

Advancing Equity in Health AI: Improving Dermatological Image Datasets for Fair and Inclusive Healthcare Solutions

Published in **Nature Scientific Data** (2025)



Kumar Abhishek



Aditi Jain



Ghassan Hamarneh

2025 CS Diversity Award Presentation Day

Content Warning

Contains clinical images of **skin conditions**
that some viewers may find sensitive.

AI for Skin Disease Diagnosis

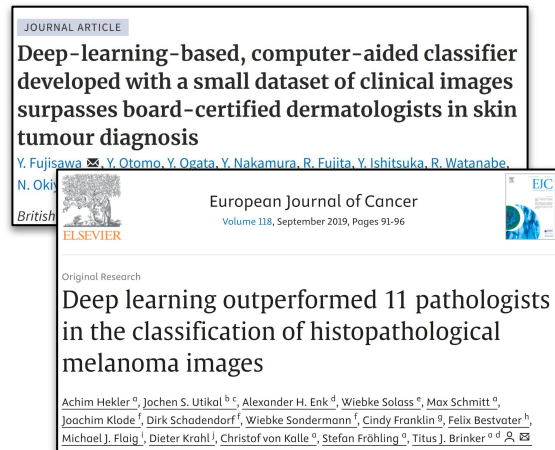
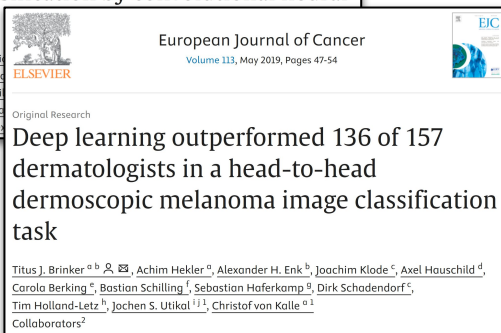
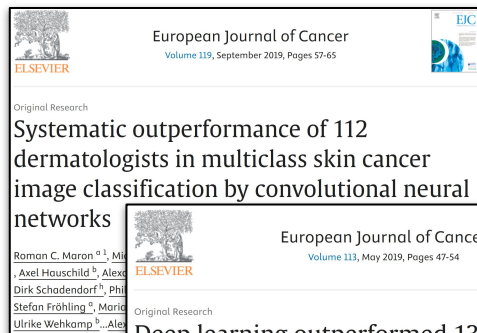
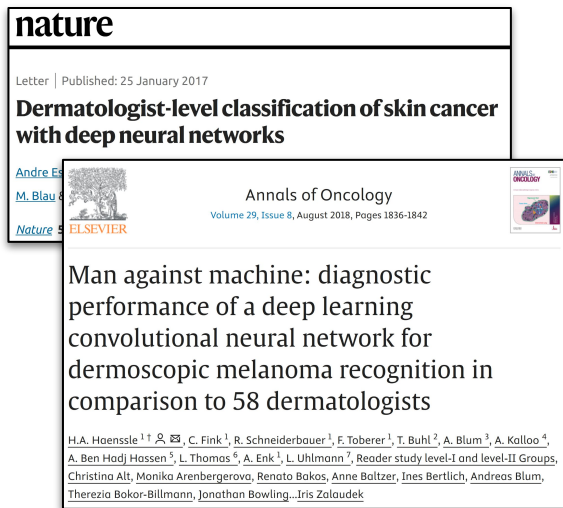
In studied populations, skin diseases are the **most common reason** for GP visits^[1].

4th leading cause of non-fatal morbidity worldwide^[2], yet ~**3 billion people lack access** to dermatological care^[3].

AI for Skin Disease Diagnosis

In studied populations, skin diseases are the **most common reason** for GP visits^[1].

4th leading cause of non-fatal morbidity worldwide^[2], yet ~**3 billion people lack access** to dermatological care^[3].



Variability in Skin Disease Appearance Affects Diagnosis

- Skin conditions may **appear different** across different skin tones.



Kawasaki Disease^[5]



Erythema Annulare Centrifugum^[6]

Variability in Skin Disease Appearance Affects Diagnosis

- Skin conditions may **appear different** across different skin tones.
- Doctors and AI perform poorly on darker-skinned patients.
 - 31%-61% **melanoma misdiagnosis on dark skin**, compared to 7%-13% on light skin^[3].
 - *"state-of-the-art dermatology AI models exhibit substantial limitations ... particularly on dark skin tones and uncommon diseases"*^[4].



Kawasaki Disease^[5]



Erythema Annulare Centrifugum^[6]

Variability in Skin Disease Appearance Affects Diagnosis

- Skin conditions may **appear different** across different skin tones.
- Doctors and AI perform poorly on darker-skinned patients.
 - 31%-61% **melanoma misdiagnosis on dark skin**, compared to 7%-13% on light skin^[3].
 - *"state-of-the-art dermatology AI models exhibit substantial limitations ... particularly on dark skin tones and uncommon diseases"*^[4].
- Errors in datasets \Rightarrow models \Rightarrow healthcare.



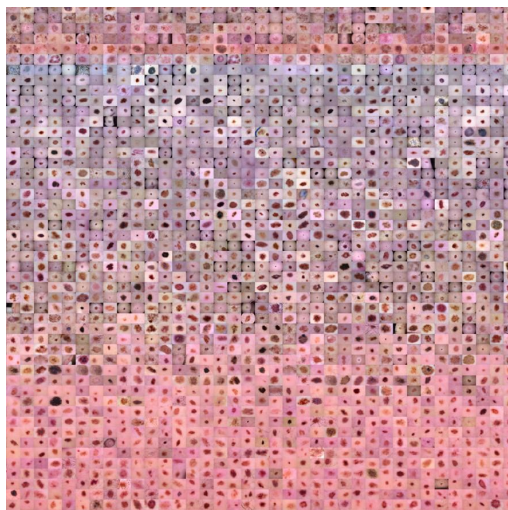
Kawasaki Disease^[5]



Erythema Annulare Centrifugum^[6]

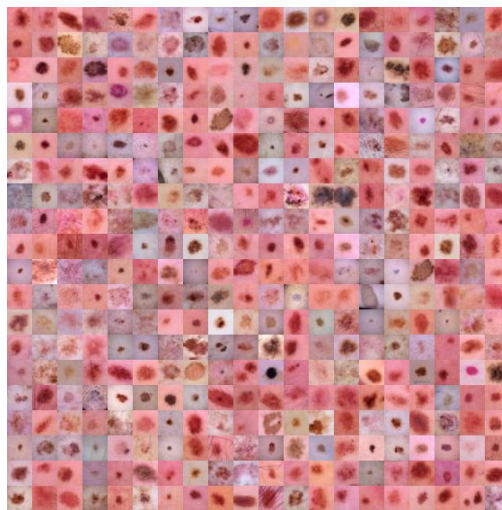
3 Widely Used Skin Image Datasets

HAM10000



~3500 citations

DermaMNIST



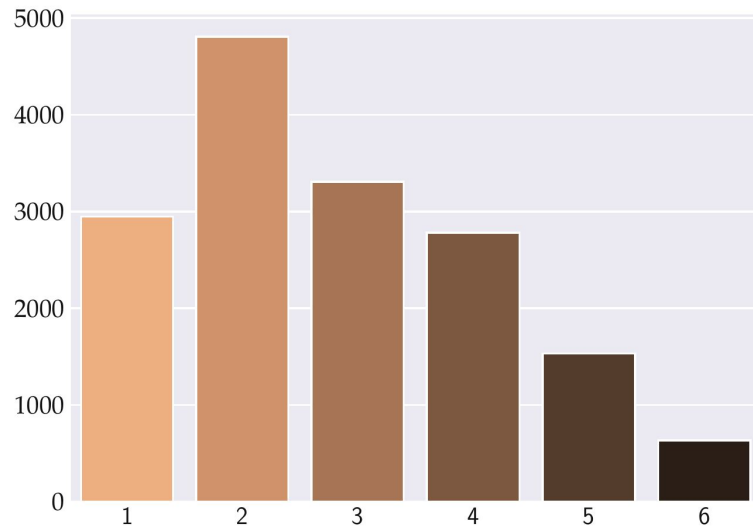
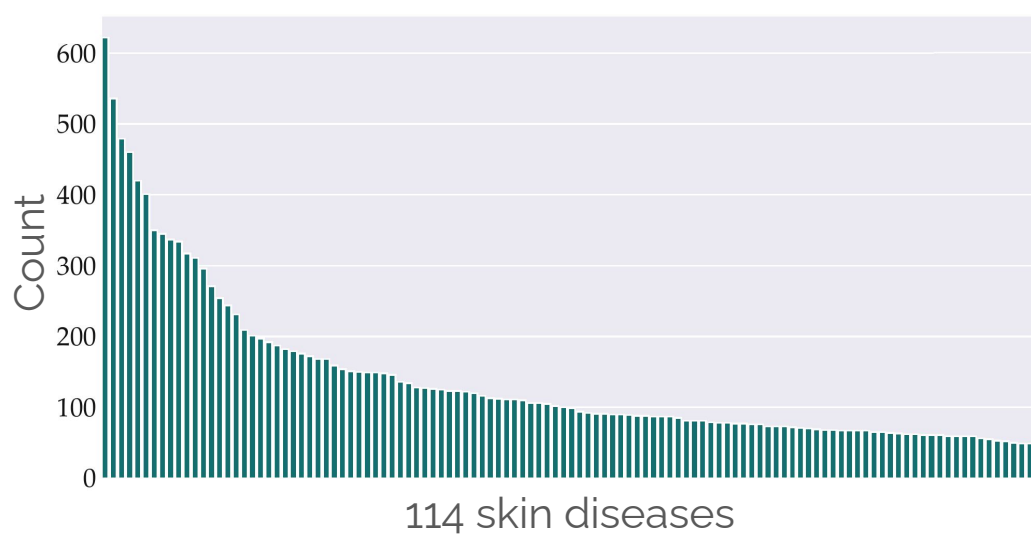
~1200 citations

Fitzpatrick17k

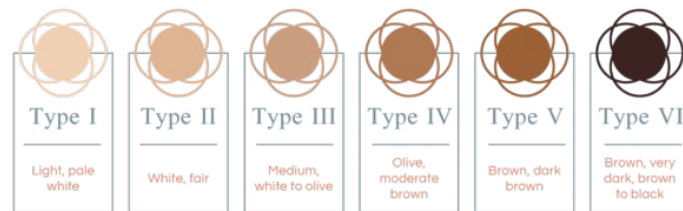


~250 citations

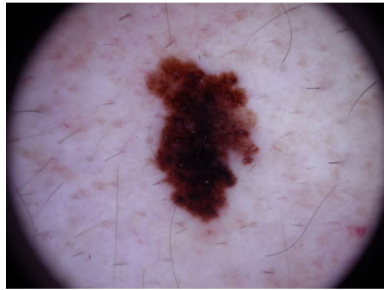
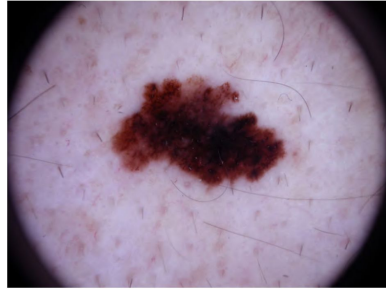
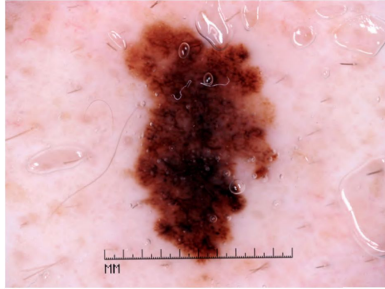
Underrepresented Diseases and Skin Tones



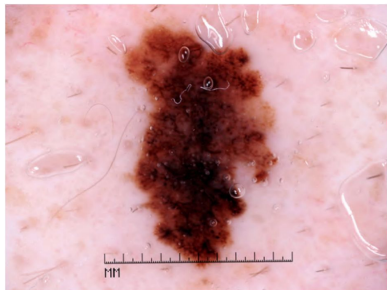
THE FITZPATRICK SCALE



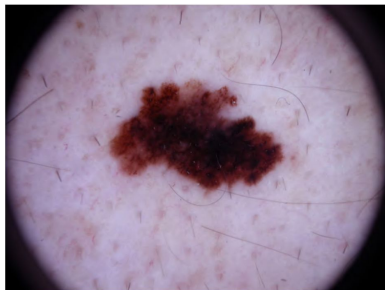
Metadata-based Analysis



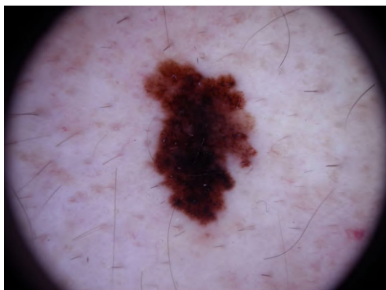
Metadata-based Analysis



HAM_0007343



HAM_0007343



HAM_0007343



HAM_0002364



HAM_0002364

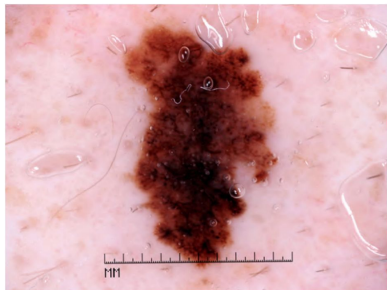


HAM_0002364

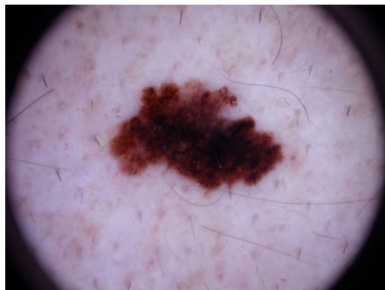


HAM_0002364

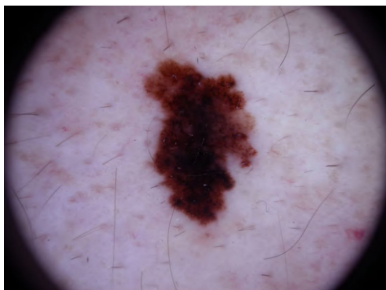
Metadata-based Analysis



HAM_0007343



HAM_0007343



HAM_0007343



HAM_0002364



HAM_0002364

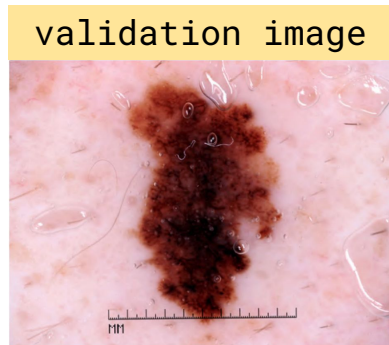


HAM_0002364



HAM_0002364

Duplicate Images Cause Data Leakage



HAM_0007343



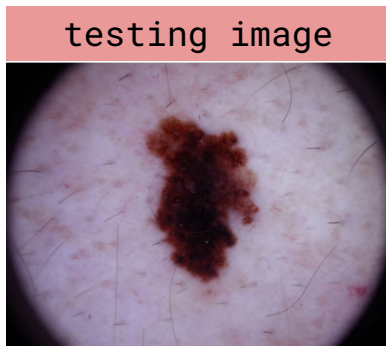
HAM_0007343



HAM_0002364



HAM_0002364



HAM_0007343

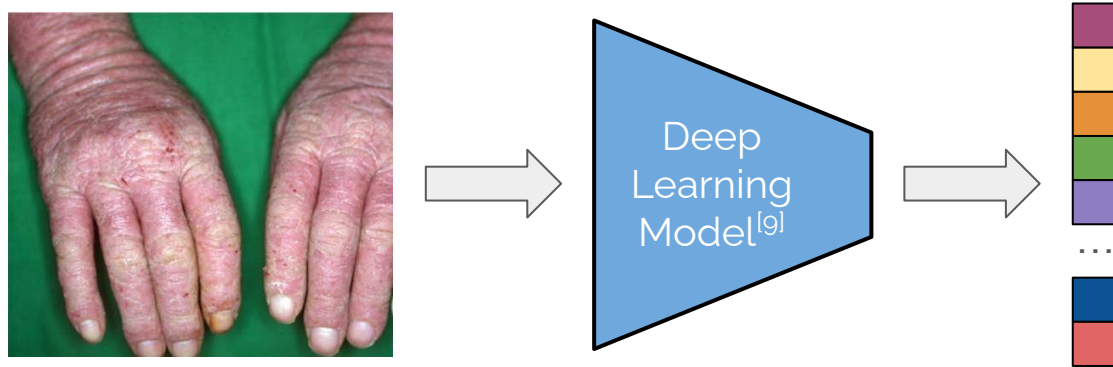


HAM_0002364

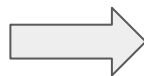
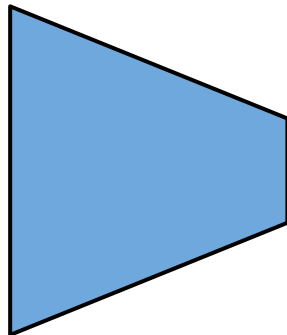
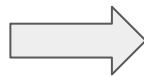
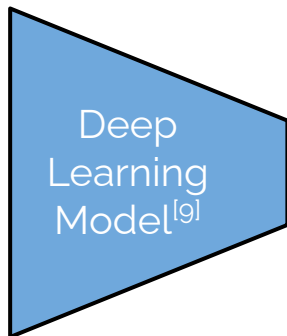


HAM_0002364

Content-based Analysis: Image Embeddings



Content-based Analysis: Image Embeddings

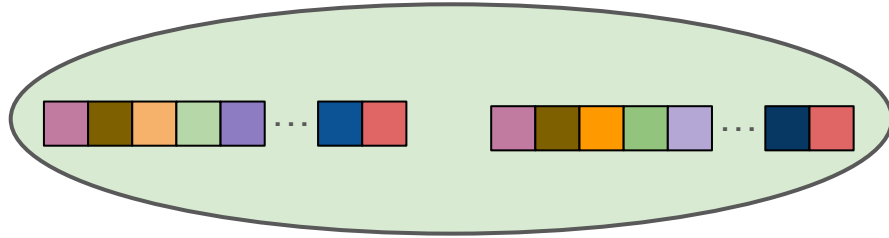


Calculate
embeddings for
all the images

Detecting Duplicate Pairs

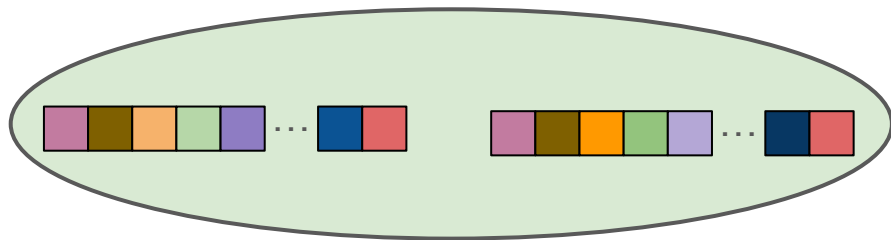


Detecting Duplicate Pairs



Pairwise embedding similarity

Detecting Duplicate Pairs

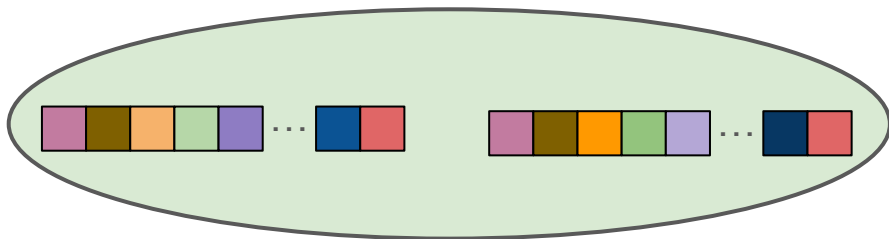


Pairwise embedding similarity



Different lighting but the same lesion

Detecting Duplicate Pairs



Different lighting but the same lesion

Different crops/zooms



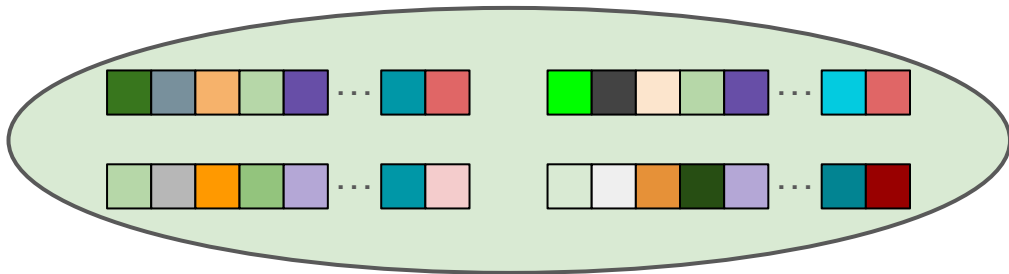
Mirrored images



Different lighting + crop

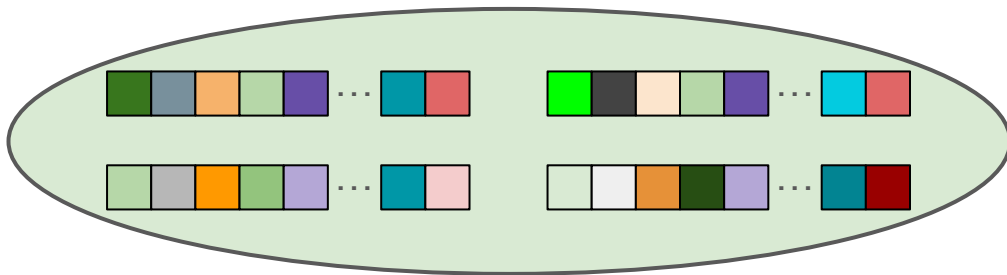


Detecting Duplicate Clusters

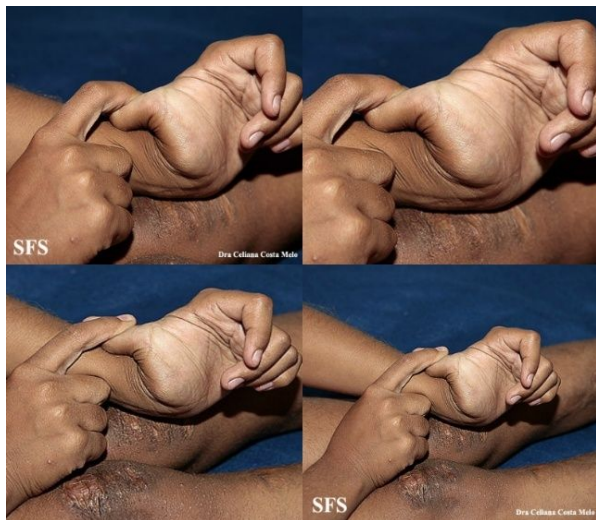


Cluster embeddings based on similarity

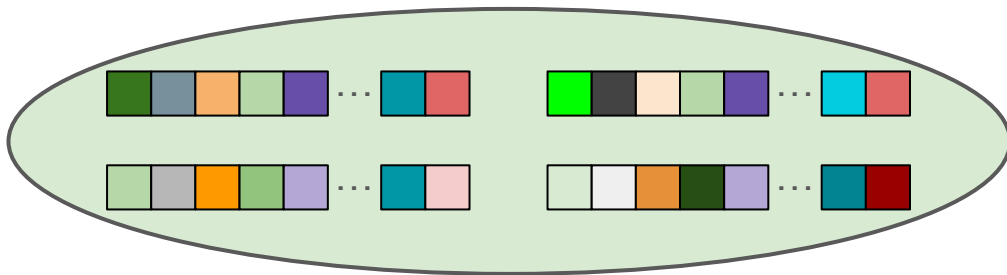
Detecting Duplicate Clusters



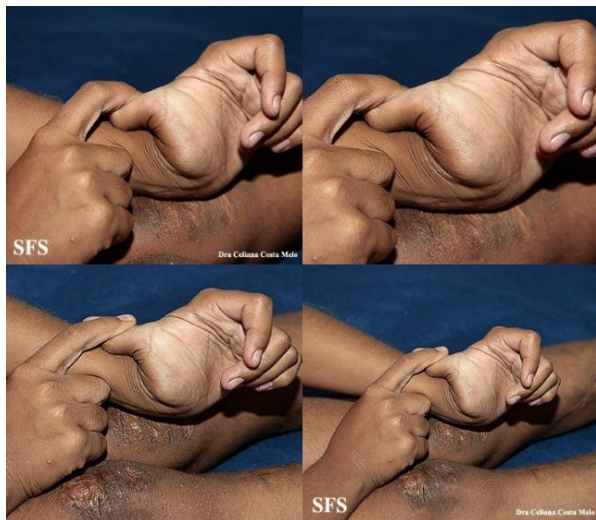
Cluster embeddings based on similarity



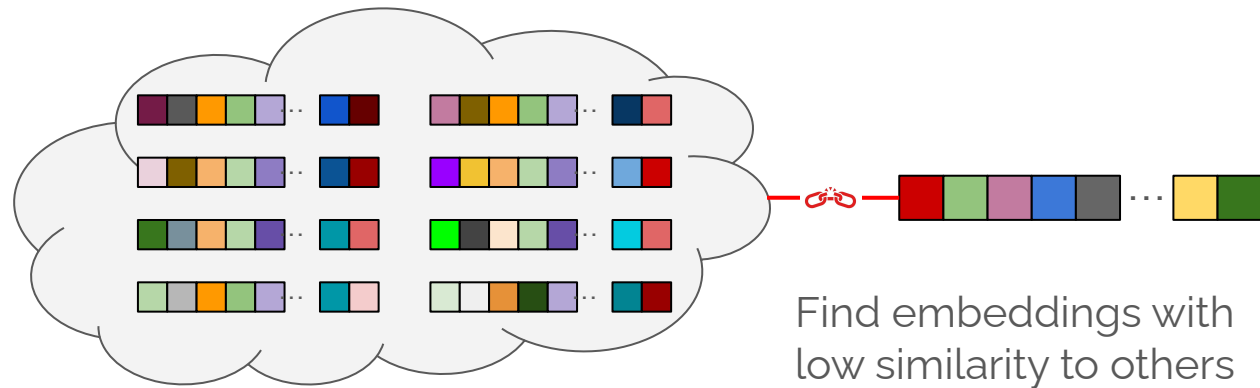
Detecting Duplicate Clusters



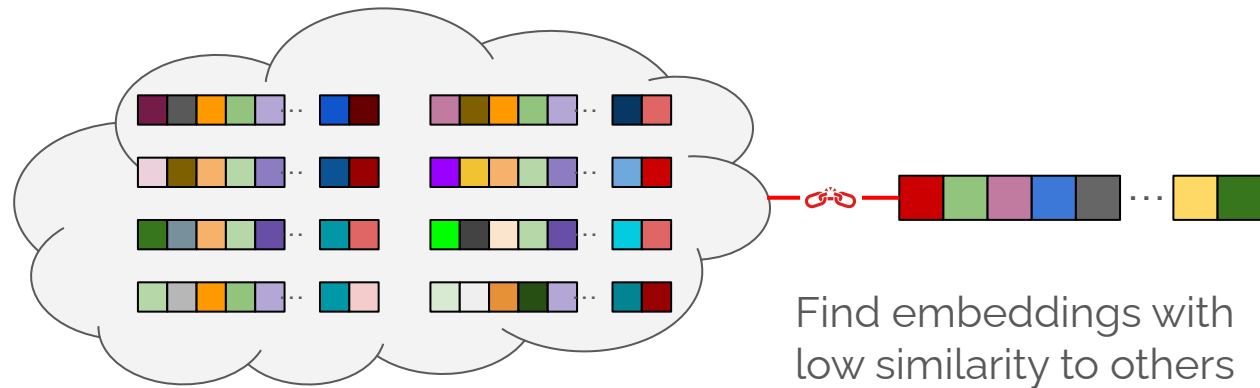
Cluster embeddings based on similarity



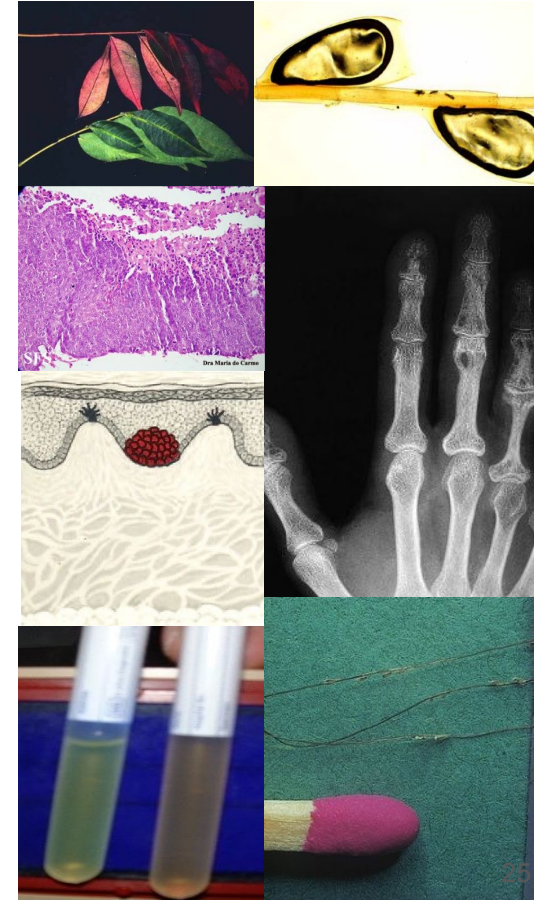
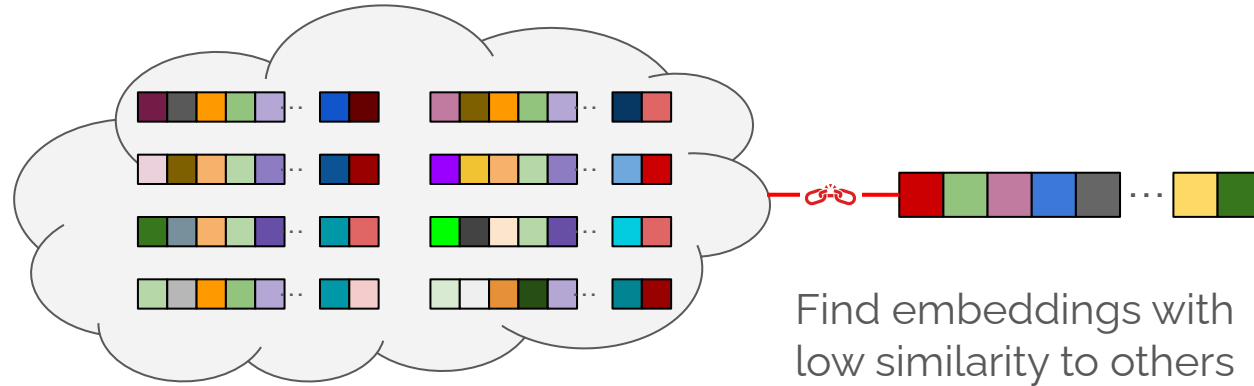
Discovering Outliers



Discovering Outliers



Discovering Outliers



Label Errors Impact Equity & Inclusivity



FST: 4



FST: 1



FST: 4



FST: 1



Label Errors Impact Equity & Inclusivity



FST: 4



FST: 1



FST: 4



FST: 1



pustular
psoriasis

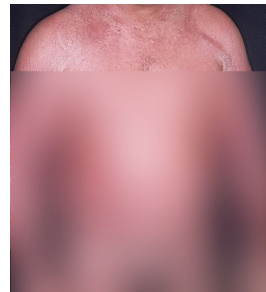


psoriasis

FST: 4



FST: 1



eczema



allergic contact
dermatitis

FST: 3



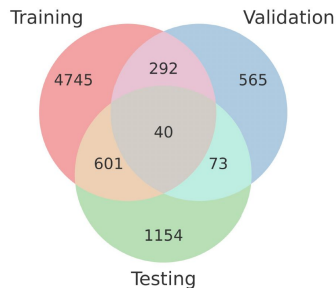
FST: 1



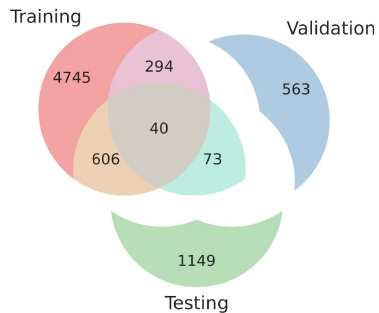
Fixing Dataset Issues

Clean data partitioning

DermaMNIST

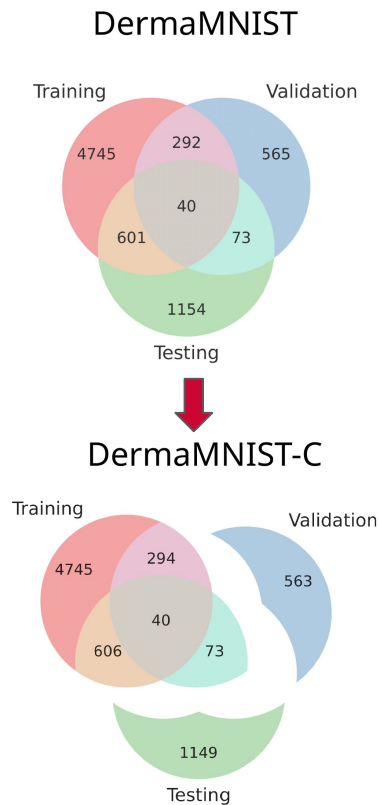


DermaMNIST-C



Fixing Dataset Issues

Clean data partitioning



Removing duplicates and outliers

Detect all duplicates.



If any label conflicts, remove all. Else, keep one.



Remove outlier images.



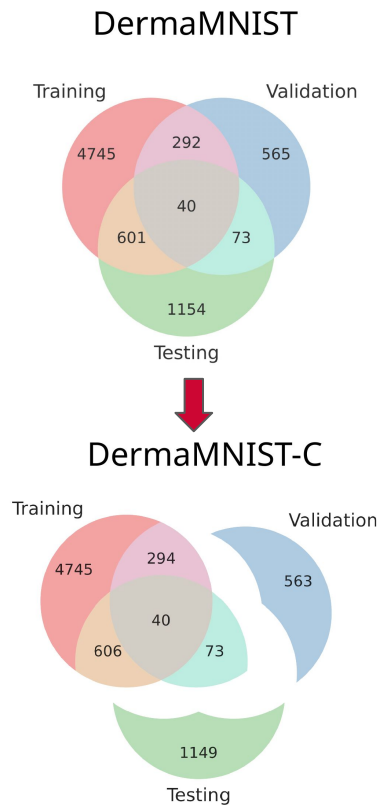
Disjoint train-valid-test sets.



#images
16577 -> 11394

Fixing Dataset Issues

Clean data partitioning



Removing duplicates and outliers

Detect all duplicates.



If any label conflicts, remove all. Else, keep one.



Remove outlier images.



Disjoint train-valid-test sets.



#images
16577 -> 11394

**New benchmarks
for future research.**

Analysis Pipeline is Public

Corrected-Skin-Image-DatasetsPublic

UnpinUnwatch1Fork0Star5

main

Go to fileCode

kakumarabhishekAdd publication detailsac655ac · 2 months ago

DermaMNIST	Update README.md	11 months ago
Fitzpatrick17k	Update README.md	11 months ago
.gitignore	Update README	11 months ago
LICENSE	Update README	11 months ago
README.md	Add publication details	2 months ago

README

Apache-2.0 license

PaperWebsiteZenodo DatasetsLicenseApache 2.0code styleblack

Investigating the Quality of DermaMNIST and Fitzpatrick17k Dermatological Image Datasets

This repository contains the code accompanying our paper titled "[Investigating the Quality of DermaMNIST and Fitzpatrick17k Dermatological Image Datasets](#)", published in Nature Scientific Data.

Repository Structure

The repository is structured as:

About

Data quality analysis of DermaMNIST (MedMNIST), HAM10000, and Fitzpatrick17k datasets

[derm.cs.sfu.ca/critique/](#)

benchmarkdeep-learningpytorchmedical-imagingmnistclassificationdatasetsmedical-image-computingdermatologydata-qualityfitzpatrickcomputer-aided-diagnosisdata-leakagemelanoma-detectionham10000skin-disease-classificationmedmnistdata-duplication

ReadmeApache-2.0 licenseActivity5 stars1 watching0 forks

Languages

jupyter Notebook83.5%

zenodo

Published June 11, 2024 | Version v1.1DatasetOpen

Investigating the Quality of DermaMNIST and Fitzpatrick17k Dermatological Image Datasets

Abhishek, Kumar (Contact person)¹; Jain, Aditi²; Hamarneh, Ghassan¹

Investigating the Quality of DermaMNIST and Fitzpatrick17k Dermatological Image Datasets

Kumar Abhishek¹, Aditi Jain², Ghassan Hamarneh¹

¹School of Computing Science, Simon Fraser University, Burnaby, BC, Canada
²Department of Mathematics, Indian Institute of Technology, Delhi, India

Nature Scientific Data 12 (1), 196 (2025)

Paper

arXiv

Zenodo Datasets

Code

DermaMNIST Analysis

- DermaMNIST-C:** A corrected version of DermaMNIST that does not have any data leakage. Downloadable 28×28 dataset, 224×224 dataset, and the metadata CSV file.
- DermaMNIST-E:** An extended and arguable more challenging version of DermaMNIST that is almost the same as the ISIC 2018 Challenge dataset except the images are resized and the "easter egg" image is removed from the test partition. Downloadable 28×28 dataset, 224×224 dataset, and the metadata CSV file.

Data Analysis Reports: <https://derm.cs.sfu.ca/critique>

Duplicate Pairs Report

Showing pairs of similar images

Each row shows a pair of duplicates. The images visualized are the "From" image, the "To" image, and the "Blended" image. The "Blended" image is the output of α -blending of the two images with $\alpha = 0.5$.



Info	
Similarity	1.0
From	/po-ac_183/po-ac_f3_130_0e2a24d2.jpg
To	/po-ac_183/po-ac_f3_106_5041af68.jpg



Info	
Similarity	1.0
From	/ur_151/ur_f1_34_0455b31f.jpg
To	/st_67/st_f1_16_fcb25afb.jpg

https://derm.cs.sfu.ca/critique/Fitzpatrick17k/duplicate_pairs.html

Duplicate Clusters Report

Showing clusters of similar images

Each row shows a cluster of duplicates along with the cluster ID, the number of images in the cluster, and the intra-cluster similarity.



Info	
Cluster ID	11357
num_images	3
mean_similarity	0.9826



Info	
Cluster ID	2234
num_images	3
mean_similarity	0.9726

https://derm.cs.sfu.ca/critique/Fitzpatrick17k/duplicate_clusters.html

Thank you!



kabhishe@sfu.ca

References

1. Schofield et al., "Skin conditions are the commonest new reason people present to general practitioners in England and Wales", *British Journal of Dermatology*, 2011.
2. Seth et al., "Global Burden of Skin Disease: Inequities and Innovations", *Current Dermatology Reports*, 2017.
3. Lyman et al., "A dermatological questionnaire for general practitioners in England with a focus on melanoma; misdiagnosis in black patients compared to white patients", *Journal of the European Academy of Dermatology and Venereology*, 2016.
4. Daneshjou et al., "Disparities in dermatology AI performance on a diverse, curated clinical image set", *Science Advances*, 2022.
5. Black & Brown Skin, <https://www.blackandbrownskin.co.uk/further-resources>.
6. DermNet, <https://dermnetnz.org/>.
7. Pakzad et al., "CIRCLE: Color Invariant Representation Learning for Unbiased Classification of Skin Lesions", *European Conference on Computer Vision (ECCV) ISIC Skin Image Analysis Workshop*, 2022.
8. Fitzpatrick Skin Type Chart, <https://www.tribecamedspa.com/fitzpatrick-skin-type-chart/>.
9. fastdup, <https://github.com/visual-layer/fastdup>.
10. What's Your Type?, <https://suttonderm.com/blog/whats-your-type/>.