



<https://www.elsevier.es/gastroenterologia>

DISEÑO Y VALIDACIÓN DE UN SISTEMA DE INTELIGENCIA ARTIFICIAL PARA DETECTAR LA CALIDAD DE LA LIMPIEZA DE COLON ANTES DE LA COLONOSCOPIA

Federica Benítez Zafra¹, Antonio Z. Gimeno-García¹, Silvia Alayón-Miranda², Domingo Hernández-Negrín³, David Nicolás-Pérez¹, José Luis Baute Dorta³, Álvaro Peralta³, Rocío del Castillo³, Ana Romero³, Melissa Piñero³, Gloria Lorenzo³, Alejandro Jiménez⁴, Marco A. Navarro-Dávila⁵ y Manuel Hernández-Guerra¹

¹Servicio de Aparato Digestivo, Hospital Universitario de Canarias, Tenerife. ²Departamento de Ingeniería Informática y de Sistemas, Universidad de La Laguna. ³Servicio de Aparato Digestivo, Sección de Enfermería de Endoscopias, Hospital Universitario de Canarias, Tenerife. ⁴Departamento de Estadística, Hospital Universitario de Canarias, Tenerife. ⁵Servicio de Farmacia Hospitalaria, Hospital Universitario de Canarias, Tenerife.

Resumen

Introducción y objetivos: La percepción de los pacientes de su calidad de limpieza ha demostrado ser un potente predictor de la limpieza intestinal evaluada durante la colonoscopia y podría guiar estrategias de limpieza de rescate. El objetivo principal de este estudio fue diseñar, entrenar y validar una Red Neuronal Convolucional (CNN) para clasificar el efluente rectal durante la ingesta de preparación intestinal en una preparación de calidad “adecuada” o “inadecuada”.

Métodos: Se solicitó a los pacientes remitidos a una colonoscopia ambulatoria que proporcionaran imágenes de sus efluentes rectales durante el proceso de preparación intestinal. Las imágenes fueron categorizadas por 3 observadores según una escala de calidad predefinida con 4 imágenes. Se diseñó, entrenó y validó una CNN utilizando las imágenes proporcionadas por los pacientes. Se incluyeron en el estudio un total de 660 pacientes y 1.203 imágenes, distribuidos en una cohorte inicial de 400 pacientes (799 imágenes: 360 etiquetadas como “adecuadas” y 439 como “inadecuadas”) y en una segunda cohorte de 260 pacientes (404 imágenes: 151 “adecuadas” y 253 “inadecuadas”). El conjunto de datos inicial se utilizó para el entrenamiento de la CNN. Los datos de la segunda cohorte de pacientes se utilizaron para realizar un testeo de la precisión de la red.

Resultados: La precisión global del modelo CNN sobre el conjunto de datos inicial fue del 97,49%, con una sensibilidad del 98,17% y una especificidad del 96,66%. La precisión global sobre el segundo conjunto de datos fue del 95%, con una sensibilidad del 99,60% y una especificidad del 87,41%.

Conclusiones: Se ha obtenido una CNN capaz de clasificar imágenes de “limpieza adecuada” y “limpieza inadecuada” con una elevada precisión.