ORIGINAL ARTICLE

Evaluation of Melanoma Thickness with Clinical Close-up and Dermoscopic Images Using a Convolutional Neural Network

Martin GILLSTEDT[1,2], Ludwig MANNIUS[1], John PAOLI[1,2], Johan DAHLÉN GYLLENCREUTZ[1], Julia FOUGELBERG[1,2], Eva JOHANSSON BACKMAN[1,2], Jenna PAKKA[1,2], Oscar ZAAR[1,2] and Sam POLESIE[1,2]

[1]Department of Dermatology and Venereology, Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg and [2]Region Västra Götaland, Sahlgrenska University Hospital, Department of Dermatology and Venereology, Gothenburg, Sweden

Convolutional neural networks (CNNs) have shown promise in discriminating between invasive and *in situ* melanomas. The aim of this study was to analyse how a CNN model, integrating both clinical close-up and dermoscopic images, performed compared with 6 independent dermatologists. The secondary aim was to address which clinical and dermoscopic features dermatologists found to be suggestive of invasive and *in situ* melanomas, respectively. A retrospective investigation was conducted including 1,578 cases of paired images of invasive ($n = 728$, 46.1%) and *in situ* melanomas ($n = 850$, 53.9%). All images were obtained from the Department of Dermatology and Venereology at Sahlgrenska University Hospital and were randomized to a training set ($n = 1,078$), a validation set ($n = 200$) and a test set ($n = 300$). The area under the receiver operating characteristics curve (AUC) among the dermatologists ranged from 0.75 (95% confidence interval 0.70–0.81) to 0.80 (95% confidence interval 0.75–0.85). The combined dermatologists' AUC was 0.80 (95% confidence interval 0.77–0.86), which was significantly higher than the CNN model (0.73, 95% confidence interval 0.67–0.78, $p = 0.001$). Three of the dermatologists significantly outperformed the CNN. Shiny white lines, atypical blue-white structures and polymorphous vessels displayed a moderate interobserver agreement, and these features also correlated with invasive melanoma. Prospective trials are needed to address the clinical usefulness of CNN models in this setting.

Key words: artificial intelligence; clinical decision-making; melanoma; neural network, computer; supervised machine learning.

Accepted Sep 29, 2022; Epub ahead of print Sep 29, 2022

Acta Derm Venereol 2022; 102: adv00790.

DOI: 10.2340/actadv.v102.2681

Corr: Sam Polesie, Department of Dermatology and Venereology, Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Gröna stråket 16, SE-413 45 Gothenburg, Sweden. E-mail: sam.polesie@vgregion.se

SIGNIFICANCE

In this investigation, a machine learning algorithm, using both close-up and dermoscopic images of melanomas, was developed to assess melanoma thickness. The algorithm was trained and validated on 1,278 images and ultimately tested on 300 images. Six dermatologists were invited to independently evaluate the same test images. The dermatologists collectively achieved a significantly better accuracy compared with the machine learning algorithm in correctly estimating melanoma thickness. Future clinical trials are necessary to determine whether the use of these algorithms can enhance decision-making in assessment of melanoma thickness.

Machine learning (ML) algorithms, including convolutional neural networks (CNNs), have recently made significant advances in dermatology research. The ultimate aim of these algorithms is to assist physicians, enhance decision-making and improve patient care (i.e. augmented intelligence) (1). While attitudes towards ML tools are generally positive (2, 3), a major challenge for broad implementation into clinical practice is to convert algorithm predictions into effective and safe software (4). Moreover, prospective clinical trials are needed to clarify whether to put algorithmic interpretation before or after physician evaluation (5), and to identify specific clinical problems for which use of these tools is suitable. Before broad implementation, there is a need for prospective clinical trials that can demonstrate usefulness in prespecified and specific domains. These should identify situations that are often faced in clinical practice that might impact patient care and are frequently challenging for the physician.

One illustrative example that can be employed as a use case is to distinguish between invasive and *in situ* melanomas in a preoperative setting in order to choose optimized surgical margins (6). Although all suspected melanomas require an excision, a single surgical procedure for melanoma *in situ* rather than, potentially, 2 procedures can result in considerable economic savings and convenience for patients.

While dermoscopy might facilitate this binary classification problem, the assessment of melanoma thickness is challenging even for experienced dermatologists (7), and only a few dermoscopic features have both an acceptable interobserver agreement and discriminatory power when differentiating between invasive and *in situ* melanoma (8). In 2 previous studies we developed ML algorithms for this setting (9, 10). In the first publication, only dermoscopic images were used, whereas the second included only clinical close-up images. To better

imitate the clinical setting, the primary aim of the current study was to develop a CNN that includes both the clinical close-up and dermoscopic images and to compare its performance with that of 6 dermatologists. The secondary aim was to identify which predefined clinical and dermoscopic features were useful for dermatologists when clinically discriminating between invasive and *in situ* melanomas.

MATERIALS AND METHODS

A retrospective single-centre investigation was conducted. The data-set was comprised of all melanomas diagnosed at the Department of Pathology at Sahlgrenska University Hospital in the time-period January 2015 to July 2021 with available images obtained from the Department of Dermatology and Venereology at the same hospital. All cases displaying a recurrent lesion (e.g. growth in a scar) or previously punch-biopsied melanomas were excluded. Overall, 1,578 lesions had both available clinical close-up and dermoscopic images and were included in this study. One image pair (1 dermoscopic and 1 clinical close-up image) was used for each lesion. All pairs were randomized into a training set (n=1,078), a validation set (n=200) and a test set (n=300). To avoid any recall bias from prior investigations, none of the cases included in the test set had been used previously.

Each dermatologist was given a list of predefined criteria relating to clinical and dermoscopic features for each lesion (Table SI). Among the 6 included dermatologists, the length of dermoscopy experience ranged from 4.5 to 19 years. The dermatologists were required to provide their answers in the time-period 20 October 2021 to 28 November 2021 (40 days) and were asked to use a single computer set-up for the evaluation.

For all included image pairs, the dermatologists were asked to report a level of confidence ranging from 1 (very uncertain) to 5 (very certain). From these confidence levels a certainty score ranging from 0 to 1 was defined as described in a previous investigation (10), where 0 represents complete certainty for *in situ* and 1 for invasive melanoma. The complete test set as presented to the dermatologists is shown in Appendix S1). The study was reviewed and approved by the Regional Ethics Review Board in Gothenburg (approval number 283–18).

The same image resolution (i.e. 600×1,200) was used for the dermatologists and the CNN. Image imperfections, such as covering hair and skin markers, were allowed. All clinical images were cropped to maintain patient anonymity. The original image resolution (height × width) for the cropped clinical close-up images ranged from 78 × 73 to 1,985 × 1,985 pixels, whereas the dermoscopy images ranged from 1,200 × 1,600 to 3,318 × 4,416 pixels. The anatomical location of the lesions was not provided, and lesion diameter was not included.

Hardware and software

The Keras library (version 2.3.1) using the Tensorflow backend (version 1.14.0) was used running on Python version 3.6.9. Model construction was performed using R version 3.5.3 (The R Foundation for Statistical Computing, Vienna, Austria) and the R-package Keras was used to call Python and its above libraries. XnView version 2.20 was used to scale and crop all images to quadratic shape with a resolution of 600 × 600 (preserving the aspect ratio). The computer running the training used the GPU version of the Keras/Tensorflow routines. The graphics card used was an Nvidia Geforce GTX 1070 with 8 GB GPU memory using CUDA version

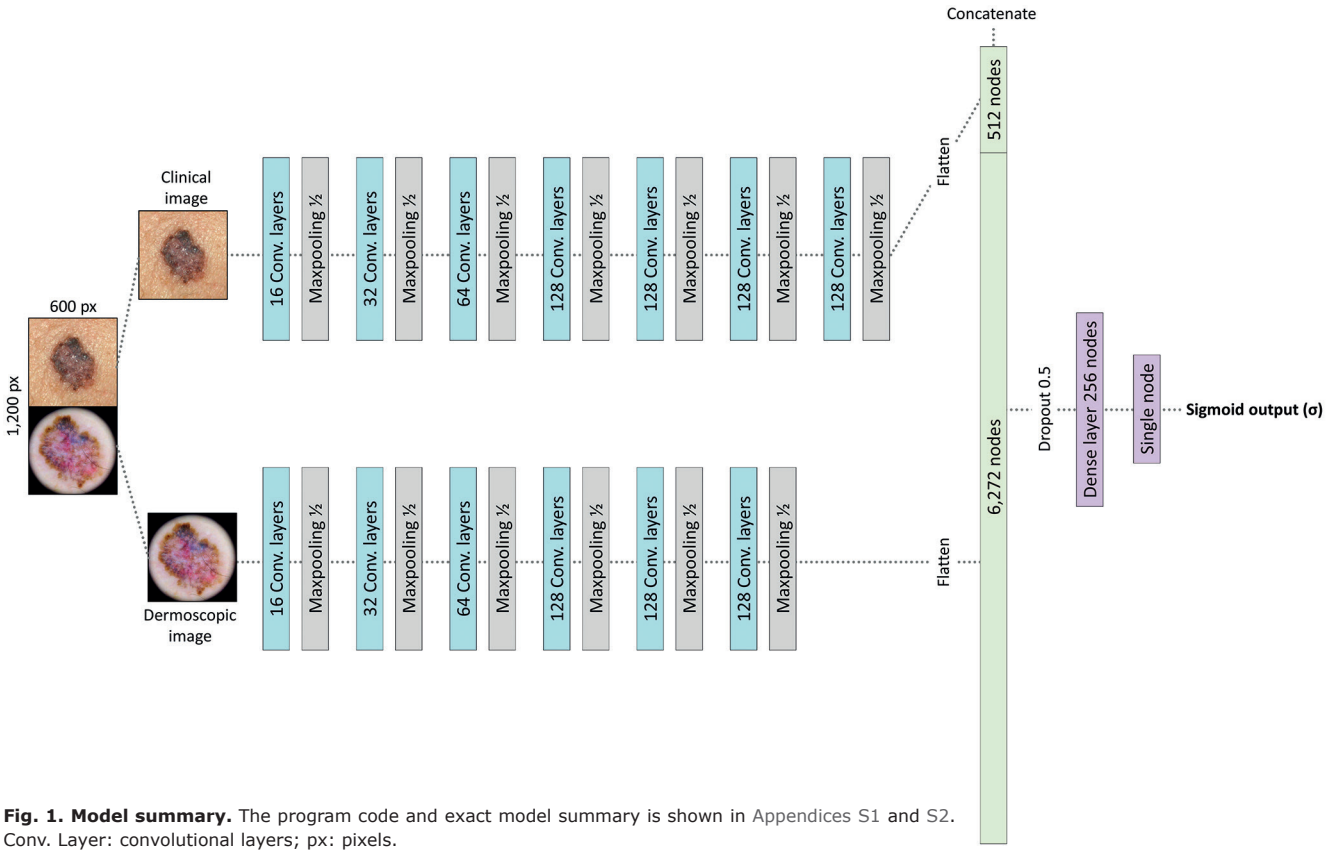

Fig. 1. Model summary. The program code and exact model summary is shown in Appendices S1 and S2. Conv. Layer: convolutional layers; px: pixels.

10.0 and cudnn version 7.6.3.30. The processor used was an Intel Core i5-2400 @ 3.10 GHz and the amount of RAM was 24 GB.

Model training

Several *de novo* CNN models (i.e. models with no pre-trained parameters) with varying architectures were employed for the training and validation phase. The final CNN architecture consisted of 6 convolutional layers for the dermoscopy image and 7 convolutional layers for the clinical close-up image (**Fig. 1**) (Appendices S2 and S3). The final model using both dermoscopic and clinical images was trained for 17 epochs, which took 15 min.

Statistical analysis

All data were analysed using R version 3.5.3 (https://www.r-project.org/). DeLong's test for 2 correlated receiver operating characteristics (ROC) curves was used to compare the performance of dermatologists and the CNN and to compare performance between different CNNs. The CNN output ranged from 0 to 1, where higher values indicated invasive melanoma and lower values indicated *in situ* melanoma. The interobserver agreement between the dermatologists was calculated using Fleiss' kappa (κ) (11). The agreement (κ-value) was interpreted as poor (<0), slight (0–0.2), fair (>0.2–0.4), moderate (>0.4–0.6), substantial (>0.6–0.8) or almost perfect (>0.8) (12). Univariate logistic regression and odds ratios (ORs) were used to assess whether the current predefined clinical (raised: yes/no) and dermoscopic features (i.e. only brown and/or black dermoscopic structures, >50% regression, atypical blue-white structures, shiny white lines, and polymorphous vessels) correlated with invasive or *in situ* melanomas as well as melanomas less than or greater than 1.0 mm in thickness, respectively. For these analyses, each lesion was given 6 scores (1 score per feature) pertaining to the proportion of dermatologists that included that specific feature in their assessment (i.e. ranging from 0 to 1). All tests were 2-sided and $p<0.05$ was considered statistically significant.

RESULTS

In total, 728 (46.1%) invasive and 850 (53.9%) *in situ* melanomas were included. The median age (interquartile range; IQR) at diagnosis was 67 years (54–76 years) and 839 lesions (53.2%) were found in males. Most lesions were located on the trunk ($n=757$, 48.0%), upper ($n=358$, 22.7%) and lower extremities ($n=320$, 20.3%). The test set ($n=300$) comprised 139 (46.3%) invasive and 161 (53.7%) *in situ* melanomas (**Table I**).

Table I. Distribution of melanomas included in the test set

	Frequency (%)
Melanoma *in situ*	161 (53.7)
Invasive melanoma	139 (46.3)
≤1.0 mm	99 (33)
Ulcerated	2
Not ulcerated	97
>1.0 mm	40 (13.3)
Ulcerated	14
Not ulcerated	26

Convolutional neural networks vs dermatologists

An ROC curve for each dermatologist was defined using the certainty scores and their AUCs ranged from 0.75 (95% CI 0.70–0.81) to 0.80 (95% CI 0.75–0.85). A combined certainty score for the dermatologists was defined by taking the mean of the 6 dermatologists' certainty scores. This combined score yielded a higher AUC than using only the mean of the dermatologists' dichotomous scores, where 0=*in situ* and 1=invasive melanoma (Fig. S1). The combined dermatologists' AUC was 0.80 (95% CI 0.77–0.86), which was significantly higher than the merged CNN model 0.73 (95% CI 0.67–0.78, $p=0.001$) (**Fig. 2**A). For the CNN, the AUC for correctly classifying invasive melanomas >1.0 mm ($n=40$) as invasive melanoma was 0.93 (95% CI 0.87–0.99). The corresponding AUC for invasive melanomas ≤1.0 mm ($n=99$) was 0.64 (95% CI 0.57–0.72). The AUC for the dermatologists were 0.97 (95% CI 0.93–1.0, $p=0.007$) for invasive melanomas >1.0 mm and 0.74 (95% CI 0.67–0.80, $p=0.004$) for invasive melanomas ≤1.0 mm (Figs. 2B and C). Three of the dermatologists significantly outperformed the CNN (**Table II**).

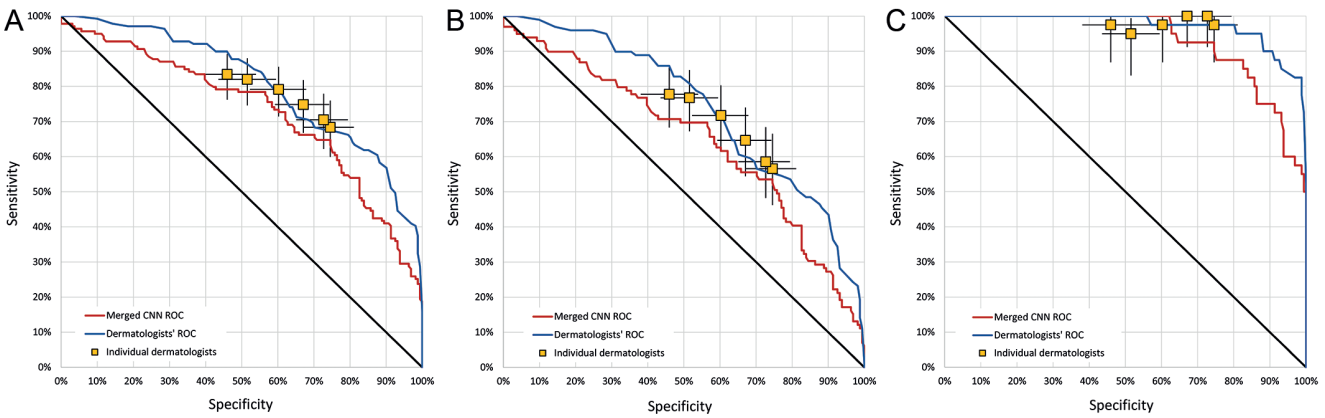

Fig. 2. Area under the receiver operating characteristics curves (AUC). (A) AUC for the merged convolutional neural network (CNN) model and the 6 dermatologists when including all lesions ($n=300$). (B) AUC for the merged CNN model and the 6 dermatologists when including only *in situ* melanomas ($n=161$) and thin invasive melanomas (i.e. with a Breslow thickness ≤1.0 mm) ($n=99$). (C) AUC for the merged CNN model and the 6 dermatologists when including only thick invasive melanomas (i.e. Breslow thickness >1.0 mm) ($n=40$). ROC: receiver operating characteristic.

Table II. Area under the receiver operating characteristic curve (AUC) for dermatologists' combined and convolutional neural network (CNN) with regards to discriminating between invasive and in situ melanomas

		Lower	Upper	
AUC		95 % CI		*p*-value
Dermatologists combined vs CNN	0.80	0.75	0.86	0.001
	0.73	0.67	0.78	
Reader 1 vs CNN	0.79	0.74	0.84	0.017
Reader 2 vs CNN	0.76	0.71	0.82	0.18
Reader 3 vs CNN	0.80	0.75	0.85	0.008
Reader 4 vs CNN	0.80	0.74	0.85	0.010
Reader 5 vs CNN	0.77	0.71	0.82	0.15
Reader 6 vs CNN	0.75	0.70	0.81	0.32

95% CI: 95% confidence interval.

Merged convolutional neural networks vs convolutional neural networks based only on dermoscopic images and those based only on clinical close-up images

The AUC for the merged CNN performed on par with the corresponding CNN model that only included dermoscopic images 0.72 (95% CI 0.66–0.78%, $p=0.20$), but outperformed the CNN model that included only clinical close-up images 0.68 (95% CI 0.62–0.74, $p=0.036$) (**Fig. 3**). The same results were obtained when comparing thin invasive melanomas (i.e. with a Breslow thickness ≤1.0 mm ($n=99$)) with melanoma *in situ* ($n=161$), and when comparing thick invasive melanomas (i.e. invasive melanomas with a Breslow thickness >1.0 mm ($n=40$)) with melanoma *in situ* ($n=161$) (Fig. S2A and B).

Dermatologists' assessment

Shiny white lines, atypical blue-white structures, and polymorphous vessels all exceeded 90% specificity for invasive melanomas. Lesions determined to be raised in the clinical images also displayed a high level of specificity (96.7%, 95% CI 95.3–97.8%) (**Fig. 4**) (Table SII). The overall interobserver agreement of classifying the melanomas as invasive or *in situ* among the 6 dermatologists ranged from moderate to substantial (κ=0.58, 95% CI 0.55–0.61). The agreement regarding whether a lesion was raised or flat was also in the same range (κ=0.60, 95% CI 0.57–0.62). Shiny white lines (κ=0.58, 95% CI 0.55–0.60), atypical blue-white structures (κ=0.57, 95% CI 0.54–0.60), and polymorphous vessels (κ=0.55, 95% CI 0.52–0.58) had the highest interobserver agreement among the included dermoscopic features (**Fig. 5**). These 3 features also correlated with invasive melanomas as well as invasive melanomas ≥1.0 mm Breslow (**Fig. 6**).

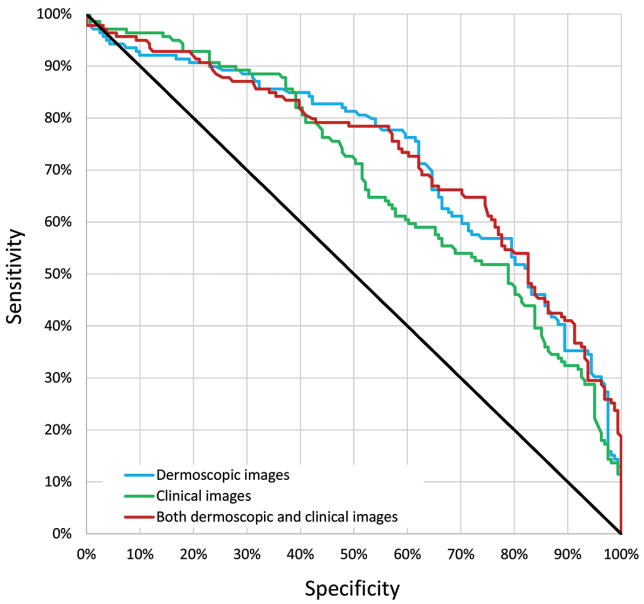

Fig. 3. Area under the receiver operating characteristics (AUC) curves for different convolutional neural network (CNN) models. The graphs represent the AUCs for 3 different CNN models: a merged CNN integrating both dermoscopic and clinical close-up image (primary model); a CNN based on dermoscopic images only, and a CNN based on clinical close-up images only.

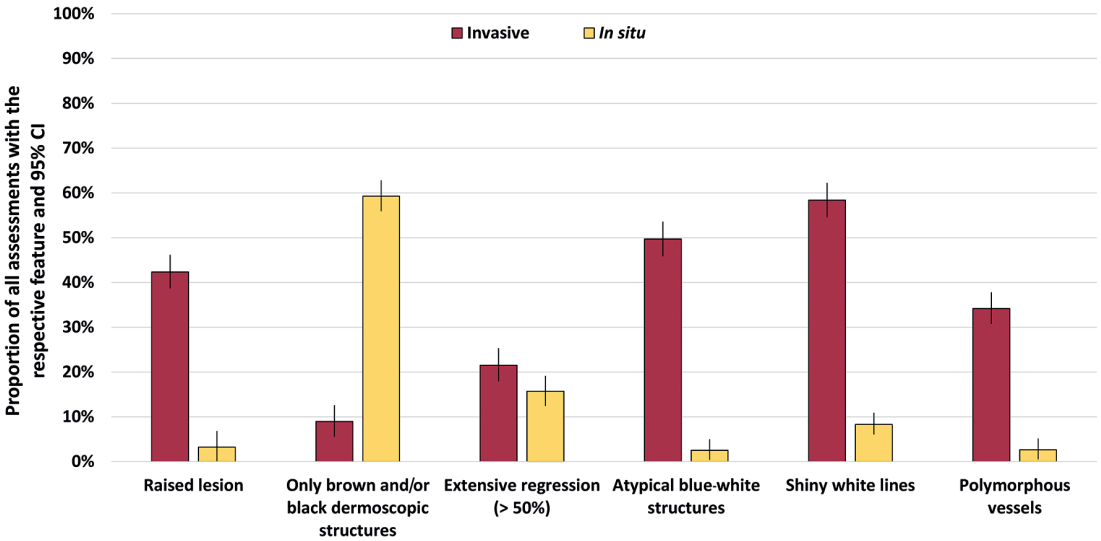

Fig. 4. Distribution of features. Distribution of all features using all 1,800 assessments (i.e. 6 readers and 300 lesions). The sensitivity and specificity for all features is shown in Table SII. 95% CI: 95% confidence interval.

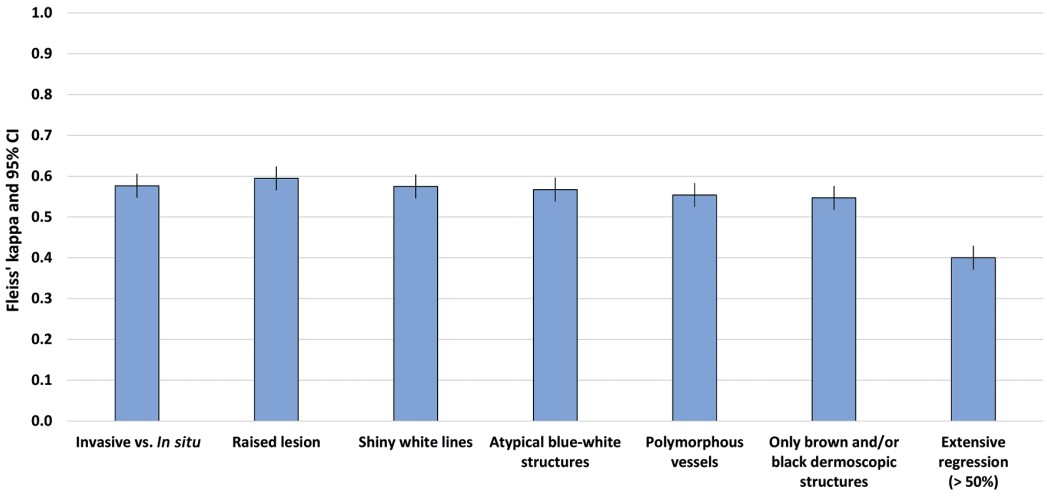

Fig. 5. Interobserver agreement of dermoscopic features. Interobserver agreement for all dermoscopic features among the 300 included melanomas. 95% CI: 95% confidence interval.

DISCUSSION

The merged CNN model, using both clinical close-up and dermoscopic images, was outperformed by combined dermatologists in deciding whether a melanoma is invasive or *in situ*. Moreover, the merged CNN performed on par with a corresponding CNN model that included only dermoscopic images, but outperformed a CNN that integrated only clinical close-up images.

In a real-life setting, the piece(s) of information that exactly affect whether physicians consider a melanoma to be invasive or *in situ* will vary. However, the current study provides *in silico* evidence that the dermoscopic image, most likely, provides more discriminatory information compared with the clinical close-up image. This is also in agreement with a study by Tschandl et al. (13), who concluded that correctly determining the malignancy status for non-pigmented melanomas was higher using a CNN based on dermoscopic images (51%, 95% CI 45–56%) compared with a CNN based on clinical close-up images (23%, 95% CI 19–28%). Although speculative, we believe that, if dermatologists had to choose between close-up or dermoscopic images when determining melanoma thickness, most would opt for the latter. For selected cases, the clinical close-up image does not align with the dermoscopic image. In our experience, for most of these cases, the dermoscopic features will more often be preferred to the details provided by a clinical close-up image.

Feature	Odds ratio (95% CI)	*P*-value
Raised lesion	30 (12 - 77)	*P* < 0.0001
Atypical blue-white structures	14 (6 - 30)	*P* < 0.0001
Polymorphous vessels	8.8 (3.8 - 20)	*P* < 0.0001
Shiny white lines	7.1 (3.7 - 14)	*P* < 0.0001
Extensive regression (>50%)	0.45 (0.19 - 1.0)	*P* = 0.062
Only brown and/or black dermoscopic structures	0.11 (0.051 - 0.22)	*P* < 0.0001
Raised lesion	110 (33 - 361)	*P* < 0.0001
Polymorphous vessels	24 (9.1 - 62)	*P* < 0.0001
Atypical blue-white structures	18 (7.3 - 47)	*P* < 0.0001
Shiny white lines	18 (6.7 - 50)	*P* < 0.0001
Extensive regression (>50%)	0.15 (0.027 - 0.81)	*P* = 0.027
Only brown and/or black dermoscopic structures	0.0026 (0.00010 - 0.065)	*P* = 0.0003

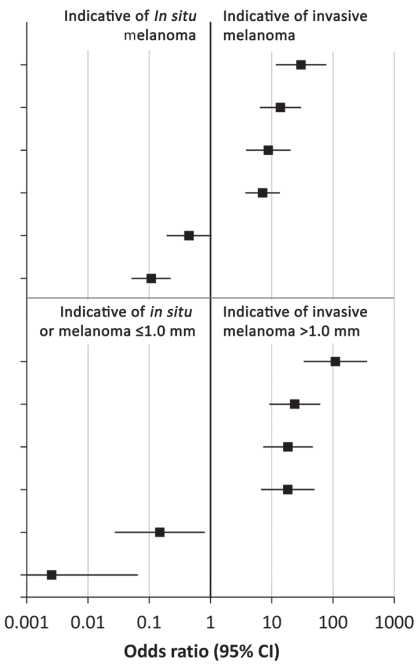

Fig. 6. Odds ratios (ORs) for dermoscopic features indicative of *in situ* or invasive melanomas among the 300 melanomas (161 *in situ* (53.7%) and 139 invasive (46.3%) melanomas). ORs for dermoscopic features indicative of *in situ* and thin (≤1.0 mm) melanomas combined ($n = 260$) or invasive melanomas >1.0 mm ($n = 40$). 95% CI: 95% confidence interval.

The overall interobserver agreement between dermatologists in terms of both clinical and dermoscopic features ranged from moderate to substantial. To be useful for physicians, clinical scoring systems and/or mnemonics ideally need to be easy to remember and, perhaps more importantly, display a high level of interobserver agreement.

Future research aims to refine and test a clinical scoring system comprising both clinical and dermoscopic features in a larger setting to determine how useful it is for dermatologists in the clinical setting. A scoring system, such as the blotches, ridge pattern, asymmetry of structures, asymmetry of colours, furrow pattern and fibrillar pattern (BRAAFF) algorithm for acral melanomas (14), might prove useful in this setting.

Importantly, for the current test set, lesions considered to be raised by the dermatologists were more indicative of invasive melanomas (sensitivity 42.3% and specificity 96.7%). This clinical feature also showed moderate to substantial agreement between the 6 dermatologists. A clinical evaluation of whether a lesion is raised from evaluation of an image is, of course, inferior to assessment in a real-life setting where the clinician may feel the lesion. However, even then, the agreement might not be perfect. A prospective clinical trial that targets solely the interobserver agreement with respect to whether clinicians consider a lesion is raised would be a valuable addition to the literature, particularly since this feature, intrinsically, may provide important metadata for future ML algorithms.

This study has some limitations. In a clinical preoperative setting, a dermatologist integrates more metadata other than the clinical close-up image and the dermoscopic features when estimating melanoma thickness. Moreover, there are melanomas that are obviously invasive for which there is no need to consult a CNN model. However, for thinner lesions (i.e. *in situ* melanomas and thin invasive melanomas ≤ 1.0 mm) the differentiation between invasive and *in situ* melanomas is often more challenging. For this subset, a CNN model may prove useful. While a CNN model integrating only dermoscopic images outperformed a corresponding model including only clinical close-up images, it is worth mentioning that the image resolution and quality was generally lower for the latter. While dermoscopic images are captured in a fairly uniform way, there is no general rule on how to standardize clinical close-up images (e.g. the exact distance between the camera and the lesion). Moreover, the main purpose of obtaining a clinical close-up image of a suspected lesion in routine healthcare is most often to provide the physician with an idea of how to locate the suspected lesion rather than providing exact diagnostic information. It is not ruled out that image standardization for our clinical close-up images would have improved the CNN models. It is noteworthy that the current dataset included only melanomas of varying thickness. In a preoperative setting, dysplastic naevi are often an important differential diagnosis. Further research aims to investigate how a CNN model trained, validated, and tested on melanomas performs on an "out-of model" test set consisting of dysplastic naevi.

The combined dermatologist's assessment outperformed the CNN. However, in a real-life setting there are usually insufficient resources for a combined assessment, and the clinical decisions rely mostly on a single physician's assessments. In this context, it is noteworthy that only 3 of the 6 dermatologists significantly outperformed the CNN. Lesion diameter was not included, but since the dermoscopy ruler was depicted in most of the lesions, the dermatologists were able to integrate this metadata in their assessment. This might have provided an advantage for the dermatologists compared with the CNN models. Future research should preferably include more clinical metadata, including lesion diameter. Alternatively, it would be interesting to develop CNN models that focus on the presence or absence of predefined dermoscopic features rather than melanoma thickness *per se*. In a clinical setting, this might interfere less with the normal workflow for dermatologists.

Inclusion of a certainty score improved the AUC for the dermatologists. Other research groups with similar ideas should consider including a level of confidence (very uncertain to very certain) as an extra parameter in their assessments. While this score may seem contrived at first, we believe that a clinician can relate to the level of certainty when making clinical decisions. This would also be feasible in a prospective setting.

The great interest around use of ML in dermatology has not been matched with prospective clinical trials. Fundamentally, it is only after ML tools have been assessed in clinical trials that we can better understand how they will integrate into routine healthcare. Even then, there is often a considerable discrepancy between the environment in a trial compared with everyday clinical practice. Ultimately, the implementation phase of these tools is very much a cognitive endeavour, and one of the challenges is whether to consider the algorithmic output before or after an initial clinician assessment. Finally, the sensitivity and specificity threshold must be carefully balanced in order to control false-positive rates.

ACKNOWLEDGEMENTS

This study was reviewed and approved by the Regional Ethics Review Board in Gothenburg (approval number 283–18).

The study was financed by grants from the Swedish state under the agreement between the Swedish Government and the county councils, the ALF-agreement (ALFGBG-728261).

The authors have no conflicts of interest to declare.

REFERENCES

1. Kovarik C, Lee I, Ko J, Ad Hoc Task force on augmented I. Commentary: position statement on augmented intelligence (AuI). J Am Acad Dermatol 2019; 81: 998–1000.
2. Polesie S, McKee PH, Gardner JM, Gillstedt M, Siarov J, Neittaanmaki N, et al. Attitudes toward artificial intelligence within dermatopathology: an international online survey. Front Med (Lausanne) 2020; 7: 591952.
3. Polesie S, Gillstedt M, Kittler H, Lallas A, Tschandl P, Zalaudek I, et al. Attitudes towards artificial intelligence within dermatology: an international online survey. Br J Dermatol 2020; 183: 159–161.
4. Emanuel EJ, Wachter RM. Artificial intelligence in health care: will the value match the hype? Jama 2019; 321: 2281–2282.
5. Janda M, Soyer HP. Can clinical decision making be enhanced by artificial intelligence? Br J Dermatol 2019; 180: 247–248.
6. Polesie S, Jergeus E, Gillstedt M, Ceder H, Dahlen Gyllencreutz J, Fougelberg J, et al. Can dermoscopy be used to predict if a melanoma is in situ or invasive? Dermatol Pract Concept 2021; 11: e2021079.
7. Polesie S, Gillstedt M, Kittler H, Rinner C, Tschandl P, Paoli J. Assessment of melanoma thickness based on dermoscopy images: an open, web-based, international, diagnostic study. J Eur Acad Dermatol Venereol 2022 Jul 16. [Online ahead of print].
8. Polesie S, Sundback L, Gillstedt M, Ceder H, Dahlen Gyllencreutz J, Fougelberg J, et al. Interobserver agreement on dermoscopic features and their associations with in situ and invasive cutaneous melanomas. Acta Derm Venereol 2021; 101: adv00570.
9. Gillstedt M, Hedlund E, Paoli J, Polesie S. Discrimination between invasive and in situ melanomas using a convolutional neural network. J Am Acad Dermatol 2022; 86: 647–649.
10. Polesie S, Gillstedt M, Ahlgren G, Ceder H, Dahlen Gyllencreutz J, Fougelberg J, et al. Discrimination between invasive and in situ melanomas using clinical close-up images and a de novo convolutional neural network. Front Med (Lausanne) 2021; 8: 723914.
11. Fleiss JL. Measuring nominal scale agreement among many raters. Psychol Bull 1971; 76: 378.
12. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33: 159–174.
13. Tschandl P, Rosendahl C, Akay BN, Argenziano G, Blum A, Braun RP, et al. Expert-level diagnosis of nonpigmented skin cancer by combined convolutional neural networks. JAMA Dermatol 2019; 155: 58–65.
14. Lallas A, Kyrgidis A, Koga H, Moscarella E, Tschandl P, Apalla Z, et al. The BRAAFF checklist: a new dermoscopic algorithm for diagnosing acral melanoma. Br J Dermatol 2015; 173: 1041–1049.