

Segmentando imágenes gastrointestinales usando ensamble ponderado U-NET++ y 2D-HMM

Jairo Enrique Ramírez Sánchez¹, Pedro Martínez Barrón²,
Hannia Medina Aguilar², Romeo Sánchez Nigenda²

¹ Tecnológico de Monterrey,
México

² Universidad Autónoma de Nuevo León,
México

{alfonso.martinezbrn,hannia.medinaglr,
romeo.sanchezng}@uanl.edu.mx, A01750443@tec.mx

Resumen. Uno de los tratamientos más utilizados para el cáncer del tracto gastrointestinal (GI) es la radioterapia que requiere de la segmentación manual de los órganos afectados para el suministro de radiación sin afectar células sanas. Para automatizar este proceso, se han utilizado técnicas de aprendizaje profundo, especialmente variantes de U-Net. Sin embargo, la segmentación de los órganos del tracto GI permanece desafiante por la alta capacidad que tienen de deformarse por el movimiento corporal y la función respiratoria. Este trabajo propone una metodología que desarrolla un ensamble ponderado integrando modelos de U-Net++ y Modelos Ocultos de Markov (2D-HMM) para una segmentación semántica del estómago y los intestinos. Los experimentos reportan una precisión de 0.811 del coeficiente de Dice usando Leave-One-Out Cross-Validation, otorgando robustez a los resultados.

Palabras clave: Segmentación de imágenes, arquitectura U-NET, aprendizaje máquina, modelos ocultos de Markov.

Gastrointestinal Image Segmentation Using a Weighted U-NET++ and 2D-HMM Ensemble

Abstract. One of the most widely used treatments for cancer of the gastrointestinal tract is radiotherapy, which requires manual segmentation of the affected organs to deliver radiation without affecting healthy cells. To automate this process, deep learning techniques have been used, especially variants of U-Net. However, the segmentation of the GI tract organs remains challenging, due to their high capacity to deform due to body movement and respiratory function. This work proposes a methodology that develops a weighted ensemble integrating U-Net++ models and Hidden Markov Models (2D-HMM) for semantic segmentation of the stomach and intestines. From our experiments, we obtained a precision of 0.811 for Dice coefficient by means of Leave-One-Out Cross-Validation, which provides robustness to the results.

Keywords: Image segmentation, U-NET architecture, machine learning, hidden Markov models.

1. Introducción

A nivel mundial, en el 2018, alrededor de 4.8 millones de personas fueron diagnosticadas con cáncer del tracto gastrointestinal, representando el 26 % de la incidencia mundial de cáncer. Las proyecciones basadas en las tendencias actuales predicen un incremento del 58 % a 7.5 millones para 2040 [1]. La mitad de estos pacientes son elegibles para radioterapia [13].

Durante la radioterapia, un acelerador lineal médico (LINAC) administra altas dosis de radiación a las células cancerosas para matarlas, y con cierta probabilidad daña al mismo tiempo a las células sanas cercanas. El daño a las células sanas provoca efectos secundarios asociados con este tratamiento, como pérdida de audición, vómitos, y cansancio extremo, entre otros efectos [18]. Para disminuir los daños colaterales, los oncólogos tratan de dirigir los rayos X hacia los tumores evitando los órganos en riesgo.

El Acelerador Lineal Guiado por Imágenes de Resonancia Magnética (MR-Linac), permite observar tumores y órganos en tiempo real para ajustar la dirección de la radiación; sin embargo, los oncólogos deben segmentar manualmente los órganos, extendiendo las sesiones de tratamiento hasta en una hora, tiempo en el que el paciente debe permanecer inmóvil.

En los últimos años, técnicas de Inteligencia Artificial como las redes neuronales convolucionales han sido capaces de realizar auto-segmentación en casos de tumores cerebrales [6], cáncer de cuello [11] y cáncer de próstata [9, 8], reduciendo a la mitad el tiempo de las sesiones de tratamiento [3]; no obstante, hay pocos avances en la segmentación de los órganos del tracto gastrointestinal (GI), principalmente porque los órganos abdominales están rodeados de tejido blando y pueden variar en forma y ubicación a lo largo del día debido a los movimientos digestivos y respiratorios [10].

En este trabajo proponemos una metodología basada en aprendizaje profundo para el pre-procesamiento y segmentación de imágenes de resonancia magnética del tracto digestivo. La arquitectura de nuestro enfoque es un ensamble ponderado basado en modelos de U-Net y Modelos Ocultos de Markov en dos dimensiones (2D-HMM) que realiza una segmentación semántica del estómago y los intestinos delgado y grueso.

La metodología propuesta tiene el potencial de ayudar a implementar tratamientos más efectivos y eficientes para los pacientes al acelerar el proceso de segmentación. La metodología presentada se evaluó sobre el conjunto de imágenes de la UW-Madison Carbone Center, proporcionadas de manera pública en la plataforma Kaggle como parte de la competencia UW-Madison GI Tract Image Segmentation³ sin poner en riesgo el tiempo de ejecución y los requerimientos de espacio en memoria del proceso de segmentación.

El presente trabajo está organizado de la siguiente manera. En la siguiente sección presentamos revisión de la literatura. La sección 3 describe la metodología propuesta ilustrando las diferentes etapas del proceso. La sección 4 discute los resultados obtenidos de los modelos generados. En la última sección presentamos nuestras conclusiones.

³ kaggle.com/competitions/uw-madison-gi-tract-image-segmentation

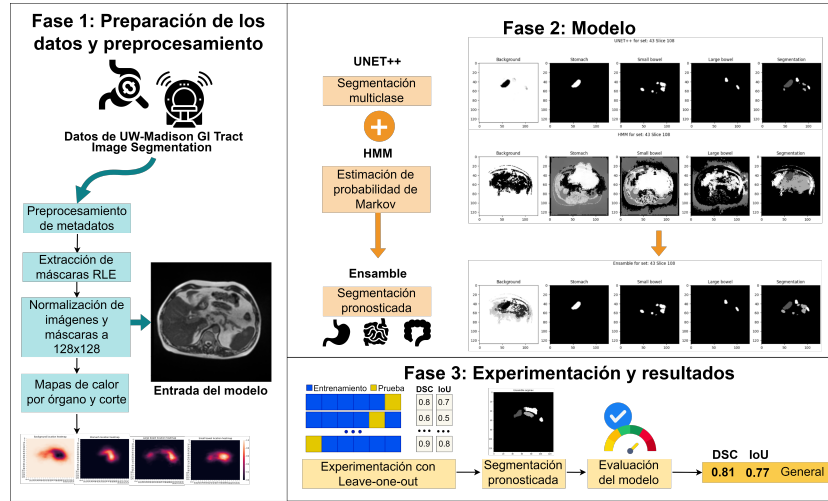


Fig. 1. Metodología para el Diseño y Validación de Modelos de Segmentación.

2. Trabajo relacionado

Estudios recientes en el área médica han empleado técnicas de aprendizaje profundo de Inteligencia Artificial para asistir en la segmentación automática de imágenes médicas en procesos de diagnóstico y tratamiento [16]. En particular, se han considerado variantes de arquitecturas U-Net para la segmentación de imágenes biomédicas.

En este tipo de arquitecturas se consideran modelos de aprendizaje profundo con buen rendimiento en la auto-segmentación de imágenes debido a que tienen la capacidad de combinar simultáneamente información de alto y bajo nivel para extraer características complejas.

Sin embargo, la segmentación de los órganos del tracto GI sigue siendo una tarea desafiante [7], ya que estos órganos tienen una alta capacidad de deformarse por el movimiento corporal y la función respiratoria del individuo.

Debido a lo anterior, existen pocos estudios sobre el uso exitoso y amplio de MR-Linac para casos de cáncer de estómago [21], y sobre la aplicación de arquitecturas U-Net para éste tipo de imágenes, la mayoría de los estudios se basan en modelos complejos como 3D U-Net.

En 2020, [12] propusieron U-Net para segmentar el hígado, el estómago, el duodeno y el riñón en imágenes de tomografía computarizada (TC) basadas en parches 3D. Sus resultados fueron prometedores para el estómago, alcanzando un 0.813 para el Coeficiente de Similitud de Dice (DSC), pero menos significativos para el duodeno donde obtuvieron 0.595.

[19] propuso un enfoque similar para segmentar los órganos del tracto GI en el 2022. En un reporte preliminar compara el rendimiento de diferentes codificadores para una arquitectura U-Net clásica, siendo el codificador Resnet34 el que reporta mejores resultados.

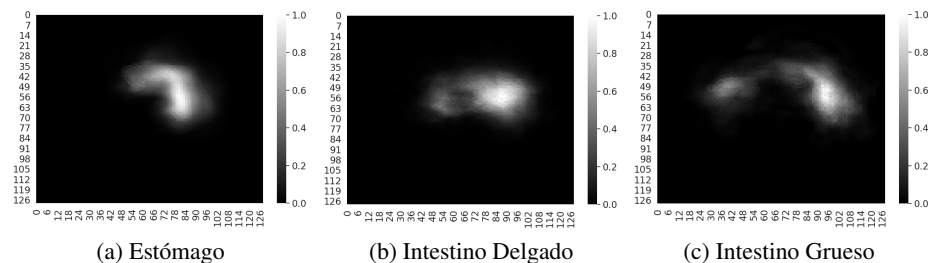


Fig. 2. Mapas de calor para cada órgano.

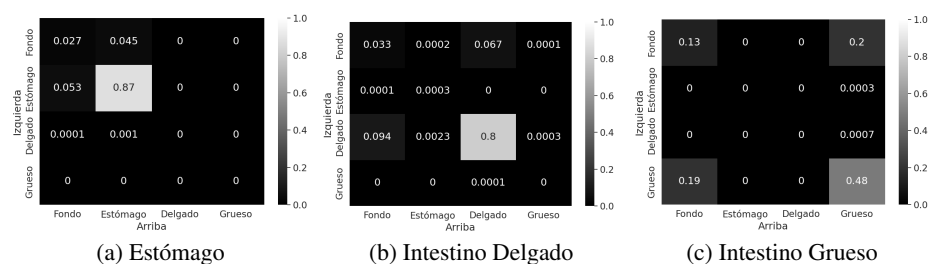


Fig. 3. Matrices de transición.

El trabajo por [5] presenta una U-Net y métodos de Redes Neuronales Convolucionales por Regiones (Mask R-CNNs) para realizar la segmentación de los órganos del tracto GI en el mismo conjunto de datos de UW-Madison que utilizamos en este trabajo. Los autores reportan que su modelo Mask R-CNN obtuvo una puntuación del DSC de 0.73 en sus datos de validación.

Otros trabajos utilizan transformadores de visión (Vision Transformers) para segmentar, de igual manera, las imágenes de UW-Madison [15]. La solución que se propone es híbrida ya que utilizan en su modelo una arquitectura LeViT como el codificador, y una U-Net++ como el decodificador. Los resultados de su modelo alcanzan una puntuación de 0.79 para DSC y de 0.72 para IoU.

En [7], se describe un método de refinamiento automático de contornos (ACR) basado en mapas de probabilidad para corregir contornos auto-segmentados en radioterapia guiada por resonancia magnética. La auto-segmentación fue generada por una arquitectura CNN profunda en 3D (una 3D-ResUNet modificada), el DSC cambió de 0.44 a 0.56, de 0.33 a 0.55, y de 0.34 a 0.54, en el estómago, intestino delgado e intestino grueso, respectivamente.

[17] desarrollaron un método basado en Modelos Ocultos de Markov (HMM) llamado Complete Enumeration Propagation para la segmentación de imágenes multi-clase, en el cual los estados ocultos de un modelo Markoviano representan la verdadera segmentación de la imagen.

[2] utilizaron modelos de Markov bidimensionales (2D-HMM) para la segmentación efectiva de radiografías, imágenes multiespectrales y sintéticas. A pesar del potencial de HMMs, no existen estudios comprensivos de su aplicación en la segmentación de imágenes de resonancias magnéticas.

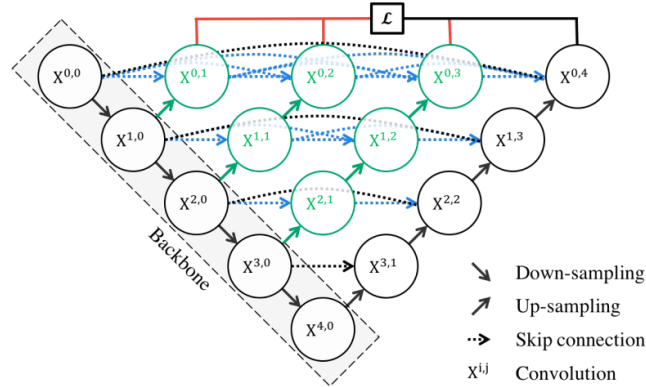


Fig. 4. Arquitectura original de U-Net++ de [22].

Existen trabajos recientes en la literatura que combinan el uso de operadores convolucionales con HMMs adaptativos para segmentar imágenes cerebrales [14, 20]. Sin embargo, hasta donde conocemos, no existe un método que incorpore HMMs en la segmentación de imágenes del tracto GI como proponemos en este trabajo.

En resumen, los enfoques de aprendizaje profundo, especialmente variantes de U-Net, son los métodos más explorados en la literatura para analizar imágenes biomédicas [16]. La aplicación de estos métodos para segmentar imágenes del tracto gastrointestinal sigue siendo un desafío y una área de investigación abierta.

3. Metodología

En esta sección presentamos la metodología propuesta la cual consta de tres fases. La primera fase incluye pre-procesamiento de las imágenes del conjunto de datos (3.1), la segunda es el diseño y construcción de los modelos de segmentación (3.2), y por último, la fase de validación de los modelos a través de experimentación y análisis de resultados (4). En la Fig.1 se pueden observar las etapas generales de la metodología propuesta.

3.1. Pre-procesamiento de los datos

Como puede observarse en el diagrama de la Fig.1, la primera fase de la metodología consiste en la preparación de los datos. El conjunto de datos, utilizado en esta investigación es público y fue proporcionado por el UW-Madison Carbone Cancer Center.

El repositorio de datos consiste de 272 conjuntos de resonancias magnéticas en formato PNG en escala de grises de 16 bits de 85 pacientes con cáncer durante su tratamiento por radiación. Cada escaneo tiene 144 cortes, lo que da un total de 39,168 imágenes. Las anotaciones de entrenamiento son las máscaras codificadas por RLE (Run-Length Encoding) para la segmentación de tres órganos del tracto GI: el estómago, el intestino grueso y el intestino delgado.

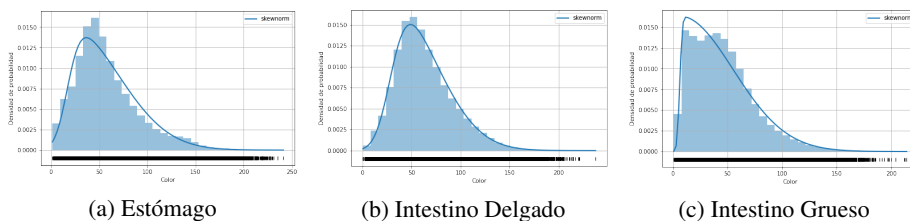


Fig. 5. Distribuciones de probabilidad en función del color del píxel.

Algorithm 1 2D-HMM Algoritmo de Segmentación.

```

Calcular  $P_e$ ,  $P_O$  y  $P_T$ 
 $\omega[128][128] : float$  ▷ Imagen a procesar
 $M[128][128][4] : float$  ▷ Matriz de segmentación predicha
 $S \leftarrow [0, 1, 2, 3]$  ▷ Clasificaciones de los estados
for  $i, j \in \omega$  do
  for  $l \in S$  do
     $O_{i,j} = \omega[i][j]$ 
     $P_{i,j} = P_O(l|O_{i,j}) \cdot P_e(l|i, j)$ 
     $temp[4][4] : float$ 
    for  $m, n \in S$  do
       $temp[m][n] = P_{i,j} \cdot P_T(l|m, n)$ 
     $M[i][j][l] = max(temp)$ 
return  $arg \max_l M$  ▷ Segmentación final

```

Las imágenes son de diferentes dimensiones por lo que fue necesario redimensionarlas a un tamaño estandarizado. En consecuencia, se decidió normalizar todas las imágenes y sus respectivas máscaras codificadas en RLE a un tamaño de 128 x 128 px. Para visualizar el patrón en la distribución de los órganos en la muestra, trazamos el mapa de calor para cada órgano que se presenta en Fig.2.

3.2. Diseño y construcción de los modelos de segmentación

Como se mencionó anteriormente, nuestra propuesta metodológica incluye la creación de 2 modelos para la segmentación de los órganos, y un ensamble que integra ambos modelos. El primer modelo se basa en una arquitectura de tipo U-Net++, y el segundo modelo está basado en HMMs bidimensionales (2D-HMM). A continuación se describen los procesos individuales para la construcción y entrenamiento de ambos modelos, así como el proceso de su integración para el ensamble.

Modelo U-Net++. Dicha arquitectura fue diseñada para resolver limitaciones del modelo U-Net base en la segmentación de imágenes médicas [22] al incluir una serie de conexiones adicionales a la U-Net original para la recuperación efectiva de los detalles de granularidad fina de los objetos, incluyendo supervisión profunda que permite establecer diferentes configuraciones de sus parámetros.

Tabla 1. Valores de Ponderación α de ensamble 2D-HMM . U-Net++ evaluados.

Métrica	Ponderación α					
	0.05	0.10	0.20	0.30	0.40	
Dice	General	0.811	0.799	0.788	0.771	0.742
	Estómago	0.888	0.885	0.872	0.844	0.749
	Intestino Delgado	0.812	0.791	0.759	0.701	0.601
	Intestino Grueso	0.817	0.814	0.804	0.786	0.747
	IoU	0.777	0.770	0.748	0.709	0.628

Tabla 2. Resultados de la experimentación con los modelos propuestos.

Métrica	Modelos				
	2D-HMM . U-Net++	2D-HMM . U-Net	U-Net++	U-Net	
Dice	General	0.811 (32 %)	0.723 (34 %)	0.610	0.538
	Estómago	0.888 (10 %)	0.803 (26 %)	0.808	0.635
	Intestino Delgado	0.812 (38 %)	0.711 (29 %)	0.585	0.548
	Intestino Grueso	0.817 (5.6 %)	0.774 (43 %)	0.773	0.538
	IoU	0.777 (18 %)	0.696 (36 %)	0.657	0.511

Las conexiones adicionales de la U-Net++ siguen una regla piramidal, donde la forma de **U** se llena con bloques convolucionales, cada uno de los cuales consta de un cierto número de capas que varía según los nodos de la red. El diagrama original de la U-Net++ tomado de [22] se muestra en la figura 4.

En este trabajo la red se implementó en Python 3.8 siguiendo la versión propuesta en [22]. Los hiper-parámetros del modelo fueron ajustados con la búsqueda por cuadrícula propia de la API de Keras, seleccionando relu como función de activación en las capas ocultas, 0.1 como dropout rate, 5×10^{-4} como learning rate durante 50 épocas y Adam como optimizador.

Finalmente, se utilizó sigmoid como función de activación en la última capa en lugar de softmax con la finalidad de asignar probabilidades a cada clase en lugar de distribuirlas. Para la determinación de los hiper-parámetros se realizó una partición del 80 % del total de imágenes para entrenamiento y 20 % para validación.

Como función de pérdida se optimizó el coeficiente de DICE por órgano, integrándose por medio de una suma ponderada debido al desbalanceo de clases. Sea $y \in \mathbb{R}^{128 \times 128 \times 4}$ la matriz de segmentación real, $\hat{y} \in \mathbb{R}^{128 \times 128 \times 4}$ la segmentación predicha por la red, sea $S \in \{\text{estómago, intestino delgado, intestino grueso, fondo}\}$ el conjunto de estados de la clasificación. Con lo cual $\hat{y}_l \in \mathbb{R}^{128 \times 128}$ hace referencia a la segmentación correspondiente al órgano $l \in S$. Finalmente, sea α_l la frecuencia inversa de la clase del órgano l . La ecuación 1 muestra matemáticamente el proceso:

$$\mathcal{L}(y, \hat{y}) = \sum_{l \in S} \alpha_l \frac{2|y_l \cap \hat{y}_l|}{|y_l| + |\hat{y}_l|}. \quad (1)$$

Modelo bidimensional de Markov (2D-HMM). Los Modelos Ocultos de Markov son una técnica estadística que permite crear un modelo con eventos observados y ocultos como factores causales en un modelo probabilístico.

Tabla 3. Comparación con Modelos Recientes de Segmentación del Tracto GI Estos modelos utilizan datos diferentes a los usados por este trabajo en su evaluación.

Métrica	Modelos en la Literatura						
	2D-HMM U-Net U-Net++	Mask [5]	Resnet34 R-CNN[5]	LeViT384- [19]	3D-ResUnet UNet++ [15]	3D U-Net [7]*	3D U-Net [12]*
General	0.811	0.51	0.73			0.79	
Dice	Estómago	0.888			0.813	0.77	0.813
	Int Delgado	0.812				0.75	
	Int Grueso	0.817				0.76	
IoU	0.777			0.852		0.728	

Un HMM consta de dos procesos estocásticos, un proceso invisible de estados ocultos y un proceso visible de símbolos observables, donde los estados ocultos forman una cadena de Markov y la distribución de probabilidad del símbolo observado depende de los estados subyacentes.

En el caso de segmentación de imágenes, la intuición es que los píxeles en una imagen presentan dependencia de aquellos aledaños; es decir, comparten características en común como color y ubicación espacial. Debido a esto es posible tratar a esta dependencia como un Markov Random Field, el cual relaciona dos probabilidades principales: de transición (P_T) y de observación (P_O). La intuición indica que los píxeles (i, j) de una imagen se relacionan con sus vecinos.

La probabilidad de transición P_T indica que el estado s al que pertenece un píxel, expresado como $s_{i,j}$, está relacionado con el estado del píxel lateral izquierdo $s_{i-1,j}$ y el superior $s_{i,j-1}$. Es decir, $P_T = P(s_{i,j} = l | s_{i-1,j} = n, s_{i,j-1} = m)$. Sea S el conjunto de estados en los que un píxel puede ser clasificado y sea Ω el conjunto total de imágenes. El cálculo de P_T se expresa en la ecuación 2, siendo $I(\cdot)$ la función contadora que retorna 1 si se cumple la condición y 0 en caso contrario:

$$P_T(l | n, m) = \frac{1}{|\Omega|} \sum_{\omega \in \Omega} \frac{\sum_{i,j} I(s_{i,j} = l, s_{i-1,j} = n, s_{i,j-1} = m)}{\sum_{i,j} I(s_{i,j} = l)} \quad \forall l, m, n \in S. \quad (2)$$

Las probabilidades de transición calculadas se muestran en las matrices de la Fig.3. Por ejemplo, podemos observar en la Fig.?? que si el estado del píxel actual correspondiera al estómago existiría una probabilidad del 4.5 % que el píxel superior fuese el estómago y el izquierdo el fondo de la imagen; mientras que existe una probabilidad de un 87 % que ambos correspondan al mismo órgano.

Por otro lado, la probabilidad de que los píxeles aledaños correspondan a los otros órganos es prácticamente nula. La probabilidad de observación es calculada con base al color de cada píxel ($O_{i,j}$) medido entre 0 y 255. Se realizó un ajuste de una función de distribución de probabilidad skewnorm a los colores de cada uno de los órganos.

Sea $P_O(s_{i,j} = l | O_{i,j})$ la función que toma como entradas el color del píxel i, j y regresa la probabilidad de que pertenezca al estado $l \in S$. Dichas funciones se muestran gráficamente en la figura 5. Como se puede observar, realizar una estimación por máxima verosimilitud sería poco precisa debido a que las probabilidades de observación para el estómago y el intestino delgado son similares.

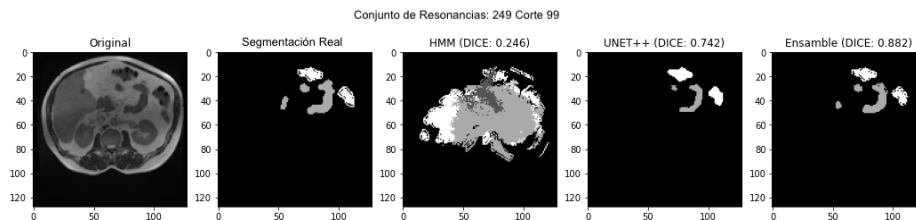


Fig. 6. Ejemplo de segmentación en conjunto 249 para un determinado corte.

Es necesario integrar más información. En el trabajo actual, el estado de un píxel además de estar condicionado por las probabilidades anteriormente descritas, también lo está por la posición espacial en la imagen; es decir, existen zonas de alta probabilidad en las que un órgano puede aparecer como se muestra en los mapas de calor de la Fig.2. Así, la probabilidad P_e de que un píxel (i,j) pertenezca a un estado $s_{i,j} = l$ se calcula integrando las probabilidades antes descritas como se muestra en la ecuación 3:

$$P_e(l|i,j) = \frac{1}{\|\Omega\|} \sum_{\omega \in \Omega} I(s_{i,j} = l). \quad (3)$$

En consecuencia, el cálculo final de la probabilidad de que el estado del píxel (i,j) sea l se muestra en la ecuación 4. Naturalmente, P_e y P_T son calculadas con anterioridad durante la fase de entreno y almacenadas para su consulta. En el caso de P_O se guardan los parámetros de las distribuciones y se calcula su valor:

$$P(s_{i,j} = l) = P_T(l|n,m) \cdot P_O(l|O_{i,j}) \cdot P_e(l|i,j). \quad (4)$$

El cálculo propuesto toma en cuenta el sentido espacial, observacional y de transición. La forma de incorporar los cálculos para la segmentación de una nueva imagen se muestra en el algoritmo 1.

En el presente trabajo, la multiplicación de probabilidades fue sustituida por la suma logarítmica de la probabilidad para evitar un problema de desbordamiento negativo o underflow. Es importante destacar que el propósito del 2D-HMM no es la segmentación por si misma, si no el cálculo de probabilidades eficientes para mejorar el rendimiento de U-Net++. Por este motivo se omite incluir una adaptación del algoritmo de Viterbi en 2D.

Ensamble. Con el propósito de integrar la información de los modelos U-Net++ y 2D-HMM, se propone una capa de ensemble ponderada la cual utiliza las probabilidades otorgadas por ambos modelos para potenciar la clasificación. Sean $H, U \in \mathbb{R}^{128 \times 128 \times 4}$ las matrices de probabilidades calculadas por 2D-HMM y U-Net++ respectivamente. La integración ponderada se muestra por la ecuación 5:

$$E(H, U) = \alpha U + (1 - \alpha)H. \quad (5)$$

El valor de α es seleccionado del espacio de búsqueda $\{0.05, 0.10, 0.20, 0.30, 0.40\}$ que se determinó experimentalmente.

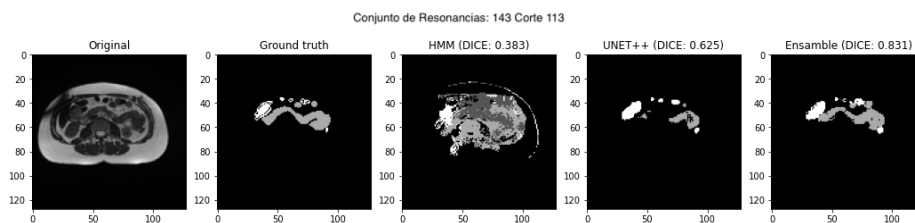


Fig. 7. Ejemplo de segmentación en conjunto 143 para un determinado corte.

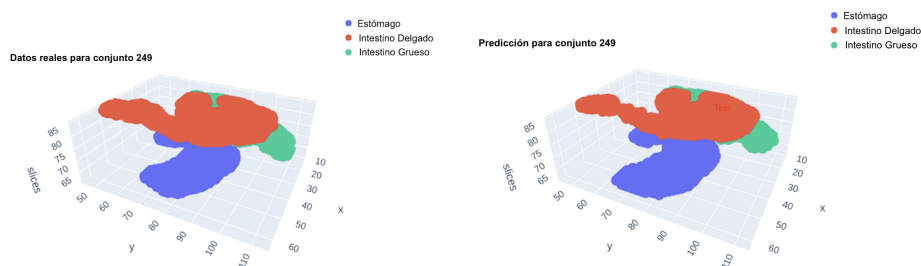


Fig. 8. Ejemplo de segmentación 3D para conjunto 249.

Los resultados de las métricas para cada valor de α se muestran en la siguiente sección. En la segmentación final se aplica máxima verosimilitud sobre el ensamble para obtener el estado l de mayor probabilidad en cada píxel en la matriz.

4. Resultados y análisis

La experimentación se llevó a cabo en la plataforma Colab de Google utilizando una Colabnotebook con Intel(R) Xeon(R) CPU @ 2.20GHz de 6 núcleos, GPU NVIDIA A100-SXM y 12 GB de memoria RAM. Para la fase de entrenamiento y validación de los modelos se siguió el método de Leave-One-Out Cross-Validation que es uno de los métodos recomendados en ciencias biomédicas para mejorar la tasa predictiva de modelos para estudios clínicos [4]. El método consiste en realizar las pruebas del modelo sobre un conjunto de imágenes ω y entrenar tanto los parámetros del 2D-HMM como la U-Net++ con $\Omega \setminus \omega$.

Con esto, el entrenamiento consiste de 271 conjuntos de 39,024 imágenes en total y el de prueba de 1 conjunto de 144 imágenes. Una de las primeras tareas de la fase de evaluación fue ajustar el parámetro de ponderación α del ensamble propuesto, los resultados se muestran en la tabla 1. El caso $\alpha = 0$ haría referencia a la U-Net++ y $\alpha = 1$ al 2D-HMM. Podemos observar que los mejores resultados para el ensamble en las métricas de evaluación se obtienen cuando $\alpha = 0.05$.

Podemos observar en la Tabla2 los resultados de todos los modelos propuestos. Es de resaltar que los modelos U-Net, que incorporan la información del proceso de Markov, reportan mejores resultados en ambas métricas de evaluación satisfaciendo nuestra intuición para su integración.

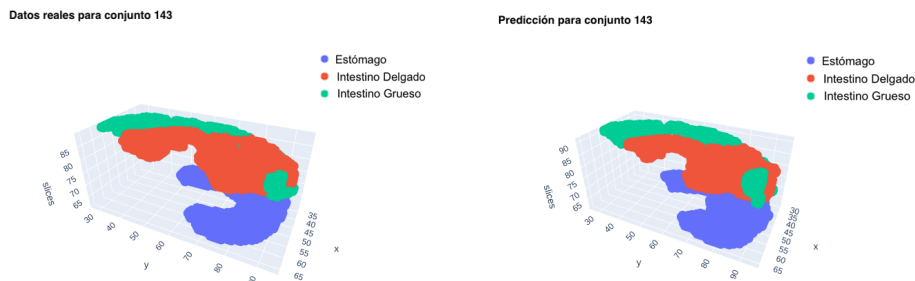


Fig. 9. Ejemplo de segmentación 3D para conjunto 143.

Por ejemplo, en el Dice general para el ensemble de la U-Net++ se tiene un porcentaje de mejora del 32 %, mientras que el ensemble para la U-Net obtiene una mejora del 34 % con respecto a sus modelos individuales.

En el caso del IoU el porcentaje de mejora es del 18 % para el ensemble de la U-Net++ y del 36 % para la U-Net. Finalmente, la Tabla 3 compara los resultados del modelo propuesto 2D-HMM U-Net++ con trabajos recientes en la literatura en el problema de segmentación de imágenes biomédicas del tracto GI. Como lo mencionamos en la sección 1, existen pocos avances en la segmentación de órganos GI debido a su estructura fisiológica.

Podemos observar que nuestro enfoque supera a la mayoría de los trabajos en las métricas de evaluación, a excepción del modelo Resnet34 que solo reporta resultados respecto al IoU; sin embargo, este trabajo siguió una metodología tradicional de particiones 80 - 20 para evaluar, lo cual puede hacer que el resultado sea altamente dependiente de la partición utilizada.

En las figuras 6 - 8 y 7 - 9 se muestran ejemplos visuales 2D y 3D de la segmentación para un corte específico de un conjunto de resonancias. En estos ejemplos el ensemble permite potenciar las predicciones de la U-Net++ hasta en un 19 %. Por ejemplo, la predicción de la U-Net++, ilustrada por la cuarta imagen de la Fig. 7, pierde múltiples detalles de los órganos si se compara con la imagen verdadera.

Sin embargo, el ensemble ponderado es capaz de restaurar estos detalles, como puede observarse en la última columna de la misma figura. En general, podemos observar como el ensemble propuesto incrementa notablemente la calidad de la segmentación.

En resumen, aunque el modelo U-Net++ ha demostrado ser una arquitectura muy eficiente para la segmentación de órganos, como lo demuestra la revisión de la literatura, ésta presenta deficiencias para segmentar ciertas secciones del tracto GI al contener dos o más clases de órganos con una alta probabilidad.

Este trabajo integra las probabilidades del modelo oculto de Markov para discernir mejor aquellos casos donde el modelo base no logra segmentar correctamente, ya que toma en cuenta las probabilidades espaciales y de transición, constituyendo la principal diferencia del trabajo propuesto respecto al trabajo relacionado.

5. Conclusiones

La segmentación de órganos para el tratamiento del cáncer del tracto gastrointestinal es una labor de alta importancia que necesita precisión y rapidez. Es vital contar con algoritmos que puedan auxiliar en la automatización del proceso de segmentación, como apoyo para los especialistas médicos, con el fin de disminuir daños colaterales a células sanas sin incrementar los tiempos de tratamiento.

Sin embargo, la segmentación de los órganos del tracto GI sigue siendo una tarea desafiante, debido a las deformaciones que sufren por el movimiento corporal y la función respiratoria. En este trabajo se propone una metodología, basada en aprendizaje profundo, que desarrolla un ensamble ponderado integrando modelos de U-Net++ y 2D-HMM para una segmentación semántica del estómago y los intestinos. A pesar que el 2D-HMM por sí mismo no otorga una segmentación con gran precisión, permite potenciar las predicciones de la U-Net++ hasta en un 32 % en el Dice general, y un 18 % en el IoU.

La precisión final de 0.811, obtenida por el ensamble en el Dice general, es mejor a los resultados reportados en la literatura. Además, al utilizar el método de evaluación Leave-One-Out, la métrica proporcionada cuenta con un alto nivel de confiabilidad sobre el dataset utilizado. La arquitectura propuesta tiene el potencial de ayudar a implementar tratamientos más efectivos y eficientes para los pacientes con cáncer al acelerar el proceso de segmentación y minimizar los riesgos.

Parte del trabajo futuro consistirá en la integración de técnicas de refinamiento automático de contornos, la cuales creemos podrían mejorar la calidad otorgada por las probabilidades espaciales y de transición del ensamble propuesto. Aunado a lo anterior, se replicará la metodología a otros datasets para determinar su generalización.

Referencias

1. Arnold, M., Abnet, C. C., Neale, R. E., Vignat, J., Giovannucci, E. L., McGlynn, K. A., Bray, F.: Global burden of 5 major types of gastrointestinal cancer. *Gastroenterology*, vol. 159, no. 1, pp. 335–349 (2020) doi: 10.1053/j.gastro.2020.02.068
2. Baumgartner, J., Flesia, A., Gimenez, J., Pucheta, J.: A new approach to image segmentation with two-dimensional hidden Markov models. In: *Congress on Computational Intelligence and 11th Brazilian Congress on Computational Intelligence*, pp. 213–222 (2013) doi: 10.1109/BRICS-CCI-CBIC.2013.43
3. Bertelsen, A., Schytte, T., Moller, P.: First clinical experiences with a high field 1.5 T Mr linac. *Acta Oncologica*, vol. 58, no. 10, pp. 1352–1357 (2019) doi: 10.1080/0284186X.2019.1627417
4. Chicco, D., Jurman, G.: The ABC recommendations for validation of supervised machine learning results in biomedical sciences. *Frontiers in Big Data*, vol. 5 (2022) doi: 10.3389/fdata.2022.979465
5. Chou, A., Li, W., Roman, E.: Gi tract image segmentation with U-Net and mask R-CNN. *Technical Report 164*, Stanford University (2022)
6. Despotovic, I., Goossens, B., Philips, W.: MRI segmentation of the human brain: Challenges, methods, and applications. *Computational and Mathematical Methods in Medicine*, vol. 2015, pp. 1–23 (2015) doi: 10.1155/2015/450341

7. Ding, J., Zhang, Y., Amjad, A., Xu, J., Thill, D., Li, X. A.: Automatic contour refinement for deep learning auto-segmentation of complex organs in MRI-guided adaptive radiation therapy. *Advances in Radiation Oncology*, vol. 7, no. 5, pp. 100986 (2022) doi: 10.1016/j.adro.2022.100968
8. Eppenhof, K., Maspero, M., Savenije, M., de Boer, J., van der Voort van Zyp, J., Raaymakers, B., Raaijmakers, A., Veta, M., van den Berg, C., Pluim, J.: Fast contour propagation for MR-guided prostate radiotherapy using convolutional neural networks. *Medical Physics*, vol. 47, no. 3, pp. 1238–1248 (2020) doi: 10.1002/mp.13994
9. Fransson, S., Tilly, D., Strand, R.: Patient specific deep learning based segmentation for magnetic resonance guided prostate radiotherapy. *Physics and Imaging in Radiation Oncology*, vol. 23, pp. 38–42 (2022) doi: 10.1016/j.phro.2022.06.001
10. Johansson, A., Balter, J. M., Cao, Y.: Gastrointestinal 4D MRI with respiratory motion correction. *Medical Physics*, vol. 48, no. 5, pp. 2521–2527 (2021) doi: 10.1002/mp.14786
11. Kawahara, D., Tsuneda, M., Ozawa, S., Okamoto, H., Nakamura, M., Nishio, T., Nagata, Y.: Deep learning-based auto segmentation using generative adversarial network on magnetic resonance images obtained for head and neck cancer patients. *Journal of Applied Clinical Medical Physics*, vol. 23, no. 5, pp. e13579 (2022) doi: 10.1002/acm2.13579
12. Kim, H., Jung, J., Kim, J., B., C., Kwak, J., Yun, J., Lee, S., Lee, J., Yoon, S.: Abdominal multi-organ auto-segmentation using 3D-patch-based deep convolutional neural network. *Scientific Reports*, vol. 10, pp. 6204 (2020) doi: 10.1038/s41598-020-63285-0
13. Lee, S. L., Li, Y., Meudt, J. J., Strang, J., Hebel, D., Alfson, A., Olson, S. J., Kruser, T. R., Smilowitz, J. B., Borchert, K., Loritz, B., Bayouth, J., Bassetti, M.: UW-Madison GI tract image segmentation (2022)
14. Li, G., Sun, J., Song, Y.: Segmentation of medical images with a combination of convolutional operators and adaptive hidden Markov model. In: *IEEE 5th International Conference on Computer and Communications*, pp. 1782–1786 (2019) doi: 10.1109/ICCC47050.2019.9064034
15. Nemani, P., Vollala, S.: Medical image segmentation using LeViT-UNet++: A case study on GI tract data. In: *26th International Computer Science and Engineering Conference* (2022) doi: 10.1109/ICSEC56337.2022.10049343
16. Punn, N. S., Agarwal, S.: Modality specific U-Net variants for biomedical image segmentation: A survey. *Artificial Intelligence Review*, vol. 55, no. 7, pp. 5845–5889 (2022) doi: 10.1007/s10462-022-10152-1
17. Pyun, K., Lim, J., Won, C. S., Gray, R. M.: Image segmentation using hidden Markov Gauss mixture models. *IEEE Transactions on Image Processing*, vol. 16, no. 7, pp. 1902–1911 (2007) doi: 10.1109/TIP.2007.899612
18. Schaeue, D., McBride, W. H.: Opportunities and challenges of radiotherapy for treating cancer. *Nature Reviews Clinical Oncology*, vol. 12, no. 9, pp. 527–540 (2015) doi: 10.1038/nrclinonc.2015.120
19. Sharma, M.: Automated GI tract segmentation using deep learning (2022) doi: 10.48550/ARXIV.2206.11048
20. Song, Y., Adobah, B., Qu, J., Liu, C.: Segmentation of ordinary images and medical images with an adaptive hidden Markov model and viterbi algorithm. *Current Signal Transduction Therapy*, vol. 15, no. 2, pp. 109–123 (2020) doi: 10.2174/1574362413666181109113834
21. Song, Y., Li, Z., Wang, H., Zhang, Y., Yue, J.: MR-LINAC-guided adaptive radiotherapy for gastric MALT: Two case reports and a literature review. *Radiation*, vol. 2, no. 3, pp. 259–267 (2022) doi: 10.3390/radiation2030019
22. Zhou, Z., Rahman-Siddiquee, M. M., Tajbakhsh, N., Liang, J.: UNet++: A nested U-Net architecture for medical image segmentation. In: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pp. 3–11 (2018)