

En sus estudios doctorales, el investigador Saúl Calderón se enfocó a estudiar el impacto y mitigación de las distribuciones diferentes entre el conjunto de datos etiquetado y no etiquetado, en aplicaciones de análisis de imágenes médicas. Parte de su investigación lo hizo con imágenes de rayos X de tórax para la detección de covid-19. Imagen cortesía S. Calderón.

Inteligencia artificial

## Investigación sobre detección de covid-19 usando rayos X produce destacadas publicaciones

14 de Diciembre 2021 Por: [Johan Umaña Venegas](#) [1]

**Dos artículos fueron publicados en revistas científicas del primer cuartil de impacto**

**Es una investigación de doctorado para mejorar modelos semisupervisados de *deep learning* para detectar enfermedades usando imágenes médicas**

La utilización de **inteligencia artificial para detectar enfermedades a partir de imágenes médicas** es un campo sumamente prometedor para mejorar los sistemas de salud en el mundo, con **utilidades que podrían haberse aplicado en la respuesta a la pandemia provocada por el covid-19**. Sin embargo, quedan muchos retos por delante para que estas técnicas sean de amplio uso.

**El primer reto por solucionar es la recopilación de datos.** Luego, que los **modelos de inteligencia artificial tengan la capacidad de operar con alta exactitud** a pesar de trabajar con conjuntos de datos de distintas fuentes y distribuciones.

Este es el planteamiento de la **investigación del M.Sc. Saúl Calderón Ramírez**, quien está por terminar sus estudios de doctorado en el **Instituto de Inteligencia Artificial, de De Montfort University**, [2] en Leicester, Inglaterra.

**“El enfoque del trabajo fue averiguar qué sucede cuando entrenamos un modelo semisupervisado que usa datos no etiquetados y etiquetados en el contexto de imágenes médicas.** Usar datos etiquetados es muy caro. Cuando entrenas un modelo, por ejemplo con datos del Hospital México, es necesario usar datos de ese mismo hospital, porque cuando usamos datos de otros hospitales para su uso cotidiano, el modelo tiende a tener un rendimiento mucho menor. Esto, dado que el modelo se entrenó con datos de pacientes con fisonomía diferente o con imágenes adquiridas con distintos parámetros.

“Entonces, por eso es que **es atractivo usar datos no etiquetados porque no ocupas de mucha supervisión**”, detalló Calderón.

**Según el investigador, la ventaja final de utilizar datos de diferentes conjuntos o lotes de procedencia es que mejora la capacidad de “aprendizaje” de los modelos computacionales y, con esto, el porcentaje de acierto.**

Calderón forma parte del **grupo especializado de investigación Parma** [3] (*Pattern Recognition and Machine Learning*), del TEC, en el cual ya ha liderado investigaciones para detectar cáncer de mama por medio de mamografías.

En sus estudios doctorales, se enfocó a estudiar el impacto y mitigación de las distribuciones diferentes entre el conjunto de datos etiquetado y no etiquetado en aplicaciones de análisis de imágenes médicas. Parte de su investigación la hizo con imágenes de rayos X de tórax para la detección de covid-19.

“Las aplicaciones son muy prometedoras. La academia tiene la posibilidad de promover y generar una cultura de investigación orientada a los datos en los hospitales, de que realmente los equipos de informática en los hospitales construyan conjuntos de datos que utilicen para entrenar los modelos de forma continua y que realmente se cercioren de la calidad de los datos, que sean representativos, con mitigación de sesgos. Entonces, se podrían generar estos modelos, que realmente van a ser muchísimo más efectivos y van a ser de gran utilidad a los médicos, tanto para detectar enfermedades como para pronosticar el avance de las mismas o incluso definir una ruta de tratamiento”. ”

M.Sc. Saúl Calderón Ramírez, especialista en inteligencia artificial.

## Producción académica

Como parte de los resultados de ese estudio, Calderón publicó dos artículos en prominentes revistas científicas, ubicadas en el primer cuartil de impacto dentro de su ámbito.

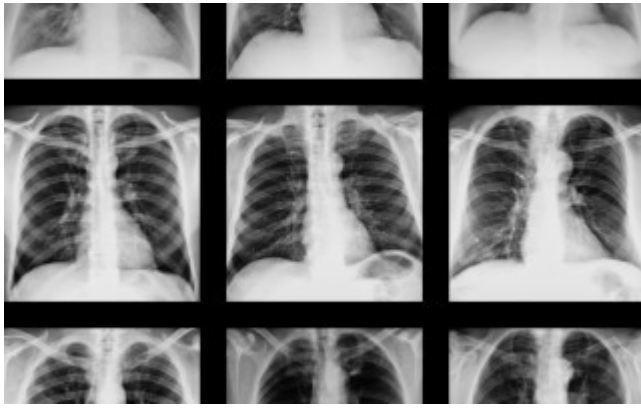
Los *papers* son los siguientes:

- **Correcting data imbalance for semi-supervised Covid-19 detection using X-ray chest images** <sup>[4]</sup> (Corrección del desequilibrio de datos para la detección de covid-19 semisupervisada mediante imágenes de rayos X de tórax). Publicado en Applied Soft Computing, revista de la World Federation on Soft Computing.
- **Improving Uncertainty Estimation With Semi-supervised Deep Learning for COVID-19 Detection Using Chest X-ray Images** <sup>[5]</sup> (Mejorando la estimación de la incertidumbre con aprendizaje profundo semisupervisado para la detección de covid-19 mediante imágenes de rayos X de tórax). Publicado en IEEE Access, revista del Instituto de Ingenieros Eléctricos y Electrónicos.

“Son casos de uso reales del problema que en general se enfrenta cuando usamos aprendizaje semisupervisado y se tiene un conjunto de datos no etiquetado o etiquetado y tienes otras fuentes de conjuntos de datos no etiquetado. Así que el enfoque en general de la tesis fue proponer técnicas que nos permitan medir la calidad de un conjunto de datos para imágenes médicas, para análisis de imagen analítica. Entonces, de acuerdo a eso, implementar técnicas que nos permitan crear un conjunto de datos más relevante para entrenar un modelo semisupervisado”, ahondó Calderón.

El doctorado de Calderón en la Universidad De Montfort también sirvió para promover el desarrollo académico de estudiantes del TEC, pues dos alumnos de Ingeniería en Computación <sup>[6]</sup> hicieron su práctica en esa casa de estudios inglesa e incluso publicaron en una conferencia de renombre.

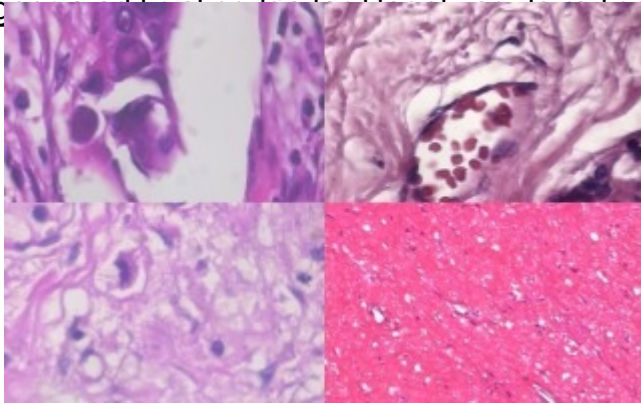
Los estudiantes **Diego Murillo Hernández y Kevin Rojas Salazar**, publicaron el *paper*—junto a Calderón y Luis Alexander Calvo Valverde, también docente del TEC— titulado **Improving Uncertainty Estimations for Mammogram Classification using Semi-Supervised Learning** <sup>[7]</sup> (Mejorando las estimaciones de incertidumbre para la clasificación de mamografías mediante el aprendizaje semisupervisado), que fue publicado en la International Joint Conference on Neural Networks <sup>[8]</sup>(IJCNN 2021).



[9]

## Inteligencia artificial podría facilitar el uso de radiografías de pecho para diagnosticar COVID-19

[9]



[10]

## Cáncer de mama: inteligencia artificial agilizará el diagnóstico [10]

**Source URL (modified on 02/17/2022 - 19:10):** <https://www.tec.ac.cr/hoyeneltec/node/4057>

### Enlaces

[1] <https://www.tec.ac.cr/hoyeneltec/users/johan-umana-venegas>

[2] <https://www.dmu.ac.uk/research/centres-institutes/iai/index.aspx>

[3] <https://www.tec.ac.cr/grupo-investigacion/parma>

[4] <https://www.sciencedirect.com/science/article/pii/S156849462100613X?v=s5>

[5] <https://ieeexplore.ieee.org/document/9445026>

[6] <https://www.tec.ac.cr/programas-academicos/bachillerato-ingenieria-computacion>

[7] [https://www.researchgate.net/profile/Saul-Calderon-Ramirez-2/publication/351528900\\_Improving\\_Uncertainty\\_Estimations\\_for\\_Mammogram\\_Classification\\_using\\_Semi-Supervised\\_Learning/links/609c6779299bf1259ece26e7/Improving-Uncertainty-Estimations-for-Mammogram-Classification-using-Semi-Supervised-Learning.pdf](https://www.researchgate.net/profile/Saul-Calderon-Ramirez-2/publication/351528900_Improving_Uncertainty_Estimations_for_Mammogram_Classification_using_Semi-Supervised_Learning/links/609c6779299bf1259ece26e7/Improving-Uncertainty-Estimations-for-Mammogram-Classification-using-Semi-Supervised-Learning.pdf)

[8] <https://www.ijcnn.org/>

[9] <https://www.tec.ac.cr/hoyeneltec/2020/08/06/inteligencia-artificial-podria-facilitar-uso-radiografias-pecho-diagnosticar-covid-19>

[10] <https://www.tec.ac.cr/hoyeneltec/2019/09/02/cancer-mama-inteligencia-artificial-agilizara-diagnostico>