Deep learning and rules-based hybrid approach to improve the accuracy of early detection of skin cancer

Varin Senthil[1], Vanju Shreyaa[2], and Venkateshwaran Kothandapany[2]

[1]Affiliation not available
[2]MD Dermatology

December 29, 2022

Deep learning and rules-based hybrid approach to improve the accuracy of early detection of skin cancer.

- *Dr. Shreyaa Vanju Venkatesh MD Dermatology, Varin Senthil, Troy Highschool, Michigan*

Abstract:

Skin cancer is one of the most dangerous diseases in the world. As per the American Cancer Society, about 99,780 new melanomas will be diagnosed in the United States and about 7,650 people are expected to die of melanoma. Correctly classifying skin lesions at an early stage could aid clinical decision-making by providing an accurate disease diagnosis, potentially increasing the chances of cure before cancer spreads. Artificial intelligence (AI) has wide applications in healthcare, including dermatology. Machine learning (ML) is a subfield of AI involving statistical models and algorithms that can progressively learn from data to predict the characteristics of new samples and perform a desired task. AI can be of use for the early detection of skin cancer. For example, the use of deep convolutional neural networks can help to develop a system to evaluate images of the skin to diagnose skin cancer. However, there is likelihood of errors and it depends on a lot of factors, including the amount and quality of the data used to train the algorithms, the environment in which machine learning operates may itself evolve or differ from what the algorithms were developed to face, the complexity of the overall systems it's embedded in. While, in the deep learning approach the knowledge is developed based on sample picture, there are rule based traditional approach which has a business logic built in the application. Business logics could be based on teat results, skin colors, etc. This paper presents a comprehensive hybrid approach combining the AI methodology along with the color pigment analysis to reduce the errors and improve the accuracy and a discussion on how it can be implemented in the field of diagnosing skin cancer.

We reviewed the latest research and key discoveries in ML encompassing various subfields of dermatology related cancers. Literature review was performed to screen the articles published in "Skincancer.org, mdpi.com, iopscience.iop.org, hbr.org, medium.com, PubMed and Google Scholar" through August 2022. The search words included "Artificial intelligence AND skin cancer" "Machine learning AND skin cancer" and "Deep learning AND skin cancer".-Relevant references of the screened articles were also included for qualitative analysis. Important websites related to skin cancer and related AI resources were also browsed to gather information on the topic.

Types of skin cancer:

Skin cancer is the out-of-control growth of abnormal cells in the epidermis, the outermost skin layer, caused by unrepaired DNA damage that triggers mutations. These mutations lead the skin cells to multiply rapidly and form malignant tumors. The main types of skin cancer are basal cell carcinoma (BCC), squamous cell carcinoma (SCC), melanoma and Merkel cell carcinoma (MCC). The two main causes of skin cancer are the sun's harmful ultraviolet (UV) rays and the use of UV tanning beds. Melanoma is the most dangerous of the three most common forms of skin cancer. Melanoma is a cancer

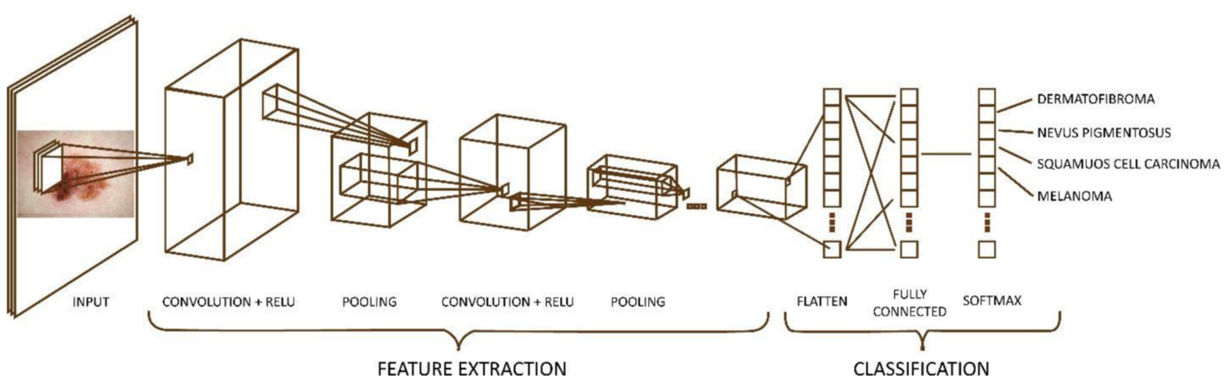
that develops from melanocytes, the skin cells that produce melanin pigment, which gives skin its color. Melanomas often resemble moles and sometimes may arise from them. They can appear on any area of the body, even in areas that are not typically exposed to the sun. Melanoma is often triggered by the kind of intense, intermittent sun exposure that leads to sunburn. Tanning bed use also increases risk for melanoma. In 2022, an estimated 197,700 new cases of melanoma are expected to occur in the U.S. Of those, 97,920 cases will be in situ (noninvasive), confined to the epidermis

(the top layer of skin), and 99,780 cases will be invasive, penetrating the epidermis into the skin's second layer (the dermis).
Melanomas can be curable when caught and treated early. In 2022, melanoma is projected to cause about 7,650 deaths [1].

Machine learning in Cancer diagnostics:
[2] Skin tumors can be difficult to recognize from common benign skin lesions, and melanoma has a particularly varied look. AI can aid in the early detection of skin cancer, lowering the burden of morbidity and mortality associated with the disease [3]. In addition to reducing the workload, AI-based systems can also help by improving skin lesion diagnostics [4,5].

Artificial intelligence (AI), a branch of computer science that uses machines and programs to mimic intelligent human behavior via a constellation of technologies, is a key driver of the fourth industrial revolution. Machine learning (ML) is a division of AI where computer systems learn from their experiences without having to be explicitly programmed. A supervised, semi-supervised, or unsupervised process can be used. The machine is fed datasets of problems and answers in a supervised configuration. Through trial and error, machines learn to select the correct response. Machines analyze incoming data with no predetermined solution in unsupervised learning. Semi-supervised learning is a method that uses both labelled and unlabeled data [6]. Deep learning is a category of ML that recruits multiple layered deep neural networks, each of which can recognize and learn distinct features particular to the dataset [7]. An artificial neuron network (ANN) is a computational model based on the structure and operations of biological neural networks [8]. The most basic type of ANN is the feedforward neural network. It has three layers, namely input, hidden, and output, where data goes in via the input layer, crosses the hidden layer, and comes out via the output nodes. Multiple hidden levels are possible [9,10]. Convolutional Neural Network is a variant of deep, feed-forward ANN that is most typically used to analyze visual imagery. It is made up of convolutional as well as pooling layers that allow the network to encode picture characteristics [11].

[12] The Architecture of CNN Convolutional Neural Network (CNN) is a development of the Multilayer Perceptron (MLP) which is designed to process two-dimensional data. CNN is included in the type of Deep Neural Network because it has a high network depth and has been widely applied to image data [13]. CNN has an architecture as like as neural networks in general, neurons in CNN have a weight, bias, and activation function. CNN architecture as shown in



Figure, which consists of the convolution layer with ReLU activation, pooling layer as feature extraction layer, and fully connected layer with softmax activation as classification layer. In the Convolution layer, the convolution process is the main process that underlies CNN. Convolution layer is the first layer that

will process the image as an input system model. The image will be convoluted with a filter to extract features from the input image that is called the feature map.
ReLU (Rectified Linear Unit) is an activation layer in CNN to increase the training stage on neural networks that have advantages to minimize errors. Rel-U activation makes all pixel values to be zero when a pixel image has a value of less than zero [14]. Rel-U Polling layers in the CNN method usually will be inserted regularly after several convolution layers, There are several advantages of the pooling layer, which can progressively reduce the size of the output volume on the Feature Map so that it can control over-fitting [13]. Pooling Layer is used to reduce data using max-pooling or mean Pooling. The max-pooling will select the maximum value, whereas the main pooling finds the average value.
The Fully-Connected Layer is the layer at the end of the architecture used in the multilayer perceptron. This layer will connect all the neurons of the previous activation layer. In this stage, all neurons in the input layer need to be transformed into one-dimensional data (flatten process) [15]. After that softmax activation as another form of the logistic regression algorithm can be used to classify more than two classes.
Hyperparameter has variable values that remain during the model training process and can affect the performance of the model trains. In this study, the hyperparameter that used is an optimizer, such as Stochastic Gradient Descent (SGD), Root Mean Square Propagation (RMSprop), Adaptive Moment Estimation (Adam), and Nesterov accelerated Adaptive Moment Estimation (Nadam). Stochastic Gradient Descent (SGD) is a repetitive optimization method that functions to optimize the model using better functions such as differential or subdifferential [16]. SGD uses each training sample as a new parameter. Root Mean Square Propagation (RMSprop) is widely used in the design of deep learning models [17]. This optimizer is an improvisation from Root Propagation (Rprop). Initially, Rprop cannot be used on files with large amounts of data. The essence of RMSprop is moving the average gradient at the time of the model. Adam optimizer is a combination of RMSprop and momentum. This optimizer also uses an average gradient of weight [18]. The advantage of Adam over other optimizers is efficient in computing time, consume less memory, and can handle sparse gradients on noisy problems. Nadam (Nesterov-accelerated Adaptive Moment Estimation) thus combines Adam and NAG (Nesterov accelerated gradient).
[2] AI is poised to bring transformation in healthcare because of its advantages over traditional analytical techniques. There is rising optimism regarding applications of AI in healthcare, ranging from assistance in medical diagnostics, treatment and administrative support to reduce timelines of new drug development. It may also be of benefit as an adjuvant in clinical decision making [19]. Dermatology, as a visually intensive field, is at the precipice of an AI revolution. The association for the advancement of AI defines it as “the scientific knowledge of the mechanisms underlying mind and intelligent behavior and its implementation in machines” [20]. AI has the potential to exceed humans, due to its endless processing power and storage capacity [21]. Because skin disease diagnosis is mostly based on visual perception, computer vision algorithms may be able to recognize skin lesions based on their morphology.
By September 2018, the US Food and Drug Administration (FDA) had authorized AI approaches for clinical usage, including devices to detect skin cancer from clinical photos obtained via a smartphone app [22].

Limitations in machine learning based diagnostics:
[23] The big difference between machine learning and the digital technologies that preceded it is the ability to independently make increasingly complex decisions and continuously adapt in response to new data. But these algorithms don’t always work smoothly. They don’t always make ethical or accurate choices. There are three fundamental reasons for this.

One is simply that the algorithms typically rely on the probability that someone will have a disease. The likelihood of errors depends on a lot of factors, including the amount and quality of the data used to train the algorithms, the specific type of machine-learning method chosen, and whether the system uses only explainable algorithms (meaning humans can describe how they arrived at their decisions), which may not allow it to maximize accuracy.

Second, the environment in which machine learning operates may itself evolve or differ from what the algorithms were developed to face. While this can happen in many ways, two of the most frequent are concept drift and covariate shift. With the former the relationship between the inputs the system uses and its outputs isn't stable over time or may be mis specified. Consider a machine-learning algorithm for stock trading. If it has been trained using data only from a period of low market volatility and high economic growth, it may not perform well when the economy enters a recession or experiences turmoil—say, during a crisis like the Covid-19 pandemic. As the market changes, the relationship between the inputs and outputs—for example, between how leveraged a company is and its stock returns—also may change. Similar misalignment may happen with credit-scoring models at different points in the business cycle.

In medicine, an example of concept drift is when a machine-learning-based diagnostic system that uses skin images as inputs in detecting skin cancers fails to make correct diagnoses because the relationship between, say, the color of someone's skin (which may vary with race or sun exposure) and the diagnosis decision hasn't been adequately captured. Such information often is not even available in electronic health records used to train the machine-learning model.

Covariate shifts occur when the data fed into an algorithm during its use differs from the data that trained it. This can happen even if the patterns the algorithm learned are stable and there's no concept drift. For example, a medical device company may develop its machine-learning-based system using data from large urban hospitals. But once the device is out in the market, the medical data fed into the system by care providers in rural areas may not look like the development data. The urban hospitals might have a higher concentration of patients from certain sociodemographic groups who have underlying medical conditions not commonly seen in rural hospitals. Such disparities may be discovered only when the device makes more errors while out in the market than it did during testing. Given the diversity of markets and the pace at which they're changing, it's becoming increasingly challenging to foresee what will happen in the environment that systems operate in, and no amount of data can capture all the nuances that occur in the real world.

The third reason machine learning can make inaccurate decisions has to do with the complexity of the overall systems it's embedded in. Consider a device used to diagnose a disease on the basis of images that doctors input—such as IDx-DR, which identifies eye disorders like diabetic retinopathy and macular edema and was the first autonomous machine-learning-based medical device authorized for use by the U.S. Food and Drug Administration. The quality of any diagnosis depends on how clear the images provided are, the specific algorithm used by the device, the data that algorithm was trained with, whether the doctor inputting the images received appropriate instruction, and so on. With so many parameters, it's difficult to assess whether and why such a device may have made a mistake, let alone be certain about its behavior.

[2] In a recent study, Jutzi and colleagues [24] conducted a survey-based study to under- stand the views of the patients in Germany towards AI in the diagnosis of melanoma. In this study involving 298 participants, 154 (51.7%) had a diagnosis of melanoma. Inter- estingly, most of the respondents (94%) supported the use of AI in healthcare. This is a very encouraging finding, as the acceptance of the patient plays an important role in the success of any healthcare related decisions. In line with the point discussed before, sharing data is necessary for better results from the AI. The results of this survey suggested that 88% of participants were aggregable to provide their own health related records for the

development of AI-based applications. Another important finding of this study was that patients with a history of melanoma were more inclined to use AI applications for early diagnosis. Another qualitative study involving 48 patients reported a receptive approach towards the use of AI for the screening of skin cancer, provided it was used in a way that ensured the integrity of the doctor−patient relationship [25]. A study by Oh et al. [26] reported good familiarity of AI by only 5.9% of physicians out of 669 who completed the questionnaire-based study. Interestingly, 83.4% considered AI useful in healthcare. Similarly, a number of participants agreed that disease diagnosis is the most promising area for the use of AI in healthcare. According to 43.9% of participants, AI is superior in the diagnosis of disease compared with doctors. We did not find any study specifically evaluating the acceptance of ML in skin cancer diagnosis by dermatologists. We feel that if increased accuracy and early detection is possible with ML, it may be well accepted by the clinicians. AI will not replace clinicians, but rather assist them in better evaluation and hence management of patients. Although the literature suggests the usefulness of ML in skin cancer diagnosis, its applications will largely depend on acceptance by the clinicians.

<u>Color pigment-based cancer diagnostics:</u>

[27] Human skin color tones fall within a small gamut of the complete spectrum of colors. It’s difficult to describe skin tones in verbal expression. However, color can be expressed in numerical entity. These values re useful in comparing tones for medical research. Commission Internationale de I’ eclairage (CIE) L*a*b is the international standard for numerical representation of color. Skin color varies among varies sites within the body. Even among twins there are color differences. Skin color of gluteal region represents constructive color whereas cheek color represents facultative color. Exposed area like cheek have increase a* value because of vascularization. Epidermal melanin melanosome size influences the L* value. a* indicates erythema while b* value indicate taming and the transition from erythema to taming can be measured objectively. Color values of newborn can give information about their health status. Color values of aged person is useful to know about their nutritional status. Comparing the color of ulcers over a period of time indicate their healing status. There was a 5mmol/L increase in 25 Hydroxy Vitamin D3 levels for every 10-degree lower skin value in the forearm indicating that forearm color can be a reliable indicator for Vitamin D3 status.

There is a 8.2% increase in skin reflectance values for every 10 degree increase in latitude in the Northern hemisphere in males whereas only 3.3% increase was noted for the latitude in the Southern hemisphere indicating the skin color is darker in Southern hemisphere for the corresponding latitude. Minimal Erythema Dose (MED) was defined as an increase in a* value by 2.5 unit. Significant risk for melanoma was found with increasing a* values in Italian patients. Objective assessment of erythema was assessed in 22 psoriatic patients and it was observed that L*a*b values of erythematous plaques were useful in determination of Psoriasis Area and Severity Index (PASI) scores. Fitzpatrick skin types correlated with the MED values. There was significant relationship between the L* values and MED values, but this did not apply for the a* and b* values. The MED value of Narrow Band Ultra Violet B (NBUVB) rays has become a basic criterion for the phototherapy protocol. Hence skin type and L* values are useful for predicting the response to NBUVB therapy.

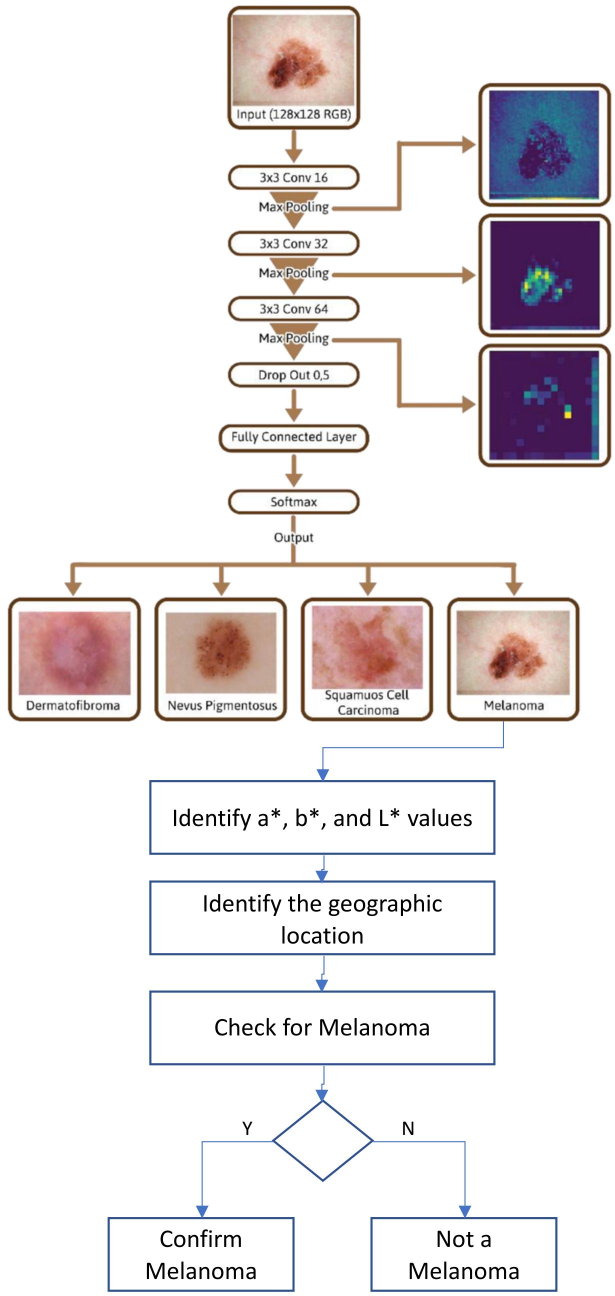
Establishing a rules-based algorithm for the a*, b*, L* along with the geographical location and monitoring the variation over a period of time can indicate the melanoma precisely.

<u>Hybrid approach:</u>

[28] Machine learning is taking the world by storm, and many companies that use rules engines for making business decisions are starting to leverage it. However, the two technologies are geared towards different problems. Rules engines are used to execute discrete logic that needs to have 100% precision. Machine learning on the other hand, is focused on taking a number of inputs and trying to predict an outcome. It’s important to understand the strengths of both technologies so you can identify the right

solution for the problem. In some cases, it's not one or the other, but how you can use both together to get maximum value.

[29] We can have ML models use various features to come to a conclusion which is then used as one of the inputs to a rule-based system. This again works by splicing a complicated into two parts – the more intelligent/complex part uses ML to process complex data patterns. Once the data is reduced to a simple conclusion, the rule-based system can jump in to make further, simpler decisions



Hybrid approach flow [12] - Picture 1

<u>Conclusion:</u>

With the development of science and technology, the diagnosis accuracy and efficiency for skin cancer classification are constantly improving. In the previous clinical diagnosis scenarios of skin cancer, the final diagnosis often depends on the number of samples, image quality and the devices used in the diagnostics, which is highly subjective and has a high rate of misdiagnosis. Recently, with the success of deep learning in medical image analysis, several researchers have applied deep learning methods for skin cancer classification in an end-to-end manner, however, ability to accurately diagnostic melanoma need rules-based approach that leverages colors and its changes. It is expected that in the future, artificial intelligence and the diagnosis of skin cancer diseases would become closely associated.

In this study, we present a comprehensive overview of the most recent breakthroughs in deep learning algorithms for skin cancer classification. Firstly, we introduced different types of dermatological images used in diagnosis and some commonly used datasets, several aspects in the skin cancer deep learning-based approaches that limits the diagnostics accuracy. Next, we present the applications of color, its changes over a period of time in skin cancer. Finally, we provide a hybrid approach that combined deep learning and rules-based approach to improve the diagnostics accuracy.

In comparison to other comparable reviews, this paper presents a unique approach in the topic of skin cancer diagnostics with a focus on contemporary deep learning applications and traditional color-based approach.

References:

[1] https://www.skincancer.org/skin-cancer-information/

[2] https://www.mdpi.com/1660-4601/18/24/13409

[3] Jutzi, T.B.; Krieghoff-Henning, E.I.; Holland-Letz, T.; Utikal, J.S.; Hauschild, A.; Schadendorf, D.; Sondermann, W.; Fröhling, S.; Hekler, A.; Schmitt, M.; et al. Artificial Intelligence in Skin Cancer Diagnostics: The Patients' Perspective. *Front. Med.* **2020**, *7*, 233. [**Google Scholar**] [**CrossRef**]

[4] Sengupta, S.; Mittal, N.; Modi, M. Improved skin lesions detection using color space and artificial intelligence techniques. *J. Dermatolog. Treat.* **2020**, *31*, 511–518. [**Google Scholar**] [**CrossRef**]

[5] Haenssle, H.A.; Fink, C.; Schneiderbauer, R.; Toberer, F.; Buhl, T.; Blum, A.; Kalloo, A.; Hassen, A.B.H.; Thomas, L.; Enk, A.; et al. Man against machine: Diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. *Ann. Oncol.* **2018**, *29*, 1836–1842. [**Google Scholar**] [**CrossRef**]

[6] Du-Harpur, X.; Watt, F.M.; Luscombe, N.M.; Lynch, M.D. What is AI? Applications of artificial intelligence to dermatology. *Br. J. Dermatol.* **2020**, *183*, 423–430. [**Google Scholar**] [**CrossRef**][**Green Version**]

[7] Bini, S.A. Artificial Intelligence, Machine Learning, Deep Learning, and Cognitive Computing: What Do These Terms Mean and How Will They Impact Health Care? *J. Arthroplast.* **2018**, *33*, 2358–2361. [**Google Scholar**] [**CrossRef**]

[8] Lin, J.W. Artificial neural network related to biological neuron network: A review. *Adv. Stud. Med. Sci.* **2017**, *5*, 55–62. [**Google Scholar**] [**CrossRef**]

[9] Bebis, G.; Georgiopoulos, M. Feed-forward neural networks. *IEEE Potentials* **1994**, *13*, 27–31. [**Google Scholar**] [**CrossRef**]

[10] Schmidhuber, J. Deep learning in neural networks: An overview. *Neural. Netw.* **2015**, *61*, 85–117. [**Google Scholar**] [**CrossRef**][**Green Version**]

[11] Gu, J.; Wang, Z.; Kuen, J.; Ma, L.; Shahroudy, A.; Shuai, B.; Liu, T.; Wang, X.; Wang, G.; Cai, J.; et al. Recent advances in convolutional neural networks. *Pattern Recognit.* **2018**, *77*, 354–377. [**Google Scholar**] [**CrossRef**][**Green Version**]

[12] https://iopscience.iop.org/article/10.1088/1757-899X/982/1/012005/pdf

[13] P.Kim, 2017 MATLAB Deep Learning: With Machine Learning, Neural Networks, and Artificial Intelligence.

[14] A. F. Agarap 2008 Deep Learning using Rectified Linear Units (ReLU) 1 p 2–8 [Online]. Available: http://arxiv.org/abs/1803.08375.
[15] S. Khan, H. Rahmani, S. A. A. Shah, M. Bennamoun, G. Medioni and S. Dickinson 2018 A Guide to Convolutional Neural Networks forComputer Vision. Morgan Claypool [Online] Available: https: //ieeexplore.ieee.org/document/8295029
[16] H. Robbins and S. Monro 1985 A Stochastic Approximation Method" in Springer
[17] J. Duchi, E. Hasan and Y. Singer 2011 Subgradient Methods for Online Learning and Stochastic Optimization J. of Machine Learning Research
[18] Y. Yunlong, L. Fuxian. 2019 Effective Neural Network Training with a New Weighting Mechanism-Based Optimition Algorithm IEEE Access.
[19] Bohr, A.; Memarzadeh, K. The rise of artificial intelligence in healthcare applications. *Artif. Intell. Healthc.* **2020**, 25–60. [**Google Scholar**] [**CrossRef**]
[20] Hernández-Orallo, J.; Martínez-Plumed, F.; Schmid, U.; Siebers, M.; Dowe, D.L. Computer models solving intelligence test problems: Progress and implications. *Artif. Intell.* **2016**, *230*, 74–107. [**Google Scholar**] [**CrossRef**][**Green Version**]
[21] Goyal, M. Artificial Intelligence in Dermatology|DermNet NZ [Internet]. DermNet NZ. Available online: **https://dermnetnz.org/topics/artificial-intelligence/** (accessed on 22 June 2021).
[22] Chan, S.; Reddy, V.; Myers, B.; Thibodeaux, Q.; Brownstone, N.; Liao, W. Machine learning in dermatology: Current applications, opportunities, and limitations. *Dermatol. Ther.* **2020**, *10*, 365–386. [**Google Scholar**] [**CrossRef**][**Green Version**]
[23] https://hbr.org/2021/01/when-machine-learning-goes-off-the-rails
[24] Jutzi, T.B.; Krieghoff-Henning, E.I.; Holland-Letz, T.; Utikal, J.S.; Hauschild, A.; Schadendorf, D.; Sondermann, W.; Fröhling, S.; Hekler, A.; Schmitt, M.; et al. Artificial Intelligence in Skin Cancer Diagnostics: The Patients' Perspective. *Front. Med.* **2020**, *7*, 233. [**Google Scholar**] [**CrossRef**]
[25] Nelson, C.A.; Perez-Chada, L.M.; Creadore, A.; Li, S.J.; Lo, K.; Manjaly, P. Patient perspectives on the use of artificial intelligence for skin cancer screening: A qualitative study. *JAMA Dermatol.* **2020**, *156*, 501–512. [**Google Scholar**] [**CrossRef**]
[26] Oh, S.; Kim, J.H.; Choi, S.-W.; Lee, H.J.; Hong, J.; Kwon, S.H. Physician confidence in artificial intelligence: An online mobile survey. *J. Med. Internet Res.* **2019**, *21*, e12422. [**Google Scholar**] [**CrossRef**]
[27]
[28] https://medium.com/capital-one-tech/a-modern-dilemma-when-to-use-rules-vs-machine-learning-61cc908769b0
[29] https://kislayverma.com/programming/combining-rule-systems-and-machine-learning/
[30] https://www.frontiersin.org/articles/10.3389/fonc.2022.893972/full