

Aprendizaje Automático y Equidad en Dermatología

Disparidades en representación con respecto al tono de piel

Celia Cintas PhD

Investigadora - Kenya Lab

Marzo 2021



Photo: TOTTO RENNA

El equipo

Los trabajos aquí presentados fueron desarrollados gracias a la excelente colaboración entre lxs investigadorxs, estudiantes de maestría y doctorandxs.



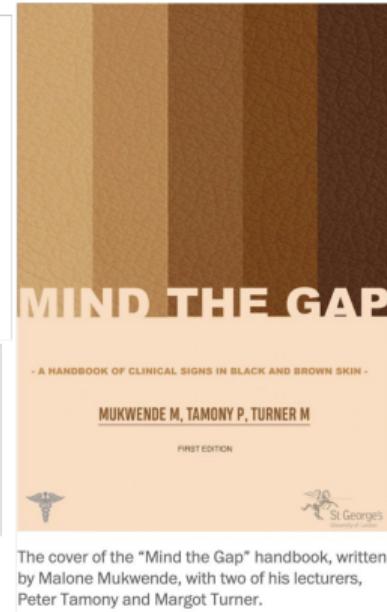
IBM Research
Carnegie Mellon University
Africa College of Engineering
Stanford
University

IBM

Disparidades en Dermatología

- En poblaciones afrodescendientes, el melanoma es comúnmente diagnosticado en etapas tardías [MSL⁺17, WEK⁺11].
- La escasez de imágenes de las manifestaciones cutáneas de COVID-19 en pacientes latinxs y afrodescendientes es un problema, ya que dificulta la identificación de COVID-19 tanto para los dermatólogos como para el público [LJZ⁺20].

The screenshot shows a news article from The New York Times. At the top, there's a navigation bar with three horizontal lines, the text "The New York Times", and an "Account" dropdown. Below the header, the main title is "Dermatology Has a Problem With Skin Color". A subtitle follows: "Common conditions often manifest differently on dark skin. Yet physicians are trained mostly to diagnose them on white skin." The article is categorized under "HEALTH". The author's name is "By USHA LEE MCFARLING @ushamcfarling / JULY 21, 2020". There are also "Responses" and "Report a problem" links.



Algunas preguntas que surgen ...

- ❑ Las imágenes que usamos para **entrenar modelos** están **sesgados respecto a tonos de piel?**
- ❑ Podemos **cuantificar** de forma automática cuál es la **representación** de distintos tonos de piel en **materiales educativos** utilizado para formar dermatólogos y médicos de atención primaria?

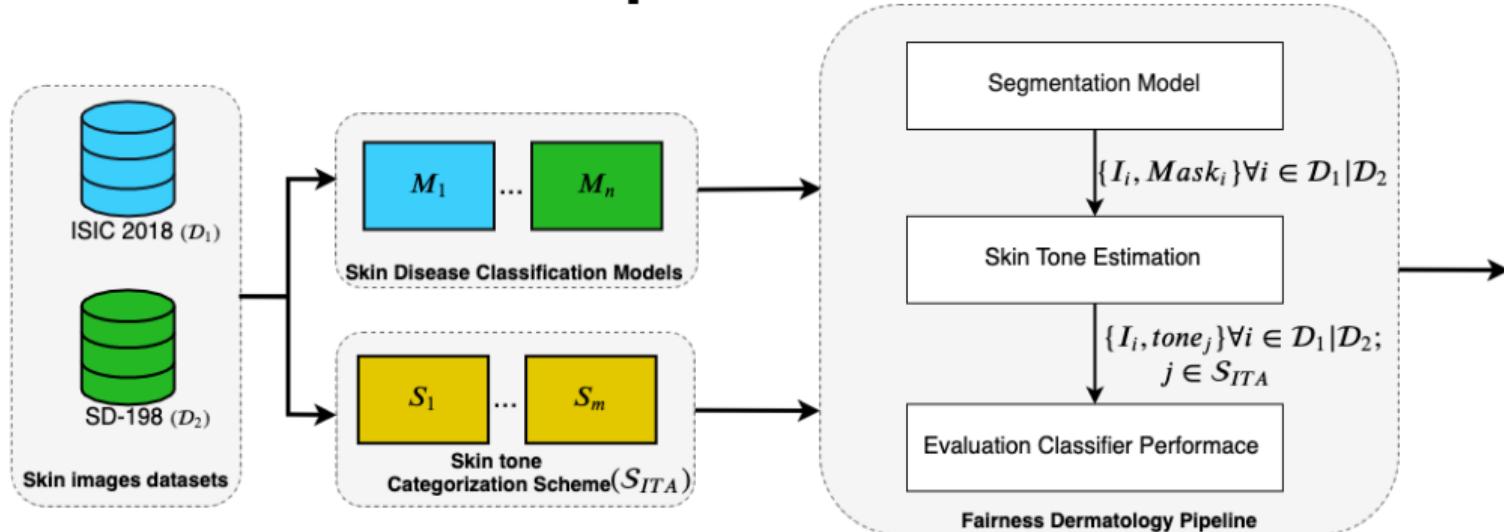


Trabajos relacionados

- Diagnóstico de enfermedades de la piel mediante aprendizaje automático
 - 1 Aprendizaje Automático para el diagnóstico de melanoma supera la performance de dermatólogos [CNP⁺16]
 - 2 ISIC challenges (<https://www.isic-archive.com/>)
- Estudios de inequidad en varias aplicaciones de visión computacional con respecto al tono de piel
 - 1 Análisis automatizado de imágenes faciales para la clasificación de género [BG18]
 - 2 Sistemas de detección de peatones en vehículos autónomos [WHM19]



Cómo evaluamos estos modelos respecto a distintos tonos de piel?



Kinyanjui, N.M., Odonga, T., Cintas, C., Codella, N.C., Panda, R., Sattigeri, P. and Varshney, K.R. Fairness of Classifiers Across Skin Tones in Dermatology. In MICCAI 2020 - International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 320-329).





Advertencia de contenido gráfico

Advertencia de contenido gráfico de enfermedades de piel

Tengan en cuenta que mostraremos **ejemplos de enfermedades de piel** que podrían herir la **sensibilidad** de algunos espectadores.

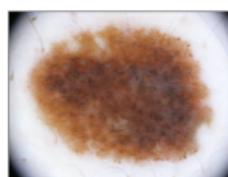
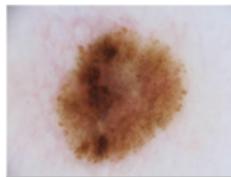
Mencionamos esto, para que los espectadores que así lo deseen se desconecten durante algunos minutos por su propio bienestar.



Qué datos usamos para estudiar los modelos entrenados?

ISIC 2018 [Int18]

- 10015 imágenes dermatoscópicas
- 7 clases de enfermedades
- 2594 imágenes con máscaras de segmentación manuales



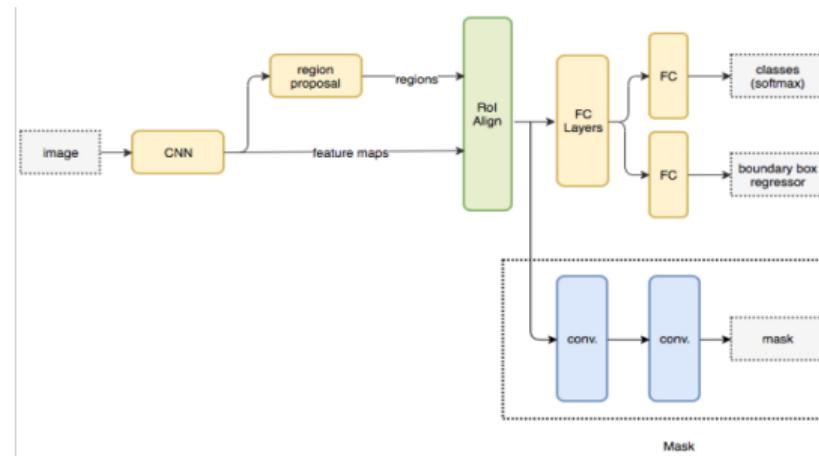
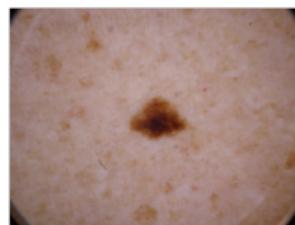
SD-198 [LGL⁺19]

- 6548 imágenes clínicas
- 198 clases de enfermedades
- Sin máscaras de segmentación



Segmentación Automática para obtener la región de la piel sin lesiones

- 1 Adaptamos un modelo pre-entrenado llamado Mask R-CNN [HGDG17].
- 2 Aplicamos umbrales para obtener máscaras binarias y extraer el contorno para ambos conjuntos de datos.



Cómo estimamos el tono de piel de la región sin lesiones?

- 1 Dados los pixeles de regiones sanas, le asignamos un valor asociado al tono de piel.
 - 1 Usamos *Individual Typology Angle* (ITA) [WWdPR15]. Este valor está altamente correlacionado con el índice de melanina.
 - 2 $\text{ITA} = \tan^{-1} \left(\frac{L-50}{b} \right) \times \frac{180}{\pi}$, donde L indica la luminancia y b cuantifica la cantidad de amarillo.
- 2 Finalmente, agrupamos estos valores en categorías [CSD⁺15].

| ITA Range | Skin Tone Category | Abbreviation |
|---|--------------------|--------------|
| $\text{ITA} > 55^\circ$ | Very Light | very_lt |
| $48^\circ < \text{ITA} \leq 55^\circ$ | Light 2 | lt2 |
| $41^\circ < \text{ITA} \leq 48^\circ$ | Light 1 | lt1 |
| $34.5^\circ < \text{ITA} \leq 41^\circ$ | Intermediate 2 | int2 |
| $28^\circ < \text{ITA} \leq 34.5^\circ$ | Intermediate 1 | int1 |
| $19^\circ < \text{ITA} \leq 28^\circ$ | Tanned 2 | tan2 |
| $10^\circ < \text{ITA} \leq 19^\circ$ | Tanned 1 | tan1 |
| $\text{ITA} \leq 10^\circ$ | Dark | dark |

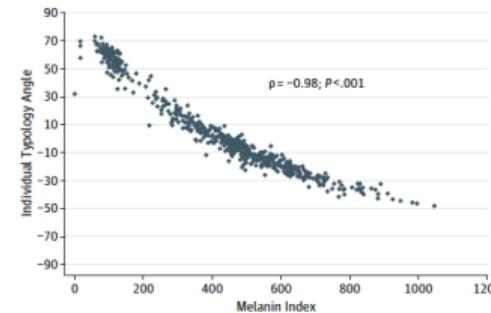


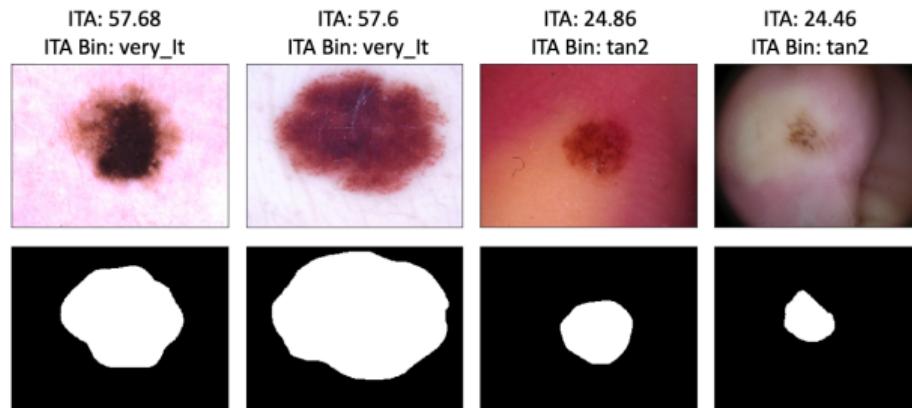
Figure from [WWdPR15].



Resultados

Modelo de segmentación y estimación de tono de piel en ISIC 2018

El modelo Mask R-CNN tiene una precisión **0.956**, una tasa de falsos negativos de **0.024**, y un error absoluto medio en la estimación de los valores ITA de **0.428** grados [KOC⁺19].

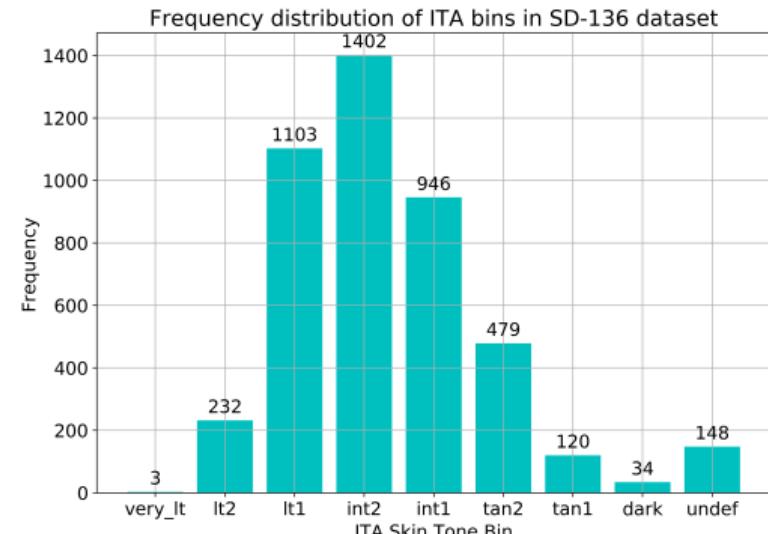
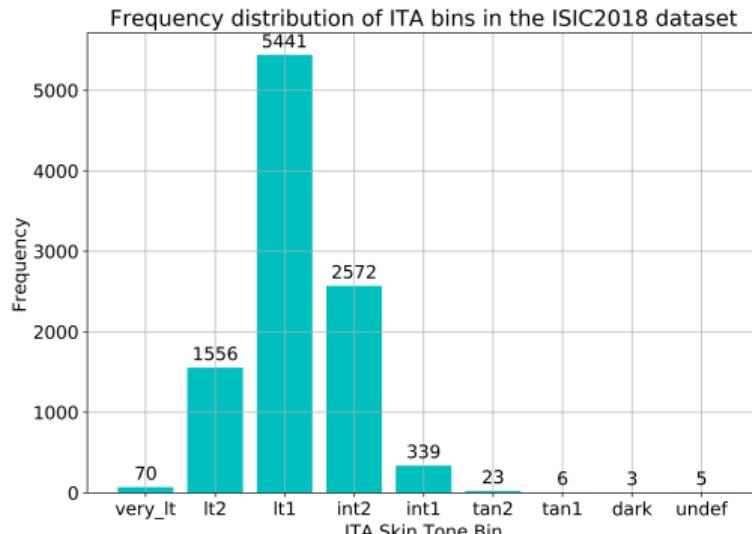


Resultados (Cont.)



Distribución con respecto a tonos de piel

En los conjuntos de datos explorados para ML y dermatología, las poblaciones afrodescendientes están subrepresentados ya que la mayor parte de los datos corresponden a poblaciones caucásicas.





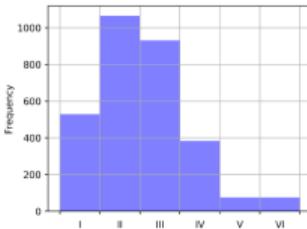
Sabemos con que datos las computadoras están entrenadas .. pero con qué se preparan a los profesionales?



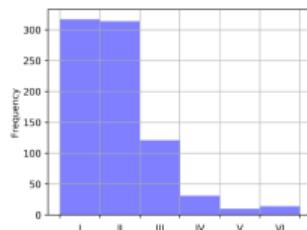
Ahora sabemos que los datos utilizados para **entrenar modelos** en ML están formados **mayoritariamente por población caucásica**.

Podemos **cuantificar** de forma automática cuál es la **representación** de distintos tonos de piel en **materiales educativos** utilizados para formar dermatólogos y médicos de atención primaria?

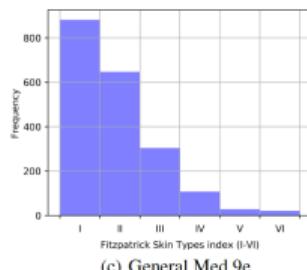
Resultados Preliminares



(a) Bologna 4e



(b) Atlas 8e



(c) General Med 9e

Como validación utilizamos tres libros de dermatología etiquetados por profesionales.



Bologna 4e

Atlas 8e

General Med 9e

Podemos notar la falta de imágenes de piel oscura en todos los libros evaluados

| | Bologna | Atlas | General Med |
|----------|---------|-------|-------------|
| GT | 11.92% | 9.56% | 11.78% |
| Proposed | 17.42% | 6.82% | 7.84% |

Tadesse ,A. G, Kim ,H., Daneshjou ,R., Cintas ,C., Varshney ,K., Adelekun, A., Lipoff ,J., Onyekab, G., Rotemberg, V., Zou, J. Automated Evaluation of Representation in Dermatology Educational Materials. In AAAI 2021 Workshop: Trustworthy AI for Healthcare.



Mis 2 centavos para crear soluciones con ML

- 1 Es crucial que **los grupos que desarrollan soluciones** tecnológicas para sectores como educación, salud, etc, sean **grupos interdisciplinarios**.
- 2 Los investigadorxs y desarrolladorxs tienen que ser **tan (o más) diversos** que **sus usuarios** finales.
- 3 Los modelos a utilizar en producción deben **explicitar** en qué contexto funcionan, ser **transparentes**, clarificar qué **sesgos fueron evaluados** y cuáles son las **técnicas de mitigación** utilizadas.



Gracias, Asante, Thanks!



IBM Research

Carnegie Mellon University
Africa College of Engineering



celia.cintas@ibm.com

@RTFMCelia

IBM

References I

-  Joy Buolamwini and Timnit Gebru, *Gender shades: Intersectional accuracy disparities in commercial gender classification*, Proc. Conf. Fair. Account. Transp., February 2018, pp. 77–91.
-  Noel C. F. Codella, Quoc-Bao Nguyen, Sharath Pankanti, David A. Gutman, Brian Helba, Allan C. Halpern, and John R. Smith, *Deep learning ensembles for melanoma recognition in dermoscopy images*, IBM J. Res. Dev. 61 (2016), no. 4/5, 5.
-  Giuseppe R. Casale, Anna Maria Siani, Henri Diémoz, Giovanni Agnesod, Alfio V. Parisi, and Alfredo Colosimo, *Extreme UV index and solar exposures at Plateau Rosà (3500 m a.s.l.) in Valle d'Aosta Region, Italy*, Sci. Total Environ. 512–513 (2015), 622–630.
-  Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross B Girshick, *Mask r-cnn. corr abs/1703.06870 (2017)*, arXiv preprint arXiv:1703.06870 (2017).



References II

-  International Skin Imaging Collaboration, *ISIC 2018: Skin lesion analysis towards melanoma detection*, 2018, Available:
<https://challenge2018.isic-archive.com/>.
-  Newton M Kinyanjui, Timothy Odonga, Celia Cintas, Noel CF Codella, Rameswar Panda, Prasanna Sattigeri, and Kush R Varshney, *Estimating skin tone and effects on classification performance in dermatology datasets*, arXiv preprint arXiv:1910.13268 (2019).
-  J. Lin, Z. Guo, D. Li, X. Hu, and Y. Zhang, *Automatic classification of clinical skin disease images with additional high-level position information*, 2019 Chinese Control Conference (CCC), 2019, pp. 8606–8610.



References III

-  JC Lester, JL Jia, L Zhang, GA Okoye, and E Linos, *Absence of skin of colour images in publications of covid-19 skin manifestations*, British Journal of Dermatology (2020).
-  Krishnaraj Mahendararaj, Komal Sidhu, Christine S. M. Lau, Georgia J. McRoy, Ronald S. Chamberlain, and Franz O. Smith, *Malignant melanoma in African–Americans: A population-based clinical outcomes study involving 1106 African–American patients from the surveillance, epidemiology, and end result (SEER) database (1988–2011)*, Medicine 96 (2017), no. 15, e6258.

References IV

-  Xiao-Cheng Wu, Melody J. Eide, Jessica King, Mona Saraiya, Youjie Huang, Charles Wiggins, Jill S. Barnholtz-Sloan, Nicolle Martin, Vilma Cokkinides, Jacqueline Miller, Pragna Patel, Donatus U. Ekwueme, and Julian Kim, *Racial and ethnic variations in incidence and survival of cutaneous melanoma in the United States, 1999-2006*, J. Am. Acad. Dermatol. 65 (2011), no. 5, S26.e1–S26.e13.
-  Benjamin Wilson, Judy Hoffman, and Jamie Morgenstern, *Predictive inequity in object detection*, arXiv:1902.11097, February 2019.
-  Marcus Wilkes, Caradee Y. Wright, Johan L. du Plessis, and Anthony Reeder, *Fitzpatrick skin type, individual typology angle, and melanin index in an African population*, JAMA Dermatol. 151 (2015), no. 8, 902–903.

