Deep learning for skin image analysis
Beyond more data and faster GPUs

Ghassan Hamarneh
hamarneh@sfu.ca

17 June 2019
ISIC Skin Image Analysis Workshop at CVPR
Acknowledgements

Current and former students

- Dr. Jeremy Kawahara
- Dr. Hengameh Mirzaalian
- Dr. Aicha Bentaieb
- Saeid Asgari
- Zahra MiriKharaji
- Saeed Izadi
- Yiqi Yan
- Kumar Abhishek
- Chris Yoon

Funding: research, infrastructure, computing

Disclosure
Scientific Advisor and Shareholder: Triage Technologies Inc.
Skin image analysis system
Image analysis via deep learning

- **input**: $x_1, x_2, x_3$
- **prediction model**: parameterized computational layers
- **predicted output**: $\hat{y}_1, \hat{y}_2, \hat{y}_3$
- **true output**: $y_1, y_2, y_3$
- **loss**: $\Delta$
- **update parameters**
- **hyperparameter tuning**
- **training, validation, testing (hold out), deployment**
derm7pt

~2000 images
~1000 cases: dermoscopic + clinical images
7 point criteria (2 to 8 classes each)
diagnosis (melanoma vs not)
meta data (e.g. age, sex, lesion location)

PH2
200 derm. images

DermoFit
1,300 clinical images
10 classes of skin lesions

ISIC challenge
25,332 derm. images
8 classes
metadata

SD198
6,584 clinical images
198 classes
10-60 images per class
(from DermQuest website)
AI at par with humans' papers

Jan. 2017

Dermatologist-level classification of skin cancer with deep neural networks

André Esteva, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau & Sebastian Thrun

Nature 542, 115–118 (02 February 2017) | Download Citation

https://www.nature.com/articles/nature21056

Feb. 2019

Deep-learning-based, computer-aided classifier developed with a small dataset of clinical images surpasses board-certified dermatologists in skin tumour diagnosis


Aug. 2018

“CNN ROC AUC greater than mean ROC area of dermatologists... higher specificity”

Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists

H. A. Heinsdle, C. Fink, R. Schneiderbauer, F. Toberer, T. Buhf, A. Blum, A. Kalco, A. Ben Hadj Hassen, L. Thomas, A. Enk & L. Uhmann


May 2019

Deep learning outperformed 36 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task


Each point represents an experiment showing the performance of either a human or a machine.

- Accuracy
- Number of Classes (Log Scale)

**Visual Diagnosis of Dermatological Disorders: Human and Machine Performance**

Jeremy Kawahara, Ghassan Hamarneh


June 2019

<table>
<thead>
<tr>
<th>Year</th>
<th>Dataset</th>
<th>N.Images</th>
<th>N.Test</th>
<th>Dem. Clinic.</th>
<th>Meta</th>
<th>H vs. M</th>
<th>Classes</th>
<th>Acc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[89]</td>
<td>2015</td>
<td>Internal</td>
<td>-</td>
<td>65</td>
<td>✓</td>
<td>human</td>
<td>2</td>
<td>63.35</td>
</tr>
<tr>
<td>[89]</td>
<td>2015</td>
<td>Internal</td>
<td>273</td>
<td>65</td>
<td>✓</td>
<td>machine</td>
<td>2</td>
<td>63.08</td>
</tr>
<tr>
<td>[90]</td>
<td>2017</td>
<td>ISIC-100</td>
<td>-</td>
<td>100</td>
<td>✓</td>
<td>human</td>
<td>2</td>
<td>70.30</td>
</tr>
<tr>
<td>[90]</td>
<td>2017</td>
<td>ISIC-100</td>
<td>1000</td>
<td>100</td>
<td>✓</td>
<td>machine</td>
<td>2</td>
<td>76.00</td>
</tr>
<tr>
<td>[96]</td>
<td>2018</td>
<td>Internal</td>
<td>-</td>
<td>100</td>
<td>✓</td>
<td>human</td>
<td>2</td>
<td>74.40</td>
</tr>
<tr>
<td>[96]</td>
<td>2018</td>
<td>Internal</td>
<td>-</td>
<td>100</td>
<td>✓</td>
<td>machine</td>
<td>2</td>
<td>78.30</td>
</tr>
<tr>
<td>[96]</td>
<td>2018</td>
<td>Internal</td>
<td>100</td>
<td>100</td>
<td>✓</td>
<td>human</td>
<td>2</td>
<td>81.60</td>
</tr>
<tr>
<td>[94]</td>
<td>2018</td>
<td>Avan</td>
<td>-</td>
<td>1133</td>
<td>✓</td>
<td>human</td>
<td>2</td>
<td>75.80</td>
</tr>
<tr>
<td>[94]</td>
<td>2018</td>
<td>Avan</td>
<td>49,567</td>
<td>1133</td>
<td>✓</td>
<td>machine</td>
<td>2</td>
<td>80.00</td>
</tr>
<tr>
<td>[97]</td>
<td>2019</td>
<td>ISIC-100</td>
<td>13,737</td>
<td>100</td>
<td>✓</td>
<td>machine</td>
<td>2</td>
<td>84.02</td>
</tr>
<tr>
<td>[98]</td>
<td>2019</td>
<td>ISIC-100</td>
<td>-</td>
<td>100</td>
<td>✓</td>
<td>human</td>
<td>2</td>
<td>62.82</td>
</tr>
<tr>
<td>[90]</td>
<td>2019</td>
<td>MED-MODE</td>
<td>-</td>
<td>100</td>
<td>✓</td>
<td>human</td>
<td>2</td>
<td>69.40</td>
</tr>
<tr>
<td>[94]</td>
<td>2017</td>
<td>Stanford</td>
<td>-</td>
<td>180</td>
<td>✓</td>
<td>human</td>
<td>3</td>
<td>65.78</td>
</tr>
<tr>
<td>[44]</td>
<td>2017</td>
<td>Stanford</td>
<td>127,463</td>
<td>127,463</td>
<td>✓</td>
<td>machine</td>
<td>3</td>
<td>72.30</td>
</tr>
<tr>
<td>[52]</td>
<td>2013</td>
<td>Dermott</td>
<td>960</td>
<td>960</td>
<td>✓</td>
<td>machine</td>
<td>5</td>
<td>74.30</td>
</tr>
<tr>
<td>[52]</td>
<td>2015</td>
<td>Atlas</td>
<td>2,018</td>
<td>2,018</td>
<td>✓</td>
<td>machine</td>
<td>5</td>
<td>71.10</td>
</tr>
<tr>
<td>[16]</td>
<td>2018</td>
<td>Atlas</td>
<td>3,945</td>
<td>3,945</td>
<td>✓</td>
<td>machine</td>
<td>5</td>
<td>73.70</td>
</tr>
<tr>
<td>[16]</td>
<td>2017</td>
<td>Internal</td>
<td>348</td>
<td>30</td>
<td>✓</td>
<td>human</td>
<td>6</td>
<td>74.00</td>
</tr>
<tr>
<td>[19]</td>
<td>2017</td>
<td>Internal</td>
<td>348</td>
<td>50</td>
<td>✓</td>
<td>machine</td>
<td>6</td>
<td>69.00</td>
</tr>
<tr>
<td>[102]</td>
<td>2002</td>
<td>Internal</td>
<td>-</td>
<td>250</td>
<td>✓</td>
<td>human</td>
<td>7</td>
<td>46.62</td>
</tr>
<tr>
<td>[103]</td>
<td>2002</td>
<td>Internal</td>
<td>-</td>
<td>250</td>
<td>✓</td>
<td>human</td>
<td>7</td>
<td>47.27</td>
</tr>
<tr>
<td>[139]</td>
<td>2011</td>
<td>Internal</td>
<td>-</td>
<td>256</td>
<td>✓</td>
<td>human</td>
<td>7</td>
<td>54.60</td>
</tr>
<tr>
<td>[141]</td>
<td>2017</td>
<td>Stanford</td>
<td>-</td>
<td>180</td>
<td>✓</td>
<td>machine</td>
<td>9</td>
<td>54.20</td>
</tr>
<tr>
<td>[141]</td>
<td>2017</td>
<td>Stanford</td>
<td>137,463</td>
<td>137,463</td>
<td>✓</td>
<td>machine</td>
<td>9</td>
<td>55.40</td>
</tr>
<tr>
<td>[155]</td>
<td>2019</td>
<td>Dermott</td>
<td>1,300</td>
<td>1,300</td>
<td>✓</td>
<td>machine</td>
<td>10</td>
<td>60.40</td>
</tr>
<tr>
<td>[155]</td>
<td>2019</td>
<td>Dermott</td>
<td>1,300</td>
<td>1,300</td>
<td>✓</td>
<td>machine</td>
<td>10</td>
<td>60.40</td>
</tr>
<tr>
<td>[156]</td>
<td>2018</td>
<td>Dermott</td>
<td>2,069</td>
<td>1,060</td>
<td>✓</td>
<td>machine</td>
<td>10</td>
<td>56.70</td>
</tr>
<tr>
<td>[156]</td>
<td>2018</td>
<td>Dermott</td>
<td>-</td>
<td>1,060</td>
<td>✓</td>
<td>human</td>
<td>12</td>
<td>57.80</td>
</tr>
<tr>
<td>[156]</td>
<td>2018</td>
<td>Dermott</td>
<td>-</td>
<td>1,060</td>
<td>✓</td>
<td>human</td>
<td>12</td>
<td>57.80</td>
</tr>
<tr>
<td>[156]</td>
<td>2018</td>
<td>Dermott</td>
<td>-</td>
<td>1,060</td>
<td>✓</td>
<td>machine</td>
<td>12</td>
<td>57.80</td>
</tr>
<tr>
<td>[156]</td>
<td>2018</td>
<td>Dermott</td>
<td>-</td>
<td>1,060</td>
<td>✓</td>
<td>machine</td>
<td>12</td>
<td>57.80</td>
</tr>
<tr>
<td>[156]</td>
<td>2018</td>
<td>Dermott</td>
<td>-</td>
<td>1,060</td>
<td>✓</td>
<td>machine</td>
<td>12</td>
<td>57.80</td>
</tr>
<tr>
<td>[156]</td>
<td>2018</td>
<td>Dermott</td>
<td>-</td>
<td>1,060</td>
<td>✓</td>
<td>machine</td>
<td>12</td>
<td>57.80</td>
</tr>
<tr>
<td>[156]</td>
<td>2018</td>
<td>Dermott</td>
<td>-</td>
<td>1,060</td>
<td>✓</td>
<td>machine</td>
<td>12</td>
<td>57.80</td>
</tr>
<tr>
<td>[156]</td>
<td>2018</td>
<td>Dermott</td>
<td>-</td>
<td>1,060</td>
<td>✓</td>
<td>machine</td>
<td>12</td>
<td>57.80</td>
</tr>
<tr>
<td>[156]</td>
<td>2018</td>
<td>Dermott</td>
<td>-</td>
<td>1,060</td>
<td>✓</td>
<td>machine</td>
<td>12</td>
<td>57.80</td>
</tr>
<tr>
<td>[156]</td>
<td>2018</td>
<td>Dermott</td>
<td>-</td>
<td>1,060</td>
<td>✓</td>
<td>machine</td>
<td>12</td>
<td>57.80</td>
</tr>
</tbody>
</table>
Skin conditions

1029 skin conditions excluding melanoma, neoplasms

https://icd.who.int/browse10/2016/en
Which classes to predict?

Large number of hierarchical conditions
Which conditions, at what level of granularity?

max specificity \( f \)
subject to:
accuracy \( (f) \geq 1 - \epsilon \)

Romero-Lopez et al.
SOCML 2019

Deng, Krause, Berg, Fei-Fei
CVPR 2012

133 diagnosis nodes parents to 588 different skin conditions
Other prediction tasks

localize

Mirzaalian, Hamarneh, Lee
CVPR 2009
https://doi.ieeecomputersociety.org/10.1109/CVPR.2009.5206725

segment

Mirikharaji, Hamarneh
MICCAI 2018
https://link.springer.com/chapter/10.1007/978-3-030-00937-3_84

Izadi, Mirikharaji, Kawahara, Hamarneh
ISBI 2018

hair removal

Mirzaalian, Hamarneh, Lee
IEEE TIP 2014

manual predicted

Taskonomy CVPR 2018
Zamir, Savarese, Malik

https://doi.ieeecomputersociety.org/10.1109/CVPR.2009.5206725
Other prediction tasks

Joint skin lesion localization and segmentation

Vesal, Patil, Ravikuma, Maier
ISIC 2018
https://link.springer.com/chapter/10.1007/978-3-030-01201-4_31

Blue: detected bounding box
Green: GT lesion boundary
Yellow: SkinNet
Red: Faster-RCNN+SkinNet
Real-life skin images

- viewpoint
- blur
- background
- over/under-exposure
Annotate:
time money variability imprecision

Fully-weakly-un-supervised

100 images with clean labels vs 1000 images with noisy labels?

Adaptively handle noisy annotations via modified deep model optimization
1. learn weight map \( W \) to control pixel contribution to loss
2. \( \uparrow W \) \( \leftrightarrow \) agreement with clean data agreement in loss gradient

Mirikharaji, Yan, Hamarneh, 2019

<table>
<thead>
<tr>
<th></th>
<th>data</th>
<th>original</th>
<th>modified</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100 clean</td>
<td>78.6</td>
<td>80.7</td>
</tr>
<tr>
<td></td>
<td>1500 noisy</td>
<td>76.1</td>
<td>80.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>75.0</td>
<td>79.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>73.0</td>
<td>73.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>70.5</td>
<td></td>
</tr>
</tbody>
</table>
Annotations

with only ~100 clean annotations, the reweighting outperforms the fully-supervised with 2000 clean annotations.

leveraging 10 clean annotations: 21.79% improvement in test Dice

U-Net results when all annotation are noisy

U-Net results when all annotation are Clean
Loss choice

1. Basal cell carcinoma (BCC)
2. Nevus (multiple) (NEV)
3. Melanoma (multiple) (MEL)
4. Seborrheic keratosis (SK)
5. Miscellaneous (MISC)

Diagnosis

1. Pigment Network (PN)
2. Blue Whitish Veil (BWV)
3. Vascular Structures (VS)
4. Pigmentation (PIG)
5. Streaks (STR)
6. Dots and Globules (DaG)
7. Regression Structures (RS)

7-point Criteria

7 Criteria

The j-th criterion from the 7-point Criteria

1. Pigment Network (PN)
2. Blue Whitish Veil (BWV)
3. Vascular Structures (VS)
4. Pigmentation (PIG)
5. Streaks (STR)
6. Dots and Globules (DaG)
7. Regression Structures (RS)

CNN parameters

weighted cross-entropy

\[ L(x, y, z; \theta) = \ell(x, y; \theta) + \sum_{j=1}^{7} \ell(x, z_j; \theta) \]
Weighted cross-entropy loss

\[ \ell(x, c; \theta) = -\frac{1}{N_b} \sum_{i=1}^{N_b} \sum_{j=1}^{N_c} w(c)_j \cdot c_j^{(i)} \cdot \log \left( \phi(x^{(i)}; \theta)_{c,j} \right) \]

Given input data e.g., image

Given class labels (e.g., diagnosis or 7-point criteria)

Number of labels in a class

Number of samples in a mini-batch

Weighting function

Standard cross-entropy loss

Predicted log-likelihood of the j-th label for the c-th class of the i-th sample
Class imbalance

Context: Melanoma diagnosis; 7 dermoscopic criteria

Ensure each mini-batch includes \( \geq k \) (random) samples from each label

1. Pigment Network (PN)
2. Blue Whitish Veil (BWV)
3. Vascular Structures (VS)
4. Pigmentation (PIG)
5. Streaks (STR)
6. Dots and Globules (DaG)
7. Regression Structures (RS)

Higher cross entropy weights assigned to infrequent labels in a mini-batch
Data augmentation

Original
Saturation
Brightness
Contrast

Hue

Affine

Flip

Crop

Erase

Elastic
Lesion Mix

F→D→E→C

F→G→D→E→C

F→D→H→E→C

I→F→D→E→C

Data Augmentation for Skin Lesion Analysis
Perez, Vasconcelos, Avila, Valle
ISIC 2018
https://link.springer.com/chapter/10.1007/978-3-030-01201-4_33
Data augmentation

Simulation via Physically- and Statistically-based Warps

**DeformIt**

MICCAI 2008
Hamarneh, Jassi, Tang, Booth
https://link.springer.com/chapter/10.1007/978-3-540-85988-8_55

\[ I = \bar{I} + \alpha Pb + (1 - \alpha) \Phi u \]

- variational
- PCA
- Vibrational
- FEM

↑data ➔ ↑α

rely more on statistical model and less on knowledge-based models
Augmentation

Hair Occlusion Simulator

**HairSim**
IEEE TIP 2014
Mirzaalian, Lee, Hamarneh
https://ieeexplore.ieee.org/document/6918479

- medial A-B curve synthesizer
- hair-thickening: dilation radius $\propto$ geodesic distance to A and B

$$r(p) = \min\{T, \alpha \Gamma(p, A), \alpha \Gamma(p, B)\}$$

- New image (H): blending of clean image I with a colored C hair mask M, Hair color C

$$\begin{bmatrix}
H_R \\
H_G \\
H_B
\end{bmatrix} = I (1 - G_\sigma \ast \mathbf{M}) + \begin{bmatrix}
C_R \\
C_G \\
C_B
\end{bmatrix} (G_\sigma \ast \mathbf{M})$$
Augmentation
Generative Adversarial Networks GANS

Generating Highly Realistic Images of Skin Lesions with GANs (MelanoGan)
Baur, Albarqouni, Navab | ISIC 2018
https://link.springer.com/chapter/10.1007/978-3-030-01201-4_28

Skin Lesion Synthesis with GANs
Bissoto, Perez, Valle, Avila | ISIC 2018
https://link.springer.com/chapter/10.1007/978-3-030-01201-4_32

Augmenting data with GANs to segment melanoma skin lesions
Pollastri, Bolelli, Paredes, Grana | Multimedia Tools and Applications 2019
Augmentation

GAN-based **Mask2Lesion** translation
Abhishek, Hamarneh. 2019

Network architecture

Deep auto-context FCN for skin lesion segmentation
Mirikharaji, Izadi, Kawahara, Hamarneh  ISBI2018
https://ieeexplore.ieee.org/document/8363711

Generative adversarial networks to segment skin lesions
Izadi, Mirikharaji, Kawahara, Hamarneh  ISBI 2018

diagnosis & 7-point derm. criteria
Network architecture search (NAS)

Arrangement of layers or blocks of layers

Neural Architecture Search: A Survey
Thomas Elsken, Jan Hendrik Metzen, Frank Hutter; 20(55):1–21, 2019

© Fjodor Van Veen, The Asimov Institute, 2017
https://www.asimovinstitute.org/author/fjodorvanveen
Network architecture search (NAS)

Exploring Randomly Wired Neural Networks for Image Recognition

Saining Xie  Alexander Kirillov  Ross Girshick  Kaiming He
Facebook AI Research (FAIR)

Exploring Randomly Wired Neural Networks for Image Recognition
Xie, Kirillov, Girshick, He  2019
Layer design

**Layers:** [de]conv, fully conn., [un]pool, sequence eg LSTM, activation eg RELU

Radial Basis Function RBF layer with learnable width $\beta$, center $c$, transformation $\Psi$, bias $b$

$$e^{-\beta (f-c)'} \Psi (f-c) + b$$

Asgari, Abhishek, Azizi, Hamarneh
CVPR 2019
https://arxiv.org/abs/1903.01015
Layer design

Radial Basis Function RBF layer with learnable width $\beta$, center $c$, transformation $\Psi$, bias $b$
Adversarial attacks

Imperceptible changes to images crafted to make DNNs produce specific output

Finlayson, Bowers, Ito, Zittrain, Beam, Kohane
Science 2019
https://science.sciencemag.org/content/363/6433/1287

Asgari, Abhishek, Azizi, Hamarneh
CVPR 2019
https://arxiv.org/abs/1903.01015
Machine Learning and Health Care Disparities in Dermatology

Unfortunately, most ML programs are largely learning on light skin. For example, in the International Skin Imaging Collaboration: Melanoma Project, which is one of the largest and often-used, open-source, public-access archives of pigmented lesions, much of the patient data are heavily collected from fair-skinned populations in the United States, Europe, and Australia. Thus, no matter how advanced the ML algorithm, it may underperform on images of lesions in skin of color.

JAMA Dermatology 2018; 154(11)
Adamson, Smith
https://jamanetwork.com/journals/jamadermatology/article-abstract/2688587

Classification of the Clinical Images for Benign and Malignant Cutaneous Tumors Using a Deep Learning Algorithm

Seung Seog Han, Myoung Shin Kim, Woohyung Lim, Gyeong Hun Park, Ilwoo Park, and Sung Eun Chang

Because of the different patient demographics in the three validation datasets we tested with our algorithm, the sensitivity and specificity of these datasets were analyzed over a change in threshold from 0.0000 to 1.0000 (Figure 4). The sensitivities of the Asan and Hallym test dataset over this threshold were similar. However, the specificities for BCC, squamous cell carcinoma, and melanoma between the Asan test dataset and Edinburgh dataset showed substantial differences, which may have been due to malignancy subtypes and the skin colors around the lesions. It may be necessary, therefore, to choose different thresholds or generate different models for different ethnic groups.

Journal of Investigative Dermatology 2018; 138(7)
Han et al.
https://jamanetwork.com/journals/jamadermatology/article-abstract/2688587
**Dataset shift**

**Bias and fairness**

Test dataset of European population:
10 classes - 1300 images

**Train and test on same dataset**

Deep features to classify skin lesions
Kawahara, BenTaieb, Hamarneh
ISBI 2016
https://ieeexplore.ieee.org/document/7493528

**Train on Asian, test on European**

Classification of the Clinical Images for Benign and Malignant Cutaneous Tumors Using a Deep Learning Algorithms
Han, Kim, Lim, Park, Park, Chang
Journal of Investigative Dermatology
https://www.jidonline.org/article/S0022-202X(18)30111-8/
Dataset shift

MICCAI 2019  Yoon, Hamarneh, Garbi

7 Domains:
1 primary:  HAM10000
6 secondary: Dermofit+MSK+UDA+ONIC+Derm7pt+PH2
n_s samples/class

CCSA loss: classification & contrastive semantic alignment
[Motiian ICCV 2017] CE loss + feature alignment/separation losses

Class imbalance:
Intra-domain  P(nevus) >> P(melanoma)
Inter-domain  dermatofibroma \notin Domain2

Dynamic sampling
two image-label pairs across domain: \((x_1, y_1), (x_2, y_2)\)
Adaptive weighting
of CCSA loss based on \(P(y = c_i)\) and \(P(y_1 = c_i, y_2 = c_j)\)
Interpretability / explainability

Activation and attention maps

Selvaraju, Cogswell, Das, Vedantam, Parikh, Batra. Grad-CAM. ICCV 2017

Melanoma Recognition via Visual Attention
Yan, Kawahara, Hamarneh. IPMI 2019
https://link.springer.com/chapter/10.1007/978-3-030-20351-1_62

Guide (add prior to) the attention maps to ROIs known to discriminatory:

\[ \mathcal{L}_D(\mathcal{A}, \tilde{\mathcal{A}}) = 1 - D(\mathcal{A}, \tilde{\mathcal{A}}) \]

\[ \mathcal{L} = \mathcal{L}_{focal} + \lambda_1 \mathcal{L}_D(\mathcal{A}^{(3)}, \tilde{\mathcal{A}}^{(3)}) + \lambda_2 \mathcal{L}_D(\mathcal{A}^{(4)}, \tilde{\mathcal{A}}^{(4)}) \]
Longitudinal tracking

Mirzaalian, Lee, Hamarneh
https://ieeexplore.ieee.org/document/7164139
Longitudinal tracking
Longitudinal tracking
Visual communication

Beyond reporting predicted class and probabilities

Image Content-Based Navigation of Skin Conditions
Kawahara, Hamarneh. WCD 2015

Graph Geodesics to Find Progressively Similar Skin Lesion Images
Kawahara, Moriarty, Hamarneh. MICCAI GRAIL 2017
https://link.springer.com/chapter/10.1007/978-3-319-67675-3_4
Multi-modal input

Clinical images
Dermoscopic images
Meta-data

\[ L(x, y, z; \theta) = \ell(x, y; \theta) + \sum_{j=1}^{7} \ell(x, z_j; \theta) \]

\[ \mathcal{L}(x_d, x_c, x_m, y, z; \theta) = \mathcal{L}((x_d, x_c, x_m), y, z; \theta_{dcm}) + \mathcal{L}(x_d, y, z; \theta_d) + \mathcal{L}((x_d, x_m), y, z; \theta_{dm}) + \mathcal{L}(x_c, y, z; \theta_c) + \mathcal{L}((x_c, x_m), y, z; \theta_{cm}) \]
Multi-modal input

Lesion metadata
body location, roughness / elevation (flat, palpable, nodular)

Patient data: age, gender, race, history

ARE NEURAL NETWORKS EFFECTIVE IN DETECTING MELANOMA USING GENOMIC DATA?

Abder-Rahman Ali, Sally Jane O'Shea, Jingpeng Li
1. Faculty of Natural Sciences, Computing Science and Mathematics, University of Stirling, UK.
2. Dermatology Department, Mater Private Hospital Cork, Ireland.
3. Faculty of Medicine, University College Cork, Ireland.

EBioMedicine 43 (2019) 107-113
Skin cancer detection by deep learning and sound analysis algorithms:
A prospective clinical study of an elementary dermoscope

A. Dascalu *, E.O. David
* Department of Physiology and Pharmacology, Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel.

Pathologist-level classification of histopathological melanoma images with deep neural networks

### What’s next?

<table>
<thead>
<tr>
<th>Category</th>
<th>Next Step</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease classes</td>
<td>&lt;10 → 1000s</td>
</tr>
<tr>
<td>Datasets</td>
<td>100s/1000s → millions of images</td>
</tr>
<tr>
<td>Training</td>
<td>full-supervision → leveraging weak/no supervision</td>
</tr>
<tr>
<td>Data sources</td>
<td>homogenous → highly heterogenous sources</td>
</tr>
<tr>
<td></td>
<td>controlled → real-world</td>
</tr>
<tr>
<td>Dimensions</td>
<td>2D + static → 3D + longitudinal/dynamic</td>
</tr>
<tr>
<td>Modalities</td>
<td>unimodal → multi-modal</td>
</tr>
<tr>
<td>Deep modes</td>
<td>hand-crafted → automatic</td>
</tr>
<tr>
<td></td>
<td>data-driven → hybrid knowledge- &amp; data-driven models</td>
</tr>
<tr>
<td></td>
<td>black-box → interpretable</td>
</tr>
<tr>
<td></td>
<td>susceptible → resilient to adversarial attacks</td>
</tr>
<tr>
<td>Beyond technical</td>
<td>communities → tighter computational-clinical collaboration</td>
</tr>
<tr>
<td></td>
<td>legal, ethical, societal, economic challenges</td>
</tr>
</tbody>
</table>
Thank you!

hamarneh@sfu.ca