

IDENTIFICACIÓN AUTOMÁTICA DE CÁNCER DE PIEL APLICANDO MACHINE LEARNING

AUTOMATIC IDENTIFICATION OF SKIN CANCER USING MACHINE LEARNING

MSc. Anderson Smith Florez Fuentes , Ph.D. Rafael Guzmán Cabrera ,
Ph.D. Everardo Vargas Rodriguez 

Universidad de Guanajuato

Doctorado en Ciencias de Ingeniería

Tel.: +52-7295416779, +52 4641239321

E-mail: {as.florezfuentes, guzmanc, evr}@ugto.mx,

Orcid: 0000-0001-6176-1071

Orcid: 0000-0002-9320-7021

Orcid: 0000-0001-5480-3384

Resumen: En la actualidad la implementación del machine learning en la medicina, especial en la oncología, se ha convertido en una herramienta clave para la detección de malformaciones, infecciones, problemas cardiacos, cáncer, entre otros, anomalías del cuerpo humano que al aplicar aprendizaje automático, estos procesos se convierten en herramientas que fortalecen el diagnostico médico. En este trabajo se presenta una revisión de la literatura y los resultados obtenidos de una implementación en Python para llevar a cabo la clasificación de imágenes médicas con presencia de cáncer de piel. Se utilizaron cuatro métodos de aprendizaje distintos: máquinas de vectores de soporte, vecinos más cercanos, redes neuronales y árboles de decisión. Se utilizó la base de datos HAM10000, la cual contiene imágenes médicas manualmente etiquetadas, lo que asegura que el proceso de entrenamiento sea efectivo. En la implementación de una muestra balanceada, la que permite demostrar la eficiencia de la implementación realizada en la identificación del cáncer de piel, esta implementación tiene una gran variabilidad en los resultados, alcanzando valores cercanos al 100% de precisión en algunos casos.

Palabras clave: Cáncer de piel, Machine learning, procesamiento, data.

Abstract: Currently the implementation of machine learning in medicine, especially in oncology, has become a key tool for the detection of malformations, infections, heart problems, cancer, among others, anomalies of the human body that by applying machine learning, these processes become tools that strengthen the medical diagnosis. This work presents a literature review and the implementation with Python to analyze the behavior of the implemented data, performing an execution for four types of models such as support vector machine, k-nearest neighbor, Neural Network and Decision Tree, these models are applied to the HAM10000 data, which contains manually labeled medical images, which ensures that the delivery process is effective. In the implementation of a balanced sample, which allows to demonstrate the efficiency of the implementation of machine learning in the identification of skin cancer, this implementation has a great variability in the results, reaching values up to 100% accuracy in the implementation.

Keywords: Skin cancer, Machine learning, processing, data.

1. INTRODUCCION

El término Machine Learning (ML) fue acuñado por primera vez en 1959 por Arthur Samuel (Hernández, 2020). Su objetivo es desarrollar algoritmos que permitan a los sistemas de cómputo tomar decisiones y aprender de los resultados; es decir, los sistemas serían capaces de aprender a realizar una tarea sin haber sido provisto explícitamente con la programación requerida (Dorado-Díaz et al., 2019; McClendon & Meghanathan, 2015). De igual manera (Mitchel, 2017) define “Un programa de computadora se dice que aprende desde la experiencia E con respecto a alguna clase de tareas T y medida de performance P, si mejora su performance con las tareas en T, con respecto a la medida P, basado en la experiencia E”, expresando de manera clara y puntual como se aprende desde la experiencia.

Al hablar de ML, es indispensable también agregar los conceptos de Aprendizaje Supervisado y No Supervisado, donde (Maisueche Cuadrado, 2019; Russel & Norvig, 2010) describen de manera sencilla el aprendizaje supervisado como, el agente que observa pares de datos de entrada y salida a modo de ejemplo para aprender una función que modele la salida según la entrada, así como también amplían el concepto sobre la resolución de problemas por medio de los algoritmos de regresión y de clasificación. De igual forma describen el aprendizaje no supervisado como el agente encargado de detectar patrones existentes en los datos de entrada sin necesidad de observar la salida. Es decir, en el aprendizaje supervisado se requieren instancias etiquetadas como entrenamiento, que son usadas por el sistema automático para aprender, mientras que en el aprendizaje no supervisado no se requiere de estas instancias.

Por otro lado, el ML también contribuye con el Deep Learning (DL) o aprendizaje profundo, el cual (Deng & Yu, 2019) describen como, una clase de técnicas de aprendizaje automático que explotan muchas capas de procesamiento de información no lineal para la extracción y transformación de características supervisadas o no supervisadas, para el análisis y clasificación de patrones. De igual forma amplían, el DL que trata de aprender múltiples niveles de representación y abstracción que ayudan a dar sentido a datos como imágenes, sonido y texto. También refiere el DL

como una clase bastante amplia de técnicas y arquitecturas de ML, con el sello distintivo de utilizar muchas capas de procesamiento de información no lineal que son de naturaleza jerárquica, el cual se puede cómodamente clasificar en: Redes profundas para el aprendizaje generativo o no supervisado, Redes profundas para el aprendizaje supervisado y Redes profundas híbridas.

La inteligencia artificial y su aplicación en múltiples áreas es una realidad, si analizamos profundamente, se puede encontrar inteligencia artificial, en dispositivos o aplicaciones que utilizamos día a día, como teléfonos móviles, lavadoras, televisores, neveras, parlantes entre otros, y si bien este uso se da en este tipo de dispositivos, de igual manera se da en muchas otras áreas, como: construcción, entidades financieras, casa de apuestas o salud, por mencionar algunos. En años recientes se ha hecho uso de la inteligencia artificial en diversas aplicaciones dada su versatilidad y aplicada correctamente puede generar diagnósticos a enfermedades entre los que destacan: síntomas (Leiva Vazquez et al., 2018), estados de ánimo (Garay et al., 2002), prevención de suicidios por medio de análisis a las redes sociales (Fontecha Pardo et al., 2020), entre muchas otras aplicaciones que se le puede dar a la inteligencia artificial. En este trabajo se presenta una breve revisión de artículos relacionados con el procesamiento de imágenes aplicado en el diagnóstico, tratamiento y seguimiento del cáncer de piel.

(Jofre & Abeldaño, 2021) plantean como hipótesis la posibilidad de que la inteligencia artificial reemplazará al dermatólogo para hacer diagnósticos, aunque la idea no es absurda ni prematura, dado el gran avance que se tiene, los grandes bancos de datos de imágenes capturadas, los diagnósticos ya realizados sobre cada una de las imágenes por especialistas, y unido a estos, la gran evolución que tiene la inteligencia artificial como el caso que plantea Gómez Rivas (Gómez Rivas et al., 2021) en donde hace mención que el entrenamiento quirúrgico ha evolucionado gracias al diseño de aplicaciones que permiten evitar complicaciones y riesgos para el paciente al facilitar no solo la detección temprana, si no, además, la evolución de la enfermedad en el tiempo. Anudado a esto y de manera puntual el DL sobre el campo de la medicina, hace que no

solo los diagnósticos dermatológicos tengan alto grado de certeza, sino también la detección de múltiples enfermedades, como el cáncer de mama (Daneshjou et al., 2021; Koh et al., 2021), depresiones (Abdul Hamid et al., 2021), enfermedades de vías urinarias (Yin et al., 2020), entre otras, que pueden ser diagnosticadas de manera eficiente, por sistemas inteligentes.

Según la Organización Mundial de la Salud (Organización Mundial de La Salud, 2021) el cáncer es la principal causa de muertes en el mundo, llegando solo en el 2020 a casi 10 millones de muertos, siendo el cáncer pulmonar con 1.8 millones de casos registradas durante el 2020, por otro lado el número de casos registrados también es alto, por ejemplo se presentan en el cáncer de mama, con un total de 2.26 millones de casos registrados durante el 2020, seguidos por los cáncer de pulmón, colorrectal, de próstata, de piel y gástrico, de manera puntual el cáncer de piel (no melanoma), alcanzo los 1.20 millones de casos durante este año, es por ende que, al igual que todos los casos de cáncer, es necesario realizar esfuerzos desarrollando aplicaciones que permitan una detección temprana y facilitar una atención adecuada y oportuna.

En este procedimiento se realiza también una implementación con python para poder analizar el comportamiento de las predicciones en cada uno de los modelos, como K-Nearest-Neighbor (kNN), Support vector machine (SVM), Decisión Tree (Tree), Neural Network, modelos que permiten realizar un análisis del comportamiento del sistema ante las imágenes manualmente etiquetadas que se implementan como muestra, para este proceso se tomó una muestra por cada uno de los 7 tipos de cáncer de piel que trae la Data analizada para este trabajo, como pauta inicial.

2. METODOLOGÍA

Para la construcción de este documento de revisión, se realizaron búsquedas en bases de datos de información, con las palabras skin cáncer y cáncer de piel, descartando aquellos resultados que no abarquen estas palabras desde su resultado o que su temática principal no sean los de interés.

De igual forma se realiza la implementación de la data HAM10000 (P et al., 2018), que consta de 10015 imágenes médicas con presencia de cáncer de piel, manualmente etiquetadas y clasificadas por expertos acordes al tipo de cáncer, al ser esta base de datos una colección estándar y disponible de manera gratuita, permite a los investigadores realizar la comparación de resultados y

metodologías. Para este experimento se tienen en cuenta los tipos de cáncer que contienen entre 230 imágenes y 460, se toman 200 imágenes de train y 30 imágenes de test, en el segundo grupo se toman 400 imágenes de train y 60 imágenes de test, para poder analizar el comportamiento de los modelos implementados y verificar la efectividad del entrenamiento en un mayor número de imágenes de train.

3. CÁNCER DE PIEL E INTELIGENCIA ARTIFICIAL

El presente trabajo se centra en analizar el comportamiento de los siguientes siete tipos de cáncer de piel: Melanocytic nevi (nv), Melanoma (mel), Benign keratosis-like lesions (bkl), Basal cell carcinoma (bcc), Actinic keratoses (akiec), Vascular lesions (vas), Dermatofibroma (df) en estos tipos de cáncer se tiene una muestra de 10000 (P et al., 2018) imágenes dermatológicas manualmente etiquetadas, las cuales se encuentran clasificada como se relacionan en la Fig 1., es de aclarar que el número de imágenes que no se tiene en cuenta en este ejercicio es considerable, este procedimiento se realiza para verificar la efectividad del entrenamiento utilizando cantidad imágenes equivalentes tanto para los train como los test.

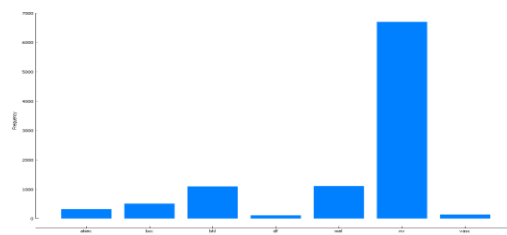


Fig. 1 Muestra de imágenes dermatológicas fuente: propia

Tomando como base la cantidad de se limitó un poco más la búsqueda realizada para obtener documentos de calidad, adicional de poder realizar las mediciones efectivas de cuáles son los estudios realizados sobre el cáncer de piel y cuál ha sido la efectividad de la implementación de los sistemas inteligentes. Es por esto, que se realiza de manera inicial una revisión que nos permite conocer cuáles son los avances del estudio de la detección del cáncer de piel. Con la finalidad de mostrar la complejidad de la tarea de clasificar las imágenes, en la Fig 2 se muestra una imagen de una base de datos de imágenes de cáncer de piel estándar.

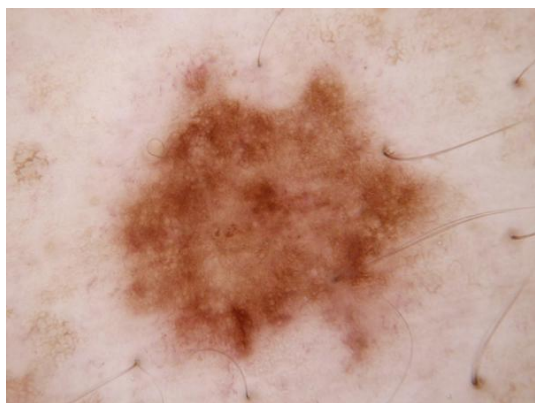


Fig. 2 Cáncer de piel tipo Melanocytic nevi(nv)
fuente:

<https://www.kaggle.com/snginh/skinclassification-ham10k-project1>

Implementación de Sistemas Inteligentes

(Yan et al., 2019) realiza la evaluación de dos conjuntos de lesiones cutáneas, Edimburgo Investigación e Innovación (Dermofit) y el Servicio de Dermatología del Hospital Pedro Hispano (PH2), principalmente se centra en realizar de manera precisa en la extracción de características, implementando diferentes métodos como el GLRLM, LBP y HOG, esto le permite realiza primeramente una clasificación de lesiones benignas y malignas, implementa grupos pequeños de entrenamiento y control, para poder clasificar las lesiones cutáneas, en este proceso implementan clasificadores de conjuntos basados en kNN y SVM y un modelo basado en redes neuronales convolucionales (convolutional neural network CNN), además de la implementación de algoritmos de búsqueda y la implementación de unos algoritmos de optimización de enjambre de partículas (Particle Swarm Optimization, PSO) (Hekler et al., 2019) presentan un trabajo integrado entre la inteligencia artificial y la inteligencia humana, realizando de manera conjunta, para ello utiliza 11,444 imágenes dermatoscópicas, que se dividieron en cinco categorías de diagnóstico, se utilizaron técnicas novedosas de aprendizaje profundo para entrenar a una sola CNN. Luego, 112 dermatólogos de 13 hospitales universitarios alemanes y la CNN entrenada clasificaron de forma independiente un conjunto de 300 lesiones cutáneas verificadas por biopsia en cinco clases. Teniendo en cuenta la certeza de las decisiones, los dos diagnósticos determinados de forma independiente se combinaron en un nuevo clasificador con la ayuda de un método de aumento de gradiente. El criterio de valoración principal del estudio fue la clasificación correcta de las imágenes en cinco

categorías designadas, mientras que el criterio de valoración secundario fue la clasificación correcta de las lesiones como benignas o malignas (clasificación binaria).

En el trabajo presentado por (Han et al., 2020) muestran el alto rendimiento de los algoritmos de aprendizaje profundo, realizan un análisis de 220680 imágenes de 174 trastornos de piel, divididas en 120780 imágenes en la data ASAN, normal posee 48271 imágenes, MED-NODE con 170, Web cuenta con 51459, SNU con 2201 y Edinburgh la cual consta 1300 imágenes y la data SNU que consta de 2201 imágenes; 134 trastornos. Dicho algoritmo les permite predecir con precisión la malignidad, para así poder ayudar a decidir qué tipo de tratamiento que se puede implementar. También muestran como los clasificadores basados en aprendizaje automático superan a los expertos humanos en el diagnóstico de lesiones cutáneas pigmentadas tomadas desde la dermatoscopia.

Dentro de la clasificación de las anomalías de la piel (Saravanan et al., 2020) describe como el crecimiento anormal de piel y la alta exposición del sol es una de las principales causas de cáncer de piel, así como también expresa que un tratamiento efectivo temprano es la mejor forma de curar este cáncer. El método que propusieron utiliza el módulo de desarrollo de visión en LabVIEW para procesar imágenes de dermatoscopia. Los pasos de preprocesamiento, como el filtrado de la mediana y el estiramiento del contraste, se realizan para eliminar el ruido y mejorar la imagen. Adoptaron una segmentación basada en umbrales para segmentar el área afectada.

(Al-masni et al., 2020) realiza la evaluación de 4 clasificadores conocidos en el estado del arte: Inception-v3, ResNet-50, Inception-ResNet-v2 y DenseNet-201, para distinguir entre diferentes enfermedades de la piel, después de realizar la implementación de los clasificadores International Skin Imaging Collaboration (ISIC) y CIU, en los años 2016 y 2017 lograron concluir que el clasificador que mejor se desempeño fue el ResNet-50, arrojando resultados de 81.57% y 89.28% en las dos ISIC aplicadas.

Por otro lado, y teniendo en cuenta el aumento drástico en la tasa de fatalidades por cáncer de piel dadas las variedades de la sintomatología, confirma que la detección temprana en la etapa prematura es necesaria para prevenir la propagación del cáncer. Por ejemplo, Krishna Monika et al. (Krishna Monika et al., 2020) narra la creación de un proyecto que trata sobre la detección y clasificación de varios tipos de cáncer

de piel utilizando herramientas de procesamiento de imágenes y aprendizaje automático. En la etapa de preprocesamiento, las imágenes dermatoscópicas se consideran como entrada. El método de la maquinilla de afeitar sin filo se utiliza para eliminar todas las partículas de pelo no deseadas en la lesión de la piel, luego se utiliza el filtro gaussiano para suavizar la imagen. Para filtrar el ruido y preservar los bordes de la lesión se utiliza un filtro de mediana. Dado que el color es una característica importante en el análisis del tipo de cáncer, la agrupación de k-medias basada en el color se realiza en la fase de segmentación.

4. RESULTADOS

En el proceso de implementación de la muestra, se realiza por medio de python, utilizando una muestra de 115 imágenes, de estas se usaron 100 de entrenamiento y 15 de test, este procedimiento se realiza duplicando el número de imágenes de entrenamiento al igual que las de test, pasando por 200, 400 y 800 imágenes de entrenamiento y 30, 60, 120 imágenes de test, para cada uno de los tipos de cáncer manualmente etiquetados que cuentan con la cantidad de imágenes para pertenecer a cada grupo, que tiene la data mencionada, en este proceso se realiza de manera controlada sobre la detección de cáncer de piel.

En este proceso se aplican kNN, SVM, TREE, Neural Network, cuyos resultados de predicción se complementan con el resultado del análisis de AUC, AC y predicted el modelo de la fig.3, para poder analizar correctamente se realiza para la implementación de las 200 y 400 imágenes de entrenamiento (train) y 30 y 60 imágenes de prueba (test), obteniendo buenos resultados pese a resultados tienen una gran variación, obteniendo los mejores resultados de manera inicial al aplicar el modelo de maquina de soporte vectorial SVM, en donde se puede visualizar en la fig 4, el resultado más alto de predicción lo tiene el melanoma (mel).

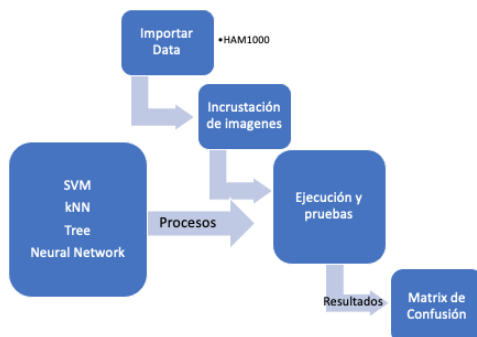


Fig. 3 Implementación del modelo de aprendizaje. Fuente: autor

En este procedimiento se encuentra en la ejecución del segundo grupo de imágenes correspondientes a las 400 imágenes un mayor desempeño en el SVN, arrojando predicciones de hasta el 100% en la ejecución de test and score, para el modelo de la maquina de soporte vectorial SVM. Si bien se presentan resultados de 100% en la ejecución, estos resultados varían acorde a las muestras implementadas y el tipo de muestra seleccionada, esto permite ajustar las formulas para encontrar una precisión en constante mejora, y que la variación entre las muestras no afecte los altos porcentajes de predicción calculados.

Actual	Predicted								Σ
	bcc/test	bcc/train	bkl/test	bkl/train	mel/test	mel/train	nv/test	nv/train	
bcc/test	58.8%	8.1%	0.0%	2.1%	0.0%	0.2%	0.0%	0.4%	60
bcc/train	29.4%	65.4%	25.0%	13.9%	0.0%	4.4%	0.0%	1.6%	400
bkl/test	5.9%	3.3%	50.0%	5.9%	0.0%	2.2%	0.0%	1.8%	60
bkl/train	5.9%	12.7%	0.0%	54.7%	0.0%	25.8%	0.0%	3.0%	400
mel/test	0.0%	3.1%	0.0%	1.6%	100.0%	4.8%	0.0%	1.4%	60
mel/train	0.0%	5.4%	25.0%	18.9%	0.0%	58.8%	0.0%	5.7%	400
nv/test	0.0%	0.2%	0.0%	0.5%	0.0%	0.9%	0.0%	10.7%	60
nv/train	0.0%	1.7%	0.0%	2.4%	0.0%	2.0%	100.0%	75.4%	400
Σ	17	480	4	375	10	458	1	495	1840

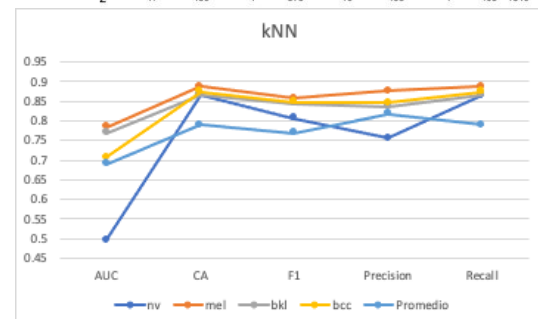


Fig. 4 kNN ejecutado para 4 tipos de cáncer.

Fuente: Autor.

4. CONCLUSIONES

La generación de nuevo conocimiento, por medio de la exploración de la historia escrita o recopilada por otros investigadores, facilita y orienta de una mejor manera la evolución de los saberes, de los procesos experimentales y de la mejora en los procesos ya probados, permitiendo a los nuevos investigadores a construir sobre cimientos y fortalecer de la mejor manera la construcción de nuevas herramientas para el saber, es por ello que la recopilación de implementaciones de inteligencia artificial en el sector salud, y de manera más específica, la implementación en el cáncer de piel, abre las puertas a una investigación con bases para mejorar y/o complementar los trabajos ya realizados.

Esta implementación permitió poder tener de manera sólida una descripción de apoyo a la aplicación de herramientas para verificar la predicción que se genera en la implementación de modelos de inteligencia artificial, generando predicciones que permiten clasificar con un alto grado de precisión, que permita detectar si, las muestras escogidas hacen parte de un cáncer de piel o no, de esta forma brindarle al dermatólogo

o al oncólogo tener un primer análisis de las muestras analizadas.

5. REFERENCIAS

- Abdul Hamid, D. S. Bin, Goyal, S. B., & Bedi, P. (2021). Integration of Deep Learning for Improved Diagnosis of Depression using EEG and Facial Features. *Materials Today: Proceedings*.
<https://doi.org/10.1016/J.MATPR.2021.05.659>
- Al-masni, M. A., Kim, D., & Kim, T. (2020). Computer Methods and Programs in Biomedicine Multiple skin lesions diagnostics via integrated deep convolutional networks for segmentation and classification. *Computer Methods and Programs in Biomedicine*, 190, 105351.
<https://doi.org/10.1016/j.cmpb.2020.105351>
- Daneshjou, R., He, B., Ouyang, D., & Zou, J. Y. (2021). How to evaluate deep learning for cancer diagnostics – factors and recommendations. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, 1875(2), 188515.
<https://doi.org/10.1016/J.BBCAN.2021.188515>
- Deng, L., & Yu, D. (2019). Deep Learning Methods and Applications. *SpringerBriefs in Computer Science*, 31–42. https://doi.org/10.1007/978-981-13-3459-7_3
- Dorado-Díaz, P. I., Sampedro-Gómez, J., Vicente-Palacios, V., & Sánchez, P. L. (2019). Applications of Artificial Intelligence in Cardiology. The Future is Already Here. In *Revista Espanola de Cardiologia* (Vol. 72, Issue 12, pp. 1065–1075). Elsevier Doyma.
<https://doi.org/10.1016/j.recesp.2019.05.016>
- Fontecha Pardo, N. A., Quiroga Sarmiento, J. S., nicolas.fontechap@campusucc.edu.co, & juan.quirogas@campusucc.edu.co. (2020). Implementación de algoritmos de inteligencia artificial enfocados en el análisis de los trastornos del estado de ánimo para prevenir futuros suicidios por medio de la red social Facebook. *Aprendeia*. (22 de 11 de 2020). Obtenido de <https://Aprendeia.Com/Naive-Bayes-Teoria-Machine-Learning/>.
- Garay, N., Abascal, J., & Gardeazabal, L. (2002). Mediación emocional aplicada en sistemas de comunicación aumentativa y alternativa. *Inteligencia Artificial*, 6(16).
<https://doi.org/10.4114/ia.v6i16.743>
- Gómez Rivas, J., Toribio Vázquez, C., Ballesteros Ruiz, C., Taratkin, M., Marengo, J. L., Cacciamani, G. E., Checcucci, E., Okhunov, Z., Enikeev, D., Esperto, F., Grossmann, R., Somani, B., & Veneziano, D. (2021). Inteligencia artificial y simulación en urología. *Actas Urológicas Españolas*.
<https://doi.org/10.1016/J.ACURO.2020.10.012>
- Han, S. S., Park, I., Chang, S. E., Lim, W., & Kim, M. S. (2020). Augmented Intelligence Dermatology : Deep Neural Networks Empower Medical Professionals in Diagnosing Skin Cancer and Predicting Treatment Options for 134 Skin Disorders. *Journal of Investigative Dermatology*, 140(9), 1753–1761.
<https://doi.org/10.1016/j.jid.2020.01.019>
- Hekler, A., Utikal, J. S., Enk, A. H., Hauschild, A., Weichenthal, M., Maron, R. C., Berking, C., Haferkamp, S., Klode, J., Schadendorf, D., Schilling, B., Holland-Letz, T., Izar, B., von Kalle, C., Fröhling, S., Brinker, T. J., Schmitt, L., Peitsch, W. K., Hoffmann, F., ... Thiem, A. (2019). Superior skin cancer classification by the combination of human and artificial intelligence. *European Journal of Cancer*, 120, 114–121.
<https://doi.org/10.1016/J.EJCA.2019.07.019>
- Hernández, J. C. (2020). *PROBLEMAS ÉTICO-JURÍDICOS DE LAS DECISIONES ALGORÍTMICAS Y EL BIG DATA* * Joslay Polanco Medina.
- Jofre, M., & Abeldaño, A. (2021). Artificial intelligence: Will replace the dermatologist making diagnoses? *Piel*, 36(5), 331–335.
<https://doi.org/10.1016/j.piel.2020.06.003>
- Koh, J., Yoon, Y., Kim, S., Han, K., & Kim, E. K. (2021). Deep Learning for the Detection of Breast Cancers on Chest Computed Tomography. *Clinical Breast Cancer*.
<https://doi.org/10.1016/J.CLBC.2021.04.015>
- Krishna Monika, M., Arun Vignesh, N., Usha Kumari, C., Kumar, M. N. V. S. S., & Laxmi Lydia, E. (2020). Skin cancer detection and classification using machine learning. *Materials Today: Proceedings*, 33, 4266–4270.
<https://doi.org/10.1016/J.MATPR.2020.07.366>
- Leiva Vazquez, M., Gonzalez Benitez, N., Hechavarria Hernandez, J., Rivero Peña, Y., & Daher Nader, J. E. (2018). *El diagnóstico de enfermedades desde el Análisis Inteligente de los Datos*.
- Maisueche Cuadrado, A. (2019). *Utilización Del Machine Learning En La Industria 4.0*.
- McClendon, L., & Meghanathan, N. (2015). Using Machine Learning Algorithms to Analyze Crime Data. *Machine Learning and Applications: An International Journal*, 2(1), 1–12.
<https://doi.org/10.5121/mlaij.2015.2101>
- Mitchel, T. M. (2017). Machine Learning. In *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR*.
<https://doi.org/10.1109/ICDAR.2019.00014>
- Organización Mundial de La Salud, O. (2021). *Cáncer*. WHO.
- P, T., C, R., & H, K. (2018). The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific Data*, 5.
<https://doi.org/10.1038/SDATA.2018.161>

- Russel, S., & Norvig, P. (2010). Artificial intelligence, A Modern Approach. In *2010 The 2nd International Conference on Computer and Automation Engineering, ICCAE 2010* (Vol. 4). <https://doi.org/10.1109/ICCAE.2010.5451578>
- Saravanan, S., Heshma, B., Shanofer, A. V. A., & Vanithamani, R. (2020). Skin cancer detection using dermoscope images. *Materials Today: Proceedings*, xxx. <https://doi.org/10.1016/j.matpr.2020.08.388>
- Yan, T., Zhang, L., & Peng, C. (2019). Intelligent skin cancer diagnosis using improved particle swarm optimization and deep learning models. *Applied Soft Computing Journal*, 84, 105725. <https://doi.org/10.1016/j.asoc.2019.105725>
- Yin, S., Peng, Q., Li, H., Zhang, Z., You, X., Fischer, K., Furth, S. L., Fan, Y., & Tasian, G. E. (2020). Multi-instance Deep Learning of Ultrasound Imaging Data for Pattern Classification of Congenital Abnormalities of the Kidney and Urinary Tract in Children. *Urology*, 142, 183–189. <https://doi.org/10.1016/J.UROLOGY.2020.05.019>