ZZ/Alamy Stock Photo

▶ EARLY DETECTION

The algorithm will see you now

> **the performance of the algorithm in correctly classifying these skin lesions was found to match that of 21 dermatologists**

Early detection of skin cancer is crucial, as survival rates for melanoma drop substantially if it is detected at a late stage. Diagnosis of skin cancer primarily relies upon visual inspection from a trained dermatologist which is then confirmed by biopsy and histological examination. Assessment of a photographic image of a skin lesion by dermatologists has been proved to substitute for a physical examination in providing an initial diagnosis. Now, a new study by Esteva *et al.* has taken this one step further by training a computer to develop artificial intelligence in image recognition for automation of skin cancer diagnosis.

Scientists had tried to develop computer-aided classification systems previously but encountered problems with the amount of variability in both the appearance of skin lesions and photographic images. To overcome these issues, Esteva and colleagues used a data-driven algorithmic approach known as deep machine learning or deep convolutional neural networks (CNNs). To train the algorithm, the authors used a data set of 129,450 clinical digital images of skin lesions — 100 times larger than previously reported data sets — comprising 2,032 different diseases, each labelled with the name of the condition by a dermatologist. These training images, which effectively functioned as reference diagnoses, were collated from 18 different online resources curated by clinicians as well as clinical data from Stanford University Medical Center.

To test the efficiency of the algorithm for accurately diagnosing skin cancers, the authors presented the algorithm and two dermatologists with test images representing benign lesions, malignant lesions and non-neoplastic lesions. The algorithm identified these distinct skin lesions with an accuracy of approximately 72% compared with an approximate 66% accuracy score for the dermatologists. Although this test demonstrated that the algorithm was learning, Esteva *et al.* reasoned that it was not sufficiently robust, as the images were labelled by dermatologists and not necessarily confirmed by biopsy. To address this challenge, the authors tested the ability of the algorithm again, but this time using previously unseen biopsy-verified images to distinguish between benign and malignant from two types of skin lesion with different cellular origins — epidermal (benign seborrhoeic keratosis versus basal and squamous cell carcinomas) or melanocytic (benign nevus versus malignant melanoma). In this case, the performance of the algorithm in correctly classifying these skin lesions was found to match that of 21 dermatologists.

Although at present this study is only proof of principle that automated skin cancer diagnosis is possible, and has yet to be validated in the clinical setting, the implications of such technology for clinical decision-making are far-reaching. The authors envisage that mobile devices could potentially be fitted with such an algorithm in the form of a smartphone app, which would make diagnostic health care low-cost, easy and accessible to many more people.

Anna Dart

ORIGINAL ARTICLE Esteva, A. *et al.* Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542**, 115–118 (2017)