The Quest for Reliable Melanoma Detection

7 Years at the Vanguard of Skin Lesion Analysis

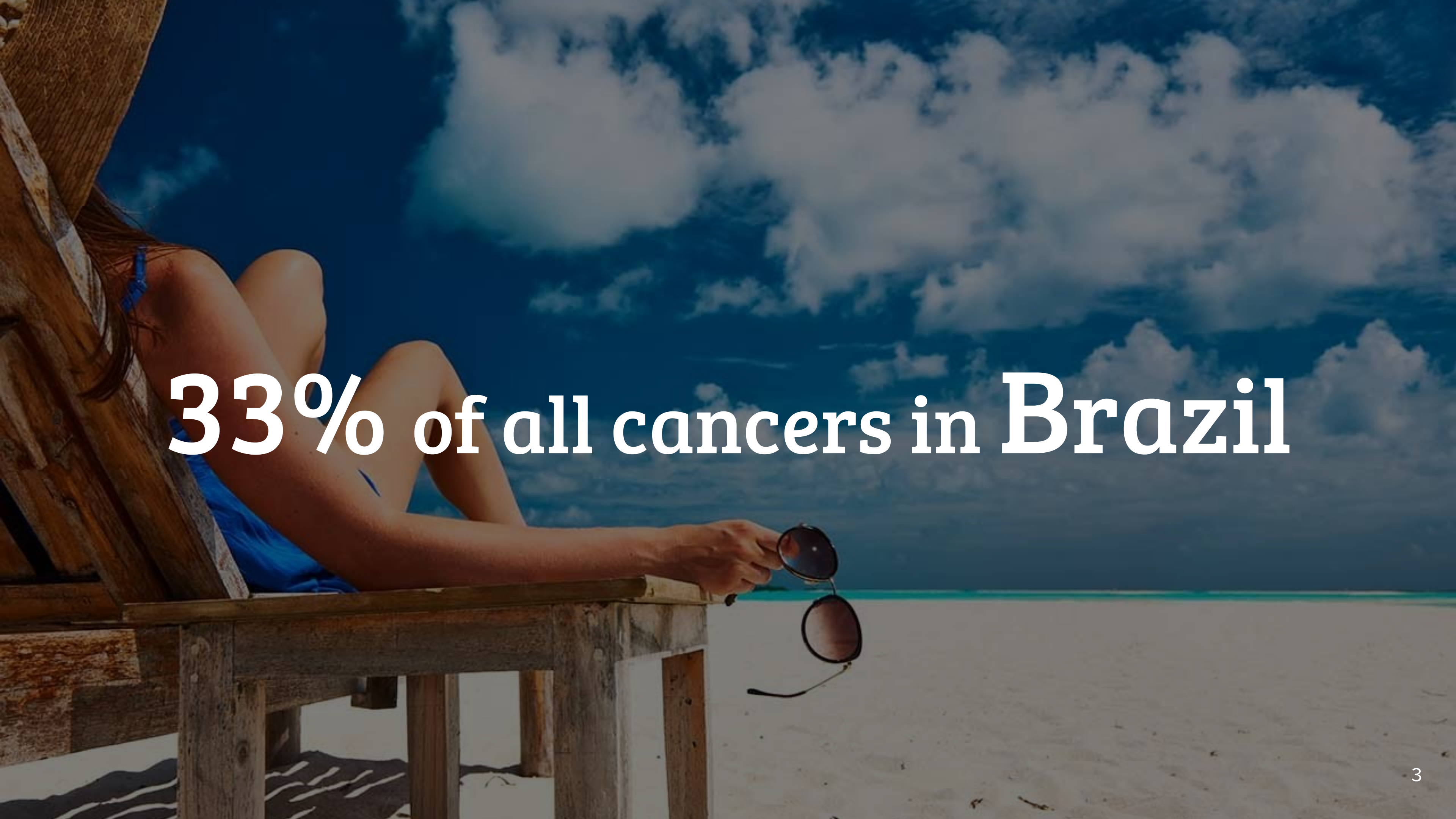
Sandra Avila

Institute of Computing
University of Campinas, Brazil

ISIC Workshop, CVPR 2021

Skin Cancer

Why do we care?


33% of all cancers in Brazil

RECOD Lab.

melanoma research
7 years 2014–2021


Brazil
2016
#GoogleAwardsLatAm
Research at Google
Research Awards for Latin America 2016
Reliable Automated Melanoma Screening for the Real World
Eduardo Alves do Valle Junior
Michel Silva Fornaciali
Universidade Estadual de Campinas
Summary


#JornaldaCultura
#JornaldaCultura
SAÚDE
INTELIGÊNCIA ARTIFICIAL: AUXÍLIO NO DIAGNÓSTICO
IAS DO DESCASO GANANCIOSO QUE CURTE A IMPUNIDADE DAS LEIS BRASILEIRAS #JC


CIDADES CORREIO POPULAR A7
Técnica agiliza diagnóstico de câncer
Software desenvolvido na Unicamp atinge precisão de 86% na detecção do câncer de pele
Brasil registra 6.260 novos casos por ano, segundo Instituto
Google premia estudo pelo 4º ano consecutivo


O ESTADO DE S. PAULO
SEGUNDA-FEIRA, 13 DE JANEIRO DE 2020 A13
Metrópole
Belo Horizonte
Polícia apura demissão em cervejaria suspeita de contaminação. Pág. A14
Saúde. Projeto de cientistas da Unicamp treina computadores para identificar melanoma; taxa de acerto do algoritmo é de 86%, ante 67% na avaliação de dermatologistas. Ideia é usar algoritmo para compreender padrões de malignidade não percebidos pelos humanos
Inteligência artificial pode diagnosticar tumor de pele que até médico não vê
Fabiana Cambricoli

Systematic Evaluation

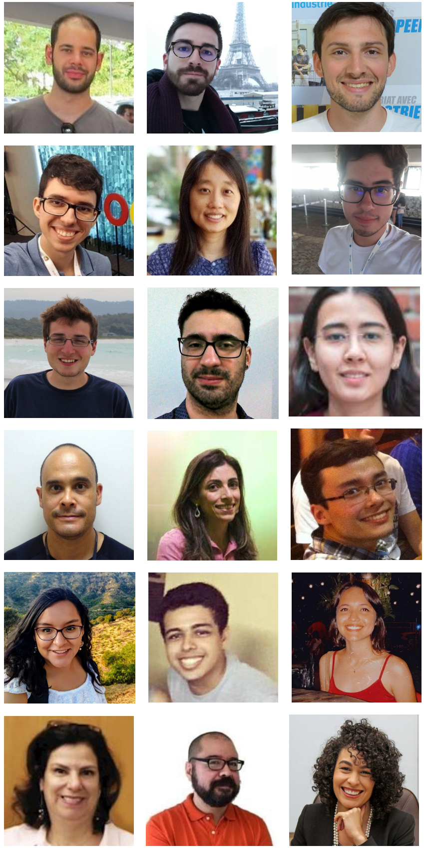
Classification

Segmentation

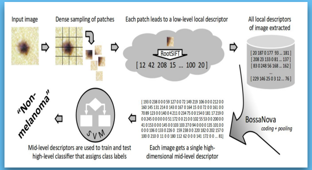
Synthesis

Debiasing

Data Augmentation

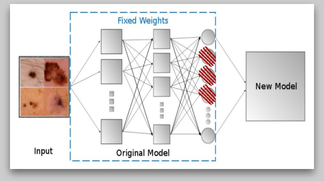


2014



Statistical learning approach for robust melanoma screening

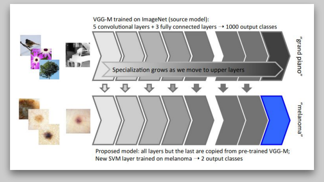
2015



Towards robust melanoma screening: A case for enhanced mid-level features

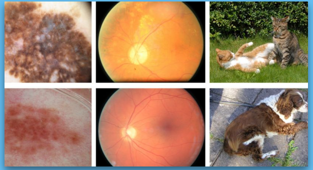
Transfer schemes for deep learning in image classification

2016



Towards automated melanoma screening: Proper computer vision & reliable results

2017

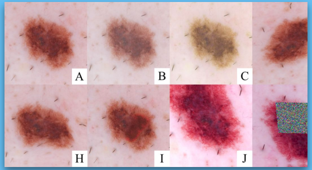


Knowledge transfer for melanoma screening with deep learning

RECOD Titans at ISIC challenge 2017

Data, depth, and design: Learning reliable models for skin lesion analysis

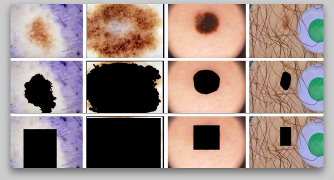
2018



Data augmentation for skin lesion analysis

Skin lesion synthesis with generative adversarial networks

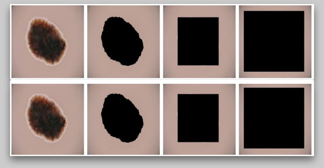
2019



(De)Constructing bias on skin lesion datasets

Solo or ensemble? Choosing a CNN architecture for melanoma classification

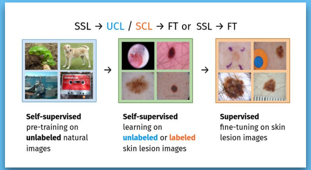
2020



Debiasing skin lesion datasets and models? Not so fast

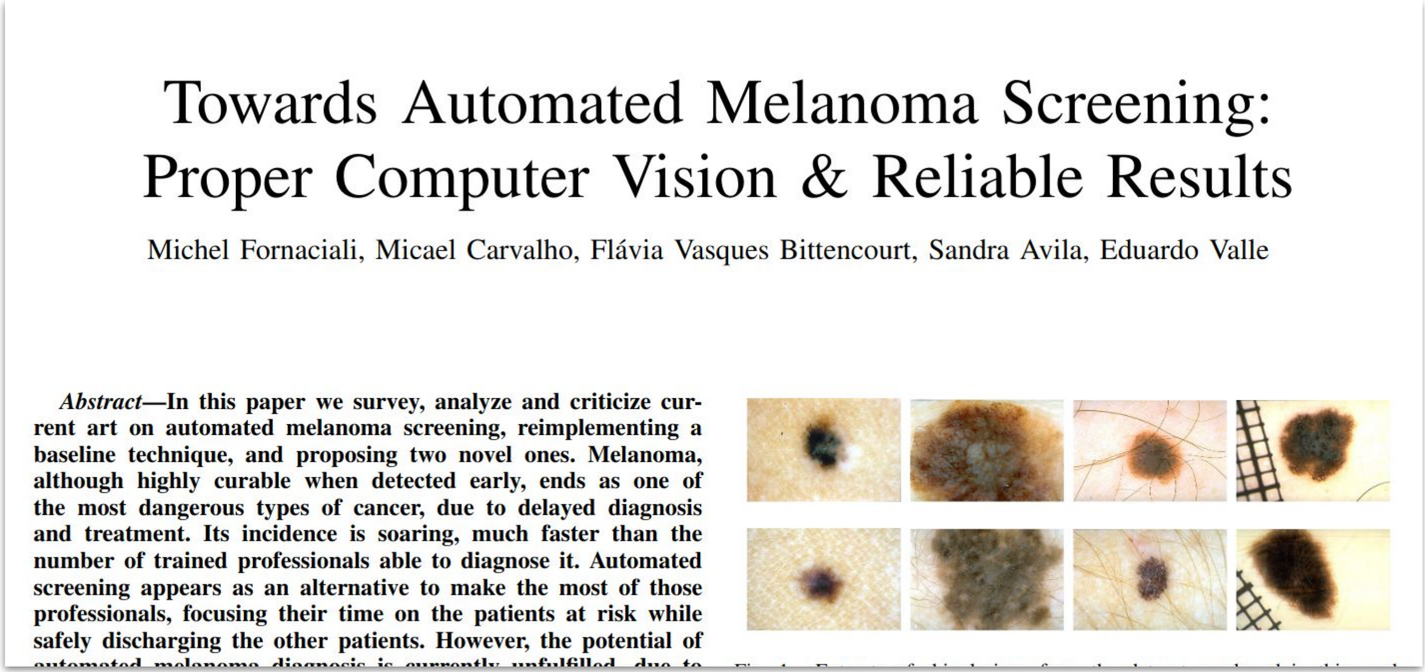
Less is more: Sample selection and label conditioning improve skin lesion segmentation

2021



GAN-based data augmentation and anonymization for skin-lesion analysis: A critical review

An evaluation of self-supervised pre-training for skin-lesion analysis



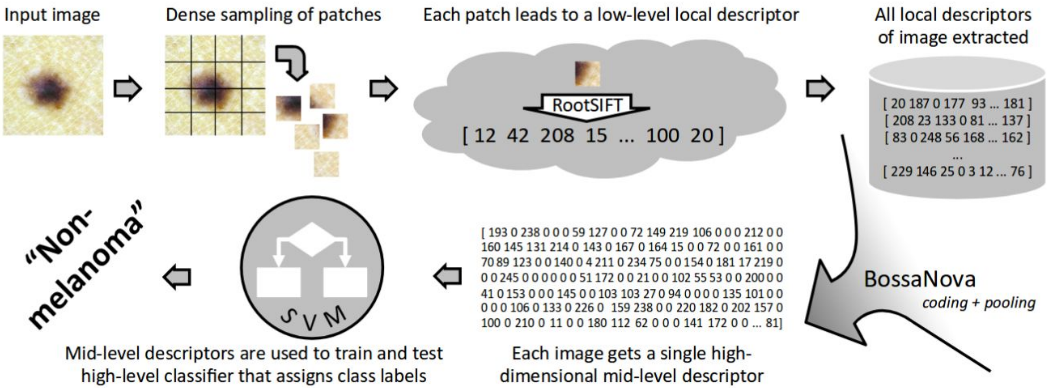
Towards Automated Melanoma Screening: Proper Computer Vision & Reliable Results

Michel Fornaciali, Micael Carvalho, Flávia Vasques Bittencourt, Sandra Avila, Eduardo Valle

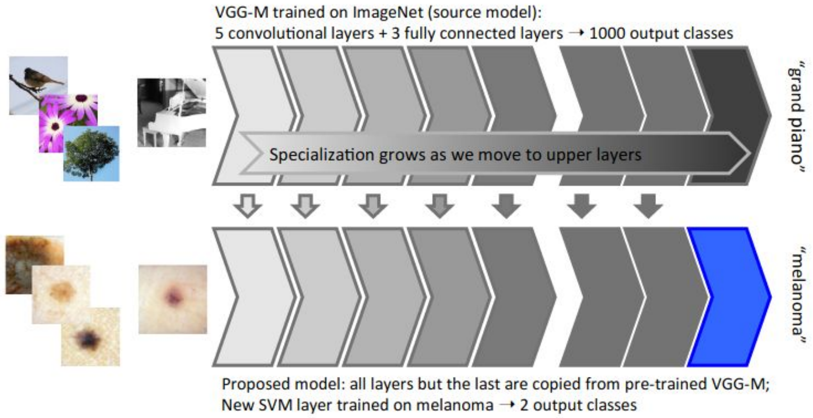
Abstract—In this paper we survey, analyze and criticize current art on automated melanoma screening, reimplementing a baseline technique, and proposing two novel ones. Melanoma, although highly curable when detected early, ends as one of the most dangerous types of cancer, due to delayed diagnosis and treatment. Its incidence is soaring, much faster than the number of trained professionals able to diagnose it. Automated screening appears as an alternative to make the most of those professionals, focusing their time on the patients at risk while safely discharging the other patients. However, the potential of

Paper, Code & Data: https://github.com/learningtitans/melanoma-screening

Bag-of-Visual-Words pipeline



Deep Transfer Learning pipeline



M. Fornaciali, M. Carvalho, F. V. Bittencourt, S. Avila, E. Valle, "Towards automated melanoma screening: Proper computer vision & reliable results", 2016.

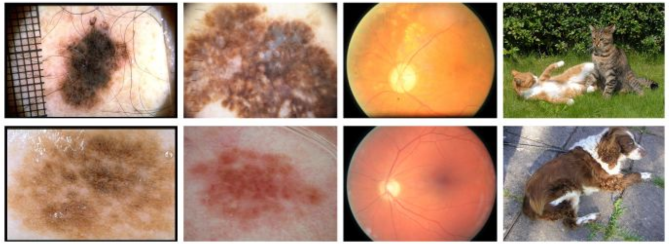
Knowledge Transfer for Melanoma Screening with Deep Learning

Afonso Menegola[†‡], Michel Fornaciali[†‡], Ramon Pires[°],
Flávia Vasques Bittencourt[•], Sandra Avila[†], Eduardo Valle[†*]

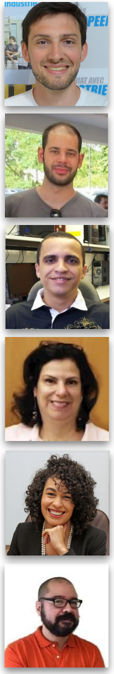
[†]RECOD Lab, DCA, FEEC, University of Campinas (Unicamp), Brazil
[°]RECOD Lab, IC, University of Campinas (Unicamp), Brazil
[•]School of Medicine, Federal University of Minas Gerais (UFMG), Brazil

ABSTRACT

Knowledge transfer impacts the performance of deep learning — the state of the art for image classification tasks, including automated melanoma screening. Deep learning's greed for large amounts of training data poses a challenge for medical tasks, which we can alleviate by recycling knowledge from models trained on different tasks, in a scheme called *transfer learning*. Although much of the best art on automated melanoma screening employs some form of transfer learning, a systematic evaluation was missing. Here we investigate the presence of transfer, from which task the transfer is sourced, and



(a) (b) (c) (d)



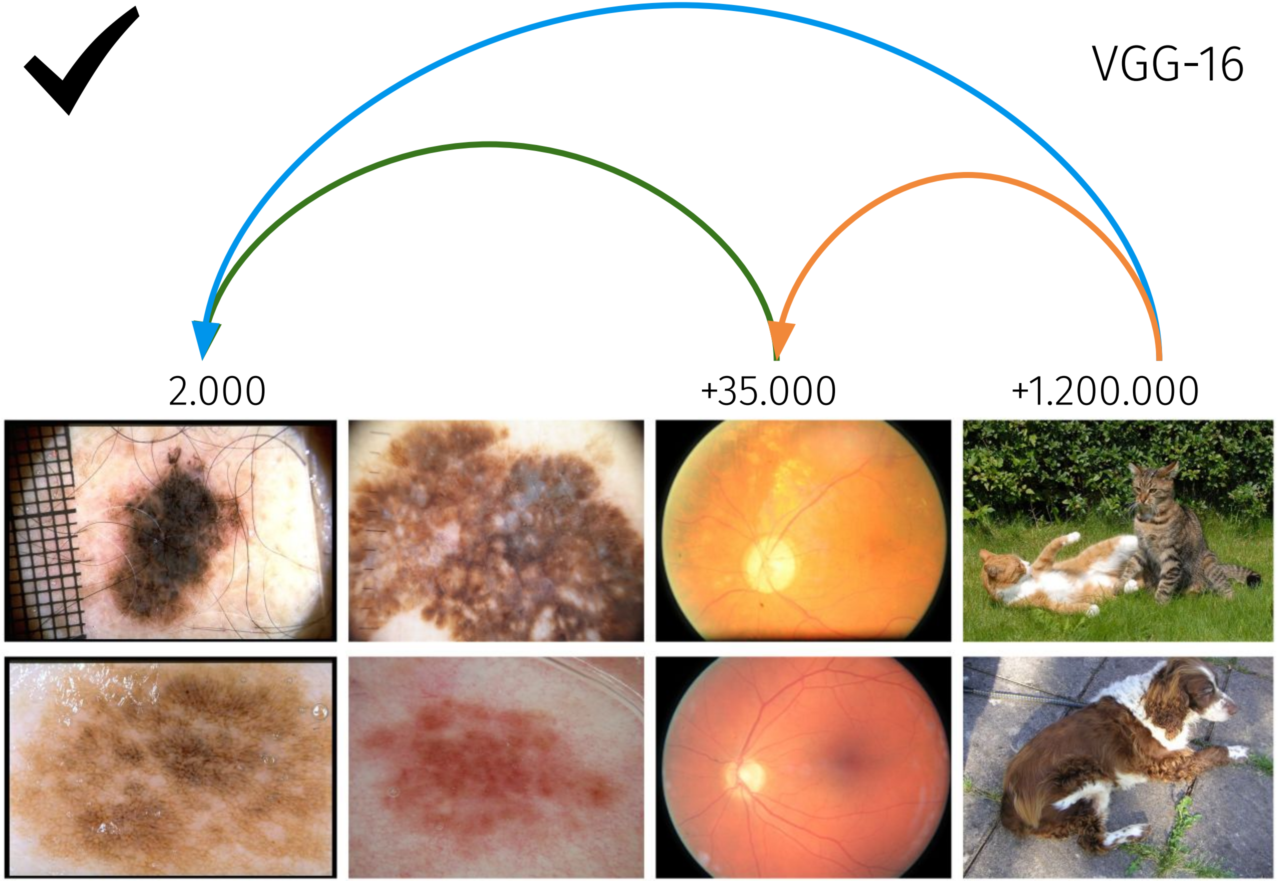
Paper, Code & Data: https://github.com/learningtitans/melanoma-transfer

ImageNet -> Melanoma ✓

Double Transfer:

ImageNet -> Retina &

Retina -> Melanoma



A. Menegola, M. Fornaciali, R. Pires, F. V. Bittencourt, S. Avila, E. Valle, “Knowledge transfer for melanoma screening with deep learning”, ISBI 2017.

RECOD Titans at ISIC Challenge 2017

Afonso Menegola†, Julia Tavares†, Michel Fornaciali, Lin Tzy Li, Sandra Avila, Eduardo Valle*

HISTORY

Our team has worked on melanoma classification since early 2014 [1], and has employed deep learning with transfer learning for that task since 2015 [2]. Recently, the community has started to move from traditional techniques towards deep learning, following the general trend of computer vision [3]. Deep learning poses a challenge for medical applications, due to the need of very large training sets. Thus, transfer learning becomes crucial for success in those applications, motivating our paper for ISBI 2017 [4].

Our team participated in Parts 1 and 3 of the ISIC Challenge 2017, described below in that order. Although our team has a long experience with skin-lesion classification (Part 3), this Challenge was the very first time we worked on skin-lesion segmentation (Part 1).

The code needed to reproduce our results is at our **code repository**[4].

C. Data Augmentation

We used online image augmentation, with up to 10% horizontal and vertical shifts, up to 20% zoom, and up to 270° degrees rotation. Images were first resized — we tried 256×256 and 128×128, ultimately keeping the latter, which was faster and resulted in similar performance. Transforming the images before resizing them was slower and did not improve the results.

D. Experiments

Our first attempt was a model based on the VGG network [6]. The first part of the model consisted of the VGG-16 layers

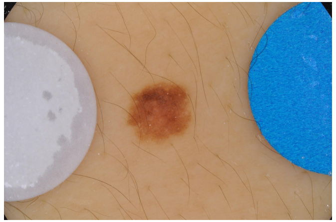


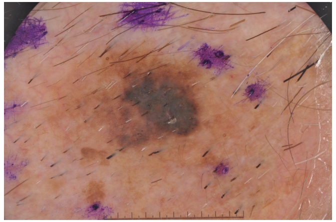
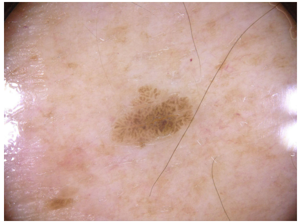



Melanoma Classification

Paper, Code & Data: https://arxiv.org/abs/1703.04819

ISIC Challenge **2017**



Training data
2000 images
95.1%
(internal validation)

Validation data
150 images
90.8%

Test data
600 images
87.4%

A. Menegola, J. Tavares, M. Fornaciali, L. T. Li, S. Avila, E. Valle, "RECOD Titans at ISIC Challenge 2017", 2017.

RECOD Titans at ISIC Challenge 2017

Afonso Menegola†, Julia Tavares†, Michel Fornaciali, Lin Tzy Li, Sandra Avila, Eduardo Valle*

HISTORY

Our team has worked on melanoma classification since early 2014 [1], and has employed deep learning with transfer learning for that task since 2015 [2]. Recently, the community has started to move from traditional techniques towards deep learning, following the general trend of computer vision [3]. Deep learning poses a challenge for medical applications, due to the need of very large training sets. Thus, transfer learning becomes crucial for success in those applications, motivating our paper for ISBI 2017 [4].

Our team participated in Parts 1 and 3 of the ISIC Challenge 2017, described below in that order. Although our team has a long experience with skin-lesion classification (Part 3), this Challenge was the very first time we worked on skin-lesion segmentation (Part 1).

The code needed to reproduce our results is at our **code repository**[4].

C. Data Augmentation

We used online image augmentation, with up to 10% horizontal and vertical shifts, up to 20% zoom, and up to 270° degrees rotation. Images were first resized — we tried 256×256 and 128×128, ultimately keeping the latter, which was faster and resulted in similar performance. Transforming the images before resizing them was slower and did not improve the results.

D. Experiments

Our first attempt was a model based on the VGG network [6]. The first part of the model consisted of the VGG-16 layers






Melanoma Classification

Paper, Code & Data: https://arxiv.org/abs/1703.04819

ISIC Challenge 2017

Models + data ✓

Image resolution

Class/sample-weighting schemes

Curriculum learning

SVM decision layer

Training and test augmentation ✓

Patient data

Per-image normalization ✓

Segmentation information

Stacking models and meta-learning ✓

A. Menegola, J. Tavares, M. Fornaciali, L. T. Li, S. Avila, E. Valle, “RECOD Titans at ISIC Challenge 2017”, 2017.

2^9 factors × 5 datasets = 2560 experiments

Neurocomputing 383 (2020) 303–313

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Data, depth, and design: Learning reliable models for skin lesion analysis

Eduardo Valle[a,*], Michel Fornaciali[a], Afonso Menegola[a], Julia Tavares[a], Flávia Vasques Bittencourt[b], Lin Tzy Li[c,d], Sandra Avila[c]

[a] School of Electrical and Computing Engineering, University of Campinas (UNICAMP), Av. Albert Einstein 400, Campinas, SP 13083-852, Brazil
[b] School of Medicine, Federal University of Minas Gerais (UFMG), Alamedalvaro Celso 55, Belo Horizonte, MG 30150-260, Brazil
[c] Institute of Computing, University of Campinas (UNICAMP), Av. Albert Einstein 1251, Campinas, SP 13083-852, Brazil
[d] Samsung R&D Institute Brazil (SRBR), Campinas, SP, Brazil

Paper, Code & Data: https://github.com/learningtitans/data-depth-design

"Amount of train data has disproportionate influence, explaining **46%** of the variation in performance."

"Other than data, the most important factor was the **use of data augmentation on test**."

Table 3
Selected lines from the 176-line ANOVA table; most of the omitted lines (126) had p-values $\geq$ 0.05. Absolute explanation based on η^2-measure, relative explanation ignores residuals and choice of test dataset (j).

	Factor	*p*-value	Explanation (%)		Best AUC (%)		Worst AUC (%)	
			Abs.	Rel.	Treatment	Mean	Treatment	Mean
a	Model architecture	< 0.001	0	1	resnet	84	inception	83
b	Train dataset	< 0.001	5	46	full	85	challenge	81
c	Input resolution	< 0.001	1	5	598	84	299–305	82
d	Data augmentation	0.17	0	0	default	83	custom	83
e	Input normalization	0.001	0	0	default	83	erase mean	83
f	Use of segmentation	< 0.001	0	2	no	84	yes	83
g	Duration of training	0.003	0	0	full	83	half	83
h	SVM layer	< 0.001	0	4	no	84	yes	83
i	Augmentation on test	< 0.001	1	12	yes	84	no	82
j	Test dataset	< 0.001	75		full.split	96	edra.clinical	66

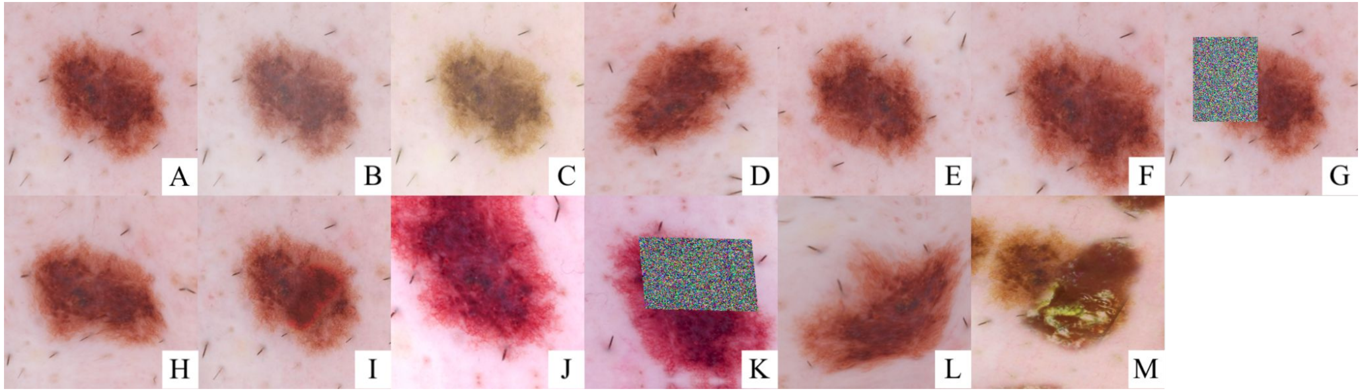
Paper, Code & Data: https://github.com/learningtitans/data-depth-design

E. Valle, M. Fornaciali, A. Menegola, J. Tavares, F. V. Bittencourt, L. T. Li, S. Avila, "Data, depth, and design: Learning reliable models for skin lesion analysis", 2020.

Data Augmentation for Skin Lesion Analysis

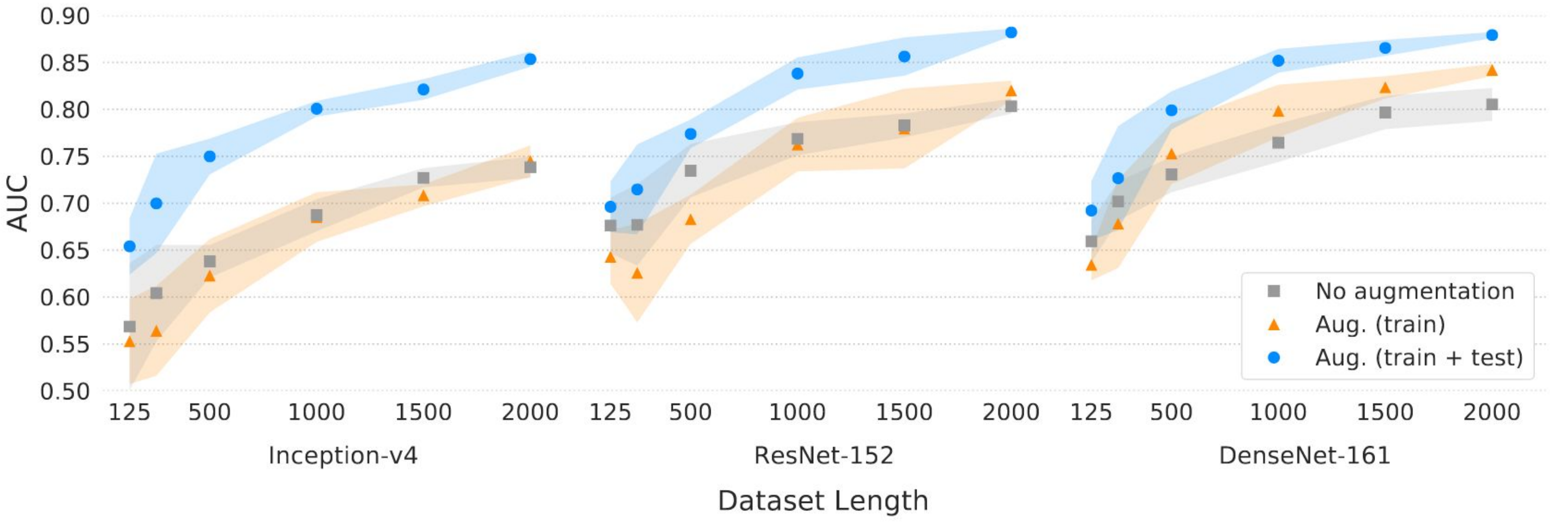
Fábio Perez[1], Cristina Vasconcelos[2], Sandra Avila[3], and Eduardo Valle[1]

[1]RECOD Lab, DCA, FEEC, University of Campinas (Unicamp), Brazil
[2]Computer Science Department, IC, Federal Fluminense University (UFF), Brazil
[3]RECOD Lab, IC, University of Campinas (Unicamp), Brazil



Paper, Code & Data: https://github.com/fabioperez/skin-data-augmentation

Augmentation on Training & Testing



Paper, Code & Data: https://github.com/fabioperez/skin-data-augmentation

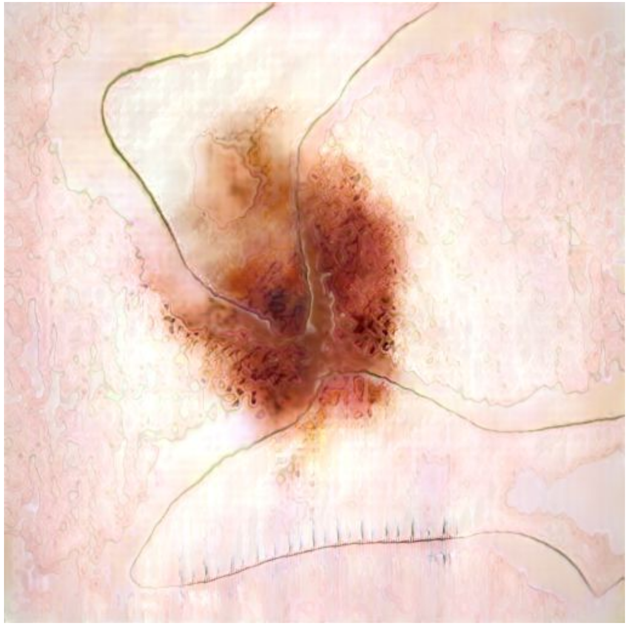
F. Perez, C. Vasconcelos, S. Avila, E. Valle, "Data augmentation for skin lesion analysis", ISIC Workshop, MICCAI, 2018. 🌟 Best paper award.

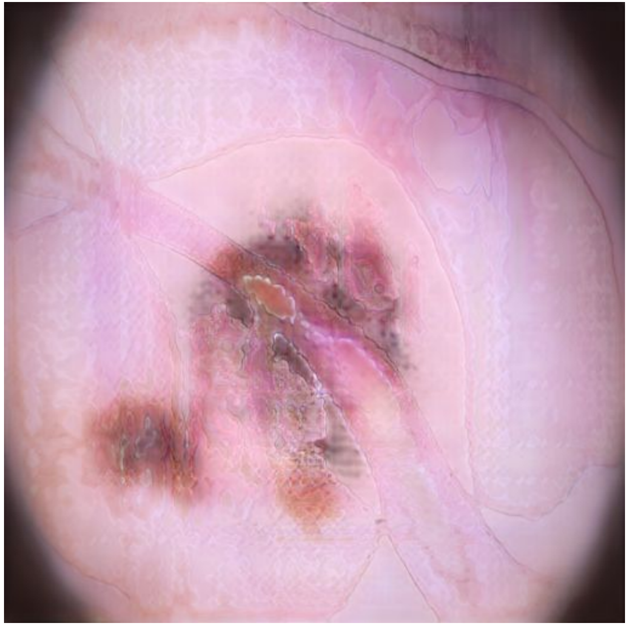
Let's generate data!

Generative Adversarial Networks (GANs)

I SEE BAD DATA

Spoiler alert! The film: “The Sixth Sense”, the line: “I see dead people.”



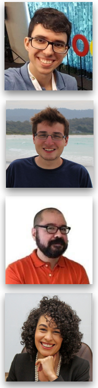


… using Progressive Growing GANs (PGAN)

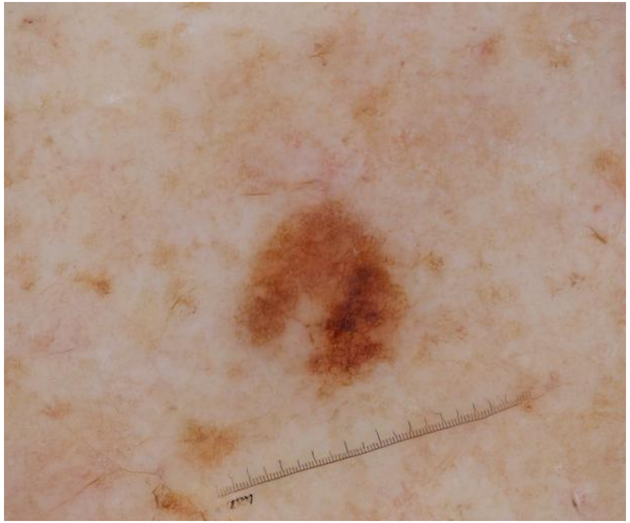
Skin Lesion Synthesis with Generative Adversarial Networks

Alceu Bissoto[1], Fábio Perez[2], Eduardo Valle[2], and Sandra Avila[1]

[1]RECOD Lab, IC, University of Campinas (Unicamp), Brazil
[2]RECOD Lab, DCA, FEEC, University of Campinas (Unicamp), Brazil

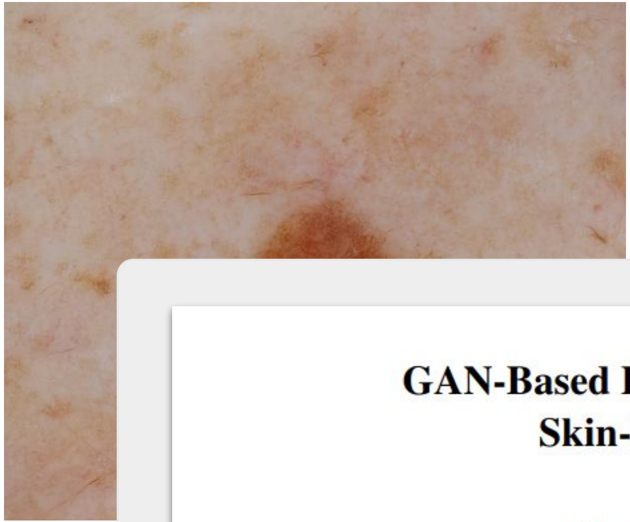


Paper, Code & Data: https://github.com/alceubissoto/gan-skin-lesion



… using pix2pixHD

A. Bissoto, F. Perez, E. Valle, S. Avila, “Skin lesion synthesis with generative adversarial networks”, ISIC Workshop, MICCAI, 2018.



GAN-Based Data Augmentation and Anonymization for Skin-Lesion Analysis: A Critical Review

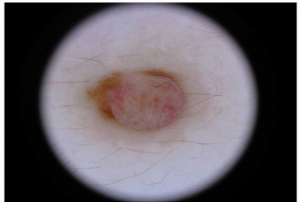
Alceu Bissoto[1] Eduardo Valle[2] Sandra Avila[1]
[1]Institute of Computing (IC) [2]School of Electrical and Computing Engineering (FEEC)
RECOD Lab., University of Campinas (UNICAMP), Brazil

Paper, Code & Data: https://github.com/alceubissoto/gan-aug-analysis

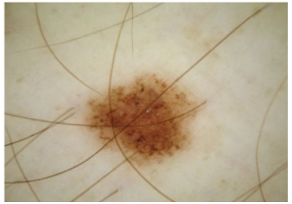
… using pix2pixHD

A. Bissoto, F. Perez, E. Valle, S. Avila, "Skin lesion synthesis with generative adversarial networks", ISIC Workshop, MICCAI, 2018.

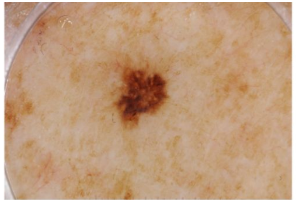
Dataset **biases** may **inflate** the performance of models!



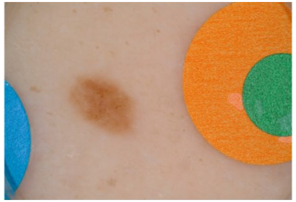
dark border



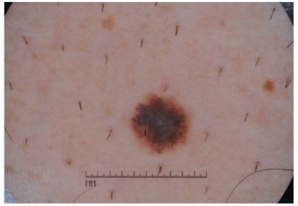
hair



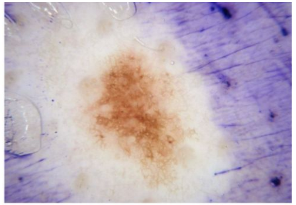
gel border



color marker

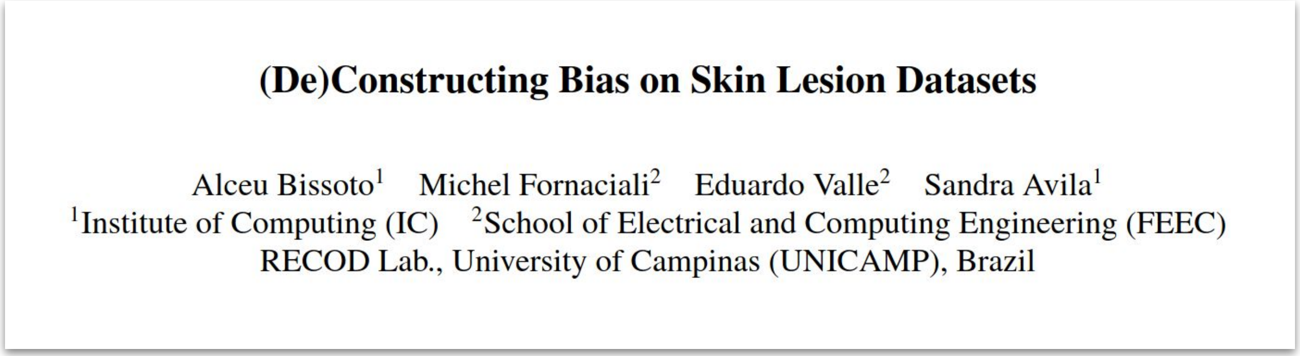


ruler

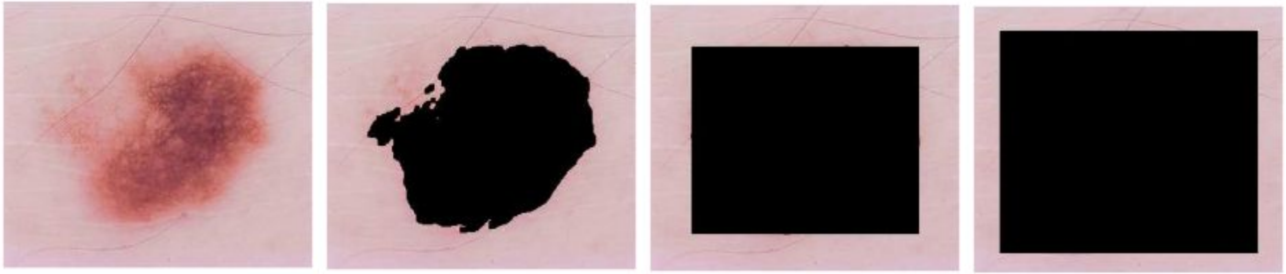


ink markings

A. Bissoto, M. Fornaciali, E. Valle, S. Avila, "(De)Constructing bias on skin lesion datasets", ISIC Workshop, CVPR, 2019. 🌟 Best paper award.


(De)Constructing Bias on Skin Lesion Datasets

Alceu Bissoto[1] Michel Fornaciali[2] Eduardo Valle[2] Sandra Avila[1]
[1]Institute of Computing (IC) [2]School of Electrical and Computing Engineering (FEEC)
RECOD Lab., University of Campinas (UNICAMP), Brazil

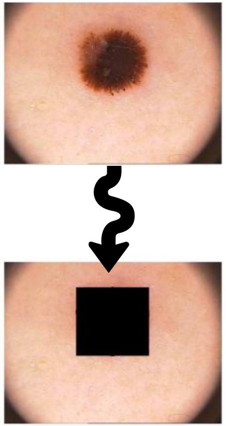


traditional | only skin | bounding box | bounding box 70%

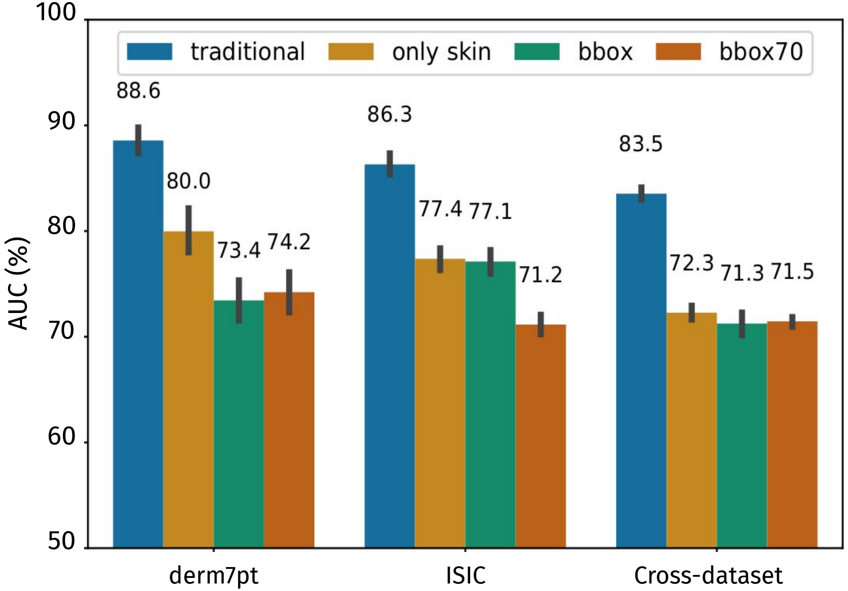
Paper, Code & Data: https://github.com/alceubissoto/deconstructing-bias-skin-lesion

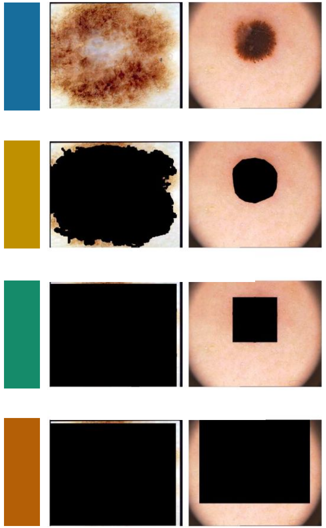
Is it possible to be completely bias free?

But, if we destroy the "important" information in the data?



Performance[1] of **157** dermatologists on ISIC: **67% AUC**



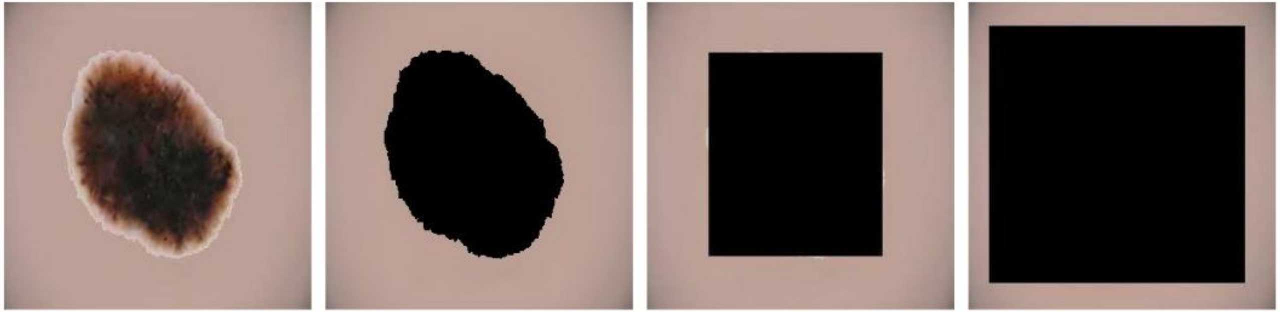


[1]Brinker et al., "The melanoma classification benchmark", European Journal of Cancer, 2019.

A. Bissoto, M. Fornaciali, E. Valle, S. Avila, "(De)Constructing bias on skin lesion datasets", ISIC Workshop, CVPR, 2019. 🌟 Best paper award.

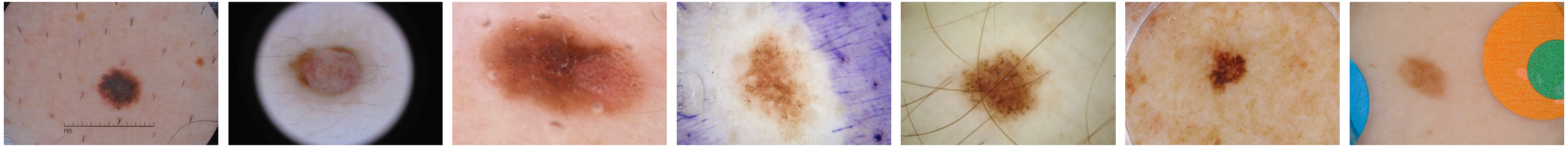
Debiasing Skin Lesion Datasets and Models? Not So Fast

Alceu Bissoto[1] Eduardo Valle[2] Sandra Avila[1]
[1]Institute of Computing (IC) [2]School of Electrical and Computing Engineering (FEEC)
RECOD Lab., University of Campinas (UNICAMP), Brazil



Paper, Code & Data: https://github.com/alceubissoto/debiasing-skin

Trap Sets



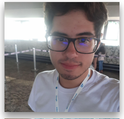
	ruler	dark corner	gel bubble	ink markings	hair	gel border	color marker
Train Spearman Correlation	0.41	0.30	-0.18	0.21	-0.26	0.12	-0.11
Test Spearman Correlation	-0.67	-0.39	0.47	-0.42	0.34	-0.51	-0.16

A. Bissoto, E. Valle, S. Avila, “Debiasing skin lesion datasets and models? Not so fast”, 2020. 🌟 Runner-up paper award.

An Evaluation of Self-Supervised Pre-Training for Skin-Lesion Analysis

A Preprint

Levy Chaves[1] **Alceu Bissoto**[1] **Eduardo Valle**[2] **Sandra Avila**[1]
[1]Institute of Computing (IC) [2]School of Electrical and Computing Engineering (FEEC)
RECOD Lab., University of Campinas (UNICAMP), Brazil



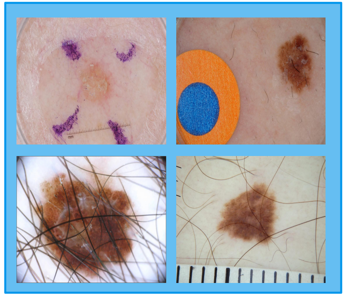

Paper, Code & Data: https://github.com/VirtualSpaceman/ssl-skin-lesions

L. Chaves, A. Bissoto, E. Valle, S. Avila, “An evaluation of self-supervised pre-training for skin-lesion analysis”, 2021.

SUP
↓
FT

Supervised training on **labeled** natural images

↓

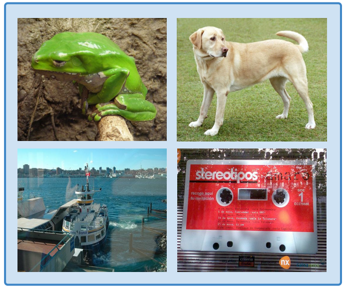


Supervised fine-tuning on skin lesion images

SSL
↓
FT

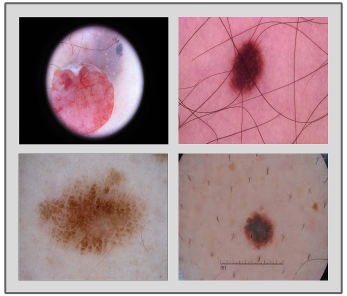
or

SSL
↓
UCL
/
SCL
↓
FT



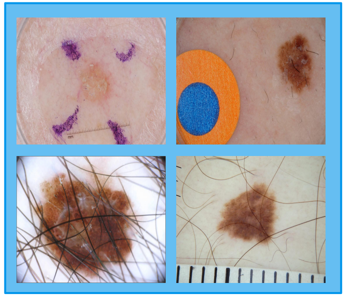
Self-supervised pre-training on **unlabeled** natural images

↓



Self-supervised learning on unlabeled or labeled skin lesion images

↓



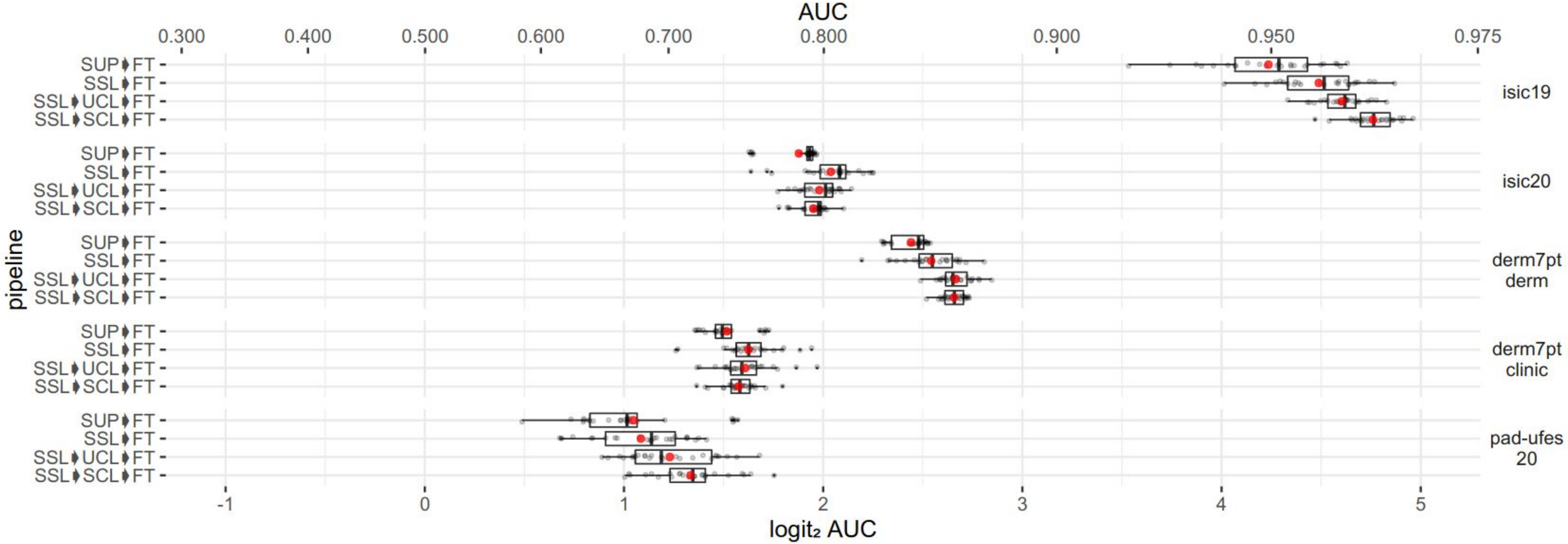
Supervised fine-tuning on skin lesion images

SSL: Self-Supervised Learning
UCL/SCL: Unsupervised/Supervised Contrastive Learning
FT: Fine-tuning

L. Chaves, A. Bissoto, E. Valle, S. Avila, “An evaluation of self-supervised pre-training for skin-lesion analysis”, 2021.

100%
of labeled images

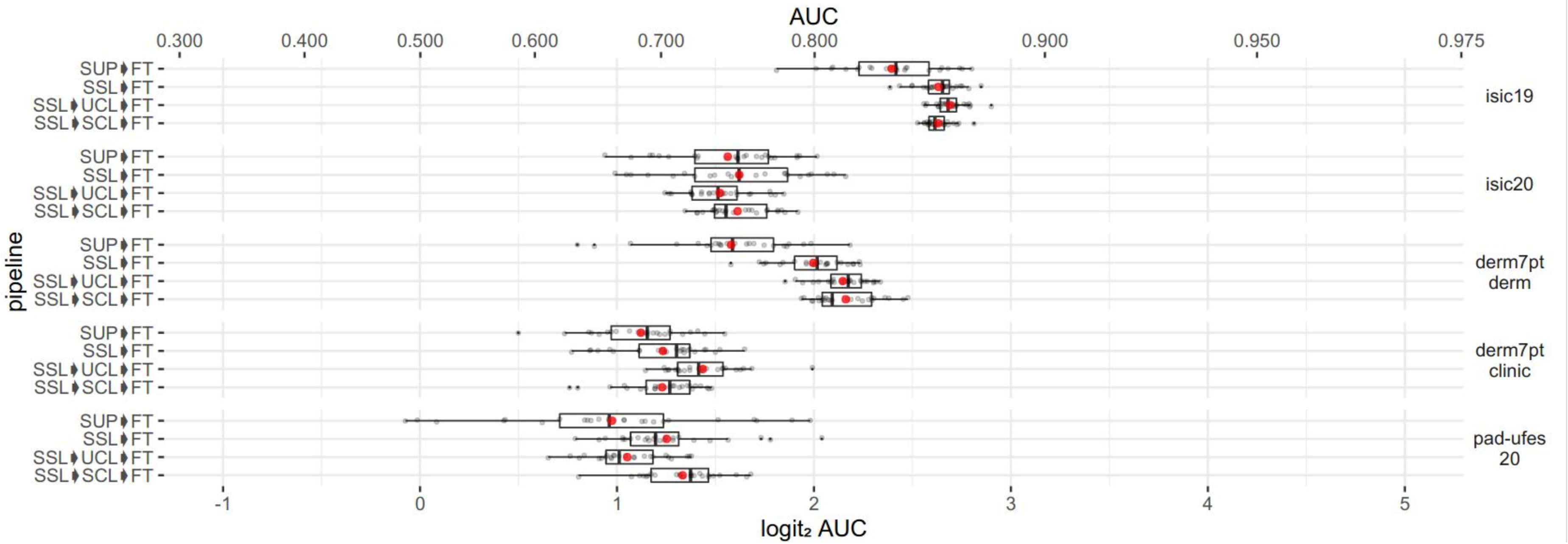
Training on the ISIC 2019 dataset **(~15,000 images)**



L. Chaves, A. Bissoto, E. Valle, S. Avila, “An evaluation of self-supervised pre-training for skin-lesion analysis”, 2021.

10%
of labeled images

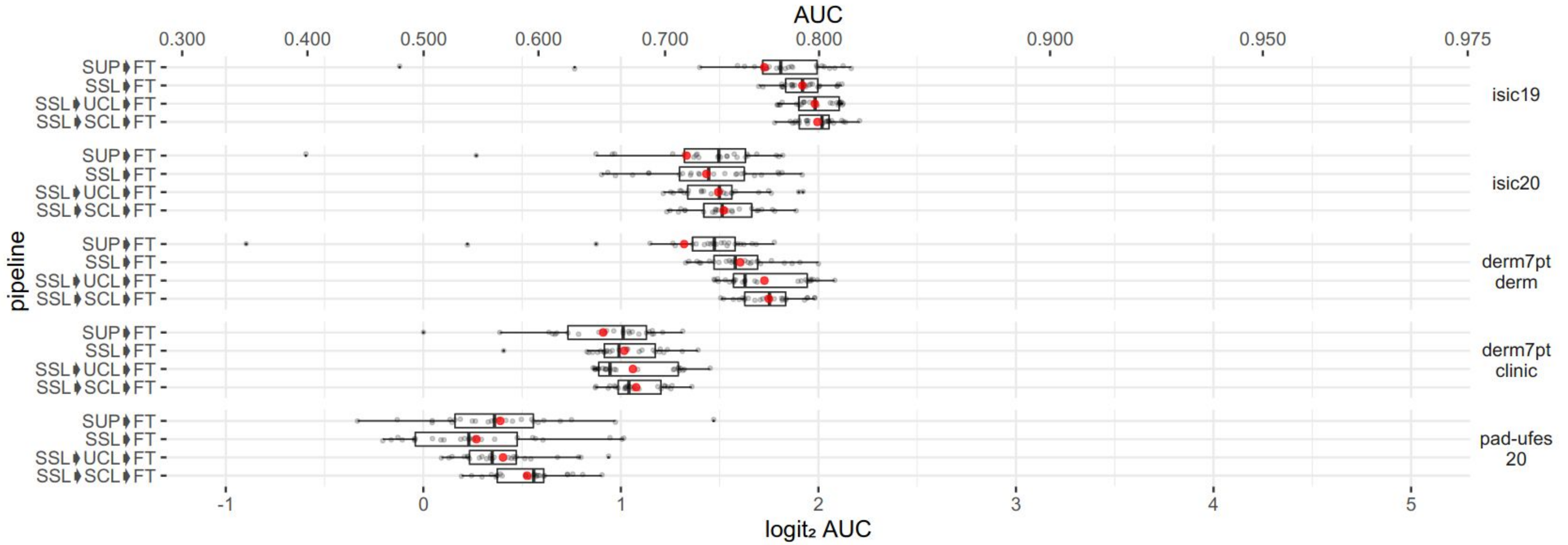
Training on the ISIC 2019 dataset **(~1500 images)**



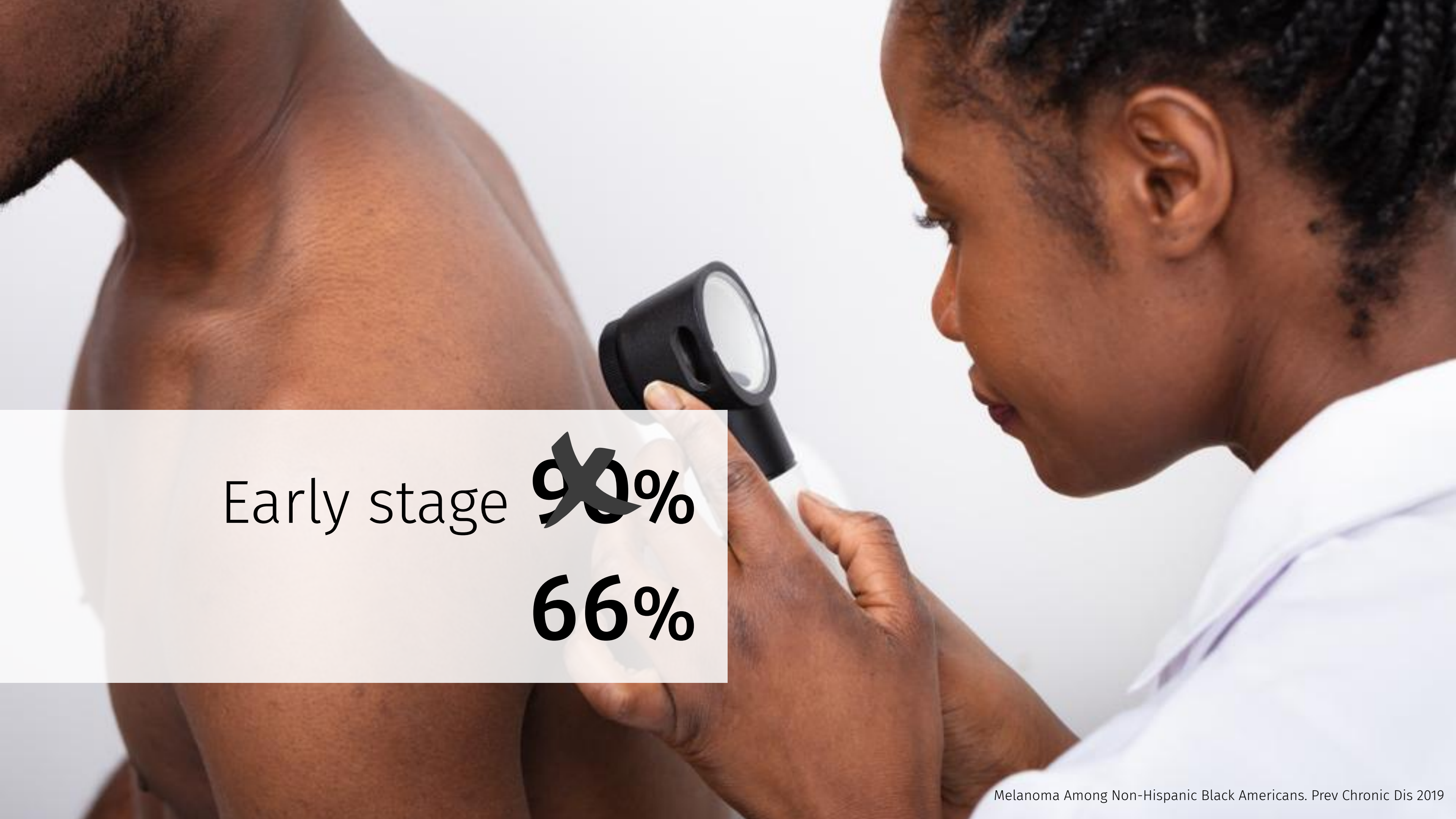
L. Chaves, A. Bissoto, E. Valle, S. Avila, "An evaluation of self-supervised pre-training for skin-lesion analysis", 2021.

Training on the ISIC 2019 dataset **(~150 images)**

1%
of labeled images



L. Chaves, A. Bissoto, E. Valle, S. Avila, "An evaluation of self-supervised pre-training for skin-lesion analysis", 2021.



Melanoma Among Non-Hispanic Black Americans. Prev Chronic Dis 2019

The New York Times

Dermatology Has a Problem With Skin Color

Common conditions often manifest differently on dark skin. Yet physicians are trained mostly to diagnose them on white skin.

By Roni Caryn Rabin

Aug. 30, 2020

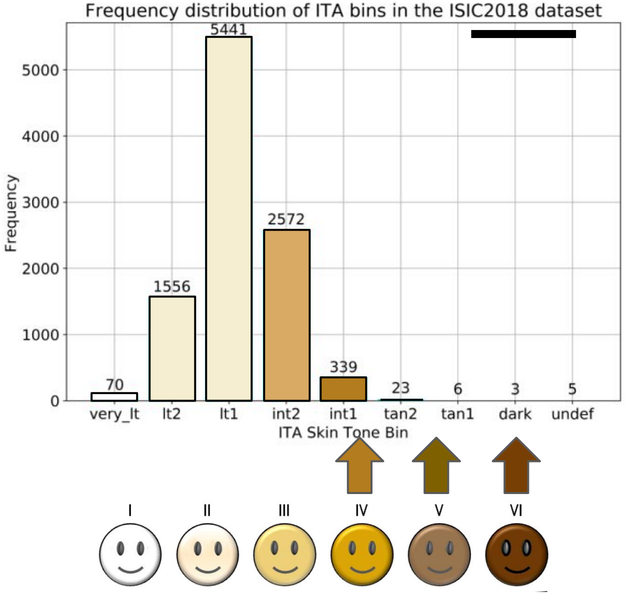
Dr. Jenna Lester, director of the Skin of Color Program at University of California, San Francisco.

Fairness of Classifiers Across Skin Tones in Dermatology

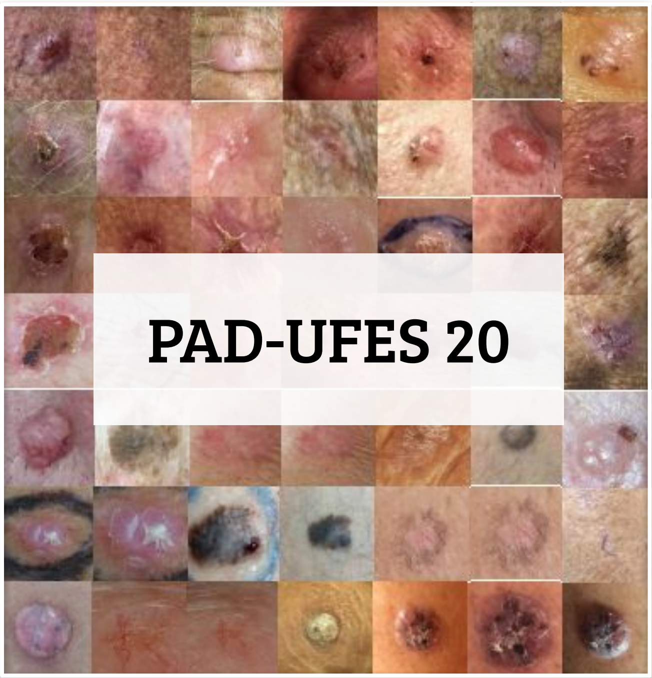
Newton M. Kinyanjui[1,4], Timothy Odonga[1,4], Celia Cintas[1], Noel C. F. Codella[2], Rameswar Panda[3], Prasanna Sattigeri[2], and Kush R. Varshney[1,2](✉)

[1] IBM Research – Africa, Nairobi 00100, Kenya
krvarshn@us.ibm.com
[2] IBM Research – T. J. Watson Research Center, Yorktown Heights, NY 10598, USA
[3] IBM Research – Cambridge, Cambridge, MA 02142, USA
[4] Carnegie Mellon University Africa, Kigali, Rwanda

MICCAI 2020



Fitzpatrick Scale

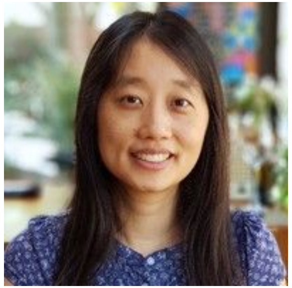


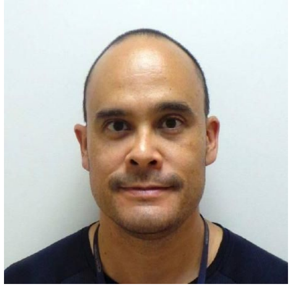
Groh et al., Evaluating Deep Neural Networks Trained on Clinical Images in Dermatology with the Fitzpatrick 17k Dataset, CVPR 2021.
Pacheco et al., PAD-UFES-20: A skin lesion dataset composed of patient data and clinical images collected from smartphones, Data in Brief, 2020.




industrie











Thanks!

Sandra Avila

www.ic.unicamp.br/~sandra

sandra@ic.unicamp.br

@sandraavilabr

UNICAMP recod CNPq CAPES LARA FAPESP CCES UNICAMP