ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11,Iss 6, Sep 2022

Machine Learning in Dermatology: Current Applications, Opportunities and Limitations

Dr. V.K. Garg[1]*, Dr. Gulshant Panesar[2], Dr. Sonali Saxena[3]

1. Professor, Department of Dermatology, Venereology & Leprosy, Santosh Medical College & Hospital, Santosh Deemed to be University, Ghaziabad.
2. Assistant Professor, Department of Dermatology, Venereology & Leprosy, Santosh Medical College & Hospital, Santosh Deemed to be University, Ghaziabad.
3. Senior Resident, Department of Dermatology, Venereology & Leprosy, Santosh Medical College & Hospital, Santosh Deemed to be University, Ghaziabad.

***Dr. V.K. Garg – Corresponding Author**

ABSTRACT

The practise of dermatologists, from diagnosis to individualised care, has the potential to be enhanced by machine learning (ML). Recent developments in faster computers, less expensive data storage, and access to big datasets (such as electronic medical records, picture databases, and omics) have stimulated the development of ML algorithms with human-like intelligence in dermatology. This article provides an overview of machine learning fundamentals, current uses, potential drawbacks, and ideas to keep in mind as machine learning technology develops. The five current areas of use for ML in dermatology that we have identified are: (1) disease classification using clinical images; (2) disease classification using dermatopathology images; (3) assessment of skin diseases using mobile applications and personal monitoring devices; (4) facilitation of large-scale epidemiology research; and (5) precision medicine. To assist dermatologists better assess the possible benefits and drawbacks of machine learning, this overview aims to demystify the foundations of ML and its vast variety of applications.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11,Iss 6, Sep 2022

Keywords: Artificial intelligence; Convolutional neural network; Deep learning; Dermatology; Image classification; Machine learning; Mobile applications; Personal monitoring devices; Precision medicine

INTRODUCTION

The abundance of data in clinical records, patient demographic data, imaging examination results, and survey data in dermatology and medicine as a whole represents a wealth of data that has the potential to transform customised therapy [1]. With data from the genome, epigenome, transcriptome, proteome, and microbiome-areas of study frequently referred to by the abbreviation "omics"—translational research in dermatology is already abounding [2]. Dermatology has many uses for machine learning (ML) algorithms with human-like intelligence that have recently been made possible by improvements in quicker processing and less expensive storage [3-5]. Dermatologists must have a fundamental understanding of AI and ML in order to evaluate the efficacy of these new technologies. In this review, we first give a general introduction of ML, AI, and the creation of algorithms.We then look at the existing ML applications that dermatologists should be aware of.

Finally, we consider potential obstacles and restrictions to ML's further progress. This review serves as a resource for dermatologists to assist in demystifying the foundations of machine learning (ML) and its broad variety of applications in order to properly assess its possible benefits and drawbacks.

METHODS AND MATERIALS

This review is based on a search of the Medline, Embase, and Web of Science databases for journal publications that discuss the use of ML and AI in dermatology. In December 2019, the investigation was carried out. To concentrate on new approaches, articles from 2000 to 2019 were incorporated. Repetitive articles were removed, and only English-language articles were included. The main search terms were "artificial intelligence," "machine learning," and "dermatology." Following a review of the preliminary findings, the secondary keywords "personalised medicine," "teledermatology," "smartphone apps," and others were combined

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11,Iss 6, Sep 2022

using the Boolean operators "AND" and "OR." cancer of the skin nonmelanocytic skin cancer, melanocytic skin cancer, psoriasis, atopic dermatitis, and dermatopathology. A total of 889 articles were found in our literature search, of which 70 were deemed pertinent to this review.

This article is based on studies that have already been done; none of the authors have undertaken any experiments involving humans or animals.

Overview of artificial intelligence and machine learning

Artificial intelligence, a subfield of computer science, simulates intelligent human behaviour using machines and programmes. The 1950s saw the beginning of artificial intelligence thanks to Alan Turing's inquiry, "Can computers think?" [6]. By the 1970s, algorithms with clear guidelines for how computers should interpret data had been developed by software developers. However, it was difficult to translate the heuristics of human decision-making in medicine into concrete principles.

To accomplish the objectives of artificial intelligence, ML is a technique that is a subset of artificial intelligence . Recently, ML has drawn attention for the wide range of applications it has in daily life, from individualised online news and video suggestions to self-driving cars.

Deep learning, logistic regression, random forests, and other statistical techniques are all included in the field of machine learning (ML). Although machine learning (ML) can initially appear mysterious, it is often closely similar to conventional statistical models that are familiar to most dermatologists.

Machine Learning Approaches

Supervised learning, unsupervised learning, and reinforcement learning are the three basic categories into which machine learning approaches can be separated [7]. A dataset must be provided as inputs (referred to as features) and outputs (referred to as labels) for supervised learning [7]. For instance, in an algorithm that determines whether a pigmented lesion is benign or malignant, pictures of pigmented lesions are "features" and categorical data indicating whether they are malignant or benign are "labels." The system is first taught using labelled photos of

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11,Iss 6, Sep 2022

benign pigmented lesions and melanoma, after which the computer generalises this knowledge to a fresh collection of skin images. The most typical method of learning utilised in dermatology is supervised learning.

Unsupervised learning, on the other hand, simply needs inputs (unlabeled data), and this method can spot undiscovered groupings or anomalies in data [7]. Reinforcement learning, a cross between supervised and unsupervised learning, acquires knowledge through trial and error as well as environmental input [7]. The AlphaGo algorithm is a prime example of reinforcement learning [8]. Dermatology has not yet studied reinforcement learning.

Machine Learning Algorithms

Different ML algorithms are frequently employed in dermatology. The majority of machine learning (ML) algorithms are examples of statistical learning; for instance, some of the most popular statistical learning techniques are natural language processing, support vector machines (SVM), k-nearest neighbour (k-NN), logistic regression, and linear regression (NLP). Based on the presence or absence of k neighbours, k-NN is utilised for data classification and regression [9, 10]. Data is classified using SVMs by locating a hyperplane that distinguishes between groups [11]. To discover the most frequent result among all the randomly created decision trees, RFs create a network of random decision trees [12]. Large amounts of text are analysed by NLP in order to find patterns [13].

Neural Networks and Deep Learning

Statistical and mathematical models are used in deep learning, a subset of machine learning, to simulate how neurons process information. The foundation of neural networks (NNs), also known as artificial neural networks (ANNs), is a group of interconnected elements, such as nodes, neurons, or process layers [14]. The network of neurons in the human brain serves as an inspiration for ANNs. The ANN's neurons, or nodes, are arranged into linear arrays known as layers [14].

Every node gets input through connections with associated weights [14]. A typical ANN has input layers, output layers, and hidden layers, and choosing the number of nodes in each layer,

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11,Iss 6, Sep 2022

the number of layers in the network, and the path of connections between nodes are all part of creating an ANN [14]. Through a learning process, ANNs are taught to carry out particular tasks, including categorization. Both supervised and unsupervised learning are possible in ANNs, however supervised learning is more typical. A training set for supervised learning includes examples of input targets and output targets [14]. The weights of the inputs are changed while the ANN is trained to reduce the error between the network output and the desired output [14].

Convolutional NNs (CNNs) are a unique subclass of ANNs that include one or more convolutional unit layers (also known as pooling units). Two-dimensional or three-dimensional inputs are fed into CNNs after being processed through numerous hidden layers. A group of pixels known as a motif or a collection of motifs can be used to analyse an image. Each component of an input image is compared against a small sub-image in the first few levels of the CNN [5]. Convolution is used to determine how closely an area of the image resembles each of the features allocated to each node (such as colour, shape, size, etc.) before the node outputs to the next layer [5]. Following these convolutional layers, pooling layers—a type of NN—classify the entire image [5].

The first time CNNs were used to classify medical images was at the illustrious 2012 ImageNet Large Scale Visual Recognition (ILSVRC) conference.The percentage of photos for which the right class was not among the top five predicted classes was 15.3% for a CNN named AlexNet, which was trained to classify 1.2 million images into 1000 different categories [15]. For the first time, CNN had such a low error rate. Tens of thousands of photos were all that were present in earlier image databases [16–18]. At the 2016 International Symposium on Biomedical Imaging, CNNs were utilised to classify medical pictures in all approaches [19].

Metrics for Assessment and Validation of Machine Learning Models

The number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) from an ML prediction is used as one of several measures used to evaluate machine learning models. These metrics include accuracy [(TP? TN)/(TP? TN? FP? FN)], sensitivity [TP/(TP? FN)], specificity [TN/(TN? FP), positive predictive value [TP/(TP? FP),

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11,Iss 6, Sep 2022

negative predictive value [TN/(TN? FN)], and others. Additionally, the area under the receiver operating characteristic (AUCROC) is used to evaluate ML models (AUROC or AUC). The sensitivity versus 1-specificity plot is used to create the receiver operating characteristic (ROC) curve [20]. The algorithm performs better the further the ROC curve deviates from the diagonal and the greater the AUROC. Perfect classification is indicated by an AUROC of 1.0, and classification that is only slightly better than random chance is shown by an AUROC of 0.5.

Algorithms run the danger of overfitting the data if they are trained and tested using the same dataset or set of images. However, the photos that are now available are limited, and research may train and validate their models using the same collection of images. Overfitting can be avoided by validating or cross-validating datasets' predicted accuracies. Underfitting, on the other hand, happens when an ML model is unable to accurately reflect the training dataset and is unable to generalise to new datasets [7]. In ML models, overfitting is more frequent than underfitting.

DERMATOLOGY AND MACHINE LEARNING

The dermatologist's office offers numerous promising ML opportunities. The ability of CNN's picture categorization to expand access to skin cancer screenings and improve dermatologists' workflow has attracted the most interest.

In numerous other disciplines, including ophthalmology, pathology, and radiology, CNN has already demonstrated its viability. After being validated in a critical prospective clinical trial involving more than 900 patients, the US Food and Drug Administration (FDA) approved a CNN in the IDx-DR diagnostic system (IDx Technologies Inc., Coralville, IA, USA) [22] for independent use in diabetic retinopathy screening in 2018. This diagnostic approach was compared to an independent, superior gold standard of imaging techniques and attained a sensitivity of 87.2% and a specificity of 90.7% [23]. However, prospective clinical trials have not validated CNN for the screening of melanoma or non-melanoma skin cancers, and it still has significant limitations. Beyond the straightforward task of telling benign lesions from malignant ones, ML can be used in mobile applications, dermatopathology, differential diagnosis,

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11,Iss 6, Sep 2022

teledermatology, and personalised medicine. Regulating and enhancing these technologies for patients requires a thorough understanding of the many deep learning techniques now in use and how they are developing.

In research using ML in dermatology, the majority of attention is given to categorising skin lesions for a range of conditions, such as melanoma, non-melanoma skin cancer (NMSC), psoriasis, atopic dermatitis, onychomycosis, and rosacea.

For the purposes of image recognition and classification, these research predominantly use CNN. In the beginning, only features were extracted using a pretrained CNN (AlexNet), and these features were subsequently classified using a more straightforward ML method, like k-nearest neighbours or SVMs [24–26].Currently, most CNNs can identify images and extract features using end-to-end learning.

Psoriasis

Despite the substantial clinical trial evidence on the effectiveness of the 11 biologics that have received FDA approval, selecting a biologic for a patient still relies on trial and error. A clinical response typically takes 12 to 16 weeks to become significant, and the drug's effectiveness can range between 30 and 80% [13]. This results in a "assessment gap" between the time when a patient's reaction to a treatment can be clinically established and the time when that response may be partially determined by biology. A faster response to treatment can be seen in changes in biopsies or gene expression patterns than in clinical improvement. Therefore, by forecasting the long-term effects of biologics in psoriasis patients, ML can close this assessment gap.

To ascertain the long-term treatment response to biologics, several research have developed ML prediction models [13-16]. Two ML models were developed in the initial investigation to evaluate this gap in order to analyse the gene expression data from skin biopsies [13]. By analysing the molecular profile of the short-term (2-4 weeks) treatment, the models were able to accurately predict the PASI 75 response, or a C 75% increase in PASI score from baseline, following 12 weeks of treatment. Both of these models significantly reduced the psoriasis assessment gap by 2 months and accurately predicted the PASI 75 response (AUC[0.80]). An

ISSN PRINT 2319 1775 Online 2320 7876

Research paper **© 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11,Iss 6, Sep 2022**

approach that was previously regarded as unreliable may now be used to predict a patient's response to biologic therapy by analysing baseline samples of gene expression from skin biopsies [14]. The major histocompatibility complex (MHC) and tumour necrosis factor (TNF) signalling genes and pathways were shown to show signs of therapy response in a different study using multi-omics to analyse individuals with severe psoriasis on etanercept [16].

Skin Cancer

Melanoma [11] and NMSC [102] cancer risk models have been created using machine learning. Because it included information on more than four million dermatology patients in the USA from a cloud-based dermatology-specific EMR called Modernizing Analytics for Melanoma [11], the dataset utilised to develop the risk model for melanoma was distinctive. Given the amount of the data, the authors adopted a hybrid approach combining non-distributed computing (using just one computer) and distributed computing (using numerous machines to maximise efficiency). Data collection and formatting were done using distributed computing, whereas machine learning was done using non-distributed computing. A study conducted by Roffman et al. in 2018 [21] is a nice illustration of how ML and large data can be utilised to evaluate a novel idea.

These authors sought to develop an ANN to predict personal NMSC risk solely based on 13 parameters of personal health data: gender, age, basal metabolic index, diabetic status, smoking status, emphysema, asthma, Hispanic ethnicity, hypertension, heart diseases, vigorous exercise habits, and history of stroke. While ultraviolet radiation exposure and family history are important associated risk factors for NMSC, they were not the authors' primary focus.

DISCUSSION

We have described the fundamental concepts of ML and some of the current dermatology applications in this paper. We have seen that ML has a wide range of possible uses in the workflow of a dermatologist, from diagnosis to treatment.

We have identified five areas of ML applications in dermatology based on the literature.The emergence of potent deep learning algorithms that can examine big datasets has aided all five fields. In the first application, dermoscopic, DSLR, and smartphone photos can be used to

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11,Iss 6, Sep 2022

identify melanoma, NMSC, and other dermatological illnesses. The most thoroughly researched ML algorithms in dermatology are those for melanoma and NMSC classification, which may help to address poor skin cancer screening rates by facilitating access. Other illness classification is still in its infancy, and it will probably entail more complicated algorithms to evaluate disease severity and generate precise differential diagnoses. The importance of including dermatologists and dermatopathologists in the development of ML studies has been highlighted by deep learning to classify histopathology, a field that is still in its early stages. Dermatopathologists' image curation improves the performance of the models in this second application area. The fact that the majority of these research weren't published with dermatologists named as authors leads one to believe that more dermatologists ought to be engaged in the creation of these ML models. The third sector involves unregulated mobile applications that are currently too inaccurate or sensitive to be effective screening tools for lesions and skin cancer. Mobile applications can make these screening instruments available to populations in need once categorization algorithms are more accurate and have undergone proper clinical validation. In the fourth category, ML may examine big datasets that are helpful for epidemiological investigations, such as EMR data and insurance claims. Finally, ML has the potential to aid in the diagnosis of individuals with psoriasis, psoriatic arthritis, and skin cancer as well as predict how well they will respond to treatment.

CONCLUSIONS

Dermatology has a lot of potential for machine learning, from diagnosis to forecasting safer and more successful therapies.Dermatologists will need to learn how ML functions and when and how it should be utilised correctly in a clinical environment as this technology develops. Although ML techniques are strong, they are nevertheless similar to earlier clinical tools in that a doctor's interpretation is essential for their application in a practical situation. Additionally, we need to be aware of how these algorithms' inherent "black box" character may be hampered by potential biases. The inclusion of people with skin of colour in these technologies is equally crucial. By making algorithms and datasets available to the public for additional validation and testing, future ML research should be transparent. Thorough prospective clinical trials that have

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11,Iss 6, Sep 2022

been peer-reviewed should be carried out before to going on the market. Overall, increasing the participation of dermatologists in the design and testing of ML is essential for producing technology that is both practical and clinically applicable.

ACKNOWLEDGEMENTS

REFERENCES

1. Marka A, Carter JB, Toto E, Hassanpour S. Automated detection of nonmelanoma skin cancer using digital images: a systematic review. BMC Med Imaging. 2019;19:21.

2. Thomsen K, Iversen L, Titlestad TL, Winther O. Systematic review of machine learning for diagnosis and prognosis in dermatology. J Dermatol Treat. 2019;31:1–15. https://doi.org/10.1080/09546634. 2019.1682500.

3. Nichols JA, Herbert Chan HW, Baker MAB. Machine learning: applications of artificial intelligence to imaging and diagnosis. Biophys Rev. 2018;11:111–8.

4. Turing AM. Computing machinery and intelligence. Mind. Dordrecht: Springer; 1950.

5. Murphy KP. Machine learning: a probabilistic perspective. Cambridge: MIT Press; 2012.

6. Silver D, Huang A, Maddison CJ, et al. Mastering the game of Go with deep neural networks and tree search. Nature. 2016;529:484–9.

7. Coomans D, Massart DL. Alternative k-nearest neighbour rules in supervised pattern recognition:part 1 k-Nearest neighbour classification by using alternative voting rules. Anal Chim Acta. 1982;136:15–27.

11. Altman NS. An introduction to kernel and nearestneighbor nonparametric regression. Am Stat.1992;46:175–85.

12. Hearst MA, Dumais ST, Osuna E, Platt J, Scholkopf B. Support vector machines. IEEE Intell Syst Appl.1998;13:18–28.

13. Breiman L. Random forests. Mach Learn. 2001;45: 5–32.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper **© 2012 IJFANS. All Rights Reserved,** **UGC CARE Listed (Group -I) Journal Volume 11,Iss 6, Sep 2022**

14. Manning CD, Schu¨tze H. Foundations of statistical natural language processing. Cambridge: MIT Press; 1999.

15. Zou J, Han Y, So S-S. Overview of artificial neural networks. In: Livingstone DJ, editor. Artificial neural networks: methods and applications. Totowa: Humana Press; 2009. https://doi.org/10.1007/978-1-60327-101-1_2.

16. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks.Commun ACM. 2017;60:84–90.

17. Fei-Fei L, Fergus R, Perona P. Learning generative visual models from few training examples: an incremental Bayesian approach tested on 101 object categories. In: 2004 Conference on Computer Vision and Pattern Recognition Workshop. Washington, DC. 2004, p. 178. https://ieeexplore. ieee.org/document/1384978. Accessed 18 Dec 2019.

18. Griffin G, Holub A, Perona P. Caltech-256 Object Category Dataset. 2007. https://resolver.caltech.edu/CaltechAUTHORS:CNS-TR-2007-001. Accessed 20 Feb 2020.

19. LeCun Y, Fu Jie Huang, Bottou L. Learning methods for generic object recognition with invariance to pose and lighting. In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC. 2004, p. 2–104. https://ieeexplore.ieee.org/document/ 1315150. Accessed 18 Dec 2019.

20. Marchetti MA, Codella NCF, Dusza SW, et al. Results of the 2016 International Skin Imaging Collaboration International Symposium on Biomedical Imaging challenge: Comparison of the accuracy of computer algorithms to dermatologists for the diagnosis of melanoma from dermoscopic images. J Am Acad Dermatol. 2018;78(270–277):e1.

21. Abra`moff MD, Lou Y, Erginay A, et al. Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. Invest Ophthalmol Vis Sci. 2016;57: 5200–6.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11,Iss 6, Sep 2022

22. Abra`moff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. npj Digit Med. 2018;1:1–8.

23. Dascalu A, David EO. Skin cancer detection by deep learning and sound analysis algorithms: a prospective clinical study of an elementary dermoscope. EBioMedicine. 2019;43:107–13.

24. Chen SC, Bravata DM, Weil E, Olkin I. A comparison of dermatologists' and primary care physicians'accuracy in diagnosing melanoma: a systematic review. Arch Dermatol. 2001;137:1627–34.

25. Brinker TJ, Hekler A, Enk AH, et al. Deep learning outperformed 136 of 157 dermatologists in a headto- head dermoscopic melanoma image classification task. Eur J Cancer. 2019;113:47–544.

26. Brinker TJ, Hekler A, Enk AH, et al. Deep neural networks are superior to dermatologists in melanoma image classification. Eur J Cancer. 2019;119: 11–7.