

Universidad del Valle de Guatemala
Facultad de ingeniería



**Detección de estructuras microvasculares
en tejidos de riñón humano sanos**

Lucía Alejandra Guzmán Domínguez 20262
Jorge Caballeros Pérez 20009
Pablo Emmanuel Escobar 2096
Eduardo Ramírez Herrera 19946

Guatemala 16 de noviembre del 2023

I. Introducción:

El cuerpo humano, está compuesto por 37 trillones de células, las cuales representan una compleja red de especialización e interacción celular donde se juegan roles cruciales. Estas células, forman parte de un organismo multicelular, no son idénticas una de la otra. Cada una se especializa en hacer funciones específicas., el cual es un proceso conocido como diferenciación o especialización celular. Esta especialización es vital durante el desarrollo embrionario y continúa siendo relevante en la edad adulta, donde células madre se especializan para reemplazar células antiguas o desgastadas. (BioExplorer, 2017)

La especialización celular puede alterar aspectos como tamaño, forma, estructura , actividades metabólicas y respuestas fisiológicas. (BioExplorer, 2017) Esta diferenciación es crucial no solo para la producción de células con funciones especializadas, sino también para la comunicación celular, preservación del material genético y posibles aplicaciones médicas, como en el tratamiento de enfermedades como la diabetes y afecciones cardiovasculares. (BioExplorer, 2017)

Por lo cual el marco de referencias de coordenadas vasculares (VCCF, por sus siglas en inglés) juega un papel fundamental. VCCF utiliza la vasculatura sanguínea como un sistema de navegación primario en el cuerpo humano, lo que permite localizar posiciones celulares a través de estructuras capilares. Este enfoque es parte de esfuerzos internacionales para mapear el cuerpo humano a nivel celular, combinando muestras de tejido heterogéneas en un mapa de referencia que permita una exploración detallada y multidimensional del cuerpo humano, similar a Google Maps. (*Vascular Common Coordinate Framework (VCCF) Workshop*, 2021)

Sin embargo, a pesar de los avances en el conocimiento de la vasculatura y su uso en la localización celular, la microvasculatura sigue siendo un campo menos explorado. La disfunción microvascular (MVD) contribuye a varias condiciones que aumentan la morbilidad y la mortalidad, como enfermedades cardíacas isquémicas, insuficiencia cardíaca, demencia, enfermedad renal crónica e hipertensión. Esta disfunción impone una carga sustancial a los sistemas de salud en todo el mundo, y aún no se ha determinado completamente si la MVD en un órgano constituye una patología distinta o una manifestación de un trastorno sistémico. (Emfietzoglou et al., 2022)

II. Objetivos Generales

Desarrollar e implementar modelos de aprendizaje automático, específicamente redes neuronales convolucionales (CNN), para identificar y segmentar con precisión vasos microvasculares en datos de tejido humano.

III. Objetivos Específicos

- a. Desarrollo de una CNN para la segmentación de vasos sanguíneos: crear un modelo CNN sólido capaz de identificar y segmentar vasos sanguíneos microvasculares en imágenes de tejidos.
- b. Mejorar la precisión y la capacidad del modelo antes mencionado, a través de diferentes técnicas y validación con el conjuntos de datos dado por Kaggle.
- c. Utilizar los datos obtenidos para explorar como la disposición de los vasos sanguíneos microvasculares afecta las relaciones entre células.

IV. Marco Teórico

El procesamiento de imágenes médicas puede ser considerada la revolución en los campos de la medicina y la biología. Estas técnicas permiten una exploración detallada y no invasiva de la anatomía interna, principalmente utilizando conjuntos de datos de imágenes 3D obtenidos de tomografía computarizada (CT) o escáneres de imágenes por resonancia magnética (MRI). (*What Is Medical Image Processing* | Synopsys, 2023) Todo esto debe iniciar con la recopilación de datos de imágenes sin procesar y su reconstrucción en un formato adecuado para su uso en el software correspondiente. Esto permite la creación de modelos 3D para mejorar los resultados del tratamiento, desarrollar dispositivos médicos avanzados y realizar diagnósticos más informados. (*What Is Medical Image Processing* | Synopsys, 2023)

Además, las arquitecturas de aprendizaje profundo, como las redes neuronales artificiales (ANNs), las redes neuronales convolucionales (CNNs) y las redes neuronales recurrentes (RNNs), han demostrado ser particularmente efectivas en la segmentación de imágenes médicas. Estos modelos han logrado resultados significativos en estudios recientes, destacando tanto sus éxitos como sus limitaciones. Un desafío notable en el entrenamiento de estos modelos es la necesidad de grandes conjuntos de datos y poder computacional, así como la complejidad en su interpretación y ajuste. (Seo et al., 2020)

En definitiva, la combinación de técnicas avanzadas de procesamiento de imágenes y algoritmos de aprendizaje automático abrirá nuevas fronteras en la medicina diagnóstica y terapéutica. Estas tecnologías no sólo mejoran la precisión y eficiencia del diagnóstico y el

tratamiento, sino que también facilitan una comprensión más profunda de la anatomía y fisiología humana, contribuyendo así significativamente al avance de la medicina personalizada y la investigación biomédica.

V. Metodología

a. Exploración de Datos (`analisisExploratorio.ipynb`):

- Herramientas y Lenguajes: Utilización de Python, con bibliotecas como pandas, matplotlib y seaborn para la manipulación y visualización de datos.
- Proceso: Se cargaron y exploraron metadatos de imágenes, probablemente relacionados con imágenes de microvasculatura. Se realizaron análisis descriptivos y visualizaciones para comprender las características de los datos.

b. Desarrollo del Modelo de Aprendizaje Automático (`main.ipynb`):

- Herramientas y Lenguajes: Uso de Python y TensorFlow. Se emplearon módulos específicos como Sequential, Conv2D, MaxPooling2D, Flatten, Dense, y Dropout para construir arquitecturas de red.
- Selección del Grupo de Entrenamiento y Prueba: Se dividieron utilizando keras para dividir el dataset.
- Elección del Algoritmo: Se seleccionó una Red Neuronal Convolutiva (CNN) para el procesamiento de imágenes, una elección común para tareas de clasificación y segmentación de imágenes debido a su eficiencia en la captura de patrones espaciales y temporales en datos visuales. Solo se implementó una al esta ser la mejor opción además de que se necesito bastante computo para procesar los 4gb de fotos.
- Razones de la Elección: Las CNN son particularmente adecuadas para el análisis de imágenes médicas debido a su capacidad para manejar la complejidad y variabilidad de este tipo de datos. Por lo tanto esta era nuestra mejor opción.

c. Recursos de Cómputo:

- Plataforma de Desarrollo: El uso de Jupyter Notebook sugiere un enfoque interactivo y flexible para la codificación y visualización.
- Capacidad de Cómputo: Se necesitaron recursos significativos de GPU, debido a la extensa cantidad de tiempo y fotos que habían, por lo cual el modelo tomo varias horas.

VI. Resultados y análisis de los mismos

Los resultados obtenidos en nuestro proyecto indican un progreso notable en la tarea de identificar vasos sanguíneos en imágenes médicas. Con un mloss (pérdida en el modelo) de 0.3860 y una accuracy (precisión) de 0.8574 en el conjunto de entrenamiento, el modelo mostró una capacidad considerable para aprender de los datos.

El conjunto de datos original consistía en imágenes TIFF de 512x512 píxeles, con máscaras de segmentación poligonal proporcionadas en formato JSONL. Se realizaron tareas de limpieza y análisis exploratorio para entender y preparar estos datos para el entrenamiento del modelo. Esto incluyó la identificación de diferentes estructuras anatómicas y la comprensión de las características demográficas asociadas.

El análisis exploratorio del dataset en el proyecto se realizó de la siguiente manera, basado en las primeras etapas del proceso reflejadas en el cuaderno analisisExploratorio.ipynb:

Importación de Herramientas y Bibliotecas: Se utilizó Python con bibliotecas como pandas para la manipulación de datos, matplotlib y seaborn para visualización, y json para manejar datos en formato JSON.

Carga de Datos: Se cargaron los datos de metadatos de imágenes de microvasculatura, específicamente los archivos wsi_meta.csv y tile_meta.csv, que contienen información detallada sobre las imágenes y las baldosas (tiles) derivadas de las mismas.

Análisis Preliminar: Se realizó un análisis inicial de los datos cargados utilizando métodos como .head() para ver las primeras entradas y .describe() para obtener estadísticas descriptivas, lo que proporcionó una comprensión básica de la estructura y características de los datos.

Este proceso inicial indica un enfoque sistemático y detallado en el análisis exploratorio, fundamental para comprender el conjunto de datos antes de proceder con el modelado y la aplicación de técnicas de aprendizaje automático.

En cuanto a la aplicación creada, permite a los usuarios cargar imágenes y visualizar los vasos sanguíneos analizados por el modelo. Esto no solo demuestra la viabilidad técnica del proyecto, sino que también proporciona una herramienta útil para la investigación médica y posibles aplicaciones clínicas. La alta precisión del modelo y la interfaz interactiva de la aplicación resaltan su potencial en el campo de la imagenología médica.

V. Conclusiones

En este proyecto se logró desarrollar e implementar modelos de aprendizaje automático, específicamente redes neuronales convolucionales (CNN), para identificar y segmentar con precisión vasos microvasculares en datos de tejido humano. La CNN desarrollada mostró una habilidad considerable en la detección de vasos sanguíneos, evidenciada por una alta precisión en el conjunto de entrenamiento. Este logro resalta la eficacia de las CNN en el procesamiento complejo y detallado requerido para el análisis de imágenes médicas, marcando un avance significativo en la aplicación de técnicas de aprendizaje automático en el campo de la imagenología médica. A pesar de no poder implementar más de un modelo debido a la falta de nivel de cómputo nuestro modelo es considerablemente eficaz.

Bibliografia

BioExplorer. (2017, June 9). *Cell Specialization | Mechanisms, Examples & Importance* | BioExplorer. Bio Explorer.
<https://www.bioexplorer.net/cell-specialization.html/>

Vascular Common Coordinate Framework (VCCF) Workshop. (2021). Github.io.
https://cns-iu.github.io/workshops/2021-03-29_vccf_workshop/

Emfietzoglou, M., Dimitrios Terentes-Printzios, Kotronias, R. A., Marin, F., Montalto, C., Luigi, G., & Banning, A. P. (2022). The spectrum and systemic associations of microvascular dysfunction in the heart and other organs. *Nature Cardiovascular Research*, 1(4), 298–311.
<https://doi.org/10.1038/s44161-022-00045-5>

What is Medical Image Processing | Synopsys. (2023). Synopsys.com.
<https://www.synopsys.com/glossary/what-is-medical-image-processing.html>

Seo, H., Masoud Badiei Khuzani, Vasudevan, V., Huang, C., Ren, H., Xiao, R., Xiao, J., & Li, X. (2020). Machine learning techniques for biomedical image segmentation: An overview of technical aspects and introduction to state-of-art applications. *Medical Physics*, 47(5). <https://doi.org/10.1002/mp.13649>

