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)
1956: Dartmouth conference: birth of A.I.

1970: Medical research 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 understanding of images drawn from a certain dataset.
HOW NEURAL NETWORKS RECOGNIZE A DOG IN A PHOTO

TRAINING
During the training phase, a neural network is fed thousands of labeled images of various animals, learning to classify them.

INPUT
An unlabeled image is shown to the pretrained network.

DEPLOY
FIRST LAYER
The neurons respond to different simple shapes, like edges.

HIGHER LAYER
Neurons respond to more complex structures.

TOP LAYER
Neurons respond to highly complex, abstract concepts that we would identify as different animals.

OUTPUT
The network predicts what the object most likely is, based on its training.

10% WOLF
90% DOG
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
LIGHTS...
AND
SHADOWS
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\textsuperscript{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 Esteva¹, Brett Kuprel¹, Roberto A. Novoa², Justin Ko², Susan M. Swetter²,⁴, Helen M. Blau⁵, Sebastian Thrun⁶

Skin lesion image
Deep convolutional neural network (Inception v3)

Training classes (757)
- Acral-lentiginous melanoma
- Amelanotic melanoma
- Lentigo melanoma
- ...

Inference classes (varies by task)
- 92% malignant melanocytic lesion
- 8% benign melanocytic lesion

and disease labels as inputs. We train a CNN using a dataset of 129,450 clinical images—two orders of magnitude larger than previous datasets¹²—consisting of 2,032 different diseases.

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

= average 63.7 images per diagnosis!!!
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]).
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 metrics **treated all diagnoses equally** (the algorithms did not consider that it is more detrimental to mistake a malignant for a benign lesion than vice versa)
- **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**)
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.
LIMITATIONS...

- **Etherogeneous studies** with high risk for bias
- **Experimental setting** (smartphone, 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
“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.”
Artificial intelligence and melanoma diagnosis: ignoring human nature may lead to false predictions

Aimilios Lallas¹, Giuseppe Argenziano³

“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

A new algorithm for border description of polarized light surface microscopic images of pigmented skin lesions.
Grana C¹, Pellacani G, Cucchiara R, Seldenari S.

Computer description of colours in dermoscopic melanocytic lesion images reproducing clinical assessment.
Seldenari S¹, Pellacani G, Grana C.

Automated description of colours in polarized-light surface microscopy images of melanocytic lesions.
Pellacani G¹, Grana C, Seldenari S.

ARTIFICIAL INTELLIGENCE APPROACH
AND CONTINUED WITH…


**Pigment distribution in melanocytic lesion images: a digital parameter to be employed for computer-aided diagnosis.**
Seidenari S², Pellacani G, Grana C.


**Asymmetry in dermoscopic melanocytic lesion images: a computer description based on colour distribution.**
Seidenari S², Pellacani G, Grana C.


**Algorithmic reproduction of asymmetry and border cut-off parameters according to the ABCD rule for dermoscopy.**
Pellacani G¹, Grana C, Seidenari S.

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

**Colour clusters for computer diagnosis of melanocytic lesions.**
Seidenari S¹, 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 DERMOSCOPY FEATURES AND TO REPRODUCE THE ABCD RULE
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

DCNN
CONVOLUTED NEURAL NETWORK APPROACH

FEATURE EXTRACTION: AUTO-ENCODER
<table>
<thead>
<tr>
<th>Rank</th>
<th>Team</th>
<th>Approach Name</th>
<th>Manuscript</th>
<th>Used External Data</th>
<th>Primary Metric Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DAISYLab Hamburg University of Technology/University Medical Center Hamburg-Eppendorf</td>
<td>Ensemble of Multi-Res EfficientNets + SEN154 2</td>
<td><img src="image" alt="Icon" /></td>
<td>Yes</td>
<td>0.636</td>
</tr>
<tr>
<td>2</td>
<td>DysionAI</td>
<td>Ensemble of EfficientNetB3-B4-Seresnext101</td>
<td><img src="image" alt="Icon" /></td>
<td>No</td>
<td>0.607</td>
</tr>
<tr>
<td>3</td>
<td>AlmageLab &amp; PRHLT Unimore &amp; UPV</td>
<td>ensemble, odd threshold 100%</td>
<td><img src="image" alt="Icon" /></td>
<td>No</td>
<td>0.593</td>
</tr>
<tr>
<td>4</td>
<td>DermaCode</td>
<td>13 models + hierarchical approach to select outliers</td>
<td><img src="image" alt="Icon" /></td>
<td>No</td>
<td>0.578</td>
</tr>
<tr>
<td>5</td>
<td>Nurithm Labs</td>
<td>Densenet-161 with heavy use of random crops</td>
<td><img src="image" alt="Icon" /></td>
<td>Yes</td>
<td>0.569</td>
</tr>
<tr>
<td>6</td>
<td>Torus Actions</td>
<td>Simple test approach</td>
<td><img src="image" alt="Icon" /></td>
<td>No</td>
<td>0.563</td>
</tr>
</tbody>
</table>
ISIC Challenge 2019

ISIC 2019

Aggregate Metrics

<table>
<thead>
<tr>
<th>Metric</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Balanced Multiclass Accuracy</td>
<td>0.593</td>
</tr>
</tbody>
</table>

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.

ROC

ensemble, ood threshold 100%
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