

PASSION for Dermatology: Bridging the Diversity Gap with Pigmented Skin Images from Sub-Saharan Africa

Philippe Gottfrois^{7,8}, Fabian Gröger^{5,8}, Faly Herizo Andriambololoniaina³, Ludovic Amruthalingam⁵, Alvaro Gonzalez-Jimenez⁸, Christophe Hsu⁷, Agnes Kessy⁵, Simone Lionetti⁵, Daudi Mavura⁶, Wingston Ng’ambi³, Dingase Faith Ngongonda¹, Marc Pouly⁵, Mendrika Fifaliana Rakotoarisaona⁴, Fahafahantsoa Rapelanoro Rabenja⁴, Ibrahima Traoré², and Alexander A. Navarini^{7,8}

¹ Bwaila Hospital Department of Dermatology

² Clinique dermatologique, Conakry

³ : Health Economics and Policy Unit, Department of Health Systems and Policy, Kamuzu University of Health Sciences, Lilongwe, Malawi

⁴ Laboratoire d’Accueil et de Recherche en santé publique spécialisé en TIC

⁵ Lucerne University of Applied Sciences and Arts

⁶ Regional Dermatology Training Center, Moshi

⁷ University Hospital Basel

⁸ University of Basel

Abstract. Africa faces a huge shortage of dermatologists, with less than one per million people. This is in stark contrast to the high demand for dermatologic care, with 80% of the paediatric population suffering from largely untreated skin conditions. The integration of AI into healthcare sparks significant hope for treatment accessibility, especially through the development of AI-supported teledermatology. Current AI models are predominantly trained on white-skinned patients and do not generalize well enough to pigmented patients. The PASSION project aims to address this issue by collecting images of skin diseases in Sub-Saharan countries with the aim of open-sourcing this data. This dataset is the first of its kind, consisting of 1,653 patients for a total of 4,901 images. The images are representative of telemedicine settings and encompass the most common paediatric conditions: eczema, fungals, scabies, and impetigo. We also provide a baseline machine learning model trained on the dataset and a detailed performance analysis for the subpopulations represented in the dataset. The project website can be found at <https://passionderm.github.io/>.

Keywords: Dermatology · Telemedicine Dataset · Sub-Saharan Africa.

1 Introduction

Access to adequate health care remains a global challenge, largely due to a shortage of health workers [18]. This is particularly acute in Africa, which is the

continent with the lowest number of health workers per patient, with a shortage of up to 2.4 million doctors and nurses [18]. In addition, Africa accounts for 25% of the global demand for health workers, while having only 1.3% of the workforce [18]. This phenomenon spans across all medical specialities. In the case of dermatology, many African countries have less than one dermatologist per million patients [17], making skincare inaccessible to the majority of people. However, the prevalence of skin conditions can be as high as 87% in the paediatric population [14], a situation that naturally extends to the adult population. The consequences of skin diseases are far-reaching and can be physical, psychological and socio-economic, especially in the case of chronic diseases [32]. Teledermatology is a solution that has already been attempted in sub-Saharan Africa [29] to address the lack of access to healthcare. However, it does not increase the number of specialists able to provide primary care.

Artificial Intelligence (AI) holds great promise for scalable diagnosis support, and facilitating access to proper medication and treatment. Although the field of medical imaging has made significant advances in recent years [20], a major shortcoming is that publicly available dermatological datasets are significantly biased towards lighter skin tones [4,2]. This bias persists even in private databases, which primarily include local patients, while images of individuals from diverse geographical origins are conspicuously absent [31,2]. Such a pronounced skew in pigmentation levels does not accurately reflect the diversity of the global population, particularly those without access to traditional dermatologic care. Moreover, common ailments in Africa —namely eczematous dermatitis, fungal infections, bacterial infections, and scabies [21,14]— differ significantly in prevalence from those in Europe [22] and North America [15], which are the usual sources of patient data for public datasets. Efforts were made to address this disparity, but existing datasets are often small [4] or fail to capture the spectrum of prevalent skin conditions in Africa, particularly Neglected Tropical Diseases (NTDs) [4].

This paper aims to fill a critical gap in dermatological research by collecting a comprehensive and inclusive dataset from regions where access to dermatologists is difficult. To achieve this goal, a prospective collection of dermatological cases in sub-Saharan Africa was initiated in 2020. Several countries have collaborated in a global effort to create this dataset, which is intended to serve not only as a valuable resource for training machine learning algorithms and evaluating existing models in these underrepresented populations, but also as a resource for training medical personnel and helping to enable AI-driven teledermatology systems. The dataset consists of patient images, with phototypes IV, V and VI on the Fitzpatrick skin type (FST) scale, acquired in conditions similar to teledermatology consultation, and representative of the most common skin conditions in sub-Saharan Africa. These conditions cover up to 80% of the conditions seen in the paediatric population [25]. During the data collection process, substantial effort was devoted to including the paediatric population, as this is the demographic group most affected by skin disease. This results in a dataset that represents the diversity observed in these countries and effectively covers the dermatological conditions of children. Additionally, we provide

a baseline machine learning model with a corresponding in-depth performance analysis for subpopulations, showcasing evaluation possibilities with this dataset. By releasing the PASSION dataset, we aim to contribute to the advancement of dermatological research and the equity of healthcare solutions.

2 Related Work

Existing datasets predominantly contain cases with lighter skin tones from North American and European populations [4]. In addition, only a minority of datasets include skin type labels, such as the FST scale, or other information related to ethnicity [13,28,3,23,8,16,27]. For example, the Fitzpatrick17k dataset [10] contains 16,577 clinical images annotated with FST labels. However, concerns were raised about the quality of the dataset as it was found to contain many duplicates, irrelevant samples and label errors [11]. Despite its emphasis on skin type labels, it contains predominantly lighter skin types, illustrating the challenges of achieving true diversity. On the other hand, the Diverse Dermatology Images (DDI) dataset [4] explicitly focuses on different skin tones and contains 656 clinical images evenly distributed across FST types.

Other work has explored the use of generative networks [1,24] to generate synthetic images of phototypes IV, V and VI, aiming to increase the diversity of dermatological datasets and improve the robustness of classifiers. However, these efforts face inherent limitations since they may not fully capture clinical realism, potentially introducing inaccuracies. For example, a generative model might produce an image of a melanoma on highly pigmented skin. As a sign of insufficient realism, the image might contain easily visible erythema and a hypopigmented border that does not occur in real life. Conversely, an image of acne on lightly pigmented skin could contain residual hyperpigmentation when in reality this occurs almost exclusively on darker skin types.

The spectrum of skin conditions covered by dermatology datasets illustrates another critical facet of diversity. The majority of existing resources focus narrowly on pigmented lesions, leaving a vast array of non-pigmented skin diseases and rare conditions underrepresented [13,28,3,23,8,16,27]. In addition, most of the images are dermoscopy images, which require a specific tool and some expertise to capture. Such images tend to be highly standardised. This not only limits the scope of research, but also the applicability of the resulting diagnostic models. Initiatives such as Fitzpatrick17k [10], DDI [4], and SKINCON [5] are the only ones encompassing a wider range of skin conditions in European and North American populations, both non-pigmented and rare, in the form of non-standardised clinical images.

3 The PASSION dataset

Data collection. Data were prospectively collected from 2020 to 2023 in Madagascar, Guinea, Malawi, and Tanzania. Madagascar’s cases (850 cases, 2,185 images) were nationwide, Guinean cases (384 cases, 480 images) were in Conakry,

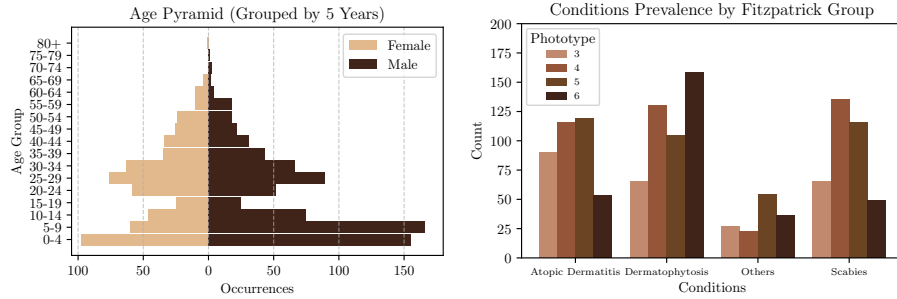


Fig. 1. Age distribution per gender. **Fig. 2.** Conditions per skin phototypes.

Malawi’s cases (261 cases, 2,150 images) were in Lilongwe, and Tanzanian cases (23 case, 86 images) were in Moshi. Malawi and Tanzania data were aggregated as the Eastern African Set (EAS) due to dataset size and geographical proximity. Local board-certified dermatologists acquired the collections, annotated and uploaded the data in each country, either individually or in teams. In addition to images and diagnoses, clinically relevant information such as age, sex and body localisation was collected. Diagnoses were converted to ICD-10 and ICD-11 codes [30], the latest two WHO standards for disease classification, to facilitate integration and interoperability with existing and future datasets. The data collection process reflects a teledermatology setup, as image capture was performed using mobile phones commonly used in the collection areas. Multiple images of the same case were taken to present lesions from different angles to ensure a comprehensive and detailed presentation and to mitigate the risks associated with low-quality images. Diagnoses were made by clinical examination and patient follow-up.

Ethical approval. Data collection was performed with patients’ informed consent in accordance to the respective legislation and ethics guidelines of each contributing country. Analysis of anonymized images was performed according to Swiss ethics statement EKNZ-2018-01074.

Patient privacy. After the collection phase, manual anonymization was conducted to ensure individuals’ privacy, recognizing that clinical dermatology images can include identifiable features. This involved removing identifiable features from the images by inpainting or cropping the images. In addition, depictions of non-disease-involved private body locations were minimized as much as possible.

Data cleaning. In order to ensure the quality and accuracy of the dataset, extensive data cleaning was carried out after acquisition. However, traditional

Table 1. Number of cases per country and body localization.

Conditions	Country			Body Localization								
	Guinea	Madagascar	EAS	Arm	Back	Pelvic	Face	Foot	Hair	Hand	Leg	Torso
Eczema	57	301	56	41	38	11	76	37	24	23	29	72
Fungal	230	299	50	32	33	23	33	35	198	36	41	109
Scabies	58	356	57	13	147	54	1	21	0	76	10	126
Others	39	29	121	14	61	8	41	3	13	1	21	16

data cleaning is very resource intensive as it usually involves checking each sample or pair of samples. Therefore, we resort to a more resource-efficient data cleaning protocol consisting of the use of an algorithmic data cleaning strategy followed by manual verification [9]. We checked for irrelevant samples, near duplicates and label errors. For irrelevant samples and near duplicates, two expert labellers performed manual verification. Similarly, verification for label errors was performed by two board-certified dermatologists with experience practising in sub-Saharan Africa. All experts involved in data curation were not involved in data collection. Annotations were aggregated by unanimous agreement. Verified irrelevant samples were removed, near duplicates were combined by assignment to the same case, and label errors were corrected.

Distributions. Skin phototypes vary between collection centres and are representative of the local population, with a majority of type VI on the FST scale in Guinea. EAS and Madagascar have a fairly similar distribution from type III to type V or VI. The age distribution is shown in fig. 1, with the EAS images including only paediatric cases. The gender distribution of the collection shows a slight tendency towards male cases: Guinea 62%, Madagascar 56% and EAS 65%. The largest difference is in the 5-10 year age group, where the number of male cases is 2 to 3 times higher. The cases collected represent 44 conditions grouped into 4 categories: Eczema, Fungal, Scabies (a skin NTD [7]) and Others. The grouping is based on the ICD-10 ontology. In addition, a binary column “Impetiginized” is added, as impetigo may occur alone or in combination with other conditions. There are a total of 191 (11.5%) impetiginised cases. Recognising its presence is clinically relevant as the prescription of drugs such as corticosteroids for eczema could increase the development of the bacteria by reducing the immune response. The prevalence of conditions is shown in table 1.

Data split. The dataset has a pre-defined 80/20 stratified development-test split to ensure reproducibility and fair comparison of future results. The split is performed at the patient level to prevent information leakage leading to overly optimistic results.

Availability. The dataset can be downloaded at <https://passionderm.github.io/> and is released under the Creative Commons Attribution-NonCommercial License (CC BY-NC).

4 Methodology

Model architecture. We use transfer learning to train a ResNet-50 [12] pre-trained on ImageNet [6]. We replace the last fully connected classification layer with the following sequence of layers: a dropout layer with a 30% chance of dropping, batch normalisation, and a single linear layer generating the class activations. The models are trained using Adam optimisation to minimise the weighted cross-entropy loss. The optimal learning rate was found using the learning rate range test described in [26]. All models are trained by first unfreezing only the classification head for 10 epochs, followed by unfreezing all layers until the end of training. Images are normalized using mean and standard deviation of ImageNet [6]. To improve the generalisation of our models and to avoid overfitting, we perform different data augmentations: random resizing and cropping the images to 224×224 pixels, random horizontal and vertical flipping, and random rotation between 0 and 90 degrees.

Evaluation protocols. For hyperparameter tuning, 5-fold cross-validation is used on the development set, while the test set was kept separate. We evaluate the classifiers with different schemes in order to provide an in-depth analysis.

1. Performance in predicting skin conditions and detecting impetiginised cases on the dedicated hold-out test set.
2. Generalisation from two centres to the wider population available in the dedicated hold-out test set, which contains samples from both the two known centres and one completely unknown centre.
3. Generalisation from different age groups (i.e. paediatric and adolescent cases) to the broader population available in the dedicated hold-out test set, which contains samples from the known age group and a completely unseen one.

All models are trained until the validation loss does not improve over twenty consecutive epochs. Performance is measured in terms of balanced accuracy and macro-averaged precision and sensitivity for detecting skin conditions, and binary F_1 , precision, and sensitivity for detecting impetigo.

Implementation. The implementation is based on PyTorch 1.9 [19]. The experiments are run on an Nvidia DGX station with eight V100 GPUs, each with 32 GB of memory, 512 GB of system memory, and 40 CPU cores. Code to replicate our results is available at <https://passionderm.github.io/>.

Table 2. Hold-out test performance of different models for predicting skin conditions (i.e. Eczema, Fungal, Scabies, and Others) or detecting impetiginized cases (i.e. binary classification).

Model Name	Conditions			Impetigo		
	bal. Acc.	Precision	Sensitivity	F ₁	Precision	Sensitivity
Random Prediction	0.25	0.25	0.25	0.27	0.18	0.56
Constant Prediction	0.25	0.06	0.25	0.28	0.16	1.00
Fine-tuned ImageNet	0.70	0.70	0.70	0.63	0.52	0.79

5 Results

In table 2, we compare the results of a classifier with two uninformed baselines, i.e. random prediction and constant class prediction (the majority class for detecting conditions and the positive class for detecting impetigo). For skin condition prediction, the fine-tuned ImageNet model achieved a balanced accuracy score of 0.70, a precision of 0.70 and a sensitivity of 0.70, significantly outperforming both uninformed baselines. In the specific task of impetigo detection, the fine-tuned ImageNet model achieved a binary F₁ score of 0.63, with precision and sensitivity rates of 0.52 and 0.79, respectively. Models were also compared at subject level by aggregating the prediction for each image of the subject using majority voting. This evaluation is more similar to the task of tele-dermatology. Here, the fine-tuned ImageNet achieves a balanced accuracy of 0.64 for skin disease prediction and binary F₁ of 0.47 for impetigo detection, which is 6% and 16% lower, respectively, than in the sample-level evaluation. This difference highlights the importance of evaluating at both sample and subject-level to get a more complete picture of a classifier’s performance. In table 3, we detail the performance of the classifier for each skin type (i.e. FST III-VI) and gender present in the dataset. The results show that the model has small performance variations across skin types, with a performance gap between the best and worst performing class of 7% (FST IV vs. VI) and 11% (FST IV vs. III) in terms of balanced accuracy for predicting skin conditions and binary F₁ for detecting impetigo, respectively. The performance for female and male cases is only slightly different for both prediction tasks.

To investigate whether a model is able to generalise across collection centres, we compare the performance of a model trained only on the subset of samples from two collection centres with a model trained on all three collection centres on the dedicated test set. To ensure an equal number of samples between the subsets, we randomly resample the training dataset to the same size. Models trained on only two centres achieve a balanced accuracy of 0.61 ± 0.05 , while models trained on all centres achieve 0.68 ± 0.01 . The lowest generalisation was obtained by training only on cases from Guinea and Madagascar, probably due to the predominance of paediatric cases in EAS. The difference in performance between

Table 3. Detailed hold-out test performance for predicting skin conditions or detecting impetiginized cases for a fine-tuned ImageNet model. Performance is shown for each skin type and gender in the dataset separately.

Demography	Support	Conditions			Impetigo		
		bal. Acc.	Precision	Sensitivity	F ₁	Precision	Sensitivity
FST III	303	0.72	0.71	0.72	0.68	0.61	0.77
FST IV	254	0.66	0.66	0.66	0.57	0.47	0.72
FST V	341	0.70	0.69	0.70	0.62	0.49	0.85
FST VI	88	0.73	0.74	0.73	0.62	0.52	0.79
Female	425	0.72	0.72	0.72	0.64	0.52	0.82
Male	561	0.68	0.69	0.68	0.61	0.52	0.76

the two groups highlights the problem of generalising across collection centres and also highlights the importance of collecting data from different sources.

Similarly to the above, we investigate whether a model is able to generalise across age groups, i.e. between paediatric and adolescent cases. Here we compare the performance of a model trained only on paediatric images with a model trained on both paediatric and adolescent cases and vice versa. Resampling is used to ensure an equal number of samples within the groups. Models trained only on paediatric images (i.e. age ≤ 16) achieve a balanced accuracy of 0.66, while models trained on all age groups achieve a score of 0.70, indicating possible generalisation from paediatric images to the wider population. On the other hand, models trained only on adolescent images (i.e. age > 16) achieve a balanced accuracy of 0.43, while models trained on all age groups achieve a score of 0.68. This indicates that generalisation from paediatric cases is higher than from adolescent cases, further highlighting the importance of focusing on the collection of paediatric datasets such as PASSION.

6 Conclusion

This paper presents the PASSION dataset, a pioneering step towards bridging the significant diversity gap in dermatological datasets, specifically targeting pigmented skin types prevalent in sub-Saharan Africa. Our efforts culminate in the compilation of a comprehensive dataset that not only improves the representation of darker skin tones in dermatological studies but also facilitates the development of AI-driven diagnostic tools that are inclusive. In addition, we provide an in-depth analysis of a baseline machine learning model trained and evaluated on PASSION. Future work will focus on expanding the dataset to include a wider range of skin conditions and exploring innovative machine learning algorithms that can further adapt to the challenge of diagnosing skin diseases across a diverse range of skin types.

Acknowledgments. This study was funded by grant REG-19-024 awarded by Fondation Botnar, Basel, Switzerland.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Akrou, M., Gyepesi, B., Holló, P., Poór, A., Kincső, B., Solis, S., Cirone, K., Kawahara, J., Slade, D., Abid, L., Kovács, M., Fazekas, I.: Diffusion-based data augmentation for skin disease classification: Impact across original medical datasets to fully synthetic images (2023)
2. Charrow, A., Di Xia, F., Joyce, C., Mostaghimi, A.: Diversity in dermatology clinical trials: A systematic review. *JAMA Dermatology* (2017)
3. Combalia, M., Codella, N.C., Rotemberg, V., Helba, B., Vilaplana, V., Reiter, O., Carrera, C., Barreiro, A., Halpern, A.C., Puig, S., et al.: Bcn20000: Dermoscopic lesions in the wild. *arXiv:1908.02288* (2019)
4. Daneshjou, R., Vodrahalli, K., Novoa, R.A., Jenkins, M., Liang, W., Rotemberg, V., Ko, J., Swetter, S.M., Bailey, E.E., Gevaert, O., Mukherjee, P., Phung, M., Yekrang, K., Fong, B., Sahasrabudhe, R., Allerup, J.A.C., Okata-Karigane, U., Zou, J., Chiou, A.S.: Disparities in dermatology ai performance on a diverse, curated clinical image set. *Science Advances* (2022)
5. Daneshjou, R., Yuksekgonul, M., Cai, Z.R., Novoa, R., Zou, J.: Skincon: A skin disease dataset densely annotated by domain experts for fine-grained model debugging and analysis (2023)
6. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: ImageNet: A large-scale hierarchical image database. In: *CVPR* (2009)
7. Engelman, D., Marks, M., Steer, A.C., Beshah, A., Biswas, G., Chosidow, O., Cof-feng, L.E., Lardizabal Dofitas, B., Enbiale, W., Fallah, M., Gasimov, E., Hopkins, A., Jacobson, J., Kaldor, J.M., Ly, F., Mackenzie, C.D., McVernon, J., Parnaby, M., Rainima-Qaniuci, M., Sokana, O., Sankara, D., Yotsu, R.R., Yajima, A., Cantey, P.T.: A framework for scabies control. *PLoS Neglected Tropical Diseases* (2021)
8. Giotis, I., Molders, N., Land, S., Biehl, M., Jonkman, M.F., Petkov, N.: Med-node: A computer-assisted melanoma diagnosis system using non-dermoscopic images. *Expert systems with applications* **42**(19), 6578–6585 (2015)
9. Gröger, F., Lionetti, S., Gottfrois, P., Gonzalez-Jimenez, A., Groh, M., Daneshjou, R., Consortium, L., Navarini, A.A., Pouly, M.: Towards reliable dermatology evaluation benchmarks. In: *Proceedings of the 3rd Machine Learning for Health Symposium. Proceedings of Machine Learning Research*, PMLR (2023)
10. Groh, M., Harris, C., Soenksen, L., Lau, F., Han, R., et al.: Evaluating Deep Neural Networks Trained on Clinical Images in Dermatology with the Fitzpatrick 17k Dataset. *IEEE Computer Society* (2021)
11. Gröger, F., Lionetti, S., et al.: SelfClean: A Self-Supervised Data Cleaning Strategy (2023)
12. He, K., Zhang, X., Ren, S., Sun, J.: Deep Residual Learning for Image Recognition (2016)
13. Kawahara, J., Daneshvar, S., Argenziano, G., Hamarneh, G.: Seven-point checklist and skin lesion classification using multitask multimodal neural nets. *IEEE journal of biomedical and health informatics* **23**(2), 538–546 (2018)
14. Kiprono, S.K., Muchunu, J.W., Masenga, J.E.: Skin diseases in pediatric patients attending a tertiary dermatology hospital in northern tanzania: a cross-sectional study. *BMC Dermatology* **15**, 16 (Sep 2015)

15. Laughter, M.R., Maymone, M.B.C., Karimkhani, C., et al.: The burden of skin and subcutaneous diseases in the united states from 1990 to 2017. *JAMA Dermatology* **156**(8), 874–881 (2020)
16. Mendonça, T., Ferreira, P.M., Marques, J.S., Marcal, A.R., Rozeira, J.: Ph 2-a dermoscopic image database for research and benchmarking. In: EMBC (2013)
17. Mosam, A., Todd, G.: Dermatology training in africa: Successes and challenges. *Dermatol Clin* (2021)
18. Naicker, S., Plange-Rhule, J., Tutt, R.C., Eastwood, J.B.: Shortage of healthcare workers in developing countries—afrika. *Ethn Dis* **19**(1 Suppl 1), S1–60–4 (2009)
19. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., et al.: PyTorch: An Imperative Style, High-Performance Deep Learning Library (2019)
20. Pham, T.C., Luong, C.M., Hoang, V.D., Doucet, A.: Ai outperformed every dermatologist in dermoscopic melanoma diagnosis, using an optimized deep-cnn architecture with custom mini-batch logic and loss function. *Scientific Reports* (2021)
21. Ranaivo, I.M., Sendrasoa, F.A., Andrianarison, M., Sata, M., Raharolahy, O., Ralandison, S.D., Ramarozatovo, L.S., Rapelanoro Rabenja, F.: Clinicoepidemiology of skin diseases in children seen at the university hospital center morafeno, toamasina, madagascar. *Dermatology Research and Practice* (Sep 2021)
22. Richard, M.A., Paul, C., Nijsten, T., Gisondi, P., Salavastru, C., Taieb, C., Trakatelli, M., Puig, L., Stratigos, A., EADV Burden of Skin Diseases Project Team: Prevalence of most common skin diseases in europe: a population-based study. *JEADV* **36**(7), 1088–1096 (Jul 2022)
23. Rotemberg, V., Kurtansky, N., Betz-Stablein, B., Caffery, L., Chousakos, E., Codella, N., Combalia, M., Dusza, S., Guitera, P., Gutman, D., et al.: A patient-centric dataset of images and metadata for identifying melanomas using clinical context. *Scientific data* **8**(1), 1–8 (2021)
24. Sagers, L.W., Diao, J.A., Groh, M., Rajpurkar, P., Adamson, A.S., Manrai, A.K.: Improving dermatology classifiers across populations using images generated by large diffusion models. *NeurIPS* (2022)
25. Satimia, F.T., McBride, S.R., Leppard, B.: Prevalence of Skin Disease in Rural Tanzania and Factors Influencing the Choice of Health Care, Modern or Traditional. *Archives of Dermatology* **134**(11), 1363–1366 (11 1998)
26. Smith, L.N.: Cyclical Learning Rates for Training Neural Networks. In: WACV (Apr 2017)
27. Sun, X., Yang, J., Sun, M., Wang, K.: A benchmark for automatic visual classification of clinical skin disease images. In: ECCV. pp. 206–222. Springer (2016)
28. Tschandl, P., Rosendahl, C., Kittler, H.: The ham10000 dataset, a large collection of multi-source dermoscopic images of common pigmented skin lesions. *Scientific data* **5**(1), 1–9 (2018)
29. Weinberg, J., Kaddu, S., Gabler, G., Kovarik, C.: The african teledermatology project: Providing access to dermatologic care and education in sub-saharan africa. *Pan Afr Med J* **3**, 16 (Nov 2009)
30. World Health Organization (WHO): International classification of diseases, eleventh revision (ICD-11). <https://icd.who.int/browse11> (2019/2021)
31. Xie, B., He, X., Zhao, S., Li, Y., Su, J., Zhao, X., et al.: Xiangyaderm: A clinical image dataset of asian race for skin disease aided diagnosis. In: MICCAIW. Springer (2019)
32. Yew, Y.W., Kuan, A.H.Y., Ge, L., Yap, C.W., Heng, B.H.: Psychosocial impact of skin diseases: A population-based study. *PLoS One* **15**(12), e0244765 (Dec 2020)