

FACULTAD DE MATEMÁTICA Y  
COMPUTACIÓN UNIVERSIDAD DE LA  
HABANA

TESIS DE DIPLOMA DE LA ESPECIALIDAD CIENCIA  
DE LA COMPUTACIÓN

**Segmentación de Úlceras de Pie  
Diabético (UPD) en secuencias  
de imágenes RGB mediante el  
Segment Anything Model  
(SAM).**

Autor: Abdel Fregel Hernández  
Tutor: Dr. José Alejandro Mesejo Chiong

La Habana  
17 de diciembre de 2024

## 1. Agradecimientos

## 2. Resumen

### 3. Índice

## Índice

1. Agradecimientos	I
2. Resumen	II
3. Índice	III
4. Introducción	1
4.1. Objetivos . . . . .	3
4.2. Estructura de la tesis . . . . .	4
5. Capítulo 1	5
5.1. Estado del arte . . . . .	5
6. Capítulo 2	6
6.1. Segment Anything Model(SAM) . . . . .	6
7. Conclusiones	7
8. Recomendaciones	8
9. Bibliografía	9
10. Anexos	10

## 4. Introducción

La diabetes(diabetes mellitus), es una enfermedad crónica que afecta la forma en la que el cuerpo utiliza la glucosa, una fuente clave de energía. De acuerdo a la *Federación Internacional de Diabetes*, en el año 2021 se reportaron 6.7 millones de muertes a causa de esta enfermedad (Federation, 2024). En nuestro país, según el *Anuario Estadístico de Salud 2022* (Ministerio de Salud Pública, 2023), la prevalencia es de 66,50% de enfermos por cada 1000 habitantes.

Cerca del 86% de personas que padecen de diabetes sufren de úlcera de pie diabético <sup>1</sup>(UPD), y corren el riesgo de amputación. La cicatrización de estas úlceras puede tardar semanas, meses e incluso años, deteriorando la calidad de vida de los pacientes. Actualmente, los médicos cubanos especializados en el tema no cuentan con una herramienta cuantitativa efectiva que valore la severidad y el proceso de curación de las UPD. La medición regular de las úlceras es crucial para evaluar la efectividad del tratamiento y realizar ajustes cuando sea necesario. Un seguimiento adecuado puede prevenir la progresión de la úlcera y reducir el riesgo de amputaciones.

Existen métodos de segmentación y medición de heridas que han conseguido grandes avances en esta área. La calidad de estas es esencial para varios análisis de heridas, como por ejemplo la clasificación de tejidos, Reconstrucción 3D, y la evaluación de la cicatrización(Filko & Nyarko, 2023). Se pueden clasificar en dos tipos: aquellos que requieren contacto y los que no. Los métodos de contacto son invasivos y presentan un alto margen de error; por ello, este trabajo se centra en los métodos no invasivos.

Varios estudios han abordado esta problemática. Por ejemplo, en (Filko & Nyarko, 2023) hacen uso de un sofisticado brazo robótico de 7 grados de libertad(DoF), equipado con una cámara RGB-D <sup>2</sup> y un escáner 3D de alta precisión, para la segmentación y medición de heridas. Este artículo aporta un nuevo algoritmo de segmentación que utiliza una combinación de pro-

---

<sup>1</sup>Hace referencia a una complicación grave de la diabetes que se manifiesta como una herida o llaga abierta en el pie

<sup>2</sup>RGB-D, se refiere a una camara capaz de captar imágenes a color(RGB, formato Red(rojo),Green(verde),Blue(azul) y un sensor de profundidad (D(depth),por su sigla en inglés)

cedimientos 2D y 3D para segmentar correctamente un modelo de herida en 3D. La segmentación se realiza a partir de múltiples fotografías 2D por herida, impulsada por una red neuronal profunda en forma del clasificador MobileNetV2 (Filko & Nyarko, 2023). Este clasificador se combina óptimamente con un único modelo 3D y la inicialización del contorno de la herida. Este contorno inicial se optimiza y ajusta mediante un modelo de contorno activo (Filko & Nyarko, 2023), que envuelve estrechamente la superficie real de la herida utilizando la curvatura de la superficie para alcanzar su objetivo.

En otro estudio realizado por Filko et al., 2018 se exploró la medición y reconstrucción de heridas crónicas de lenta curación, utilizando cámara RGB-D. Con la llegada de cámaras RGB-D económicas, la comunidad de visión por computadora ha ganado una forma más accesible para innovar y crear aplicaciones en diversos campos, incluyendo la medicina, Filko et al., 2018. Estas cámaras, que combinan información de color y profundidad, permiten un análisis más detallado y preciso de imágenes, facilitando el desarrollo de tecnologías que mejoran diagnósticos y tratamientos. El sistema desarrollado en dicho artículo detecta automáticamente heridas analizando bloques de imagen según la similitud del histograma de color utilizando un enfoque de vecinos más cercanos(KNN,por sus siglas en inglés). Este método permite identificar características específicas de las heridas en imágenes, facilitando su segmentación y medición.

Es fundamental el desarrollo de una herramienta capaz de realizar la segmentación y la medición de estas heridas lo más preciso posible, por lo que en este trabajo se propone , para la tarea de segmentar las imágenes, utilizar una herramienta desarrollada por Meta AI llamada Segment Anything Model(SAM)(Meta AI, 2023) que permite identificar y segmentar objetos en imágenes de manera eficiente. Esta herramienta ha ganado popularidad en cuanto a las tareas de segmentación y en el presente trabajo será usada para hacer una segmentación de las imágenes de UPD, las cuales serán tomadas con una cámara RGB-D para asegurar su calidad.

La cámara que se utiliza es una cámara Intel Realsense D435i RGB-D. La cámara utiliza luz estructurada infrarroja y métodos binoculares para obtener información de profundidad, que es la tecnología más común y madura utilizada en las cámaras RGB-D de grado consumidor actuales. La cámara de profundidad puede emitir un flujo de video de hasta 720p a 90 FPS, y la

cámara RGB puede emitir un flujo de video de hasta 1080p a 30 FPS (Zhang et al., 2023). Luego de la segmentación se realizará una clasificación de los tejidos para poder realizar una evaluación de la cicatrización.

La importancia de los tejidos en el proceso de cicatrización radica en que cada tipo de tejido desempeña un papel crucial en la reparación y regeneración de la herida. Durante la cicatrización, se forman diferentes tipos de tejidos, siendo el tejido de granulación uno de los más fundamentales (de Navarra, 2023). Este tejido no solo es esencial para el cierre de la herida, sino que también prepara el lecho para la epitelización, proporcionando un entorno adecuado para la migración celular y la formación de nuevos vasos sanguíneos.

Sin embargo, la presencia de tejido necrótico y biofilm bacteriano puede complicar este proceso. El tejido necrótico, formado por células muertas y detritos, actúa como una barrera que impide la formación de tejido de granulación saludable (Cl, 2024). Esto no solo retrasa el proceso de curación, sino que también aumenta el riesgo de infección, lo que puede llevar a complicaciones graves, como la amputación. Por lo tanto, es crucial realizar un desbridamiento adecuado para remover el tejido necrótico y permitir que la herida progrese hacia las fases de proliferación y cicatrización. Por otro lado, el biofilm bacteriano se forma cuando microorganismos se adhieren a la superficie del lecho de la herida, creando microcolonias protegidas por una matriz polimérica. Esta estructura no solo protege a las bacterias de los tratamientos antibióticos convencionales, sino que también interfiere con la respuesta inmune del cuerpo y dificulta la migración celular necesaria para regenerar tejido sano. La identificación y tratamiento adecuados del tejido necrótico y el biofilm son esenciales para facilitar una curación efectiva y mejorar la calidad de vida del paciente.

#### **4.1. Objetivos**

Este trabajo tiene como objetivo desarrollar una herramienta que permita realizar la segmentación automática de úlceras y los tejidos que las componen a partir de una secuencia de imágenes RGB-D, facilitando así el tratamiento por parte del médico.

Para lograr este objetivo general se tiene los siguientes objetivos específicos:

- 1 Hacer un estudio de la literatura sobre los métodos de segmentación
- 2 Estudiar sobre la utilización del Segment Anything Model.
- 3 La creación de un dataset para la posterior evaluación del modelo
- 4 La evaluación del modelo en cuanto métricas de calidad.

## **4.2. Estructura de la tesis**

La tesis está organizada en cuatro capítulos. El Capítulo 1 presenta una revisión literaria donde se discuten algunos algoritmos existentes así como datasets relevantes. El Capítulo 2 ofrece una introducción al SAM y su aplicación en imágenes médicas. En el Capítulo 3 se explican los detalles estructurales e implementativos del sistema propuesto. Finalmente, el Capítulo 4 muestra los resultados obtenidos y compara estos con otros modelos existentes utilizando medidas cualitativas. Se concluirá con recomendaciones basadas en esta investigación para futuras continuaciones del trabajo. Esta versión busca mantener un hilo conductor claro entre las ideas expuestas, eliminando saltos innecesarios y mejorando la cohesión general del texto.



## 5. Capítulo 1

### 5.1. Estado del arte

## 6. Capítulo 2

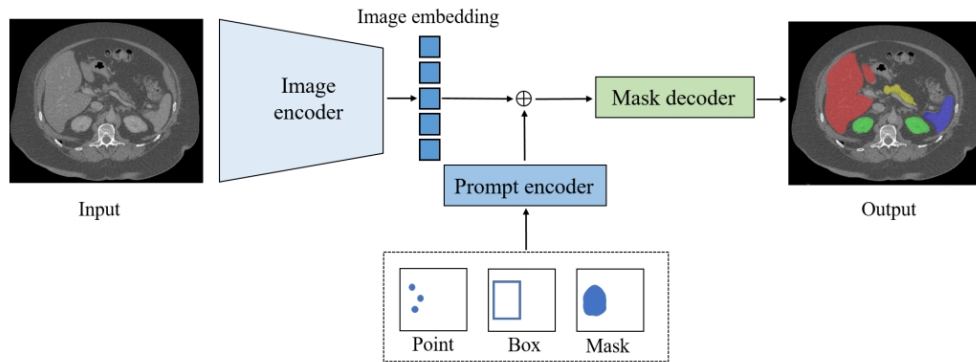
### 6.1. Segment Anything Model(SAM)

SAM utiliza una arquitectura transformer-based, a la cual se le ha probado su eficiencia en el procesamiento de lenguaje natural y en tareas de reconocimiento de imágenes. Específicamente, SAM contiene un codificador de imagen(image encoder) basado en un Vision Transformer(ViT), con este extrae las características de la imagen, también utiliza un prompt encoder para integrar las interacciones del usuario y por último un mask decoder con el objetivo de predecir las máscaras de segmentación con la fusión de las características de la imagen con las entradas del usuario.

En la figura 1 se aprecia la estructura de SAM:

- 1 Image Encoder: SAM utiliza un image encoder basado en un Vision Transformer preentrenado en el esquema de Masked Autoencoder(MAE), el cual está adaptado para procesar imágenes de alta resolución. Este toma imágenes de 1024 x 1024 y da como salida image embeddings bajando su tamaño a un mapa de características de 64 x 64.
- 2 Prompt Encoder: para esto dos tipos de ellos son considerados, uno incluyendo puntos y rectángulos y otro que trabaja con máscaras.
- 3 Mask Decoder : consiste en dos transformer layers con una máscara de predicción dinámica y una puntuación de regresión Intersection-over-Union(IoU, por sus siglas en inglés)

Figura 1: Imagen adaptada de



## 7. Conclusiones

## 8. Recomendaciones

en la figura 1

## 9. Bibliografía

- Cl, Ú. (2024). *El proceso de cicatrización de las heridas crónicas* [Accedido: diciembre 2024]. <https://ulceras.cl/monografico/el-proceso-de-cicatrizacion-de-las-heridas-cronicas/>
- de Navarra, C. U. (2023). *Tejido de granulación* [Accedido: diciembre 2024]. <https://www.cun.es/diccionario-medico/terminos/tejido-de-granulacion>
- Federation, I. D. (2024). *Diabetes Atlas*. Consultado el 9 de diciembre de 2024, desde <https://diabetesatlas.org/>
- Filko, D., Cupec, R., & Nyarko, E. K. (2018). Wound measurement by RGB-D camera. *Machine Vision and Applications*, 29, 633-654. <https://doi.org/10.1007/s00138-018-0920-4>
- Filko, D., & Nyarko, E. K. (2023). 2D/3D Wound Segmentation and Measurement Based on a Robot-Driven Reconstruction System. *Sensors*, 23(6), 3298. <https://doi.org/10.3390/s23063298>
- Meta AI. (2023). *Segment Anything* [Accedido el 14 Nov 2024]. <https://segment-anything.com/>
- Ministerio de Salud Pública. (2023). Anuario Estadístico de Salud 2022 [La Habana: Dirección Nacional de Registros Médicos y Estadísticas de Salud]. <https://salud.msp.gob.cu/wp-content/Anuario/Anuario-2022.pdf>
- Zhang, P., Zhang, Y., & Li, Q. (2023). RGB-D Camera-Based Automatic Wound-Measurement System. *IEEE Transactions on Instrumentation and Measurement*.

## 10. Anexos