



PRISMA ODS
REVISTA MULTIDISCIPLINARIA
SOBRE DESARROLLO SOSTENIBLE

**DISEÑO DE UN INSTRUMENTO
PARA LA EVALUACIÓN DE
IMÁGENES EN ENDOSCOPIA Y
LAPAROSCOPIA PARA
INTELIGENCIA ARTIFICIAL
(ENDOLAP-IA)**

*INSTRUMENT DESIGN FOR
EVALUATION OF IMAGES FROM
ENDOSCOPY AND LAPAROSCOPY
FOR ARTIFICIAL INTELLIGENCE
(ENDOLAP-IA)*

AUTOR

*NORMAN ALEJANDRO RENDÓN MEJÍA
UNIVERSIDAD AUTÓNOMA DE CHIHUAHUA -
HOSPITAL GENERAL - SALVADOR ZUBIRÁN
ANCHONDO
CHIHUAHUA - MÉXICO

Volumen 3 - Número 1
2024

Diseño de un Instrumento para la Evaluación de Imágenes en Endoscopia y Laparoscopia para Inteligencia Artificial (ENDOLAP-IA)

Instrument Design for Evaluation of Images from Endoscopy and Laparoscopy for Artificial Intelligence (ENDOLAP-IA)

Norman Alejandro Rendón Mejía

norman-rendon@hotmail.com

<https://orcid.org/0000-0003-2998-0981>

Universidad Autónoma de Chihuahua – Hospital General - Salvador Zubirán Anchondo
Chihuahua - México

Artículo recibido: 19 de abril del 2024
Aceptado para publicación: 22 de mayo 2024
Conflictos de Intereses: Ninguno que declarar

RESUMEN

El aumento en el número de publicaciones relacionadas con el uso de inteligencia artificial (IA) en los últimos años ha incrementado el interés por el usar esta novedosa tecnología en desarrollo de nuevos métodos diagnósticos. El uso de imágenes para la elaboración de algoritmos de IA conlleva una serie de pasos complejos, que en aquellos no familiarizados en el área de las ciencias computacionales puede ocasionar confusión. Se trata de un estudio instrumental para la elaboración de una lista de cotejo, para delimitar los pasos y un control de calidad del material empleado para la elaboración de algoritmos de IA en la detección automática de objetos. Se realizó el presente estudio en 2 fases: validez de facie y validez de contenido. Se enviaron a 3 expertos un cuestionario con 20 ítems para valorar la pertinencia y comprensión de los ítems. Para la validez de contenido se reenvió el cuestionario a 5 expertos para valorar su pertinencia y redacción, se mantuvieron 19 ítems para la versión final del instrumento. El instrumento presentado "ENDOLAP-IA", es el primero en su clase, el crear instrumentos como el presente, facilitan la introducción e integración a aplicar estas nuevas tecnologías en el área de la medicina.

Palabras clave: laparoscopia, inteligencia artificial, imagen, estudio instrumental, lista de cotejo

ABSTRACT

The increase in the number of publications related to the use of artificial intelligence (AI) in recent years has increased interest in using this novel technology in the development of new diagnostic methods. The use of images for the development of AI algorithms involves a series of complex steps, which can be confusing for those unfamiliar with computer science. This is an instrumental study for the elaboration of a checklist, to delimit the steps and quality control of the material used for the elaboration of AI algorithms in automatic object detection. The present study was conducted in 2 phases: face validity and content validity. A questionnaire with 20 items was sent to 3 experts to assess the relevance and comprehension of the items. For content validity, the questionnaire was sent back to 5 experts to assess its relevance and wording; 19 items were kept for the final version of the instrument. The instrument presented, "ENDOLAP-IA", is the first of its kind. The creation of instruments like this one facilitates the introduction and integration of these new technologies in the medical field.

Keywords: information, knowledge, skills, SFIA, profile, cybersecurity

INTRODUCCIÓN

El procesamiento digital de imágenes consiste en una serie de procesos y técnicas para la representación digital de una imagen, un objeto a destacar en una escena, con el fin de facilitar su análisis (Medina & Bellera, 2003). En el trabajo de (Restrepo-Velázquez, 1999), las imágenes obtenidas a través de este proceso pueden ser utilizadas para realizar diagnósticos, tratamiento y pronóstico en un contexto clínico específico (CJ Acuña et al, 2007), sin embargo, la obtención de las imágenes para la realización de algoritmos de detección automática de objetos tiene sus obstáculos, las imágenes extraídas de sistemas de captura de video de una cámara de laparoscopia, presentan "ruido", una distorsión de la imagen ocasionada por elementos externos al sistema de registro digital de las imágenes y video, causado por la lente de la cámara, iluminación, la escala o dimensión, rotación y ausencia de enfoque y nitidez apropiado, que pueden comprometer con el resultado final de la calidad de la imagen obtenida (Selman, 2004).

En el área de la cirugía laparoscópica y endoscópica, el uso de sistemas de almacenamiento de video y uso de cámaras especializadas, provee de imágenes digitales para realizar diversos procedimientos diagnósticos y terapéuticos (Mascagni et al, 2022), sin embargo, durante un procedimiento quirúrgico es usual utilizar energía proveniente del calor por medio del electrocauterio, el cual durante el procedimiento genera humo al ser utilizado para la disección, causando distorsión de la imagen y pérdida del campo de visión, así mismo el sangrado secundario a una lesión o disección en tejido altera la anatomía quirúrgica, cuestiones técnicas previas a la configuración y calibración de la cámara laparoscópica pueden ocasionar falta de enfoque de la visión llegando a generar un riesgo durante el procedimiento (Torres, Serra & Marrecos, 2009). Para resolver esta clase de problemas, existen una serie de técnicas y algoritmos de preprocesamiento digital, estas técnicas incluyen normalización de cuadros, adecuado contraste, compresión de la imagen, escala apropiada, rotación y color (Mereles, 2012). En el contexto de la inteligencia artificial y el aprendizaje computacional automatizado, se requiere de datos de calidad y un robusto proceso de validación para crear un algoritmo para la detección de objetos con fines médicos (Arias et al, 2019).

En el proceso de recopilación se requieren de imágenes distintivas sobre el fenómeno de estudio. En la cirugía endoscópica, para una adecuada aplicación del potencial de inteligencia artificial es necesario contar con imágenes con adecuada nitidez, de adecuada calidad para ser utilizadas en el proceso del aprendizaje del algoritmo (Guo & Meng, 2024). Un paso

importante para lograr la detección automática, es realizar la anotación e identificación de las estructuras anatómicas de interés. La división y segmentación de las imágenes, eliminar áreas que no son necesarias para la tarea planeada, es una parte fundamental de este proceso (Iluidi et al, 2022). Posterior al preprocesamiento y compresión imágenes, los píxeles de cada una de las imágenes son representadas como un vector, de esta forma el área de interés (ADI) es seleccionada y las características que son de utilidad para el estudio son extraídas por medio de segmentación de la región que se desea analizar posteriormente, debido a que la conversión de la imagen a un sistema binario permitirá que el algoritmo computacional pueda realizar el aprendizaje y realizar la detección de objetos (Ali et al, 2020).

Es necesario realizar las correspondientes anotaciones del ADI para crear clases, las cuales servirán para delimitar si el objeto o no se encuentra presente al momento de la detección. Se deben contar con los elementos necesarios para el desarrollo de bases de datos con información apropiada, debido a que, si la base de datos es inadecuada, los resultados finales del algoritmo no serán lo suficiente para realizar la detección automática. (Ward et al, 2021). El desarrollo de la segmentación, creación de las clases y anotaciones dentro del procesamiento deben ser realizados por expertos en el área correspondiente y además cuenten con un entrenamiento en el uso de sistema digital de anotaciones, que estén familiarizados con la elaboración de anotaciones y creación de las clases correspondientes (Hashimoto et al, 2018). Su importancia radica en la calidad final del algoritmo, debido a que, si se cuentan con la señalización de los objetos por un experto, los resultados tendrán una mayor validez y precisión (Ward et al, 2021), (Ali et al, 2020), (Guo & Meng, 2024).

Una calidad óptima de las imágenes extraídas de videos de cirugías laparoscópicas permite crear una mejor precisión en algoritmos de inteligencia artificial para la detección de objetos, segmentación y detección por medio de píxeles sobre aspectos anatómicos, variaciones y enfermedad (Hashimoto et al, 2018). Permite tener una base de datos limpia, ordenada y lista para utilizarse para realización de algoritmos de inteligencia artificial para el diagnóstico de patologías con imágenes médicas, además de crear un repertorio para su uso en otros países para continuar con el desarrollo de nuevas tecnologías aplicadas en la cirugía laparoscópica (Fichtel et al, 2023).

En la literatura existen listas de cotejo empleadas en el área de la inteligencia artificial, una de ellas consiste en una serie de pasos para la publicación de artículos científicos relacionados con el procesamiento de imágenes médicas, como es el caso de la metodología CLAIM (Checklist for Artificial Intelligence in Medical Imaging), (Mongan et al, 2020), tal

metodología se creó con el fin de generar pautas y el cumplimiento de requisitos necesarios para la elaboración de modelos de inteligencia artificial para detección de imágenes. Existe una lista de verificación en investigación científica relacionada con la inteligencia artificial, se trata de MAIC-10 (Must AI Criteria-10), (Cerdá-Alberich et al, 2023), se desarrolló como una medida objetiva y cuantitativa de valorar la calidad de los estudios realizados cuyo objetivo es la detección de objetos por medio de visión y aprendizaje computacional. Se trata de una lista de cotejo que se compone de 10 ítems, que describen aspectos sobre la metodología realizada en el estudio, se creó con fines de medir la calidad de las investigaciones en el área de inteligencia artificial, su ventaja en comparación con otros instrumentos es su simplicidad y facilidad para utilizarse y emitir una calificación numérica sobre la calidad de la investigación. Sin embargo, debido a su simplicidad no contempla las dimensiones específicas que se deben de tener para valorar con precisión la calidad de los datos extraídos que se utilizaran. En cuanto a los parámetros o métricas habitualmente utilizadas para valorar la calidad de la imagen (VCI) se utiliza la representación de la imagen por el número de píxeles, capacidad de reproducir la luz, colores y texturas (Rodrigues et al, 2022). El parámetro más importante para valorar la calidad final es la percepción subjetiva del ojo humano de identificar los rasgos y características que se encuentran, para la extracción de los componentes de su imagen para su posterior interpretación y análisis (Jamil, 2024); (Obuchowicz et al, 2024). La visualización y la correcta identificación de los componentes que forman parte de la anatomía intraabdominal son de suma importancia, ya que a partir de su visión directa se pueden realizar distintos procedimientos para la reparación o extracción de un órgano abdominal, como lo es en el caso de la vesícula biliar. operador que manipula la cámara laparoscópica para el enfoque y documentación de las estructuras (Carsterns et al, 2023). Para el análisis de imágenes médicas con fines de predicción, detección de objetos, se requiere de una base de datos que tenga las características necesarias a investigar y probar por medio del algoritmo de inteligencia artificial. la identificación por medio de etiquetas o “mascaras” de estructuras u objetos de interés del estudio, los cuales se desea que se identifiquen y clasifiquen de manera automática con el uso de conocimiento supervisado de inteligencia artificial (Moore et al, 2023). Requiere de trabajo manual el realizar las anotaciones de cada imagen para formar parte de la recopilación que se utilizara para la detección, sin embargo, para que se tenga una mayor precisión y validación se requiere que un experto en el área este presente o realice estas anotaciones para tener mejores resultados en el procesamiento de las imágenes, es de importancia utilizar una metodología y sistemas computacionales para realizar estas anotaciones y tener una retroalimentación a

distancia para maximizar el tiempo y disminuir costos de trabajo (Mascagni et al, 2022). La segmentación proporcionará la unidad a medir del algoritmo, ya que la detección de las etiquetas y anotaciones permitirá por medio de uso de métodos matemáticos y estadísticos usando el sistema de coordenadas de los píxeles para detectar de manera correcta los límites y bordes de interés (Mirikharaji et al, 2023).

Las imágenes utilizadas en el diagnóstico médico son de vital importancia para garantizar la precisión y confiabilidad de los resultados. Una imagen de baja calidad puede llevar a interpretaciones erróneas, diagnósticos incorrectos y, en última instancia, a decisiones clínicas inapropiadas que afectan la salud del paciente (Ali et al, 2020). Por lo tanto, es fundamental establecer un método de evaluación riguroso para asegurar que las imágenes cumplan con los estándares necesarios. El objetivo del instrumento es proporcionar un conjunto de criterios específicos y objetivos que permitan a los profesionales de la salud determinar si una imagen es adecuada para su uso clínico. Esta herramienta de evaluación busca garantizar que las imágenes cumplen con los requisitos necesarios en términos de nitidez, contraste, resolución, ausencia de artefactos y otras características relevantes para su interpretación precisa. Permitirá una evaluación sistemática y consistente de la calidad de las imágenes minimizando el riesgo de errores de interpretación y mejora la confiabilidad de los diagnósticos.

MATERIALES Y MÉTODOS

Tipo de estudio

Se realizó un estudio instrumental (Montero & León, 2002), en el cual se diseñó un instrumento para la recolección de información en relación con las imágenes obtenidas de diversos medios digitales para la realización de protocolos con inteligencia artificial en el área de cirugía (ENDOLAP-IA). Se diseñó una lista de cotejo para valorar la calidad procedimental para la recopilación de datos, para la ejecución de algoritmos de inteligencia artificial en la detección de objetos (Cerdá-Alberich et al, 2023), (Egaña et al, 2014), (Alarcón et al, 2017) (Juárez-Hernández et al, 2017).

Procedimiento

En la primera etapa se diseñó una propuesta, por medio de la recopilación de ítems a partir de una revisión de la literatura relacionado con la valoración de la calidad de imágenes médicas para su procesamiento digital. El instrumento inicial contiene 20 ítems, que incluían preguntas con respuestas binarias de sí y no, los cuales están representados en 5

características: Calidad, Preprocesamiento, Identificación de los componentes anatómicos, Segmentación y Almacenamiento,

Figura 1 Cuantificación de validez de contenido

Tabla 1. Propuesta de instrumento para la valoración de la calidad del procesamiento de imágenes médicas digitales para su aplicación en inteligencia artificial				
Criterios a evaluar		Cumplimiento		Puntaje
		Cumple	No cumple	
Calidad				
1	¿Se tiene una resolución suficiente para identificar aspectos anatómicos relevantes?			
2	¿Las dimensiones de la imagen permiten ver con claridad los elementos que la componen?			
3	¿Los detalles de los elementos de la imagen permanecieron sin alteración posterior a la compresión del archivo?			
4	¿La imagen se encuentra en el mismo formato que el resto de las imágenes que serán utilizadas?			
Preprocesamiento				
5	¿La intensidad de la luz es la apropiada para evitar la sobreexposición o la subexposición de la imagen?			
6	¿Se redujo la variación del brillo y el color en las imágenes?			
7	¿Se ajusto correctamente el contraste para la identificación de los bordes de los distintos objetos en la imagen?			
8	¿Se simplifico el color de las imágenes en blanco y negro, para resaltar formas y contrastes sin distracciones provenientes del color?			
Identificación de los componentes anatómicos				
9	¿Previo a la laparoscopia, se realizó el balance de blancos?			
10	¿Durante la laparoscopia diagnostica se observan los elementos anatómicos?			
11	¿Hay distorsiones que impiden la identificación de las estructuras?			
12	¿Se identifican la vesícula biliar, hígado y adherencias en la laparoscopia?			
Segmentación				
13	¿Se definieron de los bordes para la detección individual de cada uno de los objetos?			
14	¿Se colocaron las etiquetas e identificación de los pixeles de cada imagen para su clase correspondiente?			
15	¿Se realizo el análisis y valoración por experto de las etiquetas de la imagen?			
16	¿Los límites de las distintas estructuras dentro de la imagen se distinguen o delimitan apropiadamente una de otra?			
17	¿Se clasifico la imagen segmentada en la clase correspondiente?			
Almacenamiento				
18	¿Se realizo el almacenamiento de la imagen digital dentro de la carpeta correspondiente?			
19	¿Se utilizo un nombre distintivo para la imagen procesada?			
20	¿Se almacenaron las imágenes en el formato correcto para su uso en el algoritmo de inteligencia artificial?			

Análisis de validez de facie

Posterior a su construcción, se sometió al instrumento de investigación por 3 expertos, con el fin de verificar los ítems pertenecientes al constructo y su relevancia, además de su redacción y comprensión de cada uno de sus ítems (Ruiz, 2002), (Abad et al, 2011) (Lloret-Segura et al, 2014). El panel de expertos se compuso de 3 integrantes: 2 cirujanos generales laparoscopistas y 1 académico experto en ciencias computacionales de la Universidad Autónoma de Chihuahua (UACH), todos ellos con mínimo de maestría (2) y doctorado (1) con experiencia en investigación y estudios relacionados con el uso de inteligencia artificial. De los cuales 2 de los expertos son de género masculino y 1 del género femenino

Figura 2 Caracterización de los expertos

Tabla 2. Caracterización de los expertos

Perfil de los expertos	Experto 1	Experto 2	Experto 3
Cargo actual	Medico adscrito	Docente	Medico adscrito
Maestría	Si	Si	Si
Doctorado	No	Si	No
Áreas de experiencia	Cirugía General	Procesamiento de imágenes y desarrollo de modelos de machine learning	Cirugía General
Número de años de experiencia profesional	8	19	30
Años de experiencia en docencia	7	19	27
Numero de artículos publicados	60	21	30
Número de libros publicados	0	0	0
Numero de capítulos de libros publicados	1	4	2
Numero de ponencias presentadas en congresos	15	12	50
¿Tiene experiencia en la revisión, diseño y/o validación de instrumentos de investigación?	Si	Si	Si
Número de años de experiencia en la revisión, diseño y/o validación de instrumentos de investigación	7	6	10

Fuente: Elaboración propia

De manera individual cada experto analizó cada uno de los ítems. La selección de los expertos se basó en relación con su experiencia profesional en el área de investigación, grado académico, experiencia en la revisión de instrumentos (Egaña et al, 2014), (Valenzuela et al, 2022). Se consultaron a través de un cuestionario semiestructurado y autoadministrado por medio de Google Forms, compuesto de 20 ítems. A cada uno de los ítems los expertos podían realizar observaciones de cada uno de los ítems, calificando la redacción y pertinencia de cada uno de los ítems presentados en el cuestionario, se les indicó a los expertos la definición conceptual de cada constructo y las indicaciones para evaluar el nivel de importancia de cada uno de los ítems de acuerdo a una evaluación cualitativa a) Si, b) Moderadamente y c) No (Juárez-Hernández et al, 2017).

Análisis de validez de contenido

Posterior a la valoración por el panel de expertos, verificando las observaciones planteadas por los distintos expertos, se procedió a realizar las modificaciones pertinentes. Se envió nuevamente un formulario en línea por medio de la plataforma Google Forms a 6 expertos, de los cuales participaron 5 expertos en la evaluación de la lista de cotejo. Para la evaluación por parte de los expertos, se utilizó la escala de jueces expertos planteada por CIFE (2018), que se compone de una sección cualitativa y otra de tipo cuantitativa (Juárez-Hernández et al, 2017). En la sección cualitativa, los expertos se les solicitaba realizar recomendaciones y/o sugerencias respecto a la pertinencia y redacción de cada uno de los ítems que componen al instrumento. Se busco obtener la validación de contenido por medio de la Razón de Validez de Contenido (CVR), la cual fue desarrollada por Lawshe en 1975, de esta manera los instrumentos pueden ser evaluados y analizados, cada experto emite su opinión y observación de cada ítem por medio de las opciones: a) esencial, b) útil pero no esencial y c) no necesario (Puerta y Marín, 2015), se homologaron las puntuaciones de dicha escala de la siguiente manera: a) esencial (4 y 3), b útil pero no esencial (2) y c) necesario pero no esencial (1). Por medio de este procedimiento se obtiene el total de coincidencias por cada uno de los ítems en cada categoría, esperando que la opción esencial presente un valor superior al 50% por medio de la ecuación presentada, con un valor mínimo de CVR de 0.49 (Araya et al, 2018).

$$CVR = \frac{ne/2}{N/2}$$

Ecuación 1. Razón de validez de contenido

ne =Número de expertos de acuerdo con la categoría esencial

N =Numero de expertos

Además, se empleó, una modificación, por medio de una simplificación con un nuevo indicador, con base en una normalización por medio de la ecuación 2, (Tristán, 2008), en la cual se establece un valor mínimo aceptable es de un CVR” 0.58:

$$CVR = \frac{CVR + 1}{2}$$

Ecuación 2. Modificación de CVR por Tristán, 2008.

Para la evaluación cuantitativa de la validez de contenido, se utilizó la ecuación 1 y 2 para las observaciones realizadas por los expertos, de esta forma obteniendo la validez de contenido para cada uno de los ítems dentro del instrumento (Parra, 2020).

RESULTADOS

Revisión por expertos

Los expertos sugirieron realizar modificaciones respecto a la redacción y simplificación de conceptos con el objetivo de evitar confusiones y errores de interpretación. En el ítem 2 se modificó la redacción del ítem para tener una mejor comprensión del objetivo a evaluar, se modificó y cambio su planteamiento a (¿Las dimensiones y resolución de la imagen permiten ver con claridad los elementos que la componen?). En el ítem 3 se modificó la redacción de la pregunta por sugerencia de uno de los expertos, debido al tecnicismo del lenguaje empleado y las posibles confusiones que pudiera plantear más adelante, cambiando la pregunta (¿Los detalles de los elementos de la imagen permanecieron sin alteración posterior a la compresión del archivo?). También se decidió realizar modificaciones al ítem 5 en cuestión de su redacción para facilitar la comprensión del objetivo a evaluar de dicho ítem, se sustituyó el termino iluminación, por intensidad de la luz para evitar sobreexposición de la imagen. En el ítem 6 se sustituyó dicha pregunta por un nuevo planteamiento debido a la ambigüedad de la misma, debido a que el termino "ruido" presentaba mucha confusión como parte de las observaciones de los expertos, por lo cual se utilizó una definición simple y objetiva para representar dicho concepto en el contexto de análisis digital de imágenes. En el ítem 7 se sustituyeron palabras dentro del planteamiento de la pregunta de dicho ítem para evitar confusión en la ambigüedad de los términos, se modificó para expresar mejor la opción respecto a que el contraste se corroboró para distinguir los bordes de los distintos elementos que componen a la imagen. En cuanto al ítem 8 se eliminó y se sustituyó por un nuevo ítem, debido a que los 3 expertos dentro de sus observaciones determinaron que el concepto era confuso y difícil de interpretar, además que no aportaba información útil al instrumento. En

el ítem número 16 se modificó el planteamiento de la pregunta del ítem, debido a que dentro de la redacción de la misma se encontraba redactada con ambigüedad, se modificó la redacción para simplificar el termino empleado. Se hicieron además modificaciones no estructurales de los ítems como revisión de los errores ortográficos y puntuaciones y espacios de cada uno de los ítems.

Validez de contenido

En la segunda versión de la lista de cotejo se incluyeron las sugerencias de los jueces expertos, además de realizar las modificaciones pertinentes a los ítems evaluados con bajo nivel de pertinencia y redacción. Como se observa en la (Tabla 3), la gran mayoría de los ítems obtuvieron puntuaciones superiores a los límites aceptables, solo el ítem 5 presentó valor inferior de 0.49, eliminando este ítem de la segunda propuesta del instrumento, por lo cual se mantuvieron 19 ítems del instrumento (Tabla 4), cuya razón de validez de contenido constituye como aceptables.

Tabla 1 Cuantificación de validez de contenido

Dimensión	Ítem	Pertinencia		Redacción	
		CVR	CVR''	CVR	CVR''
Calidad	1	1	1	1	1
	2	1	1	1	1
	3	1	1	1	1
	4	1	1	1	1
	5	0.2	0.6	0.2	0.6
Preprocesamiento	6	1	1	1	1
	7	1	1	1	1
	8	1	1	1	1
Identificación de componentes anatómicos	9	1	1	1	1
	10	1	1	1	1
	11	1	1	1	1
	12	1	1	1	1
Segmentación	13	1	1	1	1
	14	1	1	1	1
	15	1	1	1	1
	16	1	1	1	1
Almacenamiento	17	1	1	1	1
	18	1	1	1	1
	19	1	1	1	1
	20	1	1	1	1

Fuente: Elaboración propia del autor

El ítem 5 ¿La intensidad de la luz es la apropiada para evitar la sobreexposición o la subexposición de la imagen?, se eliminó debido a que presentó un CVR para pertinencia de 0.2 y una redacción de 0.2, con un CVR” para pertinencia y redacción de 0.6. En cambio, el resto de los ítems (1-4, 6-20) presentaron valores de 1.0 tanto para la pertinencia y redacción, teniendo valores óptimos.

Tabla 2 Modificación del instrumento posterior a la revisión por expertos

Criterios a evaluar	Cumplimiento		Puntaje
	Cumple	No cumple	
Calidad			
1	¿Se tiene una resolución suficiente para identificar aspectos anatómicos relevantes?		
2	¿Las dimensiones de la imagen permiten ver con claridad los elementos que la componen?		
3	¿Los detalles de los elementos de la imagen permanecieron sin alteración posterior a la compresión del archivo?		
4	¿La imagen se encuentra en el mismo formato que el resto de las imágenes que serán utilizadas?		
Preprocesamiento			
5	¿Se redujo la variación del brillo y el color en las imágenes?		
6	¿Se ajustó correctamente el contraste para la identificación de los bordes de los distintos objetos en la imagen?		
7	¿Se simplificó el color de las imágenes en blanco y negro, para resaltar formas y contrastes sin distracciones provenientes del color?		
Identificación de los componentes anatómicos			
8	¿Previo a la laparoscopia, se realizó el balance de blancos?		
9	¿Durante la laparoscopia diagnóstica se observan los elementos anatómicos?		
10	¿Hay distorsiones que impiden la identificación de las estructuras?		
11	¿Se identifican la vesícula biliar, hígado y adherencias en la laparoscopia?		

Segmentación			
12	¿Se definieron de los bordes para la detección individual de cada uno de los objetos?		
13	¿Se colocaron las etiquetas e identificación de los pixeles de cada imagen para su clase correspondiente?		
14	¿Se realizo el análisis y valoración por experto de las etiquetas de la imagen?		
15	¿Los límites de las distintas estructuras dentro de la imagen se distinguen o delimitan apropiadamente una de otra?		
16	¿Se clasifico la imagen segmentada en la clase correspondiente?		
Almacenamiento			
17	¿Se realizo el almacenamiento de la imagen digital dentro de la carpeta correspondiente?		
18	¿Se utilizo un nombre distintivo para la imagen procesada?		
19	¿Se almacenaron las imágenes en el formato correcto para su uso en el algoritmo de inteligencia artificial?		
Resultados de valoración			
Valoración de los criterios	Nivel de desempeño	Valoración numérica	
>18 criterios demostrados	Excelente		
14 a 17 criterios demostrados	Bueno		
11 a 13 criterios demostrados	Regular		
Menos de 10 criterios demostrados	Pobre		
Fuente: Elaboración propia			

DISCUSIÓN

El propósito de este estudio fue diseñar un instrumento para delimitar los pasos y la calidad de las imágenes que serán utilizadas para la realización de algoritmos de detección

automática con IA por medio de una lista de cotejo sin aspectos complejos ni técnicos de por medio. Al realizar la revisión de la literatura, se encontró una falta de existencia de instrumentos relacionados para mantener una estandarización y una adecuada metodología para el desarrollo de algoritmos diagnósticos que puedan ser utilizados en el ámbito clínico (Arias et al, 2019). En la actualidad el creciente uso y adopción de la inteligencia artificial en la cirugía ha crecido y se ha iniciado su introducción dentro de las salas de operaciones (Hashimoto et al, 2020). Sin embargo, aún se encuentra en una etapa muy temprana para su aplicación de manera establecida dentro de las unidades hospitalarias y las salas de quirófanos. La creación de un método de evaluación objetiva brinda mayor información y aportación para mejorar los aspectos diagnósticos; aumentando la precisión diagnóstica para la detección patologías (Ali et al, 2019); (Geleijnse et al, 2022); (Reijers et al, 2017); (Schiavetti et al, 2020); (Concha-Torre et al, 2020).

La lista de cotejo creada en el presente estudio aborda la necesidad de crear modelos de evaluación de calidad del material empleado con el fin de crear aplicaciones con fines diagnósticos para automatizar procesos que de forma tradicional consumen recursos y tiempo (Kulp et al, 2020). Se debe realizar una validación del instrumento, con el fin de crear una herramienta de fácil acceso y comprensión para la elaboración de algoritmos de detección automática por medio de inteligencia artificial con fines diagnósticos, dada la pesada tarea de comprender los pasos necesarios para su elaboración, un instrumento validado generara un mayor acercamiento a aquellos del personal de salud que estén interesado en incursionar en estas nuevas tecnologías.

En la revisión por expertos se hicieron cambios relacionados con la redacción de los ítems, debido a que inicialmente se utilizaron conceptos ambiguos y técnicos que no se consideraban comprensibles. Los ítems a los cuales se les realizaron modificaciones por sugerencia de los expertos fueron los ítems 2, 3, 5, 6, 7, 8 y 16. Dentro de las categorías de los ítems a los cuales se les realizaron modificaciones están Calidad, Preprocesamiento y Segmentación. En cuanto a las observaciones realizadas por los expertos en la categoría Calidad, fueron sobre simplificar la redacción de las preguntas y ser concisos respecto al objetivo de cada una de las preguntas, con el fin de tener una mayor comprensión sin necesidad de realizar una búsqueda previa de información. En la categoría de Preprocesamiento fue donde existieron el mayor número de observaciones relacionadas con la redacción y comprensión, esto relacionado con la redacción de las preguntas planteadas, dada la naturaleza de esta área, al tratarse de la forma en la cual se representan los colores, el

brillo, la intensidad de la luz, el contraste, se realizaron modificaciones pertinentes para ser concisos y tener una mejor comprensión del ítem, sin embargo el ítem 8 termino por ser eliminado al no tener una sustitución debido a ausencia de traducción al español. El resto de categorías permanecieron sin modificaciones, debido a que se tuvo un alto nivel de valoración por parte de los expertos, ya que las consideraban como esenciales para formar parte del instrumento.

En el juicio de expertos realizado por 5 expertos, fue necesario eliminar el ítem 5 del instrumento debido a que presentó un CVR para pertinencia y redacción con 0.2, además de un CVR” en pertinencia y redacción de 0.6, según (Araya et al, 2018), una razón de contenido con este valor, no puede ser aceptable dentro del instrumento, ya que condiciona a que este elemento o ítem no cuenta con la pertinencia ni redacción necesaria para ser parte de este instrumento y de no ser eliminado puede ocasionar una diferencia significativa en la interpretación del instrumento. El resto de los ítems del instrumento presentaron valores óptimos de CVR en la pertinencia y redacción, lo que establece (Parra, 2020), que tal valor representa que los ítems de cada una de las categorías dentro del instrumento son representativos y su construcción para la obtención y recopilación de datos es la óptima.

Dentro de las limitaciones del instrumento, el contar con un número reducido de expertos para la validación del instrumento ocasiona que los resultados no sean muy representativos estadísticamente, al tratarse de un área de reciente crecimiento e integración en el área de la medicina, aún hay pocos expertos en el área disponibles. El desarrollo de este instrumento es innovador, se trata del primer instrumento creado para evaluar el contenido que forma parte del proceso de elaboración de algoritmos de inteligencia artificial para la detección de objetos, con el fin de mejorar la precisión y validez (Ward et al, 2020), para la compleja tarea de desarrollo de proyectos en aprendizaje automatizado (Hashimoto et al, 2018).

CONCLUSIONES

El instrumento presentado ENDOLAP-IA, es el primero en su clase, por lo cual se puede utilizar como referencia para futuros proyectos en innovación tecnológica aplicada en la medicina y la inteligencia artificial. En este instrumento se incluyen conceptos relacionados con la programación y edición digital de forma clara, concisa y práctica, para introducir no solo a cirujanos, sino también personal de salud interesado en la aplicación de la inteligencia artificial como una herramienta en su práctica diaria. Estos conceptos comprendan los pasos y puntos necesarios a tener en cuenta al momento de procesar imágenes de cirugías

laparoscópicas, procedimientos endoscópicos y fotografías de piezas extraídas de cirugías. Se presentó una metodología simple y rápida de aplicar con la cual se obtendrán resultados apropiados y con mayor precisión conforme lo establecido en la literatura. La tecnología y la medicina continúan en crecimiento, por ello es importante adoptar y aplicar nuevas herramientas para realizar diagnósticos, toma de decisiones, entrenamiento a residentes y educación al paciente, sin tener complicaciones con tecnicismos y miedo al cambio, por ello el crear instrumentos como el presente, facilitan la introducción e integración a aplicar estas nuevas tecnologías en desarrollo en el ámbito hospitalario.

REFERENCIAS BIBLIOGRÁFICAS

Abad, F. J., Olea, J., Ponsoda, J., y García, C. (2011). *Medición en ciencias sociales y de la salud*. Madrid. Síntesis

Acuña, C. J., Marcos, E., De Castro, V., Hernández, J. A., & López-Sanz, M. (2007). Gestión de Imágenes Médicas a través de la Web. *Revista Colombiana de Computación*, 8(1), 1-11.

<http://revistas.unab.edu.co/index.php?journal=rcc&page=article&op=viewArticle&path%5B%5D=1040>

Alarcón, L. A. G., Trápaga, J. A. B., & Navarro, R. E. (2017). Content validity by experts' judgment: Proposal for a virtual tool. *Apertura*, 9(2), 42-53.

<https://doi.org/10.32870/ap.v9n2.993>

Ali, O., Ally, M., Clutterbuck, P., & Dwivedi, Y. (2020). The state of play of blockchain technology in financial services sector: A systematic review. *International Journal of Information Management*, 54. [doi:10.1016/j.ijinfomgt.2020.102199](https://doi.org/10.1016/j.ijinfomgt.2020.102199)

Araya, S. C. B., Egaña, M. J. U., & Salamanca, M. I. C. (2018). Validez de contenido de un banco de ítems en el área de salud del niño. *Educación Médica Superior*, 31(4), 1-9.

http://scielo.sld.cu/pdf/ems/v31n4/a13_1171.pdf

Arias, V., Salazar, J., Gariciano, C., Contreras, J., Chacón, G., Chacín-González, M., Añez, R., Rojas, J., & Bermúdez-Pirela, V. (2019). Una introducción a las aplicaciones de la

inteligencia artificial en Medicina: Aspectos históricos. *Revista Latinoamericana de Hipertensión*, 14(5), 590-600.

<https://biblat.unam.mx/hevila/Revistalatinamericanadehipertension/2019/vol14/no5/13.pdf>

Carretero-Dios, H., & Pérez, C. (2005). Normas para el desarrollo y revisión de estudios instrumentales. DOAJ (DOAJ: Directory of Open Access Journals).

<https://doaj.org/article/b4c0167b550847178361de17bad327be>

Carstens, M., Rinner, F. M., Bodenstedt, S., Jenke, A. C., Weitz, J., Distler, M., Speidel, S., & Kolbinger, F. R. (2023). The Dresden Surgical Anatomy Dataset for Abdominal Organ Segmentation in Surgical Data Science. *Scientific Data*, 10(1).

<https://doi.org/10.1038/s41597-022-01719-2>

Cerdá-Alberich, L., Solana, J., Mallol, P., Ribas, G., García-Junco, M., Alberich-Bayarri, A., & Marti-Bonmati, L. (2023). MAIC-10 brief quality checklist for publications using artificial intelligence and medical images. *Insights Into Imaging*, 14(1).

<https://doi.org/10.1186/s13244-022-01355-9>

Concha-Torre, A., Alonso, Y. D., Blanco, S. Á., Allende, A. V., Mayordomo-Colunga, J., & Barrio, B. F. (2020). The checklists: ¿A help or a hassle? *Anales de Pediatría*, 93(2), 135.e1-135.e10. <https://doi.org/10.1016/j.anpede.2020.05.003>

Egaña, M. U., Araya, S. B., Núñez, M. G., & Camus, M. M. (2014). Métodos óptimos para determinar validez de contenido. *Revista Cubana de Educación Médica Superior*, 28(3), 547-558. <http://scielo.sld.cu/pdf/ems/v28n3/ems14314.pdf>

Fichtel, L., Erbacher, D., Heller, L., Fruhwald, A., Hösch, L., & Bachmeir, C. (2023). Analysis of Object Detection Datasets for Machine Learning with Small and Tiny Objects. *Civil-comp Conferences*. <https://doi.org/10.4203/coc.2.7.4>

Geleijnse, G., Veder, L. L., Hakkesteegt, M. M., & Metselaar, R. M. (2022). The Objective Measurement and Subjective Perception of Flexible ENT Endoscopes; Image

- Quality. *The Journal of Imaging Science and Technology/The Journal of Imaging Science and Technology*, 66(3), 030508-6.
<https://doi.org/10.2352/j.imagingsci.technol.2022.66.3.030508>
- Guo, F., & Meng, H. (2024). Application of artificial intelligence in gastrointestinal endoscopy. *Arab Journal of Gastroenterology /Arab Journal of Gastroenterology*.
<https://doi.org/10.1016/j.ajg.2023.12.010>
- Hashimoto, D. A., Rosman, G., Rus, D., & Meireles, O. R. (2018). Artificial Intelligence in Surgery: Promises and Perils. *Annals Of Surgery*, 268(1), 70-76.
<https://doi.org/10.1097/sla.0000000000002693>
- Jamil, S. (2024). Review of Image Quality Assessment Methods for Compressed Images. *Journal Of Imaging*, 10(5), 113. <https://doi.org/10.3390/jimaging10050113>
- Juarez-Hernandez, L. G., & Tobon, S. (2018). Análisis de los elementos implícitos en la validación de contenido de un instrumento de investigación. *Revista ESPACIOS*, 39(53). <https://es.revistaespacios.com/cited2017/cited2017-23.html>
- Kulp, L., Sarcevic, A., Zheng, Y., Cheng, M., Alberto, E., & Burd, R. (2020). Checklist Design Reconsidered: Understanding Checklist Compliance and Timing of Interactions. *CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3313831.3376853>
- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada. *Anales de Psicología*, 30(3). <https://doi.org/10.6018/analesps.30.3.199361>
- Mascagni, P., Alapatt, D., Sestini, L., Altieri, M. S., Madani, A., Watanabe, Y., Alseidi, A., Redan, J. A., Alfieri, S., Costamagna, G., Boškoski, I., Padoy, N., & Hashimoto, D. A. (2022). Computer vision in surgery: from potential to clinical value. *Npj Digital Medicine*, 5(1). <https://doi.org/10.1038/s41746-022-00707-5>

- Medina, R., & Bellera, J. (2003). Bases del procesamiento de imágenes médicas [Tesis de maestría, Universidad de Los Andes, Facultad de Ingeniería, Grupo de Ingeniería Biomédica de la ULA]. http://www.saber.ula.ve/redtelemedicina/TallerTelemedicina/ponencia-j_bellera.html
- Mereles, L. (2012). Preprocesamiento de imágenes digitales a través de su Transformada de Fourier [Tesis de Maestría, Universidad Nacional del Sur, Argentina]. <http://lcr.uns.edu.ar/fvc/NotasDeAplicacion/FVC-Leonardo%20Mereles.pdf>
- Mirikharaji, Z., Abhishek, K., Bissoto, A., Barata, C., Avila, S., Valle, E., Celebi, M. E., & Hamarneh, G. (2023). A survey on deep learning for skin lesion segmentation. *Medical Image Analysis*, 88, 102863. <https://doi.org/10.1016/j.media.2023.102863>
- Mongan, J., Moy, L., & Kahn, C. E. (2020). Checklist for Artificial Intelligence in Medical Imaging (CLAIM): A Guide for Authors and Reviewers. *Radiology. Artificial Intelligence*, 2(2), e200029. <https://doi.org/10.1148/ryai.2020200029>
- Montero, I., & León, O. G. (2002). Clasificación y descripción de las metodologías de investigación en Psicología. *Directory Of Open Access Journals*. <https://doaj.org/article/098b74fb3e494b3dabfeb4b897dc09ad>
- Moor, M., Banerjee, O., Abad, Z. S. H., Krumholz, H. M., Leskovec, J., Topol, E. J., & Rajpurkar, P. (2023). Foundation models for generalist medical artificial intelligence. *Nature*, 616(7956), 259-265. <https://doi.org/10.1038/s41586-023-05881-4>
- Obuchowicz, R., Strzelecki, M., & Piórkowski, A. (2024). Clinical Applications of Artificial Intelligence in Medical Imaging and Image Processing—A Review. *Cancers*, 16(10), 1870. <https://doi.org/10.3390/cancers16101870>
- Parra, R. I. M. (2020). Validez de contenido de un instrumento de medición de derechos humanos en México. *Revista de Ciencias Sociales*, 168, 203-232. <https://doi.org/10.15517/rcs.v0i168.43981>

- Puerta Sierra, L. M., & Martín Vargas Ma. Enselmina. (2015). Análisis de validez de contenido de un instrumento de transferencia de tecnología universidad-industria de baja california, México.
<https://investigacion.fca.unam.mx/docs/memorias/2015/2.02.pdf>
- Reijers, H. A., Leopold, H., & Recker, J. (2017). Towards a Science of Checklists. Proceedings Of The . . . Annual Hawaii International Conference on System Sciences/Proceedings of The Annual Hawaii International Conference on System Sciences. <https://doi.org/10.24251/hicss.2017.696>
- Restrepo, A. (1998). Procesamiento de imágenes médicas. Revista Universidad EAFIT, 34(110), 86-92. <https://biblat.unam.mx/en/revista/revista-universidad-eafit/articulo/procesamiento-de-imagenes-medicas>
- Rodrigues, R., Lévêque, L., Gutiérrez, J., Jebbari, H., Outtas, M., Zhang, L., Chetouani, A., Al-Juboori, S., Martini, M., & Pinheiro, A. M. G. (2022). Objective quality assessment of medical images and videos: Review and challenges. arXiv (Cornell University). <https://doi.org/10.48550/arxiv.2212.07396>
- Ruiz Bolívar, C. (2002). Instrumentos y técnicas de investigación educativa: Un enfoque cuantitativo y cualitativo para la recolección y análisis de datos (2.a ed., Vol. 1). Venezuela: Fedupel.
- Schiavetti, B., Wynendaele, E., Melotte, V., Van Der Elst, J., De Spiegeleer, B., & Ravinetto, R. (2020). A simplified checklist for the visual inspection of finished pharmaceutical products: a way to empower frontline health workers in the fight against poor-quality medicines. Journal Of Pharmaceutical Policy and Practice, 13(1). <https://doi.org/10.1186/s40545-020-00211-9>

- Selman, R., & Miguel, J. (2004). Aplicaciones clínicas del procesamiento digital de imágenes médicas. *Revista Médica Clínica las Condes*, 15(2), 0. <https://www.elsevier.es/es-revista-revista-medica-clinica-las-condes-202-pdf-X0716864004320006>
- Tristán, A. (2008). Modificación al modelo de Lawshe para el dictamen cuantitativo de la validez de contenido de un instrumento objetivo. *Avances en Medición Instituto de Evaluación e Ingeniería Avanzada*, 1 (6), 37-48
- Valenzuela, R. R., Ramos, D. N. M., Torres, A. G., Rodríguez, R. J., & Montero, F. y. M. (2022). Análisis de validez de contenido por criterio de jueces de un instrumento para evaluar un manuscrito. *RIDE Revista Iberoamericana Para la Investigación y el Desarrollo Educativo*, 12(24). <https://doi.org/10.23913/ride.v12i24.1183>
- Ward, M. D., Zimmerman, M. I., Meller, A., Chung, M., Swamidass, S. J., & Bowman, G. R. (2021). Deep learning the structural determinants of protein biochemical properties by comparing structural ensembles with DiffNets. *Nature Communications*, 12(1). <https://doi.org/10.1038/s41467-021-23246-1>