ARTIFICIAL INTELLIGENCE IN TELEMEDICINE FOR DERMATOLOGICAL CARE

Anonymous authors
Paper under double-blind review

ABSTRACT

In this paper we report our efforts to expand the dermatological knowledge to the refugee camps of the Sahrawi Arab Democratic Republic (SADR) through telemedicine and present our project to evaluate a clinical decision support tool based on state of the art artificial intelligence for the diagnosis of skin lesions in medically undersupplied locations.

1 INTRODUCTION

Telemedicine, which consists of providing patient surveillance through consultations with medical specialists, improves patient access to medical care and reduces waiting time to obtain a consultation (Brinker et al., 2018). Dermatology is particularly suited for this care system as skin disorders are visible to the naked eye. Through teledermatology (TD), diagnoses and treatment recommendations can then be rendered and implemented through the remote analysis of diagnostic images of skin disorders with accompanying clinical histories. The evidence to date supports the efficiency and cost-effectiveness of TD and its ability to improve dermatologic care (Lee & English, 2018). Patients inhabiting in developing countries experience a large proportion of the world's burden of skin disease, which is compounded by minimal public health resources and little or no access to specialists with dermatologic expertise. The World Health Report of 2006 summarises the problem, "Africa has 24% of the World's burden of disease but only 3% of health workers commanding less than 1% of world health expenditure (Guilbert, 2006)." Population growth is surpassing the production of doctors, with the population of Africa forecast to more than double in the next 40 years (Nations et al., 2013). TD has the potential to make a profound impact in regions where expert dermatological care would be unattainable (Tran et al., 2011).

At the same time, artificial intelligence (AI) is rapidly progressing and has the potential to reduce skin cancer-associated mortality, morbidity, and healthcare costs by improving diagnostic accuracy and screening efficiency. Multiple recent studies have demonstrated the ability of AI algorithms to match, if not outperform, clinicians in the diagnosis of individual skin lesion images in reader studies (Tschandl et al., 2019; Haenssle et al., 2018). However, prospective evaluation of an AI algorithm for skin cancer diagnosis has not yet been studied. To validate the real-world utility of medical AI systems, prospective clinical studies that evaluate the systems' performance in clinical settings are needed (Yu & Kohane, 2019).

In this project, we aim to validate an AI-enabled diagnostic tool in TD for skin cancer diagnosis and determine its impact for the first time in the clinician decision making for the diagnosis of skin lesions. We will study the feasibility of the usage of artificial intelligence for dermatological care in the Sahrawi Arab Democratic Republic (SADR) refugee camps.

2 METHODS AND PRELIMINARY RESULTS

In 2017, our hospital signed a collaboration agreement with the Ministry of Health (MS) of the Sahrawi Arab Democratic Republic (SADR) and Medicus Mundi Mediterrània (MMM) to support health issues in the population living in refugee camps. One of the projects was the deployment of a TD consultation system based on a portable software application. This software allowed the evaluation of skin lesions and the creation of a photographic archive based on the most prevalent dermatological diseases in the area to help diagnose future cases and treat them accordingly to the most efficient therapeutic guidelines.

In May 2020, four pharmacology and dermatology specialists will once again relocate to the SADR camps to support their population and doctors. During this time, a clinical decision support tool based on AI will be incorporated into the TD software. The algorithm is based on the state of the art technology and will analyze the images from the lesions to provide a diagnostic prediction to our specialists.

The neural network is based on an ensemble of EfficientNets architecture (Tan & Le, 2019) and has been trained on the HAM10000 dataset (Tschandl et al., 2018), which contains over 10000 dermoscopic images of skin lesions from seven different categories. The model has been cross-validated using five splits, with a mean balanced accuracy of 0.845. The retrospective balanced accuracy of the model has been scored using the live leaderboard from the ISIC 2018 Challenge (Codella et al., 2019; isi), obtaining the 20th best worldwide performance. The performance metrics for the test split (from the live leaderboard) are reported on table 1.

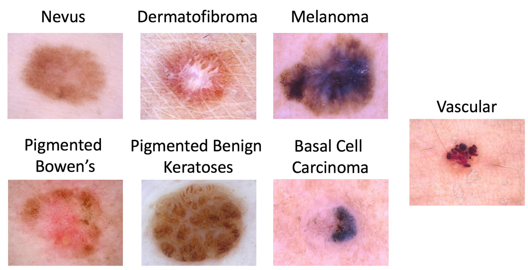

Figure 1: Example image for each one of the diagnosis categories in the HAM10000 dataset (Tschandl et al., 2018).

Table 1: Performance of the neural network which will be prospectively studied as decision support system during the 2020 expedition, for each diagnostic category: MEL (melanoma), NEV (nevus), BCC (Basal Cell Carcinoma), AK (Actinic Keratosis), BK (Benign Keratosis), DF (Dermatofibroma) and VASC (Vascular Lesion)

Metric	**Mean Value**	**MEL**	**NEV**	**BCC**	**AK**	**BK**	**DF**	**VASC**
Accuracy	0.962	0.931	0.916	0.979	0.979	0.947	0.992	0.993
Sensitivity	0.827	0.789	0.909	0.860	0.767	0.802	0.773	0.886
Specificity	0.973	0.949	0.927	0.987	0.986	0.971	0.999	0.995

The predictions of the clinical decision support tool for the diagnosis of skin cancer will be compared to the clinical diagnosis provided by the specialists in the refugee camp. The performance of the algorithm will be reported in terms of diagnostic accuracy. This work will represent the first feasibility study of an AI-enabled clinical decision support tool for the diagnosis of skin cancer for medically undersupplied locations.

REFERENCES

Isic challenge live leaderboard. https://challenge2018.isic-archive.com/live-leaderboards/. Accessed: 2020-02-21.

Titus Josef Brinker, Achim Hekler, Christof von Kalle, Dirk Schadendorf, Stefan Esser, Carola Berking, Martina T Zacher, Wiebke Sondermann, Niels Grabe, Theresa Steeb, et al. Teledermatology: comparison of store-and-forward versus live interactive video conferencing. *Journal of medical Internet research*, 20(10):e11871, 2018.

Noel Codella, Veronica Rotemberg, Philipp Tschandl, M Emre Celebi, Stephen Dusza, David Gutman, Brian Helba, Aadi Kalloo, Konstantinos Liopyris, Michael Marchetti, et al. Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic). *arXiv preprint arXiv:1902.03368*, 2019.

J Guilbert. The world health report 2006 1: Working together for health 2. *Education for health*, 19 (3):385–387, 2006.

Holger A Haenssle, Christine Fink, R Schneiderbauer, Ferdinand Toberer, Timo Buhl, A Blum, A Kalloo, A Ben Hadj Hassen, Luc Thomas, A Enk, et al. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. *Annals of Oncology*, 29(8):1836–1842, 2018.

Jonathan J Lee and Joseph C English. Teledermatology: a review and update. *American journal of clinical dermatology*, 19(2):253–260, 2018.

United Nations, Department Affairs, and Population Division. *World population prospects: the 2012 revision*. 01 2013.

Mingxing Tan and Quoc V Le. Efficientnet: Rethinking model scaling for convolutional neural networks. *arXiv preprint arXiv:1905.11946*, 2019.

Kathleen Tran, Mohamed Ayad, Jennifer Weinberg, Augustin Cherng, Mridul Chowdhury, Saadeddin Monir, Mohamed El Hariri, and Carrie Kovarik. Mobile teledermatology in the developing world: implications of a feasibility study on 30 egyptian patients with common skin diseases. *Journal of the American Academy of Dermatology*, 64(2):302–309, 2011.

Philipp Tschandl, Cliff Rosendahl, and Harald Kittler. The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific data*, 5: 180161, 2018.

Philipp Tschandl, Noel Codella, Bengü Nisa Akay, Giuseppe Argenziano, Ralph P Braun, Horacio Cabo, David Gutman, Allan Halpern, Brian Helba, Rainer Hofmann-Wellenhof, et al. Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study. *The Lancet Oncology*, 20(7):938–947, 2019.

Kun-Hsing Yu and Isaac S Kohane. Framing the challenges of artificial intelligence in medicine. *BMJ Qual Saf*, 28(3):238–241, 2019.