



ARTÍCULO ORIGINAL



**Díaz-Flores García, Víctor**  
Profesor Contratado Doctor. Departamento de Odontología Pre-Clinica I. Facultad de Ciencias Biomédicas y de la Salud. Universidad Europea de Madrid.

**García Moneo, Natalia**  
Profesora de Grado en Odontología. Departamento de Odontología Pre-Clinica II. Facultad de Ciencias Biomédicas y de la Salud. Universidad Europea de Madrid.

**Llorente de Pedro, María**  
Profesora de Grado en Odontología. Departamento de Odontología Pre-Clinica II. Facultad de Ciencias Biomédicas y de la Salud. Universidad Europea de Madrid.

**Freire, Yolanda**  
Profesora Contrada Doctor. Departamento de Odontología Pre-Clinica II. Facultad de Ciencias Biomédicas y de la Salud. Universidad Europea de Madrid.

**Suárez, Ana**  
Profesora Titular de Odontología. Vicedecana de Odontología. Facultad de Ciencias Biomédicas y de la Salud. Universidad Europea de Madrid.

#### correspondencia:

Natalia García Moneo.  
Facultad de Ciencias Biomédicas y de la Salud. Universidad Europea de Madrid. C. Tajo, s/n, 28670 Villaviciosa de Odón, Madrid.  
natalia.garcia@universidadeuropea.es

Fecha de recepción: 5 de marzo de 2026.  
Fecha de aceptación para su publicación: 14 de abril de 2026.

ARTÍCULO GANADOR PREMIO CIENTÍFICO ANUAL  
"PROFESOR DR. D. PEDRO RUIZ DE TEMIÑO MALO" 2025

## DETECCIÓN DE MANIPULACIONES RADIOGRÁFICAS EN ENDODONCIA: LA INTELIGENCIA ARTIFICIAL EN LA LUCHA CONTRA EL FRAUDE RADIOGRÁFICO

Díaz-Flores García V, García Moneo N, Llorente de Pedro M, Freire Y, Suárez A.  
Detección de manipulaciones radiográficas en Endodoncia: la inteligencia artificial en la lucha contra el fraude radiográfico.  
Cient. Dent. 2026; 23; 1; 18-24

### RESUMEN

**Introducción:** La radiografía digital ha mejorado la precisión diagnóstica en odontología, pero también ha introducido el riesgo de manipulaciones indetectables, con posibles implicaciones clínicas, forenses y legales. Este estudio evaluó la capacidad de ChatGPT-4o para diferenciar entre radiografías intraorales originales y radiografías alteradas digitalmente.

**Métodos:** Ocho radiografías se modificaron intencionadamente para simular alteraciones frecuentes y se analizaron junto con sus originales. Cada par se evaluó mediante ChatGPT-4o en 30 sesiones independientes (240 consultas en total), calculando la exactitud diagnóstica y la repetibilidad intramodelo, esta última mediante los coeficientes kappa de Fleiss, AC1 de Gwet y alfa de Krippendorff.

**Resultados:** ChatGPT-4o alcanzó una exactitud global del 37,4% (IC 95%: 30,9–44,0), significativamente inferior a lo esperado por azar ( $p < 0,001$ ) y menor que el 56% reportado previamente para observadores humanos. El rendimiento varió notablemente entre los pares de imágenes (0%–100%). A pesar de la baja exactitud, los índices de repetibilidad indicaron una consistencia interna moderada a sustancial, lo que sugiere respuestas estables, pero frecuentemente incorrectas. ChatGPT-4o no logró detectar de forma fiable la manipulación radiográfica y obtuvo un rendimiento inferior al de un clínico.

**Conclusiones:** Aunque el modelo muestra potencial como herramienta accesible, su exactitud actual resulta insuficiente para aplicaciones clínicas, forenses o académicas. Estos hallazgos subrayan la necesidad de conjuntos de datos más amplios, entrenamiento multimodal e integración con sistemas de verificación consolidados para mejorar su fiabilidad.

**Palabras clave:** Radiografía dental; Endodoncia; Reconocimiento de Patrones Automatizado; ChatGPT; Fraude; Seguridad del paciente.

### DETECTING RADIOGRAPHIC MANIPULATION IN ENDODONTICS: ARTIFICIAL INTELLIGENCE IN THE FIGHT AGAINST RADIOGRAPHIC FRAUD

#### ABSTRACT

**Introduction:** Digital radiography has enhanced diagnostic accuracy in dentistry but introduced the risk of undetectable image manipulation, with potential clinical, forensic, and legal consequences. This study examined the ability of ChatGPT-4o to differentiate between original intraoral radiographs and digitally altered counterparts.

**Methods:** Eight radiographs were deliberately modified to replicate common alterations and assessed alongside their originals. Each pair was analysed by ChatGPT-4o in 30 independent sessions (240 queries), and diagnostic accuracy was calculated. Intramodel repeatability was evaluated using Fleiss' kappa, Gwet's AC1, and Krippendorff's alpha.

**Results:** ChatGPT-4o achieved an overall accuracy of 37.4% (95% CI: 30.9–44.0), significantly lower than chance ( $p < 0.001$ ) and below the 56% accuracy previously reported for human observers. Performance varied substantially across image pairs (0%–100%). Despite poor accuracy, repeatability indices indicated moderate to substantial internal consistency, suggesting stable yet frequently incorrect responses. ChatGPT-4o was unable to reliably detect radiographic manipulation and performed worse than clinicians.

**Conclusions:** Although the model shows promise as an accessible tool, its current accuracy is inadequate for clinical, forensic, or academic use. These findings highlight the need for larger datasets, multimodal training, and integration with established verification systems to improve reliability.

**Key words:** Dental Radiography; Endodontics; Pattern Recognition Automated; ChatGPT; Fraud; Patient Safety.

## INTRODUCCIÓN

La radiografía es una herramienta esencial en endodoncia, ya que proporciona información clave para el diagnóstico y el seguimiento de las estructuras dentarias y óseas que no son accesibles mediante la exploración clínica. En muchos casos, se convierte en la única vía para revelar patologías ocultas o justificar la necesidad de tratamiento<sup>1</sup>.

El desarrollo de la radiografía digital, inicialmente mediante la digitalización de la película convencional y posteriormente a través de sensores directos, transformó radicalmente la práctica clínica al permitir una adquisición de imágenes más rápida, una menor exposición a la radiación y una mejor gestión del almacenamiento y la preservación. También introdujo la posibilidad de procesar las imágenes para optimizar su visualización, ampliando así las capacidades diagnósticas del profesional<sup>2</sup>.

Sin embargo, esta evolución tecnológica trajo consigo un nuevo desafío. El formato digital de las imágenes abrió la puerta a un riesgo significativo: la posible manipulación intencionada de su contenido. Una radiografía que antes se percibía como un registro objetivo e inmutable puede hoy alterarse con relativa facilidad. Una edición maliciosa podría modificar diagnósticos, pronósticos y planes de tratamiento, así como ocultar errores clínicos o incluso simular procedimientos nunca realizados<sup>3,4</sup>.

Hoy en día, las herramientas de edición de imágenes están al alcance de cualquier usuario. Incluso de forma gratuita, con software de código abierto como GNU Image Manipulation Program (GIMP; The GIMP Development Team, Free Software Foundation, Boston, MA, EE. UU.), es posible añadir, eliminar o modificar estructuras con un alto grado de indetectabilidad<sup>5,6</sup>.

Las consecuencias de estas alteraciones van mucho más allá del ámbito clínico: pueden influir de forma determinante en litigios, reclamaciones o procesos de acreditación profesional<sup>4,7</sup>. Y aunque en radiología médica se han desarrollado protocolos de autenticación, en odontología no existen estándares ampliamente implantados que garanticen la integridad de las imágenes.

Esta laguna en el control convierte la manipulación digital en un problema global. Trasciende la práctica clínica y se inscribe en un escenario forense más amplio, donde la manipulación digital constituye un desafío constante, descrito como una “carrera armamentística” entre los falsificadores y las técnicas diseñadas para detectarlos<sup>8</sup>. En este contexto, los métodos de detección pasiva, que operan únicamente con la información contenida en la propia imagen, adquieren especial relevancia, ya que en odontología no es habitual insertar marcas de agua especiales ni metadatos<sup>8,9</sup>.

La magnitud del problema ha quedado reflejado en estudios previos. Ya en 1997 se reconoció la dificultad de los profesionales de la odontología para identificar radiografías manipuladas<sup>10</sup>, y, veinte años después, un análisis más amplio mostró que incluso odontólogos con experiencia solo alcanzaron una tasa de acierto del 56% en la detección de alteraciones. En otras palabras, casi la mitad de las falsificaciones pasaron desapercibidas<sup>11</sup>.

Ante este panorama, en los últimos años han surgido nuevas herramientas, como los modelos de lenguaje a gran escala (LLM) con capacidades de análisis multimodal, capaces de procesar texto e imágenes de forma simultánea. Estos sistemas, están siendo evaluados como posibles herramientas de interés en la detección de alteraciones en imágenes médicas y odontológicas. Varios estudios han demostrado su utilidad, lo que abre una vía prometedora para reforzar la autenticidad de las radiografías digitales<sup>12</sup>.

No se ha explorado aún de forma sistemática si un Large Language Model (LLM), sin entrenamiento específico para la tarea, es capaz de identificar alteraciones radiográficas y con qué grado de reproducibilidad<sup>11</sup>. Por ello, este estudio se diseñó como una prueba de concepto, no con el objetivo de entrenar el modelo, sino de examinar su capacidad inicial en la detección de modificaciones, así como analizar la estabilidad de su rendimiento en la detección de radiografías intraorales manipuladas. Este enfoque pretende contribuir al desarrollo de estrategias objetivas y reproducibles que garanticen la autenticidad de las imágenes en contextos clínicos, jurídicos y forenses.

## MATERIAL Y MÉTODOS

Se llevó a cabo un estudio observacional y repetible de pruebas diagnósticas diseñado para evaluar la capacidad y la repetibilidad de un modelo de lenguaje basado en inteligencia artificial, ChatGPT (versión 4o), en la identificación de imágenes radiográficas originales frente a imágenes modificadas digitalmente.

El estudio se concibió como una prueba piloto de concepto. En lugar de maximizar el número de casos, el diseño priorizó medidas repetidas sobre una muestra más pequeña, generando así datos suficientes para estimar la exactitud con intervalos de confianza del 95% y calcular múltiples coeficientes de fiabilidad. Este enfoque nos permitió caracterizar tanto el rendimiento como la reproducibilidad a pesar del número limitado de pares de imágenes.

### 2.1 Comité ético

El estudio fue aprobado por el Comité de Ética de la Universidad Europea de Madrid (código de aprobación 2025-56) y se llevó a cabo de acuerdo con los principios éticos de la Declaración de Helsinki. Todos los participantes fueron informados de los objetivos del estudio y otorgaron su

consentimiento informado por escrito. Para garantizar la protección de datos, todas las sesiones se realizaron utilizando la función de chat temporal de ChatGPT, que impide el almacenamiento permanente de las conversaciones<sup>13</sup>.

## 2.2 Diseño del estudio

Se seleccionaron ocho radiografías intraorales de diagnóstico o tratamiento endodóntico a partir de las historias clínicas disponibles. Las radiografías se obtuvieron de la base de datos de la Clínica Universitaria Odontológica de la Universidad Europea de Madrid. El dispositivo de adquisición digital utilizado fue un VistaScan Mini (Dürr Dental SE, Bietigheim-Bissingen, Alemania).

Cada radiografía se exportó a formato JPG para su manipulación mediante el software Adobe Photoshop 2021 (Adobe Inc., San José, CA, EE. UU.). Se realizó una modificación diferente en cada radiografía. Las modificaciones efectuadas en cada radiografía se detallan en la Tabla.

### Tabla. MODIFICACIONES realizadas en cada radiografía

Número de radiografía intraoral	Modificación realizada
1	Eliminación de la extrusión de cemento endodóntico.
2	Sustitución de la zona radiolúcida en una raíz tratada endodónticamente y preparada para perno por una radiopacidad similar a la de la obturación endodóntica.
3	Relleno radiolúcido en el lado distal de la raíz con una imagen de densidad similar a la del hueso alveolar.
4	Relleno radiolúcido en el lado mesial de la raíz con una imagen de densidad similar a la del hueso alveolar.
5	Creación de una imagen radiolúcida compatible con caries en la zona distal de una corona.
6	Imagen radiolúcida compatible con caries en la zona distal de una corona con apariencia similar a dentina y esmalte.
7	Imagen radiolúcida compatible con caries en la zona distal de una corona con apariencia similar a dentina.
8	Los conductos radiculares de un molar se rellenan con un material compatible con el de obturación endodóntica, con parte del material extruido hacia la zona periapical.

Las imágenes modificadas se importaron en el software de gestión de radiografías VistaSoft (Dürr Dental SE, Bietigheim-Bissingen, Alemania) y las imágenes originales y modificadas se volvieron a importar en formato Digital Imaging and Communication in Medicine (DICOM), ya que este formato permite verificar datos como los parámetros

de adquisición, las fechas e incluso diferencias de calidad/contraste (mediante conversión directa a JPG en ChatGPT para su visualización).

Cada par de radiografías se subió de forma independiente al chatbot, solicitándole que identificara cuál de las dos radiografías presentadas correspondía a la radiografía original, para lo cual se utilizó el siguiente prompt: "De estas dos radiografías dentales, una es la original y la otra está modificada digitalmente. Necesito que me digas cuál es la original y cuál es la modificada. Lee manualmente los archivos DICOM en formato binario para comparar las imágenes y los metadatos e identificar cuál es la original y cuál está modificada y, si necesitas más información, conviértelas a JPG para continuar el análisis."

Para minimizar posibles sesgos derivados de la memoria o del contexto conversacional, cada par se evaluó en 30 sesiones independientes ("nueva conversación"), distribuidas en distintos momentos del día, siguiendo la metodología descrita por Freire y cols.<sup>14</sup>. En total, se realizaron 240 registros (8 pares × 30 repeticiones). Algunos ejemplos de las radiografías modificadas pueden observarse en la Figura 1.

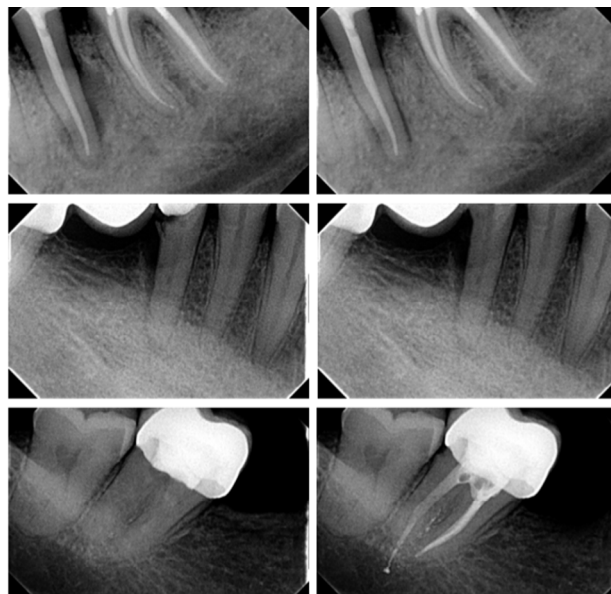


Figura 1. Ejemplos de radiografías originales (izquierda) y modificadas (derecha).

## 2.3 Análisis estadístico

La exactitud diagnóstica se evaluó globalmente y por par de imágenes calculando la proporción de aciertos (respuestas correctas/30 repeticiones) y sus intervalos de confianza del 95% (IC95%) mediante el método binomial de Wald, según la metodología previamente empleada por Freire y cols.<sup>14</sup>. La comparación con el valor esperado por

azar (0,50) se realizó mediante prueba binomial bilateral, siguiendo el enfoque descrito por Díaz-Flores y cols.<sup>11</sup> Se analizaron tanto la exactitud global (porcentaje de aciertos considerando todos los pares) como la exactitud por par (porcentaje de aciertos calculado individualmente para cada par).

La repetibilidad intramodelo se estimó a partir de las 30 repeticiones de cada par y de forma global, utilizando las siguientes métricas: porcentaje de acuerdo global, kappa de Fleiss/Conger, coeficiente de Gwet (AC1) y alfa de Krippendorff. La interpretación se realizó según la escala de Gwet (0,0–0,2 pobre; 0,2–0,4 leve; 0,4–0,6 moderada; 0,6–0,8 sustancial; 0,8–1,0 casi perfecta).

Por último, para contextualizar el rendimiento del modelo, se realizó una comparación indirecta con el referente humano previamente publicado (56% de aciertos)<sup>11</sup>, lo que permitió situar los resultados obtenidos en relación con un estándar de referencia clínicamente relevante.

## RESULTADOS

### 3.1. Exactitud

La exactitud global obtenida por el modelo fue del 37,4% (IC95%: 30,9–44,0%), una diferencia estadísticamente significativa respecto al 50% esperado por azar (prueba binomial bilateral,  $p < 0,001$ ). El análisis por par de imágenes ( $n = 8$ ) mostró una marcada heterogeneidad, con una mediana de exactitud del 30,0% (rango: 0–100%), como se muestra en la Figura 2.

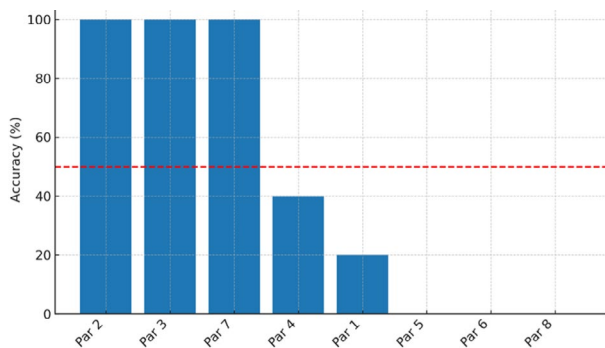


Figura 2. Exactitud global (accuracy) por par de imágenes.

La Figura 3 muestra la estimación puntual y el intervalo de confianza de la exactitud para cada par. Se observa que algunos pares fueron identificados correctamente en todas las repeticiones, mientras que otros no registraron ningún acierto.

Como referencia contextual, Díaz-Flores y cols.<sup>11</sup> indicaron que dentistas con experiencia identificaron correcta-

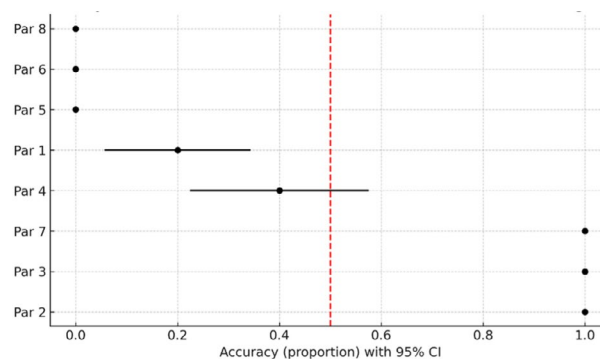


Figura 3. Estimaciones de la exactitud con IC del 95 % por cada par de imágenes.

mente radiografías manipuladas en el 56% de los casos. En comparación con este valor de referencia, la exactitud global de ChatGPT-4o (del 37,4%) indica un rendimiento sustancialmente inferior, a pesar de su elevada consistencia interna.

### 3.1. Repetibilidad

La concordancia media observada (P) fue de 0,772. La repetibilidad global, evaluada mediante el coeficiente kappa de Fleiss, alcanzó un valor de 0,61, lo que corresponde a un grado sustancial de concordancia. El coeficiente global de Gwet (AC1) fue de 0,57, interpretado como concordancia moderada, mientras que el alfa de Krippendorff fue de 0,77, lo que indica una alta consistencia en las respuestas.

A nivel de cada par, la repetibilidad intramodelo mostró valores muy elevados para el coeficiente AC1 (mediana: 1,00), lo que refleja que el modelo tendió a mantener invariante su respuesta a lo largo de las 30 repeticiones, con independencia de su corrección. La relación entre la exactitud y el coeficiente AC1 por par se muestra en la Figura 4, evidenciando que un alto grado de consistencia no implica necesaria-

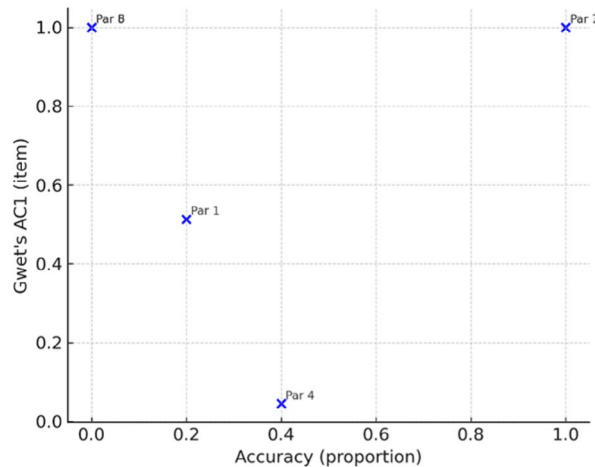


Figura 4. Relación entre la exactitud (accuracy) y la repetibilidad intramodelo.

mente mayor exactitud. En esta figura solo se analizan los pares con valores calculables tanto para la exactitud como para el AC1 de Gwet (pares 1, 3, 4 y 8).

Por el contrario, los valores de kappa por pares presentaron una mediana de aproximadamente  $-0,03$ , resultado atribuible a la conocida sensibilidad de esta métrica a distribuciones extremadamente desequilibradas entre categorías.

En el estudio de referencia con evaluadores humanos<sup>11</sup>, la tasa global de aciertos fue del 56,2 % (371/660 decisiones). En comparación, el modelo evaluado alcanzó el 37,4 % (79/211 decisiones), con una diferencia absoluta de  $-18,8$  puntos porcentuales. El contraste estadístico entre ambas proporciones mostró una diferencia significativa ( $z = -4,75$ ;  $p = 2,0 \times 10^{-6}$ ), evidenciando que el rendimiento de ChatGPT-4o fue sustancialmente inferior al observado en el grupo de referencia humano (Figura 5).

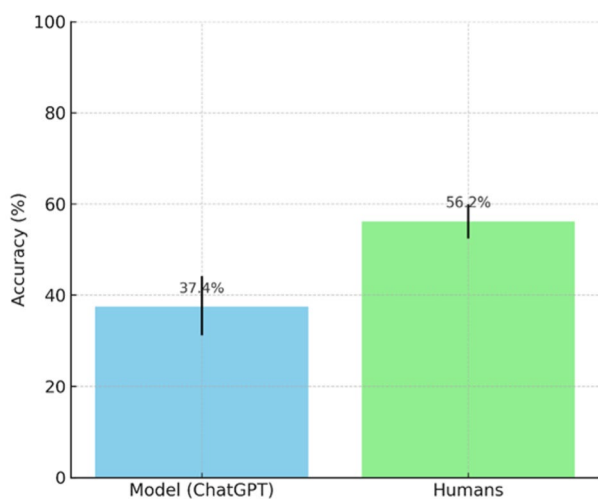


Figura 5. Comparación de la exactitud (accuracy) entre el modelo humano y ChatGPT.

## DISCUSIÓN

Los resultados demuestran que, aunque ChatGPT produjo respuestas consistentes en las repeticiones, su exactitud se mantuvo en un nivel claramente inferior a la de expertos humanos en estudios previos. Esto sugiere que el modelo tiende a reproducir errores sistemáticos más que a lograr clasificaciones válidas. El adiestramiento de los modelos requiere entrenamiento específico, recursos computacionales y pericia técnica, mientras que el enfoque analizado en esta investigación se apoya en un modelo de chatbot de uso general, sin entrenamiento adicional que, por un lado, ofrece mayor accesibilidad, pero genera problemas a costa de la exactitud.

En este contexto, la comparación con evaluadores humanos resulta especialmente relevante. La capacidad reducida encontrada en el modelo analizado contrasta con el rendimiento humano documentado previamente en el estudio de referencia, donde odontólogos con experiencia identificaron imágenes manipuladas en el 56% de los casos<sup>11</sup>. Aunque esta cifra pueda considerarse modesta, refleja una superioridad estadísticamente significativa frente a ChatGPT, lo que confirma que, en el estado actual de desarrollo, el modelo analizado no puede sustituir la pericia clínica ni garantizar una detección fiable de las alteraciones en radiografías dentales.

La paradoja metodológica observada —alta consistencia de las respuestas acompañada de baja exactitud— apunta a la presencia de un sesgo sistemático en la forma en que el modelo interpreta las imágenes. En términos prácticos, esto indica que el sistema responde de manera estable, pero lo hace en la dirección equivocada con excesiva frecuencia. Este patrón es coherente con lo descrito en otros contextos odontológicos, donde los chatbots mantienen una consistencia formal en sus respuestas, pero cometen errores clínicamente relevantes<sup>15</sup>.

La necesidad de herramientas eficaces para la detección de imágenes manipuladas es incuestionable en el escenario actual. La literatura ha documentado cómo los programas de edición de uso común permiten realizar modificaciones prácticamente indetectables, desde alteraciones estéticas hasta fraudes, los cuales son determinantes en endodoncia, ya que la radiografía es, en muchas ocasiones, la única prueba que se puede ofrecer de que se ha realizado este tipo de tratamientos<sup>6,7,16</sup>. Incluso se han comunicado casos en los que se remitieron imágenes alteradas a compañías médicas y fueron aceptadas como válidas<sup>6</sup>.

En radiología dental, se ha demostrado que incluso profesionales con experiencia pueden no reconocer adecuadamente las manipulaciones. El problema trasciende el ámbito clínico, ya que la proliferación de técnicas de falsificación digital, incluidos los conocidos como *deepfakes*, entraña riesgos sustanciales en contextos jurídicos, forenses y académicos. En este escenario, la manipulación de los registros radiográficos podría emplearse en reclamaciones legales, en el encubrimiento de situaciones iatrogénicas o en la generación de pruebas fraudulentas en publicaciones científicas, comprometiendo tanto la seguridad del paciente como la credibilidad de la investigación<sup>7,17</sup>.

En este contexto, los chatbots se presentan como una alternativa conceptualmente atractiva frente a los sistemas especializados de detección. En comparación con técnicas forenses de alto coste y acceso restringido, los modelos de lenguaje ofrecen accesibilidad, escalabilidad y bajos requisitos técnicos, lo que abre la posibilidad de aplicaciones

en entornos clínicos o académicos rutinarios sin necesidad de software propietario avanzado<sup>3,8,9</sup>. Sin embargo, los resultados obtenidos en este estudio muestran que su rendimiento actual no alcanza el nivel mínimo requerido para tareas en las que la exactitud es crítica.

Este trabajo debe interpretarse a la luz de varias limitaciones. El número de pares de imágenes fue reducido, lo que limita la generalización de los resultados. Se empleó un único modelo de lenguaje en una versión específica, sin comparación con otras arquitecturas o modalidades de entrenamiento. Las radiografías correspondían a un único tipo de proyección intraoral, lo que restringe la extrapolación a otras modalidades de imagen médica u odontológica. Además, la comparación con humanos se basa en datos previamente publicados y no en un experimento paralelo bajo condiciones idénticas, lo que introduce posibles sesgos indirectos.

A pesar de estas limitaciones, el estudio aporta nueva evidencia sobre la posibilidad de aplicar chatbots a un problema creciente en la práctica clínica y académica: la

detección de imágenes manipuladas. Los hallazgos sugieren que, aunque la exactitud sigue siendo insuficiente, la consistencia observada en las respuestas refleja un comportamiento sistemático que podría optimizarse mediante ajustes de entrenamiento, la implementación de algoritmos multimodales o la integración con herramientas consolidadas de análisis radiográfico. Las investigaciones futuras deberían ampliar el número y tipo de imágenes evaluadas, incluir comparaciones directas con profesionales humanos en condiciones homogéneas y explorar la combinación de chatbots con herramientas específicas de verificación digital.

## CONCLUSIONES

En conjunto, los resultados muestran que, aunque los modelos de lenguaje ofrecen potencial como herramientas de apoyo accesibles, su uso clínico, legal o académico requiere mejoras sustanciales. El avance continuo de las técnicas de manipulación digital hace necesario reforzar los sistemas de detección si se desea salvaguardar la fiabilidad de los registros médicos y la integridad científica.



## BIBLIOGRAFÍA

1. White SC, Heslop EW, Hollender LG, Mosier KM, Ruprecht A, Shrout MK. Parameters of radiologic care: An official report of the American Academy of Oral and Maxillofacial Radiology. *Oral Surg Oral Med Oral Pathol Oral Radiol Endodontology*. 2001;91(5):498–511.
2. Makkissi J, Pawar R. Digital radiography in the dental practice: An update. *Prim Dent J*. 2013;2(1):58–65.
3. Maruo IT, Maruo H. Digital signature of electronic dental records. *Am J Orthod Dentofacial Orthop*. 2012;141(5):662–665.
4. Madhan B, Gayathri H. Identification and prevention of digital forgery in orthodontic records. *Am J Orthod Dentofacial Orthop*. 2010;138(6):850–857.
5. Biagioli M. Before and after Photoshop: Recursive fraud in the age of digital reproducibility. *Angew Chem Int Ed*. 2019;58(46):16334–16335.
6. Tsang A, Sweet D, Wood RE. Potential for fraudulent use of digital radiography. *J Am Dent Assoc*. 1999;130(9):1325–1329.
7. Calberson FLG, Hommez GM, De Moor RJ. Fraudulent use of digital radiography: methods to detect and protect digital radiographs. *J Endod*. 2008;34(5):530–536.
8. Amiri E, Mosallanejad A, Sheikahmadi A. The optimal model for copy-move forgery detection in medical images. *J Med Signals Sens*. 2024;14(2). doi:10.4103/jmss.jmss\_35\_22
9. Qazi T, Ali M, Hayat K, Magnier B. Seamless copy-move replication in digital images. *J Imaging*. 2022;8(3):69.
10. Visser H, Krüger W. Can dentists recognize manipulated digital radiographs? *Dentomaxillofac Radiol*. 1997;26(1):67–69.
11. Díaz-Flores-García V, Labajo-González E, Santiago-Sáez A, Perea-Pérez B. Detecting the manipulation of digital clinical records in dental practice. *Radiography*. 2017;23(4):e103–107.
12. Suárez A, Arena S, Herranz Calzada A, Castillo Varón AI, Diaz-Flores García V, Freire Y. Decoding wisdom: Evaluating ChatGPT's accuracy and reproducibility in analyzing orthopantomographic images for third molar assessment. *Comput Struct Biotechnol J*. 2025;28:141–147.
13. OpenAI Help Center [Internet]. [cited 2025]. Temporary Chat FAQ. Available from: <https://help.openai.com/en/articles/8914046-temporary-chat-faq?stream=top>
14. Freire Y, Santamaría Laorden A, Orejas Pérez J, Gómez Sánchez M, Díaz-Flores García V, Suárez A. ChatGPT performance in prosthodontics: Assessment of accuracy and repeatability in answer generation. *J Prosthet Dent*. 2024;131(4):659.e1-659.e6.
15. Freire Y, Santamaría Laorden A, Orejas Pérez J, Ortiz Collado I, Gómez Sánchez M, Thuissard Vasallo IJ, et al. Evaluating the influence of prompt formulation on the reliability and repeatability of ChatGPT in implant-supported prostheses. Kolahi J, editor. *PLOS One*. 2025;20(5):e0323086.
16. Landolt S, Hirschel J, Schlienger T, Businger W, Zbinden AM. Assessing and comparing information security in swiss hospitals. *Interact J Med Res*. 2012;1(2):e11.
17. Chang HJ, Oh SN, Park MY, Rha SE, Choi BG. Fraudulent retouching of digital radiographic images – a potential risk. *Clin Radiol*. 2010;65(12):967–973.