

Man against machine: AI is better than dermatologists at diagnosing skin cancer

28 May 2018



Melanoma in skin biopsy with H&E stain — this case may represent superficial spreading melanoma. Credit: Wikipedia/CC BY-SA 3.0

Researchers have shown for the first time that a form of artificial intelligence or machine learning known as a deep learning convolutional neural network (CNN) is better than experienced dermatologists at detecting skin cancer.

In a study published in the leading cancer journal *Annals of Oncology* today, researchers in Germany, the USA and France trained a CNN to identify [skin cancer](#) by showing it more than 100,000 images of malignant melanomas (the most lethal form of skin cancer), as well as benign moles (or nevi). They compared its performance with that of 58 international dermatologists and found that the CNN missed fewer melanomas and misdiagnosed benign moles less often as malignant than the group of dermatologists.

A CNN is an artificial neural network inspired by the biological processes at work when nerve cells (neurons) in the brain are connected to each other

and respond to what the eye sees. The CNN is capable of learning fast from images that it "sees" and teaching itself from what it has learned to improve its performance (a process known as machine learning).

The first author of the study, Professor Holger Haenssle, senior managing physician at the Department of Dermatology, University of Heidelberg, Germany, explained: "The CNN works like the brain of a child. To train it, we showed the CNN more than 100,000 images of malignant and benign skin cancers and moles and indicated the diagnosis for each image. Only dermoscopic images were used, that is lesions that were imaged at a 10-fold magnification. With each training image, the CNN improved its ability to differentiate between benign and malignant lesions.

"After finishing the training, we created two test sets of images from the Heidelberg library that had never been used for training and therefore were unknown to the CNN. One set of 300 images was built to solely test the performance of the CNN. Before doing so, 100 of the most difficult lesions were selected to test real dermatologists in comparison to the results of the CNN."

Dermatologists from around the world were invited to take part, and 58 from 17 countries around the world agreed. Of these, 17 (29%) indicated they had less than two years' experience in dermoscopy, 11 (19%) said they were skilled with between two to five years' experience, and 30 (52%) were expert with more than five years' experience.

The dermatologists were asked to first make a diagnosis of [malignant melanoma](#) or benign mole just from the dermoscopic images (level I) and make a decision about how to manage the condition (surgery, short-term follow-up, or no action needed). Then, four weeks later they were given clinical information about the patient (including age, sex and position of the lesion) and

close-up images of the same 100 cases (level II) and asked for diagnoses and management decisions again.

In level I, the dermatologists accurately detected an average of 86.6% of melanomas, and correctly identified an average of 71.3% of lesions that were not malignant. However, when the CNN was tuned to the same level as the physicians to correctly identify benign moles (71.3%), the CNN detected 95% of melanomas. At level II, the dermatologists improved their performance, accurately diagnosing 88.9% of malignant melanomas and 75.7% that were not cancer.

"The CNN missed fewer melanomas, meaning it had a higher sensitivity than the dermatologists, and it misdiagnosed fewer benign moles as malignant [melanoma](#), which means it had a higher specificity; this would result in less unnecessary surgery," said Professor Haenssle.

"When dermatologists received more clinical information and images at level II, their diagnostic performance improved. However, the CNN, which was still working solely from the dermoscopic images with no additional [clinical information](#), continued to out-perform the physicians' diagnostic abilities."

The expert dermatologists performed better at level I than the less experienced dermatologists and were better at detecting malignant melanomas. However, their average ability to make the correct diagnosis was still worse than the CNN at both levels.

"These findings show that deep learning convolutional neural networks are capable of out-performing dermatologists, including extensively trained experts, in the task of detecting melanomas," he said.

The incidence of malignant melanoma is increasing, with an estimated 232,000 new cases worldwide and around 55,500 deaths from the disease each year. It can be cured if detected early, but many cases are only diagnosed when the cancer is more advanced and harder to treat.

Professor Haenssle said: "I have been involved in research projects that aim at improving the early detection of melanoma in its curable stages for almost 20 years. My group and I are focusing on non-invasive technologies that may help physicians not to miss melanomas, for instance, while performing skin cancer screenings. When I came across recent reports on deep-learning algorithms that outperform human experts in specific tasks, I immediately knew that we had to explore these artificial intelligence algorithms for diagnosing melanoma."

The researchers do not envisage that the CNN would take over from dermatologists in diagnosing skin cancers, but that it could be used as an additional aid.

"This CNN may serve physicians involved in skin cancer screening as an aid in their decision whether to biopsy a lesion or not. Most dermatologists already use digital dermoscopy systems to image and store lesions for documentation and follow-up. The CNN can then easily and rapidly evaluate the stored image for an 'expert opinion' on the probability of melanoma. We are currently planning prospective studies to assess the real-life impact of the CNN for physicians and patients."

The study has some limitations, which include the fact that the dermatologists were in an artificial setting where they knew they were not making "life or death" decisions; the test sets did not include the full range of skin lesions; there were fewer validated images from non-Caucasian skin types and genetic backgrounds; and the fact that doctors may not always follow the recommendation of a CNN they don't trust.

In an accompanying editorial Dr. Victoria Mar (Monash University, Melbourne, Australia) and Professor H. Peter Soyer (The University of Queensland, Brisbane, Australia) write: "Currently, diagnostic accuracy for melanoma is dependent on the experience and training of the treating doctor. Haenssle et al...have shown that a computer algorithm using convolutional neural networks outperformed the majority of 58 dermatologists testedThis shows that [artificial intelligence](#) (AI)

promises a more standardised level of diagnostic accuracy, such that all people, regardless of where they live or which doctor they see, will be able to access reliable diagnostic assessment."

They highlight a number of issues that would need to be addressed before AI could become standard in clinics, including the difficulty of imaging some melanomas on sites such as the fingers, toes and scalp, and how to train AI sufficiently to recognise atypical melanomas and ones that patients are unaware of.

They conclude: "Currently, there is no substitute for a thorough clinical examination. However, 2-D and 3-D total body photography is able to capture about 90 to 95% of the skin surface and given exponential development of imaging technology we envisage that sooner than later, automated diagnosis will change the diagnostic paradigm in dermatology. Still, there is much more work to be done to implement this exciting technology safely into routine clinical care."

More information: "Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists", by H.A. Haenssle et al. *Annals of Oncology*. DOI: [10.1093/annonc/mdy166](https://doi.org/10.1093/annonc/mdy166)

Provided by European Society for Medical Oncology

APA citation: Man against machine: AI is better than dermatologists at diagnosing skin cancer (2018, May 28) retrieved 19 June 2019 from <https://medicalxpress.com/news/2018-05-machine-ai-dermatologists-skin-cancer.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.