

Sistema de auditoría automática para la exploración endoscópica del estómago con Inteligencia Artificial – Gastro UNAL: Gastroendoscopy UNit for Automatic Labeling

Automatic Auditing System for Endoscopic Exploration of the Stomach with Artificial Intelligence-Gastro UNAL: Gastroendoscopy UNit for Automatic Labeling

Martín Gómez,^{1*}  Diego Bravo,²  Josué Ruano,³  María Jaramillo,²  Fabio A. González,⁴  Eduardo Romero.⁵ 

ACCESO ABIERTO

Citación:

Gómez M, Bravo D, Ruano J, Jaramillo M, González F, Romero E. Sistema de auditoría automática para la exploración endoscópica del estómago con Inteligencia Artificial – Gastro UNAL: Gastroendoscopy UNit for Automatic Labeling. *Revista. colomb. Gastroenterol.* 2024;39(2):133-145. <https://doi.org/10.22516/25007440.1163>

¹ Médico internista, gastroenterólogo, Hospital Universitario Nacional de Colombia. Profesor asociado de Medicina Interna, director de la unidad de Gastroenterología, Universidad Nacional de Colombia. Bogotá, Colombia.

² Estudiante de Doctorado en Ingeniería Eléctrica, grupo de investigación Computer Imaging and Medical Applications Laboratory (CIM@LAB), Universidad Nacional de Colombia. Bogotá, Colombia.

³ Estudiante de Doctorado en Ingeniería de Sistemas y Computación, grupo de investigación Computer Imaging and Medical Applications Laboratory (CIM@LAB), Universidad Nacional de Colombia. Bogotá, Colombia.

⁴ Doctor en Ciencias de la Computación. Director grupo de investigación Machine Learning, Perception and Discovery (MINDLAB). Profesor titular, Departamento de Ingeniería de Sistemas e Industrial, Universidad Nacional de Colombia. Bogotá, Colombia.

⁵ Doctor en Ciencias Biomédicas, director de la Maestría en Ingeniería Biomédica, director del grupo de investigación Computer Imaging and Medical Applications Laboratory (CIM@LAB). Profesor, Facultad de Medicina, Universidad Nacional de Colombia. Bogotá, Colombia.

*Correspondencia: Martín Gómez.
martinalonsogomez@gmail.com

Fecha recibido: 19/12/2023

Fecha aceptado: 10/04/2024



Resumen

Introducción: la endoscopia digestiva alta es el método estándar para diagnosticar el cáncer gástrico en etapas tempranas. Sin embargo, su precisión puede verse afectada por la variabilidad en su realización, y se estiman hasta 20% de tumores no detectados. En Colombia, la mayoría de los diagnósticos se realizan en etapas avanzadas, lo que agrava el problema. Para abordar la variabilidad, se han propuesto protocolos con el fin de asegurar la observación completa de áreas propensas a lesiones premalignas. **Objetivo:** construir y validar un sistema de auditoría automática para endoscopias usando técnicas de inteligencia artificial. **Metodología:** en este estudio, 96 pacientes de un hospital universitario se sometieron a endoscopias documentadas en video, abarcando 22 estaciones reorganizadas para minimizar solapamientos y mejorar la identificación de 13 regiones gástricas clave. Se utilizó una red convolucional avanzada para procesar las imágenes, extrayendo características visuales, lo que facilitó el entrenamiento de la inteligencia artificial en la clasificación de estas áreas. **Resultados:** el modelo, llamado *Gastro UNAL*, fue entrenado y validado con imágenes de 67 pacientes (70% de los casos) y probado con 29 pacientes distintos (30% de los casos), con lo que alcanzó una sensibilidad promedio del 85,5% y una especificidad del 98,8% en la detección de las 13 regiones gástricas. **Conclusiones:** la eficacia del modelo sugiere su potencial para asegurar la calidad y precisión de las endoscopias. Este enfoque podría confirmar las regiones evaluadas, alertando puntos ciegos en la exploración a los endoscopistas con menos experiencia o en entrenamiento, de tal forma que se aumente la calidad de estos procedimientos.

Palabras clave

Esofagogastroduodenoscopia, inteligencia artificial, puntos ciegos de diagnóstico, red neural.

Abstract

Introduction: Upper endoscopy is the standard method for diagnosing early-stage gastric cancer. However, according to estimates, up to 20% of tumors are not detected, and their accuracy may be affected by the variability in their performance. In Colombia, most diagnoses take place in advanced stages, which aggravates the problem. Protocols have been proposed to ensure the complete observation of areas prone to premalignant lesions to address variability. **Objective:** To build and validate an automatic audit system for endoscopies using artificial intelligence techniques. **Methodology:** In this study, 96 patients from a teaching hospital underwent video-documented endoscopies, spanning 22 stations rearranged to minimize overlaps and improve the identification of 13 key gastric regions. An advanced convolutional network was used to process the images, extracting visual characteristics, which facilitated the training of artificial intelligence in the classification of these areas. **Results:** the model, called *Gastro UNAL*, was trained and validated with images of 67 patients (70% of cases) and tested with 29 different patients (30% of cases), which reached an average sensitivity of 85,5% and a specificity of 98,8% in detecting the 13 gastric regions. **Conclusions:** The effectiveness of the model suggests its potential to ensure the quality and accuracy of endoscopies. This approach could confirm the regions evaluated, alerting less experienced or trained endoscopists about blind spots in the examinations, thus, increasing the quality of these procedures.

Keywords

Esofagogastroduodenoscopy, artificial intelligence, diagnostic blind spots, neural network.

INTRODUCCIÓN

El cáncer gástrico (CG) es un problema de salud pública. En el mundo se diagnosticó cerca de un millón de nuevos casos de CG en 2022, mientras que la enfermedad fue la quinta causa de muerte más común, con 660.175 personas muertas en ese mismo periodo⁽¹⁾. En Colombia, según GLOBOCAN 2022, se estimó que ocurrieron 8938 casos nuevos en ese año, que representaron el cuarto lugar en frecuencia (7,6%) y la primera causa de mortalidad por cáncer con 6901 casos (12,2%)⁽¹⁾. Esto ubica la tasa media de mortalidad por CG en Colombia en cerca de 11,5 por cada 100.000 habitantes, pero muestra cifras aún más altas en el área andina^(2,3). Además de esta alta tasa de incidencia, se le suma que la mayoría de los casos se detectan en estados avanzados, lo cual lo convierte en un verdadero problema de salud pública.

Por otro lado, se sabe que la tecnología endoscópica ha avanzado en las últimas décadas y ahora se utiliza ampliamente para el diagnóstico del cáncer gástrico temprano (CGT)⁽⁴⁾. La esofagogastroduodenoscopia (EGD) o endoscopia gastrointestinal alta es el procedimiento de elección para examinar el estómago en busca de lesiones gástricas premalignas o cáncer. Se ha reportado que entre el 11% y el 20% de los casos de cáncer gástrico temprano no se diagnostican durante una EGD⁽⁵⁾. Esta alta tasa de error se debe a muchos factores, principalmente a factores humanos que pueden afectar la eficacia de la endoscopia en la detección temprana de lesiones gástricas⁽⁶⁾.

Entre estos factores se incluyen la presencia de moco y saliva en la mucosa gástrica, la realización del examen muy rápido y el mal entrenamiento del endoscopista que no evalúa todas las áreas del estómago y deja algunos puntos ciegos⁽⁷⁾. Por tanto, evitar los puntos ciegos es un requisito previo fundamental para la eficacia de la endoscopia en la detección del cáncer gástrico en etapa temprana^(8,9). En general, es muy difícil para un médico en entrenamiento adquirir las habilidades necesarias para maniobrar el dispositivo endoscópico e interpretar la anatomía observada desde un endoscopio; este entrenamiento requiere de cientos de procedimientos supervisados⁽¹⁰⁾. En este contexto, diferentes asociaciones en gastroenterología a nivel mundial han desarrollado protocolos para mejorar la eficiencia y calidad del procedimiento⁽¹¹⁻¹³⁾, de tal forma que la toma sistemática de fotografías en áreas anatómicas específicas del estómago es recomendada como uno de los indicadores de calidad para los procedimientos de endoscopia por la Sociedad Británica de Gastroenterología (BSG)⁽¹⁴⁾ y la Sociedad Europea de Endoscopia Gastrointestinal (ESGE)⁽¹⁵⁾, que recomiendan la adquisición de imágenes en ocho ubicaciones específicas del estómago. Por una parte, la ESGE propone una tasa de captura de imágenes (fotodocumenta-

ción) mayor del 90% de todos los marcadores anatómicos como indicador de un examen completo. Por otra parte, la Sociedad de Endoscopia Gastrointestinal de Japón (JGES) implementó un protocolo de detección sistemática del estómago (SSS) más extenso⁽¹⁶⁾ que fue descrito inicialmente por el profesor Kenshi Yao y consiste en una serie de 22 fotografías endoscópicas que capturan 4 áreas que cubren el 100% de la superficie gástrica: antro gástrico, cuerpo distal, parte media y superior del cuerpo gástrico⁽⁸⁾.

Recientemente, la Organización Mundial de Endoscopia (WEO) les proporcionó una guía práctica a los profesionales para realizar una fotodocumentación exhaustiva de los procedimientos endoscópicos. En este documento se detallaron 28 áreas endoluminales distintas que abarcan la totalidad del revestimiento interno del tracto gastrointestinal superior, desde la hipofaringe hasta la segunda porción del duodeno⁽¹⁷⁾. En nuestro estudio nos enfocamos en la cavidad gástrica; por tanto, utilizamos la guía del Dr. Yao como referencia para nuestra investigación.

A pesar de la alta concordancia entre expertos, no se ha conseguido difundir ampliamente estos protocolos, y muchos endoscopistas no documentan visualmente estas regiones y probablemente tampoco las observan⁽¹⁵⁾. Esta situación plantea una dificultad para asegurar la realización de una inspección completa del estómago, lo que podría, al menos en parte, explicar la diferencia en la detección de tumores en etapas tempranas entre Japón y Colombia. En Japón, aproximadamente el 70% de los casos se detectan en esta fase, que es curable. En contraste, en Colombia, este porcentaje es mínimamente representativo, dado que, como se mencionó anteriormente, más del 90% de los casos diagnosticados en el país se encuentran en estado avanzado, cuando la enfermedad es potencialmente mortal⁽¹⁹⁻²³⁾. Por esto sería deseable que los endoscopistas practicaran una inspección sistemática del estómago documentando estas 22 zonas específicas, lo que garantizaría la eliminación de puntos ciegos o áreas no evaluadas. Sin embargo, para validar que esto se haga de manera rutinaria, sería útil contar con un sistema de supervisión (monitorización) que verificara la evaluación de todas las zonas o alertara al endoscopista sobre las áreas que podrían haber quedado sin evaluar durante el examen.

En este contexto, la integración de métodos computacionales en la práctica médica puede brindar un apoyo sustancial a los gastroenterólogos para que puedan realizar la endoscopia completa y de calidad de manera rutinaria y en todos los casos. Estos métodos pueden actuar como un segundo lector de las regiones anatómicas que deben ser analizadas por el especialista. Este segundo lector podría identificar áreas que no han sido examinadas, evaluadas de manera incompleta o puntos ciegos en la cavidad gástrica en su totalidad, lo que podría contribuir a reducir la omisión de lesiones precance-

rosas y malignas. Al verificar si se han inspeccionado todos los puntos de referencia anatómicos del protocolo del Dr. Yao, el enfoque computacional puede jugar un papel esencial. En situaciones en las que la estrategia no detecte una exploración completa, puede funcionar como una alerta que le notifique al especialista que una región anatómica específica aún no ha sido evaluada. Esto es de suma importancia, ya que un cáncer gástrico en sus primeras etapas puede pasar desapercibido, dado que a menudo se presentan como lesiones pequeñas o sutiles que podrían pasar inadvertidas si no se examina una zona específica.

MATERIALES Y MÉTODOS

En este trabajo se explora la clasificación automática de diferentes regiones del estómago que deben revisarse durante los

procedimientos endoscópicos con métodos de clasificación de imágenes con aprendizaje profundo. La **Figura 1** exhibe la disposición de las regiones de interés en el estómago, acompañadas de imágenes de muestra, e ilustra la metodología propuesta en este trabajo. La estrategia utilizada es el aprendizaje por transferencia, que aprovecha un modelo de aprendizaje profundo previamente entrenado con millones de imágenes naturales para capturar características visuales de bajo nivel (bordes), adaptar las características de alto nivel del modelo a imágenes de endoscopia y utilizarlas como entrada para un módulo de clasificación.

La primera etapa del proceso, la extracción de características, corresponde a una Red neuronal convolucional (CNN) que ha sido previamente entrenada con un gran conjunto de imágenes naturales. La segunda etapa, clasificación, corresponde a un conjunto de capas densas que se

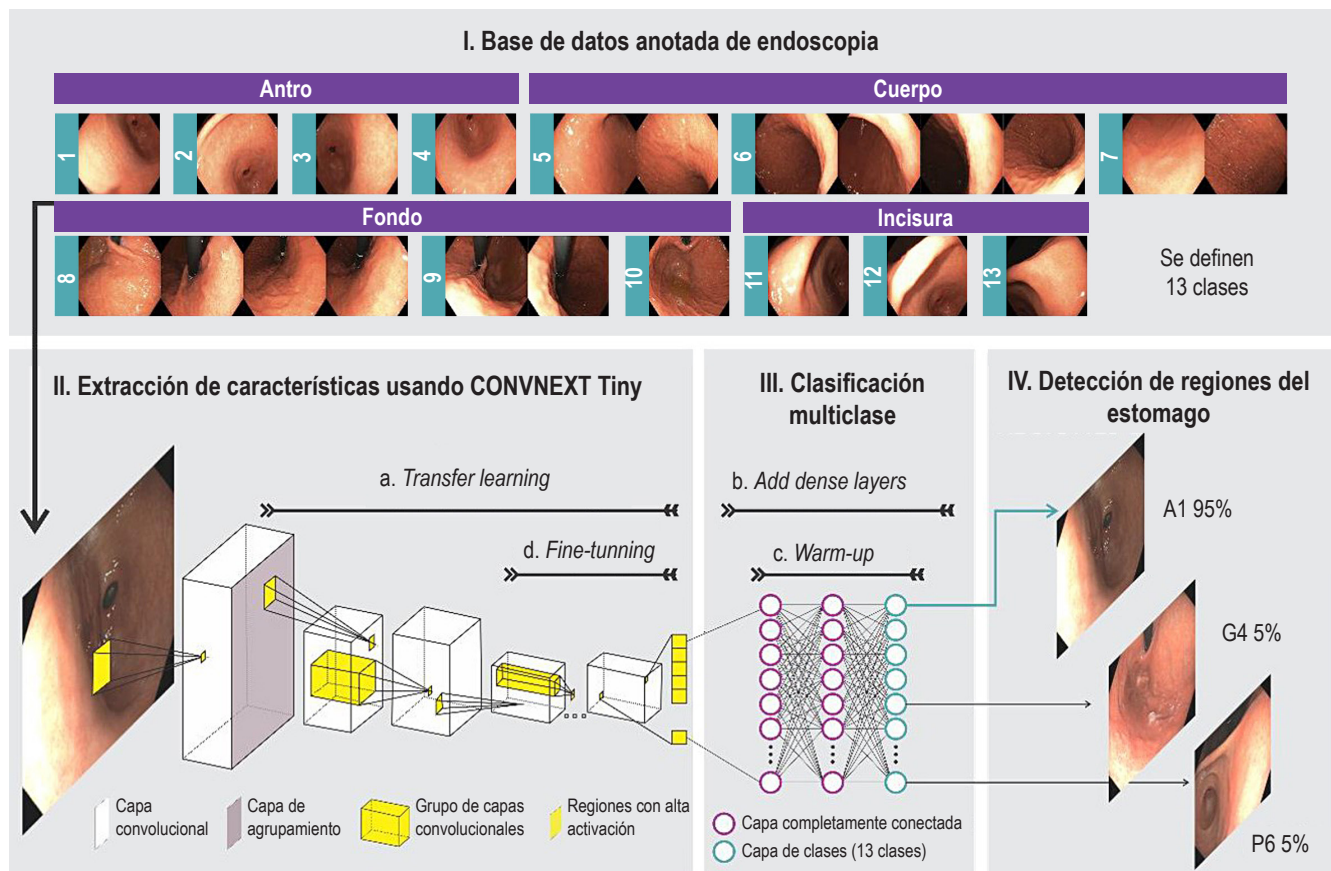


Figura 1. Flujo del método propuesto para identificar automáticamente 13 regiones gástricas. En (I) se consolida la base de datos de imágenes endoscópicas categorizada en regiones anatómicas, dividida en particiones de entrenamiento, validación y prueba. En (II), cada cuadro de entrenamiento alimenta un modelo basado en redes neuronales convolucionales. Este modelo (II-a) está preentrenado con millones de imágenes naturales, en (III-b) se añade una capa densamente conectada a la red neuronal y (III-c) se preentrena. En (II-d) se liberan capas de neuronas encargadas de la extracción de la información para aprender las relaciones contextuales. La red entrenada es evaluada en (IV) para clasificar cada cuadro en alguna de las 13 regiones del estómago. Fuente: elaboración propia.

entrena utilizando el conjunto de imágenes específicas del problema. La salida del modelo es una capa softmax con 13 neuronas correspondientes a las 13 clases diferentes de regiones del estómago, cuya función es asignar una probabilidad a todas las clases, destacando la más alta como la clasificación definitiva para la región en cuestión. Para la etapa de extracción de características se usó una arquitectura de CNN de última generación, como la ConvNext Tiny⁽²⁴⁾. Esta arquitectura fue seleccionada por presentar el mejor rendimiento en un estudio previo comparado con otras 22 redes evaluadas⁽²⁵⁾.

Protocolo de adquisición y preprocesamiento

La esofagogastroduodenoscopia y la captura de imágenes se llevan a cabo por un gastroenterólogo altamente experimentado, con más de 20 años de práctica en endoscopia y un historial de más de 50.000 procedimientos realizados. Este especialista trabaja en el Hospital Universitario Nacional de Colombia y sigue el protocolo de endoscopia sistemática desarrollado por el Dr. K. Yao⁽⁸⁾. El proceso se desarrolla de la siguiente manera: cuando un paciente está programado para una endoscopia digestiva alta antes de entrar a la sala, firma los consentimientos informados del procedimiento, la sedación, el protocolo de investigación y la captura de las imágenes; luego, se le proporciona una preparación que consiste en 10 mL de una solución que contiene 400 mg de N-acetil cisteína y 200 mg de simeticona, aproximadamente media hora antes del procedimiento. Luego, se instruye al paciente para que se recueste en posición de cubito lateral izquierdo durante 5 minutos, seguido de una espera de 20 a 30 minutos antes de ingresar a la sala de procedimientos. En la sala, se coloca una cánula en el brazo derecho del paciente, y un anestesiólogo certificado administra la sedación endovenosa con propofol. Una vez el paciente ha sido sedado, se introduce un endoscopio marca Olympus serie 190.

Este endoscopio, al ingresar a la cavidad gástrica, realiza la aspiración de residuos del contenido gástrico, distiende la cavidad y procede a ubicarse en el duodeno. Una vez este se ha inspeccionado, se ubica la posición del píloro y se retira el equipo 5 cm para iniciar la fotodocumentación del antro, iniciando por la curva mayor y en el sentido de las manecillas del reloj, y se toman cuatro fotos que se superponen entre ellas, de la siguiente manera: curvatura mayor (foto 1-G1); pared anterior (foto 2-A1); curvatura menor (foto 3-L1) y pared posterior (foto 4-P1). Desde esta posición se retira 15 cm el equipo y se ubica en el cuerpo gástrico distal, y en el mismo sentido de las manecillas del reloj se toman las siguientes fotos: curvatura mayor (foto 5-G2), pared anterior (foto 6-A2), curvatura menor (foto 7-L2) y pared posterior (foto 8-P2). A continuación, se retira el equipo 15 cm adicionales

y se ubica en el cuerpo gástrico superior-medio, para continuar el registro en el sentido de las manecillas del reloj: curvatura mayor (foto 9-G3); pared anterior (foto 10-A3); curvatura menor (foto 11-L3) y pared posterior (foto 12-P3). Posteriormente, se avanza el gastroscopio hasta la unión corporoantral, donde se realiza retroflexión para visualizar la región del cardias y el fondo gástrico, y desde esta posición se continúa la fotodocumentación así: curvatura mayor (foto 13-G4); pared anterior (foto 14-A4); curvatura menor (foto 15-L3) y pared posterior (foto 16-P4). Con el equipo en retrovisión se empuja por la curvatura menor (5 cm) para visualizar completamente la curva menor, y obtener las siguientes tres fotos: pared anterior (foto 17-A5), curvatura menor (foto 18-L5) y pared posterior (foto 19-P5). Después, se ubica la punta del equipo en la incisura y se visualiza en toda su extensión para obtener las últimas fotos así: pared anterior (foto 20-A6), curva menor (foto 21-L6) y pared posterior (foto 22-P6). El resultado de este registro se presenta en la **Figura 2**.

Una vez capturadas las imágenes de las 22 regiones y después del proceso de asignación de etiquetas en un sistema de información, se procede a iniciar la creación del modelo de inteligencia artificial que se denominó en honor a nuestra universidad GastroUNAL, evidenciamos que algunas áreas son anatómicamente muy similares y resultan solapándose con frecuencia, al punto que puede dificultarse diferenciarlas, por lo cual en conjunto con el equipo decidimos reagrupar las 22 zonas en solo 13 fusionando las que inspeccionaban áreas similares. A continuación, se describen las regiones agrupadas:

- región del cuerpo inferior en L2 y P2, y región media alta en L3 y P3;
- región del cuerpo inferior en A2 y región media alta en A3;
- región del cuerpo inferior en G2 y región media alta en G3;
- en vista retroflex, región del fondo cardias en L4, A4 y región media alta L5, A5;
- en vista retroflex, región del fondo cardias en P4 y región media alta P5.

Sistema de recolección de casos y etiquetado de regiones

El almacenamiento, organización y etiquetado de videos/imágenes endoscópicas con las 22 regiones del estómago fue realizado en un prototipo de sistema de información desarrollado especialmente para este estudio. El sistema fue diseñado de acuerdo con los requerimientos funcionales obtenidos con la asesoría de dos gastroenterólogos. Estos requerimientos incluyen la información clínica que debe solicitarse para capturar un estudio de endoscopia, las estrategias de visualización de videos/imágenes endoscópi-

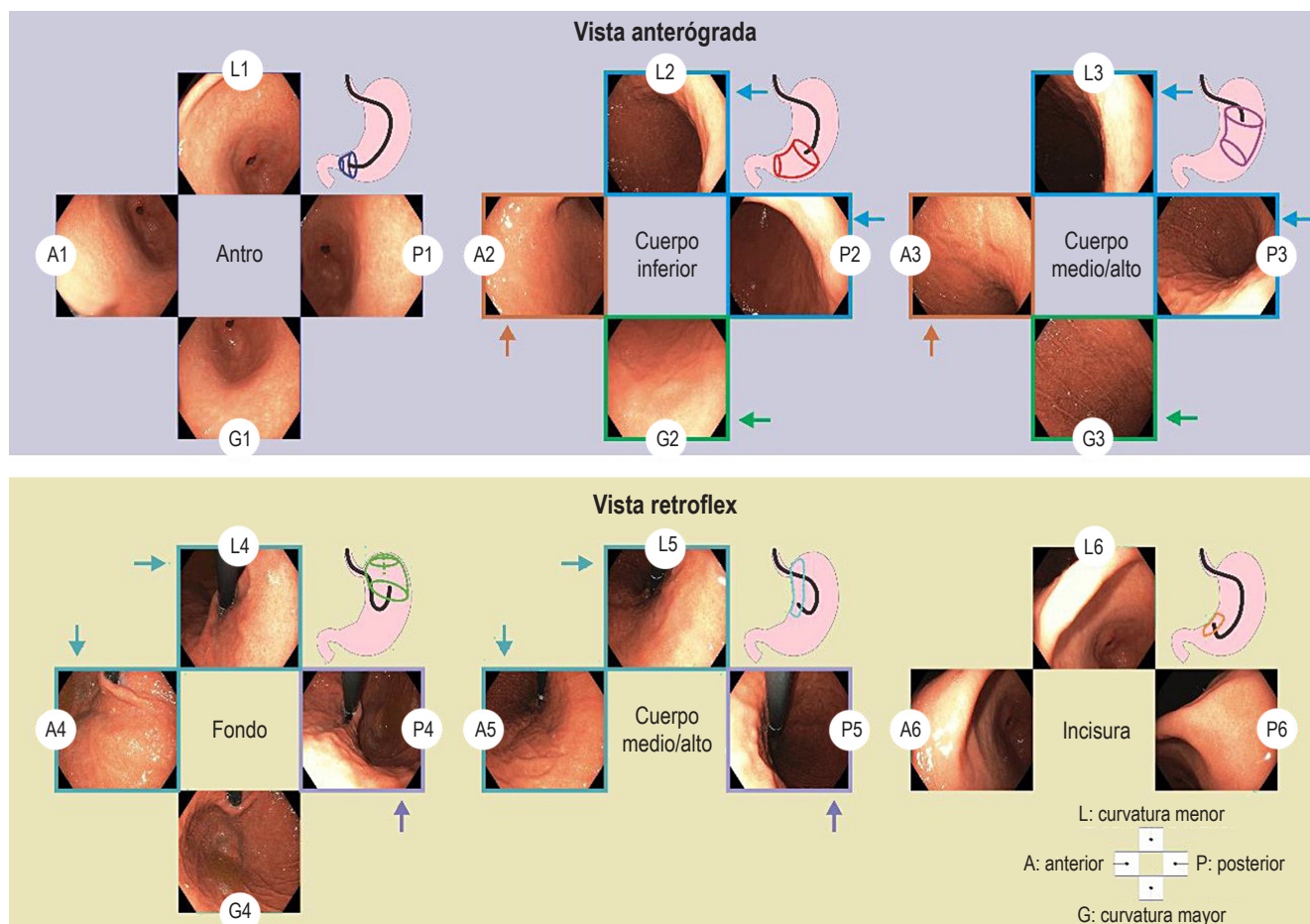


Figura 2. Protocolo de documentación fotográfica del estómago que comienza tan pronto es insertado el endoscopio en el antro gástrico. Con la vista anterógrada se toman fotografías endoscópicas de cuatro cuadrantes del antro gástrico, cuerpo inferior, medio y superior. Luego, con la vista retroflex, se toman fotografías endoscópicas de cuatro cuadrantes del fondo cardias, y de tres cuadrantes del cuerpo medio superior y la incisura gástrica. En total, la serie SSS comprende 22 fotografías endoscópicas. Las flechas representan la unificación de etiquetas para obtener 13 categorías. A: pared anterior; G: curvatura mayor; L: curvatura menor; P: pared posterior; Q: cuadrante. Fuente: elaboración propia.

cas y las herramientas requeridas para etiquetar los vídeos/ imágenes. Los requerimientos no funcionales del sistema se establecieron colaborativamente con un equipo de tres ingenieros con experiencia en prototipos clínicos. Se concentraron en la escalabilidad, seguridad, interoperabilidad y usabilidad, de modo que se garantizó un funcionamiento estable y una gestión segura de datos sensibles. Este sistema fue elaborado usando el entorno de desarrollo web Django conectado con un gestor de base de datos llamado PostgreSQL. La combinación de estas dos tecnologías establece un entorno de captura seguro y confiable para información sensible, como lo son los datos clínicos. Django tiene incorporadas estrategias de seguridad avanzadas para transmitir datos vía web, que posteriormente PostgreSQL encripta y almacena. Acorde a los requerimientos, este sistema cuenta con los siguientes módulos:

- Autenticación: controla el registro y acceso de los usuarios.
- Gestión de pacientes: visualizar, crear y editar la información de los pacientes pertenecientes al estudio.
- Gestión de procedimientos endoscópicos: se asigna uno o varios procedimientos a un paciente en específico, capturando la información clínica asociada, el video del procedimiento y el consentimiento informado.
- Visualización y etiquetado de procedimientos: el sistema dispone el video de un procedimiento en un reproductor web (**Figura 3-I**) junto con la secuencia de cuadros específica para el instante actual del video (**Figura 3-II**). Cuando el usuario, en este caso un gastroenterólogo, observa una de las 22 regiones gástricas mientras visualiza el video, este puede seleccionar y etiquetar una imagen o una secuencia de ellas que represente dicha región (**Figura 3-III**).

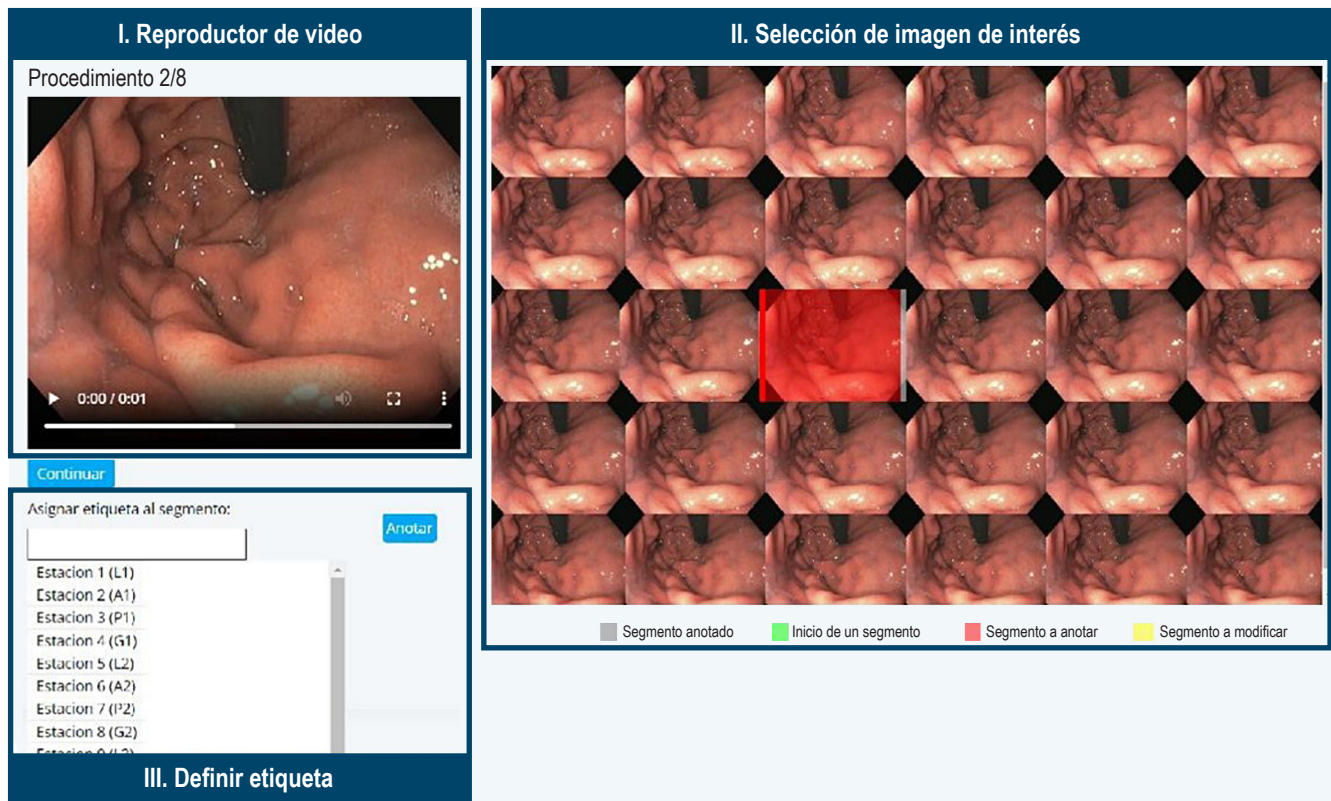


Figura 3. Sistema de etiquetado de las regiones anatómicas. Se visualiza el video del procedimiento y las imágenes asociadas al instante del video para anotar la imagen que mejor represente la región. Luego de seleccionar la imagen, esta se marca con color rojo y se habilita una opción para asignar la etiqueta de dicha imagen. Fuente: elaboración propia.

Arquitectura de la red neuronal convolucional para describir imágenes de endoscopia

Una red neuronal convolucional (CNN) construye una representación jerárquica de patrones gástricos distintivos que provienen de la apariencia; es decir, color y textura, y de estructuras anatómicas en distintas perspectivas, a saber, pliegues gástricos, esfínter esofágico bajo, píloro, incisura, entre otros. Una CNN se ensambla como un conjunto de neuronas, agrupadas por capas, que a su vez están conectadas de manera secuencial, desde la entrada hasta la salida, como se ilustra en la **Figura 1 (II-III)**. Las primeras etapas de este conjunto de capas toman una imagen como entrada y la descomponen en primitivas, que se extraen mediante la aplicación de convoluciones o conjunto de filtros diseñados para capturar patrones locales. Esta transformación inicial de la imagen se procesa secuencialmente en capas consecutivas de convoluciones adicionales, seleccionando información relevante a medida que estas señales se van procesando. A lo largo de este procedimiento, es común introducir capas que se encargan de reducir la dimensio-

nalidad o agrupar información, junto con otras que normalizan la salida de cada neurona mediante funciones de activación para mitigar la atenuación de la señal producida por el procesamiento secuencial. Este proceso, relativamente simple, se repite de capa en capa, transformando la imagen en patrones cada vez más complejos, no lineales y con menos dependencias estadísticas.

La arquitectura de la red está definida por la cantidad de filtros, capas y neuronas, componentes que a su vez determinan la cantidad de parámetros del modelo, típicamente del orden de millones. Para este trabajo se seleccionó la arquitectura ConvNext Tiny⁽²⁴⁾, una red neuronal convolucional que aprende relaciones espaciales a diferentes escalas, mientras conserva la simplicidad y eficiencia de las redes convolucionales tradicionales. La arquitectura de ConvNext se compone de una secuencia de bloques llamados *Inverted Bottleneck Blocks*. Cada bloque consta de los siguientes componentes: convolución con separación de profundidad, mecanismo de autoatención, convolución tradicional, normalización de capas y activación de GeLU (*Gaussian Error Linear Units*). Esta arquitectura incorpora

conexiones directas entre capas no contiguas, llamadas *bloques residuales*, que ayudan a reducir el problema de atenuación del gradiente. La arquitectura del modelo consta de 27 millones de parámetros, de los cuales 11.000 corresponden a las capas completamente conectadas de la red neuronal y el resto a los pesos de las capas convolucionales.

Clasificación de imágenes endoscópicas mediante transferencia de aprendizaje

El rendimiento de una red neuronal convolucional (CNN) en tareas de clasificación depende significativamente del volumen de datos empleados, entrenarla con pocos casos resulta ineficaz. Este estudio implementó aprendizaje por transferencia en la arquitectura ConvNext, utilizando como fuente inicial de información la amplia base de datos ImageNet, que incluye más de 14 millones de imágenes naturales distribuidas en cerca de 20.000 categorías. La ConvNeXt ha sido rigurosamente entrenada y validada con estas imágenes, aprovechando el gran volumen de ejemplos en ImageNet^(26,27). Así como el sistema visual humano construye la representación del mundo con primitivas visuales similares hasta la corteza primaria⁽²⁸⁻³⁰⁾, es en las regiones superiores que estos patrones aprenden asociaciones particulares de un dominio; entonces, estas redes se han entrenado para capturar las unidades informativas de la imagen y este aprendizaje se puede adaptar para este problema específico. Este proceso, aprendizaje por transferencia, consiste en utilizar los pesos de los filtros que fueron aprendidos con millones de ejemplos de imágenes naturales, mientras que algunas capas, usualmente las últimas, aprenden los pesos al presentarle ejemplos de imágenes de endoscopia de vías superiores. Este entrenamiento consta de tres etapas:

- Capas convolucionales congeladas; es decir, aquellas capas cuyos parámetros no se actualizan con imágenes de endoscopia.
- Las capas no congeladas, cuyos parámetros se ajustan al entrenar la red con imágenes de endoscopia (ajuste fino). La salida es un vector de características.
- Las capas encargadas de la clasificación, o capas completamente conectadas, se entrenan para el problema de clasificación específico que se presenta en este trabajo (**Figura 1**). La cantidad de neuronas en la capa de salida se determina como función de la cantidad de clases. Finalmente, esta última capa también entrega una confianza que establece la probabilidad de que una imagen pertenezca a una clase específica (**Figura 1-IV**). El proceso de aprendizaje de la red es iterativo y adaptativo, utilizando una tasa de aprendizaje relativamente baja durante varias iteraciones (épocas). En cada época, se entrenan lotes de datos hasta que todas las muestras del conjunto de entrenamiento se han procesado. La can-

tividad de épocas se determina en función de la complejidad del problema de clasificación. Finalmente, el entrenamiento se concluye cuando el error de la clasificación sobre los datos de validación sea menor a un umbral.

Bases de datos

La base de datos consta de 96 casos de pacientes que se sometieron a procedimientos de EGD. A partir de los videos grabados con luz blanca, se obtuvieron 2054 imágenes, cuya ubicación anatómica se estableció como alguna entre las posibles 22 del protocolo de Kenshi Yao (véase la columna *Etiqueta* de la **Tabla 1**)⁽⁸⁾ y posterior unificación en 13 regiones. Estas anotaciones fueron realizadas en consenso por el grupo de residentes y el jefe de gastroenterología con una experiencia de más de 20 años en el Hospital Universitario Nacional. Cada cuadro fue capturado con una resolución espacial de 1350 x 1080 píxeles. Este estudio se llevó a cabo siguiendo los principios de la Declaración de Helsinki y recibió la aprobación del Comité de Ética del Hospital Universitario Nacional de Colombia, con número de aprobación CEI-2019-06-10. La descripción detallada de la base de datos se presenta en la **Tabla 1**.

CONSIDERACIONES ÉTICAS

El presente trabajo está realizado de acuerdo con la resolución n.º 008430 de 1993 que establece las normas científicas, técnicas y administrativas para la investigación en humanos, artículo 11. Este proyecto se clasifica como investigación con riesgo mínimo dado que solo se requiere del uso de imágenes digitales, las cuales se generaron a partir de videos endoscópicos anonimizados; es decir, no existirá manera alguna de conocer el nombre o la identificación de los sujetos incluidos en el estudio.

Los participantes del estudio otorgaron un consentimiento informado en el que se aclara que la información obtenida en el curso de su atención médica puede utilizarse para propósitos educativos e investigativos. Este acuerdo garantiza que dichos datos, incluidos imágenes y videos anonimizados, se manejarán con los más altos estándares de confidencialidad, seguridad y custodia, cumpliendo con las normativas aplicables a este tipo de investigaciones.

RESULTADOS

En esta sección se evalúa el desempeño de la metodología propuesta que compara las predicciones de una red neuronal convolucional y las anotaciones realizadas por un experto altamente calificado. A continuación, se detalla la configuración experimental y el esquema de validación cuantitativa aplicados al enfoque propuesto.

Tabla 1. Distribución de la base de datos de EGD en imágenes para el esquema de validación

Nombre	Etiqueta*	Entrenamiento (n = 58)	Validación (n = 9)	Prueba (n = 29)	Total (n = 96)
Antro A1	1	48	7	29	84
Antro L1	2	48	5	37	90
Antro P1	3	45	7	30	82
Antro G1	4	56	7	33	96
Cuerpo A2-3	5	114	18	58	190
Cuerpo L2-3, P2-3	6	223	39	121	383
Cuerpo G2-3	7	114	19	58	191
Fondo cardias A4-5, L4-5	8	231	38	120	389
Fondo cardias P4-5	9	106	18	58	182
Fondo cardias G4	10	54	8	33	95
Incisura A6	11	46	12	31	89
Incisura L6	12	56	10	34	100
Incisura P6	13	44	10	29	83

n: cantidad de pacientes. *Etiqueta: corresponde a la estación o agrupación. De 1-7, vista anterógrada, y de 8-13, vista retroflex. Fuente: elaboración propia.

Configuración experimental

En total, nuestro estudio involucró a 96 pacientes con una edad promedio de $62 \pm 15,5$ años. De estos participantes, el 50,6% eran mujeres y el 49,4% eran hombres. El modelo desarrollado fue sometido a una evaluación utilizando un esquema de división 70-30: 70% de los casos para entrenamiento y validación (67 casos, 1383 imágenes) y el 30% restante para prueba (29 casos, 671 imágenes). Como función de pérdida, se utilizó la entropía cruzada con ponderaciones de clase para equilibrar la cantidad de muestras en relación con la clase predominante (Tabla 1). Se mantuvo la proporción desigual en el conjunto de datos de prueba.

La CNN se entrenó en dos etapas: un preentrenamiento de las capas de clasificación *warm-up* con una tasa de aprendizaje constante durante 10 épocas y un ajuste fino (*fine-tuning*) del último 20% de capas durante 100 épocas con un criterio de parada de 15 épocas sin aumento de la métrica macro F1-score. A continuación, se detallan los aspectos de la CNN y la configuración de entrenamiento:

- Pesos previamente entrenados: ImageNet.
- Optimizador: Adam.
- Función de pérdida: entropía cruzada.

- Capas densas: una capa densa, seguida de una capa de normalización por lotes; luego, una capa de coadaptación (deserción) y, finalmente, dos capas densas con 13 categorías.

Las etapas de *warm-up* y *fine-tuning* se incluyeron en una optimización de hiperparámetros en 200 pruebas (*trials*) que monitorizaron la métrica F para encontrar un tamaño de lote óptimo, una tasa de aprendizaje inicial y un programa de tasa de aprendizaje (gamma y tamaño de paso). Los valores durante la optimización fueron los siguientes:

- Tasa de aprendizaje para el preentrenamiento: 0,001 con gamma: 0,1.
- Rango de valores de hiperparámetros durante la optimización: tamaño de lote (8-128), gamma (0,1-0,5), tamaño de paso (5-10) y tasa de aprendizaje ($1e^{-3}$ a $1e^{-5}$).

Después del proceso de optimización, se determinó que los hiperparámetros que maximizaron el puntaje F1 se encontraron en la validación 132/200 con tamaño de lote de 8, tasa de aprendizaje de 0,00071, tamaño de paso de 7 y gamma de 0,30 para el programa de tasa de aprendizaje.

Evaluación cuantitativa

La red neuronal convolucional está diseñada para estimar una de 13 posibles ubicaciones anatómicas para cada imagen, como se ilustra en la **Figura 4**. Estas categorías corresponden a distintas regiones anatómicas: el antro (4 clases), el cuerpo gástrico (3 clases), el fondo cardias (3 clases) y la incisura (3 clases). El rendimiento de la red se evalúa mediante la comparación de sus predicciones con las anotaciones de un experto, utilizando una matriz de confusión que registra verdaderos positivos (VP), verdaderos negativos (VN), falsos positivos (FP) y falsos negativos (FN). Con esta matriz, se calculan métricas específicas para cada clase y un promedio global de estas métricas, tal como se presenta en la **Tabla 2**. Los resultados demuestran que el método propuesto alcanza una clasificación eficaz de las 13 clases, logrando una exactitud promedio del 86%.

Dado que el objetivo principal de este estudio es la detección de regiones anatómicas, es importante resaltar que el método logra el valor más alto de F1 (94,1%) para la región del cardias A4-5, L4-5 en la vista retroflex (**Figura 4**). En contraste, la clase con el valor más bajo de F1 es el cuerpo G2-3, con un 73,7% (**Figura 4**). Esto se debe probablemente a que, en algunos casos, la información del tejido

gástrico no presenta un patrón diferenciador que permita la ubicación precisa de la región, una limitación que se mitigaría con una cantidad mayor de casos para entrenamiento. Sin embargo, es importante resaltar que el método propuesto demuestra capacidad para manejar una alta variabilidad de las diversas estructuras anatómicas y vistas.

Tabla 2. Resultados obtenidos para la configuración propuesta

Métrica (%)	13 estaciones
Exactitud	85,99
Especificidad	98,82
Sensibilidad macro	85,53
Precisión macro	84,73
F1 macro	84,86
Sensibilidad ponderada	85,99
Precisión ponderada	86,64
F1 ponderado	86,07

Fuente: elaboración propia.

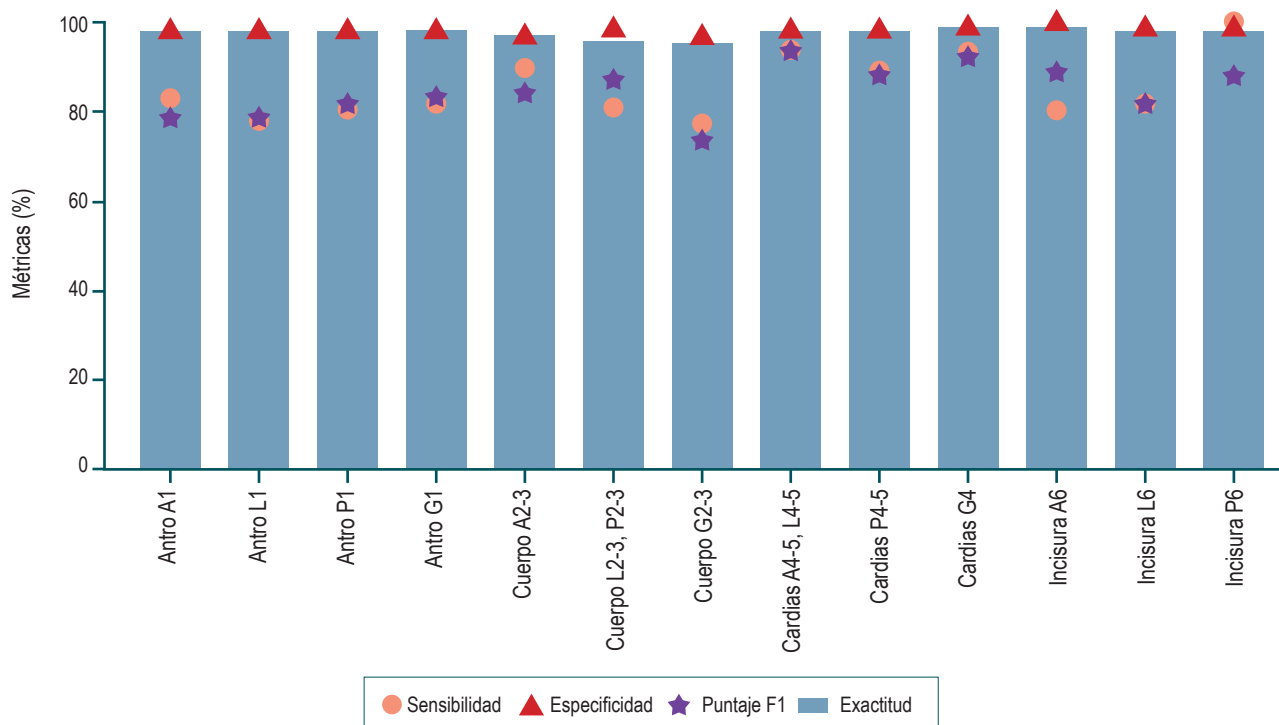


Figura 4. Métricas de sensibilidad, especificidad y exactitud de la CNN para la clasificación de regiones del estómago en imágenes de EGD en 13 ubicaciones. Fuente: elaboración propia.

Evaluación cualitativa

A través de Grad-CAM (*Gradient-weighted Class Activation Mapping*), se identifican las áreas en las imágenes que la red considera más representativas para distinguir entre diferentes regiones anatómicas. En la **Figura 5** se ilustra una evaluación de un caso de prueba con las 13 clases.

En relación con las regiones del antro (A1, L1, P1 y G1), se observa que el patrón de la ubicación del esfínter pilórico es determinante; mientras que, para las regiones del cuerpo gástrico, la red utiliza el cuadrante de la pared gástrica. En lo que respecta al fondo cardias, la estimación del posicionamiento del endoscopio y su ubicación determinan la

estructura de la clase. Finalmente, para la incisura (A6, L6 y P6), se evidencia que su captura (porción anterior, media y posterior) es determinante para su identificación.

DISCUSIÓN

La endoscopia digestiva alta, también conocida como *endoscopia gastroduodenal* (EGD), es el método preferido para diagnosticar el cáncer gástrico y las lesiones premalignas. Identificar y monitorizar estas lesiones es crucial para prevenir esta grave enfermedad, y lograrlo implica llevar a cabo endoscopias de alta calidad. La EGD también se utiliza como tamizaje para el diagnóstico temprano de cáncer

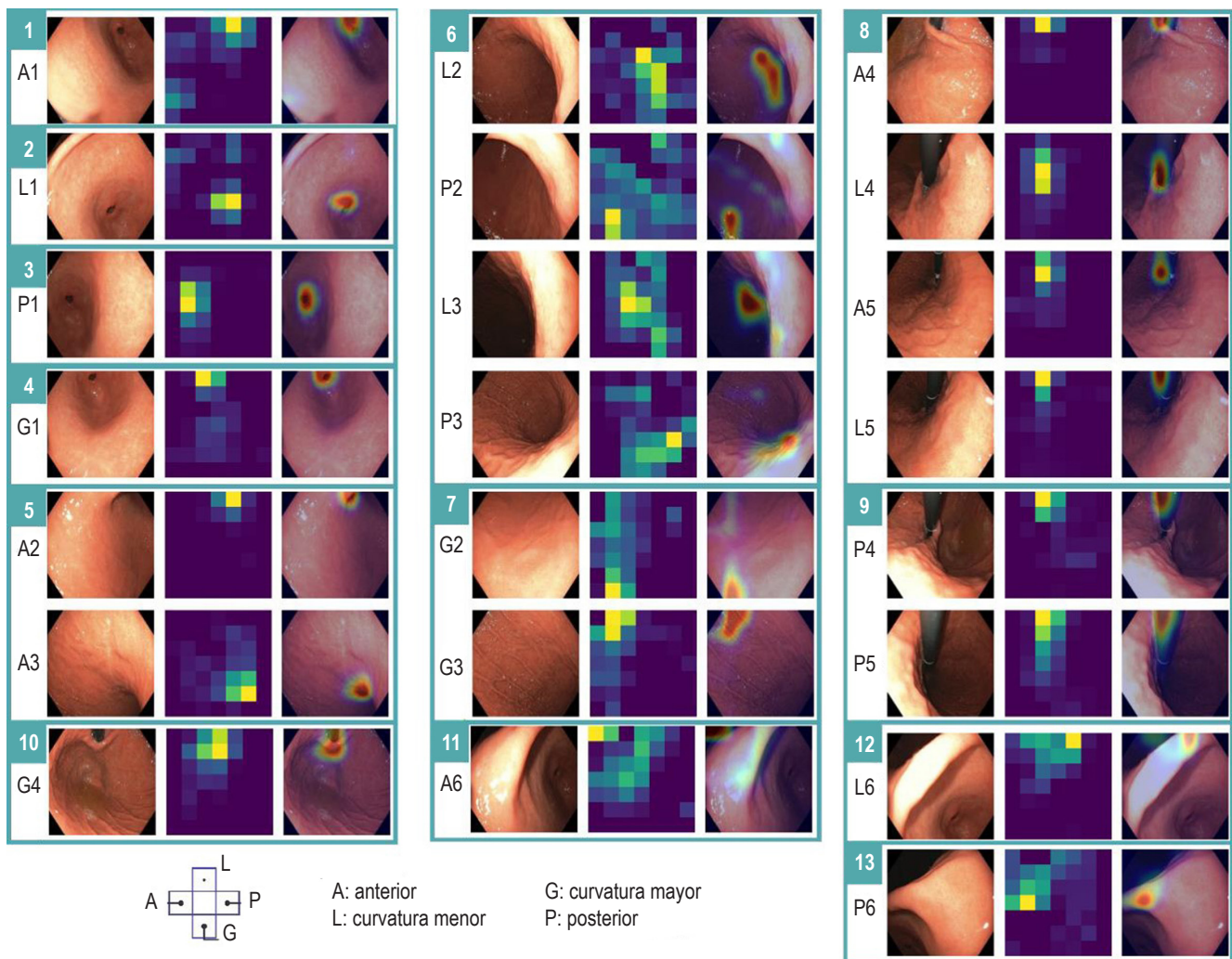


Figura 5. Análisis de una red neuronal convolucional en un caso de prueba. En la secuencia se incluye una imagen endoscópica, un mapa de calor que visualiza la importancia de las áreas de la imagen en la predicción de la red neuronal y la superposición de dicho mapa de calor en la imagen endoscópica para resaltar las zonas relevantes. En este caso, se identificaron 13 regiones (marcadas con un rectángulo verde) con una predicción de confianza superior al 95%, que coinciden con la etiqueta asignada por un experto. Dentro del rectángulo, se observa A1-P6, que corresponde al mecanismo de agrupación de regiones. A: anterior; G: curvatura mayor; L: curvatura menor; P: posterior. Fuente: elaboración propia.

gastrointestinal superior en áreas de alto riesgo⁽¹⁸⁾. En una población de estudio, Hamashima, Chisato y colaboradores⁽³¹⁾ se enfocaron en evaluar el modo en que el tamizaje endoscópico influye en la reducción de la mortalidad por cáncer gástrico. Los resultados mostraron que la mortalidad por cáncer gástrico fue significativamente más baja en el grupo sometido a tamizaje endoscópico comparado con los grupos de tamizaje radiográfico y de fotofluorografía. La conclusión del trabajo indica que el tamizaje por endoscopia para cáncer gástrico está asociado con una notable reducción del 57% en la tasa de mortalidad por esta enfermedad, lo que sugiere su eficacia en la detección temprana y el tratamiento efectivo.

En la práctica, se ha encontrado que si un gastroenterólogo realiza una endoscopia de manera sistemática es poco probable que deje áreas sin observar o que tenga puntos ciegos, pero generalmente no hay registro de cómo un especialista realizó el procedimiento, ni si visitó un mínimo de regiones o durante cuánto tiempo fueron observadas. En este sentido, la fotodocumentación se ha convertido en un aspecto importante del informe de la EGD. Un estudio reciente informó una asociación positiva entre la cantidad de imágenes endoscópicas y la tasa de detección de lesiones gástricas clínicamente significativas ($p < 0,001$). En el contexto clínico, un endoscopista capacitado puede realizar una fotodocumentación sistemática y eficiente. No obstante, esto exige una orientación endoluminal adecuada gracias a una amplia capacitación previa y una mucosa limpia para asegurar una visualización ideal. La limpieza inadecuada de la mucosa puede comprometer la calidad y la integridad del procedimiento. A nivel de registro fotográfico, una captura de imágenes exitosa exige una insuflación y control del equipo adecuados. Adicionalmente, este registro puede verse influenciado por el movimiento del paciente y la variabilidad anatómica. Por último, una barrera es el almacenamiento y el análisis de imágenes, que es una frecuente desventaja de los equipos endoscópicos⁽¹⁷⁾.

Actualmente, se requieren estudios bien diseñados que examinen la relación entre la cantidad de imágenes, un protocolo de fotodocumentación específico y la tasa de detección de neoplasias durante la EGD⁽³²⁾. No obstante, un consenso asiático sobre estándares en endoscopia digestiva alta sugiere que una observación metódica del tracto gastrointestinal superior puede mejorar la detección de neoplasias superficiales al minimizar las zonas sin visualizar. Asimismo, el panel de expertos concluyó que la evidencia actual apoya el uso de simeticona para potenciar la claridad visual durante el examen⁽³³⁾.

En este escenario, un segundo lector (inteligencia artificial) no solo debe reconocer las regiones anatómicas gástricas de interés sino también asegurarse de que se sigan durante un tiempo mínimo. La inspección exhaustiva de toda la cavidad

gástrica constituye un paso fundamental en el proceso de búsqueda sistemática del cáncer gástrico temprano. Por lo anterior, hay interés de muchos grupos a nivel mundial de aplicar la inteligencia artificial en esta área, y algunos investigadores han clasificado cuadros utilizando arquitecturas de redes neuronales convolucionales (CNN)⁽³⁴⁾. Por ejemplo, Takiyama y colaboradores⁽³⁵⁾ utilizaron una arquitectura GoogleNet para reconocer con precisión cuatro ubicaciones anatómicas (laringe, esófago, estómago y duodeno), así como tres subclasificaciones posteriores específicas para imágenes del estómago. Wu y colaboradores⁽³⁶⁾ emplearon una red VGG-16 para clasificar las ubicaciones gástricas en diez categorías y luego refinar la clasificación en 26 partes anatómicas (22 para el estómago, 2 para el esófago y 2 para el duodeno). Además, Chang y colaboradores⁽³⁷⁾ entrenaron una arquitectura ResNeSt para clasificar imágenes de EGD en ocho ubicaciones anatómicas, con una ubicación adicional específicamente para la faringe. Sin embargo, la mayoría de las arquitecturas propuestas tienen una cantidad de parámetros que dificultan la resolución de la tarea en tiempo real, por lo cual es necesario entrenar arquitecturas de menor tamaño. En un estudio previo⁽²⁵⁾, se comparó el rendimiento de 22 arquitecturas para identificar seis regiones anatómicas del estómago, incluidas arquitecturas del estado del arte, y se obtuvo el mejor rendimiento con la ConvNext Tiny. Dicha red tiene alrededor de 27 millones de parámetros, lo cual permitiría implementarla en un escenario clínico.

En el presente estudio, hemos desarrollado un sistema de auditoría automática del protocolo de exploración endoscópica del estómago con inteligencia artificial que permite la identificación de áreas gástricas, y se logró una sensibilidad macro del 85,5% en la interpretación de las 13 áreas específicas previamente descritas. Estos resultados demuestran la potencial eficacia de la inteligencia artificial como una herramienta valiosa para asistir a los endoscopistas en la realización de una exploración completa y minuciosa de la totalidad del estómago.

Diversos expertos y organizaciones han propuesto directrices con respecto a la cantidad de imágenes que deben ser capturadas durante una EGD. Una cantidad apropiada de imágenes incentiva a los endoscopistas a llevar a cabo el procedimiento con mayor atención, y esta información puede utilizarse para entrenar, validar y mejorar el rendimiento de las estrategias de inteligencia artificial. Este factor es esencial para superar el desafío que impide que estas tecnologías aborden eficazmente este problema clínico.

Una de las limitaciones más grandes de este tipo de estudio es la falta de un marco común de validación, problema frecuente en el análisis de imágenes endoscópicas y que ha limitado la comparación entre los enfoques existentes. Es difícil determinar cuáles de estos trabajos podrían tener una ventaja real en el uso clínico.

En resumen, esta propuesta es un sistema computarizado de monitorización de calidad que identifica 13 regiones gástricas anatómicas (las cuales cubren la totalidad de la cavidad gástrica) con una red neuronal convolucional. El rendimiento de esta metodología podría mejorarse y aplicarse a más regiones anatómicas con el aumento de casos anotados. Se espera que este sistema se convierta en una herramienta valiosa para el reconocimiento de regiones del estómago, alertando sobre exploraciones incompletas del tracto gastrointestinal y, en consecuencia, asegurando la calidad del procedimiento. Actualmente, se están realizando estudios adicionales para validar y mejorar aún más la eficacia de este sistema.

Finalmente, este estudio de referencia, que para nuestro conocimiento es el primero en Latinoamérica, puede

orientar investigaciones futuras sobre la aplicación de la clasificación automática sobre las 22 regiones gástricas recomendadas por K. Yao⁽⁸⁾ u otros protocolos de fotodocumentación de endoscopia superior como los presentados en otros trabajos^(14,15,17).

Agradecimientos

Este trabajo fue apoyado en parte por el proyecto con código 110192092354 y titulado *Programa de Detección Temprana de Lesiones Premalignas y Cáncer Gástrico en zonas urbanas, rurales y dispersas del Departamento de Nariño*, de la convocatoria n.º 920 de 2022 del Ministerio de Ciencias.

REFERENCIAS

1. Ferlay J, Ervik M, Lam F, Laversanne M, Colombet M, Mery L, et al. Global Cancer Observatory: Cancer Today [Internet]. Lyon, France: International Agency for Research on Cancer; 2024 [consultado el 1 de abril de 2024]. Disponible en: <https://gco.iarc.who.int/today>, accessed [01 04 2024].
2. Oliveros-Wilches R, Grillo-Ardila CF, Vallejo-Ortega M, Gil-Parada F, Cardona-Tobón M, Páramo-Hernández D, et al. Guía de práctica clínica para la prevención primaria y secundaria y diagnóstico temprano de cáncer gástrico. *Rev Colomb Cancerol*. 2022;26(1):39-96. <https://doi.org/10.35509/01239015.754>
3. Oliveros R, Navarerra LF. Diagnóstico, estadificación y tratamiento del cáncer gástrico en Colombia desde 2004 a 2008 (REGATE -Colombia). *Rev Colomb Gastroenterol*. 2012;27(4):269-74.
4. Pasechnikov V, Chukov S, Fedorov E, Kikuste I, Leja M. Gastric cancer: Prevention, screening and early diagnosis. *World J Gastroenterol WJG*. 2014;20(38):13842-62. <https://doi.org/10.3748/wjg.v20.i38.13842>
5. Kaise M. Advanced endoscopic imaging for early gastric cancer. *Best Pract Res Clin Gastroenterol*. 2015;29(4):575-87. <https://doi.org/10.1016/j.bpg.2015.05.010>
6. Pimenta-Melo AR, Monteiro-Soares M, Libânio D, Dinis-Ribeiro M. Missing rate for gastric cancer during upper gastrointestinal endoscopy: a systematic review and meta-analysis. *Eur J Gastroenterol Hepatol*. 2016;28(9):1041-9. <https://doi.org/10.1097/MEG.0000000000000657>
7. Lee SH, Park YK, Cho SM, Kang JK, Lee DJ. Technical skills and training of upper gastrointestinal endoscopy for new beginners. *World J Gastroenterol WJG*. 2015;21(3):759-85. <https://doi.org/10.3748/wjg.v21.i3.759>
8. Yao K. The endoscopic diagnosis of early gastric cancer. *Ann Gastroenterol Q Publ Hell Soc Gastroenterol*. 2013;26(1):11.
9. Yao K, Uedo N, Muto M, Ishikawa H. Development of an e-learning system for teaching endoscopists how to diagnose early gastric cancer: basic principles for improving early detection. *Gastric Cancer*. 2017;20(Suppl 1):28-38. <https://doi.org/10.1007/s10120-016-0680-7>
10. Kim JS, Kim BW. Training in Endoscopy: Esophagogastroduodenoscopy. *Clin Endosc*. 2017;50(4):318-21. <https://doi.org/10.5946/ce.2017.096>
11. Eisen GM, Baron TH, Dominitz JA, Faigel DO, Goldstein JL, Johanson JF, et al. Methods of granting hospital privileges to perform gastrointestinal endoscopy. *Gastrointest Endosc*. 2002;55(7):780-3. [https://doi.org/10.1016/S0016-5107\(02\)70403-3](https://doi.org/10.1016/S0016-5107(02)70403-3)
12. Farthing MJ, Walt RP, Allan RN, Swan CH, Gilmore IT, Mallinson CN, et al. A national training programme for gastroenterology and hepatology. *Gut*. 1996;38(3):459-70. <https://doi.org/10.1136/gut.38.3.459>
13. Beattie AD, Greff M, Lamy V, Mallinson CN. The European Diploma of Gastroenterology: progress towards harmonization of standards. *Eur J Gastroenterol Hepatol*. 1996;8(4):403-6. <https://doi.org/10.1097/00042737-199604000-00021>
14. Beg S, Rangunath K, Wyman A, Banks M, Trudgill N, Pritchard DM, et al. Quality standards in upper gastrointestinal endoscopy: a position statement of the British Society of Gastroenterology (BSG) and Association of Upper Gastrointestinal Surgeons of Great Britain and Ireland (AUGIS). *Gut*. 2017;66(11):1886-99. <https://doi.org/10.1136/gutjnl-2017-314109>
15. Bisschops R, Areia M, Coron E, Dobru D, Kaskas B, Kuvaev R, et al. Performance measures for upper gastrointestinal endoscopy: A European Society of Gastrointestinal Endoscopy quality improvement initiative. *United Eur Gastroenterol J*. 2016;4(5):629-56. <https://doi.org/10.1177/2050640616664843>

16. Yao K, Uedo N, Kamada T, Hirasawa T, Nagahama T, Yoshinaga S, et al. Guidelines for endoscopic diagnosis of early gastric cancer. *Dig Endosc Off Jpn Gastroenterol Endosc Soc.* 2020;32(5):663-98. <https://doi.org/10.1111/den.13684>
17. Emura F, Sharma P, Arantes V, Cerisoli C, Parra-Blanco A, Sumiyama K, et al. Principles and practice to facilitate complete photodocumentation of the upper gastrointestinal tract: World Endoscopy Organization position statement. *Dig Endosc.* 2020;32(2):168-79. <https://doi.org/10.1111/den.13530>
18. Kim GH, Bang SJ, Ende AR, Hwang JH. Is screening and surveillance for early detection of gastric cancer needed in Korean Americans? *Korean J Intern Med.* 2015;30(6):747-58. <https://doi.org/10.3904/kjim.2015.30.6.747>
19. Asaka M, Mabe K. Strategies for eliminating death from gastric cancer in Japan. *Proc Jpn Acad Ser B Phys Biol Sci.* 2014;90(7):251-8. <https://doi.org/10.2183/pjab.90.251>
20. Hamashima C, Systematic Review Group and Guideline Development Group for Gastric Cancer Screening Guidelines. Update version of the Japanese Guidelines for Gastric Cancer Screening. *Jpn J Clin Oncol.* 2018;48(7):673-83. <https://doi.org/10.1093/jjco/hyy077>
21. Yashima K, Shabana M, Kurumi H, Kawaguchi K, Isomoto H. Gastric Cancer Screening in Japan: A Narrative Review. *J Clin Med.* 2022;11(15):4337. <https://doi.org/10.3390/jcm11154337>
22. Goss PE, Lee BL, Badovinac-Crnjevic T, Strasser-Weippel K, Chavarri-Guerra Y, St Louis J, et al. Planning cancer control in Latin America and the Caribbean. *Lancet Oncol.* 2013;14(5):391-436. [https://doi.org/10.1016/S1470-2045\(13\)70048-2](https://doi.org/10.1016/S1470-2045(13)70048-2)
23. Bravo LE, Hernández Vargas JA, Collazos P, García LS, Valbuena AM, Acuña L. Survival in stomach cancer: analysis of a national cancer information system and a population-based cancer registry in Colombia. *Colomb Médica CM.* 2023;53(4):e2025126. <https://doi.org/10.25100/cm.v53i4.5126>
24. Liu Z, Mao H, Wu CY, Feichtenhofer C, Darrell T, Xie S. A ConvNet for the 2020s [Internet]. arXiv; 2022 [consultado el 15 de septiembre de 2023]. Disponible en: <http://arxiv.org/abs/2201.03545>. <https://doi.org/10.1109/CVPR52688.2022.01167>
25. Bravo D, Ruano J, Jaramillo M, Gallego D, Gómez M, González FA, et al. Automatic Classification of Esophagogastroduodenoscopy Sub-Anatomical Regions. En: 2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI). 2023. p. 1-5. <https://doi.org/10.1109/ISBI53787.2023.10230483>
26. Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. ImageNet: A large-scale hierarchical image database. En: 2009 IEEE Conference on Computer Vision and Pattern Recognition. 2009. p. 248-55. <https://doi.org/10.1109/CVPR.2009.5206848>
27. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, et al. ImageNet Large Scale Visual Recognition Challenge [Internet]. arXiv; 2015 [consultado el 15 de septiembre de 2023]. Disponible en: <http://arxiv.org/abs/1409.0575>
28. Serre T, Kreiman G, Kouh M, Cadieu C, Knoblich U, Poggio T. A quantitative theory of immediate visual recognition. *Prog Brain Res.* 2007;165:33-56. [https://doi.org/10.1016/S0079-6123\(06\)65004-8](https://doi.org/10.1016/S0079-6123(06)65004-8)
29. Lee H, Ekanadham C, Ng A. Sparse deep belief net model for visual area V2 [Internet]. Advances in Neural Information Processing Systems; 2007 [consultado el 24 de septiembre de 2023]. Disponible en: https://papers.nips.cc/paper_files/paper/2007/file/4daa3db355ef2b0e64b472968cb70f0d-Paper.pdf
30. Glorot X, Bordes A, Bengio Y. Deep Sparse Rectifier Neural Networks. International Conference on Artificial Intelligence and Statistics; 2011 [consultado el 24 de septiembre de 2023]. Disponible en: <https://www.semanticscholar.org/paper/Deep-Sparse-Rectifier-Neural-Networks-Glorot-Bordes/67107f78a84bdb2411053cb54e94fa226eea6d8e>
31. Hamashima C, Ogoshi K, Narisawa R, Kishi T, Kato T, Fujita K, et al. Impact of endoscopic screening on mortality reduction from gastric cancer. *World J Gastroenterol WJG.* 2015;21(8):2460-6. <https://doi.org/10.3748/wjg.v21.i8.2460>
32. Kim SY, Park JM. Quality indicators in esophagogastroduodenoscopy. *Clin Endosc.* 2022;55(3):319-31. <https://doi.org/10.5946/ce.2022.094>
33. Chiu PWY, Uedo N, Singh R, Gotoda T, Ng EKW, Yao K, et al. An Asian consensus on standards of diagnostic upper endoscopy for neoplasia. *Gut.* 2019;68(2):186-97. <https://doi.org/10.1136/gutjnl-2018-317111>
34. Renna F, Martins M, Neto A, Cunha A, Libânio D, Dinis-Ribeiro M, et al. Artificial Intelligence for Upper Gastrointestinal Endoscopy: A Roadmap from Technology Development to Clinical Practice. *Diagn Basel Switz.* 2022;12(5):1278. <https://doi.org/10.3390/diagnostics12051278>
35. 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(1):7497. <https://doi.org/10.1038/s41598-018-25842-6>
36. Wu L, Zhou W, Wan X, Zhang J, Shen L, Hu S, et al. A deep neural network improves endoscopic detection of early gastric cancer without blind spots. *Endoscopy.* 2019;51(6):522-31. <https://doi.org/10.1055/a-0855-3532>
37. Chang YY, Li PC, Chang RF, Yao CD, Chen YY, Chang WY, et al. Deep learning-based endoscopic anatomy classification: an accelerated approach for data preparation and model validation. *Surg Endosc.* 2022;36(6):3811-21. <https://doi.org/10.1007/s00464-021-08698-2>