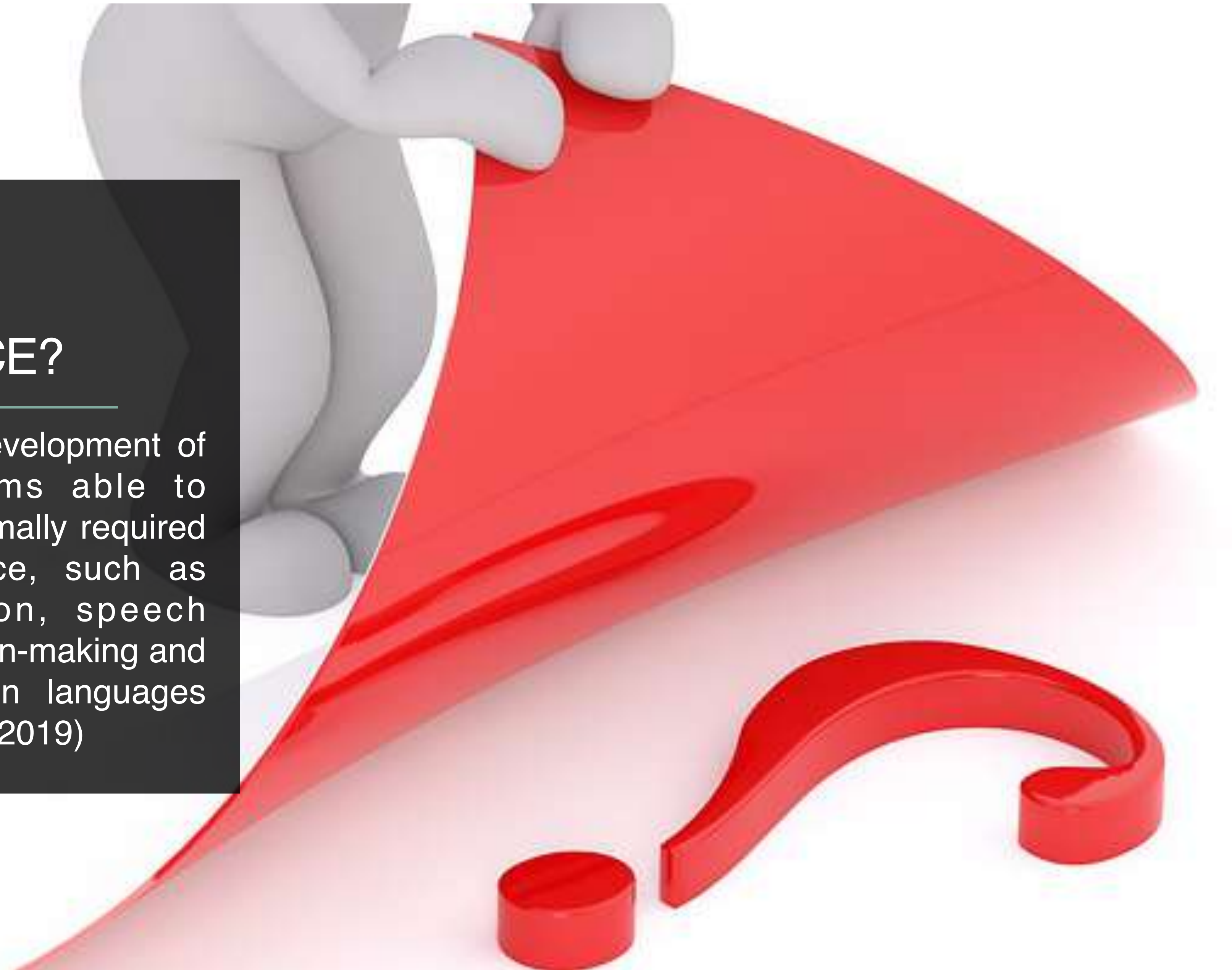
COMPUTER-ASSISTED DIAGNOSIS OF SKIN CANCER

THE DERMATOLOGY POINT OF VIEW

PROF.

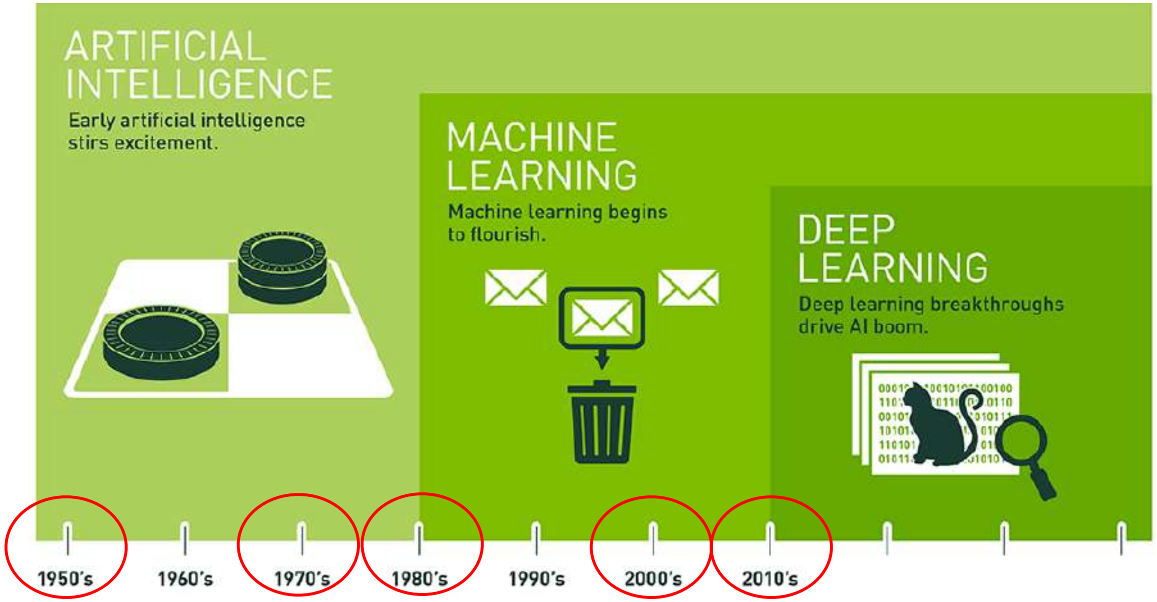
GIOVANNI PELLACANI

1. WHAT ARTIFICIAL INTELLIGENCE IS?

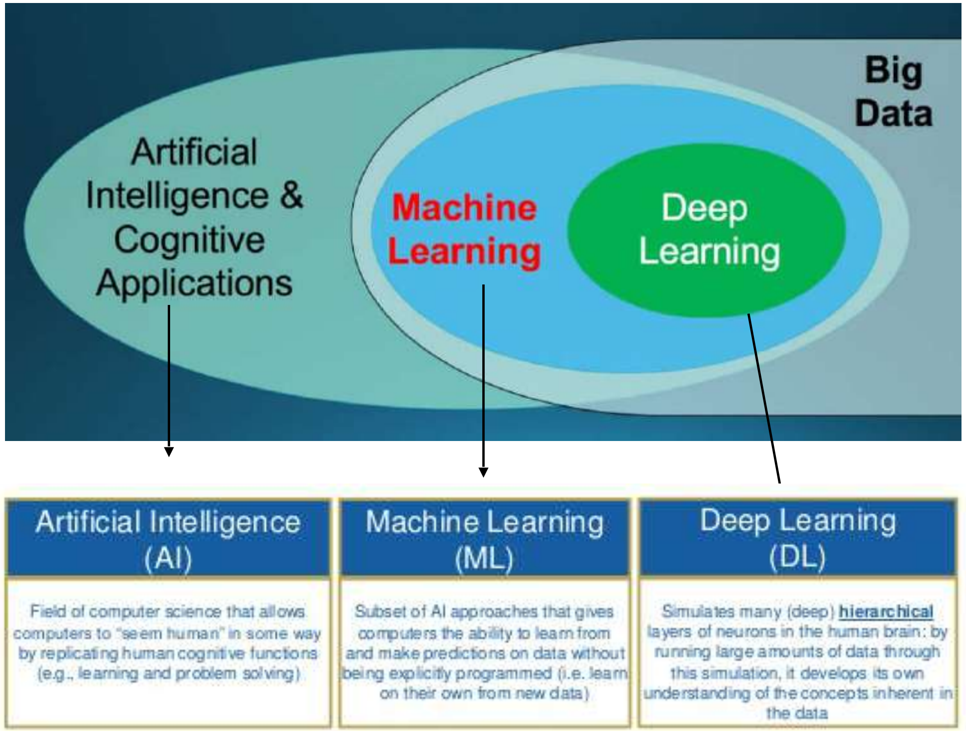


WHAT IS ARTIFICIAL INTELLIGENCE?

The theory and development of computer systems able to perform tasks, normally required human intelligence, such as visual perception, speech recognition, decision-making and translation between languages (Oxford Dictionary, 2019)



Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.



1956: Dartmouth conference: birth of A.I.

1970: medical researches discovered the applicability of AI in life sciences (limited in dermatology)

1980: machine learning

2006: "deep learning" (Hinton et al)

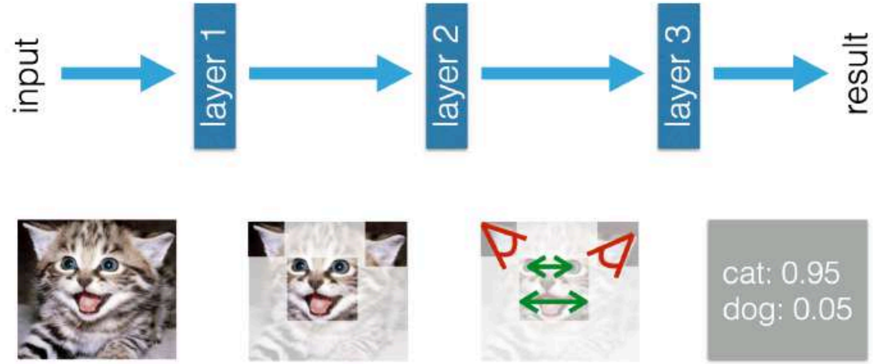
2012: "CNN" (convolutional neural network) for image processing, speech recognition, text processing and… for medical sciences (**also dermatology**)

A.I.: classical approach

- Round face
- 2 triangular ears
- 2 eyes
- Tail
- Rectangular body

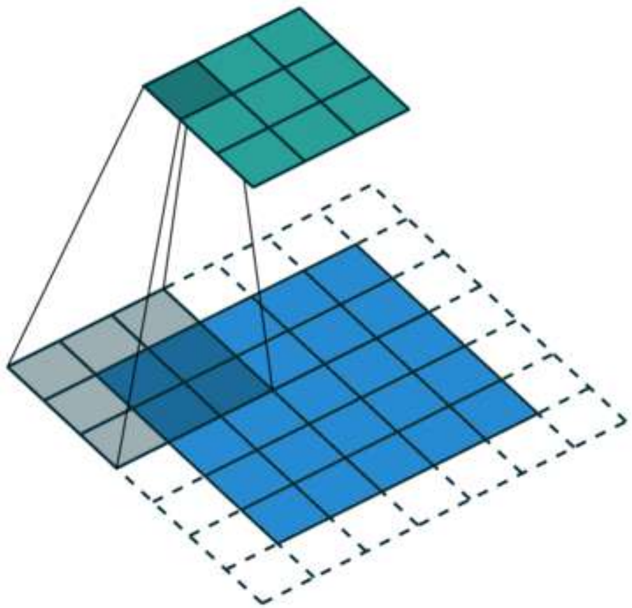
"CAT"

Deep learning (Convolutional Neural Network)



Statistical model based on brain network (neuronal units connected in different layers), to analyze automatically features (from simple features to complex features) → output (probability)

Convolutional Neural Networks



Convolutional Neural Networks are designed by stacking layers of convolutional filters

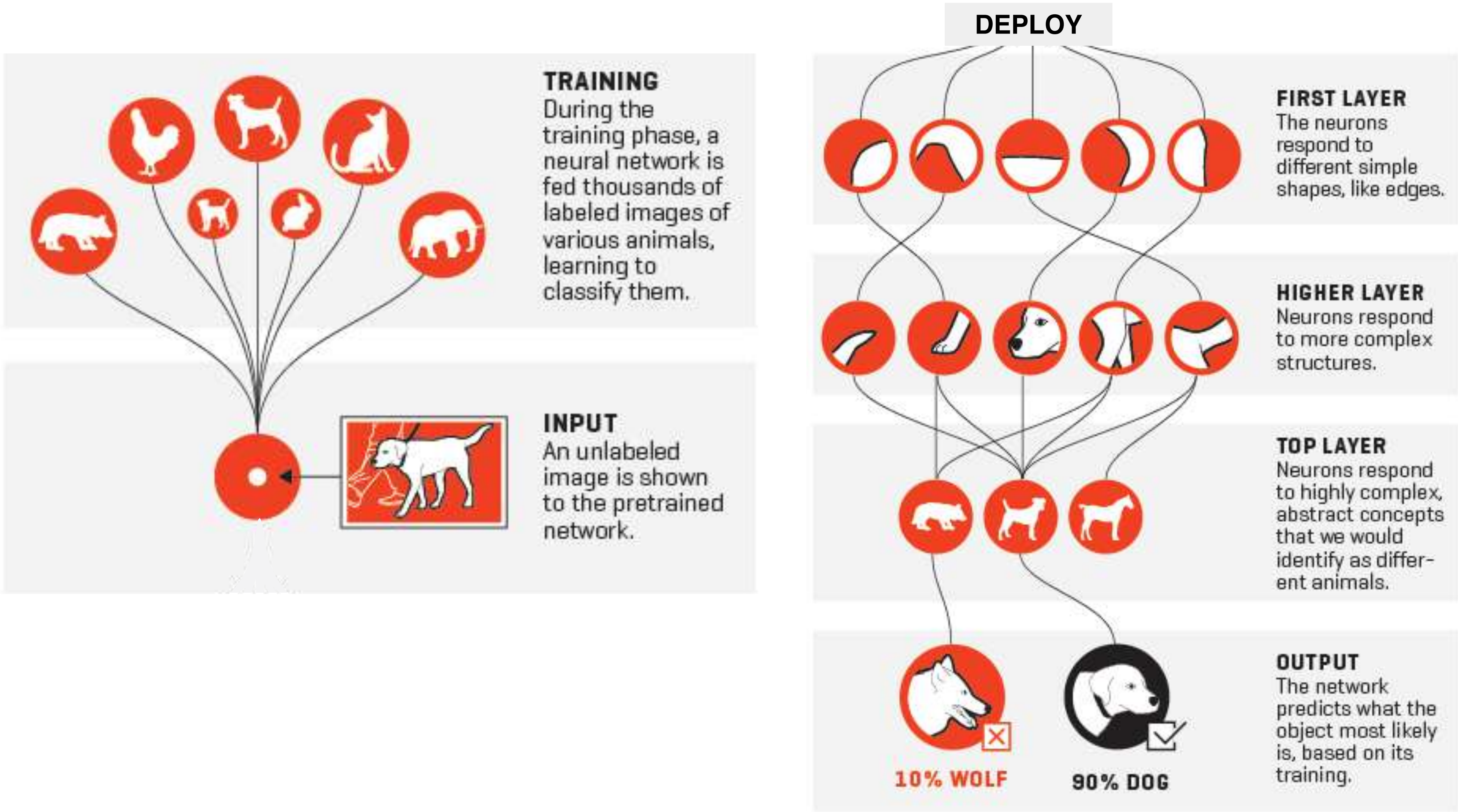
Convolutional Filters can learn to extract features from the input image

The first convolutional layer is responsible for capturing low-level features such as edges

The following layers adapt to extract high-level features such as object shapes

Convolutional Neural Networks gain a wholesome understainding of images drawn from a certain dataset

HOW NEURAL NETWORKS RECOGNIZE A DOG IN A PHOTO



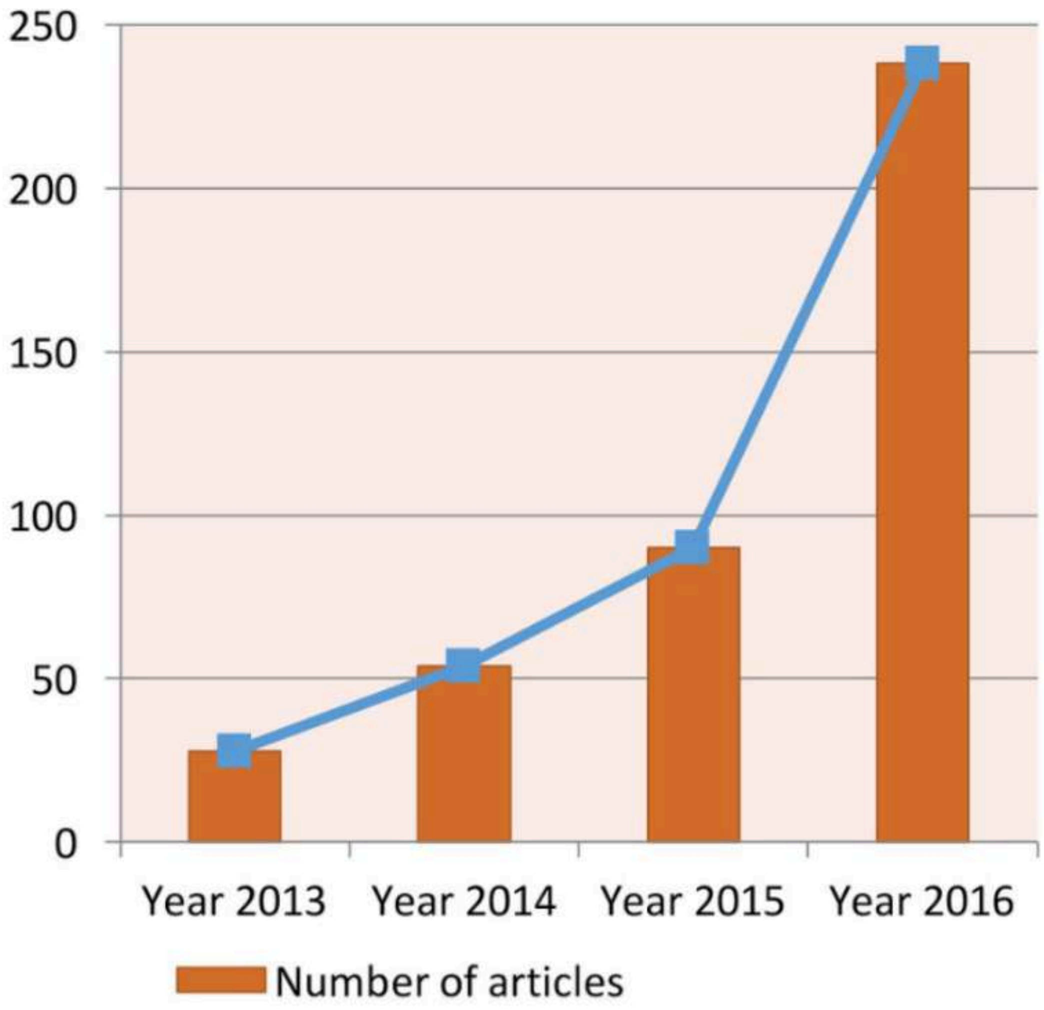
2. ARTIFICIAL INTELLIGENCE AND SKIN LESIONS

WHY ARTIFICIAL INTELLIGENCE IN DIAGNOSIS OF SKIN LESIONS?

- In most countries , primary vigilance is maintained through primary care clinics before being referred to dermatologists → primary care doctors are under a heavy burden to correctly screen patients
- Paucity of experts dermatologists
- Rising incidence of skin cancer in aging population
- Early diagnosis of skin cancer

↓

HIGH DEMAND FOR POINT-OF-CARE DECISION **SUPPORT SYSTEMS** TO DIAGNOSE SKIN LESIONS WITHOUT THE NEED OF HUMAN EXPERTISE



An improved strategy for **skin lesion** detection and classification using uniform segmentation and feature selection based approach.
Nasir M, Attique Khan M, Sharif M, Lali IU, Saba T, Iqbal T.
Microsc Res Tech. 2018 Jun;81(6):528-543. doi: 10.1002/jemt.23009. Epub 2018 Feb 21.
PMID: 29464868
Similar articles

A Review of Denoising Medical Images Using **Machine Learning** Approaches.
Kaur P, Singh G, Kaur P.
Curr Med Imaging Rev. 2018 Oct;14(5):675-685. doi: 10.2174/1573405613666170428154156. Review.
PMID: 30532667 **Free PMC Article**
Similar articles

Risk-Aware **Machine Learning** Classifier for **Skin Lesion** Diagnosis.
Mobiny A, Singh A, Van Nguyen H.
J Clin Med. 2019 Aug 17;8(8). pii: E1241. doi: 10.3390/jcm8081241.
PMID: 31426482 **Free PMC Article**
Similar articles

Rethinking **Skin Lesion** Segmentation in a Convolutional Classifier.
Burdick J, Marques O, Weinthal J, Furht B.
J Digit Imaging. 2018 Aug;31(4):435-440. doi: 10.1007/s10278-017-0026-y. Review.
PMID: 29047032 **Free PMC Article**
Similar articles

Rapid and accurate intraoperative pathological diagnosis by artificial intelligence with deep **learning** technology.
Zhang J, Song Y, Xia F, Zhu C, Zhang Y, Song W, Xu J, Ma X.
Med Hypotheses. 2017 Sep;107:98-99. doi: 10.1016/j.mehy.2017.08.021. Epub 2017 Sep 1.
PMID: 28915974
Similar articles

A novel cumulative level difference mean based GLDM and modified ABCD features ranked using eigenvector centrality approach for four **skin lesion** types classification.
Wahba MA, Ashour AS, Guo Y, Napoleon SA, Elnaby MMA.
Comput Methods Programs Biomed. 2018 Oct;165:163-174. doi: 10.1016/j.cmpb.2018.08.009. Epub 2018 Aug 24.
PMID: 30337071
Similar articles

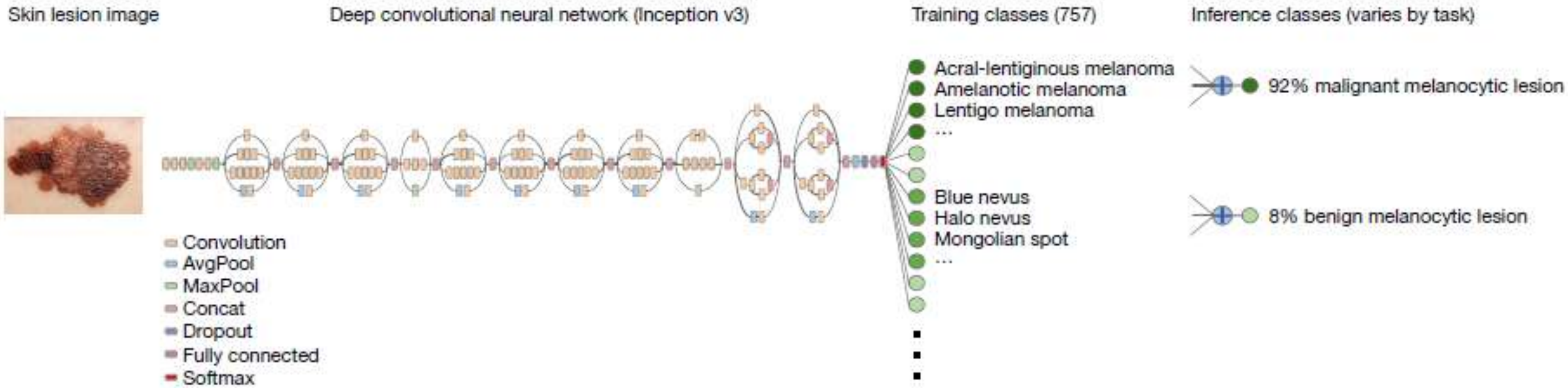
Automatic discrimination of actinic keratoses from clinical photographs.
Spyridonos P, Gaitanis G, Likas A, Bassukas ID.
Comput Biol Med. 2017 Sep 1;88:50-59. doi: 10.1016/j.compbiomed.2017.07.001. Epub 2017 Jul 3.
PMID: 28692931
Similar articles

LIGHTS…

…AND SHADOWS

Dermatologist-level classification of skin cancer with deep neural networks

Andre Esteva[1]*, Brett Kuprel[1]*, Roberto A. Novoa[2,3], Justin Ko[2], Susan M. Swetter[2,4], Helen M. Blau[5] & Sebastian Thrun[6]

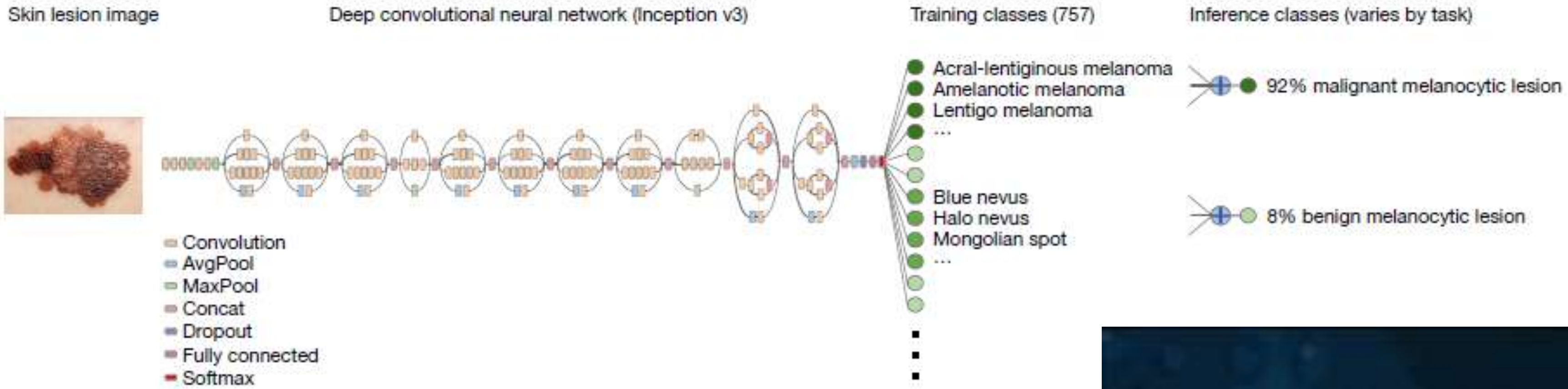


and disease labels as inputs. We train a CNN using a dataset of 129,450 clinical images—two orders of magnitude larger than previous datasets[12]—

The CNN achieves performance on par with all tested experts across both tasks, demonstrating an artificial intelligence capable of classifying skin cancer with a level of competence comparable to dermatologists.

Dermatologist-level classification of skin cancer with deep neural networks

Andre Esteva[1]*, Brett Kuprel[1]*, Roberto A. Novoa[2,3], Justin Ko[2], Susan M. Swetter[2,4], Helen M. Blau[5] & Sebastian Thrun[6]



and disease labels as inputs. We train a CNN using a dataset of 129,450 clinical images—two orders of magnitude larger than previous datasets[12]—consisting of 2,032 different diseases.

= average 63.7 images per diagnosis!!!

The C ... ts across both tasks, demonstrating an artificial intelligence capable of classifying skin cancer with a level of competence comparable to dermatologists.

THE LANCET
Oncology

ARTICLES | VOLUME 20, ISSUE 7, P938-947, JULY 01, 2019

Purchase Subscribe Save

Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study

Philipp Tschandl, PhD • Noel Codella, PhD • Bengü Nisa Akay, MD • Prof Giuseppe Argenziano, PhD • Ralph P Braun, MD • Prof Horacio Cabo, MD • et al. Show all authors

Published: June 11, 2019 • DOI: https://doi.org/10.1016/S1470-2045(19)30333-X • Check for updates

STRENGHT …

When comparing all human readers with all machine-learning algorithms, the algorithms achieved a mean of 2·01 (95% CI 1·97 to 2·04; p<0·0001) more correct diagnoses (17·91 [SD 3·42] *vs* 19·92 [4·27]).


THE LANCET
Oncology

ARTICLES | VOLUME 20, ISSUE 7, P938-947, JULY 01, 2019

Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study

Philipp Tschandl, PhD • Noel Codella, PhD • Bengü Nisa Akay, MD • Prof Giuseppe Argenziano, PhD • Ralph P Braun, MD • Prof Horacio Cabo, MD • et al. Show all authors

Published: June 11, 2019 • DOI: https://doi.org/10.1016/S1470-2045(19)30333-X •

STRENGHT …

LIMITATIONS

When comparing all human readers with all machine-learning algorithms, the algorithms achieved a mean of 2·01 (95% CI 1·97 to 2·04; $p<0\cdot0001$) more correct diagnoses (17·91 [SD 3·42] *vs* 19·92 [4·27]).

- The metrics **treated all diagnoses equally** (the algorithms did not consider that it is more detrimental to mistake a malignant for a benign lesion than viceversa)
- **Lesions** in the test set and training set were not standardized
- **Absence of additional data** (anatomical site, age, and sex)
- **Low sensitivity** if used image from source different than training data (**test set**)

JAMA Dermatology | Original Investigation

Accuracy of Computer-Aided Diagnosis of Melanoma
A Meta-analysis

Vincent Dick, CandMed; Christoph Sinz, MD; Martina Mittlböck, PhD; Harald Kittler, MD; Philipp Tschandl, MD, PhD

RESULTS The literature search yielded 1694 potentially eligible studies, of which 132 were included and 70 offered sufficient information for a quantitative analysis. Most studies came from the field of computer science. Prospective clinical studies were rare. Combining the results for automated systems gave a melanoma sensitivity of 0.74 (95% CI, 0.66-0.80) and a specificity of 0.84 (95% CI, 0.79-0.88). Sensitivity was lower in studies that used independent test sets than in those that did not (0.51; 95% CI, 0.34-0.69 vs 0.82; 95% CI, 0.77-0.86; $P < .001$); however, the specificity was similar (0.83; 95% CI, 0.71-0.91 vs 0.85; 95% CI, 0.80-0.88; $P = .67$). In comparison with dermatologists' diagnosis, computer-aided diagnosis showed similar sensitivities and a 10 percentage points lower specificity, but the difference was not statistically significant. Studies were heterogeneous and substantial risk of bias was found in all but 4 of the 70 studies included in the quantitative analysis.

CONCLUSIONS AND RELEVANCE Although the accuracy of computer-aided diagnosis for melanoma detection is comparable to that of experts, the real-world applicability of these systems is unknown and potentially limited owing to overfitting and the risk of bias of the studies at hand.

Research

JAMA Dermatology | Original Investigation

Accuracy of Computer-Aided Diagnosis of Melanoma
A Meta-analysis

Vincent Dick, CandMed; Christoph Sinz, MD; Martina Mittlböck, PhD; Harald Kittler, MD; Philipp Tschandl, MD, PhD

LIMITATIONS…

- **Etherogeneous studies** with high risk for **bias**
- **Experimental setting** (smarthpone, pc, tablet…)
- **Low sensitivity WHEN** used images **from different sources** than the training dataset
- Referral centers for melanoma
- Skin lesion images not standardized
- Often use limited to melanocytic lesions ONLY
- Lack of data for unusual presentations and rare skin tumors

Research

JAMA Dermatology | Original Investigation

Accuracy of Computer-Aided Diagnosis of Melanoma
A Meta-analysis

Vincent Dick, CandMed; Christoph Sinz, MD; Martina Mittlböck, PhD; Harald Kittler, MD; Philipp Tschandl, MD, PhD

"There is a fear that less-skilled physicians or even non-medical personnel will use such systems to deliver a service that should be restricted to dermatologists. Therefore, a successful CAD would most probably enhance and support dermatologists rather than replace them."

DERMATOLOGY PRACTICAL & CONCEPTUAL
www.derm101.com

Artificial intelligence and melanoma diagnosis: ignoring human nature may lead to false predictions

Aimilios Lallas[1], Giuseppe Argenziano[2]

> “We are convinced that AI has the potential….to become an additional precious tool in the hand of doctors… the main obstacles of this goal are the misconceptions about our role as doctors”

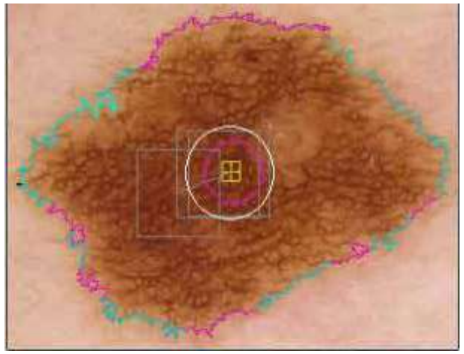
3. WHAT WE DID IN THE PAST

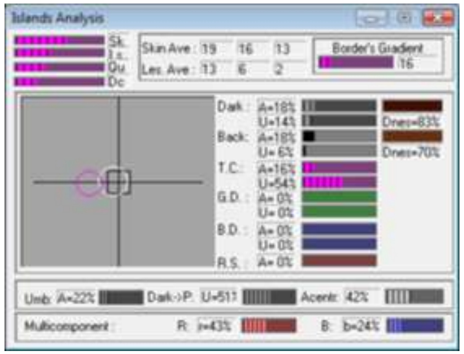
WE STARTED 20 YEARS OLD…

Melanoma Research 1999, 9, pp. 163–171

Digital videomicroscopy and image analysis with automatic classification for detection of thin melanomas

S. Seidenari*, G. Pellacani and A. Giannetti
Department of Dermatology, University of Modena, 41100 Modena, Italy. Tel: (+39) 59 422464; Fax: (+39) 59 424271; Email: seidenar@unimo.it





ARTIFICIAL INTELLIGENCE APPROACH

AND CONTINUED WITH…

Collaboration with:
Prof Costantino Grana
Dept. of Informatic Engineering, Unimore

IEEE Trans Med Imaging. 2003 Aug;22(8):959-64.

A new algorithm for border description of polarized light surface microscopic images of pigmented skin lesions.

Grana C[1], Pellacani G, Cucchiara R, Seidenari S.

Br J Dermatol. 2003 Sep;149(3):523-9.

Computer description of colours in dermoscopic melanocytic lesion images reproducing clinical assessment.

Seidenari S[1], Pellacani G, Grana C.

Melanoma Res. 2004 Apr;14(2):125-30.

Automated description of colours in polarized-light surface microscopy images of melanocytic lesions.

Pellacani G[1], Grana C, Seidenari S.

ARTIFICIAL INTELLIGENCE APPROACH

AND CONTINUED WITH…

Skin Res Technol. 2005 Nov;11(4):236-41.

Pigment distribution in melanocytic lesion images: a digital parameter to be employed for computer-aided diagnosis.

Seidenari S[1], Pellacani G, Grana C.

Acta Derm Venereol. 2006;86(2):123-8.

Asymmetry in dermoscopic melanocytic lesion images: a computer description based on colour distribution.

Seidenari S[1], Pellacani G, Grana C.

J Eur Acad Dermatol Venereol. 2006 Nov;20(10):1214-9.

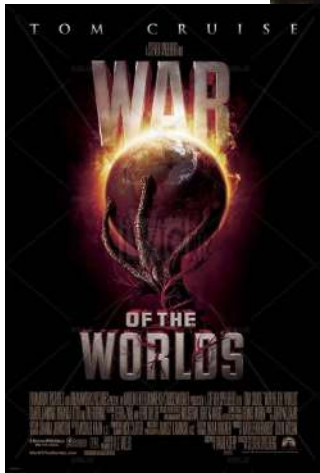
Algorithmic reproduction of asymmetry and border cut-off parameters according to the ABCD rule for dermoscopy.

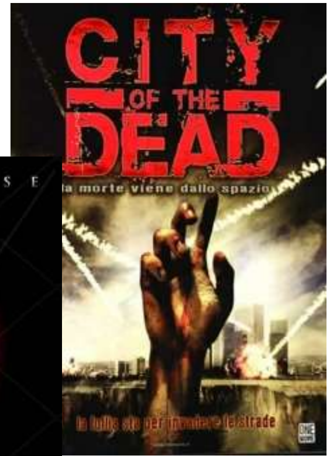
Pellacani G[1], Grana C, Seidenari S.

Dermatology. 2007;214(2):137-43.

Colour clusters for computer diagnosis of melanocytic lesions.

Seidenari S[1], Grana C, Pellacani G.

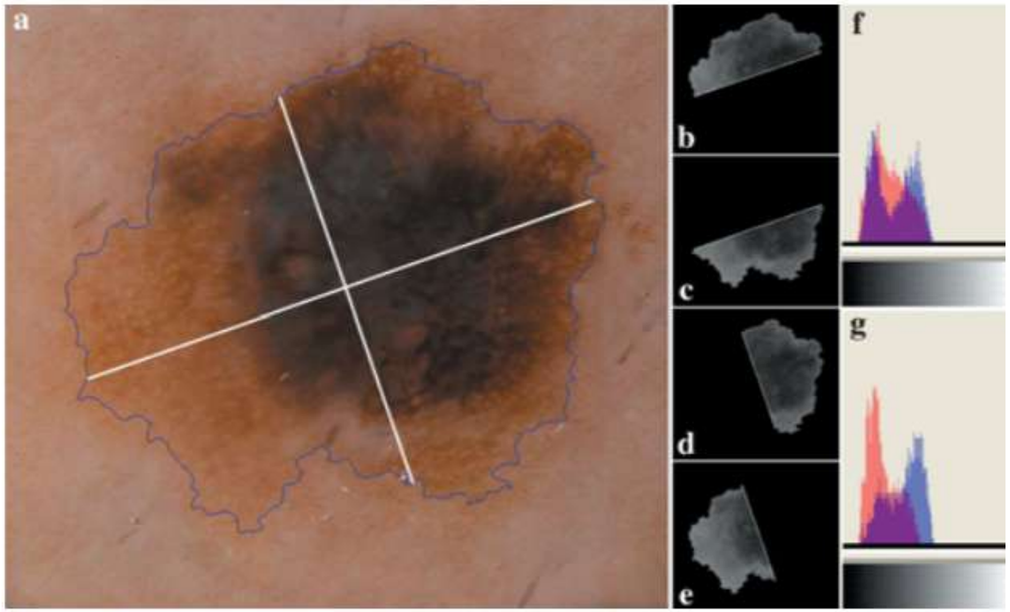




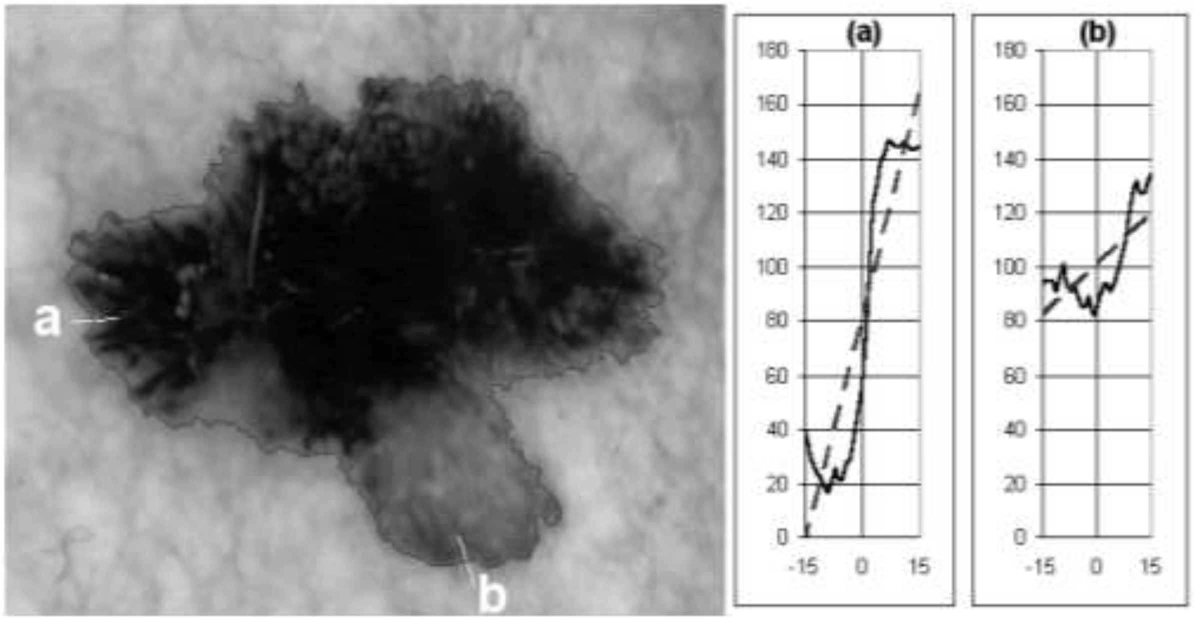
ARTIFICIAL INTELLIGENCE APPROACH

THE MAIN GOAL WAS TO IDENTIFY THE DERMOSCOPIC FEATURES AND TO REPRODUCE THE ABCD RULE

Asymmetry

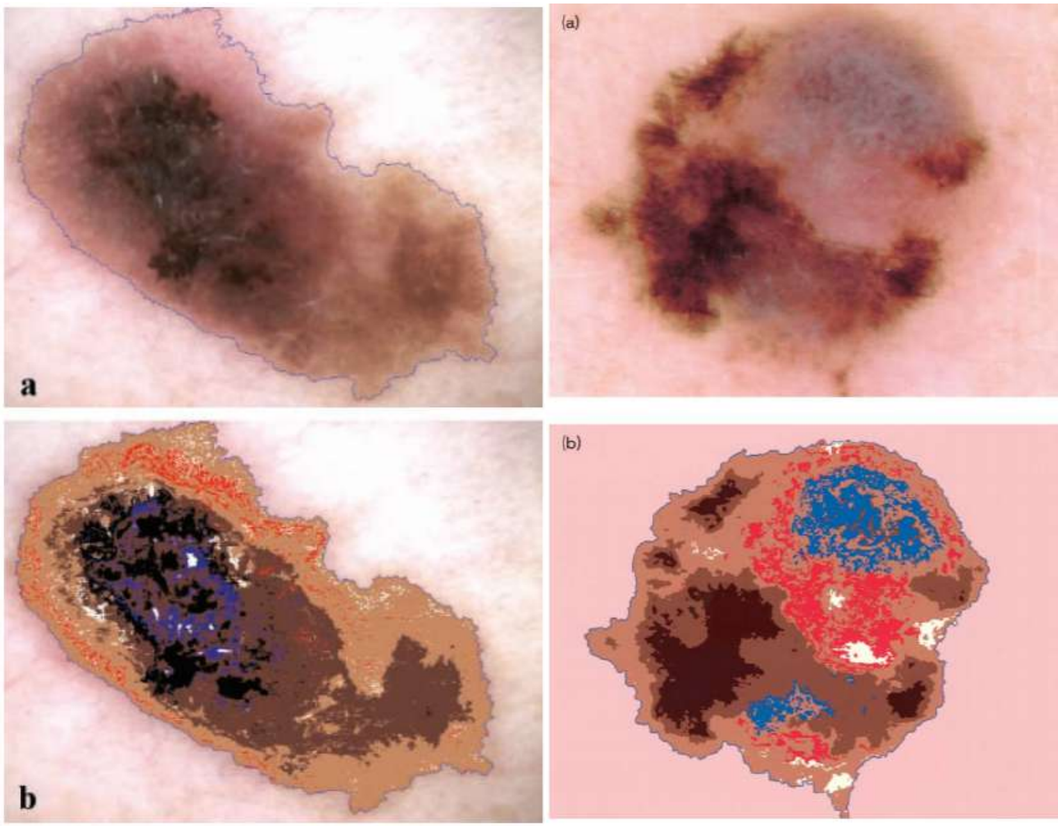


Border

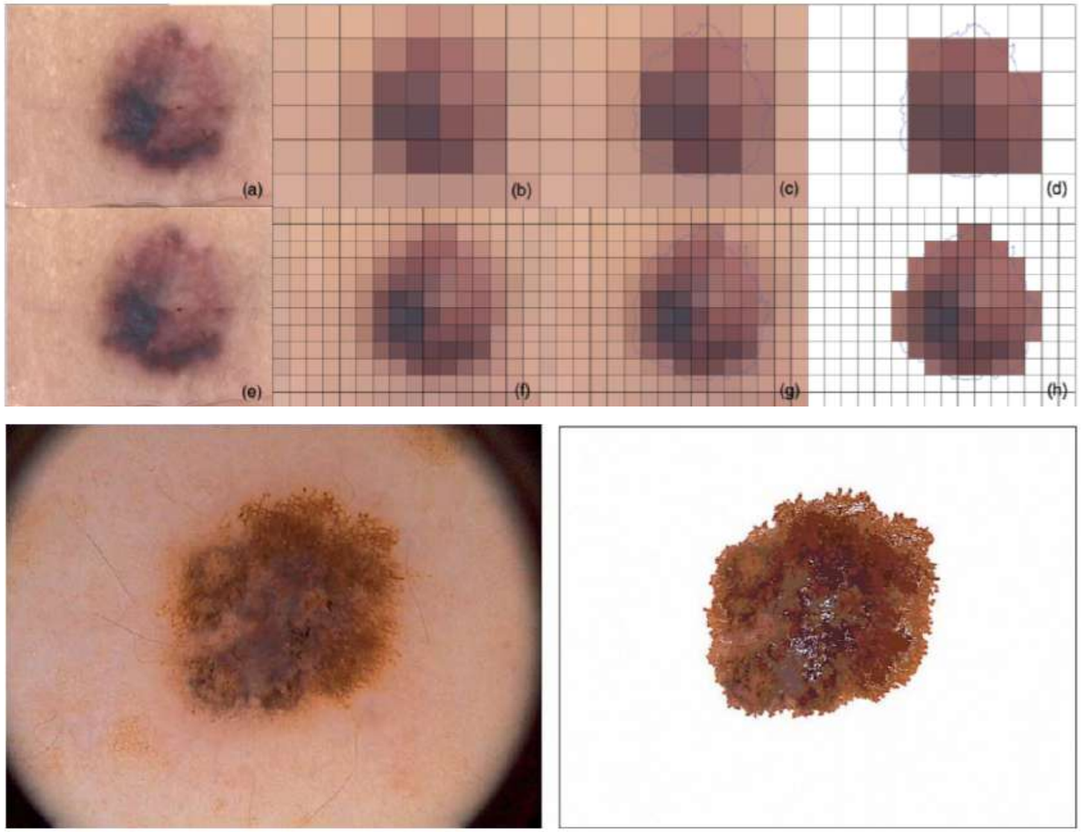


THE MAIN GOAL WAS TO IDENTIFY THE DERMOSCOPIC FEATURES AND TO REPRODUCE THE ABCD RULE

Colours



Dermoscopic structures



4. WHAT’S GOING ON

CONVOLUTED NEURAL NETWORK APPROACH

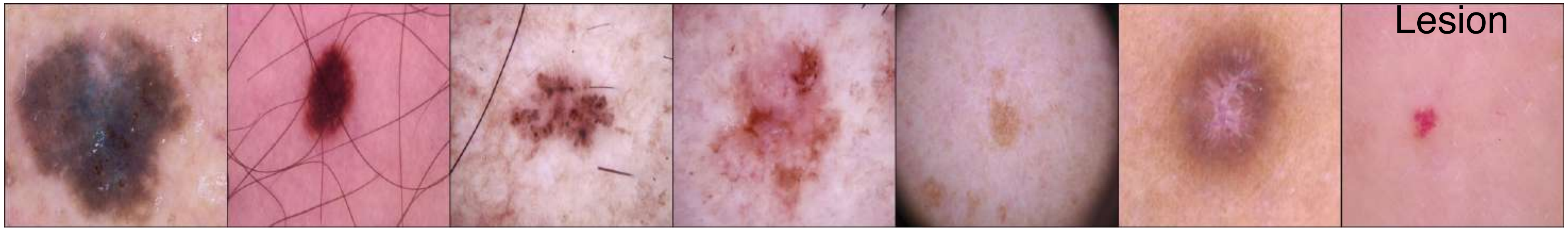
FEATURE EXTRACTION: CLASSIFICATION

Melanoma

Basal Cell Carcinoma

Benign Keratosis

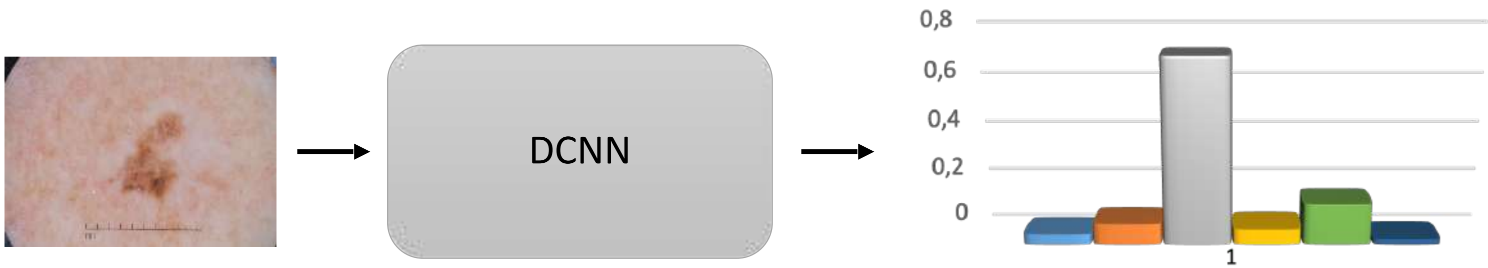
Vascular Lesion



Melanocytic Nevus

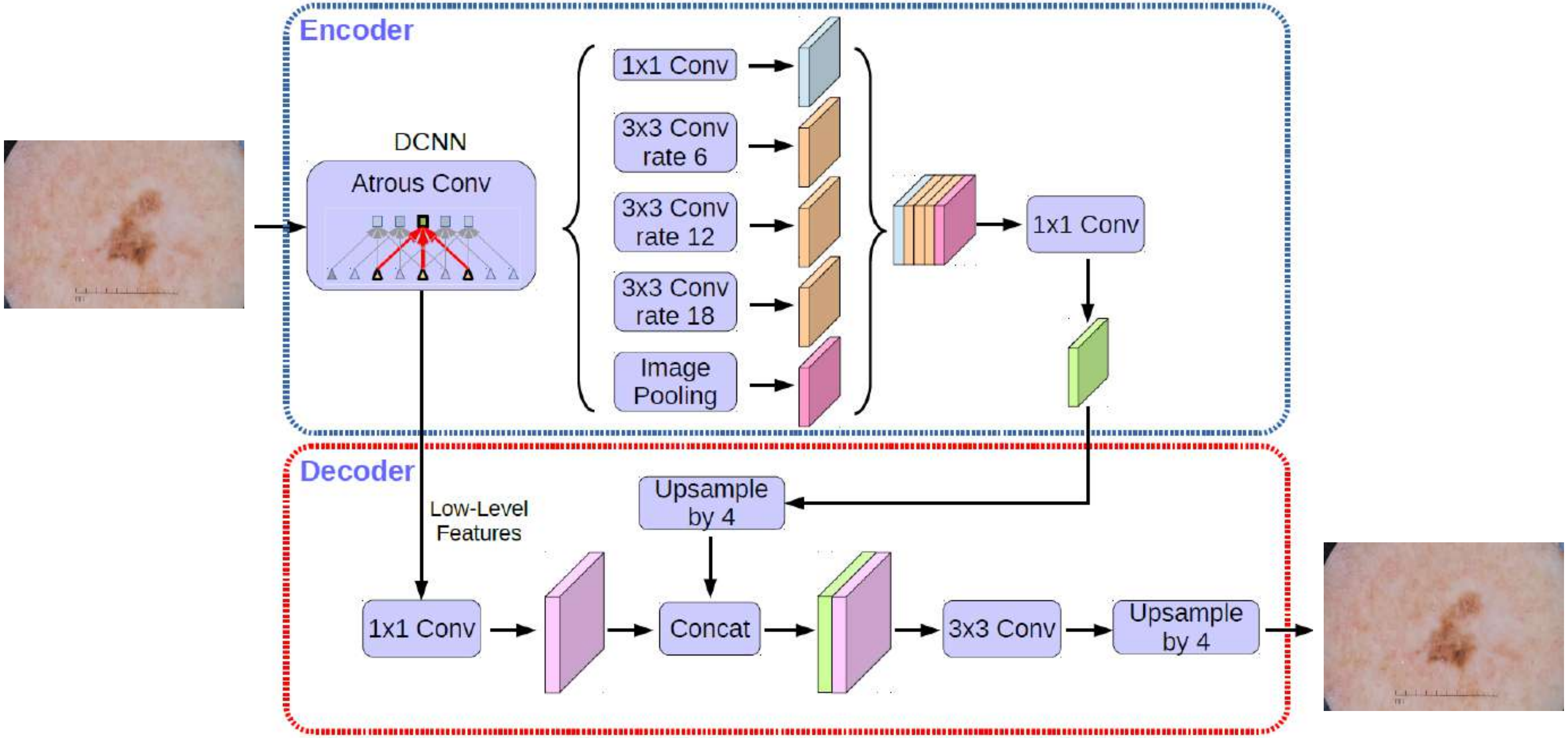
Actinic Keratosis

Dermatofibroma



CONVOLUTED NEURAL NETWORK APPROACH

FEATURE EXTRACTION: AUTO-ENCODER



ISIC Challenge 2019

ISIC 2019

Home Background Data Submission Evaluation Leaderboard Contact

LESION DIAGNOSIS: IMAGES ONLY | LESION DIAGNOSIS: IMAGES AND METADATA

Rank <64 total>	Team <64 unique teams>	Approach Name	Manuscript	Used External Data <19 yes>	Primary Metric Value <Balanced Multiclass Accuracy>	
1	DAISYLab Hamburg University of Technology/University Medical Center Hamburg-Eppendorf	Ensemble of Multi-Res EfficientNets + SEN154 2		Yes	0.636	
2	DysionAI DYSION AI, Inc, Beijing, China	Ensemble of EfficienetB3-B4-Seresnext101		No	0.607	
3	AImageLab & PRHLT Unimore & UPV	ensemble, ood threshold 100%		No	0.593	
4	DermaCode	13 models + hierarchical approach to select outliers		No	0.578	
5	Nurithm Labs Nurithm Labs	Densenet-161 with heavy use of random crops		Yes	0.569	
6	Torus Actions Torus Actions	Simple test approach		No	0.563	

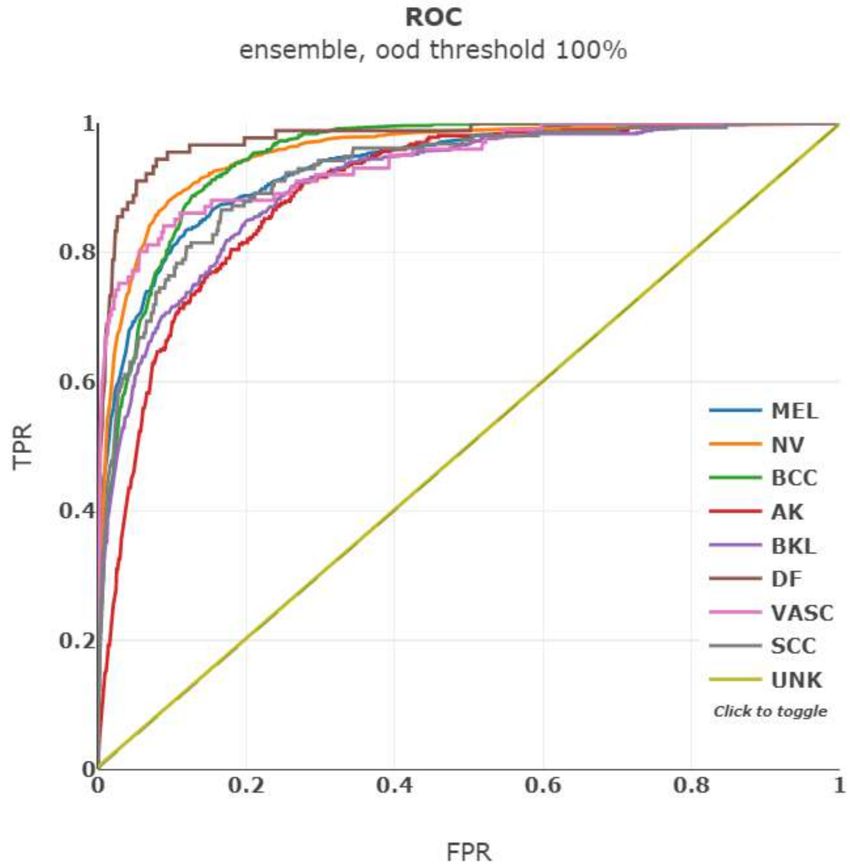
ISIC Challenge 2019

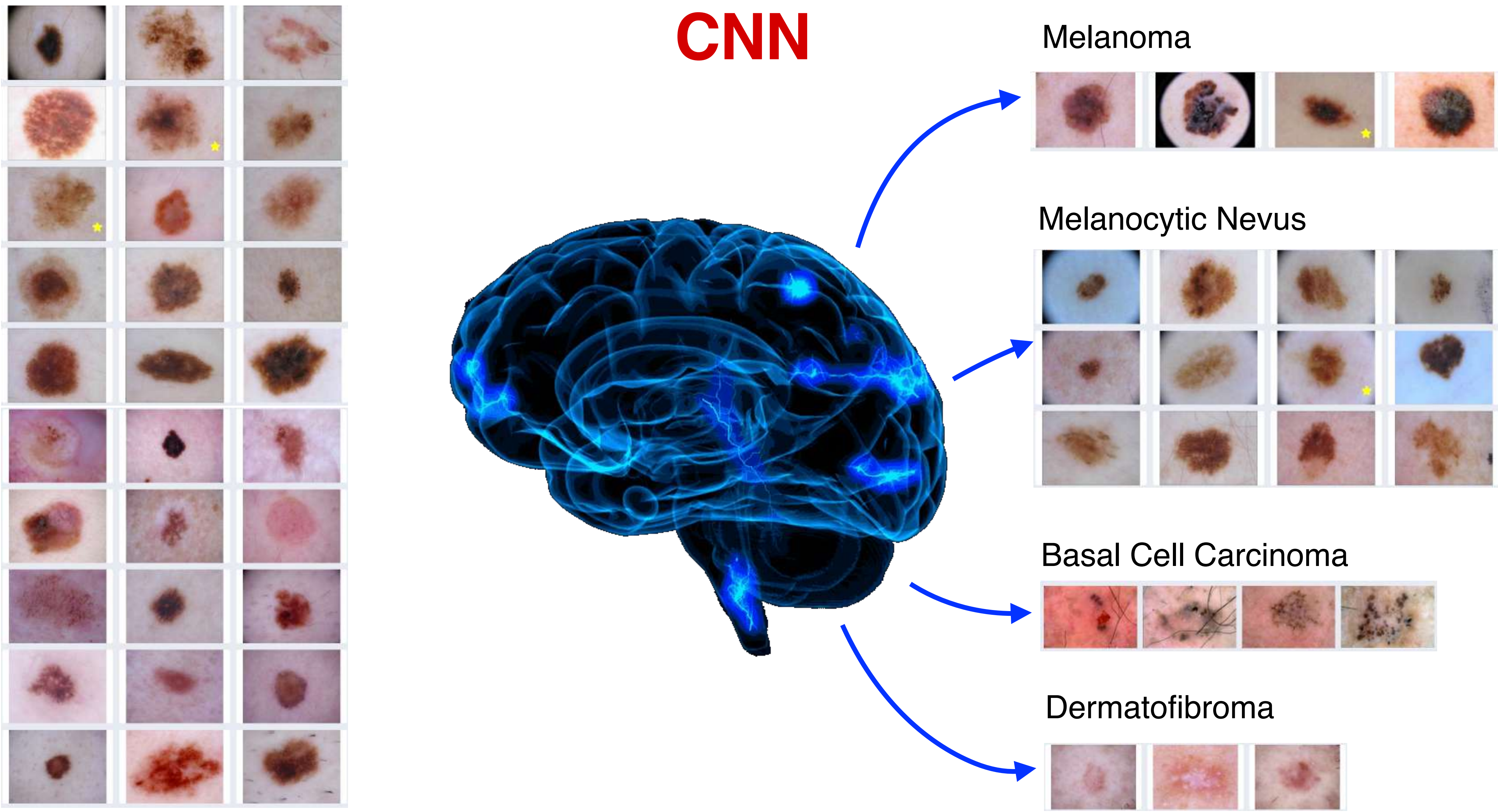
ISIC 2019

Home Background Data Submission Evaluation Leaderboard

Aggregate Metrics	Value
Balanced Multiclass Accuracy	0.593

The greatest diagnosis category score determines the category prediction for each image; the mean recall of this multiclass confusion matrix (i.e. the mean of the diagonal element-wise divided by the positive incidences) is the balanced multiclass accuracy





CNN
Melanoma
Melanocytic Nevus
Basal Cell Carcinoma
Dermatofibroma

THE QUESTION IS:

to formulate a diagnosis,
what a machine is looking at?

what a machine is looking at?

Artificial Intelligence

Neural Network

what a machine is looking at?

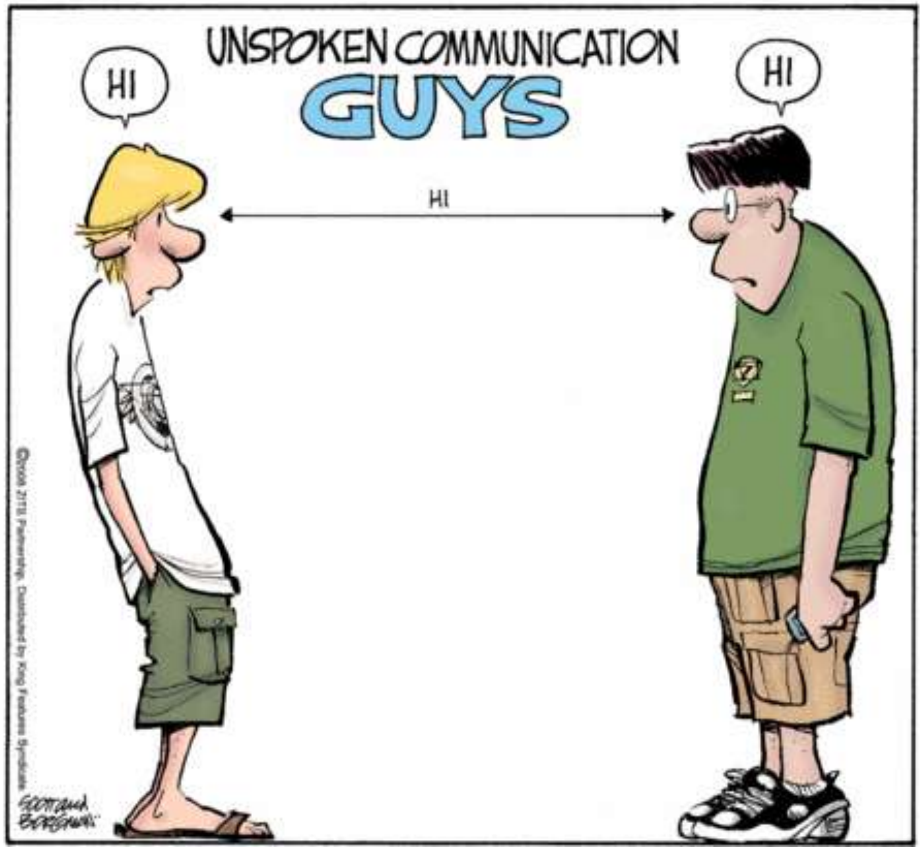
Artificial Intelligence

Neural Network

what a machine is looking at?

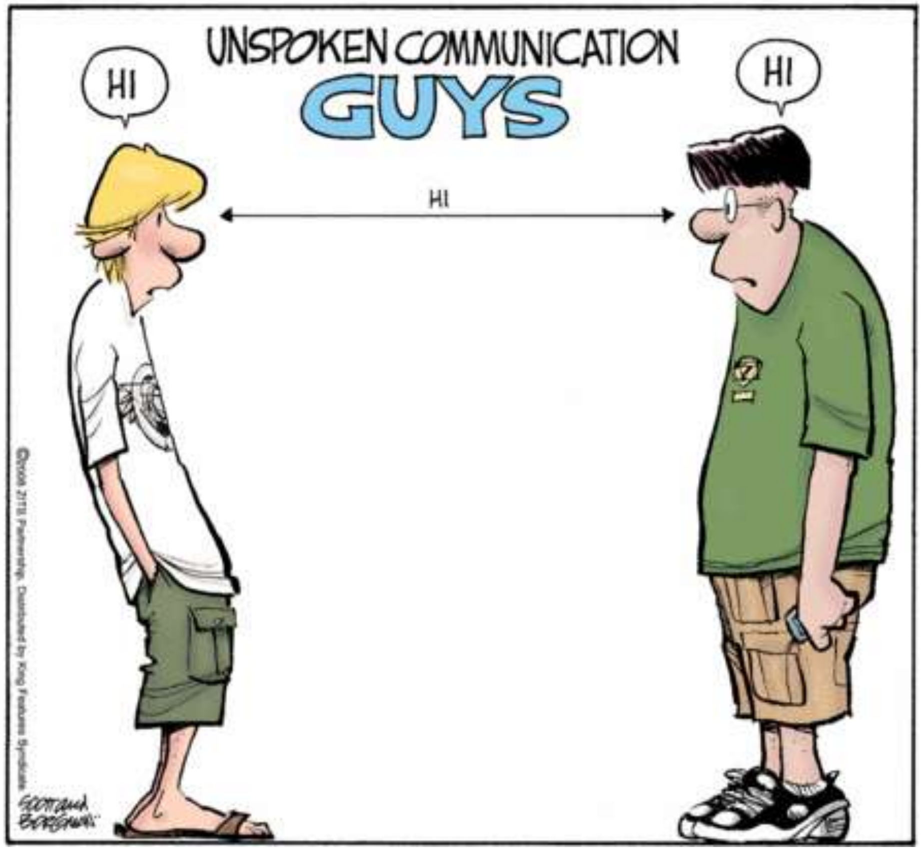
Artificial Intelligence

Neural Network

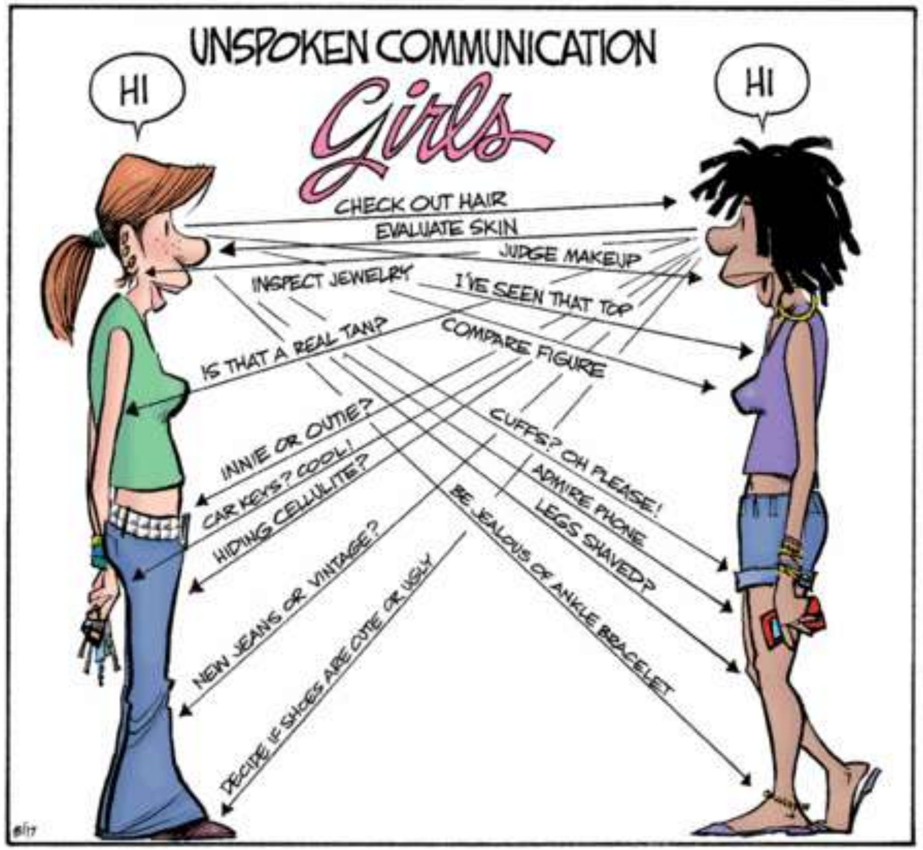


what a machine is looking at?

Artificial Intelligence

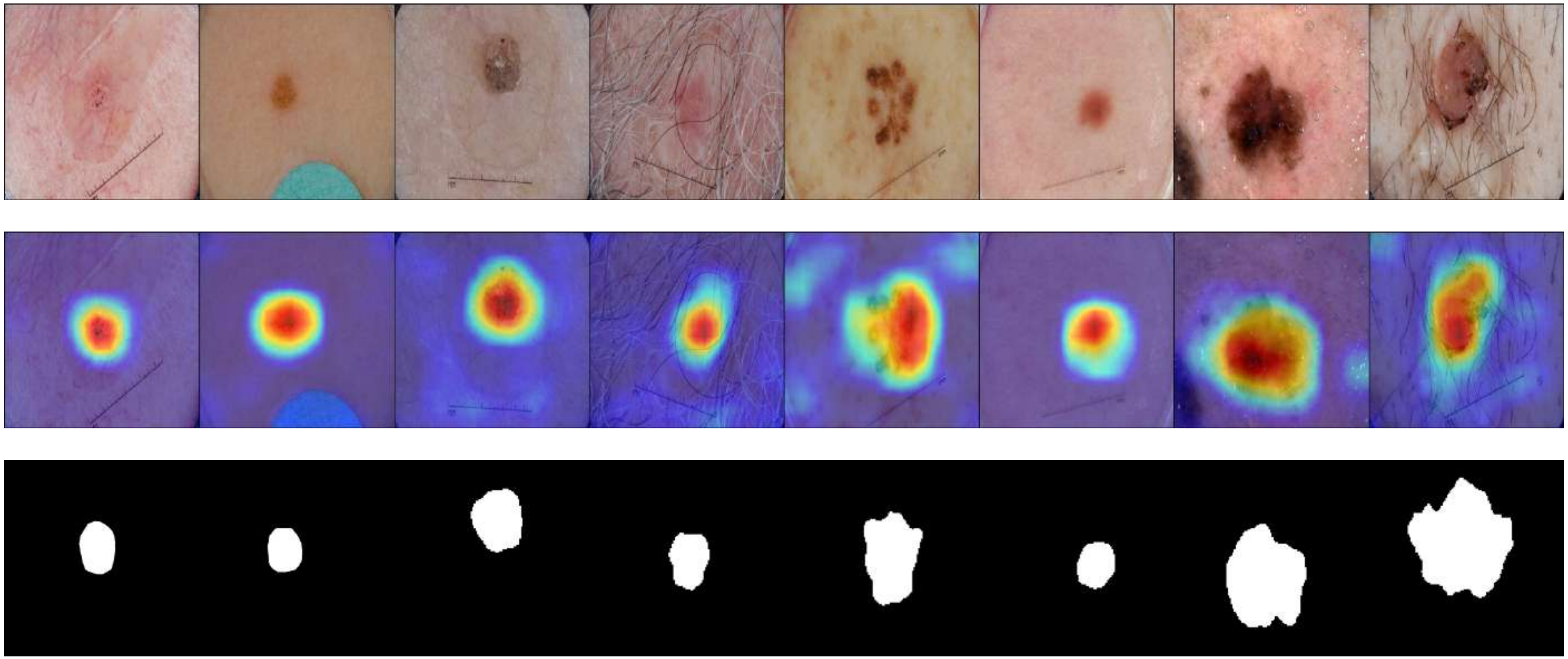


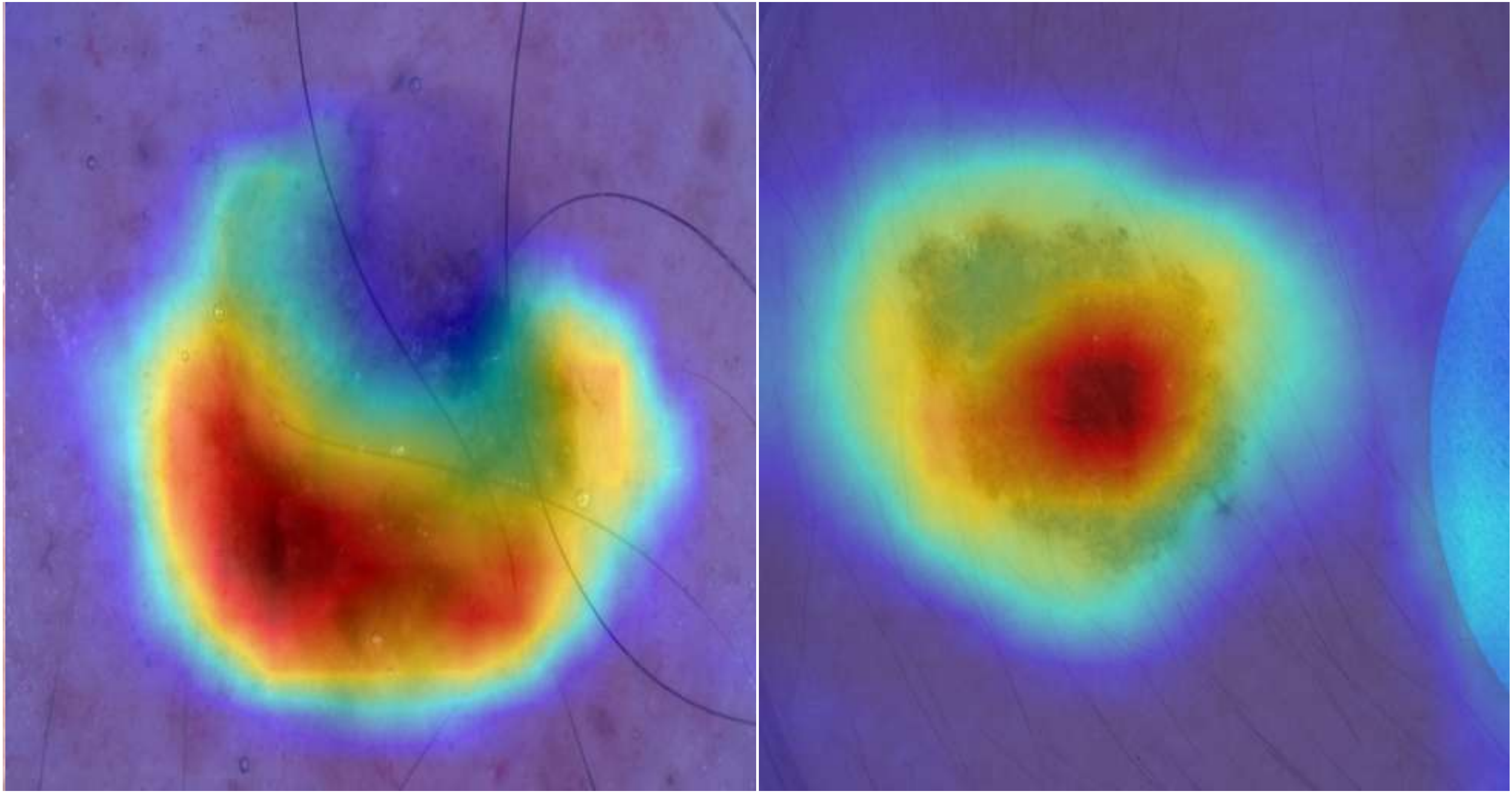
Neural Network

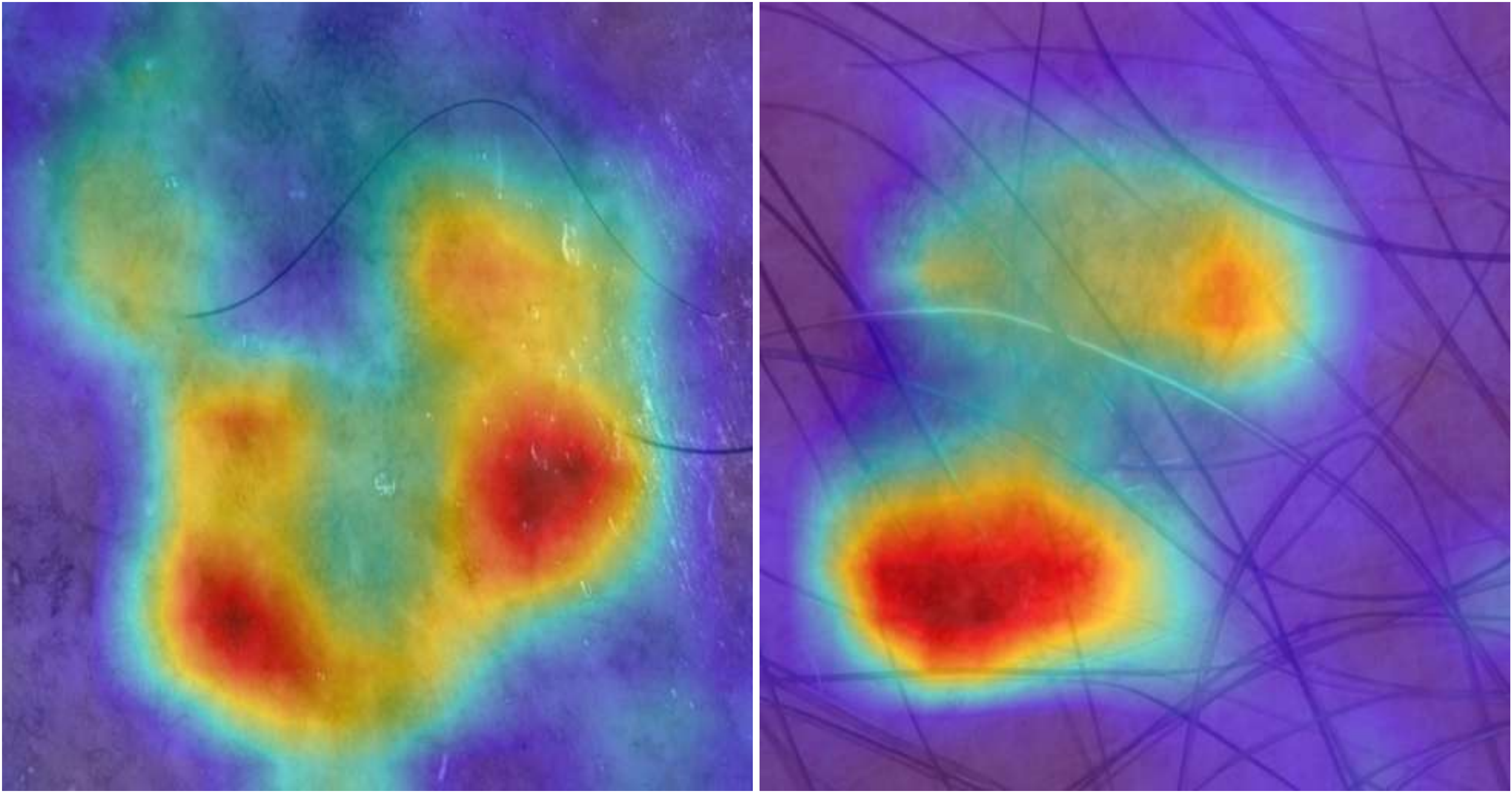


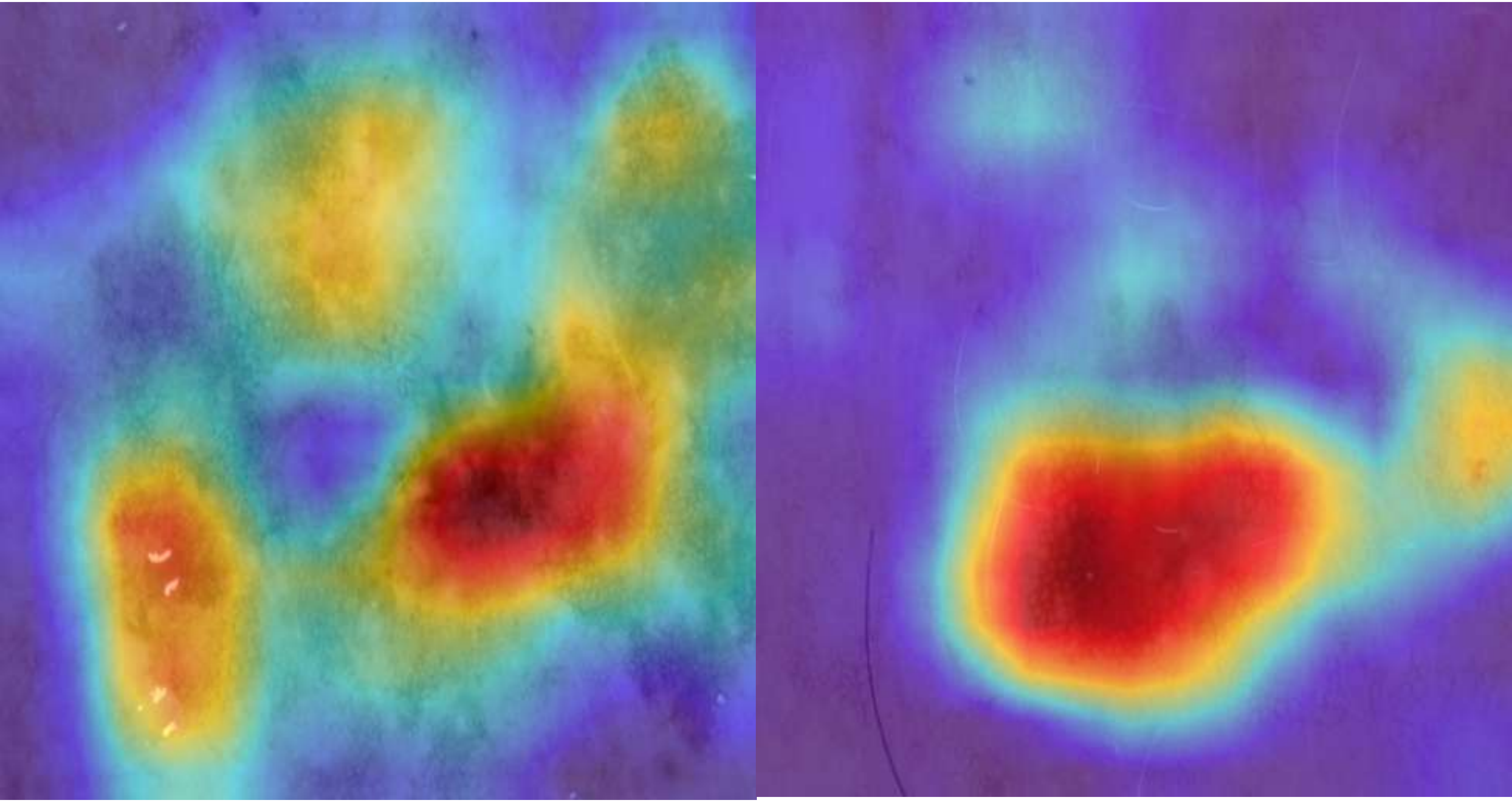
FEATURE EXTRACTION: HEATMAP

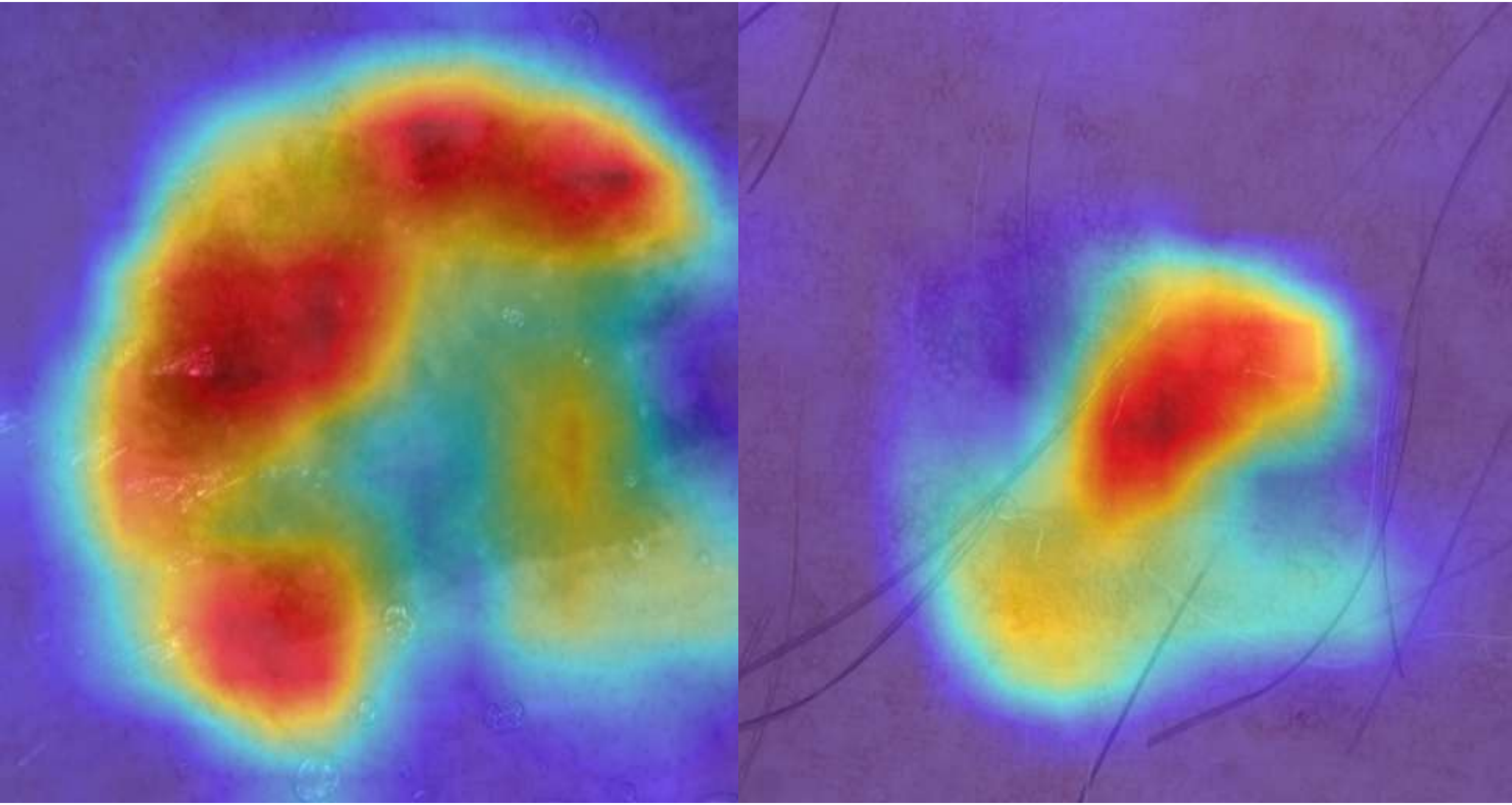
what a machine is looking at?



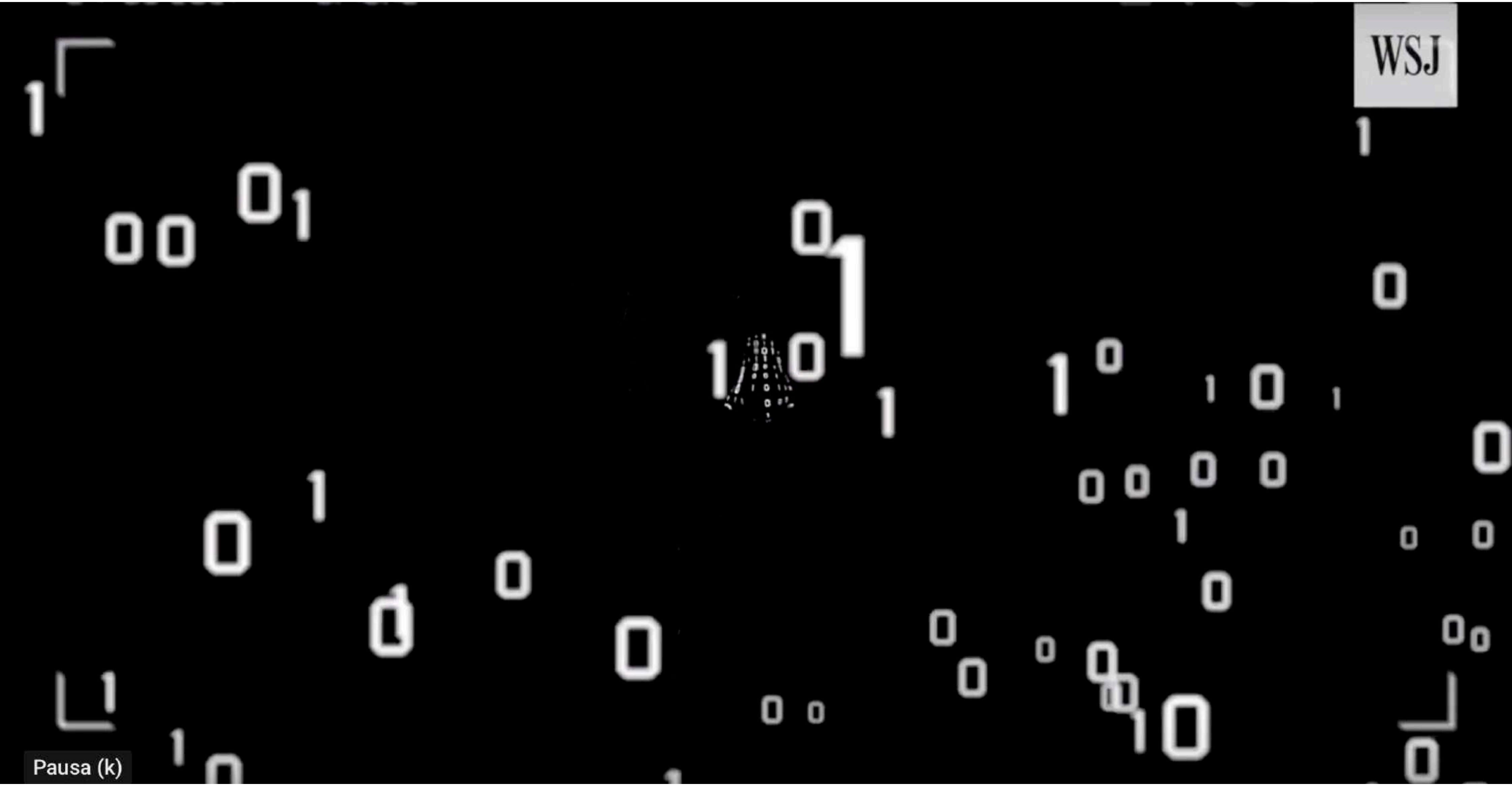




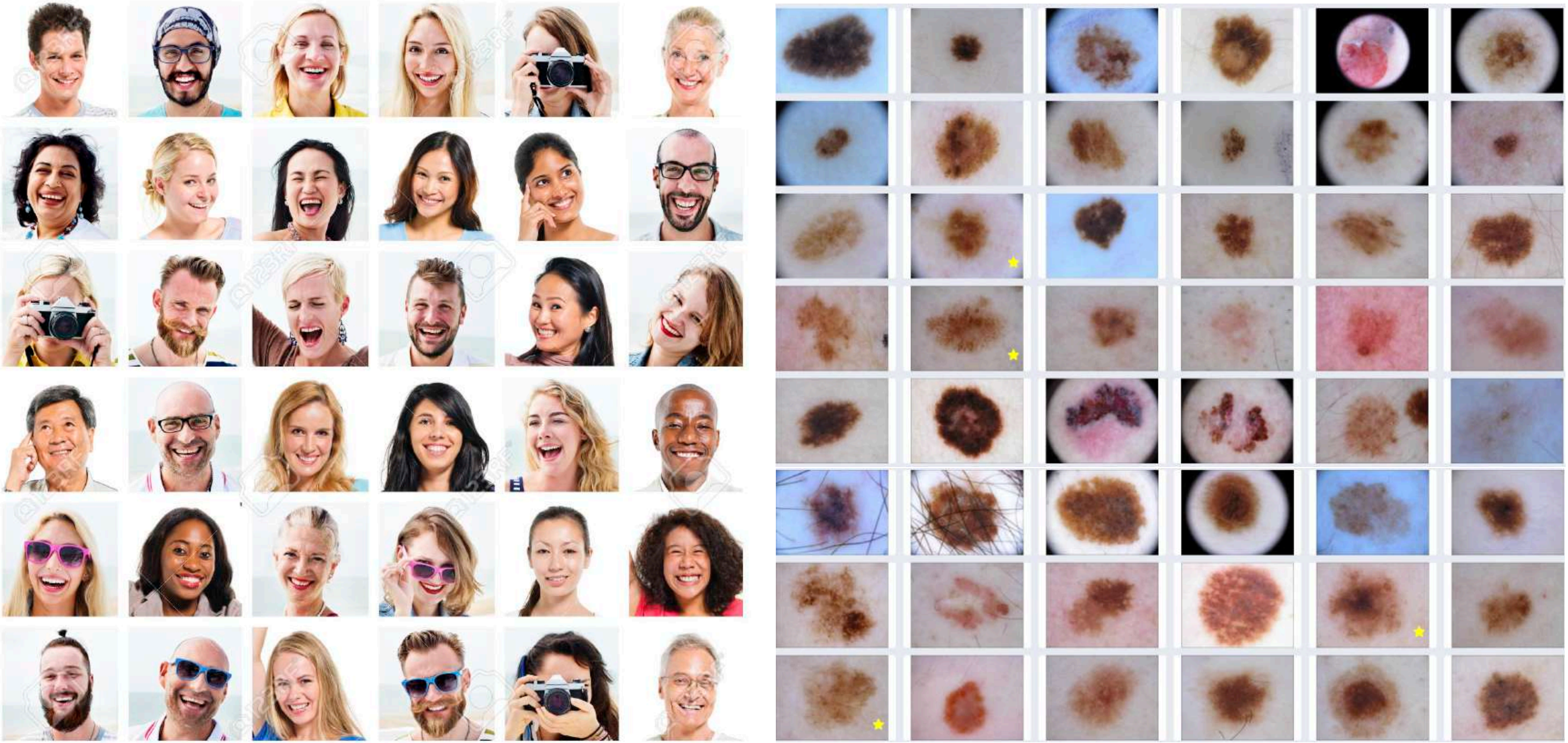




5. MAIN LIMITATION IN CNN

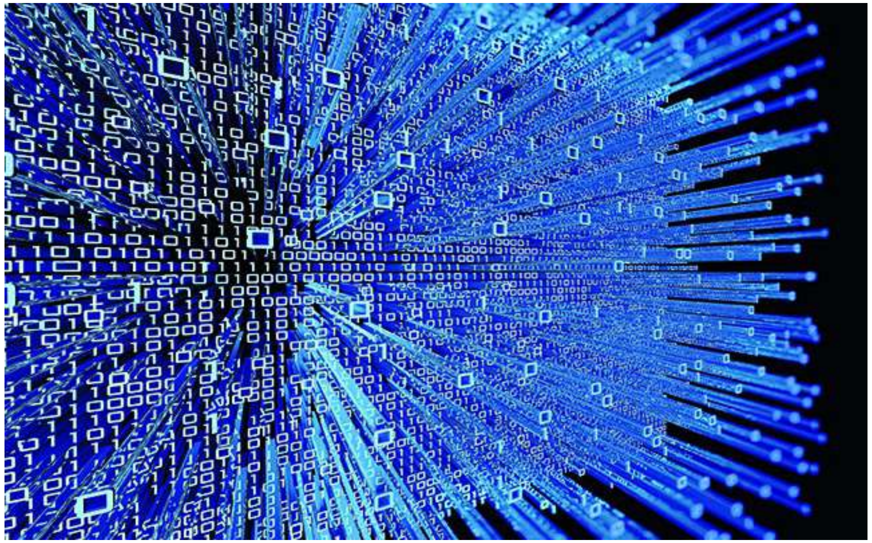


Today facial recognition is not “utopia” …



.. does melanoma have more faces than human faces?

WHICH IS THE LIMIT IN MELANOMA RECOGNITION BY CNN?



LACK OF AN ADEQUATE AMOUNT OF GOOD QUALITY DATA!!!!

6. FUTURE PERSPECTIVES

POSSIBLE FUTURE PROSPECTIVE

1. Open-access, standardized data (including also medical information)
2. Big data
3. Not only diagnosis (mobile app to digital monitoring skin lesions?)
4. Independent evaluation methodologies to accurately measure system efficacy

Thanks

To Prof Costantino Grana, Informatic Engineer - UNIMORE