. 3 de junio de 2023;23(1):358. doi:10.1186/s12903-023-03027-6
 16. Biçengil K, Kurt A, Naralan ME, Okumuş İ. Deep Learning-Based Dental Caries Diagnosis on Panoramic Radiographies: Performance of YOLOv8 Versus Human Observers. Diagnostics. 13 de abril de 2026;16(8):1150. doi:10.3390/diagnostics16081150 PubMed PMID: 42072778; PubMed Central PMCID: PMC13114539.
 17. Bonfanti-Gris M, Herrera A, Salido Rodríguez-Manzanique MP, Martínez-Rus F, Pradies G. Deep learning for tooth detection and segmentation in panoramic radiographs: a systematic review and meta-analysis. BMC Oral Health. 30 de julio de 2025;25(1):1280. doi:10.1186/s12903-025-06349-9 PubMed PMID: 40739210; PubMed Central PMCID: PMC12312236.
 18. Arijji Y, Kusano K, Fukuda M, Wakata Y, Nozawa M, Kotaki S, et al. Two-step deep learning models for detection and identification of the manufacturers and types of dental implants on panoramic radiographs. Odontology. abril de 2025;113(2):788-98. doi:10.1007/s10266-024-00989-z PubMed PMID: 39198339.
 19. Turosz N, Chęcińska K, Chęciński M, Sielski M, Sikora M. Evaluation of Dental Panoramic Radiographs by

Artificial Intelligence Compared to Human Reference: A Diagnostic Accuracy Study. J Clin Med. enero de 2024;13(22):6859.
doi:10.3390/jcm13226859

Diagnostic accuracy of AI models in classifying impacted mandibular third molars. Odontology. 1 de julio de 2026;114(3):1533-41.
doi:10.1007/s10266-025-01214-1

20. Erden MB, Kanmaz MG, Sabah GA.
Can chatbots replace experts?

DECLARACIÓN Y AUTORIZACIÓN

Yo, **Jarro Jaramillo, Kerly Arlen** con C.C: **0952502748** autora del trabajo de titulación: **Precisión diagnóstica de la inteligencia artificial comparada con la interpretación humana en radiografías panorámicas**, previo a la obtención del título de **Odontólogo** en la Universidad Católica de Santiago de Guayaquil.

1.- Declaro tener pleno conocimiento de la obligación que tienen las instituciones de educación superior, de conformidad con el Artículo 144 de la Ley Orgánica de Educación Superior, de entregar a la SENESCYT en formato digital una copia del referido trabajo de titulación para que sea integrado al Sistema Nacional de Información de la Educación Superior del Ecuador para su difusión pública respetando los derechos de autor.

2.- Autorizo a la SENESCYT a tener una copia del referido trabajo de titulación, con el propósito de generar un repositorio que democratice la información, respetando las políticas de propiedad intelectual vigentes.

Guayaquil, 7 de septiembre del 2026

Kerly Jarro

f. _____
Jarro Jaramillo, Kerly Arlen
C.C: 0952502748

REPOSITORIO NACIONAL EN CIENCIA Y TECNOLOGÍA

FICHA DE REGISTRO DE TESIS/TRABAJO DE TITULACIÓN

TEMA Y SUBTEMA:	Precisión diagnóstica de la inteligencia artificial comparada con la interpretación humana en radiografías panorámicas.		
AUTOR(ES)	Jarro Jaramillo, Kerly Arlen		
REVISOR(ES)/TUTOR(ES)	Dra. Díaz Rojas, Dennisse Fernanda		
INSTITUCIÓN:	Universidad Católica de Santiago de Guayaquil		
FACULTAD:	Facultad de Ciencias de la Salud		
CARRERA:	Carrera de Odontología		
TÍTULO OBTENIDO:	Odontólogo		
FECHA DE PUBLICACIÓN:	7 de septiembre de 2026	No. DE PÁGINAS:	18
ÁREAS TEMÁTICAS:	Radiografías panorámicas, Inteligencia artificial, Diagnóstico Asistido por Computador, Diagnóstico Odontológico.		
PALABRAS CLAVES/KEYWORDS:	Inteligencia Artificial; Radiografía Panorámica; Diagnóstico por Imagen; Sensibilidad y Especificidad; Odontología.		

RESUMEN/ABSTRACT *La inteligencia artificial se ha incorporado al análisis de imágenes odontológicas como herramienta de apoyo para identificar diferentes hallazgos. Sin embargo, su desempeño puede variar según el sistema utilizado y el tipo de alteración evaluada, por lo que es necesario compararla con la interpretación profesional. **Objetivo:** Determinar si la precisión diagnóstica de los sistemas de inteligencia artificial Gemini y NeuroVet Dental es comparable o superior a la interpretación humana en radiografías panorámicas odontológicas. **Materiales y métodos:** Estudio cuantitativo, observacional, retrospectivo y de exactitud diagnóstica. Se analizaron 140 radiografías panorámicas digitales mediante Gemini, NeuroVet Dental y una especialista en odontología. Se evaluó la presencia de caries, fracturas coronarias, restauraciones, implantes y prótesis fija, utilizando como patrón de referencia la interpretación de una especialista en Radiología Oral. Se calcularon sensibilidad, especificidad, exactitud, valores predictivos y concordancia. **Resultados:** La evaluación humana presentó la mayor exactitud y concordancia en la mayoría de los hallazgos. Gemini obtuvo resultados comparables para caries, con una exactitud del 94,3% frente al 96,4% de la evaluadora humana. En implantes no se encontraron diferencias significativas entre los métodos ($p=0,135$). En fracturas, restauraciones y prótesis fija, la evaluación humana fue significativamente más exacta que ambos sistemas de inteligencia artificial ($p<0,001$). **Conclusión:** Gemini y NeuroVet Dental no alcanzaron una precisión diagnóstica comparable o superior a la interpretación humana. Su desempeño dependió del hallazgo evaluado, por lo que pueden utilizarse como herramientas de apoyo, pero requieren supervisión profesional.*

ADJUNTO PDF:	☒ SI	☐ NO
CONTACTO CON AUTOR/ES:	Teléfono: +593 96 296 2307	E-mail: kerly.jarro@cu.ucsg.edu.ec
CONTACTO CON LA INSTITUCIÓN (COORDINADOR DEL PROCESO UTE)::	Nombre: Ocampo Poma, Estefanía Del Rocio	
	Teléfono: +593 99 675 7081	
	E-mail: estefania.ocampo@cu.ucsg.edu.ec	

SECCIÓN PARA USO DE BIBLIOTECA

Nº. DE REGISTRO (en base a datos):	
Nº. DE CLASIFICACIÓN:	
DIRECCIÓN URL (tesis en la web):	