POLITECNICO DI TORINO
Repository ISTITUZIONALE

Impact of artificial intelligencebased color constancy on dermoscopical assessment of skin lesions: A comparative study

Original
Impact of artificial intelligencebased color constancy on dermoscopical assessment of skin lesions: A comparative study / Branciforti, F., Meiburger, K.M., Zavattaro, E., Veronese, F., Tarantino, V., Mazzoletti, V., Cristo, N.D., Savoia, P., Salvi, M.. - In: SKIN RESEARCH AND TECHNOLOGY. - ISSN 0909-752X. - STAMPA. - 29:11(2023). [10.1111/srt.13508]

Availability:
This version is available at: 11583/2983388 since: 2023-10-27T10:47:59Z

Publisher:
Wiley

Published
DOI:10.1111/srt.13508

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

29 November 2026

Check for updates

Received: 24 August 2023 | Accepted: 12 October 2023

DOI: 10.1111/srt.13508

ORIGINAL ARTICLE

WILEY

Impact of artificial intelligence-based color constancy on dermoscopical assessment of skin lesions: A comparative study

Francesco Branciforti[1] | Kristen M. Meiburger[1] | Elisa Zavattaro[2] | Federica Veronese[3] | Vanessa Tarantino[3] | Vanessa Mazzoletti[2] | Nunzia Di Cristo[2] | Paola Savoia[2] | Massimo Salvi[1]

[1]Biolab, PolitoBIOMed Lab, Department of Electronics and Telecommunications, Politecnico di Torino, Turin, Italy

[2]Department of Health Science, University of Eastern Piedmont, Novara, Italy

[3]AOU Maggiore della Carità, Novara, Italy

Correspondence
Kristen M. Meiburger, Department of Electronics and Telecommunications, Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129 Torino, Italy.
Email: kristen.meiburger@polito.it

Abstract

Background: The quality of dermoscopic images is affected by lighting conditions, operator experience, and device calibration. Color constancy algorithms reduce this variability by making images appear as if they were acquired under the same conditions, allowing artificial intelligence (AI)-based methods to achieve better results. The impact of color constancy algorithms has not yet been evaluated from a clinical dermatologist's workflow point of view. Here we propose an in-depth investigation of the impact of an AI-based color constancy algorithm, called DermoCC-GAN, on the skin lesion diagnostic routine.

Methods: Three dermatologists, with different experience levels, carried out two assignments. The clinical experts evaluated key parameters such as perceived image quality, lesion diagnosis, and diagnosis confidence.

Results: When the DermoCC-GAN color constancy algorithm was applied, the dermoscopic images were perceived to be of better quality overall. An increase in classification performance was observed, reaching a maximum accuracy of 74.67% for a six-class classification task. Finally, the use of normalized images results in an increase in the level of self-confidence in the qualitative diagnostic routine.

Conclusions: From the conducted analysis, it is evident that the impact of AI-based color constancy algorithms, such as DermoCC-GAN, is positive and brings qualitative benefits to the clinical practitioner.

KEYWORDS
AI, color constancy, dermoscopy, generative adversarial networks, melanoma, non-melanoma skin cancer

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2023 The Authors. *Skin Research and Technology* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

Melanoma and non-melanoma skin cancer are common cancers worldwide, and their incidence have increased over recent years.[1] The periodic screening of skin lesions is essential for early diagnosis and hence a better prognosis of affected patients.[1]

Nowadays, the screening process is typically based on dermoscopy; dermatoscopes allow both a visual analysis and the subsequent acquisition of a high-resolution and epiluminescent digital image of the skin lesion. Teledermatology, the provision of dermatology services at a distance, particularly benefits from the acquisition of digital dermoscopic images and can increase patient access to expert dermatologic care. In this context, numerous studies in literature have also recently focused on the use of a smartphone camera for skin lesion image acquisition, which could provide a potential tool for general practitioners who can filter out the clearly benign skin lesions and consult expert dermatologists only in dubious cases.[2–5] The acquisition of digital images of skin lesions further allows for the application of computer-assisted methods, which have shown to obtain more accurate diagnostic results.[6–9] However, the quality and appearance of digital skin lesion images depend on numerous factors. In fact, illumination conditions, device calibration, and operator dexterity are the main factors affecting the quality of digital dermatoscopic, and in general skin lesion, images. There is hence a dire need to reduce image quality variability to make images acquired under different conditions and presets more similar to each other to facilitate skin lesion analysis and diagnosis. To respond to this need, color constancy algorithms have been introduced. The process of color constancy involves transforming the colors of an image acquired under unknown light as if it were acquired under defined light conditions, usually perfectly white light.[10] It's important to distinguish that color constancy is not identical to normalization; while they are closely related, color normalization specifically involves reducing or eliminating color variations across different images or within the same image, bringing all images to a common color standard or scale. Still, for simplicity purposes, we will refer to the application of a color constancy method also as "color normalization" in this paper. Figure 1 shows the intrinsic variability in color and illumination that characterizes dermatological images (Figure 1A), and the effect of color constancy on the same images (Figure 1B).

Several color constancy algorithms exist in the literature, such as Gray World,[11] Shades Of Gray,[12] and max-RGB.[13] These algorithms, while still widely employed, are limited by statistic-based assumptions used to perform color constancy, which results in suboptimal color constancy normalizations.

Recently, artificial intelligence (AI) methods have enabled more versatile, optimized, robust and fast algorithms to perform color constancy.[14,15] Within this context, DermoCC-GAN (Dermatological Color Constancy Generative Adversarial Network), the first AI method in the literature that performs a color constancy task on dermatoscopic images as an image-to-image translation problem with a Generative Adversarial Network (GAN), was developed.[15] As shown in Figure S1, DermoCC-GAN consists of two deep neural networks, the generator and the discriminator, that are trained together in an adversarial game with the goal of improving each other. The training phase ends when the generator succeeds in fooling the discriminator, generating realistic and accurate images. This results in a deep network that can perform the color constancy task quickly and efficiently on images characterized by different illumination conditions.

The use of a color constancy method as a preprocessing step has been widely demonstrated in the literature to improve the performance of AI, and specifically deep learning, methods for classification and/or segmentation tasks.[10,15,16,17] However, to the best of our knowledge, no study has yet investigated the impact of the color constancy process on skin lesions diagnostic routine or the qualitative improvement due to this process from the clinical dermatologist's perspective.

In this study, we assess the impact of the color constancy process within the diagnostic routine and specifically during the evaluation of different skin lesions. The analysis is focused on the evaluation of several parameters, such as the perceived image quality, the diagnosis for the patient, and the confidence of diagnosis, for both original and normalized dermatological images.

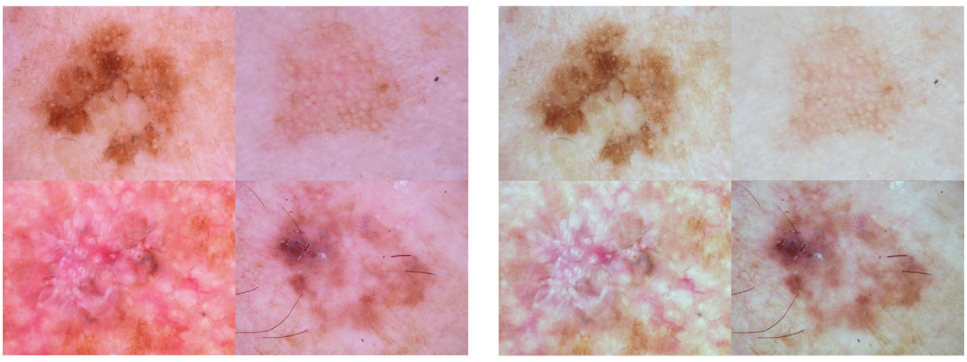

FIGURE 1 Color constancy process for dermatological images. (A) Original images, (B) Images after applying color constancy.

16000846, 2023, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/srt.13508 by CochraneItalia, Wiley Online Library on [27/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Composition of the starting dataset.

Lesion type	Labels	# Images
Actinic keratosis	AKIEC	29
Basal cell carcinoma	BCC	31
Keratosis-like	KL	32
Melanoma	MEL	26
Nevus	NV	32

2 | MATERIALS AND METHODS

The starting dataset consists of 150 dermoscopic images from the open access dataset HAM10000.[18] Images were selected from different subsets with different acquisition settings. The dataset includes five different types of lesions, and the division is shown in Table 1.

Despite the presence of distinctive clinical and dermoscopic features, the differential diagnosis between these types of lesions, and in particular the discrimination between definitely benign lesions, that do not require treatment, and suspected malignancy that need to be resected, can in fact be a challenge, not only for the general practitioner but also for the dermatologist.

The choice of lesions in this study was guided by their relative frequency in the clinical setting, as well as their diverse morphological characteristics which represent different diagnostic challenges. So, while this study focuses on performance obtained on these five lesions, the good performance of DermoCC-GAN is not limited to these lesion types. Examples of DermoCC-GAN processed images on other types of lesions (i.e., dermatofibromas and vascular lesions) are shown in the Supplementary Materials (Figure S2). Importantly, DermoCC-GAN also performs effectively across a range of skin tones as evidenced in Figure S3.

All images in the primary dataset were processed with DermoCC-GAN, our AI-based color-constancy framework.[15] This results in a final pool of 300 images consisting of 150 original images and their respective normalized images. An example of a dermoscopic image processed with DermoCC-GAN is provided in Figure S1. DermoCC-GAN is a Pix2Pix-GAN AI algorithm trained to perform a domain transfer task, from the original image color domain to the color-constancy normalized domain. Unlike heuristic color constancy algorithms in the state of the art,[11–13] DermoCC-GAN is parameter-free and not based on hypothetical starting assumptions. These features give the tool a high generalization capability, which enables it to quickly and optimally process images that have never been seen during the development phase, overperforming the state-of-the-art color constancy algorithms. Looking at Figure S4, it can be seen that color-constancy performed by DermoCC-GAN is more realistic and color-balanced than the most widely used state-of-the-art methods.

This study involved the expertise of three dermatologists, each with a differing level of experience: *E.Z.* (17 years), *F.V.* (11 years), and *V.M.* (3 years). To assess the impact of the color constancy normalization, the clinicians carried out two different assignments:

1. *Unpaired evaluation task*, in which two dermatologists, with different clinical experience, evaluated the original and normalized images independently (for a total of 300 images each). All images were provided in random order, and the DermoCC-GAN normalized images were also randomly rotated and flipped to avoid evaluation bias.
2. *Paired evaluation task*, in which two experienced dermatologists independently assessed the images using a paired approach, closely examining the original and AI-processed versions of each image side-by-side. To minimize recall bias, one dermatologist (E.Z.) conducted this evaluation 6-months after participating in the unpaired evaluation task. In this case, the normalized images did not undergo any kind of geometric transformation.

A visual overview of the two tasks is shown in Figure 2. The following subsections give a detailed explanation of these assignments.

2.1 | Unpaired evaluation task

Using a dedicated Graphical User Interface (GUI), dermatologists *E.Z.* and *V.M.* provided the following evaluations for each image:

- Image quality: using a scale from 1 to 5, where 1 indicates a low-quality image and 5 a high-quality image in terms of chromaticity, brightness, and sharpness.
- Lesion diagnosis: assigning the label related to the diagnosis based solely on a visual analysis of the image (i.e., AKIEK, BCC, KL, MEL, NV).
- Diagnosis confidence: using a scale from 1 to 5, where 1 indicates a low degree of confidence in the diagnosis and 5 denotes a high degree of confidence. Operationally, a high confidence diagnosis is one where the dermatologist believes that the given image has sufficient information to make a diagnosis, while a low confidence indicates that the dermatologist is not fully convinced by the appearance of the examined image.

The goal of this task is to test the contribution of the DermoCC-GAN color constancy normalization when the original image and the corresponding normalized image are submitted to the clinicians at different times.

2.2 | Paired evaluation task

In order to assess the effectiveness of the DermoCC-GAN color constancy normalization when displayed concurrently with the original image, the experienced dermatologists *F.V.* and *E.Z* performed a paired evaluation task. Considering *E.Z.*'s prior involvement in the unpaired evaluation task, a 6-month interval was ensured before *E.Z.* participated in this stage to minimize recall bias. For each paired image, the dermatologists provided evaluations based on the following criteria:

16000846, 2023, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/srt.13508 by CochraneItalia, Wiley Online Library on [27/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

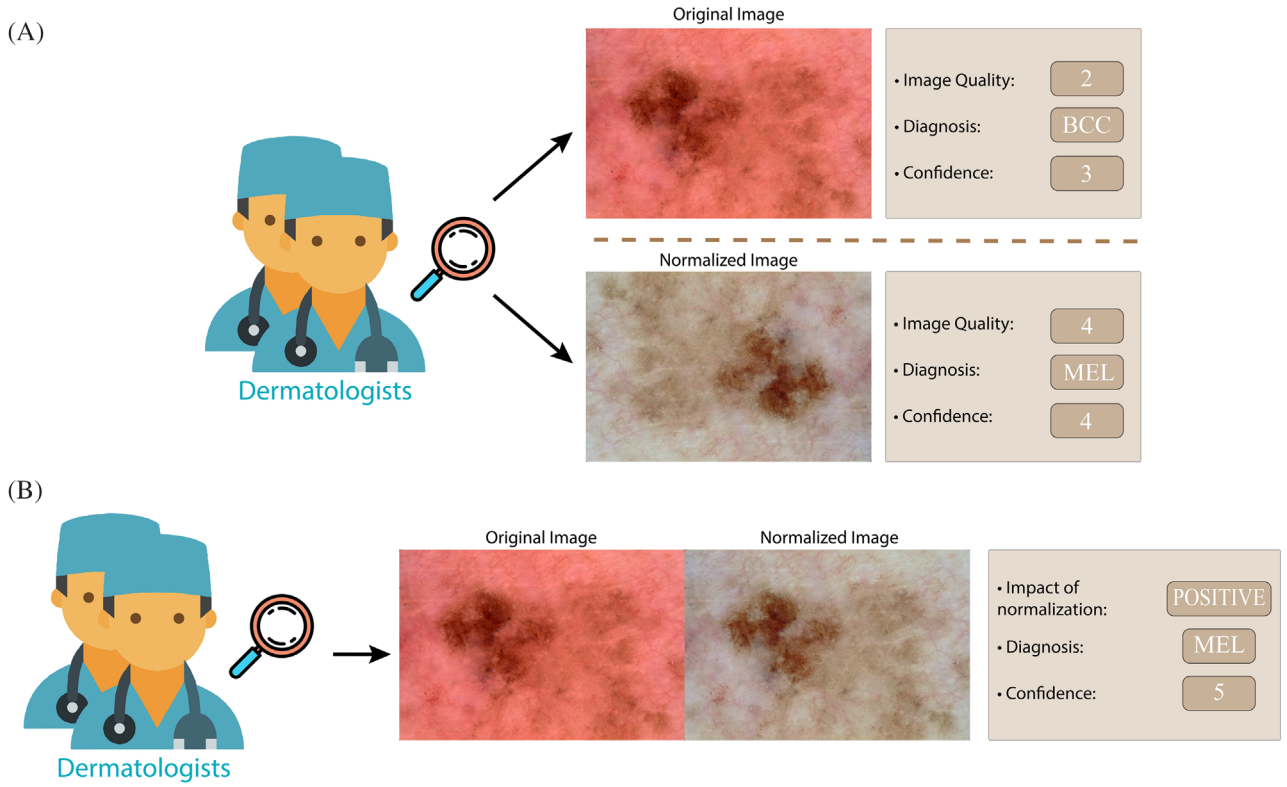

FIGURE 2 Examples of unpaired and paired evaluation tasks on a melanoma image. (A) In the unpaired evaluation task, the clinicians independently evaluate the original and normalized images at two different times, assigning the proposed scores. To avoid evaluation bias, the normalized image is randomly rotated and/or flipped. (B) In the paired evaluation task, the clinicians independently assess the original and normalized images presented side-by-side, assigning the proposed scores.

- Impact of normalization: this score quantifies the usefulness of the information brought by the normalized images. The clinician can select a score of negative if the normalized image confuses the diagnosis, neutral if it does not add any extra valuable information, and positive if the normalized image adds crucial information and simplifies the interpretation of the lesion.
- Lesion diagnosis: assigning the label related to the diagnosis based solely on a visual analysis of the paired images (i.e., AKIEK, BCC, KL, MEL, NV).
- Diagnosis confidence: using a scale from 1 to 5, where 1 represents the lowest degree of confidence and 5 the highest degree of confidence. In this context, confidence implies the degree of certainty with which the physician assigns a particular diagnosis to the skin lesion.

3 | RESULTS

3.1 | Image quality evaluation

In this section we compare the image quality scores assigned by clinicians in the unpaired evaluation task. The operators evaluated brightness, sharpness, and the chromatic component of the images. Ratings assigned by each clinician were firstly divided into original image scores and normalized image scores, in order to analyze any differences. Figure 3 shows the distribution of the original and normalized images quality for dermatologists *E.Z.* and *V.M.*: for both clinicians the overall quality of the normalized images is higher than the originals. Specifically, for the *E.Z.* operator (17 years of experience), low-quality images (score 1 and 2) account for 31.3% for original images and only 20.6% for normalized images. Similarly, the number of normalized images considered to be of high-quality (score 4 and 5) increases compared to the originals (27.7% vs. 38.6%).

For the *V.M.* operator (3 years of experience), we observe an increase in the median value of the distribution from 2.5 (for the original images) to 3 (for the normalized images). This result shows that, regardless of experience level, the images normalized by DermoCC-GAN are visually perceived as being of higher quality than the original ones.

3.2 | Impact of normalization

This section illustrates the impact of the simultaneous analysis of both original and normalized images on clinicians. Figure 4 shows the distribution of scores assigned by the dermatologists *F.V.* and *E.Z.* in the paired evaluation task. The clinicians assessed the usefulness of information provided by the normalized images, assigning low scores when the normalized image carried potentially disturbing or

16000846, 2023, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/srt.13508 by CochraneItalia, Wiley Online Library on [27/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

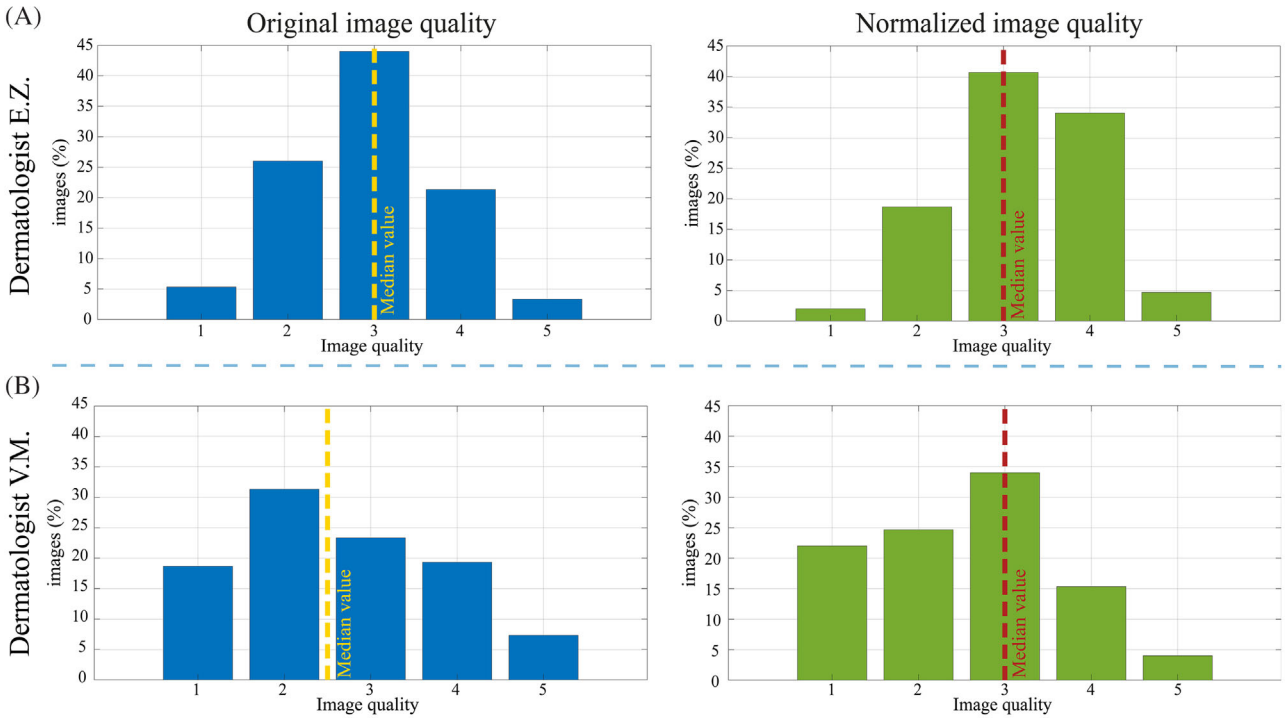

FIGURE 3 Comparison between original and normalized image quality scores assigned by dermatologist E.Z. (A) and dermatologist V.M. (B) in the unpaired evaluation task. Dashed lines represent the median value of the distributions.

misleading information, neutral when it did not provide additional information, and positive when it simplified lesion interpretation, providing additional knowledge. As can be seen from Figure 4, none of the normalized images had a negative impact on either doctor's diagnosis.

Specifically, for the *F.V.* operator (11 years of experience), in 53% of the cases the normalized image did not provide any additional information, while in 47% of the cases it had a positive impact on the lesion analysis. For the *E.Z.* operator (17 years of experience) the percentage of images with neutral scores was notably lower at 7%, compared to *F.V.*'s 53%. Moreover, the percentage of images with positive scores was significantly higher at 93%, compared to *F.V.*'s 47%.

3.3 | Diagnosis of skin lesions: Accuracy and confidence

This section shows the results obtained in the diagnosis of skin lesions and their confidence scores. In order to analyze any different trends, the assessments on the unpaired evaluation task were divided into two separate analyses dividing the original images and the normalized ones. The results of the lesion diagnosis are shown in Figure S5 in the form of a multiclass confusion matrix, a kind of contingency table that shows the predicted and the actual labels for each lesion class. Each confusion matrix is accompanied by its overall accuracy score.

Examining the overall accuracy, it can be seen that the diagnostic accuracy between original and normalized images in the unpaired evaluation task shows little difference, with both operators demonstrating similar performance. Moreover, analyzing the paired data (original + normalized) revealed that *E.Z.* reached an accuracy of 65.33% and *F.V.* achieved an overall accuracy score of 74.77%. This score is higher than the best overall scores obtained by operators *E.Z.* (67.33%) and *V.M.* (69.33%) in the unpaired task.

Figure 5 shows the confidence scores assigned in unpaired and paired evaluation tasks. The confidence values were grouped into three classes: low confidence (values of 1–2); medium confidence (value equal to 3); high confidence (values of 4–5). The assessments in the unpaired evaluation task were divided into two separate analyses: one for original images and another for normalized images, in order to assess how the confidence of the diagnosis made on the original images varies from the confidence when considering the normalized images.

In the unpaired evaluation task, both operators showed an increase in confidence level for normalized images. In particular:

- For the *E.Z.* operator, we observe a decrease in low confidence from 18.66% to 11.33%, and an increase in medium and high confidence, respectively, from 24% to 27.33% and from 57.34% to 61.34%.
- For the *V.M.* operator, probably due to less experience, there is a smaller increase on the high confidence, which goes from 42% to 42.66%.

In the paired evaluation task, dermatologist *F.V.* assigned a high confidence score to 72.67% of the images and a medium confidence score to 27.33% of the images. The confidence of the operator was never low when viewing the original image alongside the normalized image. Similarly, in the paired evaluation task, *E.Z.* assigned a high confidence score to 76.67% of the images, a medium confidence score to 20% of

16000846, 2023, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/srt.13508 by CochraneItalia, Wiley Online Library on [27/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

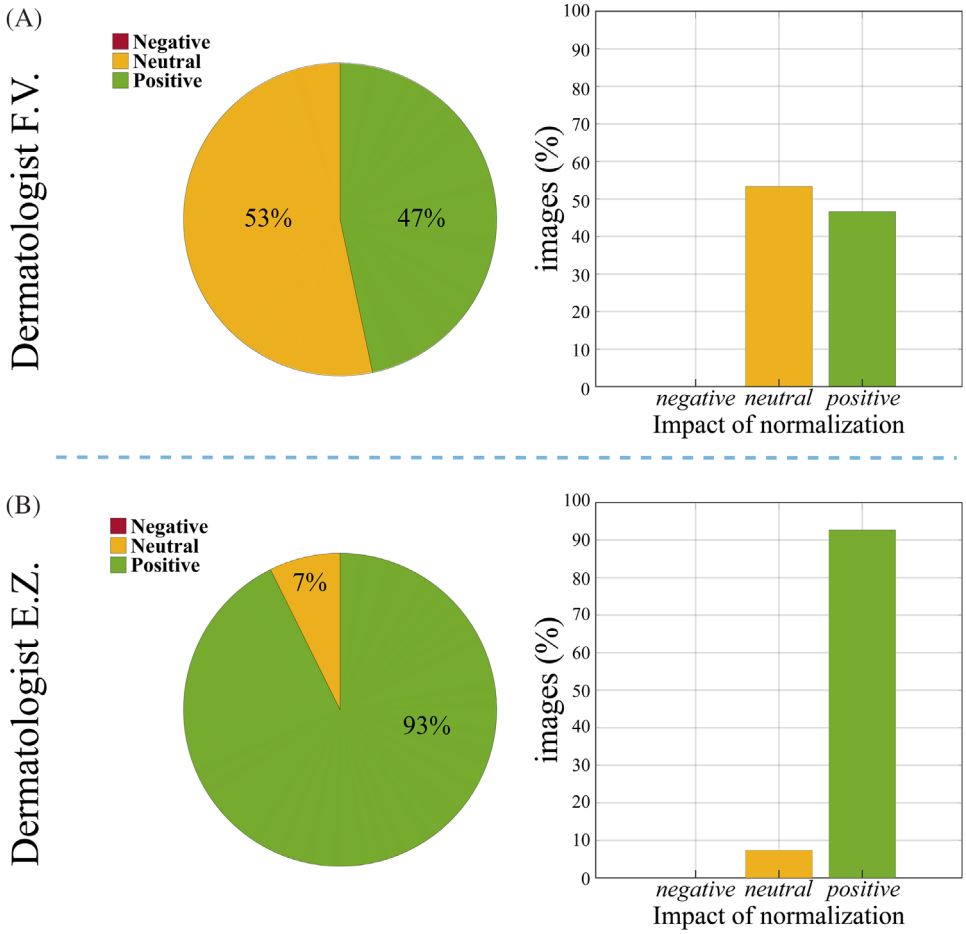

FIGURE 4 Comparison of normalization impact scores assigned by dermatologist F.V. (A) and dermatologist E.Z. (B) in the paired evaluation task. The pie and bar chart distribution illustrate negative scores in red, neutral scores in yellow and positive scores in green.

the images, and a low confidence score to only 3.33% of the images. *E.Z.*'s performance in the paired evaluated task shows a noticeable increase in high confidence scores and a decrease in low confidence scores compared to the unpaired task (Figure 5D vs. Figure 5A), indicating improved confidence when analyzing original and normalized images side-by-side.

4 | DISCUSSION AND CONCLUSIONS

In this paper, we propose a study on the impact that color constancy normalization by the AI-based DermoCC-GAN algorithm has on the diagnostic phase of dermoscopic images when evaluated in clinical practice. Three dermatology clinicians with different levels of experience were involved in this investigation and carried out two main assignments:

- *Unpaired evaluation task*, which consisted of analyzing a pool of original and normalized images presented in a random order and unbiased manner, evaluating the image quality, determining a diagnosis to the skin lesion, and grading the confidence level related to the diagnosis. This task was done by two operators (*E.Z.* and *V.M).*
- *Paired evaluation task*, in which two clinicians (*F.V.* and *E.Z.*) evaluated the original-normalized image pairs, assigning scores related to the impact of the color constancy normalization, determining a diagnosis to the skin lesion, and grading the confidence level related to the diagnosis. To minimize recall bias, *E.Z.* conducted this assessment 6 months after the unpaired evaluation task.

DermoCC-GAN is a color constancy tool based on artificial intelligence that has been shown to be quantitatively superior to state-of-the-art methods. The use color constancy tools has also brought performance benefits to segmentation and classification tasks using other artificial intelligence methods.[16,17,19] The main goal of the study presented here is to verify whether the use of normalized images by DermoCC-GAN also brings real qualitative benefit to dermatology clinicians in visual diagnosis, in terms of perceived image quality, impact of normalization, accuracy and confidence level on lesion diagnosis.

Assessing the perception of image quality, it was verified that the DermoCC-GAN normalized images are recognized to be of higher quality on average than the original images, for both operators who performed the unpaired evaluation task. As shown in Figure 3, this perception is more pronounced for operator *V.M.*, with less experience

16000846, 2023, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/srt.13508 by CochraneItalia, Wiley Online Library on [27/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

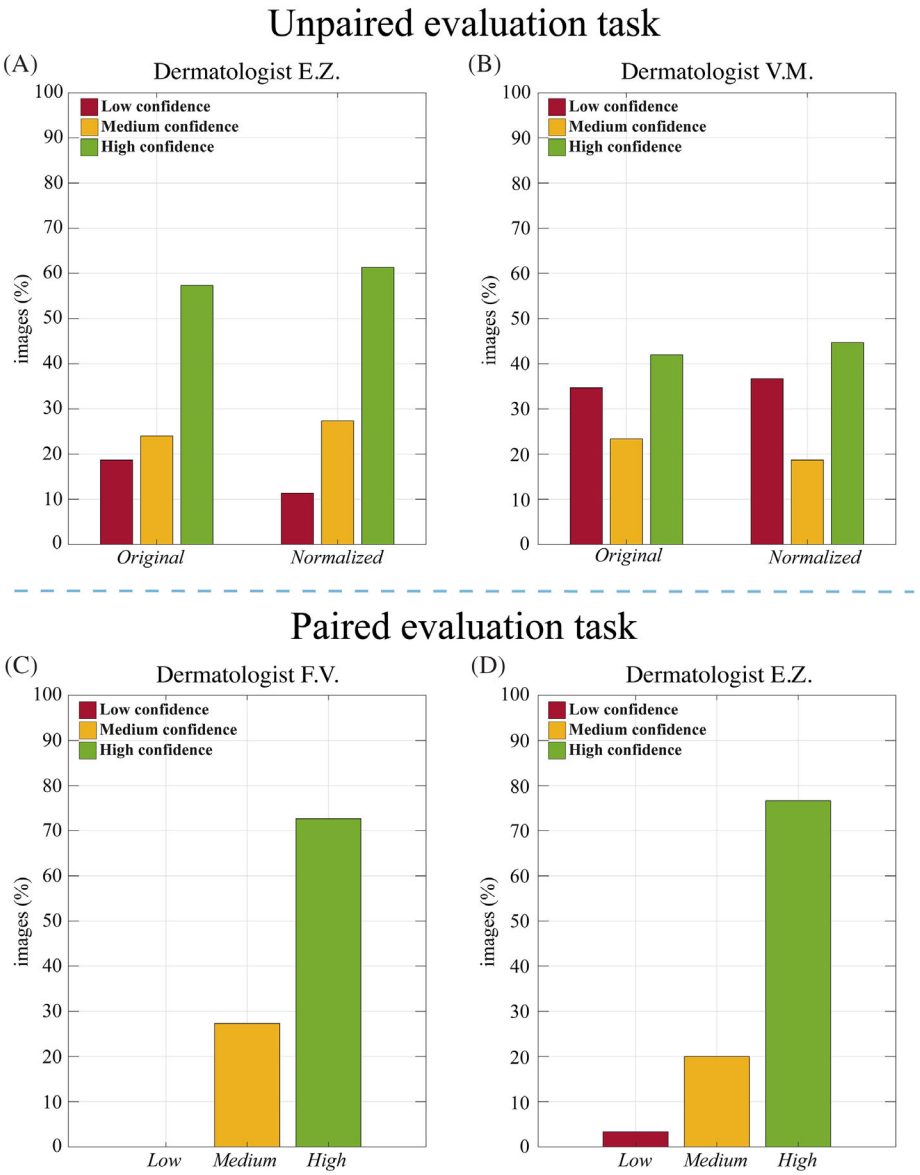

FIGURE 5 Bar chart of distribution of confidence values for unpaired (A. E.Z and B. V.M) and paired evaluation task (C. F.V., D. E.Z.). Confidence values are grouped in three classes: low-quality (1–2) represented in red, medium quality (3) represented in yellow and high-quality (4–5) represented in green.

than *E.Z.*, for whom a 20 percent increase in the median value of the entire normalized image quality distribution is observed.

To evaluate the impact of normalization, the operators who observed the paired images assigned a score related to the quality of the information brought by the normalized image. Figure 4 demonstrates how images normalized by DermoCC-GAN do not add noise to the information carried by the original images. More importantly, by looking at the corresponding normalized image, the clinician is able to obtain additional and useful information that can support them in the diagnostic phase.

Regarding the performance related to lesion diagnosis, shown in Figure S5, it is possible to observe how the diagnoses made on paired images are more accurate than on the single images: *F.V.* clinician, with an intermediate level of experience among the three operators, has an improvement in classification accuracy of up to 74.67%. Thus, we can state that, especially for a novice clinician, the impact of normalization on clinical assessment is greatest when the normalized image is provided alongside the original, leading to an increase in overall classification performance.

Finally, the level of self-assurance with which clinicians made the diagnosis was assessed for both tasks. Figure 5 shows that, for the unpaired assessment, the experienced operator *E.Z.* perceives an increase in confidence by analyzing normalized images, while for the less experienced operator *V.M.*, confidence levels remain stable. For the paired evaluation, the distribution of confidence values assigned by clinicians *F.V.* and *E.Z.* is concentrated at very high confidence levels, far higher than the unpaired task. Moreover, *E.Z.* exhibited better performance in the paired evaluation as compared to the unpaired

16000846, 2023, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/srt.13508 by CochraneItalia, Wiley Online Library on [27/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

task, highlighting the potential benefits of analyzing original and normalized images side-by-side.

Overall, it is possible to say that, in accordance with the results obtained in the lesion diagnosis, the use of normalized images coupled with the originals also increases the level of confidence for the operator. Benefits are found to a lesser extent also when the normalized image is provided individually. A possible explanation could be represented by the improvement in the visualization of certain structures (i.e., blood vessel features, depigmentation and/or regression areas, shiny white structures) in the normalized images. Such aspects can provide a further help in the diagnostic process, in addition to the more global features, that are easily recognizable in the original images.

This study is not free from limitations. One limitation is related to the limited number of images analyzed ($n = 150$), considering five main skin lesions (i.e., AKIEK, BCC, KL, MEL, NV), and the number of dermatologists involved ($n = 3$). We plan to increase these numbers in the future. In addition to the five lesion types originally included, we also processed dermatofibromas and vascular lesions with the DermoCC-GAN. Illustrative examples of this extended application are provided in Figure S2, and early observations suggest that our tool shows promise in these contexts as well. We also intend to extend the study to a more diverse range of skin lesions and pathologies (e.g., squamous cell carcinoma -SCC, keratoacanthoma—KA, vascular lesions of the skin, and rare skin tumors) and expand the use of the normalization process on different datasets (e.g., PH2,[20] derm7pt,[19] and SKINL2.[21]).

In conclusion, analyzing all used indicators, it is evident that the use of normalized images by AI-based color constancy algorithms, such as DermoCC-GAN, brings qualitative benefits to the clinical practitioner on skin lesions diagnostic routine. However, acknowledging the importance of extreme caution in this area, we always suggest the simultaneous analysis of original and normalized images. This approach enables dermatologists to extract essential information from both images, contributing to a more accurate classification. The combination of the original and normalized images not only enhances the diagnostic capability of the clinicians but also strengthens their confidence level during the diagnostic process.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, upon request.

ETHICS STATEMENT

Ethical review and approval were not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

ORCID

Francesco Branciforti https://orcid.org/0000-0002-2922-1558
Kristen M. Meiburger https://orcid.org/0000-0002-7302-6135
Elisa Zavattaro https://orcid.org/0000-0003-4537-3014
Federica Veronese https://orcid.org/0000-0001-6438-5171
Paola Savoia https://orcid.org/0000-0002-1636-8411
Massimo Salvi https://orcid.org/0000-0001-7225-7401

REFERENCES

1. Guy GP, Thomas CC, Thompson T, Watson M, Massetti G M, Richardson LC. Vital signs: melanoma incidence and mortality trends and projections—United States, 1982–2030. *MMWR Morb Mortal Wkly Rep*. 2015;64:591-596.
2. Coates SJ, Kvedar J, Granstein RD. Teledermatology: from historical perspective to emerging techniques of the modern era: Part I: history, rationale, and current practice. *J Am Acad Dermatol*. 2015;72:563-574.
3. Veronese F, Branciforti F, Zavattaro E, et al. The role in teledermoscopy of an inexpensive and easy-to-use smartphone device for the classification of three types of skin lesions using convolutional neural networks. *Diagnostics*. 2021;11:451.
4. Pala P, Bergler-Czop BS, Gwiżdż JM. Teledermatology: idea, benefits and risks of modern age—a systematic review based on melanoma. *Postep Dermatologii i Alergol*. 2020;37:159–167.
5. Veronese F, Tarantino V, Zavattaro E, et al. Teledermoscopy in the diagnosis of melanocytic and non-melanocytic skin lesions: NurugoTM derma smartphone microscope as a possible new tool in daily clinical practice. *Diagnostics*. 2022;12:1–11.
6. Rezvantalab A, Safigholi H, Karimijeshni S. Dermatologist Level Dermoscopy Skin Cancer Classification Using Different Deep Learning Convolutional Neural Networks Algorithms. arXiv Prepr arXiv. 2018.
7. Datta SK, Shaikh MA, Srihari SN, Gao M. Soft attention improves skin cancer classification performance. *Lect Notes Comput Sci*. 2021;12929:13–23.
8. Goceri E. Automated skin cancer detection: where we are and the way to the future. 2021 *44th Int Conf Telecommun Signal Process TSP 2021*. 2021;48–51.
9. Maurya A. Stanley RJ, Lama N, et al. A deep learning approach to detect blood vessels in basal cell carcinoma. *Skin Res Technol*. 2022; 28: 571–576.
10. Barata C, Marques JS, Celebi ME. Improving dermoscopy image analysis using color constancy. 2014. *IEEE Int Conf Image Process ICIP* 2014. 2014; 3527–3531.
11. Buchsbaum G. A spatial processor model for object colour perception. *J Franklin Inst*. 1980;310:1–26.
12. Finlayson GD, Trezzi E. Shades of gray and colour constancy. *Color Imaging Conference*. 2004;37–41.
13. Land EH. The Retinex Theory of Color Vision. 1977.
14. Das P, Baslamisli AS, Liu Y, Karaoglu S, Gevers T. Color constancy by GANs: an experimental survey. arXiv. 2018.
15. Salvi M, Branciforti F, Veronese F, et al. DermoCC-GAN: a new approach for standardizing dermatological images using generative adversarial networks. *Comput Methods Programs Biomed*. 2022;225:107040.
16. Ng J, Goyal M, Hewitt B, Yap MH. The effect of color constancy algorithms on semantic segmentation of skin lesions. Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging. 2019;10953:109530R.
17. Mahbod A, Schaefer G, Wang C, Dorffner G, Ecker R, Ellinger I. Transfer learning using a multi-scale and multi-network ensemble for skin lesion classification. *Comput Methods Programs Biomed*. 2020;193:105475.

16000846, 2023, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/srt.13508 by CochraneItalia, Wiley Online Library on [27/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

18. Tschandl P, Rosendahl C, Kittler H. Data descriptor: The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci Data*. 2018; 5:1–9.
19. Kawahara J, Daneshvar S, Argenziano G, Hamarneh G. Seven-point checklist and skin lesion classification using multitask multimodal neural nets. *IEEE J Biomed Heal Informatics*. 2019; 23: 538–546.
20. Mendonca T, Ferreira PM, Marques JS, Marcal ARS, Rozeira J. PH2 – a dermoscopic image database for research and benchmarking. *Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Annu Int Conf*. 2013;5437–5440.
21. De Faria SMM, Henrique M, Filipe JN, et al. Light field image dataset of skin lesions. *Annu Int Conf IEEE Eng. Med Biol Soc IEEE Eng Med Biol Soc Annu Int Conf*. 2019;3905-3908

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Branciforti F, Meiburger KM, Zavattaro E, et al. Impact of artificial intelligence-based color constancy on dermoscopical assessment of skin lesions: A comparative study. *Skin Res Technol*. 2023;29:e13508. https://doi.org/10.1111/srt.13508

16000846, 2023, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/srt.13508 by CochraneItalia, Wiley Online Library on [27/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License