

## Segmentación Automática de Cálculos Renales mediante U-Net a través de Imágenes de Tomografía Computarizada

Automatic Segmentation of Kidney Stones by U-Net through Computed Tomography Images

Salvador Calderón Uribe<sup>1</sup>, Uriel Calderón Uribe<sup>2</sup>, Luis Alberto Morales Hernández<sup>1</sup>, Irving Armando Cruz Albarrán<sup>13</sup>.

<sup>1</sup>Laboratorio de Visión Artificial y Termografía/Mecatrónica Facultad de Ingeniería, Universidad Autónoma de Querétaro, Campus San Juan del Río, San Juan del Río 76807, México.

<sup>2</sup>Departamento de Estudios Multidisciplinarios, División de Ingenierías, Campus Irapuato-Salamanca, Universidad de Guanajuato, Yuriria 38944, Guanajuato, México.

<sup>3</sup>Sistemas de Inteligencia Artificial Aplicados a Modelos Biomédicos y Mecánicos, Facultad de Ingeniería, Universidad Autónoma de Querétaro, Campus San Juan del Río, San Juan del Río 76807, México.

scalderon02@alumnos.uaq.mx<sup>1</sup>, urielcal92@gmail.mx<sup>2</sup>, lamorales@hspdigital.org<sup>1</sup>, icruz@hspdigital.org<sup>13</sup>.

### Resumen

Los cálculos renales o litiasis urinaria se definen como la presencia de depósitos duros hechos de minerales y sales en riñón, uretra y vejiga. Actualmente, los cálculos renales son considerados como una enfermedad crónica que, sin tratamiento adecuado, tiene una alta tasa de recurrencia. Debido a lo anterior, el objetivo del presente trabajo es la segmentación automática de los cálculos renales a través del análisis de imágenes de tomografía computarizada (TC). Para lograrlo, se propone un método de aprendizaje profundo (deep learning) basado en U-Net para segmentar los depósitos duros. La base de datos utilizada cuenta con un total de 1300 muestras, de las cuales se aplicó la técnica de un aumento de datos, rotando, cortando y variando el brillo de las imágenes de forma aleatoria. Los resultados muestran un índice de similitud promedio del 64%, una exactitud del 99.87% y una sensibilidad del 88%, demostrando que los algoritmos basados en deep learning pueden ser una herramienta viable para la segmentación automática de cálculos renales.

**Palabras clave:** Calculo renal, Tomografía computarizada, Segmentación automática, U-Net, Aumento de datos.

### Introducción

Los cálculos renales son un problema creciente diagnosticado cada vez con mayor frecuencia en el mundo; definiéndose como un desequilibrio entre la cantidad de inhibidores (citrato, magnesio) y de promotores (calcio, ácido úrico, fosfatos, oxalatos) de la cristalización de la orina secundario a alteraciones metabólicas (Hernández-Nájera, 2020). Esta alteración en la composición urinaria asociado a un bajo volumen urinario permite crear un medio propicio para la formación de cálculos renales. En México, Otero *et al.*, (1995) destacaron que esta enfermedad abarca el 13% de todas las hospitalizaciones por enfermedad renal en el ámbito nacional en el Instituto Mexicano del Seguro Social (IMSS); lo cual lleva a un incremento en la incidencia de otros padecimientos como infecciones urinarias, requerimiento de procedimientos quirúrgicos para la extracción de los cálculos y, en etapas más complicadas, a la progresión a la falla renal crónica.

Con el desarrollo de algoritmos basados en inteligencia artificial, en la última década se han propuesto modelos de Deep Learning (DL) basados en redes neuronales convolucionales (RNC), los cuales se han enfocado en el análisis de imágenes médicas con la capacidad de clasificar, detectar y segmentar diversas afecciones, tales como, tumores (Wablang *et al.*, 2022), esclerosis múltiple (Shoeibi *et al.*, 2021), Alzheimer (Hazarika *et al.*, 2022), entre otras. Así mismo, técnicas de visión artificial y DL se han empleado en el campo de la urología para la detección automática de cálculos ureterales y renales. Thein *et al.* (2018), por ejemplo, desarrollaron algoritmos de umbralización basado en la intensidad para la detección y segmentación de cálculos renales en imágenes de TC, alcanzando una sensibilidad del 95.24%. Por otro lado, Yildirim *et al.*, 2021, implementaron un modelo automático basado en RNC para clasificar los cálculos urinarios mediante TC, obteniendo una precisión del 96.82%. Lee *et al.*, (2010) utilizaron características basadas en la textura y la intensidad para entrenar una red neuronal artificial (RNA) para distinguir los cálculos renales de las calcificaciones vasculares. Más recientemente, Cui *et al.* (2021), utilizaron una U-Net 3D de doble etapa seguida de un umbral simple y un crecimiento de región para segmentar los cálculos



renales, alcanzando una sensibilidad del 96%. De este modo, aunque herramientas de umbralización y textura permitan la identificación de cálculos renales en TC, el diagnóstico de estos depósitos presenta un desafío en la segmentación debido a factores como, las complicadas estructuras de interés, el ruido, las similitudes entre órganos adyacentes, variación en los tamaños, localización, entre otros.

Por ello, en el presente estudio se propone un modelo basado en Deep Learning que permita la segmentación automática cálculos renales basado en el modelo U-Net. Debido a la poca cantidad de imágenes, se realizaron técnicas de aumento de datos para el entrenamiento del modelo, planteando una propuesta para proporcionar una mejor generalización de los datos de entrenamiento. En este trabajo, las TC son divididas en conjuntos de entrenamiento-validación y prueba para evaluar el modelo. La U-Net alcanzó un índice de similitud DICE del 64%, una exactitud del 99.87% y una sensibilidad del 88% utilizando dicha arquitectura.

## Metodología Propuesta

En esta sección se describe la metodología utilizada para la segmentación de cálculos renales. En la Figura 1 se muestra una visión general de la metodología propuesta. De esta figura, es posible observar que el sistema de segmentación se divide en 3 etapas: aumento de datos, entrenamiento-validación, y evaluación. La etapa de aumento de datos se basa en incrementar la cantidad y diversidad de las imágenes disponibles para el entrenamiento del modelo. En la etapa de entrenamiento-validación, las imágenes resultantes del aumento de datos son divididos en dos subconjuntos utilizados para entrenar y validar el modelo. Por último, en la etapa de evaluación, un subconjunto de imágenes es utilizado para validar el modelo.

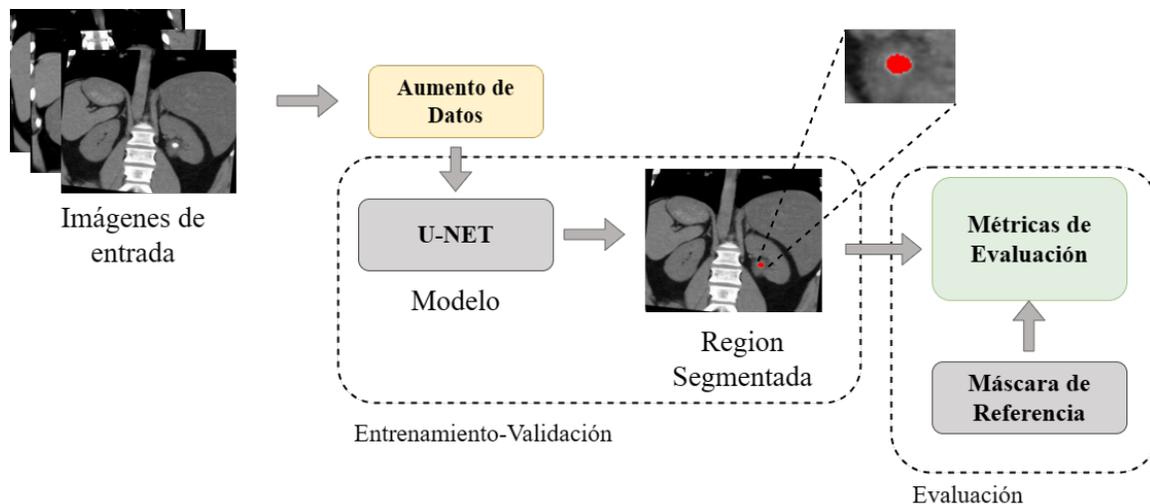


Figura 1. Diagrama general del proceso de segmentación.  
Fuente: Elaboración propia.

## Imágenes de Entrada

Para medir el rendimiento del modelo U-Net propuesto en este estudio, se utilizó el conjunto de datos tez-roi-aug dataset de la plataforma Roboflow Universe (Tez, 2023). El conjunto de datos contiene un total de 1300 imágenes de TC, que abarcan distintos tamaños, formas y posiciones de los cálculos dentro del sistema urinario. Cada imagen cuenta con un tamaño de 320x391 píxeles y va acompañada de cuadros delimitadores marcados, de los cuales se obtuvieron las máscaras mediante técnicas de umbralización tradicionales. Así mismo, el conjunto de datos está dividido en función del área que abarca los cálculos dentro de la imagen, dividiendo al conjunto en tres diferentes subconjuntos; el conjunto donde el área de interés abarca un rango de 1-20 píxeles, el conjunto donde el área abarca un rango de 20-200 píxeles y el conjunto donde el área abarca un rango superior a los 200 píxeles. En la figura 2, se muestran algunas de las

imágenes y sus máscaras correspondientes. De esta figura, se aprecia que existen diversas variaciones en los tamaños, posición y forma de los cálculos, así como también una similitud en la intensidad entre estos depósitos y diferentes órganos y/o estructuras adyacentes, por lo tanto, se requiere de técnicas de reconocimiento de patrones e inteligencia artificial para segmentar cada uno de los depósitos.

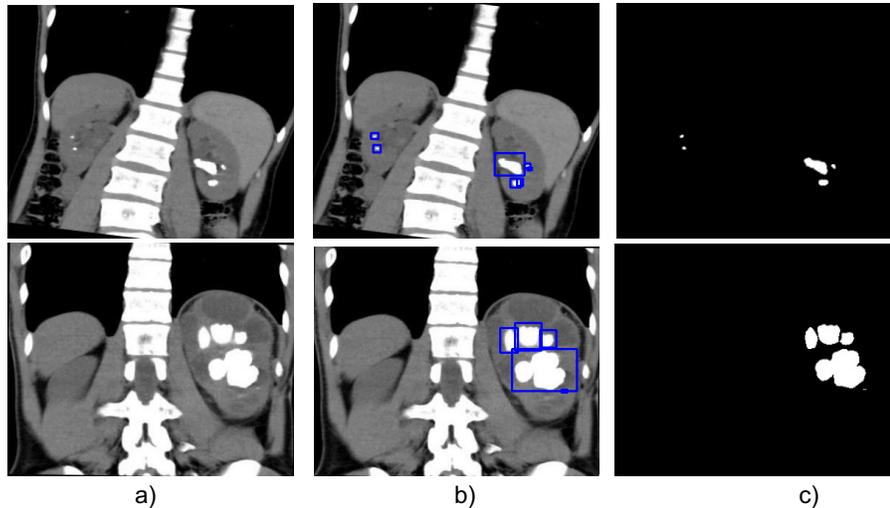


Figura 2. Imágenes de entrada, a) Tomografía computarizada de múltiples cálculos renales, b) Cuadro delimitador de las regiones de interés, c) Máscara de salida.

Fuente: Elaboración propia.

## Aumento de Datos

En DL el aumento de datos resulta ser una herramienta eficiente para abordar el problema de la escasez de información, ya que se caracteriza por aumentar el conjunto de datos de entrenamiento, disminuyendo el sobreajuste y mejorando el rendimiento cuando se validan en un conjunto de datos independientes no vistos (Hussain *et al.*, 2017, Chlap *et al.*, 2021). Sin embargo, aunque existe diferentes técnicas para el aumento de datos (ejemplo, Rotación, escalamiento, inyección de ruido, entre otras) en este estudio mediante la librería Albumentation (Buslaev *et al.*, 2020), se utilizaron las técnicas de rotación, corte, variación de brillo y variación del contenido entre canales RGB; siendo la rotación en un ángulo  $0 \pm 10^\circ$ , el corte aleatorio de 200 píxeles de ancho por 200 píxeles de largo, y la probabilidad de brillo y variación del contenido RGB de 0.2.

## U-Net

U-Net es un tipo de red neuronal perteneciente a las redes totalmente convolucionales (FCN, por sus siglas en inglés) utilizadas comúnmente en tareas de segmentación de imágenes y procesamiento de imágenes médicas (Long *et al.*, 2015, Huang *et al.*, 2022). La arquitectura U-Net se caracteriza por componerse de un codificador para la extracción de características y un decodificador para la reconstrucción de imágenes. La Figura 3 muestra la estructura utilizada para la segmentación de cálculos renales.

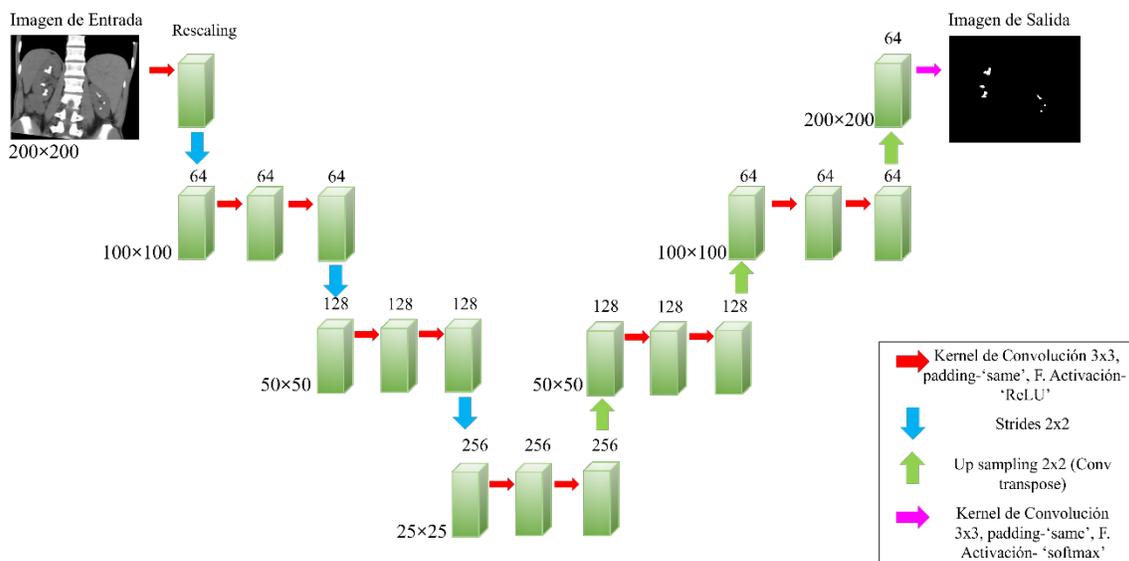


Figura 3. Arquitectura U-Net.  
Fuente: Elaboración

De la Figura 3, se puede notar que la U-Net cuenta con una capa de entrada de 200x200x3 y una capa de reescalado en un rango de [0-1], dentro de la etapa de codificación, la estructura cuenta con una secuencia de convolución con un conjunto de filtros que aumentan gradualmente utilizando una función ReLU como función de activación. La segunda mitad del modelo es una etapa de deconvolución simétrica a la etapa de codificación, la cual, permite llevar a cabo el upsampling del conjunto de características y, a su vez, reconstruir la imagen a la forma 200x200.

## Métricas de evaluación

El rendimiento de la arquitectura se evaluó sobre el conjunto de prueba utilizando el índice de similitud DICE (1) (Huang *et al.*, 2022), la exactitud (2) y la sensibilidad (3), DICE permite medir la similitud que existe entre la máscara de referencia y la predicción realizada por el modelo; la exactitud se refiere al grado en el que los resultados (pixeles) de la predicción coinciden con la segmentación del especialista (Li *et al.*, 2022), la sensibilidad, es utilizado para medir el número de cálculos renales detectados correctamente (Thein *et al.*, 2018).

$$DICE = 2 \left( \frac{|P \cap G|}{|P| \cup |G|} \right) \quad (1)$$

$$Exactitud = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$Sensibilidad = \frac{TP}{TP + FN} \quad (3)$$

Donde P es la imagen binaria predicha por el modelo y G la imagen obtenida por el especialista. TP son los verdaderos positivos, TN los verdaderos negativos, FP los falsos positivos, FN los falsos negativos.

## Resultados

En esta sección se presenta la experimentación para medir el índice de similitud, la exactitud y la sensibilidad del modelo de segmentación. El desarrollo del sistema se llevó a cabo utilizando la base de datos mencionada en la sección anterior. En primer lugar, se muestran los resultados obtenidos del aumento de datos y del entrenamiento de la Arquitectura. Por último, se compara el modelo U-Net propuesto con modelos de la literatura que utilizan una arquitectura similar.

### Aumento de Datos de Entrenamiento

Al contar con un total de 1300 imágenes, se realizó el proceso de separación entrenamiento-evaluación, en una proporción 80%-20%, (1040 imágenes para entrenamiento y 260 imágenes para evaluación), y se realizó el aumento de datos al conjunto de entrenamiento, aumentando los datos a un total de 3120 imágenes del subconjunto de entrenamiento (Figura 4).

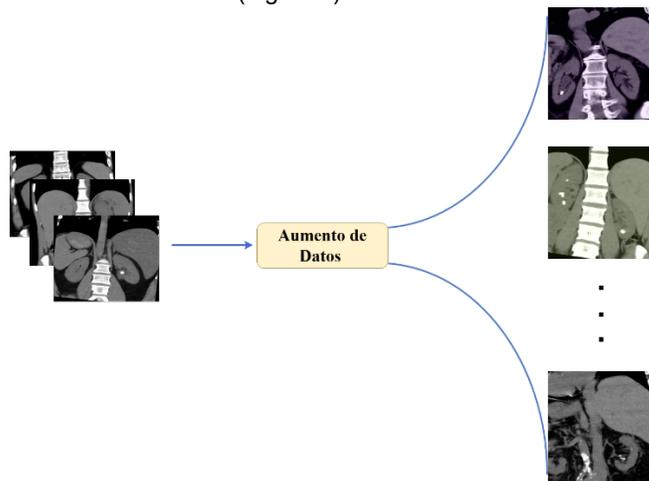


Figura 4. Representación del aumento de datos.  
Fuente: Elaboración propia

### Entrenamiento del Modelo

De la etapa de entrenamiento se utilizó el conjunto de datos formado por un total de 3120 imágenes, de los cuales 312 imágenes fueron reservadas para validar el rendimiento del modelo U-Net. La Figura 5 describe los resultados obtenidos del entrenamiento del modelo. De la Figura 5, es posible observar que el valor de la función perdida decremента en cada época de entrenamiento, alcanzando el valor mínimo de 0.0038 con la función perdida de entropía cruzada dispersa en las épocas 40-50.

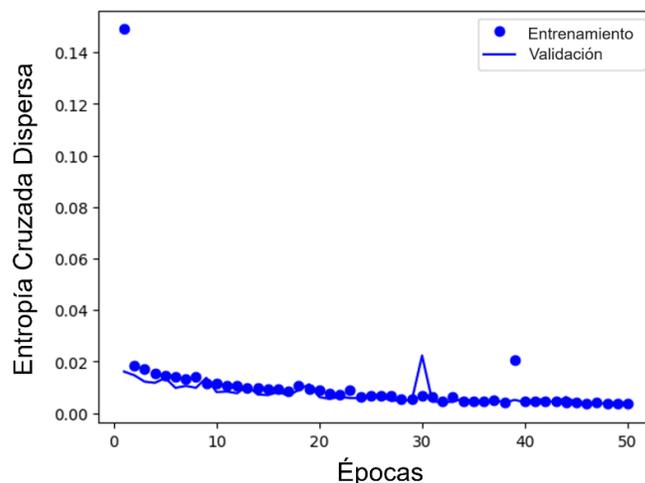


Figura 5. Resultados de la función perdida en la etapa de entrenamiento.  
Fuente: Elaboración propia

## Evaluación de la Similitud

Los resultados obtenidos del conjunto de evaluación son presentados en la tabla 1. De esta tabla, es posible apreciar que el modelo U-Net es capaz de generar máscaras con un mayor índice de similitud de imágenes cuya presencia de cálculos abarca un rango superior a los 20 píxeles. Sin embargo, presenta dificultades al segmentar depósitos cuyo rango sea inferior a este número de píxeles, esto debido a la dificultad de detectar detalles sutiles y la capacidad de distinguir entre el objeto de interés y otras estructuras o variaciones de intensidad por parte del modelo.

**Tabla 1.** Índice de similitud en imágenes con presencia de cálculos renales.

	1-20 píxeles	20-200 píxeles	>200 píxeles
Coefficiente de similitud (%)	45.45%	66.20%	79.32%

Fuente: Elaboración propia.

## Comparaciones

Se compararon los resultados obtenidos con diferentes modelos de DL. La tabla 2 muestra los diferentes métodos utilizados para segmentar cálculos renales.

De esta tabla 2, es posible observar que el desempeño alcanzado por el modelo es aceptable con respecto a los propuestos en la literatura; si bien la arquitectura propuesta no supera a los modelos más robustos, esta requiere de un menor número de imágenes para otorgar una buena segmentación de los cálculos renales, a diferencia de los modelos propuestos en la literatura basados en el uso de voxels que les permite capturar la información volumétrica de los cálculos y su entorno con mayor precisión.



**Tabla 2.** Comparación con diferentes modelos.

Modelo	DICE	Exactitud	Sensibilidad
Res U-Net (Li et al., 2022)	79.83%	99.95%	96.61%
DeepLabV3+ (Li et al., 2022)	41.09%	99.59%	70.91%
UNETR (Li et al., 2022)	61.92%	99.72%	82.02%
DenseNet (Yildirim et al., 2021)	59%	75%	-/-
<b>Propuesto</b>	<b>64.06%</b>	<b>99.87%</b>	<b>88%</b>

Fuente: Elaboración propia.

## Conclusión

En este estudio, se presenta un modelo destinado a la segmentación de cálculos renales en imágenes de TC. Se emplea la arquitectura U-Net reconocida en la segmentación semántica de imágenes médicas alcanzando un índice de similitud del 64% una precisión del 99.87% y una sensibilidad del 88%. Los resultados obtenidos demuestran que los enfoques basados en U-Net producen un mayor índice de similitud a partir del uso de aumento de datos lo que los hace una herramienta relevante en el análisis médico. Futuros trabajos estarán enfocados en mejorar el índice de similitud y la sensibilidad en la detección de cálculos renales más pequeños, esto a partir de un mayor número de imágenes y una mejor representación de los cálculos renales.

## Referencias

- Buslaev, A., Iglovikov, V. I., Khvedchenya, E., Parinov, A., Druzhinin, M., & Kalinin, A. A. (2020). *Albumentations: fast and flexible image augmentations*. *Information*, 11(2), 125. <https://doi.org/10.3390/info11020125>
- Chlap, P., Min, H., Vandenberg, N., Dowling, J., Holloway, L., & Haworth, A. (2021). *A review of medical image data augmentation techniques for deep learning applications*. *Journal of Medical Imaging and Radiation Oncology*, 65(5), 545-563. <https://doi.org/10.1111/1754-9485.13261>
- Cui, Y., Sun, Z., Ma, S., Liu, W., Wang, X., Zhang, X., & Wang, X. (2021). *Automatic detection and scoring of kidney stones on noncontrast CT images using STONE nephrolithometry: combined deep learning and thresholding methods*. *Molecular Imaging and Biology*, 23, 436-445. <https://doi.org/10.1007/s11307-020-01554-0>
- Hazarika, R. A., Kandar, D., & Maji, A. K. (2022). *An experimental analysis of different deep learning based models for Alzheimer's disease classification using brain magnetic resonance images*. *Journal of King Saud University-Computer and Information Sciences*, 34(10), 8576-8598. <https://doi.org/10.1016/j.jksuci.2021.09.003>
- Hernández Nájera, V. V. (2020). *Efectos de la obesidad y el síndrome metabólico sobre litiasis urinaria en el noreste de México*. <https://hdl.handle.net/11285/640928>



- Huang, S. Y., Hsu, W. L., Hsu, R. J., & Liu, D. W. (2022). Fully convolutional network for the semantic segmentation of medical images: A survey. *Diagnostics*, 12(11), 2765. <https://doi.org/10.3390/diagnostics12112765>
- Hussain, Z., Gimenez, F., Yi, D., & Rubin, D. (2017). Differential data augmentation techniques for medical imaging classification tasks. In *AMIA annual symposium proceedings (Vol. 2017, p. 979)*. American Medical Informatics Association. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5977656/>
- Lee, H. J., Kim, K. G., Hwang, S. I., Kim, S. H., Byun, S. S., Lee, S. E., ... & Seong, C. G. (2010). Differentiation of urinary stone and vascular calcifications on non-contrast CT images: an initial experience using computer aided diagnosis. *Journal of digital imaging*, 23, 268-276. <https://doi.org/10.1007/s10278-009-9181-0>
- Li, D., Xiao, C., Liu, Y., Chen, Z., Hassan, H., Su, L., ... & Huang, B. (2022). Deep segmentation networks for segmenting kidneys and detecting kidney stones in unenhanced abdominal CT images. *Diagnostics*, 12(8), 1788. <https://doi.org/10.3390/diagnostics12081788>
- Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3431-3440)*. [https://openaccess.thecvf.com/content\\_cvpr\\_2015/html/Long\\_Fully\\_Convolutional\\_Networks\\_2015\\_CVPR\\_paper.html](https://openaccess.thecvf.com/content_cvpr_2015/html/Long_Fully_Convolutional_Networks_2015_CVPR_paper.html)
- Otero, F., Lugo, A., & Durán, A. (1995). Las enfermedades renales en el Instituto Mexicano del Seguro Social (1982-1989). *Rev Asoc Med Int Mex*, 11, 21-29. [revistamedica.imss.gob.mx/editorial/index.php/revista\\_medica/article/view/1802](http://www.revistamedica.imss.gob.mx/editorial/index.php/revista_medica/article/view/1802)
- Shoeibi, A., Khodatars, M., Jafari, M., Moridian, P., Rezaei, M., Alizadehsani, R., ... & Acharya, U. R. (2021). Applications of deep learning techniques for automated multiple sclerosis detection using magnetic resonance imaging: A review. *Computers in Biology and Medicine*, 136, 104697. <https://doi.org/10.1016/j.compbiomed.2021.104697>
- Tez (abril de 2023), TEZ\_ROI\_AUG Computer Vision Project, Roboflow Universe, <http://universe.roboflow.com/tez-nwxf5/tez-roi-aug>
- Thein, N., Nugroho, H. A., Adji, T. B., & Hamamoto, K. (2018, November). An image preprocessing method for kidney stone segmentation in CT scan images. In *2018 International Conference on Computer Engineering, Network and Intelligent Multimedia (CENIM) (pp. 147-150)*. IEEE. Doi: 10.1109/CENIM.2018.8710933
- Wahlang, I., Maji, A. K., Saha, G., Chakrabarti, P., Jasinski, M., Leonowicz, Z., & Jasinska, E. (2022). Brain Magnetic Resonance Imaging Classification Using Deep Learning Architectures with Gender and Age. *Sensors*, 22(5), 1766. <https://doi.org/10.3390/s22051766>
- Yildirim, K., Bozdog, P. G., Talo, M., Yildirim, O., Karabatak, M., & Acharya, U. R. (2021). Deep learning model for automated kidney stone detection using coronal CT images. *Computers in biology and medicine*, 135, 104569. <https://doi.org/10.1016/j.compbiomed.2021.104569>

