

Discriminating Basal Cell Carcinoma with Only 20 Dermoscopic Images: A Few-shot, Self-supervised Approach

Kyungho PAIK^{1,2}, Sang Woong YOUN^{1,2}, Jung-Im NA^{1,2}, Chang-Hun HUH^{1,2} and Jung-Won SHIN^{1,2*}

¹Department of Dermatology, Seoul National University Bundang Hospital, Seongnam; ²Department of Dermatology, Seoul National University College of Medicine, Seoul, Republic of Korea. *E-mail: jungwonshin@snuh.org

Submitted Aug 25, 2025. Accepted after revision Sep 19, 2025

Published Oct 23, 2025. DOI: 10.2340/actadv.v105.44686. Acta Derm Venereol 2025; 105: adv44686.

To the Editor,

Recent advances in deep learning algorithms (DLAs) have shown considerable promise in assisting in dermatological diagnosis. However, most systems rely on supervised learning from thousands of meticulously labelled images (1), a requirement that remains unattainable for many institutions due to variability in dermoscopic imaging and the cost of expert annotation. To address this, we applied few-shot learning (FSL), a framework designed to enable model generalization from only a handful of labelled examples per class (2). We present a proof-of-concept study on the application of FSL in dermatology, focusing on a purposely simple binary problem: basal cell carcinoma (BCC) vs common benign lesions such as seborrheic keratosis (SK) or melanocytic nevus. BCC, the most common skin cancer, is often clinically mistaken for SK or nevus, and both lesions are routinely included in its differential diagnosis (3).

A powerful approach to enabling FSL is self-supervised learning (SSL), in which a model first learns rich visual representations from large pools of unlabelled images by solving pretext tasks (4). DINO is a notable SSL framework that trains a vision transformer using a student–teacher self-distillation scheme, encouraging consistent representations across different augmentations of the same image without any ground-truth labels (5). After unsupervised pretraining, the DINO backbone can

be fine-tuned with only a few annotated samples per class (Fig. S1).

We first pretrained a vision transformer backbone with DINO-style self-supervised learning on approximately 100,000 unlabelled cropped non-dermoscopic clinical images of skin lesions to learn visual representation. We then fine-tuned the pretrained network with only 20 dermoscopic images, 10 BCC and 10 benign lesions (Fig. S2) such as SK or nevus, to teach the model the specific distinction of interest. Model performance was evaluated on a separate test dataset of 119 dermoscopic images collected in 2024, with no patient overlap with the training dataset (Fig. S3). The subtypes of basal cell carcinoma included in the training dataset were 2 superficial (20%), 7 nodular (70%), and 1 infiltrating (10%), while the subtypes of BCC included in the test dataset were 7 superficial (31.82%), 10 nodular (45.45%), 2 micronodular (9.09%), and 3 infiltrating (13.64%).

Our DLA achieved an accuracy of 90.5%, sensitivity of 90.3%, specificity of 90.5%, and area under the ROC curve (AUROC) of 0.957, indicating excellent performance. The accuracy was 86.84% when differentiating BCC from SK and 92.23% when differentiating BCC from melanocytic nevus. The ROC curve and confusion matrix are presented in **Fig. 1**. The evaluation results of the DLA for each dermoscopic image and the code used to train the DLA are described in Table S1, and Appendix S1.

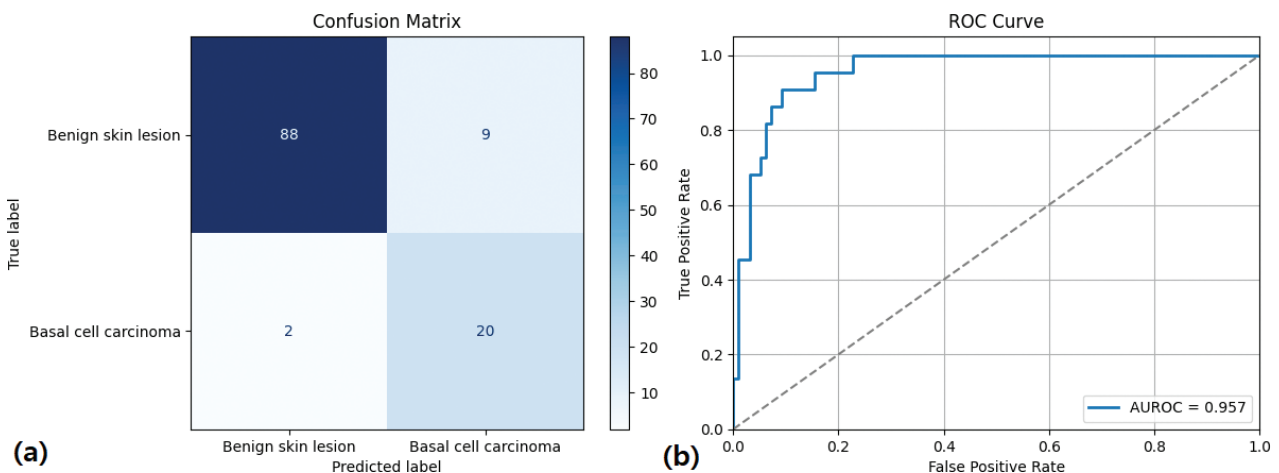


Fig. 1. (a) Confusion matrix of the binary classifier distinguishing basal cell carcinoma (BCC) from benign skin lesions. Darker blue cells correspond to higher case counts. Numbers on the diagonal indicate correctly classified images. (b) Receiver-operating characteristic (ROC) curve illustrating the trade-off between sensitivity and specificity for the classifier distinguishing basal cell carcinoma from benign skin lesions. An area under the curve (AUC) of 0.957 reflects excellent discriminatory performance, with the dashed diagonal representing random chance.

To our knowledge, this is the first report of a self-supervised, few-shot model that distinguishes BCC from common benign skin lesions using only 20 dermoscopic images. Notably, our DLA achieved performance comparable to that of large-scale supervised systems. For instance, Esteva et al. achieved 0.96, and Mei et al. reported an AUROC of 0.79, yet both models were trained on thousands of labelled images (6, 7).

Our approach has two key implications. First, it demonstrates that high-performing AI models can be developed in settings with limited dermoscopic image data. Any clinic with approximately 10 representative dermoscopic images per class could replicate the fine-tuning step to develop locally adapted models without extensive annotation. Second, the model provides an opportunity for automated triage by non-dermatologists. When integrated into mobile platforms, it could assist with early detection of BCC in underserved areas, facilitating timely referral and potentially reducing diagnostic delays and morbidity.

Our study has important limitations. Our goal is not to solve the full complexity of real-world skin cancer screening, but to test whether a deliberately minimalist data regime can still yield clinically meaningful performance. Furthermore, we anticipate that diagnostic difficulty would increase and performance would decline in real-world settings that require multi-class classification encompassing lesions beyond BCC, SK, and melanocytic nevus. Future research can generalize the proposed approach to discriminate among a broad spectrum of conditions and conduct external validation.

In conclusion, using only 10 BCC and 10 benign dermoscopic images for fine-tuning, we developed a vision transformer-based few-shot model that discriminates BCC with AUROC 0.957 and diagnostic accuracy

exceeding 90%. These findings support the feasibility of rapidly developing institution-specific DLA tools for skin cancer screening, even in data-limited environments.

ACKNOWLEDGEMENTS

IRB approval status: The study was approved by the institutional review board of SNUBH (IRB No. B-2407-915-102).

The authors have no conflicts of interest to declare.

REFERENCES

1. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. *Med Image Anal* 2017; 42: 60–88. <https://doi.org/10.1016/j.media.2017.07.005>
2. Ge Y, Guo Y, Das S, Al-Garadi MA, Sarker A. Few-shot learning for medical text: a review of advances, trends, and opportunities. *J Biomed Inform* 2023; 144: 104458. <https://doi.org/10.1016/j.jbi.2023.104458>
3. Takenouchi T. Key points in dermoscopic diagnosis of basal cell carcinoma and seborrheic keratosis in Japanese. *J Dermatol* 2011; 38: 59–65. <https://doi.org/10.1111/j.1346-8138.2010.01093.x>
4. Azizi S, Culp L, Freyberg J, Mustafa B, Baur S, Kornblith S, et al. Robust and data-efficient generalization of self-supervised machine learning for diagnostic imaging. *Nat Biomed Eng* 2023; 7: 756–779. <https://doi.org/10.1038/s41551-023-01049-7>
5. Caron M, Touvron H, Misra I, Jégou H, Mairal J, Bojanowski P, et al. Emerging properties in self-supervised vision transformers. Available from: https://openaccess.thecvf.com/content/ICCV2021/papers/Caron_Emerging_Properties_in_Self-Supervised_Vision_Transformers_ICCV_2021_paper.pdf
6. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017; 542: 115–118. <https://doi.org/10.1038/nature21056>
7. Mei LH, Cao MK, Li J, Ye XG, Liu XD, Yang G. Deep learning in assisting dermatologists in classifying basal cell carcinoma from seborrheic keratosis. *Front Oncol* 2025; 15: 1507322. <https://doi.org/10.3389/fonc.2025.1507322>