ISSN (e): 2250 – 3005 || Volume, 14 || Issue, 4|| Jul. - Aug. – 2024 ||
International Journal of Computational Engineering Research (IJCER)

Integrating Diagnostics on Skin Lesion and Classification Using Machine Learning

Kulla Naga Sowndarya[1], S. Surya Godha Devi [2]
#1. M.Tech Scholar in Department of Artificial Intelligence,
#2. Assistant Professor, Department of Artificial Intelligence, Kakinada Institute Of Engineering & Technology or Women, AP, India.

ABSTRACT
Skin cancer is the most common cancer in the existing world constituting one-third of the cancer cases. Benign skin cancers are not fatal, can be cured with proper medication. But it is not the same as the malignant skin cancers. In the case of malignant melanoma, in its peak stage, the maximum life expectancy is less than or equal to 5 years. But, it can be cured if detected in early stages. Though there are numerous clinical procedures, the accuracy of diagnosis falls between 49% to 81% and is time-consuming. So, dermoscopy has been brought into the picture. It helped in increasing the accuracy of diagnosis but could not demolish the error-prone behaviour.
In this project, an automated model for skin lesion classification using dermoscopic images has been developed with CNN(Convolution Neural Networks) as a training model. Convolution neural networks are known for capturing features of an image. So, they are preferred in analyzing medical images to find the characteristics that drive the model towards success. This survey shows that the diagnostic accuracy in image processing methods was relatively uneven, ranged between (50% to 100%). As for the methods of treating tissue features, the accuracy was of an excellent level of 94% or more. The results provide an overview of the actual relevant studies found in the literature and highlight most of which research gaps have emerged.
Keywords: *Skin Diseases, Artificial Intelligence, Machine Learning, Convolutional Neural Network, Skin Lesion Classification , Deep Learning.*

I. INTRODUCTION

Skin lesion refers to the abnormality on the skin. Skin lesions may be of type cancerous, allergic, etc. Among these, cancer-causing skin lesions are hazardous to health. Some forms of these cancerous lesions are deadly. Melanoma is considered to be the one with a high mortality rate of 8% among the cancerous lesions. The occurrence rate of melanoma is increasing day by day. Skin cancer is an abnormal growth in skin cells that sometimes occurs as a result of a specific skin injury that is not treated at first or is often caused by the ultraviolet rays of the sun. Cancer is the second cause of human death in recent years. Approximately nine million people die annually, and 70% of these deaths are recorded in countries with low levels of living income. This is due to the delay in reviewing specialized doctors at the beginning of the disease, and thus the difficulty of treatment due to the progression of the disease and its transformation into a type of cancer and it is reaching advanced stages, causing death.

Data extraction techniques from healthcare systems are useful in designing automated disease diagnostic tools using machine learning algorithms and deep learning algorithms. Researchers have used multiple types of artificial intelligence algorithms to train classifiers needed to perform machine diagnostics, using the principles of machine learning and deep learning. The medical data created in patient records can be used to research the causes of diseases and how they spread, to develop appropriate plans and policies, and to prepare and develop medicines and medical treatments. In villages and rural areas, patients do not receive adequate health care, and there is a lack of dermatologists in those places, and the diagnosis is made by trained staff . Because of the high costs of medical professionals to monitor the condition of a person with a skin disease. There must be technical systems that use images to determine the degree of patient injury. The expert systems technology has contributed to the diagnosis of skin diseases with outstanding accuracy and efficiency, as they simulate human behavior in prediction . Skin disease symptoms are a long-term and continuously changing process, so the person diagnosing the condition must provide an assessment of the extent of the changes that have occurred since the appearance of the lesion.

II. Related Work

Figure 1: Common types of skin disease

Skin diseases, due to their widespread impact, can severely affect both the physical and mental health of individuals. Therefore, precise and prompt diagnosis holds paramount importance for effective management. However, certain skin conditions still pose challenges in diagnosis and treatment. For instance, diseases like skin cancer and vitiligo can be elusive to diagnose in their early stages due to the absence of distinct pathological features. The conventional diagnostic methods heavily rely on visual inspection and subjective assessments, which may lack precision and objectivity [2], leading to potential misdiagnosis even by dermatologists. Moreover, in remote regions with limited access to dermatologists, non-specialists often handle dermatological cases with inadequate training and resources, despite the availability of reference materials. The shortage of dermatologists and the uneven distribution of healthcare resources exacerbate the difficulties in achieving accurate diagnoses in underserved regions.

AI technology based on image recognition has emerged as a promising tool for diagnosing skin diseases. These algorithms can be trained using large datasets of skin images to learn disease patterns, potentially providing more accurate diagnoses, especially in early stages. Additionally, AI algorithms can offer more objective diagnoses by avoiding human biases through careful design and debugging [3]. This technology has the potential to address some of the challenges associated with diagnosing skin diseases, particularly in underserved areas lacking dermatologists. Common AI algorithms used for this purpose include machine learning (ML) and deep learning (DL), both capable of identifying repetitive features of skin lesions for accurate diagnosis of both benign and malignant conditions. While DL typically performs better with large and complex datasets, ML methods remain useful in situations with limited data. These methods can be integrated into computer-aided diagnosis (CAD) systems, providing accurate classification results for dermatologists. Moreover, for non-specialists, such systems can mitigate errors stemming from their limited expertise [4]. Hence, it is crucial to explore the advancements and recent accomplishments of machine learning (ML) and deep learning (DL) techniques in dermatological diagnosis. This examination can help identify existing challenges and suggest suitable recommendations to propel advancements in this field. Figure 2 illustrates the typical schematic diagram of the automated skin image diagnosis procedure.

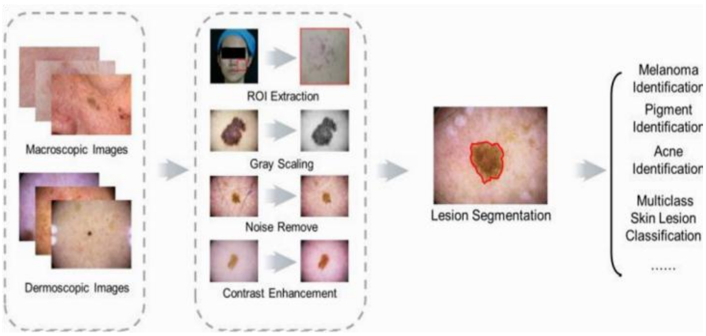


Figure 2: Schematic diagram of skin image diagnosis.

A.Skin Lesion Datasets and Image Preprocessing

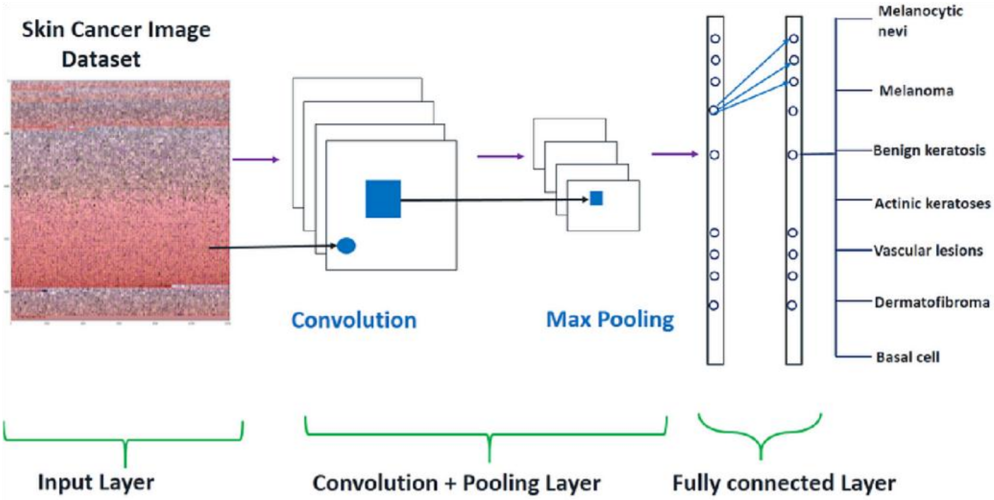
Educational institutions and medical organizations have generated datasets of various sizes and types of skin lesions to develop skin diagnosis models. These datasets double as educational platforms [5] for the public and as testing grounds for new diagnosis algorithms. However, images often contain artifacts like hair, varying illumination, circular markings, and black frames, which degrade image quality and impede analysis. To address these challenges, image pre-processing techniques are employed, including resizing, grayscale conversion, noise removal, and contrast enhancement, either individually or in combination. These techniques typically enhance both the efficiency and effectiveness of diagnosis, as well as the processes of segmentation and classification. Image resizing is a common approach to handling images with diverse sizes, allowing most images to be standardized to similar dimensions. Furthermore, deep learning (DL) methods excel at managing images with varying sizes [6]. For instance, Convolutional Neural Networks (CNNs) exhibit translation invariance, enabling trained models to accommodate different image scales without compromising performance. Consequently, researchers can confidently address skin classification and segmentation tasks across a wide range of image sizes. Image resizing involves standardizing the pixel count and cropping irrelevant portions of the image before extracting the area of interest (ROI). This not only reduces computation time but also facilitates subsequent diagnosis tasks. To mitigate the impact of varying skin tones, scaled photos are often converted to grayscale images [7]. Different filters, as depicted in Figure 3, can be applied to eliminate various types of noise artifacts. For ROI determination, a monomorphic filter can enhance contrast and equalize image brightness. Hair removal filters are selected based on hair features: thick hairs can be addressed with inpainting techniques, while more hair can be removed using the Dull Razor method. Gaussian, average, and median filters are effective for thin hair removal.

III. METHODOLOGY

- CNN Model:-

After the Extraction training model is created. training the model, the model weight and architecture can be saved with the .h5 file extension which is a Keras file. Tensorflow is needed in this process where you freeze the graph producing a .h5 file and a txt file containing the labels of the created model. To fully use the model for inference, a Tensorflow utility is used which is the Image Data Generator function and with this, the model can be successfully loaded into your application. It shows the block diagram of the necessary steps in deploying the model.

CNN Architecture:



Skin Cancer Detection using Convolution Neural Network(CNN)

- Method 1 - Training from scratch :-

The architecture is initialized with random weights and trained for a number of epochs. After each epoch, the model learns features from data and computes weights through backpropagation. This method is unlikely to produce the most accurate results if the dataset is not significantly large. However, it still can serve as a baseline for comparison against the two other methods.

- Method 2 :-

ConvNet as feature extractor Due to the relatively small number of images of skin lesion in most dermatology datasets, this method initializes the model with weights from the VGG16 trained on a larger dataset (such as ImageNet), a process known as transfer learning. The underlying assumption behind transfer learning is that the pre-trained model has already learned features that might be useful for the classification task at hand. This corresponds, in practice, to using selected layer(s) of the pre-trained ConvNet as a fixed feature extractor, which can be achieved by freezing all the convolutional blocks and only training the fully connected layers with the new dataset.

• Method 3 :-

Fine-tuning the ConvNet Another common transfer learning technique consists of not only retraining the classifier on the top of the network with the new dataset, but also applying a fine-tuning of the network by training only the higher-level portion of the convolutional layers and continuing the backpropagation. In this work, we propose to freeze the lower level layers of the network because they contain more generic features of the dataset.

True

Table: Confusion Matrix for Method 3

Malignant	59	16
Benign	12	63

Predicted

IV. MODELING AND ANALYSIS

The melanoma diagnosis can be improved with the ABCD rule based and computer assisted systems. These systems usually consist of the separate units for the image segmentation, feature extraction and classification respectively [8-12]. Studies conducted in this field are as follows: Baldrick et al. compared in their study the expert opinion and artificial neural networks when they classify the lesions. They obtained from the computer program a sensitivity of 95% and a specificity of 88%, while they measured the expert dermatological sensitivity and specificity as 95% and 90% respectively.Clinically, skin cancers are diagnosed based on ABCDE rule [2]. Where, • A - Asymmetry • B - Border irregularity • C - Color of lesion • D - Diameter of lesion • E - Enlarging lesion.

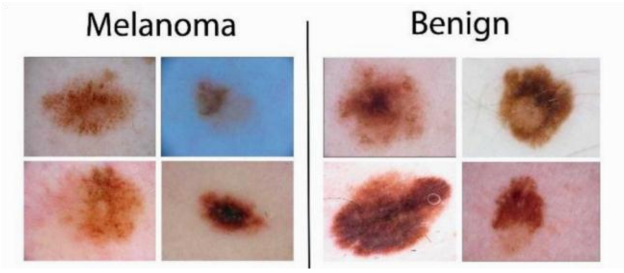


Figure : Sample images created from the ISIC Archive dataset

• MELANOMA :-

It grows rapidly and becomes risky in one or two months. It comes under malignant melanocytic lesion and it might be widening to other parts when untreated. It is the major cause of increased mortality rate due to its nature metastasis. It occurs on the covered areas which are not exposed to sun and it might be blue, red or grey in color ; it is unusual in people with dark skin and it is frequently appears under the fingernails and toe nails, on hand palms, and feet soles.

• Preprocessing :-

Input images must be preprocessed by: (i) normalizing the pixel values to a [0,1] range; (ii) cropping the image to square aspect ratio (if necessary); and (iii) resizing the image to the expected size of 224×224 pixels.

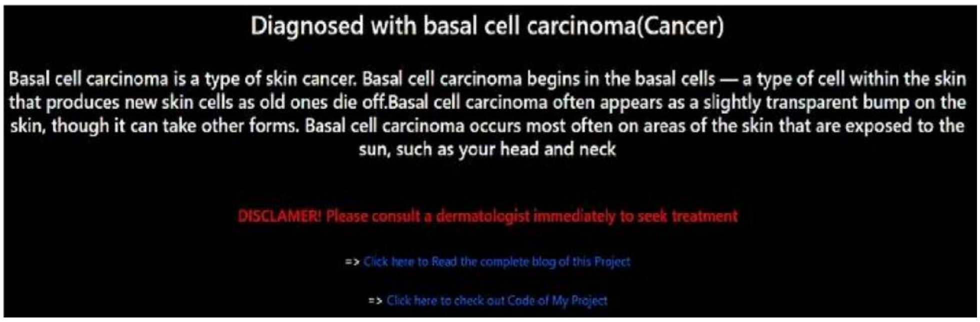
• Feature Extraction :-

At the beginning, Convolutional Neural Network (CNN) is a fixed of stacked layers regarding each nonlinear and linear processes. These layers are
discovered in a joint manner. The primary constructing blocks of any CNN version are: convolutional layer, pooling layer, nonlinear Rectified Linear Units (ReLU) layer linked to a normal multilayer neural community known as absolutely linked layer, and a loss layer on the backend. CNN hasacknowledged for its massive overall performance in packages because the visible responsibilities and herbal language processing.

• SVM :-

SVMs are nonparametric classifiers. Regarding their distribution is no preliminary information as a presupposition available. Inputs and outputs are paired in the training sets. Through the pairs, decision functions are obtained which classify the input variables in the test set and new data set.

V. RESULT/OUTPUT



VI. CONCLUSION

This model is not robust to all skin images, because, we not trained with good amount of equal class images data. Due to random oversampling it may give some wrong predictions to images. This research work has developed a ML and DL based effective automated segmentation and classification model for skin lesion images. The presented research work involves several subprocesses such as preprocessing, segmentation, feature selection, feature reduction and classification. The research work is presented under a set of four objectives. Skin treatments are more effective and less disfiguring when found early. We should point out that it is to replace the human input on analysis and intuition.

The proposed approach achieves promising results – most notably, a sensitivity value of 78.66% and a precision of 79.74% – which are significantly higher than the current state of the art on this dataset (50.7% and 63.7%, respectively). We propose a solution for assisting dermatologists during the diagnosis of skin lesions. More specifically, we have designed and implemented a two-class classifier that takes skin lesion images labeled as benign or malignant as an input, builds a model using deep convolutional neural networks, and uses this model to predict whether a (previously unseen) image of a skin lesion is either benign or malignant. Most of the researchers have tended to use the convolutional neural network algorithm for its high ability to classify large groups of skin diseases and to overcome the confusing obstacles to the process and also its success in classifying the lesions into benign and malignant, but one of the disadvantages is that it takes a long time to train and process data.

VII. FUTURE SCOPE

As a part of future scope, the presented model can be extended by the following factors:Develop an Internet of Things (IoT) and cloud based skin lesion diagnosis model to assist professionals and patients in smart healthcare applications. The research work and the future scope presented in this these may bid a new track of research which would catch the attention from many research communities and of course, may offer enhanced set of solution models in the future.

REFERENCES

[1]. Robert J Friedman, Darrell S Rigel, and Alfred W Kopf. Early detection of malignant melanoma: the role of physician examination and self-examination of the skin. CA: a cancer journal for clinicians, 35(3):130–151 1985.
[2]. Nils Gessert, Maximilian Nielsen, Mohsin Shaikh, René Werner, and Alexander Schlaefer. Skin lesion classification using ensembles of multi-resolution efficientnets with meta data. MethodsX, page 100864, 2020.
[3]. Josef Steppan, Sten Hanke arXiv:2101.03814v1 [eess.IV] 11 Jan 2021
[4]. K. J. Wernli, N. B. Henrikson, C. C. Morrison, M. Nguyen, G. Pocobelli, and P. R. Blasi, "Screening for skin cancer in adults: updated evidence report and systematic review for the us preventive services task force," Jama, vol. 316, no. 4, pp. 436–447, 2016.

[5]. L. F. di Ruffano, Y. Takwoingi, J. Dinnes, N. Chuchu, S. E. Bayliss, C. Davenport, R. N. Matin, K. Godfrey, C. O'Sullivan, A. Gulati et al., "Computer-assisted diagnosis techniques (dermoscopy and spectroscopybased) for diagnosing skin cancer in adults," Cochrane Database of Systematic Reviews, no. 12, 2018.

[6]. Adriano Lucieri 1 ,2[0000 −0003 −1473 −4745], Fabian Schmeisser1[0000 −0001 −8222 −7900], Christoph Peter Balada2[0000 −0003 −0307 −7866] .