





La era de la endoscopia inteligente: cómo la inteligencia artificial potencia la endoscopia digestiva

Jorge Baquerizo-Burgos  · María Egas-Izquierdo  · Doménica Cunto  · Carlos Robles-Medranda 

Departamento de Endoscopia, Instituto Ecuatoriano de Enfermedades Digestivas (IECED), Guayaquil, Ecuador.

Acta Gastroenterol Latinoam 2023;53(3):211-225

Recibido: 09/08/2023 / Aceptado: 14/09/2023 / Publicado online: 30/09/2023 / <https://doi.org/10.52787/agl.v53i3.339>

Resumen

La inteligencia artificial es un campo de la ciencia y la ingeniería que se ocupa de la comprensión computacional de comportamientos inteligentes y la creación de artefactos que exhiben tales comportamientos, lo que permite a las computadoras funcionar y pensar de manera similar a la de los seres humanos. Esta tecnología ayuda a superar los múltiples retos que enfrentan los profesionales de la salud y aporta al diagnóstico, al manejo y al pronóstico de los pacientes. Actualmente se están desarrollando varios modelos para la endoscopia digestiva, incluyendo algunos que permiten la detección de estructuras anatómicas que pueden ayudar en el entrenamiento de los médicos, servir como guía durante los procedimientos endoscópicos o para la estratificación de lesiones pre-

malignas y malignas. De esta forma disminuirían los falsos negativos y se proporcionarían más tratamientos oportunos. En la actualidad existen sistemas computarizados de detección de lesiones y de diagnóstico en los distintos segmentos de la vía digestiva, cada uno con funciones particulares que proporcionan asistencia durante los procedimientos. Todo esto se ha llevado a cabo con el fin de reducir los riesgos derivados de los factores humanos y ambientales, entre otros, los cuales pueden afectar el diagnóstico y el manejo de las enfermedades. Los modelos de inteligencia artificial en la endoscopia digestiva pueden, además de mejorar la impresión visual de los endoscopistas, disminuir la curva de aprendizaje a través de la aplicación de tecnologías precisas. De esta manera, se reduce la diferencia entre los endoscopistas expertos y menos expertos. En este artículo se discuten los avances tecnológicos de la inteligencia artificial en la endoscopia digestiva y los aspectos futuros relacionados.

Palabras claves. Inteligencia artificial, detección asistida por computadora, diagnóstico asistido por computadora, aprendizaje profundo, endoscopia.

The Era of Intelligent Endoscopy: How Artificial Intelligence Empowers Digestive Endoscopy

Summary

Artificial intelligence is a field of science and engineering that focuses on the computational understanding of

Correspondencia: Jorge Baquerizo-Burgos
Correo electrónico: jorgebaquerizoburgos@gmail.com

intelligent behaviors and the creation of artifacts that exhibit such behaviors, enabling computers to function and think similarly to humans. This technology assists in overcoming the various challenges that healthcare professionals face, contributing to the diagnosis, management, and prognosis of patients. Currently, several artificial intelligence models have been developed in the field of digestive endoscopy, including some that enable the detection of anatomical structures that can aid in the training of medical professionals, serve as guides during endoscopic procedures, or assist in stratifying pre-malignant and malignant lesions. This reduces false negatives, providing timely treatments. Computerized lesion detection and diagnostic systems exist for different segments of the digestive tract, each with specific functions that provide assistance during procedures. All of this has been done with the aim of reducing risks stemming from human and environmental factors, among others, which can affect early diagnosis and management of diseases. Artificial intelligence models for digestive endoscopy can not only enhance the visual impression for endoscopists but also reduce the learning curve through the application of precise technologies, thus bridging the gap between experienced and less experienced endoscopists. This article discusses the technological advancements of artificial intelligence in digestive endoscopy and related future aspects.

Keywords. Artificial intelligence, Computer-assisted detection, computer-assisted diagnosis, deep learning, endoscopy.

Abreviaturas

IA: Inteligencia artificial.

ML: Machine learning (aprendizaje de máquinas).

DL: Deep learning (aprendizaje profundo).

CADe: Computer-assisted detection device (dispositivo de detección asistida por computadora).

CADx: Computer-assisted diagnostic device (dispositivo de diagnóstico asistido por computadora).

EUS: Endoscopic ultrasound (ultrasonido endoscópico).

EGD: Esofagogastroduodenoscopia.

AUC: Area under the curve (área bajo la curva).

EB: Esófago de Barrett.

NBI: Narrow Band Imaging (imágenes de banda estrecha).

ERGE: Enfermedad por reflujo gastroesofágico.

SVM: Support Vector Machine (máquinas de vector en soporte).

CNN: Convolutional Neural Network (red neuronal convolucional).

CCR: Cáncer colorrectal.

RR: Riesgo relativo.

PDR: Polyp Detection Rate (tasa de detección de pólipos).

CPRE: Colangiopancreatografía endoscópica retrógrada.

VPP: Valor predictivo positivo.

VPN: Valor predictivo negativo.

mAP: Mean Average Precision (media de precisión).

EPS: Frames per Second (fotogramas por segundo).

NPMI: Neoplasia papilar mucinosa intraductal.

GIST: Gastrointestinal Stromal Tumors (tumores del estroma gastrointestinal).

Introducción

Inteligencia artificial: conceptos básicos

La inteligencia artificial (IA) es una rama de las ciencias de la computación cuyo objetivo es la ejecución de comprensiones inteligentes a partir de un conjunto de modelos.¹ Usando un conjunto de algoritmos, la IA es capaz de funcionar y pensar como un ser humano mediante un proceso de aprendizaje basado en el entrenamiento y tiene como ventaja el poder completarlo en menor tiempo que el ser humano.¹ Adicionalmente, esta tecnología puede integrar el aprendizaje de máquinas (machine learning, ML) y una rama de éste, el aprendizaje profundo (deep learning, DL).²

El ML es un subgrupo de la IA que se caracteriza por la utilización de modelos matemáticos para el aprendizaje de datos que posteriormente permitirán el reconocimiento de patrones.³ A partir de los algoritmos se crean modelos predictivos que permiten el análisis de datos y la resolución de problemas complejos. Asimismo, el ML puede ser supervisado, no supervisado y de reforzamiento.^{2,4,5}

- El aprendizaje supervisado está basado en el entrenamiento a partir de datos bien categorizados o etiquetados (supervisión externa). Los datos etiquetados son divididos para el entrenamiento y la validación interna. Este tipo de aprendizaje está basado en la regresión, la clasificación y la caracterización.^{2,3}

- El aprendizaje no supervisado o automatizado es el modelo que aprende de datos no categorizados y que permite al algoritmo actuar sin ningún tipo de guía, basándose en el entendimiento de patrones y, por ende, requiriendo mayor cantidad de información.^{2,3}

- El aprendizaje de reforzamiento no requiere datos ni supervisión para alimentarse sino que está basado en el aprendizaje del entorno a partir de recompensas.^{2,3}

El DL es una categoría especial del ML, el cual está basado en la arquitectura de redes neuronales análogas a las del cerebro humano.³ Consiste en una capa inicial que recibe una entrada o *input*; a esta capa la sigue un conjunto de capas medias ocultas o *hidden layers* y, a continuación, la capa final de salida u *output* (Figura 1). Cada capa de esta red consiste en un grupo de neuronas o nodos que convierten (activación) un *input* en *output* a

partir de funciones matemáticas.² El *output* de una capa anterior sirve de *input* para la capa siguiente, y así sucesivamente, hasta llegar a la capa de salida y obtener un resultado o detección.^{2,3}

El desarrollo del modelo de detección basado en DL incluye tres fases principales. En la primera fase, se realiza la recolección de datos (imágenes o videos) y se etiquetan adecuadamente las estructuras que se utilizarán en el aprendizaje del modelo. Luego, en la segunda fase, se establece la arquitectura del modelo y se crean las redes neuronales (capa de entrada, capas medias y capa de salida). Finalmente, en la tercera fase, se utilizan las muestras obtenidas en las fases anteriores para realizar el entrenamiento del modelo y su subsecuente validación interna (Figura 2).¹ De esta última fase se obtienen las métricas para la evaluación del rendimiento del modelo (Tabla 1).

Figura 1. Representación esquemática de la arquitectura de los modelos de redes neuronales convolucionales

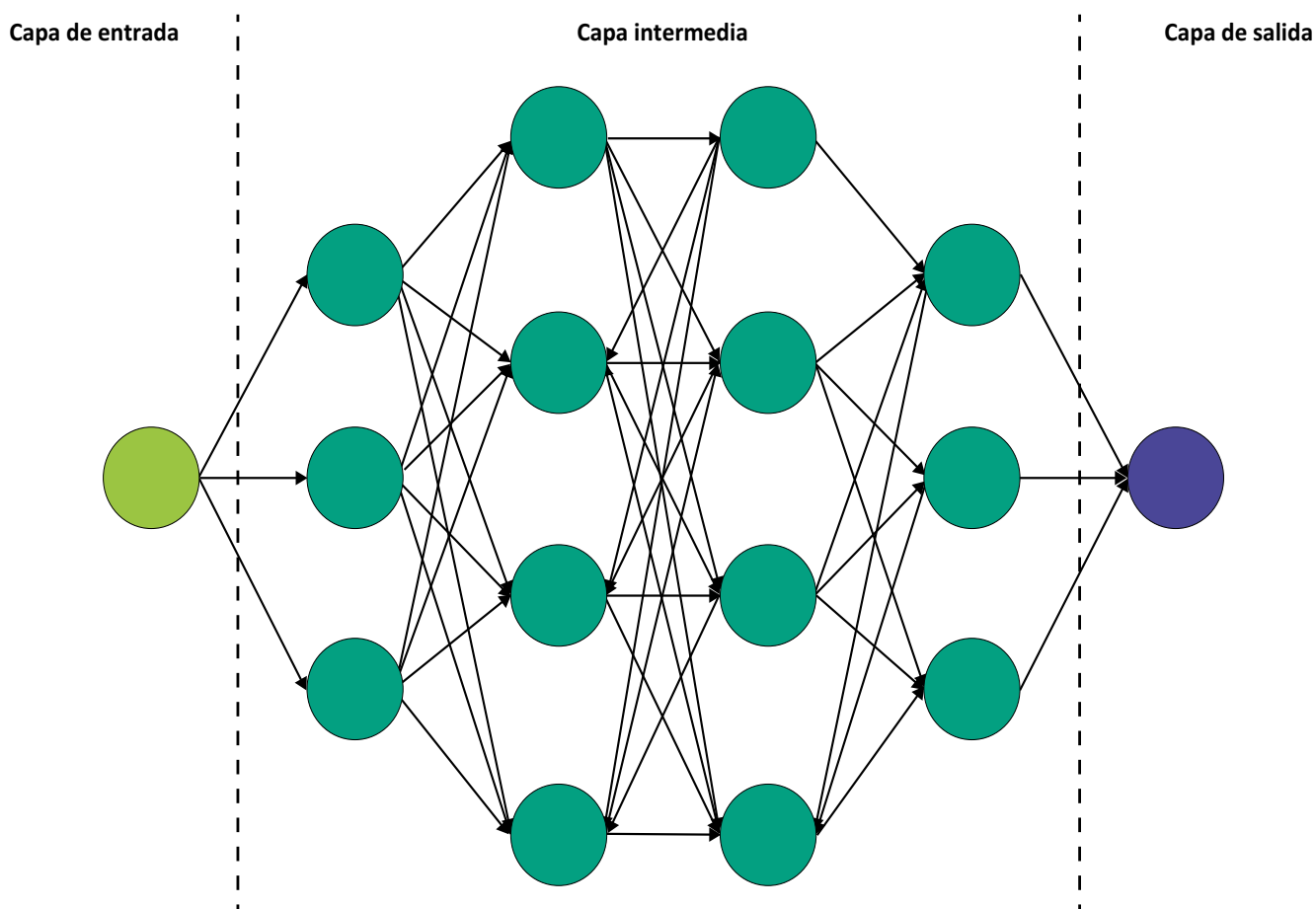
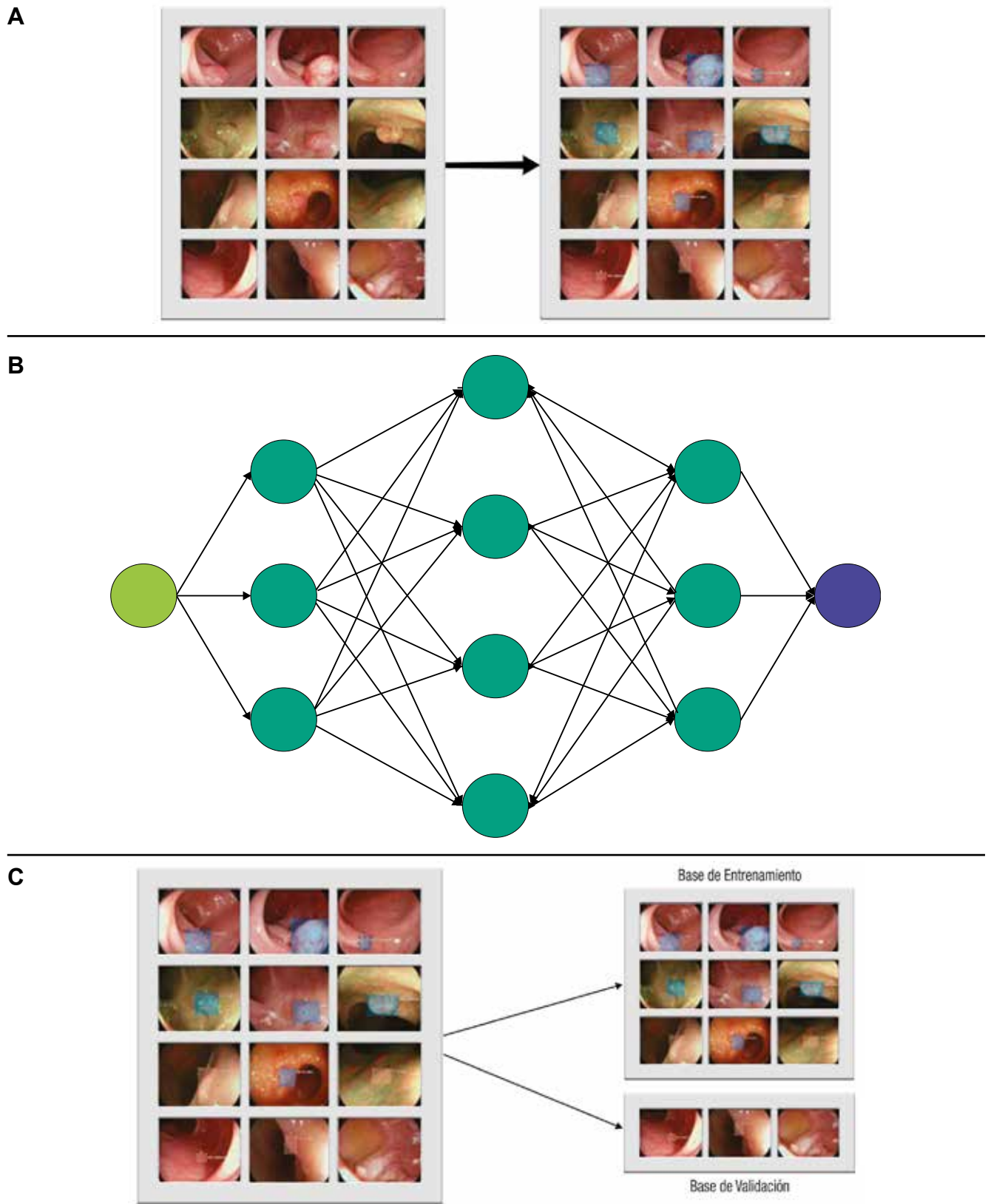


Figura 2. Fases de desarrollo de modelos de aprendizaje profundo



A) Representa la fase de recolección de datos y etiquetado. B) Establecimiento de la arquitectura del modelo. C) Fase de entrenamiento del modelo por medio del uso de las muestras obtenidas en la primera fase.

Tabla 1. Métricas obtenidas para evaluación del rendimiento de los modelos de aprendizaje profundo

Métricas	Definición
Sensibilidad:	La fracción de muestras positivas realmente clasificadas como positivas.
Especificidad:	La fracción de muestras negativas realmente clasificadas como negativas.
Precisión:	La fracción de las muestras correctamente clasificadas.
Puntaje F1:	La media armónica de sensibilidad y valor predictivo positivo.
IoU:	El rendimiento de la detección de objetos al comparar el "ground truth" con la detección del modelo.

IoU: Intersección sobre la unión.

Aplicaciones clínicas de la inteligencia artificial

Las aplicaciones clínicas de la IA han aumentado progresivamente en el área de la salud, incluida la gastroenterología. La IA ayuda a superar los múltiples retos que enfrentan los profesionales de la salud durante la adquisición, el análisis y la aplicación de conocimientos que aportan al diagnóstico, al manejo y al pronóstico de los pacientes.¹ Además, la automatización en la identificación y el reconocimiento de imágenes ayudan a la reducción de errores derivados de factores humanos (fatiga, carga laboral, entre otros).

El avance tecnológico ha permitido la realización de sistemas inteligentes que facilitan la detección o la estratificación de lesiones observadas durante procedimientos endoscópicos o imagenológicos. Estos se denominan dispositivos de detección asistida por computadora (*Computer-assisted detection device, CADe*) o de diagnóstico asistido por computadora (*Computer-assisted diagnostic device, CADx*).⁶ De esta manera, la aplicación de estos dispositivos en la endoscopia digestiva puede facilitar y aumentar la detección de lesiones durante los procedimientos y categorizar lesiones en benignas y malignas en tiempo real.

Inteligencia artificial en la endoscopia digestiva

Actualmente se están desarrollando varios modelos para la endoscopia digestiva alta y baja, y para procedimientos endoscópicos avanzados (colangioscopia y ultrasonido endoscópico, EUS) (Tabla 2). Estos modelos incluyen CADe y CADx con algoritmos característicos que permiten distintas funcionalidades como detectar estructuras anatómicas y lesiones específicas, ayudar en el entrenamiento de los médicos, servir como guía durante los procedimientos endoscópicos, reducir el número de

falsos negativos por medio de la caracterización y la estratificación de lesiones premalignas y malignas, entre otras.⁴

En los siguientes apartados se revisará información actualizada sobre los usos de la IA y su repercusión de acuerdo con el tipo de evaluación endoscópica.

Tabla 2. Aplicaciones de la inteligencia artificial en los distintos segmentos de la vía digestiva

Segmento	Aplicación de la IA
Esófago	<ul style="list-style-type: none"> • Diagnóstico temprano de EB. • Clasificación de lesiones en benignas y malignas. • Reconocimiento de carcinoma esofágico. • Invasión y profundidad de lesiones tumorales.
Estómago	<ul style="list-style-type: none"> • Monitorización de puntos ciegos. • Detección de lesiones sugestivas de cáncer. • Invasión y profundidad de lesiones tumorales. • Diferenciación entre tejido normal y patológico. • Detección de ERGE. • Detección de gastritis asociada a <i>H. pylori</i>.
Intestino delgado	<ul style="list-style-type: none"> • Detección y clasificación de lesiones (pólipos, sangrados, úlceras). • Estudio de enfermedad de Crohn y enfermedad celiaca.
Colon	<ul style="list-style-type: none"> • Detección y caracterización de pólipos • Asistencia durante procedimientos de tamizaje
Vía biliar	<ul style="list-style-type: none"> • Detección de áreas sugestivas de neoplasia • Detección de estructuras normales • Asistencia para obtención de biopsias

IA: Inteligencia artificial; EB: Esófago de Barrett; ERGE: Enfermedad por reflujo gastroesofágico.

Endoscopia digestiva alta

También conocida como esofagogastroduodenoscopia (EGD), es un procedimiento de gran importancia en el diagnóstico de las lesiones del tracto gastrointestinal superior.^{7,8} Sin embargo, la tasa de diagnóstico varía según el desempeño de cada endoscopista.⁹ Los errores durante la realización de la EGD son una de las principales causas en el fallo del diagnóstico de lesiones premalignas y de enfermedades esofagogastroduodenales severas. Los sistemas de IA han sido desarrollados para sobrepasar las barreras técnicas antes descritas. Su aplicación en el tracto digestivo superior va desde la localización anatómica hasta la detección y la evaluación de lesiones malignas y premalignas.⁷⁻¹⁰

Takiyama *et al.* desarrollaron un modelo de IA capaz de clasificar las estructuras anatómicas del tracto digestivo

superior, el cual ha mostrado un buen desempeño durante la identificación de la laringe (AUC 1,00), el esófago (AUC 1,00), el estómago (superior, medio e inferior) y el duodeno (AUC 0,99).⁷

Por otro lado, en el estudio multicéntrico realizado por Luo *et al.* se evaluó la precisión diagnóstica del modelo GRADIS para la detección de neoplasias del tubo digestivo superior (esófago y estómago). Este modelo alcanzó una precisión diagnóstica del 95,5% (IC 95%: 95,2 - 95,7) durante la validación interna. Al comparar su desempeño con el de los endoscopistas que tenían diferentes niveles de experiencia, la sensibilidad fue similar a la de los expertos (94,2% *versus* 94,5%) y mayor al compararla con los endoscopistas competentes (94,2% *versus* 85,8%) y los endoscopistas en entrenamiento (94,2% *versus* 72,2%).¹¹ Adicionalmente, la precisión diagnóstica de los expertos (92,8%) al usar el modelo de IA fue similar al grupo de endoscopistas competentes (93,4%) y endoscopistas en entrenamiento (90,4%).¹¹ Esto demuestra que la aplicación de IA disminuye la brecha entre expertos y no expertos.¹¹

Esófago

La precisión en el diagnóstico temprano del esófago de Barrett (EB) y de la neoplasia esofágica continúa siendo un reto incluso para muchos endoscopistas experimentados. Una vez identificado el EB, es necesaria la identificación de las regiones con displasia o adenocarcinoma temprano.

Los modelos de IA han sido designados con el fin de ayudar a los endoscopistas a mejorar la precisión en el diagnóstico de las lesiones mencionadas,¹² e incluyen sistemas para la clasificación de las neoplasias que aplican la magnificación en vivo con alta precisión (89,9%). Esto ha permitido el diagnóstico y la diferenciación tempranos de neoplasia en el EB.^{13,14}

El CADx desarrollado y validado por De Groof *et al.* permitió la clasificación de imágenes neoplásicas y no neoplásicas en el EB comparándola con el desempeño de 53 endoscopistas. El modelo superó el desempeño de los endoscopistas y obtuvo mayor precisión (88% *versus* 73%), sensibilidad (93% *versus* 72%) y especificidad (83% *versus* 74%).¹⁵

Por otro lado, debido a la gran importancia en el reconocimiento y el tratamiento oportuno del carcinoma de esófago, los investigadores han desarrollado sistemas que permiten la detección de las lesiones, así como también la invasión de la enfermedad.^{12,16,17} El carcinoma esofágico suele ser detectado en estadios tardíos, y las lesiones pequeñas suelen ser detectadas por endoscopistas muy experimentados.¹² La IA permite la detección de lesiones menores de 10 mm con una gran precisión (91,4%), incluso mayor que muchos endoscopistas denominados expertos

(> 15 años; 88,8%), con moderada experiencia (5 - 15 años; 81,6%), y baja experiencia (< 5 años; 77,2 %).¹⁶

La determinación de la profundidad de la lesión permite la selección del tratamiento (quirúrgico, endoscópico, farmacológico) y la determinación del pronóstico.¹⁷ Uno de los modelos con gran exactitud diagnóstica en la predicción de la profundidad de la invasión del carcinoma escamoso del esófago es el propuesto por Tokai *et al.* Los investigadores utilizaron 1751 imágenes para el entrenamiento y 291 para la validación; se obtuvo una sensibilidad del 84,1% y una exactitud diagnóstica del 80,9% en la estimación de la profundidad de la invasión del carcinoma. A su vez, al comparar el modelo con trece endoscopistas éste obtuvo mayor exactitud diagnóstica y mayor AUC.¹⁷

Estómago

El cáncer de estómago suele presentar síntomas inespecíficos durante sus estadios tempranos; debido a esto los pacientes suelen ser diagnosticados en estadios avanzados. El pronóstico del cáncer de estómago depende de la valoración de la profundidad y de la detección temprana de las lesiones. Se ha reportado que la detección temprana del cáncer de estómago puede aumentar la supervivencia a los 5 años a un 90%.¹⁸

De acuerdo con Menon *et al.*, la tasa de falsos negativos en el diagnóstico de cáncer de estómago temprano puede alcanzar el 25%.¹⁸ Los sistemas de automatización buscan disminuir este porcentaje con modelos que clasifiquen las imágenes del estómago en la EGD para monitorear puntos ciegos con gran precisión,¹⁹ modelos que detecten lesiones sugestivas de cáncer de estómago¹¹ y lesiones precancerosas,²⁰ y que además evalúen la profundidad de la invasión.²¹

La cromoendoscopia es uno de los métodos diagnósticos utilizados para la detección temprana de neoplasia gástrica. Sin embargo, durante una sesión endoscópica se pueden generar múltiples fotogramas de videos (*frames*) que convierten la revisión de ésta en una labor exhaustiva para los endoscopistas.

Para evitar pérdidas durante la evaluación, Ali *et al.* desarrollaron un CADx entrenado para clasificar fotogramas normales y patológicos a partir de la textura local y global. Este mostró una sensibilidad, especificidad, precisión y AUC del 91%, 82%, 87% y 0,91%, respectivamente.²² El modelo demostró ser una ayuda diagnóstica durante la detección temprana del cáncer gástrico ya que reduce el tiempo utilizado para la evaluación de la secuencia endoscópica.

El modelo estudiado por Wu *et al.*¹⁹ obtuvo alta precisión, especificidad y sensibilidad (92,5%, 94% y 91%, respectivamente) en la evaluación de no malignidad y superó a los endoscopistas expertos en esta tarea.¹⁹ Adicionalmente, durante procedimientos en tiempo real mostró

un gran rendimiento en la detección del adenocarcinoma gástrico temprano con monitorización de puntos ciegos. Junto a este, otros modelos de IA han mostrado una gran precisión durante el diagnóstico de la neoplasia gastrointestinal comparable con la de endoscopistas expertos.^{11,17}

Por otro lado, se han desarrollado modelos que además de identificar las lesiones neoplásicas también predicen su profundidad. Nagao *et al.* entrenaron a un modelo para predecir la invasión en profundidad del cáncer gástrico utilizando imágenes convencionales de luz blanca, imágenes de banda estrecha (*Narrow Band Imaging, NBI*), e imágenes con contraste índigo carmín.²¹ El modelo demostró alta precisión basada en los tres sistemas (imágenes de luz blanca 94,5%, NBI 94,3% e índigo carmín 95,5%).²¹

Zhu *et al.* (precisión total 89,16%) y Yoon *et al.* (sensibilidad 81,7%, especificidad 75,4%) han reportado que la precisión diagnóstica de su modelo para evaluar la profundidad de la invasión es comparable con la de otros métodos convencionales. Las ventajas de utilizar estos modelos radican en la evaluación más objetiva de las características macroscópicas de las lesiones, con la disminución del uso de otras técnicas invasivas como el EUS.^{21,23,24}

Además de la identificación de neoplasia, otras aplicaciones incluyen la detección de la enfermedad por reflujo gastroesofágico (ERGE)²⁵ y de gastritis asociada a *Helicobacter pylori*.²⁶ Modelos desarrollados para la asistencia durante la clasificación de la ERGE con NBI han alcanzado una precisión diagnóstica total del 99,2% para lesiones grado A-B, 100% para lesiones grado C-D y 100% en el grupo control. Por lo tanto, se consideran de gran utilidad para la asistencia durante la detección de lesiones compatibles con ERGE, y aumentan la precisión diagnóstica en endoscopistas en entrenamiento.²⁵

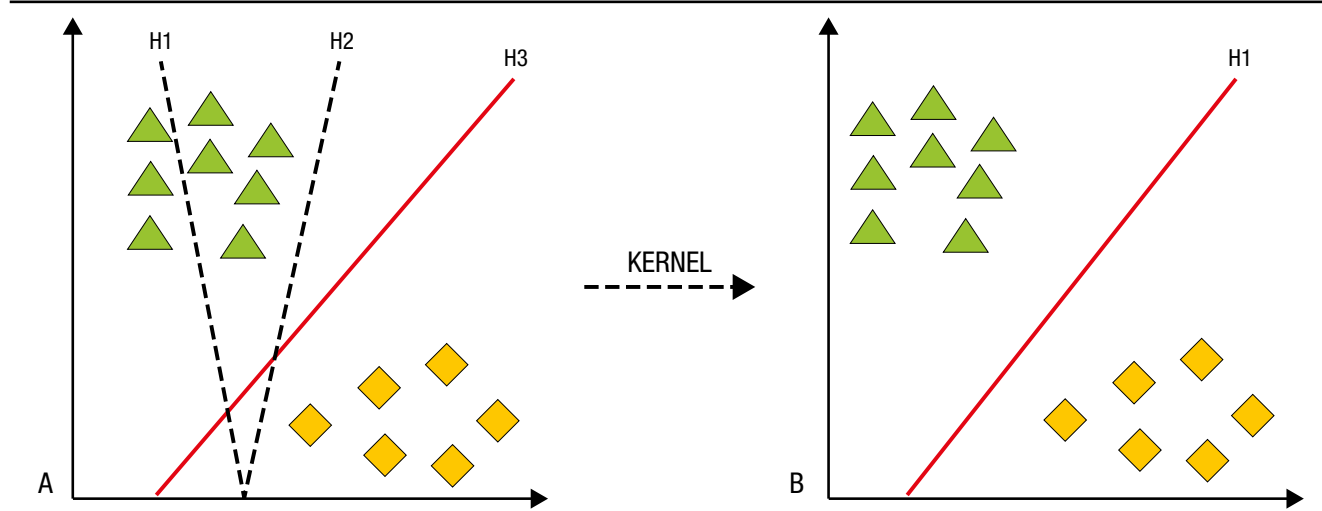
Por otro lado, la red neuronal diseñada para la predicción de *Helicobacter pylori* en imágenes endoscópicas diagnosticó correctamente el 80% de los casos negativos, el 84% de los casos erradicados y el 48% de los casos positivos. Los autores de este estudio resaltan la utilidad de este modelo en la identificación de pacientes que requieran alguna prueba de confirmación de *Helicobacter pylori* a partir de los resultados endoscópicos, y recomiendan su uso como asistencia diagnóstica.²⁶

Intestino delgado

La cápsula endoscópica, procedimiento no invasivo, permite la detección y la clasificación de lesiones (sangrado, úlceras y pólipos), la valoración de la motilidad intestinal y la evaluación de la enfermedad celíaca, entre otras patologías que afectan especialmente al intestino delgado. Sin embargo, la evaluación de la gran cantidad de imágenes obtenidas (> 60 000 imágenes) y la dificultad para el manejo direccional de la cápsula (dependiente totalmente de la peristalsis del tubo digestivo) hacen que el estudio sea prolongado (de 45 minutos a 8 horas) y tedioso. Para sobrellevar las dificultades técnicas derivadas de este procedimiento se ha buscado su automatización a partir de la IA.²⁸

Los modelos de IA para cápsula endoscópica han basado su desarrollo en DL. La clasificación y categorización de las imágenes se realiza a partir de máquinas de vector en soporte (*SVM*, por sus siglas en inglés). Por medio de estos modelos, los datos son separados por hiperplanos de dos o más dimensiones. Luego de utilizar los parámetros de Kernel, se consigue el hiperplano «óptimo» que crea «fronteras» para la categorización de los datos (Figura 3).^{27,28} Posterior a la categorización se forman los algoritmos de DL que crean así las redes neuronales artificiales.^{27,29}

Figura 3. Máquinas de vector en soporte (*SVM*) para clasificación de datos



A) Clasificación linear con hiperplanos (H1, H2, H3) por SVM. B) Datos clasificados con un hiperplano (H1) óptimo posterior a los parámetros de Kernel.

Actualmente, los modelos de IA para cápsula endoscópica incluyen el seguimiento de la cápsula, la detección de pólipos, sangrado, úlceras, y el estudio de patologías propias del intestino delgado como la enfermedad celíaca y la enfermedad de Crohn.²⁶

En general, en el tracto gastrointestinal, los modelos incorporados en la cápsula endoscópica permiten su seguimiento y localización en los distintos segmentos (boca, estómago, intestino delgado y colon), previa anulación de *frames* con «ruido» (heces, burbujas, entre otros) con una sensibilidad y especificidad > 88%.³⁰ La evaluación de las imágenes de acuerdo con su localización topográfica ahorra tiempo de estudio y aumenta la precisión diagnóstica.³⁰

En el estudio multicéntrico publicado por Ding *et al.*³⁰ se desarrolló una red neural convolucional (CNN) para la identificación de imágenes normales, inflamación, úlceras, pólipos, linfangiectasias, sangrado, enfermedad vascular, divertículos y parásitos, entre otros. Al comparar los resultados del modelo de los gastroenterólogos participantes, se obtuvo una sensibilidad del 98,9% (IC 95%: 99,7-99,9) *versus* 74,6% (IC 95%: 73,1-76) en la identificación de anomalías por paciente y una sensibilidad del 99,9% (IC 95%: 99,6 - 99,9) *versus* 76,9% (IC 95%: 75,6 - 78,2) en el análisis por lesión. Asimismo, el tiempo de lectura por paciente fue menor en el grupo de la CNN *versus* los gastroenterólogos (5,9 ± 2,23 minutos *versus* 96,6 ± 22,53 minutos, $p < 0,001$). A partir de estos resultados los investigadores concluyeron que la aplicación de la IA en la cápsula endoscópica es una herramienta importante para ayudar a los gastroenterólogos a analizar las imágenes capturadas con este dispositivo de manera más eficiente y precisa.³¹

Los modelos para la detección de sangrado en tiempo real alcanzan una precisión diagnóstica de hasta el 99%.³²⁻³⁴ El modelo de Aoki *et al.* detectó sangrado gastrointestinal con una sensibilidad, especificidad y precisión del 96,6%, 99,9% y 99,9%, respectivamente.³¹ Otros modelos desarrollados ayudan a la estratificación y predicción del riesgo de hemorragias recurrentes con el fin de brindar un tratamiento oportuno y evitar endoscopias innecesarias.^{35,36}

Para la detección de úlceras y erosiones, el modelo de Wang *et al.* alcanzó una precisión diagnóstica del 92,1%.³⁷ Para la identificación de tumores la precisión diagnóstica parte del 86%, con una sensibilidad que oscila entre el 88% y el 97%, y una especificidad entre el 81% y el 96%.³⁸

En un metaanálisis reciente sobre la CNN en cápsulas endoscópicas, se obtuvo una sensibilidad agrupada

y una especificidad del 96% (IC 95%: 91-98) y 97% (IC 95%: 93-99) en la detección de úlceras y erosiones; 97% (IC 95%: 93-99) y 100% (IC 95%: 99-100) en la identificación de sangrado gastrointestinal; 97% (IC 95%: 82-99) y 98% (IC 95%: 92-99) en la detección de cáncer y pólipos.³⁷

Adicionalmente, los modelos desarrollados para la identificación de enfermedad inflamatoria intestinal por medio de cápsula endoscópica alcanzan en la actualidad altos niveles de precisión (83,3% al 90,8%)³⁹ y permiten también el reconocimiento de patrones ocultos de la enfermedad.⁴

En comparación con los endoscopios actuales con los que se obtienen imágenes de muy alta calidad, la calidad de las imágenes de la cápsula endoscópica es baja.²⁷ Sin embargo, los modelos de IA para la cápsula endoscópica tienen la ventaja de contar con una base de datos robusta, alimentada con una gran cantidad de imágenes útiles para la creación de la CNN.

Endoscopia digestiva baja

Intestino grueso

El cáncer colorrectal (CCR) se considera en la actualidad una de las principales causas de muerte relacionadas con el cáncer en hombres y mujeres.⁴⁰ Para la identificación temprana de lesiones premalignas (pólipos y adenomas) la colonoscopia sigue siendo un estudio imprescindible. Sin embargo, de acuerdo con la literatura, se puede perder la identificación de aproximadamente el 25% de estas lesiones, incluso en manos expertas.⁴¹ Las lesiones premalignas no detectadas aumentan el riesgo de desarrollar CCR.

Para la detección y caracterización de los pólipos se han desarrollado sistemas automatizados. Los primeros sistemas desarrollados utilizaban una cantidad reducida de imágenes y tenían como consecuencia una precisión diagnóstica baja (72%).⁴² Posteriormente, se han entrenado nuevos modelos con mayor precisión (> 95%), que utilizan un mayor número de imágenes. Esto permite que en la actualidad se pueda evaluar mejor los pólipos y los adenomas diminutos (≤ 5 mm); y predecir el pronóstico del CCR, la supervivencia de los pacientes y la presencia de invasión a distancia.^{42,43}

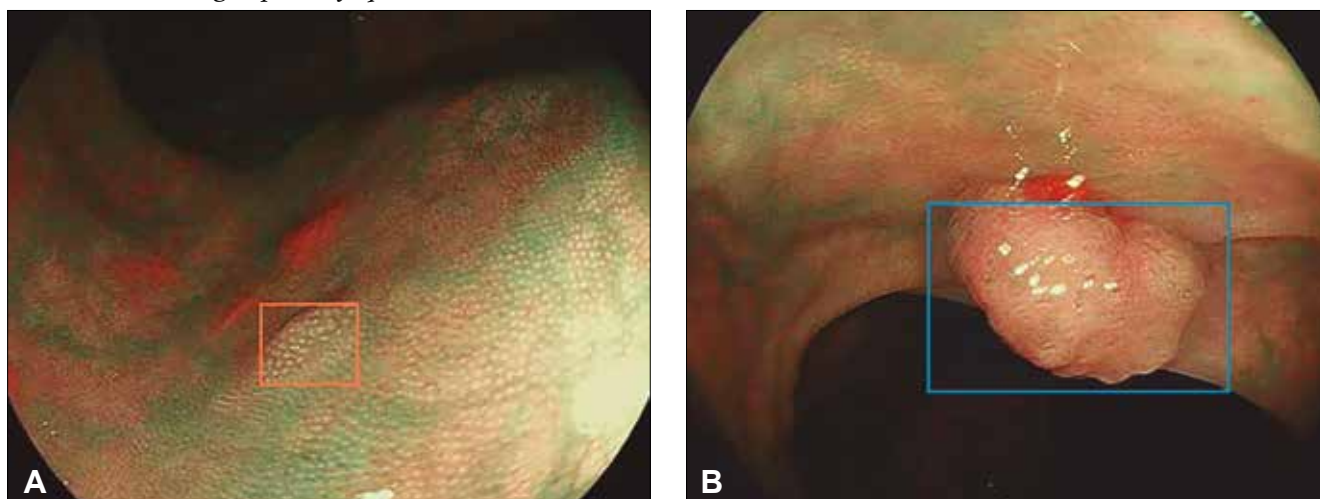
Un metaanálisis evaluó el rendimiento de los CADE en la detección de neoplasia colorrectal.⁴⁴ Los autores obtuvieron mayor tasa de detección de adenomas (*adenoma detection rate, ADR*) en los grupos en los que se usaron los CADE en comparación con sus grupos control (36,6% *versus* 25,2%; RR: 1,44; IC 95%: 1,27-1,62; $p < 0,01$). Asimismo, la detección de adenomas por colonoscopia fue superior en los grupos en los que se usaron CADE en

comparación con el grupo control (50,3% *versus* 34,6%; RR: 1,70; IC 95%: 1,53-1,89; $p < 0.01$). Los autores no evidenciaron una diferencia significativa en la eficiencia de la colonoscopia (tiempo de retirada) entre los grupos en que se usaron CADe y el grupo control.⁴⁴

Robles-Medranda *et al.* estudiaron la eficiencia de la colonoscopia asistida por IA para la detección de pólipos y adenomas durante los estudios de *screening* de CCR (Figura 4).⁴⁵ Los resultados obtenidos por los autores

fueron comparados de acuerdo con el nivel de experiencia de los endoscopistas (expertos *versus* no expertos). Con la asistencia de la IA durante los procedimientos endoscópicos, se obtuvo un aumento de la ADR y tasa de detección de pólipos del 16,5% al 18,2% y del 50,4% al 60%, respectivamente. De acuerdo con el nivel de experiencia, el aumento de la ADR fue evidente en el grupo *junior* (10,8% a 16,2%), que se aproximó al nivel de los expertos.⁴⁵

Figura 4. Detección de lesión polipoide en colon utilizando un dispositivo de detección asistido por computadora (CADe) (AITROL, *mdconsgroup*, Guayaquil, Ecuador)



A: Lesión polipoide tipo I según la clasificación de NICE (delimitación naranja). B: Lesión polipoide tipo II según la clasificación de NICE (delimitación azul).

Además, se han diseñado CADx con capacidad de caracterizar de manera inmediata los pólipos utilizando tecnología de imágenes más allá de la endoscopia con luz blanca y NBI magnificada, y la endomicroscopia confocal. En este campo, un modelo desarrollado por Sánchez-Montes *et al.* para la predicción de la clasificación histológica de los pólipos, alcanzó una precisión diagnóstica, una sensibilidad y una especificidad del 91,1%, 92,3%, y 89,2%, respectivamente.⁴⁶

Adicionalmente, características como la depresión de la lesión, la convergencia de los pliegues, y el patrón capilar irregular y heterogéneo se asocian con la invasión en profundidad de las lesiones premalignas. Los modelos actuales de los CADe para la identificación de las características antes descritas resultan atractivos antes de determinar el tipo de tratamiento a realizar (por ejemplo, la resección endoscópica de la mucosa).⁴

Técnicas endoscópicas avanzadas

A pesar de que en la actualidad se encuentran disponibles algunos métodos diagnósticos, como colangioscopia, colangiopancreatografía endoscópica retrógrada (CPRE)

o EUS para el estudio del sistema biliopancreático, existen dificultades en la diferenciación de lesiones y discrepancia entre los evaluadores.

Colangioscopia

Como la colangioscopia es una técnica endoscópica avanzada, relativamente nueva y sin guías de entrenamiento establecidas, la impresión visual entre los operadores es muy variable.⁴⁷ Se han creado varias clasificaciones para detectar malignidad basándose en las características macroscópicas de las lesiones de la vía biliar durante el procedimiento,⁴⁸⁻⁵¹ con la intención de disminuir esta variabilidad entre los observadores. Sin embargo, estas clasificaciones no han logrado ese objetivo.⁴⁷

Recientemente, se han empezado a desarrollar modelos de IA con la finalidad de ayudar a los operadores a detectar lesiones malignas y a obtener biopsias. Los primeros modelos de IA en la colangioscopia fueron desarrollados para la detección de vasos sanguíneos tortuosos, pero tenían la desventaja de estar limitados a imágenes y no era posible su utilidad en casos en vivo.^{52,53} Otra limitación de dichos modelos es su falta de validación clínica.

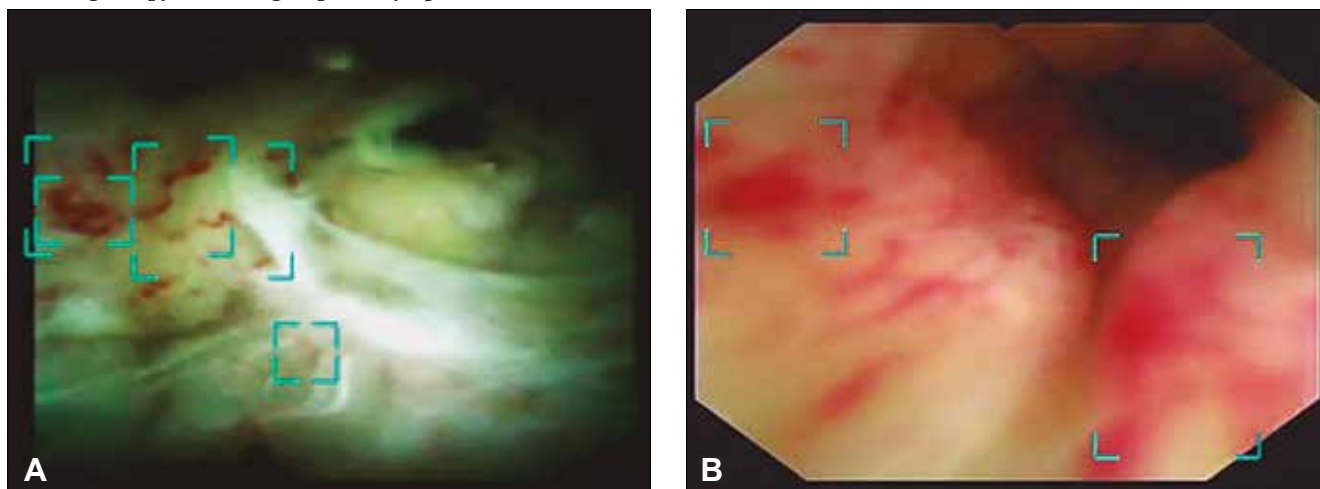
Sin embargo, obtuvieron métricas de validación interna bastante altas. En su primer estudio, Mascarenhas *et al.* desarrollaron un modelo con 6475 imágenes de colangioscopia obtenidas de 85 pacientes. Durante la validación con *frames* se obtuvo una sensibilidad, una especificidad, un valor predictivo positivo (VPP) y valor predictivo negativo (VPN) de 99,3%, 99,4%, 99,6% y 98,7%, respectivamente.⁵³ Luego, desarrollaron un nuevo trabajo en el que duplicaron el número de imágenes obtenidas del mismo número de pacientes (de 6475 a 11855); esta vez evaluaron la precisión del modelo para distinguir entre lesiones benignas y malignas. Las métricas obtenidas fueron: precisión diagnóstica (94,9%), sensibilidad (94,7%), especificidad (92,1%), VPP (94,8%) y VPN (84,2%).⁵² Se resalta que a pesar de que el modelo obtuvo excelentes parámetros de validación interna, estos no deben ser extrapolados a su utilidad real en casos clínicos en vivo.

Posteriormente, se realizaron dos estudios con modelos de IA en tiempo real. El primer estudio, realizado por

Marya *et al.*, evalúa la aplicación clínica del modelo y lo compara con los resultados de las biopsias y la citología. En el estudio se observa que el modelo de CNN obtuvo mayores sensibilidad (93,3%), especificidad (88,2%) y precisión (90,6%) en comparación con las biopsias, que obtuvieron 35,7% de sensibilidad, 100% de especificidad y 60,9% de precisión; y con la citología, que obtuvo 40%, 100% y 62,5% de sensibilidad, de especificidad y de precisión, respectivamente.⁵⁴ Una de las limitaciones del estudio fue que sólo un operador estuvo a cargo de realizar las anotaciones de alrededor de 2 millones de imágenes, lo que puede conllevar a errores por fatiga y a la aparición de un sesgo.

Robles-Medrandra *et al.* desarrollaron su propio modelo de CNN capaz de detectar lesiones neoplásicas en videos pregrabados y en tiempo real (Figura 5). Posterior a las etapas de implementación de modelos de IA (recolección, anotación y diseño del modelo), se realizó la validación interna y luego una comparación clínica entre el modelo y los endoscopistas (expertos y no expertos).⁵⁵

Figura 5. Detección de imágenes sugestivas de neoplasia en colangioscopia por medio de la inteligencia artificial (AIWorks-Cholangioscopy, Mdconsgroup, Guayaquil, Ecuador)



A: Área potencialmente neoplásica delimitada por el sistema de IA usando un colangioscopio Spyglass (Boston Scientific, Marlborough, MA, USA). B: Área potencialmente neoplásica delimitada por el sistema de IA usando un colangioscopio de 7Fr eyeMAX (Micro-Tech, Nanjing, China).

Este estudio multicéntrico se llevó a cabo en dos fases. La primera fase fue la de desarrollo y validación del modelo de IA, denominado AIworks-Cholangioscopy (mdconsgroup, Guayaquil, Ecuador). La primera versión de este modelo fue desarrollada a partir de 81080 imágenes de 23 pacientes, y obtuvo una media de precisión (mAP) de 0,298, un puntaje F1 de 0,280, la intersección sobre la unión fue de 32,3% y la pérdida total fue de 0,1034. A pesar de ser resultados aceptables, la detección de fotogramas por segundo (*frames per second, FPS*) fue baja y llegó a un promedio de 5. Esta baja detección de FPS

impidió que el modelo se utilice para la detección de lesiones en tiempo real. Su validación interna obtuvo una sensibilidad, una especificidad, un VPP y un VPN de 98%, 95%, 98%, y 94%, respectivamente. Posteriormente, al aumentar el número de casos e imágenes disponibles para el entrenamiento (de 81080 a 198941), junto con una mejor calidad de imagen, las métricas de validación interna aumentaron drásticamente: mAP de 0,298 a 0,880, puntaje F1 de 0,280 a 0,738, IoU de 32,3% a 83,2%, pérdida total de 0,1034 a 0,0975. La sensibilidad, la especificidad, el VPP y el VPN para detectar

lesiones neoplásicas en imágenes tuvieron resultados similares a los obtenidos por Mascarenhas *et al.*: 98,6%, 98%, 89,2%, y 99,2%, respectivamente.^{52,53,55} Sin embargo, al realizar la segunda fase para la validación clínica en 170 pacientes, se observó que estos valores de precisión diagnóstica disminuyeron y se aproximaron a los de los endoscopistas. Durante la comparación del modelo con la impresión visual de los endoscopistas (expertos y no expertos) que usaron dos clasificaciones de lesiones neoplásicas (clasificación CRM y de Mendoza),^{48,50} se evidenció que el modelo de IA fue superior tanto al de los expertos como al de los no expertos.⁵⁵ Este estudio demuestra la importancia de realizar una validación clínica y de no extrapolar los resultados de una validación interna a la clínica como resultado final.

Ultrasonido endoscópico

Para el manejo diagnóstico y terapéutico de patologías biliopancreáticas, el EUS se considera superior a la tomografía y a la resonancia magnética debido a su mayor precisión diagnóstica y a su capacidad de obtener imágenes de mejor calidad.⁵⁶ Sin embargo, estos procedimientos tienen limitaciones como la baja sensibilidad para diferenciar la neoplasia papilar mucinosa intraductal (NPMI) benigna de la maligna, y la baja especificidad para diferenciar las lesiones malignas de la pancreatitis crónica.⁵⁷ Otra limitación de este procedimiento es su dependencia de los operadores, por lo que los endoscopistas menos experimentados pueden no apreciar las diferencias entre pancreatitis crónica y malignidad pancreática.⁵⁷ Por esta razón, la aplicación de IA en este tipo de procedimientos sería beneficiosa y podría influenciar la calidad del EUS realizado por endoscopistas expertos o en entrenamiento.⁵⁸

Se han desarrollado varios estudios para evaluar y comparar la precisión diagnóstica del EUS asistido por IA contra el EUS tradicional en la detección del cáncer pancreático, y distinguir entre lesiones crónicas y tejido normal.⁵⁵ Un estudio realizado por Norton *et al.*, demostró que su modelo de IA tuvo mayor sensibilidad (100%) para la diferenciación entre malignidad e inflamación, sin embargo, la precisión diagnóstica fue similar entre el modelo (80%), el endoscopista ciego a los resultados del procedimiento (83%) y el procedimiento tradicional (85%).⁵⁹ Los autores demostraron la posibilidad de aplicar modelos de IA para la interpretación de imágenes en el EUS y la capacidad de distinción entre malignidad y condiciones crónicas; de esta forma resolvieron una de las limitaciones mencionadas anteriormente.

Con el paso del tiempo, se desarrollaron nuevas técni-

cas de EUS que se incluyeron como parte del manejo de los pacientes. En paralelo, se diseñaron diferentes tipos de IA para diferenciar las lesiones pancreáticas benignas de las malignas. Săftoiu *et al.* evaluaron la aplicación de un modelo de IA que permite diferenciar las lesiones benignas de las malignas usando la elastografía, una técnica que evalúa la rigidez y la elasticidad de los tejidos. Obtuvieron altas sensibilidad (91,4%), especificidad (87,9%) y precisión diagnóstica (89,7%).⁶⁰ Estos resultados indican la posibilidad de aplicar modelos de IA en la elastografía en casos en los que la aspiración con aguja fina guiada por EUS tenga resultados negativos.

Posteriormente, los mismos autores compararon el uso de la elastografía con IA y la elastografía sin IA; observaron que la precisión diagnóstica del primer procedimiento (EUS + IA) fue mayor (AUC: 94%) que la elastografía tradicional (AUC: 85%). Esto sugiere que los modelos de elastografía basados en CNN pueden proporcionar soporte para la toma de decisiones al proveer una interpretación de imágenes rápida y precisa, cuando se los compara con la elastografía sin IA.⁶¹

Otra de las limitaciones del EUS que ha sido evaluada es la diferenciación de las NPMI benignas de las malignas. Las NPMI son las precursoras de los adenocarcinomas pancreáticos, y una vez que las lesiones progresan a lesiones invasivas, el pronóstico de los pacientes disminuye.⁶²

Kuwahara *et al.* desarrollaron un modelo de DL e investigaron si el análisis preoperatorio de imágenes de NPMI por EUS que utiliza IA puede predecir malignidad. Los autores compararon la interpretación de las lesiones con los diagnósticos preoperatorios de los endoscopistas, las técnicas predictivas convencionales y otras técnicas de EUS.⁶³ Como resultado, observaron que el modelo de IA obtuvo una AUC de 91% para predecir malignidad. En la comparación de la precisión diagnóstica entre el modelo y el endoscopista, se observó que el modelo fue superior (94% *versus* 56%, respectivamente).⁶³

Otra aplicación de la IA en el EUS es el diagnóstico diferencial de lesiones subepiteliales. Hirai *et al.* evaluaron un modelo de DL que permitía su clasificación en imágenes de EUS. Recolectaron imágenes de lesiones subepiteliales de la vía digestiva alta, incluyendo tumores del estroma gastrointestinal (GIST), leiomiomas, schwannomas, tumores neuroendocrinos y páncreas ectópico. El modelo pudo clasificar las lesiones anteriormente mencionadas con una precisión del 86,1%, la cual fue mucho mayor que la de los endoscopistas participantes.⁶⁴ La sensibilidad, la especificidad y la precisión diagnóstica para diferenciar los GIST de las demás lesiones fue del 98,8%, 67,6%, y 89,3%, respectivamente.⁶⁴ Otros estu-

dios, incluyendo metaanálisis que compararon la precisión de otro modelo de CNN para diferenciar los GIST de otras lesiones, tuvieron resultados similares.⁶⁵

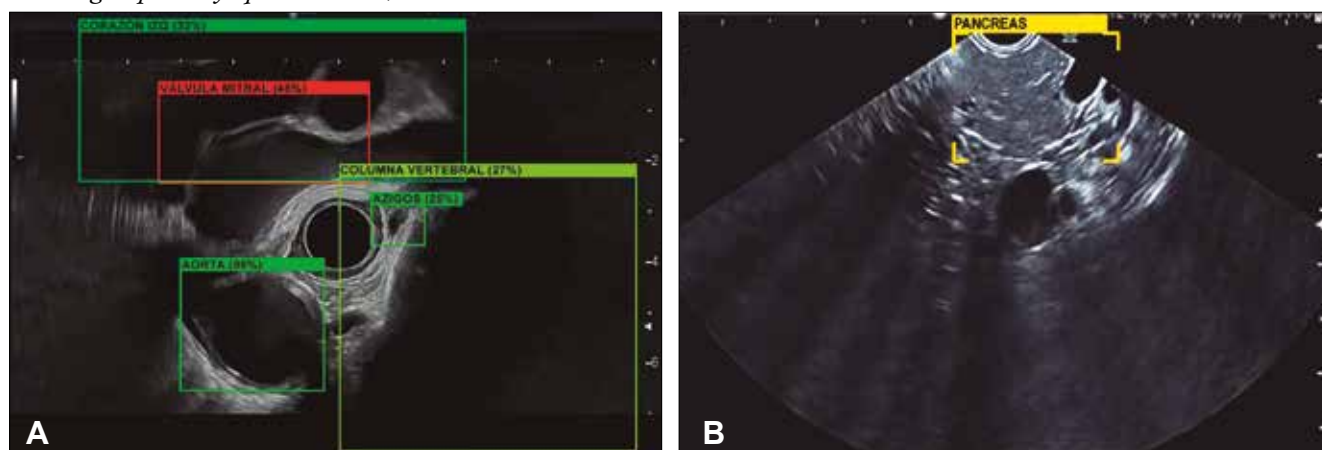
Debido a la tendencia de aumento de neoplasias biliopancreáticas a nivel mundial, es importante poder realizar una adecuada diferenciación entre las lesiones tumorales malignas y las benignas o tejidos normales. La aplicación de la IA en la endoscopia biliopancreática ha sido evaluada por mucho tiempo y mostró resultados prometedores. La IA y su aplicación en medicina se deben considerar beneficiosas porque ayudarán a sobrepasar las limitaciones de estos procedimientos complejos (colangioscopia y EUS).

Aplicaciones futuras de la IA en la endoscopia digestiva

La IA ha demostrado ser útil para detectar y clasificar lesiones durante los distintos tipos de procedimientos endoscópicos disponibles y estudiados. Sin embargo, la aplicación de modelos inteligentes puede ir más allá de la detección y el diagnóstico.

Carlos Robles-Medranda *et al.*, desarrollaron un sistema de EUS basado en modelos de CNN entrenados para detectar estructuras anatómicas normales en las distintas ventanas evaluadas durante este procedimiento avanzado (mediastinal, gástrica y duodenal) (Figura 6), lo que permitió la identificación de 20 estructuras anatómicas con alta sensibilidad y especificidad (Tabla 3).⁶⁶

Figura 6. Modelo de inteligencia artificial de ultrasonido endoscópico detectando estructuras anatómicas (AIWorks-EUS, mdconsgroup, Guayaquil, Ecuador)



A: Dispositivo de detección asistida por computadora (CADE) detectando el corazón izquierdo, la válvula mitral, la columna vertebral, la vena ácigos y la aorta usando una sonda de ultrasonido endoscópico radial; B: Dispositivo de detección asistida por computadora (CADE) que detecta el páncreas usando una sonda de ultrasonido endoscópico lineal.

Tabla 3. Estructuras anatómicas observadas por el sistema de detección asistido por computadora AIWorks-EUS (Mdconsgroup, Guayaquil, Ecuador)

Ventanas estudiadas	Estructuras anatómicas
Mediastínico	Aorta, Columna Vertebral, Arco Aórtico, Tráquea, Espacio Aórtico-Pulmonar
Gástrica	Riñón izquierdo, Hígado, Bazo, Cuerpo y Cola del Páncreas, Tronco celíaco, Arteria y Vena Esplénicas, Vena Cava Inferior, Glándula Suprarrenal
Duodenal	Riñón derecho, Vesícula biliar, Colédoco, ampolla de Vater, vena porta

La precisión diagnóstica del modelo para detectar dichas estructuras fue mayor al 95%. Esto indica que los modelos de IA no solo tienen utilidad para detectar patologías, sino también para detectar estructuras normales, lo cual puede beneficiar el entrenamiento de los endoscopistas y reforzar el conocimiento de aquellos con más experiencia. La aplicación de la IA en el entrenamiento de los endoscopistas fue evaluada por Zhang *et al.*, quienes demostraron que aquellos endoscopistas entrenados con IA tuvieron una menor curva de aprendizaje y mejores resultados que aquellos entrenados tradicionalmente.⁶⁷

Adicionalmente, por medio de modelos inteligentes, es posible automatizar el proceso de obtención de imágenes para la generación de reportes y como medidor de calidad de las pruebas.⁶⁸ Esto se puede llevar a cabo por medio de la unión de varias funciones y su aplicación simultánea, como la detección del órgano valorado junto con la detección de lesiones dentro de ese órgano. De esta

manera, al final del procedimiento, se puede generar de manera automática el reporte del estudio realizado.^{68,69} Modelos como el AI-EARS y el ISRGS han demostrado buena precisión diagnóstica para identificar los órganos y las lesiones, generando un reporte automático utilizando IA tanto en vía digestiva alta⁶⁸ como en baja.⁶⁹

La aplicación de modelos de IA para entrenar a endoscopistas en procedimientos endoscópicos avanzados puede aumentar la eficacia de dichos endoscopistas y disminuir el número de procedimientos necesarios para alcanzar competencia.

Conclusión

A pesar de que la automatización de los modelos durante la endoscopia busca reducir los riesgos derivados de factores humanos y ambientales, entre otros, el éxito los modelos de IA dependerá de la calidad y la cantidad de información utilizada durante su entrenamiento y validación. Asimismo, cobra gran importancia la validación externa por medio de estudios multicéntricos e internacionales con endoscopistas expertos antes de la generalización y la universalización de los resultados del entrenamiento.

Los modelos de IA en la endoscopia digestiva tienen el potencial de mejorar la percepción visual de los endoscopistas, y acercar la precisión entre aquellos con menos experiencia y los considerados expertos. Además, pueden ser de gran ayuda en la detección de lesiones y en la evaluación de la invasión en los tejidos. Se vislumbra un futuro en el cual los enfoques de formación tradicionales serán superados y los centros médicos en todo el mundo podrán implementar mejoras en la capacitación de los profesionales, mediante la aplicación de tecnologías precisas que reduzcan la curva de aprendizaje en estos procedimientos. Esto beneficiará a los profesionales menos experimentados y reducirá la brecha entre los endoscopistas expertos y los de menor experiencia, mejorando la competencia de los endoscopistas.

Propiedad intelectual. Los autores declaran que los datos, las figuras y las tablas presentes en el manuscrito son originales y se realizaron en sus instituciones pertenecientes.

Financiamiento. Los autores declaran que no hubo fuentes de financiación externas.

Conflicto de interés. Carlos Robles-Medranda es un líder de opinión y consultor para Pentax Medical, Steris, Microtech, G-Tech Medical Supply, CREO Medical, EndoSound y Mdconsgroup. Los demás autores no declaran ningún conflicto de interés.

Aviso de derechos de autor



© 2023 *Acta Gastroenterológica Latinoamericana*. Este es un artículo de acceso abierto publicado bajo los términos de la Licencia Creative Commons Attribution (CC BY-NC-SA 4.0), la cual permite el uso, la distribución y la reproducción de forma no comercial, siempre que se cite al autor y la fuente original.

Cite este artículo como: Baquerizo-Burgos J, Egas-Izquierdo M, Cunto D y col. La era de la endoscopia inteligente: cómo la inteligencia artificial potencia la endoscopia digestiva. *Acta Gastroenterol Latinoam*. 2023;53(3):211-225. <https://doi.org/10.52787/agl.v53i3.339>

Referencias

1. Ramesh AN, Kambhampati C, Monson JRT, Drew PJ. Artificial Intelligence in Medicine. *Ann R Coll Surg Engl*. 2004;86(5):334-8.
2. Tonozuka R, Mukai S, Itoi T. The Role of Artificial Intelligence in Endoscopic Ultrasound for Pancreatic Disorders. *Diagnostics*. 2021;11(1).
3. Minoda Y, Ihara E, Fujimori N, Nagatomo S, Esaki M, Hata Y, et al. Efficacy of Ultrasound Endoscopy with Artificial Intelligence for the Differential Diagnosis of Non-Gastric Gastrointestinal Stromal Tumors. *Sci Rep*. 2022;12(1).
4. Kröner PT, Engels MM, Glicksberg BS, Johnson KW, Mzaik O, van Hooft JE, et al. Artificial Intelligence in Gastroenterology: A State-Of-The-Art Review. *World J Gastroenterol*. 2021;27(40):6794-6824.
5. Yu KH, Beam AL, Kohane IS. Artificial Intelligence in Healthcare. *Nat Biomed Eng*. 2018;2(10):719-31.
6. Firmino M, Angelo G, Morais H, Dantas MR, Valentim R. Computer-Aided Detection (CADe) and Diagnosis (CADx) System for Lung Cancer with Likelihood of Malignancy. *Biomed Eng Online*. 2016;15(1):2.
7. Takiyama H, Ozawa T, Ishihara S, Fujishiro M, Shichijo S, Nomura S, et al. Automatic anatomical classification of esophagogastroduodenoscopy images using deep convolutional neural networks. *Sci Rep*. 2018;8:7497.
8. Van der Sommen F, Zinger S, Curvers WL, Bisschops R, Pech O, Weusten BL, et al. Computer-Aided Detection of Early Neoplastic Lesions in Barrett's Esophagus. *Endoscopy*. 2016;48:617-624.
9. Kumagai Y, Takubo K, Kawada K, Aoyama K, Endo Y, Ozawa T, et al. Diagnosis Using Deep-Learning Artificial Intelligence Based on the Endocytoscopic Observation of the Esophagus. *Esophagus*. 2019;16:180-187.
10. Horie Y, Yoshio T, Aoyama K, Yoshimizu S, Horiuchi Y, Ishiyama A, et al. Diagnostic Outcomes of Esophageal Cancer by Artificial Intelligence Using Convolutional Neural Networks. *Gastrointest Endosc*. 2019;89:25-32.

11. Luo H, Xu G, Li C, He L, Luo L, Wang Z, *et al.* Real-Time Artificial Intelligence for Detection of Upper Gastrointestinal Cancer by Endoscopy: A Multicentre, Case-Control, Diagnostic Study. *Lancet Oncol.* 2019;20(12):1645-1654.
12. Cao JS, Lu ZY, Chen MY, Zhang B, Juengpanich S, Hu JH, *et al.* Artificial Intelligence in Gastroenterology and Hepatology: Status and Challenges. *World J Gastroenterol.* 2021;27(16):1664-1690.
13. Hashimoto R, Requa J, Dao T, Ninh A, Tran E, Mai D, *et al.* Artificial Intelligence Using Convolutional Neural Networks for Real-Time Detection of Early Esophageal Neoplasia in Barrett's Esophagus (With Video). *Gastrointest Endosc* 2020;91: 1264-1271.e1
14. Ebigbo A, Mendel R, Probst A, Manzeneder J, Prinz F, de Souza LA Jr, *et al.* Real-Time Use of Artificial Intelligence in The Evaluation of Cancer in Barrett's Oesophagus. *Gut.* 2020;69:615-616.
15. De Groof AJ, Struyvenberg MR, van der Putten J, van der Sommen F, Fockens KN, Curvers WL, *et al.* Deep-Learning System Detects Neoplasia in Patients With Barrett's Esophagus With Higher Accuracy Than Endoscopists in a Multistep Training and Validation Study With Benchmarking. *Gastroenterology.* 2020;158(4):915-929.e4.
16. Cai SL, Li B, Tan WM, Niu XJ, Yu HH, Yao LQ, *et al.* Using a Deep Learning System in Endoscopy for Screening of Early Esophageal Squamous Cell Carcinoma (With Video). *Gastrointest Endosc* 2019;90: 745-753.e2
17. Tokai Y, Yoshio T, Aoyama K, Horie Y, Yoshimizu S, Horiuchi Y, *et al.* Application of Artificial Intelligence Using Convolutional Neural Networks in Determining the Invasion Depth of Esophageal Squamous Cell Carcinoma. *Esophagus.* 2020;17:250-256.
18. Menon S, Trudgill N. How Commonly is Upper Gastrointestinal Cancer Missed at Endoscopy? *Endosc Int Open.* 2014;2:E46-E50
19. Wu L, Zhang J, Zhou W, An P, Shen L, Liu J, *et al.* Randomised Controlled Trial of WISENSE, a Real-Time Quality Improving System for Monitoring Blind Spots During Esophagogastroduodenoscopy. *Gut.* 2019;68:2161-2169.
20. Ali H, Yasmin M, Sharif M, Rehmani MH. Computer Assisted Gastric Abnormalities Detection Using Hybrid Texture Descriptors for Chromoendoscopy Images. *Comput Methods Programs Biomed.* 2018;157:39-47.
21. Nagao S, Tsuji Y, Sakaguchi Y, Takahashi Y, Minatsuki C, Niimi K, *et al.* Highly Accurate Artificial Intelligence Systems to Predict the Invasion Depth of Gastric Cancer: Efficacy of Conventional White-Light Imaging, Nonmagnifying Narrow-Band Imaging, and Indigo-Carmine Dye Contrast Imaging. *Gastrointest Endosc* 2020;92:866-873.e1.
22. Sakai Y, Takemoto S, Hori K, Nishimura M, Ikematsu H, Yano T, *et al.* Automatic Detection of Early Gastric Cancer in Endoscopic Images Using a Transferring Convolutional Neural Network. *Conf Proc IEEE Eng Med Biol Soc.* 2018;2018:4138-4141.
23. Tsujii Y, Kato M, Inoue T, Yoshii S, Nagai K, Fujinaga T, *et al.* Integrated Diagnostic Strategy for the Invasion Depth of Early Gastric Cancer by Conventional Endoscopy and EUS. *Gastrointest Endosc.* 2015;82(3):452-9
24. Wang CC, Chiu YC, Chen WL, Yang TW, Tsai MC, Tseng MH. A Deep Learning Model for Classification of Endoscopic Gastroesophageal Reflux Disease. *Int J Environ Res Public Health.* 2021;18(5):2428.
25. Shichijo S, Endo Y, Aoyama K, Takeuchi Y, Ozawa T, Takiyama H, *et al.* Application of Convolutional Neural Networks for Evaluating Helicobacter Pylori Infection Status on the Basis of Endoscopic Images. *Scand J Gastroenterol.* 2019;54(2):158-163.
26. Le Berre C, Sandborn WJ, Aridhi S, Devignes MD, Fournier L, Smail-Tabbone M, *et al.* Application of Artificial Intelligence to Gastroenterology and Hepatology. *Gastroenterology.* 2020;158(1):76-94.e2.
27. Seguí S, Drozdal M, Pascual G, Radeva P, Malagelada C, Azpiroz F, *et al.* Generic Feature Learning for Wireless Capsule Endoscopy Analysis. *Comput Biol Med.* 2016;79:163-172.
28. Pannala R, Krishnan K, Melson J, Parsi MA, Schulman AR, Sullivan S, *et al.* Artificial Intelligence in Gastrointestinal Endoscopy. *VideoGIE.* 2020;5(12):598-613.
29. Chen H, Wu X, Tao G, Peng Q. Automatic Content Understanding with Cascaded Spatial-Temporal Deep Framework for Capsule Endoscopy Videos. *Neurocomputing.* 2017;229:77-87.
30. Ding Z, Shi H, Zhang H, Meng L, Fan M, Han C, *et al.* Gastroenterologist-Level Identification of Small-Bowel Diseases and Normal Variants by Capsule Endoscopy Using a Deep-Learning Model. *Gastroenterology.* 2019;157(4):1044-1054.e5.
31. Aoki T, Yamada A, Kato Y, Saito H, Tsuboi A, Nakada A, *et al.* Automatic Detection of Blood Content in Capsule Endoscopy Images Based on a Deep Convolutional Neural Network. *J Gastroenterol Hepatol.* 2020;35(7):1196-1200.
32. Xiao Jia, Meng MQ. A Deep Convolutional Neural Network for Bleeding Detection in Wireless Capsule Endoscopy Images. *Annu Int Conf IEEE Eng Med Biol Soc.* 2016;2016:639-642
33. Usman MA, Satrya GB, Usman MR, Shin SY. Detection of Small Colon Bleeding in Wireless Capsule Endoscopy Videos. *Comput Med Imaging Graph.* 2016;54:16-26.
34. Leenhardt R, Vasseur P, Li C, Saurin JC, Rahmi G, Cholet F, *et al.* A Neural Network Algorithm for Detection of GI Angiectasia During Small-Bowel Capsule Endoscopy. *Gastrointest Endosc.* 2019;89:189-194.
35. Sengupta N, Tapper EB. Derivation and Internal Validation of a Clinical Prediction Tool for 30-Day Mortality in Lower Gastrointestinal Bleeding. *Am J Med* 2017;130: 601.e1-601.e8.
36. Soffer S, Klang E, Shimon O, Nachmias N, Eliakim R, Ben-Horin S, *et al.* Deep Learning For Wireless Capsule Endoscopy: A Systematic Review and Meta-Analysis. *Gastrointest Endosc.* 2020;92:831-839.e8.
37. Qin K, Li J, Fang Y, Xu Y, Wu J, Zhang H, *et al.* Convolution Neural Network for the Diagnosis of Wireless Capsule Endoscopy: A Systematic Review and Meta-Analysis. *Surg Endosc.* 2022;36(1):16-31.
38. Pannala R, Krishnan K, Melson J, Parsi MA, Schulman AR, Sullivan S, *et al.* Artificial Intelligence in Gastrointestinal Endoscopy. *VideoGIE.* 2020;5(12):598-613.
39. Khorasani HM, Usefi H, Peña-Castillo L. Detecting Ulcerative Colitis From Colon Samples Using Efficient Feature Selection and Machine Learning. *Sci Rep.* 2020;10:13744.
40. Wang P, Berzin TM, Glissen Brown JR, Bharadwaj S, Becq A, Xiao X, *et al.* Real-Time Automatic Detection System Increases Colonoscopic Polyp and Adenoma Detection Rates: A Prospective Randomized Controlled Study. *Gut.* 2019;68(10):1813-1819.
41. Zhao S, Wang S, Pan P, Xia T, Chang X, Yang X, *et al.* Magnitude, Risk Factors, and Factors Associated with Adenoma Miss Rate of Tandem Colonoscopy: A Systematic Review and Meta-analysis. *Gastroenterology.* 2019;156 (6):1661-1674.e11.
42. Gohari MR, Biglarian A, Bakhshi E, Pourhoseingholi MA. Use of an Artificial Neural Network to Determine Prognostic Factors in Colorectal Cancer Patients. *Asian Pac J Cancer Prev.* 2011;12:1469-1472.

43. Biglarian A, Bakhshi E, Gohari MR, Khodabakhshi R. Artificial Neural Network for Prediction of Distant Metastasis in Colorectal Cancer. *Asian Pac J Cancer Prev.* 2012;13:927-930.
44. Hassan C, Spadaccini M, Iannone A, Maselli R, Jovani M, Chandrasekar VT. Performance of Artificial Intelligence in Colonoscopy for Adenoma and Polyp Detection: A Systematic Review and Meta-Analysis. *Gastrointest Endosc.* 2021;93(1):77-85.e6
45. Robles-Medranda C, Cifuentes-Gordillo C, Arevalo-Mora M, Egas-Izquierdo M, Puga-Tejada M, Cunto D, et al. Real-Time Computer-Aided Polyp and Adenoma Detection During Screening Colonoscopy in Expert And Non-Expert Endoscopists: A Single Center Study. *Endoscopy.* 2023;55(S 02):S232-S233.
46. Khorasani HM, Usefi H, Peña-Castillo L. Detecting Ulcerative Colitis from Colon Samples Using Efficient Feature Selection and Machine Learning. *Sci Rep.* 2020;10:13744.
47. Kahaleh M, Gaidhane M, Raijman I, Tyberg A, Sethi A, Slivka A, et al. Digital Cholangioscopic Interpretation: When North Meets the South. *Dig Dis Sci.* 2022;67(4):1345-1351.
48. Robles-Medranda C, Valero M, Soria-Alcivar M, Puga-Tejada M, Oleas R, Ospina-Arboleda J, et al. Reliability and Accuracy of a Novel Classification System Using Peroral Cholangioscopy for the Diagnosis of Bile Duct Lesions. *Endoscopy.* 2018;50(11):1059-70.
49. Robles-Medranda C, Oleas R, Sánchez-Carriel M, Olmos JJ, Alcivar-Vásquez J, Puga-Tejada M, et al. Vascularity Can Distinguish Neoplastic From Non-Neoplastic Bile Duct Lesions During Digital Single-Operator Cholangioscopy. *Gastrointest Endosc.* 2021;93(4):935-41.
50. Kahaleh M, Gaidhane M, Shahid HM, Tyberg A, Sarkar A, Ardengh JC, et al. Digital Single-Operator Cholangioscopy Interobserver Study Using A New Classification: The Mendoza Classification (With Video). *Gastrointest Endosc.* 2022;95(2):319-26.
51. Sethi A, Tyberg A, Slivka A, Adler DG, Desai AP, Sejjal D V., et al. Digital Single-operator Cholangioscopy (DSOC) Improves Interobserver Agreement (IOA) and Accuracy for Evaluation of Indeterminate Biliary Strictures: The Monaco Classification. *J Clin Gastroenterol.* 2022;56(2).
52. Saraiva MM, Ribeiro T, Ferreira JPS, Boas FV, Afonso J, Santos AL, et al. Artificial Intelligence for Automatic Diagnosis of Biliary Stricture Malignancy Status in Single-Operator Cholangioscopy: A Pilot Study. *Gastrointest Endosc.* 2022;95(2):339-48.
53. Pereira P, Mascarenhas M, Ribeiro T, Afonso J, Ferreira JPS, Vilas-Boas F, et al. Automatic Detection of Tumor Vessels in Indeterminate Biliary Strictures in Digital Single-Operator Cholangioscopy. *Endosc Int Open.* 2022;10(03):E262-8.
54. Marya NB, Powers PD, Petersen BT, Law R, Storm A, Abusaleh RR, et al. Identification of Patients with Malignant Biliary Strictures Using a Cholangioscopy-Based Deep Learning Artificial Intelligence (With Video). *Gastrointest Endosc.* 2022.
55. Robles-Medranda C, Baquerizo-Burgos J, Alcivar-Vásquez J, Kahaleh M, Raijman I, Kunda R, et al. Artificial Intelligence for Diagnosing Neoplasia on Digital Cholangioscopy: Development and Multicentric Validation of a Convolutional Neural Network Model. *Endoscopy.* 2023.
56. Dahiya DS, Al-Haddad M, Chandan S, Gangwani MK, Aziz M, Mohan BP, et al. Artificial Intelligence in Endoscopic Ultrasound for Pancreatic Cancer: Where Are We Now and What Does the Future Entail? *J Clin Med.* 2022;11(24):7476.
57. Shahidi N, Ou G, Lam E, Enns R, Telford J. When Trainees Reach Competency in Performing Endoscopic Ultrasound: A Systematic Review. *Endosc Int Open.* 2017;05(04):E239-43.
58. Goyal H, Mann R, Gandhi Z, Perisetti A, Zhang Z, Sharma N, et al. Application of Artificial Intelligence in Pancreaticobiliary Diseases. *Ther Adv Gastrointest Endosc.* 2021; 14:1-12.
59. Norton ID, Zheng Y, Wiersema MS, Greenleaf J, Clain JE, DiMaggio EP. Neural Network Analysis of EUS Images to Differentiate Between Pancreatic Malignancy and Pancreatitis. *Gastrointest Endosc.* 2001;54(5):625-9.
60. Săftoiu A, Vilman P. Endoscopic ultrasound elastography Endoscopic Ultrasound Elastography-a New Imaging Technique for the Visualization of Tissue Elasticity Distribution. *J Gastrointest Liver Dis.* 2006;15(2):161-5.
61. Săftoiu A, Vilmann P, Gorunescu F, Janssen J, Hocke M, Larsen M, et al. Efficacy of an Artificial Neural Network-Based Approach to Endoscopic Ultrasound Elastography in Diagnosis of Focal Pancreatic Masses. *Clin Gastroenterol Hepatol.* 2012;10(1).
62. Brosens LAA, Hackeng WM, Offerhaus J, Hruban RH, Wood LD. Pancreatic Adenocarcinoma Pathology: Changing "Landscape." *J Gastrointest Oncol.* 2015;6(4):358-74.
63. Kuwahara T, Hara K, Mizuno N, Okuno N, Matsumoto S, Obata M, et al. Usefulness of Deep Learning Analysis for the Diagnosis of Malignancy in Intraductal Papillary Mucinous Neoplasms of the Pancreas. *Clin Transl Gastroenterol.* 2019;10(5).
64. Hirai K, Kuwahara T, Furukawa K, Kakushima N, Furune S, Yamamoto H, et al. Artificial Intelligence-Based Diagnosis of Upper Gastrointestinal Subepithelial Lesions on Endoscopic Ultrasonography Images. *Gastric Cancer.* 2022;25(2):382-91.
65. Minoda Y, Ihara E, Fujimori N, Nagatomo S, Esaki M, Hata Y, et al. Efficacy of Ultrasound Endoscopy with Artificial Intelligence for the Differential Diagnosis of Non-Gastric Gastrointestinal Stromal Tumors. *Sci Rep.* 2022;12(1).
66. Robles-Medranda C, Oleas R, Del Valle R, Mendez JC, Alcivar-Vásquez J, Puga-Tejada M, et al. ID: 3521955 Application of Artificial Intelligence For Real-Time Anatomical Recognition During Endoscopic Ultrasound Evaluation: A Pilot Study. *Gastrointest Endosc.* 2021;93(6):AB221.
67. Zhang J, Zhu L, Yao L, Ding X, Chen D, Wu H. Deep Learning-Based Pancreas Segmentation and Station Recognition System in EUS: Development and Validation of A useful Training Tool (With Video). *Gastrointest Endosc.* 2020;92(4):874-885.e3
68. Zhang L, Lu Z, Yao L, Dong Z, Zhou W, He C, et al. Effect of a Deep Learning-based Automatic Upper GI Endoscopic Reporting System: A Randomized Crossover Study (With Video). *Gastrointest Endosc.* 2023; 98(2): 181 - 190
69. Qu JY, Li Z, Su JR, Ma MJ, Xu CQ, Zhang AJ, et al. Development and Validation of an Automatic Image-Recognition Endoscopic Report Generation System: A Multicenter Study. *Clin Transl Gastroenterol.* 2022; 12(1): e00282.

The Era of Intelligent Endoscopy: How Artificial Intelligence Empowers Digestive Endoscopy

Jorge Baquerizo-Burgos  · Maria Egas-Izquierdo  · Doménica Cunto  · Carlos Robles-Medrandá 

Endoscopy Division, Instituto Ecuatoriano de Enfermedades Digestivas (IECED), Guayaquil, Ecuador.

Acta Gastroenterol Latinoam 2023;53(3):226-240

Received: 09/08/2023 / Accepted: 14/09/2023 / Published online: 30/09/2023 / <https://doi.org/10.52787/agl.v53i3.339>

Summary

Artificial intelligence is a field of science and engineering that focuses on the computational understanding of intelligent behaviors and the creation of artifacts that exhibit such behaviors, enabling computers to function and think like humans. This technology assists in overcoming the multiple challenges faced by healthcare professionals and contributes to the diagnosis, management, and prognosis of patients. Currently, several artificial intelligence models have been developed for digestive endoscopy, including models that allow the detection of anatomical structures that can assist in the training of physicians, serve as a guide during endoscopic procedures, or assist in stratifying pre-

malignant and malignant lesions. This would reduce false negatives and provide more timely treatments. Computerized systems for lesion detection and diagnosis exist for different segments of the digestive tract, each with specific functions that provide assistance during procedures. All of this has been aimed at reducing risks stemming from human and environmental factors, among others, which can affect the diagnosis and management of diseases. Artificial intelligence models for digestive endoscopy can not only enhance the visual impression of endoscopists but also reduce the learning curve through the application of precise technologies. In this way, the gap between experienced and less experienced endoscopists is reduced. In this article, the technological advancements of artificial intelligence in digestive endoscopy and related future aspects are discussed.

Keywords. Artificial intelligence, computer-assisted detection, computer-assisted diagnosis, deep learning, endoscopy.

La era de la endoscopia inteligente: cómo la inteligencia artificial potencia la endoscopia digestivas

Resumen

La inteligencia artificial es un campo de la ciencia e ingeniería que se ocupa de la comprensión computacional de comportamientos inteligentes y la creación de artefactos que

Correspondence: Jorge Baquerizo-Burgos
Mail: jorgebaquerizoburgos@gmail.com

exhiben tales comportamientos, lo que permite a las computadoras funcionar y pensar de manera similar a la de los seres humanos. Esta tecnología ayuda a superar los múltiples retos que enfrentan los profesionales de la salud aportando favorablemente en el diagnóstico, manejo y pronóstico de los pacientes. Actualmente se han desarrollado varios modelos a nivel de endoscopia digestiva, incluyendo algunos que permiten la detección de estructuras anatómicas que pueden ayudar en el entrenamiento de médicos, como guía durante procedimientos endoscópicos o para la estratificación de lesiones pre-malignas y malignas, disminuyendo falsos negativos y proporcionar tratamientos oportunos. En la actualidad existen sistemas computarizados de detección de lesiones y de diagnóstico en los distintos segmentos de la vía digestiva, cada uno con funciones particulares que proporcionan asistencia durante los procedimientos. Todo esto se ha llevado a cabo con el fin de reducir riesgos derivados por factores humanos, ambientales, entre otros, los cuales pueden afectar el diagnóstico y manejo de enfermedades. Los modelos de inteligencia artificial para endoscopia digestiva pueden, además de mejorar la impresión visual de los endoscopistas, disminuir la curva de aprendizaje a través de la aplicación de tecnologías precisas, y de esta manera reducir la diferencia entre endoscopistas expertos y menos expertos. En este artículo se discuten los avances tecnológicos de la inteligencia artificial en endoscopia digestiva y los aspectos futuros relacionados.

Palabras claves. *Inteligencia artificial, detección asistida por computadora, diagnóstico asistido por computadora, aprendizaje profundo, endoscopia.*

Abbreviations

AI: Artificial Intelligence.

ML: Machine learning.

DL: Deep learning.

CADe: Computer-assisted detection device.

CADx: Computer-assisted diagnostic device.

EUS: Endoscopic ultrasound.

EGD: Esophagogastroduodenoscopy.

AUC: Area under the curve.

BE: Barrett's Esophagus.

SCEC: Squamous cell esophageal carcinoma.

NBI: Narrow Band Imaging.

GERD: Gastroesophageal reflux disease.

SVM: Support Vector Machine.

CNN: Convolutional Neural Network.

CRC: Colorectal cancer.

ADR: Adenoma Detection Rate.

RR: Relative risk.

APC: Adenoma per Colonoscopy.

PDR: Polyp Detection Rate.

ERCP: Endoscopic retrograde cholangiopancreatography.

PPV: Positive Predictive Value.

NPV: Negative Predictive Value.

mAP: Mean Average Precision.

FPS: Frames per Second.

IoU: Intersection over the union.

CT: Computerized tomography.

MRI: Magnetic resonance imaging.

IPMN: Intraductal papillary mucinous neoplasm.

SEL: Subepithelial Lesion.

GIST: Gastrointestinal Stromal Tumor.

NET: Neuroendocrine Tumor.

Introduction

Artificial Intelligence: Basic Concepts

Artificial intelligence (AI) is a branch of computer science whose purpose is the understanding and execution of intelligent insights from a set of computational models.¹ Using a set of algorithms, AI is capable of functioning and reasoning like a human being through a learning process based on training and has the advantage of being able to complete it in less time than a human being.¹ Additionally, this technology can incorporate machine learning (ML) and its subset, deep learning (DL).²

ML is a subgroup of AI characterized by the use of mathematical models for learning from data, which later enables pattern recognition.³ Predictive models are created from algorithms, allowing for data analysis and the resolution of complex problems. Additionally, ML can be categorized into three types: supervised, unsupervised, and reinforcement^{2,4,5}:

a) Supervised learning: this type of learning is based on training using well-categorized or labeled data (external supervision). Labeled data is divided for both training and internal validation. Supervised learning is based on regression, classification, and characterization.^{2,3}

b) Unsupervised learning: this model learns from uncategorized data, enabling the algorithm to operate without

any guidance, relying on the understanding of patterns and thus requiring a greater amount of information.^{2,3}

c) Reinforcement learning: it does not require data or supervision to learn; instead, it is based on learning from the environment through rewards.^{2,3}

DL is a specialized category of ML that is based on the architecture of neural networks resembling those in the human brain.³ It consists of an initial layer that receives input; this layer is followed by a set of hidden middle layers, and then the final output layer (Figure 1). Each layer in this network comprises a group of neurons or nodes that transform (activate) an input into an output through mathematical functions.² The output of a previous layer

serves as the input for the next layer, and so on, until reaching the output layer to obtain a final outcome or detection.^{2,3}

The development of a detection model based on DL involves three main phases. In the first phase, data (images or videos) is collected, and the structures to be used in training the model are properly labeled. Next, in the second phase, the model's architecture is established, and neural networks are created (input layer, middle layer, and output layer). Finally, in the third phase, the samples obtained in the previous phases are used to train the model and subsequently validate it internally (Figure 2).¹ The metrics for evaluating the model's performance are obtained from this final phase (Table 1).

Figure 1. Schematic representation of the architecture of convolutional neural network models

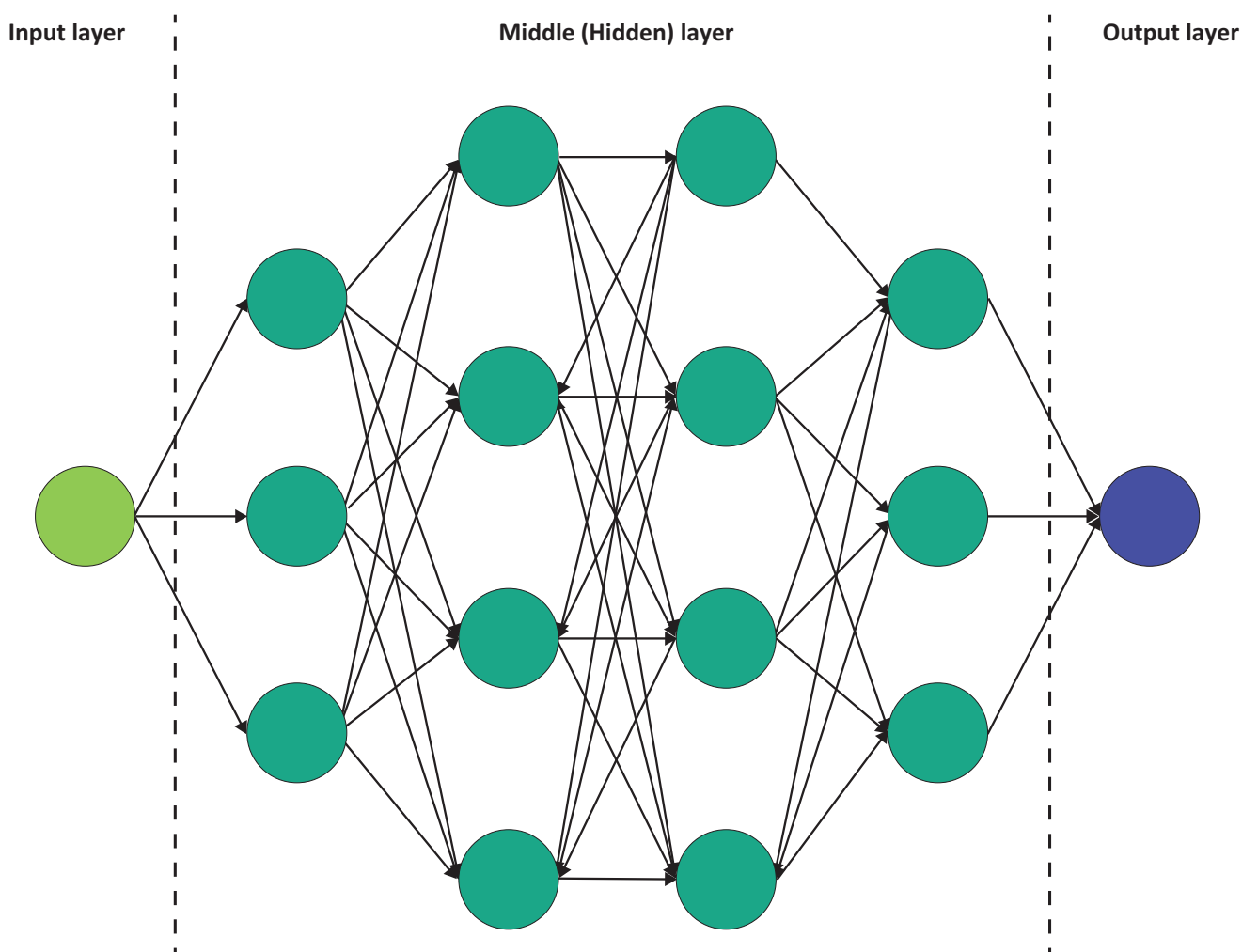
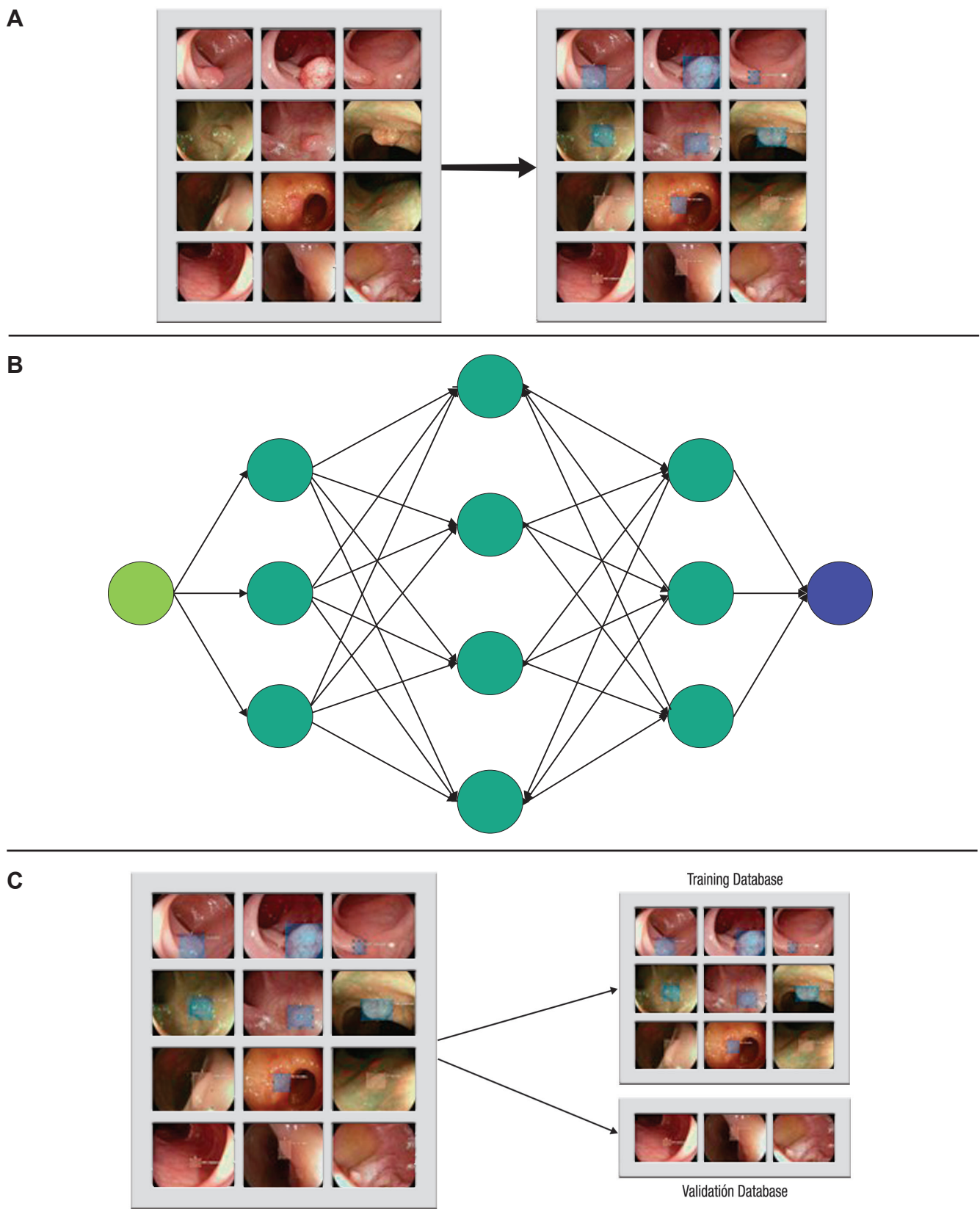


Figure 2. Phases of deep learning model development



A) Data collection and labeling phase. B) Model architecture establishment. C) Model training phase using samples obtained in the first phase.

Table 1. Metrics obtained for the performance evaluation of deep learning models

Metrics	Definition
Sensitivity:	The fraction of positive samples classified as positive by the model.
Specificity:	The fraction of negative samples classified as negative by the model.
Precision:	The fraction of positively classified samples which are truly positive.
F1-Score:	The harmonic means of precision and sensitivity.
IoU:	The performance of object detection by comparing the "ground truth" bounding box to the predicted bounding box.

IoU: Intersection over the union.

Clinical Applications of Artificial Intelligence

Clinical applications of AI have progressively increased in the field of healthcare, including gastroenterology. AI helps overcome the numerous challenges faced by healthcare professionals during data acquisition, analysis, and knowledge application that contributes to patient diagnosis, management, and prognosis.¹ Additionally, automation in image identification and recognition assists in reducing errors stemming from human factors (fatigue and workload, among others).

Technological advancements have led to the development of intelligent systems that facilitate the detection or the stratification of lesions observed during endoscopic or imaging procedures. These are referred to as computer-assisted detection device (CADE) or computer-assisted diagnostic device (CADx).⁶ Thus, the application of these devices in digestive endoscopy can facilitate and increase lesion detection during procedures and categorize lesions as benign or malignant in real-time.

Artificial intelligence in Digestive Endoscopy

Currently, several models have been developed for upper and lower gastrointestinal endoscopy, as well as for advanced endoscopic procedures such as cholangioscopy and endoscopic ultrasound (EUS) (Table 2). These models encompass both CADE and CADx systems, each employing distinct algorithms that enable different functionalities. These functionalities include identifying anatomical structures and specific lesions, assisting in physician training or serving as a guide during

endoscopic procedures, reducing the number of false negatives through characterization and stratification of premalignant and malignant lesions, among others.⁴

In the following sections, we will review updated information on the uses of AI and its impact according to the type of endoscopic assessment.

Table 2. Applications of artificial intelligence in different segments of the digestive system

Segment	AI Applications
Esophagus	<ul style="list-style-type: none"> • BE early detection and diagnosis • Lesion classification (benign or malignant) • Esophageal carcinoma detection • Tumoral lesions invasion
Stomach	<ul style="list-style-type: none"> • Blindspot surveillance • Detection of lesions suggestive of neoplasia • Tumoral lesions invasion • Differentiation between normal and pathological tissues • GERD detection • <i>H. pylori</i>-associated gastritis detection
Small Intestine	<ul style="list-style-type: none"> • Lesion detection and classification (polyps, bleeding, ulcers, among others) • Evaluation of celiac and Crohn's disease
Colon	<ul style="list-style-type: none"> • Polyp detection and characterization • Assistance during screening colonoscopy
Biliary tract	<ul style="list-style-type: none"> • Detection of areas suggestive of neoplasia • Identification of normal structures • Guidance during biopsy sampling

AI: Artificial intelligence; BE: Barrett's esophagus; GERD: Gastroesophageal reflux disease.

Upper Digestive Endoscopy

Also known as esophagogastroduodenoscopy (EGD), this procedure is of great importance in the diagnosis of upper gastrointestinal tract lesions.^{7,8} However, the diagnostic rate varies according to the performance of the endoscopist.⁹ Errors during EGD are one of the leading causes of incorrect diagnosis of premalignant lesions and severe esophagogastricoduodenal diseases. AI systems have been developed to overcome the technical challenges described above. Their use in the upper digestive tract ranges from anatomical localization to the detection and evaluation of malignant and premalignant lesions.⁷⁻¹⁰

Takiyama *et al.* developed an AI model capable of classifying anatomical structures in the upper digestive

tract, which has shown excellent performance in identifying the larynx (AUC 1.00), the esophagus (AUC 1.00), the stomach (upper, middle, and lower portions), and the duodenum (AUC 0.99).⁷

On the other hand, in the multicenter study conducted by Luo *et al.*, the diagnostic accuracy of the GRADIS model for detecting upper digestive tract neoplasms (esophagus and stomach) was evaluated. This model achieved a diagnostic accuracy of 95.5% (95% CI: 95.2 - 95.7) during internal validation. When comparing its performance with endoscopists of different expertise levels, it showed similar sensitivity to experts (94.2% vs. 94.5%), and higher sensitivity compared to competent endoscopists (94.2% vs. 85.8%) and trainees (94.2% vs. 72.2%).¹¹ Additionally, the diagnostic accuracy of experts (92.8%) when using the AI model was similar to that of the group of competent endoscopists (93.4%) and trainees (90.4%).¹¹ This demonstrates that the application of AI can narrow the gap between experts and non-experts.¹¹

Esophagus

Accuracy in the early diagnosis of Barrett's esophagus (BE) and esophageal neoplasia remains a challenge, even for many experienced endoscopists. Once BE is identified, the identification of regions with dysplasia or early adenocarcinoma becomes necessary.

AI models have been designed to assist endoscopists in improving the accuracy of diagnosing these lesions,¹² including systems for neoplasia classification using real-time magnification with high precision (89.9%), which have enabled early diagnosis and differentiation of neoplasia in BE.^{13,14}

The CADx system developed and validated by de Groof *et al.* allowed for the classification of neoplastic and non-neoplastic images in BE compared to the performance of 53 endoscopists. The model outperformed the endoscopists' performance and achieved higher accuracy (88.0% vs. 73.0%), sensitivity (93.0% vs. 72.0%), and specificity (83.0% vs. 74.0%).¹⁵

On the other hand, due to the significant importance of recognizing and treating esophageal carcinoma promptly, researchers have developed systems that enable lesion detection as well as the assessment of disease invasion.^{12,16,17} Esophageal carcinoma is often detected at advanced stages, and small lesions are usually detected by highly experienced endoscopists.¹² AI allows for the detection of lesions smaller than 10 mm with high accuracy (91.4%), even surpassing that of many so-called expert endoscopists (> 15 years of experience,

88.8%), those with moderate experience (5 - 15 years, 81.6%), and those with limited experience (< 5 years, 77.2%).¹⁶

Determining the depth of the lesion enables the selection of the appropriate treatment (surgical, endoscopic, pharmacological), and prognosis assessment.¹⁷ One of the models with high diagnostic accuracy in predicting the depth of invasion of squamous cell esophageal carcinoma (SCEC) is proposed by Tokai *et al.* Researchers used 1751 images for training and 291 images for validation, achieving a sensitivity of 84.1% and a diagnostic accuracy of 80.9% in estimating the depth of SCEC invasion. When compared to thirteen endoscopists, this model showed higher diagnostic accuracy and a greater AUC.¹⁷

Stomach

Stomach cancer usually shows nonspecific symptoms during its early stages, and that's why patients are often diagnosed at advanced stages. The prognosis of stomach cancer depends on the assessment of the depth of the lesion and its early detection. It has been reported that the early detection of stomach cancer can increase 5-year survival rates to 90.0%.¹⁸

According to Menon *et al.*, the rate of false negatives in the diagnosis of early stomach cancer can reach up to 25.0%.¹⁸ Automation systems aim to reduce this percentage with models that classify stomach images in EGD to monitor blind spots with high precision,¹⁹ models that detect lesions suggestive of stomach cancer¹¹ and precancerous lesions,²⁰ and that further assess the depth of invasion.²¹

Chromoendoscopy is one of the diagnostic methods used for the early detection of gastric neoplasm. However, during an endoscopic session, multiple video frames can be generated, making the review process an exhaustive task for endoscopists.

To prevent losses during evaluation, Ali *et al.* developed a CADx trained to classify frames as normal or pathological based on local and global texture. This model showed a sensitivity, specificity, precision, and AUC of 91.0%, 82.0%, 87.0%, and 0.91, respectively.²² The model proved to be a diagnostic aid in the early detection of gastric cancer since it reduces the time used for the evaluation of endoscopic sequence.

The model studied by Wu *et al.*¹⁹ showed high accuracy, specificity, and sensitivity (92.5%, 94.0%, 91.0%, respectively) in assessing non-malignancy, and surpassing expert endoscopists in this task.¹⁹ Additionally, during real-time procedures, it showed excellent performance in

early gastric adenocarcinoma detection with blind-spot monitoring. Additionally, other AI models have shown high accuracy in diagnosing gastrointestinal neoplasia comparable to that of expert endoscopists.^{11,17}

Furthermore, models have been developed that not only identify neoplastic lesions but also predict their depth. Nagao *et al.* trained a model to predict the depth of invasion of gastric cancer using conventional white-light images, narrow band imaging (NBI) images, and images with indigo carmine contrast.²¹ The model showed high accuracy across all three systems (white-light images 94.5%, NBI 94.3%, and indigo carmine 95.5%).²¹

Zhu *et al.* (total accuracy 89.2%) and Yoon *et al.* (sensitivity 81.7%, specificity 75.4%) have reported diagnostic accuracy of their models for assessing invasion depth, which is comparable to that of other conventional methods. The advantages of using these models lie in the more objective assessment of macroscopic lesion characteristics, reducing the need for other invasive techniques like EUS.^{21,23,24}

In addition to neoplasia identification, other applications include the detection of gastroesophageal reflux disease (GERD)²⁵ and *H. pylori*-associated gastritis.²⁶ Models developed to assist in classifying GERD with NBI have achieved a total diagnostic accuracy of 99.2% for grade A-B lesions, 100% for grade C-D lesions, and 100% in the control group. Therefore, these models are considered highly useful for assisting in the automatic detection of lesions consistent with GERD, and increase diagnostic accuracy for trainees.²⁵

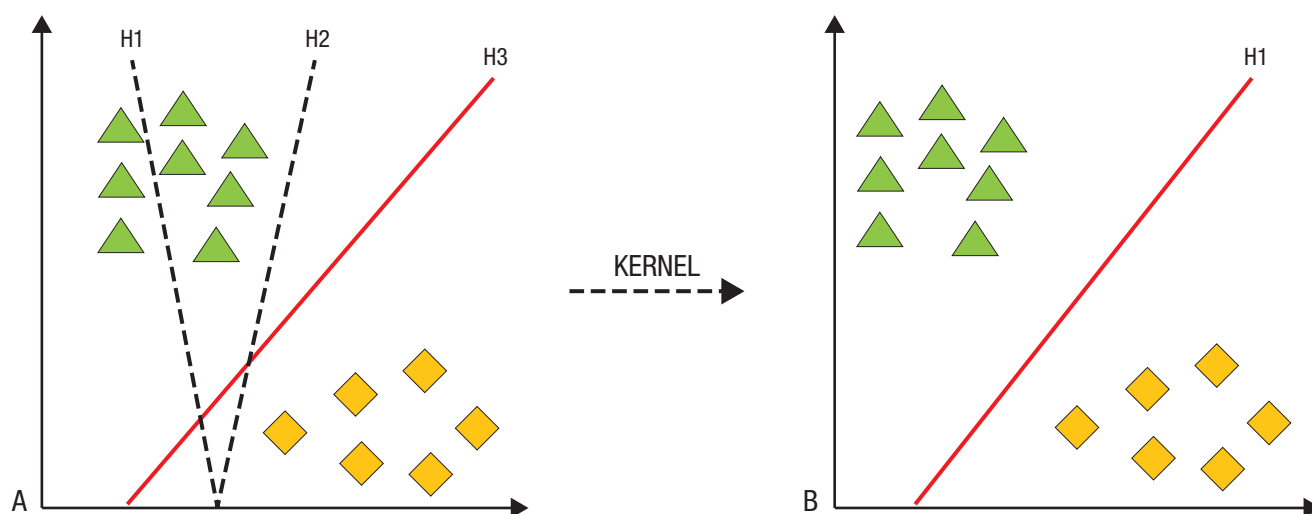
On the other hand, the neural network designed for predicting *H. pylori* in endoscopic images correctly diagnosed 80% of negative cases, 84% of eradicated cases, and 48% of positive cases. The authors of this study emphasize the utility of this model in identifying patients who may require additional confirmation testing for *H. pylori* based on endoscopic results, and recommend its use as a diagnostic aid.²⁶

Small Intestine

Capsule endoscopy, as a non-invasive procedure allows for the detection and classification of lesions (bleeding, ulcers, and polyps), assessment of intestinal motility, and evaluation of conditions like celiac disease and other pathologies that primarily affect the small intestine. However, the evaluation of the large number of obtained images (> 60 000) and the difficulty in directing the capsule (completely dependent on gastrointestinal peristalsis) make the procedure slow (from 45 minutes to 8 hours) and tedious. To address the technical difficulties associated with this procedure, automation using AI has been explored.²⁸

AI models for capsule endoscopy have been developed based on DL. Image classification and categorization are performed using support vector machines (SVMs). These models separate data using hyperplanes in two or more dimensions. After employing Kernel parameters, an “optimal” hyperplane is determined, which creates “boundaries” for data categorization (Figure 3).^{27,28} Following categorization, DL algorithms are used to create artificial neural networks.^{27,29}

Figure 3. Support vector machines (SVM) for data classification



A) Linear classification with hyperplanes (H1, H2, H3) by SVM. B) Data classified with an optimal hyperplane (H1) after Kernel parameters.

Currently, AI models for capsule endoscopy include capsule tracking, detection of polyps, bleeding, ulcers, and the study of small intestine-specific pathologies like celiac disease and Crohn's disease.²⁶

In the gastrointestinal tract, models incorporated in capsule endoscopy usually allow for its tracking and localization in different segments (mouth, stomach, small intestine, and colon), after excluding frames with "noise" (feces, bubbles, etc.), with a sensitivity and specificity of 88.0%.³⁰ Evaluating images based on their topographic location saves study time and enhances diagnostic accuracy.³⁰

In the multicenter study published by Ding *et al.*,³⁰ a convolutional neural network (CNN) was developed for the identification of normal images, inflammation, ulcers, polyps, lymphangiectasia, bleeding, vascular disease, diverticula, and parasites, among others. When comparing the model's results with those of participating gastroenterologists, a sensitivity of 98.9% (95% CI: 99.7 - 99.9) was obtained vs. 74.6% (95% CI: 73.1 - 76.0) in the identification of abnormalities per patient, and a sensitivity of 99.9% (IC 95%: 99.6 - 99.9) vs. 76.9% (IC 95%: 75.6 - 78.2) in lesion analysis. Moreover, the reading time per patient was significantly shorter in the CNN group compared to gastroenterologists (5.9 ± 2.23 minutes vs. 96.6 ± 22.53 minutes, $p < 0.001$). Based on these results, the researchers concluded that the application of AI in capsule endoscopy is an important tool to assist gastroenterologists in analyzing images captured with this device more efficiently and accurately.³¹

Models for real-time bleeding detection achieve diagnostic accuracy of up to 99.0%.³²⁻³⁴ The model developed by Aoki *et al.* detected gastrointestinal bleeding with a sensitivity, specificity, and accuracy of 96.6%, 99.9%, and 99.9%, respectively.³¹ Other models developed aid in the stratification and prediction of the risk of recurrent bleeding in order to provide timely treatment and avoid unnecessary endoscopies.^{35,36}

For the detection of ulcers and erosions, Wang *et al.*'s model achieved a diagnostic accuracy of 92.1%.³⁷ For the identification of tumors, the diagnostic accuracy starts at 86.0%, with a sensitivity ranging from 88.0% to 97.0%, and specificity between 81.0% and 96.0%.³⁸

In a recent meta-analysis of CNN in capsule endoscopy, pooled sensitivity and specificity were obtained as follows: 96.0% (95% CI: 91.0 - 98.0) and 97.0% (95% CI: 93.0 - 99.0) in the detection of ulcers and erosion; 97.0% (95% CI: 93.0 - 99.0) and 100% (95% CI: 99.0

- 100) in the identification of gastrointestinal bleeding; and 97.0% (95% CI: 82.0 - 99.0) and 98.0% (95% CI: 92.0 - 99.0) in the detection of cancer and polyps.³⁷

Furthermore, models developed for the identification of inflammatory bowel disease using capsule endoscopy currently achieve high levels of accuracy (83.3% to 90.8%),³⁹ and allow the recognition of hidden disease patterns as well.⁴

Compared to current endoscopes which provide high-quality images, the image quality of capsule endoscopy is poor.²⁷ However, AI models for capsule endoscopy have the advantage of having a robust databases fed with a large number of images useful for the creation of CNN.

Lower Digestive Endoscopy

Colon

Colorectal cancer (CRC) is currently considered one of the leading causes of cancer-related death in both men and women.⁴⁰ For the early identification of premalignant lesions (polyps and adenomas), colonoscopy remains an essential procedure. However, according to the literature, approximately 25% of these lesions can be undetected, even by expert hands.⁴¹ The undetected premalignant lesions increase the risk of developing CRC.

Automated systems have been developed for the detection and characterization of polyps. The first systems developed used a limited number of images and consequently had a poor diagnostic accuracy (72.0%).⁴² Subsequently, new models have been trained with higher accuracy (> 95.0%) using a larger number of images. This makes it possible to better evaluate polyps and tiny adenomas (≤ 5 mm), and to predict CRC prognosis, patient survival, and distant invasion.^{42,43}

A meta-analysis evaluated the performance of CADe systems in the detection of colorectal neoplasia.⁴⁴ The authors found a higher adenoma detection rate (ADR) in the groups that used CADe compared to their control groups (36.6% vs. 25.2%, relative risks (RR): 1.44; 95% CI: 1.27 - 1.62; $p < 0.001$). Additionally, the detection of adenomas during colonoscopy (APC) was superior in the CADe groups compared to the control group (50.3% vs. 34.6%; (RR): 1.70; 95% CI: 1.53 - 1.89; $p < 0.001$). The authors did not find a significant difference in colonoscopy efficiency (withdrawal time) between the group that used CADe and the control group.⁴⁴

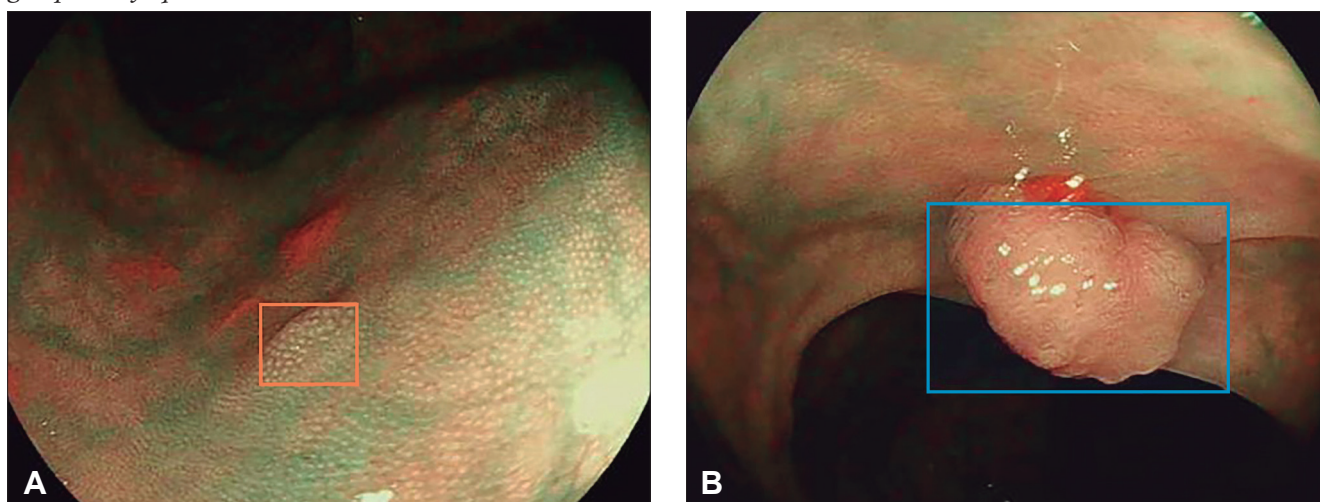
Robles-Medrandá *et al.* studied the efficiency

of AI-assisted colonoscopy for the detection of polyps and adenomas during CCR screening studies (Figure 4).⁴⁵ The results obtained by the authors were compared according to the level of experience of the endoscopists (experts vs. non-experts). With AI assistance during endoscopic procedures, an increase in the ADR and polyp detection rate (PDR) was achieved from 16.5% to 18.2% and from 50.4%

to 60.0%, respectively. Depending on the experience level, the increase in ADR was evident in the junior group (10.8% to 16.2%), which approached the level of the experts.⁴⁵

Furthermore, CADx systems have been designed with the ability to immediately characterize polyps using imaging technologies beyond white-light endoscopy and magnified NBI, such as confocal endomicroscopy.

Figure 4. Detection of polypoid lesion in the colon using a computer-aided detection device (CADE) (AITROL, mdconsgroup, Guayaquil, Ecuador)



A) Polypoid lesion type I according to NICE classification (orange delineation). B) Polypoid lesion type II according to NICE classification (blue delineation).

In this field, a model developed by Sánchez-Montes *et al.* for predicting the histological classification of polyps achieved a diagnostic accuracy, sensitivity, and specificity of 91.1%, 92.3%, and 89.2%, respectively.⁴⁶

Additionally, features such as lesion depression, fold convergence, and irregular and heterogeneous capillary pattern are associated with deep invasion of premalignant lesions. Current CADe models for the identification and detection of the above characteristics sound “attractive” to determine the type of treatment to be performed (e.g., endoscopic mucosal resection).⁴

Advanced Endoscopic Techniques

Despite the availability of diagnostic methods such as cholangioscopy, endoscopic retrograde cholangiopancreatography (ERCP), or EUS for studying the biliopancreatic system, there are difficulties in differentiating lesions and discrepancies among evaluators.

Cholangioscopy

As cholangioscopy is a relatively new advanced endoscopic technique without established training guide-

lines, the visual impressions among operators is highly variable.⁴⁷ As a result, several classifications have been created to detect malignancy based on the macroscopic characteristics of biliary lesions during the procedure,⁴⁸⁻⁵¹ with the intention of reducing this variability among observers. However, these classifications have not achieved that goal.⁴⁷

Recently, AI models have been developed to assist operators in detecting malignant lesions and obtaining biopsies. The first AI models in cholangioscopy were developed for the detection of tortuous blood vessels, but had the disadvantage of being limited to still images and were not usable in live cases.^{52,53} Another limitation of these models is their lack of clinical validation. However, these models achieved quite high internal validation metrics. In their first study, Mascarenhas *et al.* developed a model using 6475 cholangioscopy images obtained from 85 patients. During validation using frames, they obtained a sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of 99.3%, 99.4%, 99.6%, and 98.7%, respectively.⁵³ Later, they conducted a new study, in which

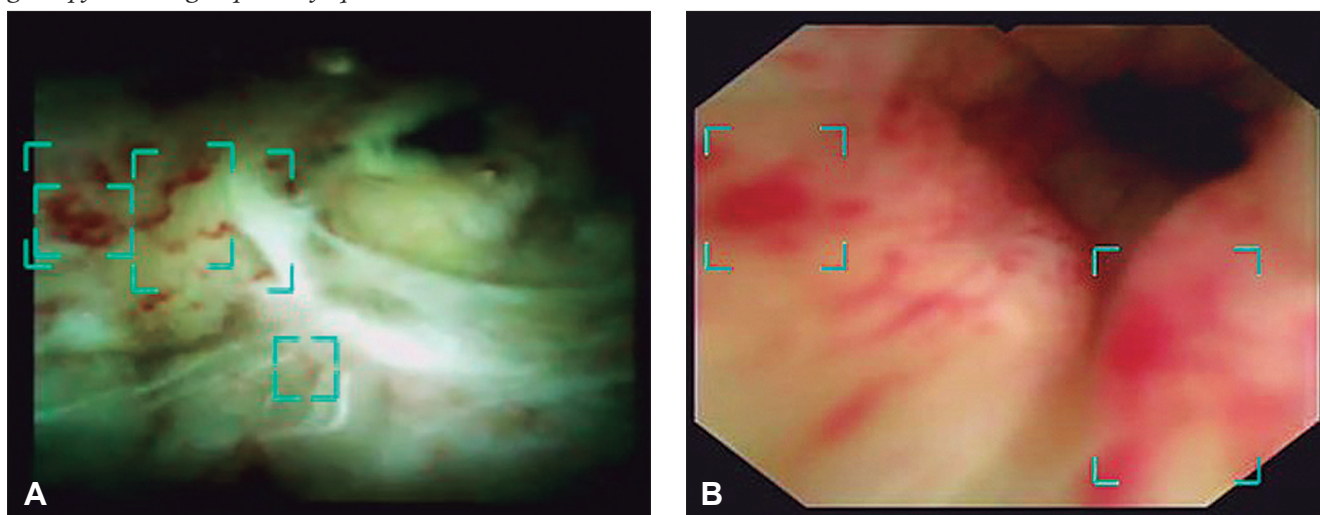
they doubled the number of images obtained from the same number of patients (from 6475 to 11855 images); this time, they assessed the model's accuracy in distinguishing between benign and malignant lesions. The metrics obtained were diagnostic accuracy (94.8%), sensitivity (94.7%), specificity (92.1%), PPV (94.8%), and NPV (84.2%).⁵² It is important to note that, even though the model achieved excellent internal validation parameters, these should not be extrapolated to its real-world utility in live clinical cases.

Subsequently, two studies with real-time AI models were conducted. The first one, conducted by Marya *et al.*, evaluated the clinical application of the CNN model and compared it to biopsy and cytology results. The study found that the CNN model achieved higher

sensitivity (93.3%), specificity (88.2%), and accuracy (90.6%) compared to biopsies (sensitivity 35.7%, specificity 100%, and accuracy 60.9%, respectively) and cytology (sensitivity 40.0%, specificity 100%, and accuracy 62.5%, respectively).⁵⁴ One limitation of the study was that only one operator was responsible for annotating around 2 million images, which could lead to errors due to fatigue or bias.

Robles-Medrandra *et al.* developed their own CNN model capable of detecting neoplastic lesions in pre-recorded and real-time videos (Figure 5). Following the AI model implementation stages (data collection, annotation, and model design), internal validation was performed, and then a clinical comparison between the model and endoscopists (experts and non-experts) was done.⁵⁵

Figure 5. Detection of images suggestive of neoplasia in cholangioscopy through artificial intelligence (AIWorks-Cholangioscopy, mdconsgroup, Guayaquil, Ecuador)



A) Potentially neoplastic area delimited by the AI system using a Spyglass cholangioscope (Boston Scientific, Marlborough, MA, USA). B) Potentially neoplastic area delimited by the AI system using a 7Fr eyeMAX cholangioscope (Micro-Tech, Nanjing, China)

This multicenter study was conducted in two phases. The first phase involved the development and validation of the AI model, named AIWorks-Cholangioscopy (mdconsgroup, Guayaquil, Ecuador). The initial version of this model was developed using 81080 images from 23 patients, and achieved a mean average precision (mAP) of 0.298, an F1-score of 0.280, an intersection over the union (IoU) of 32.3%, and a total loss of 0.1034. Despite these acceptable results, the frames per second (FPS) detection rate was low (average around 5). This low FPS detection prevented the model from being used for real-time lesion detection. Its internal validation obtained a sensitivity, specificity, PPV, and NPV of 98.0%, 95.0%, 98.0%, and 94.0%, respectively. Subse-

quently, by increasing the number of cases and images available for training (from 81080 to 198941 images), along with improved image quality, the internal validation metrics increased dramatically: mAP from 0.298 to 0.880, F1-score from 0.280 to 0.738, IoU from 32.3% to 83.2%, and total loss decreased from 0.1034 to 0.0975. Sensitivity, specificity, PPV, and NPV for detecting neoplastic lesions in images had similar results to those obtained by Mascarenhas *et al.* (98.6%, 98.0%, 89.2%, and 99.2%, respectively).^{52,53,55} However, during the second phase for clinical validation in 170 patients, it was observed that these diagnostic accuracy values decrease and approached those of the endoscopists. When comparing the AI model to the visual impression

of endoscopists (experts and non-experts) using two classifications of neoplastic lesions (CRM and Mendoza classification),^{48,50} it was evident that the AI model was superior to both experts and non-experts.⁵⁵ This study highlights the importance of conducting clinical validation and not extrapolating internal validation results to the clinical setting as the final outcome.

Endoscopic Ultrasound

For the diagnostic and therapeutic management of biliopancreatic pathologies, EUS is considered superior to CT scans and MRI due to its higher diagnostic accuracy and ability to obtain higher-quality images.⁵⁶ However, these procedures have limitations such as low sensitivity to differentiate between benign and malignant intraductal papillary mucinous neoplasia (IPMN) and low specificity to differentiate between malignant lesions and chronic pancreatitis.⁵⁷ Another limitation of this procedure is its operator-dependence, so less experienced endoscopists may not appreciate the differences between chronic pancreatitis and pancreatic malignancy.⁵⁷ For this reason, the application of AI in such procedures would be beneficial and could influence the quality of EUS performed by expert endoscopists or trainees.⁵⁸

Several studies have been developed to evaluate and compare the diagnostic accuracy of AI-assisted EUS versus traditional EUS for detecting pancreatic cancer, and to distinguish between chronic lesions and normal tissue.⁵⁵ A study by Norton *et al.* demonstrated that their AI model had higher sensitivity (100%) for the differentiation between malignancy and inflammation; however, the diagnostic accuracy was similar between the model (80%), the endoscopist blinded to procedure results (83.0%) and the traditional procedure (85.0%).⁵⁹ Their study shows the potential of applying AI models for image interpretation in EUS and the ability to differentiate between malignancy and chronic conditions; this way they solved one of the limitations mentioned above.

Over time, new EUS techniques have been developed and included as part of patient management. At the same time, different types of AI were designed to differentiate between benign and malignant pancreatic lesions. Săftoiu *et al.* evaluated the application of an AI model that distinguishes between benign and malignant lesions using elastography, a technique that assesses tissue stiffness and elasticity. Based on these parameters, high sensitivity (91.4%), specificity (87.9%), and diagnostic accuracy (89.7%) were achieved.⁶⁰ These results

indicate the potential for applying AI models in elastography in cases where EUS-guided fine-needle aspiration yields negative results.

Subsequently, a comparison was made between the use of AI-assisted elastography and traditional one. They observed that the diagnostic accuracy of the AI-assisted procedure (EUS + AI) was higher (AUC: 94.0%) than traditional one (UC: 85.0%). This suggests that CNN-based elastography models can provide decision support by offering a quick and accurate image interpretation compared to elastography without AI.⁶¹

Another limitation of EUS that has been evaluated is the differentiation between benign and malignant IPMNs. IPMNs are precursors to pancreatic adenocarcinomas, and once these lesions progress to invasive stages, patient prognosis worsens.⁶²

Kuwahara *et al.*, developed a DL model and investigated whether preoperative analysis of IPMN images by EUS using AI can predict malignancy. They compared lesion interpretation with preoperative diagnoses by endoscopists, conventional predictive techniques, and other EUS techniques.⁶³ The AI model achieved an AUC of 91.0% for predicting malignancy. In the comparison of diagnostic accuracy between the model and the endoscopist, the model was superior (94.0% vs. 56.0%, respectively).⁶³

Another application of AI in EUS is for the differential diagnosis of subepithelial lesions (SELs). Hirai *et al.* evaluated a DL model that allowed classification in EUS images. They collected images of SELs in the upper digestive tract, including gastrointestinal stromal tumors (GISTs), leiomyomas, schwannomas, neuroendocrine tumors (NETs), and ectopic pancreas. The model was able to classify the aforementioned lesions with an accuracy of 86.1%, which was much higher than that of the participating endoscopists.⁶⁴ The sensitivity, specificity, and diagnostic accuracy for distinguishing GISTs from other lesions were 98.8%, 67.6%, and 89.3%, respectively.⁶⁴ Other studies, including meta-analyses comparing accuracy of another CNN model to differentiate GISTs from other lesions had similar results.⁶⁵

Due to the increasing trend of biliopancreatic neoplasms worldwide, it's important to accurately differentiate malignant tumor lesions from benign or normal tissues. The application of AI in biliopancreatic endoscopy has been under evaluation for a long time, demonstrating promising results. AI and its application in medicine is considered beneficial as it will help overcome the limitations of these complex procedures (cholangioscopy and EUS).

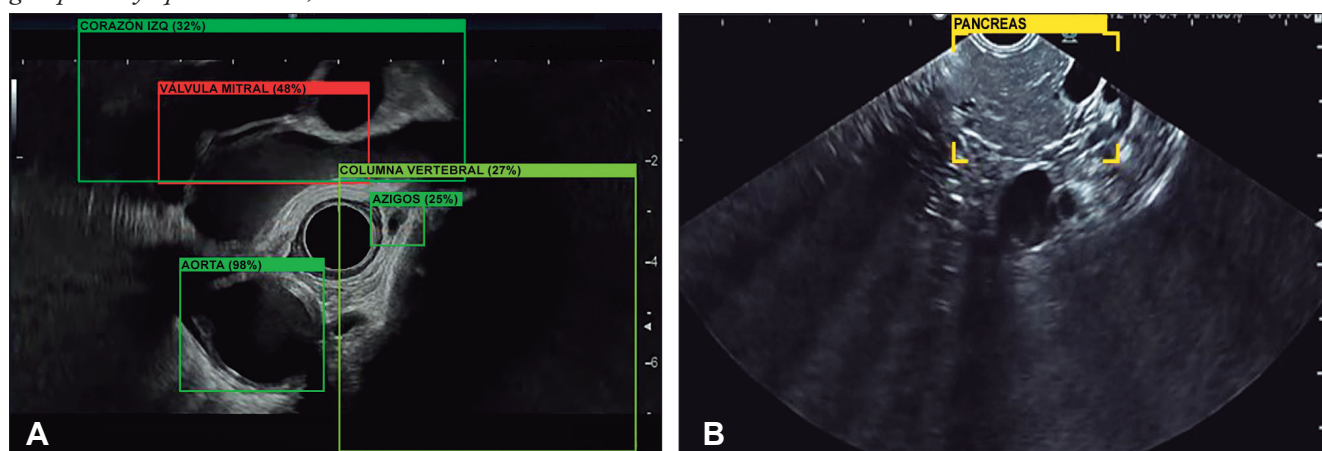
Future Applications of Artificial Intelligence in Digestive Endoscopy

AI has proven to be useful for detecting and classifying lesions during various types of available and studied endoscopic procedures. However, the application of intelligent models can go beyond detection and diagnosis.

Carlos Robles-Medranda *et al.* developed an EUS system based on CNN models trained to detect normal anatomical structures in different windows evaluated during this advanced procedure (mediastinal, gastric, and duodenal) (Figure 6). This allowed the

identification of 20 anatomical structures with high sensitivity and specificity (Table 3).⁶⁶ The diagnostic accuracy of the model for detecting these structures was higher than 95.0%. This indicates that AI models are not only useful for detecting pathologies but also for detecting normal structures, which can benefit the training of endoscopists and reinforce the knowledge of those more experienced. The application of AI in endoscopist training was evaluated by Zhang *et al.*, who demonstrated that those trained with AI had a shorter learning curve and better results than those trained traditionally.⁶⁷

Figure 6. Artificial Intelligence model for endoscopic ultrasound detecting anatomical structures (AIWorks-EUS, mdconsgroup, Guayaquil, Ecuador)



A) Computer-assisted detection (CADE) device detecting the left heart, mitral valve, spine, azygos vein, and aorta using a radial-array endoscopic ultrasound probe; B) CADE device detecting the pancreas using a linear-array endoscopic ultrasound probe

Table 3. Anatomical structures detected by the computer-assisted detection system AIWorks-EUS (mdconsgroup, Guayaquil, Ecuador)

Windows	Anatomic structure
Mediastinal	Aorta, vertebral spine, aortic arch, trachea, aortopulmonary space
Gastric	Left kidney, liver, spleen, body and tail of the pancreas, celiac trunk, splenic vein and artery, inferior vena cava, adrenal gland
Duodenal	Right kidney, gallbladder, common bile duct, ampulla of Vater, portal vein

Additionally, through intelligent models, it is possible to automate the process of obtaining images for report generation and as a measure of test quality.⁶⁸ This can be achieved by combining various functions and applying them simultaneously, such as detecting the evaluated organ along with detecting lesions within the same organ. In this way, at the end of the procedure, the study report can be automatically generated.^{68,69} Models such as AI-EARS and ISRGS have shown good diagnostic accuracy in identifying organs and lesions, automatically generating a report using AI in both upper⁶⁸ and lower⁶⁹ digestive tracts.

The application of AI models to train endoscopists in advanced endoscopic procedures can increase their effectiveness and reduce the number of procedures needed to achieve competence.

Conclusion

Despite the goal of automating models during en-

doscopy to reduce risks associated with human and environmental factors, among others, the success of AI models will depend on the quality and quantity of information used during their training and validation. Additionally, external validation through multicenter and international studies with expert endoscopists is of great importance before generalization and universalization of training results.

AI models in digestive endoscopy have the potential to improve the visual perception of endoscopists and reduce the accuracy gap between those with less experience and those considered experts. Moreover, they can be of great assistance in lesion detection and tissue invasion assessment. A future is envisioned where traditional training approaches are surpassed, and medical centers worldwide will be able to implement improvements in professional training by applying precise technologies that reduce the learning curve in these procedures. This will benefit less experienced professionals and reduce the gap between expert endoscopists and those with less experience, ultimately improving their competency.

Intellectual Property. *The authors declare that the data, figures, and tables presented in this manuscript are original and were made in their respective institutions.*

Funding. *The authors declare that there were no external sources of funding.*

Conflict of interest. *Carlos Robles-Medranda is a key opinion leader for Pentax Medical, Steris, Micro-tech, G-Tech Medical Supply, CREO Medical, EndoSound, and mdcons-group. The other authors declare no conflict of interest.*

Copyright



© 2023 Acta Gastroenterológica latinoamericana. This is an open-access article released under the terms of the Creative Commons Attribution (CC BY-NC-SA 4.0) license, which allows non-commercial use, distribution, and reproduction, provided the original author and source are acknowledged.

Cite this article as: Baquerizo-Burgos J, Egas-Izquierdo M, Cunto D *et al.* The Era of Intelligent Endoscopy: How Artificial Intelligence Empowers Digestive Endoscopy. *Acta Gastroenterol Latinoam.* 2023;53(3):226-240. <https://doi.org/10.52787/agl.v53i3.339>

References

- Ramesh AN, Kambhampati C, Monson JRT, Drew PJ. Artificial Intelligence in Medicine. *Ann R Coll Surg Engl.* 2004; 86(5): 334-8.
- Tonozuka R, Mukai S, Itoi T. The Role of Artificial Intelligence in Endoscopic Ultrasound for Pancreatic Disorders. *Diagnostics.* 2021;11(1).
- Minoda Y, Ihara E, Fujimori N, Nagatomo S, Esaki M, Hata Y, *et al.* Efficacy of Ultrasound Endoscopy with Artificial Intelligence for the Differential Diagnosis of Non-Gastric Gastrointestinal Stromal Tumors. *Sci Rep.* 2022;12(1).
- Kröner PT, Engels MM, Glicksberg BS, Johnson KW, Mzaik O, van Hooft JE, *et al.* Artificial Intelligence in Gastroenterology: A State-Of-The-Art Review. *World J Gastroenterol.* 2021;27(40): 6794-6824.
- Yu KH, Beam AL, Kohane IS. Artificial Intelligence in Healthcare. *Nat Biomed Eng.* 2018;2(10):719-31.
- Firmino M, Angelo G, Morais H, Dantas MR, Valentim R. Computer-Aided Detection (CADe) and Diagnosis (CADx) System for Lung Cancer with Likelihood of Malignancy. *Biomed Eng Online.* 2016;15(1):2.
- Takiyama H, Ozawa T, Ishihara S, Fujishiro M, Shichijo S, Nomura S, *et al.* Automatic anatomical classification of esophagogastroduodenoscopy images using deep convolutional neural networks. *Sci Rep.* 2018;8:7497.
- van der Sommen F, Zinger S, Curvers WL, Bisschops R, Pech O, Weusten BL, *et al.* Computer-Aided Detection of Early Neoplastic Lesions in Barrett's Esophagus. *Endoscopy.* 2016;48:617-624.
- Kumagai Y, Takubo K, Kawada K, Aoyama K, Endo Y, Ozawa T, *et al.* Diagnosis Using Deep-Learning Artificial Intelligence Based on the Endocytoscopic Observation of the Esophagus. *Esophagus.* 2019;16:180-187.
- Horie Y, Yoshio T, Aoyama K, Yoshimizu S, Horiuchi Y, Ishiyama A, *et al.* Diagnostic Outcomes of Esophageal Cancer by Artificial Intelligence Using Convolutional Neural Networks. *Gastrointest Endosc.* 2019;89:25-32.
- Luo H, Xu G, Li C, He L, Luo L, Wang Z, *et al.* Real-Time Artificial Intelligence for Detection of Upper Gastrointestinal Cancer by Endoscopy: A Multicentre, Case-Control, Diagnostic Study. *Lancet Oncol.* 2019;20(12):1645-1654.
- Cao JS, Lu ZY, Chen MY, Zhang B, Juengpanich S, Hu JH, *et al.* Artificial Intelligence in Gastroenterology and Hepatology: Status and Challenges. *World J Gastroenterol.* 2021;27(16):1664-1690.
- Hashimoto R, Requa J, Dao T, Ninh A, Tran E, Mai D, *et al.* Artificial Intelligence Using Convolutional Neural Networks for Real-Time Detection of Early Esophageal Neoplasia in Barrett's Esophagus (With Video). *Gastrointest Endosc.* 2020;91:1264-1271.e1
- Ebigbo A, Mendel R, Probst A, Manzeneder J, Prinz F, de Souza LA Jr, *et al.* Real-Time Use of Artificial Intelligence in The Evaluation of Cancer in Barrett's Oesophagus. *Gut.* 2020;69:615-616.
- de Groof AJ, Struyvenberg MR, van der Putten J, van der Sommen F, Fockens KN, Curvers WL, *et al.* Deep-Learning System Detects Neoplasia in Patients with Barrett's Esophagus with Higher Accuracy Than Endoscopists in a Multistep Training and Validation Study With Benchmarking. *Gastroenterology.* 2020;158(4):915-929.e4.
- Cai SL, Li B, Tan WM, Niu XJ, Yu HH, Yao LQ, *et al.* Using a Deep Learning System in Endoscopy for Screening of Early Esophageal Squamous Cell Carcinoma (With Video). *Gastrointest Endosc.* 2019;90:745-753.e2.

17. Tokai Y, Yoshio T, Aoyama K, Horie Y, Yoshimizu S, Horiuchi Y, *et al.* Application of Artificial Intelligence Using Convolutional Neural Networks in Determining the Invasion Depth of Esophageal Squamous Cell Carcinoma. *Esophagus*. 2020;17:250-256.
18. Menon S, Trudgill N. How Commonly is Upper Gastrointestinal Cancer Missed at Endoscopy? *Endosc Int Open*. 2014;2:E46-E50.
19. Wu L, Zhang J, Zhou W, An P, Shen L, Liu J, *et al.* Randomised Controlled Trial of WISENSE, a Real-Time Quality Improving System for Monitoring Blind Spots During Esophagogastroduodenoscopy. *Gut*. 2019;68:2161-2169.
20. Ali H, Yasmin M, Sharif M, Rehmani MH. Computer Assisted Gastric Abnormalities Detection Using Hybrid Texture Descriptors for Chromoendoscopy Images. *Comput Methods Programs Biomed*. 2018;157:39-47.
21. Nagao S, Tsuji Y, Sakaguchi Y, Takahashi Y, Minatsuki C, Niimi K, *et al.* Highly Accurate Artificial Intelligence Systems to Predict the Invasion Depth of Gastric Cancer: Efficacy of Conventional White-Light Imaging, Nonmagnifying Narrow-Band Imaging, and Indigo-Carmine Dye Contrast Imaging. *Gastrointest Endosc* 2020; 92:866-873:e1.
22. Sakai Y, Takemoto S, Hori K, Nishimura M, Ikematsu H, Yano T, *et al.* Automatic Detection of Early Gastric Cancer in Endoscopic Images Using a Transferring Convolutional Neural Network. *Conf Proc IEEE Eng Med Biol Soc*. 2018;2018:4138-4141.
23. Tsujii Y, Kato M, Inoue T, Yoshii S, Nagai K, Fujinaga T, *et al.* Integrated Diagnostic Strategy for the Invasion Depth of Early Gastric Cancer by Conventional Endoscopy and EUS. *Gastrointest Endosc*. 2015;82(3):452-9.
24. Wang CC, Chiu YC, Chen WL, Yang TW, Tsai MC, Tseng MH. A Deep Learning Model for Classification of Endoscopic Gastroesophageal Reflux Disease. *Int J Environ Res Public Health*. 2021; 18(5): 2428.
25. Shichijo S, Endo Y, Aoyama K, Takeuchi Y, Ozawa T, Takiyama H, *et al.* Application of Convolutional Neural Networks for Evaluating Helicobacter Pylori Infection Status on the Basis of Endoscopic Images. *Scand J Gastroenterol*. 2019;54(2):158-163.
26. Le Berre C, Sandborn WJ, Aridhi S, Devignes MD, Fournier L, Smail-Tabbone M, *et al.* Application of Artificial Intelligence to Gastroenterology and Hepatology. *Gastroenterology*. 2020;158(1):76-94.e2.
27. Seguí S, Drozdal M, Pascual G, Radeva P, Malagelada C, Azpíroz F, *et al.* Generic Feature Learning for Wireless Capsule Endoscopy Analysis. *Comput Biol Med*. 2016;79:163-172.
28. Pannala R, Krishnan K, Melson J, Parsi MA, Schulman AR, Sullivan S, *et al.* Artificial Intelligence in Gastrointestinal Endoscopy. *VideoGIE*. 2020;5(12):598-613.
29. Chen H, Wu X, Tao G, Peng Q. Automatic Content Understanding with Cascaded Spatial-Temporal Deep Framework for Capsule Endoscopy Videos. *Neurocomputing*. 2017;229:77-87.
30. Ding Z, Shi H, Zhang H, Meng L, Fan M, Han C, *et al.* Gastroenterologist-Level Identification of Small-Bowel Diseases and Normal Variants by Capsule Endoscopy Using a Deep-Learning Model. *Gastroenterology*. 2019;157(4):1044-1054.e5.
31. Aoki T, Yamada A, Kato Y, Saito H, Tsuboi A, Nakada A, *et al.* Automatic Detection of Blood Content in Capsule Endoscopy Images Based on a Deep Convolutional Neural Network. *J Gastroenterol Hepatol*. 2020; 35(7):1196-1200.
32. Xiao Jia, Meng MQ. A Deep Convolutional Neural Network for Bleeding Detection in Wireless Capsule Endoscopy Images. *Annu Int Conf IEEE Eng Med Biol Soc*. 2016;2016:639-642
33. Usman MA, Sattrya GB, Usman MR, Shin SY. Detection of Small Colon Bleeding in Wireless Capsule Endoscopy Videos. *Comput Med Imaging Graph*. 2016;54:16-26.
34. Leenhardt R, Vasseur P, Li C, Saurin JC, Rahmi G, Cholet F, *et al.* A Neural Network Algorithm for Detection of GI Angiectasia During Small-Bowel Capsule Endoscopy. *Gastrointest Endosc*. 2019;89:189-194.
35. Sengupta N, Tapper EB. Derivation and Internal Validation of a Clinical Prediction Tool for 30-Day Mortality in Lower Gastrointestinal Bleeding. *Am J Med*. 2017;130:601.e1-601.e8.
36. Soffer S, Klang E, Shimon O, Nachmias N, Eliakim R, Ben-Horin S, *et al.* Deep Learning For Wireless Capsule Endoscopy: A Systematic Review and Meta-Analysis. *Gastrointest Endosc*. 2020;92:831-839.e8.
37. Qin K, Li J, Fang Y, Xu Y, Wu J, Zhang H, *et al.* Convolution Neural Network for the Diagnosis of Wireless Capsule Endoscopy: A Systematic Review and Meta-Analysis. *Surg Endosc*. 2022;36(1):16-31.
38. Pannala R, Krishnan K, Melson J, Parsi MA, Schulman AR, Sullivan S, *et al.* Artificial Intelligence in Gastrointestinal Endoscopy. *VideoGIE*. 2020;5(12):598-613.
39. Khorasani HM, Usefi H, Peña-Castillo L. Detecting Ulcerative Colitis from Colon Samples Using Efficient Feature Selection and Machine Learning. *Sci Rep*. 2020;10:13744.
40. Wang P, Berzin TM, Glissen Brown JR, Bharadwaj S, Becq A, Xiao X, *et al.* Real-Time Automatic Detection System Increases Colonoscopic Polyp and Adenoma Detection Rates: A Prospective Randomized Controlled Study. *Gut*. 2019;68(10): 1813-1819.
41. Zhao S, Wang S, Pan P, Xia T, Chang X, Yang X, *et al.* Magnitude, Risk Factors, and Factors Associated with Adenoma Miss Rate of Tandem Colonoscopy: A Systematic Review and Meta-analysis. *Gastroenterology*. 2019;156(6):1661-1674.e11.
42. Gohari MR, Biglarian A, Bakhshi E, Pourhoseingholi MA. Use of an Artificial Neural Network to Determine Prognostic Factors in Colorectal Cancer Patients. *Asian Pac J Cancer Prev*. 2011;12:1469-1472.
43. Biglarian A, Bakhshi E, Gohari MR, Khodabakhshi R. Artificial Neural Network for Prediction of Distant Metastasis in Colorectal Cancer. *Asian Pac J Cancer Prev*. 2012;13:927-930.
44. Hassan C, Spadaccini M, Iannone A, Maselli R, Jovani M, Chandrasekar VT. Performance of Artificial Intelligence in Colonoscopy for Adenoma and Polyp Detection: A Systematic Review and Meta-Analysis. *Gastrointest Endosc*. 2021;93(1):77-85.e6.
45. Robles-Medranda C, Cifuentes-Gordillo C, Arevalo-Mora M, Egas-Izquierdo M, Puga-Tejada M, Cunto D, *et al.* Real-Time Computer-Aided Polyp and Adenoma Detection During Screening Colonoscopy in Expert and Non-Expert Endoscopists: A Single Center Study. *Endoscopy*. 2023;55(S 02):S232-S233.
46. Sánchez-Montes C, Sánchez FJ, Bernal J, Córdova H, López-Cerón M, Cuatrecasas M, *et al.* Computer-aided prediction of polyp histology on white light colonoscopy using surface pattern analysis. *Endoscopy*. 2019;51(3):261-5.
47. Kahaleh M, Gaidhane M, Rajman I, Tyberg A, Sethi A, Slivka A, *et al.* Digital Cholangioscopic Interpretation: When North Meets the South. *Dig Dis Sci*. 2022;67(4):1345-1351.
48. Robles-Medranda C, Valero M, Soria-Alcivar M, Puga-Tejada M, Oleas R, Ospina-Arboleda J, *et al.* Reliability and Accuracy of a Novel Classification System Using Peroral Cholangioscopy for the Diagnosis of Bile Duct Lesions. *Endoscopy*. 2018;50(11):1059-70.

49. Robles-Medranda C, Oleas R, Sánchez-Carriel M, Olmos JI, Alcívar-Vásquez J, Puga-Tejada M, *et al.* Vascularity Can Distinguish Neoplastic from Non-Neoplastic Bile Duct Lesions During Digital Single-Operator Cholangioscopy. *Gastrointest Endosc.* 2021;93(4):935-41.
50. Kahaleh M, Gaidhane M, Shahid HM, Tyberg A, Sarkar A, Ardengh JC, *et al.* Digital Single-Operator Cholangioscopy Interobserver Study Using a New Classification: The Mendoza Classification (With Video). *Gastrointest Endosc.* 2022;95(2):319-26.
51. Sethi A, Tyberg A, Slivka A, Adler DG, Desai AP, Sejpal D V., *et al.* Digital Single-operator Cholangioscopy (DSOC) Improves Interobserver Agreement (IOA) and Accuracy for Evaluation of Indeterminate Biliary Strictures: The Monaco Classification. *J Clin Gastroenterol.* 2022;56(2).
52. Saraiva MM, Ribeiro T, Ferreira JPS, Boas FV, Afonso J, Santos AL, *et al.* Artificial Intelligence for Automatic Diagnosis of Biliary Stricture Malignancy Status in Single-Operator Cholangioscopy: A Pilot Study. *Gastrointest Endosc.* 2022;95(2):339-48.
53. Pereira P, Mascarenhas M, Ribeiro T, Afonso J, Ferreira JPS, Vilas-Boas F, *et al.* Automatic Detection of Tumor Vessels in Indeterminate Biliary Strictures in Digital Single-Operator Cholangioscopy. *Endosc Int Open.* 2022;10(03):E262-8.
54. Marya NB, Powers PD, Petersen BT, Law R, Storm A, Abusaleh RR, *et al.* Identification of Patients with Malignant Biliary Strictures Using a Cholangioscopy-Based Deep Learning Artificial Intelligence (With Video). *Gastrointest Endosc.* 2022.
55. Robles-Medranda C, Baquerizo-Burgos J, Alcívar-Vásquez J, Kahaleh M, Raijman I, Kunda R, *et al.* Artificial Intelligence for Diagnosing Neoplasia on Digital Cholangioscopy: Development and Multicentric Validation of a Convolutional Neural Network Model. *Endoscopy.* 2023.
56. Dahiya DS, Al-Haddad M, Chandan S, Gangwani MK, Aziz M, Mohan BP, *et al.* Artificial Intelligence in Endoscopic Ultrasound for Pancreatic Cancer: Where Are We Now and What Does the Future Entail? *J Clin Med.* 2022;11(24):7476.
57. Shahidi N, Ou G, Lam E, Enns R, Telford J. When Trainees Reach Competency in Performing Endoscopic Ultrasound: A Systematic Review. *Endosc Int Open.* 2017;5(4):E239-43.
58. Goyal H, Mann R, Gandhi Z, Perisetti A, Zhang Z, Sharma N, *et al.* Application of Artificial Intelligence in Pancreaticobiliary Diseases. *Ther Adv Gastrointest Endosc.* 2021;14:1-12.
59. Norton ID, Zheng Y, Wiersema MS, Greenleaf J, Clain JE, DiMagno EP. Neural Network Analysis of EUS Images to Differentiate Between Pancreatic Malignancy and Pancreatitis. *Gastrointest Endosc.* 2001; 54(5): 625-9.
60. Săftoiu A, Vilman P. Endoscopic ultrasound elastography Endoscopic Ultrasound Elastography-a New Imaging Technique for the Visualization of Tissue Elasticity Distribution. *J Gastrointest Liver Dis.* 2006;15(2):161-5.
61. Săftoiu A, Vilman P, Gorunescu F, Janssen J, Hocke M, Larsen M, *et al.* Efficacy of an Artificial Neural Network-Based Approach to Endoscopic Ultrasound Elastography in Diagnosis of Focal Pancreatic Masses. *Clin Gastroenterol Hepatol.* 2012;10(1).
62. Brosens LAA, Hackeng WM, Offerhaus J, Hruban RH, Wood LD. Pancreatic Adenocarcinoma Pathology: Changing “Landscape.” *J Gastrointest Oncol.* 2015;6(4):358-74.
63. Kuwahara T, Hara K, Mizuno N, Okuno N, Matsumoto S, Obata M, *et al.* Usefulness of Deep Learning Analysis for the Diagnosis of Malignancy in Intraductal Papillary Mucinous Neoplasms of the Pancreas. *Clin Transl Gastroenterol.* 2019;10(5).
64. Hirai K, Kuwahara T, Furukawa K, Kakushima N, Furune S, Yamamoto H, *et al.* Artificial Intelligence-Based Diagnosis of Upper Gastrointestinal Subepithelial Lesions on Endoscopic Ultrasonography Images. *Gastric Cancer.* 2022;25(2):382-91.
65. Minoda Y, Ihara E, Fujimori N, Nagatomo S, Esaki M, Hata Y, *et al.* Efficacy of Ultrasound Endoscopy with Artificial Intelligence for the Differential Diagnosis of Non-Gastric Gastrointestinal Stromal Tumors. *Sci Rep.* 2022;12(1).
66. Robles-Medranda C, Oleas R, Del Valle R, Mendez JC, Alcívar-Vásquez J, Puga-Tejada M, *et al.* ID: 3521955 Application of Artificial Intelligence for Real-Time Anatomical Recognition During Endoscopic Ultrasound Evaluation: A Pilot Study. *Gastrointest Endosc.* 2021;93(6):AB221.
67. Zhang J, Zhu L, Yao L, Ding X, Chen D, Wu H. Deep Learning-Based Pancreas Segmentation and Station Recognition System in EUS: Development and Validation of a Useful Training Tool (With Video). *Gastrointest Endosc.* 2020; 92(4):874-885.e3.
68. Zhang L, Lu Z, Yao L, Dong Z, Zhou W, He C, *et al.* Effect of a Deep Learning-based Automatic Upper GI Endoscopic Reporting System: A Randomized Crossover Study (With Video). *Gastrointest Endosc.* 2023;98(2):181-190.
69. Qu JY, Li Z, Su JR, Ma MJ, Xu CQ, Zhang AJ, *et al.* Development and Validation of an Automatic Image-Recognition Endoscopic Report Generation System: A Multicenter Study. *Clin Transl Gastroenterol.* 2023;12(1):e00282.