

Human Skin Disease Detection using MLXG model

Dhruv Dodia ^a, Hemangi Jakharia ^a, Ritik Soni ^a, Shwetambari Borade ^a, Nilakshi Jain ^a

^a Shah & Anchor Kutchhi Engineering College, Chembur, Mumbai, Maharashtra, India

Abstract

Human skin disease detection deals with the recognition of skin type and the skin disease in a given image. Edges, textures, and colour pixels are often used to detect skin diseases. These features are invariant and are fast to process. A variant Human Skin Disease Detection model (MLXG) is put forward in this paper. The three main parameters which help to detect skin diseases are the edges, the RGB (Red, Green, Blue) colour model of the image, and the texture of the skin. The goal of the MLXG model is to improve the current human skin disease detection models in terms of accuracy and to work on different skin diseases. The model not only considers the above parameters but also uses Machine Learning to improve the accuracy.

Keywords

Skin Disease, Machine Learning, edges, image processing, RGB, dry, acne scars, vitiligo, warts, alopecia areata, oily, acne, MLXG.

1. Introduction

According to analysis human skin diseases are ranked fourth common disease affecting the world's 1/3 rd population [18]. In a study, a total count of 2701 population was engaged amongst which 1662 entrants i.e. 64.5% had at least one skin abnormality [19]. Human Skin is an important part of the human body which consists of three layers. Human Skin Disease accrued with ages. Frequency of increase of human skin disease in males (72.3%) than in females (58.0%)[19]. Skin can also be affected due to its skin type whether it is oily or dry skin. The diagnosis of skin disease is important and should not be ignored.

Clinical examination showed that just about a simple fraction of the affected entrants were unfamiliar with their aberrant skins. Human Skin diseases come up in a variety of forms, lack and inconsistent availability of dermatologists, need for timely diagnosis to prevent long term effects to the skin, calls for computer-aided diagnosis. Human skin disease detection is the process of finding the skin type and the infected part of the skin in a given image provided by the user. Human skin disease detection system uses a range of image processing, convolution neural network, and machine learning algorithms. The primary feature for skin disease recognition is the edges, but edges cannot be the only factor for the same, as there are variations in the skin types according to different conditions such as geographical conditions, climatic conditions, etc.

The following factors should be considered for determining skin diseases:

1. Varying skin tone with respect to different races will affect the RGB pixel values of the image.
2. Individual characteristics such as age and gender also affect the contours of the image.

CVMLH-2022: Workshop on Computer Vision and Machine Learning for Healthcare, April 22 – 24, 2022, Chennai, India.

EMAIL: shwetambari.borade@sakec.ac.in (Shwetambari Borade)

ORCID: 0000-0001-7547-6351 (Shwetambari Borade)



© 2022 Copyright for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

- Other factors such as watermark, shadow, blurred image, background colour affect the contours of the image.

2. Literature Survey

Many Authors and Researchers have imposed Image Processing Techniques and Different Machine learning algorithms to detect the type of skin and relate to that type of skin disease. We have reported some techniques which are as follows.

In [1], a system is proposed, image processing technique 2-D Wavelet Transform is used as a feature extractor from the images, and the detection of disease of the processed images has been done using Convolutional Neural Network (CNN). They have tested three diseases and have good accuracy.

The author of [2] applied the method Edge Detection using Gaussian Filter Adaptive Thresholding, K-Means clustering. Segmentation algorithms are used. Chickenpox, eczema, psoriasis, ringworm are the diseases detected.

The work [3] of Acne, Amyloidosis, Cherry Angioma, Eczema Lids, Halo-Nevus are the diseases detected in this technique. KNN is the algorithm used here. The image segmentation is done by the active delineation method. The feature extraction is done using a mathematical equation, then the classification of diseases takes place. The image will be grayscale converted, cropped and noise filtered. The accuracy of this method is 81.92%.

Similarly, [4] the input image is pre-processed by resizing it to the size fixed for the dataset. Feature extraction was proposed from a pre-trained convolutional neural network. A support vector machine does the classification. This model detected 3 types of skin with a 100% accuracy rate. Dataset has 80 images.

Therefore, [5] The processing of an image contains image filtering, image rotation & image segmentation. In feature extraction, 3 types of human skin diseases are detected. They are given as follows: herpes, dermatitis, and psoriasis. The average accuracy of this model is 87%. 90 images are used in the dataset.

According to the observation from Table I that shows the comparison of current work for human skin disease detection, it has been noticed that the MLXG model that uses the XGBoost algorithm as a classifier is not yet explored. In this paper, we have illustrated a machine learning system for human skin disease detection on images using different techniques such as machine learning and deep learning

Table 1: Comparison Table of Various Survey Papers

Sr no.	Year	Paper Cited	Disease count	Diseases	Accuracy	Method	Dataset
1	2020	[3]	5	<ul style="list-style-type: none"> • Acne • Amyloidosis • Cherry-Angioma • Eczema • Lids • Halo-Nevus 	81.92%	KNN	-----
2	2019	[4]	3	<ul style="list-style-type: none"> • Melanoma 	100%	AlexNet - DeepCNN, SVM	80 images

3	2019	[6]	8	<ul style="list-style-type: none"> • Eczema • Psoriasis • Acne • Actinic • keratosis • Angioedema • Blepharitis, • Granuloma facial • Pityriasis Alba • Rosacea, • Vitiligo 	88%	VGG16 – DeepCNN	6000 images
4	2019	[2]	4	<ul style="list-style-type: none"> • Eczema, • psoriasis • chicken pox, • ringworm 	-----	Kmeans	-----
5	2018	[5]	3	<ul style="list-style-type: none"> • Herpes, • Dermatitis, • Psoriasis 	85%	GLCM - Features SVM - Classification	90 images
6	2020	[1]	3	<ul style="list-style-type: none"> • Psoriasis, • Lichen • Planus, • Pityriasis • Rosea 	-----	2D Wavelet Transform algorithm: feature extraction convolutional neural network (CNN): Classification	-----
7	2019	[7]	2	<ul style="list-style-type: none"> • malignant, • melanoma or seborrheic • keratosis. 	98%	GLCM, SVM	-----
8	2019	[8]	2	<ul style="list-style-type: none"> • Acne • Boils 	96%	Hessian Matrix	400
9	2019	[9]	2	<ul style="list-style-type: none"> • malignant • melanoma or seborrheic • keratosis. 	85%	KNN, SVM	-----
10	2020	[10]	5	<ul style="list-style-type: none"> • Nummular • eczema • lichen simplex • stasis 	73%	CNN	3000

				<ul style="list-style-type: none"> • dermatitis • ulcers 			
11	2018	[11]	4	<ul style="list-style-type: none"> • Acne, • Psoriasis • Melanoma • Heat Rash 	86%	KNN, MSIN	-----



Figure 1: Disease selected for Human Skin Disease Detection System

3. Methodology

3.1. Dataset

The dataset itself means collection of data. The diseases which we have selected for our model are acne, acne scars, vitiligo, alopecia areata, warts shown in fig 1. The dataset has been generated manually using google images, we have also used images from Dermnet. The images used to build the dataset are free from watermark, the pixels of the images are in the form of RGB, the background and blurred part has been cropped. The count of images present in the dataset is shown in Table 2.

The MLXG model, images are divided in the ratio of 80:20 where 80% is used to train and the rest 20% is used to test.

Table 2: Image count in dataset

Classification Task	Number of images
Acne	58
Acne Scars	48
Vitiligo	51
Alopecia Areata	55
Warts	45

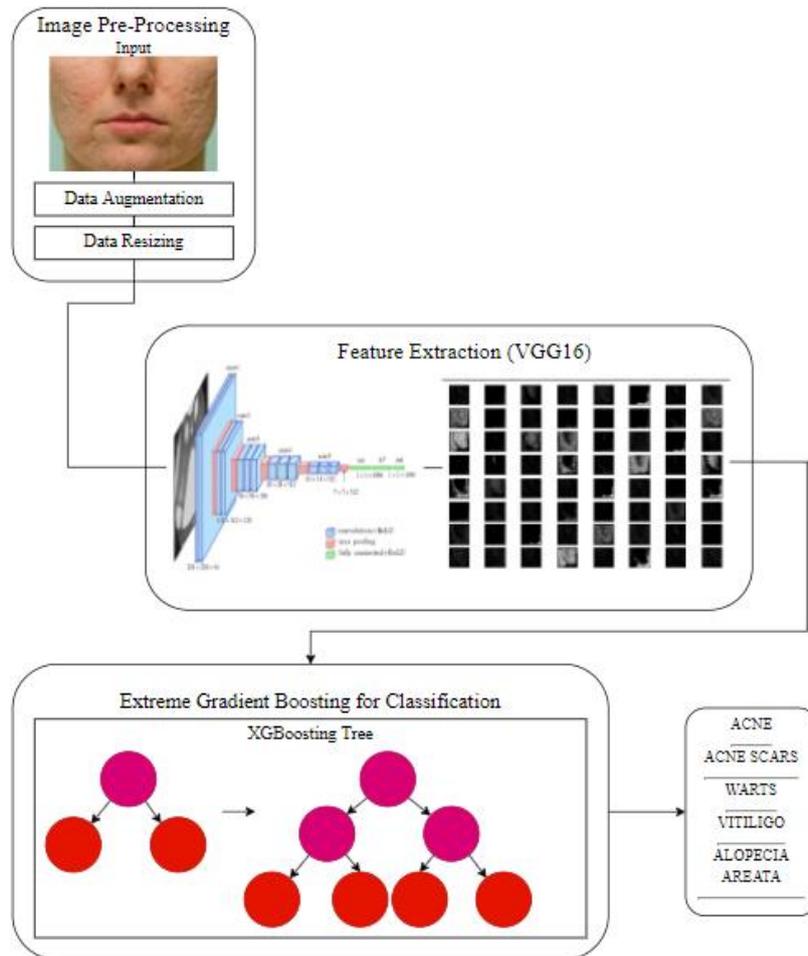


Figure 2: Proposed MLXG Methodology for classifying human skin disease

3.2. Approach

The proposed Machine Learning with XGBoost on microscopic examination of human skin diseases detection in brief. The MLXG model is classified into two prime parts, the first one is the feature extraction and the second one is the classifier. The initial step taken in the architecture is pre-processing the images for uniformity in the images. The pre-processing of images also includes resizing and resampling. The above-mentioned procedure i.e. data pre-processing is taken to maintain the uniformity and the quality of the images before passing them further into the Feature extraction or a machine learning model. The key features from the image are extracted by the network named VGG16. VGG16 executes an important role in pointing out the features from the images passed after pre-processing and also explores all the pixels of an image in depth. The final step of the MLXG model trains all the extracted features from the images using the XGBoost classification algorithm. The XGBoost algorithm is scalable and incisive to train the MLXG model to distinguish different human skin diseases. Fig 2 demonstrates the workflow of the proposed Machine Learning with XGBoost (MLXG) for the Human skin diseases detection expert system.

3.3. Pre-processing and data augmentation

Image processing plays an important role in the images before passing the images in the model. The randomly downloaded images are collected and further resampled and rearranged. To load the images, we have used the OpenCV library. Later we have resized the image by $224 * 224 * 3$ three are the channels that say the image is in the RGB form. The image is resized to reduce the calculation

of dimensions of each pixel for the feature extractor. The dataset is divided into 80% and 20%, training and testing respectively.

Data augmentation is a technique in pre-processing for small datasets. A challenge for human skin disease detection is its limited dataset so we have applied data augmentation to generate a model with high accuracy and to avoid overfitting. The objective to use data augmentation is not only to increase the dataset but also to formalize and resolve the disparity. This process will also strengthen the MLXG model.

3.4. Extraction of features

As extraction of features is an important task over which, the performance of classification of skin diseases depend. So, we have selected a deep learning perspective for the model to use each pixel of the image and automatically finds out the features from an image provided. The convolution conducts signal processing operations that may be calculated in an easy manner as discrete spatial processing operations, on the basis of principle of Convolution Neural Network (CNN). To implement the model, we need a Training model and a Classifier Model.

- **Training from Scratch**

In order for this strategy to be accurate and dependable, it will require a large number of histological breast pictures to be input. This technique will take more time and effort to achieve since it necessitates defining and fine-tuning the necessary parameters to reach the optimal results. The learning rate, number of layers, kind of convolutional layer, and other hyper parameters are all included in this hyper parameter. Aside from the difficult issue of tuning the parameter, training also necessitates high-performance GPU processing capability.

- **Transfer Learning:**

By transferring information learnt from a source domain, such as the dataset from ImageNet with a huge quantity of data, to the model targeted, we may solve issues like overfitting and create a more generic deep learning model. In a deep learning model, these models that are pre-trained can give adequate information and helps in the preparation of the limited histology dataset. This method aims on using the information gained from a group of data samples and apply it to any future samples that are not included in the data. The model will be able to fetch all of the learnt features and knowledge to generate predictions in future for fresh samples in this manner. Learning provides benefits such as accelerating network convergence, lowering processing power, and improving network performance.

As a result, transfer learning is a more effective technique than training the model from the beginning for extraction of features from an input by just randomly initialising weights. As a result, throughout this process, we used pre-trained models. Many deep learning models have enhanced the CNN model for detecting skin diseases, such as:

1. VGG16 Feature Extraction

For this system, VGG16 (Visual Geometry Group) was selected as the pre-trained feature extraction model. The VGG16 is improvised because it has less hyperparameters and is a better version of AlexNet. The VGG16 consists of convolution + ReLU, maximum pooling, and full connection + ReLU. The input size of the image passed in the first level of convolution is a fixed size $224 * 224 * 3$ (RGB channels). The image then goes further through the stack of convolution layers, with the filter size set to $3 * 3$ and the convolution step size set to 1 pixel. Spatial padding for convolutional layer inputs is such that spatial resolution is maintained after convolution, so $3 * 3$ convolutional layer padding is 1 pixel.

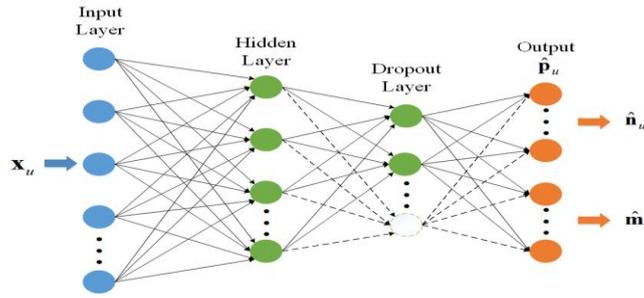


Figure 3: VGG16 Model [14]

Spatial pooling will continue to be performed using the Max-Pooling layer. Maximum pooling is performed at $2 * 2$ pixels using Stride 2 while keeping the padding the same. Maximum pooling can be derived using

$$(N + 2P - F / S) + 1.$$



Figure 4: VGG16 Architecture [14]

Therefore, this is followed by three Fully Connected layers that follow the stack of convolution layers. The first two fully connected layers contain 4096 channels and the last layer contains 1000 channels. As shown in Figure 4, the VGG16 architecture contains 16 layers. Figure 5 shows the feature extraction by VGG16.

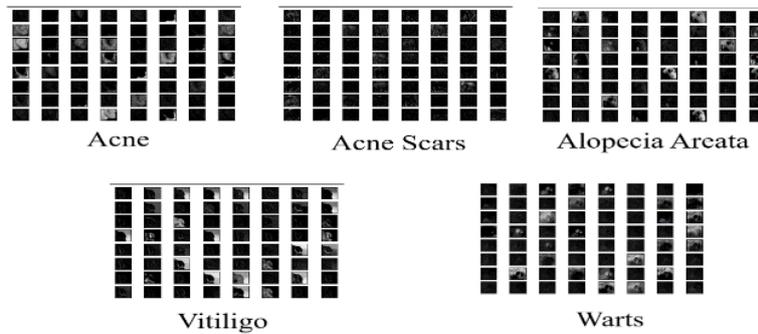


Figure 5: Training images using VGG16

2. Extreme Gradient Boosting (XGBoost) Classifier

It is a brand-new tree-primarily based totally set of rules that has these days emerge as famous in facts class and has established to be an enormously powerful facts class method. XGBoost is an enormously scalable stop-to-stop tree boosting machine utilized in system mastering for class and regression tasks. XGBoost is extreme, and that means that it is a big machine learning algorithm with a great number of parts. XGBoost was designed to be used with large, complicated data sets. We are training data consisting of five different skin diseases. From fig 7 green dots indicate that the classified skin disease was true and the red dots indicate that the classified skin disease was false. The first step in adapting XGBoost to your training data is to make an initial prediction. The prediction can be anything. In our project, the probability of observing skin disease predictions in training data is 0.5 by default, regardless of whether XGBoost is used as a regression or classification.

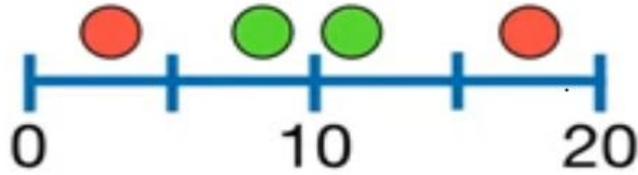


Figure 6: Skin Diseases

From fig 6 we can illustrate the initial prediction by adding a y-axis to our graph to represent the probability of classification. Drawing a thick line at 0.5 to represent a 50% chance that the classification is correct. The two green dots we will move up as it represents true classification of skin disease and the probability is one. Similarly, the two red dots we will leave them at bottom where the probability is 0 because it represents false classification. The residual, or difference between the observed and predicted values, shows how good the initial prediction is.

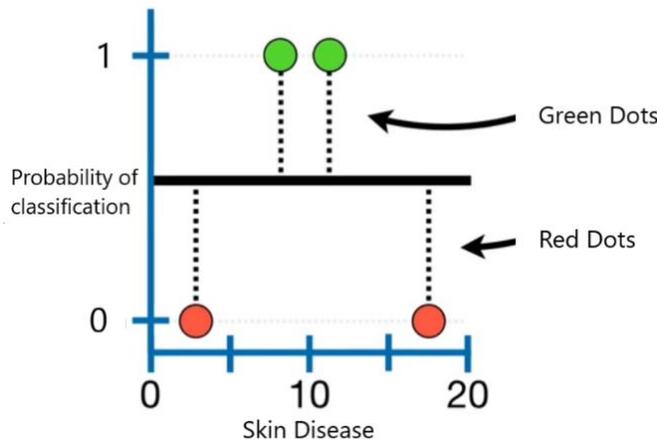


Figure 7: Probability of Classification of Skin Diseases.

The formula to calculate the similarity scores is given in fig 6. Lambda is regularization parameter.

$$\frac{(\sum \text{Residual}_i)^2}{\sum [\text{Previous Probability}_i \times (1 - \text{Previous Probability}_i)] + \lambda}$$

In the next step we tried to fit an XGBoost Tree to the Residuals. Each tree begins off evolved out as a lone leaf and all the residuals go to the leaf. So now we need to calculate the similarity score for the leaf. As we do not square the residuals before we add them together, they will cancel each other out so the similarity score = 0. Next, we need to divide the similar residuals into two groups to determine if they can be grouped better. So, our next step is to split. As we chose the threshold, classification < 15 because 15 is the average value between last observations. Thus, the three residuals with classification < 15 go to the leaf on the left and the one residual with classification > 15 moves to the leaf on the right side. To calculate the similarity score for the three residuals that ended up in the leaf on the left. So, we plug in the three residuals into the numerator (-0.5, 0.5, 0.5) and since we are building the first tree, the Previous Probability refers to the prediction from the initial leaf. So, let's keep lambda 0.

$$\frac{(-0.5 + 0.5 + 0.5)^2}{(0.5 \times (1-0.5)) + (0.5 \times (1-0.5)) + (0.5 \times (1-0.5)) + 0}$$

Thus, the similarity score = 0.33. Similarly, to calculate the similarity score for the leaf on the right is. Now calculate the Gain.

$$\frac{(-0.5)^2}{0.5 \times (1 - 0.5) + \lambda} = 1, \text{ when } \lambda = 0$$

$$\text{Gain} = \text{Left}_{\text{Similarity}} + \text{Right}_{\text{Similarity}} - \text{Root}_{\text{Similarity}}$$

$$\text{Gain} = 0.33 + 1 - 0 = 1.33$$

No other threshold gives us a larger Gain value and that means classification < 15 will be the first branch in our tree. The least number of Residuals in every leaf is determined by calculating cover. Cover is the denominator of the Similarity Score minus lambda. Gain is associated with the lowest branch and a number we pick for gamma. The Output value is given by the formula.

$$\text{Cover} = \sum [\text{Previous Probability}_i \times (1 - \text{Previous Probability}_i)]$$

$$\text{Gain} - \gamma = \begin{cases} \text{If positive, then do not prune.} \\ \text{If negative, then prune.} \end{cases}$$

$$\text{Output Value} = \frac{(\sum \text{Residual}_i)}{\sum [\text{Previous Probability}_i \times (1 - \text{Previous Probability}_i)] + \lambda}$$

Lambda is the regularization parameter and when lambda > 0, it results in greater pruning, by reducing the similarity scores, and smaller output values for the leaves. Figure gives the idea of tree built and shows the XGBoost tree.

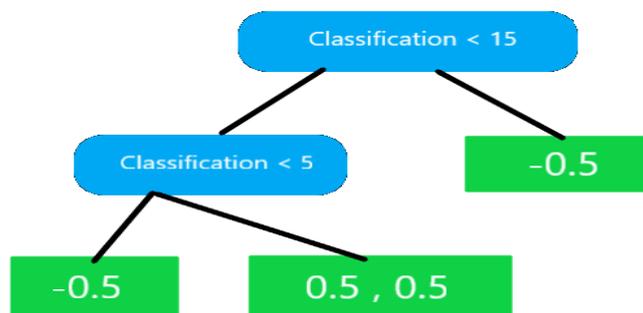
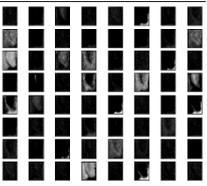
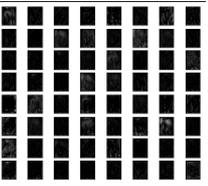
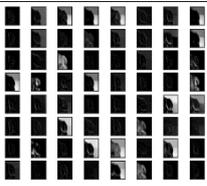


Figure 8: Tree using XGBoost

	precision	recall	f1-score	support
acne	1.00	1.00	1.00	4
acne-scars	1.00	0.75	0.86	4
alopecia	0.80	1.00	0.89	4
vitiligo	1.00	1.00	1.00	2
warts	1.00	1.00	1.00	2
accuracy			0.94	16
macro avg	0.96	0.95	0.95	16
weighted avg	0.95	0.94	0.94	16

Figure 9: Analysis of Average Result of Classification task

Table 3: Feature Extraction using VGG16 and Range of the values

	Acne	Acne Scars	Alopecia Areata	Vitiligo	Warts
Image of features extracted using VGG16					
Minimum Value	-123.68	-123.68	-123.68	-123.68	-80.939
Maximum Value	139.061	151.061	133.061	61.06	109.061

4. Experimental Outcome

In this research, we are focusing to develop a Human Skin Disease Detection using Machine Learning algorithms. The proposed model is implemented using Python 3.8.5. OpenCV library is used for the implementation of image processing. The experiment is run on Spyder IDLE. Figure 10 outlines the proposed method.

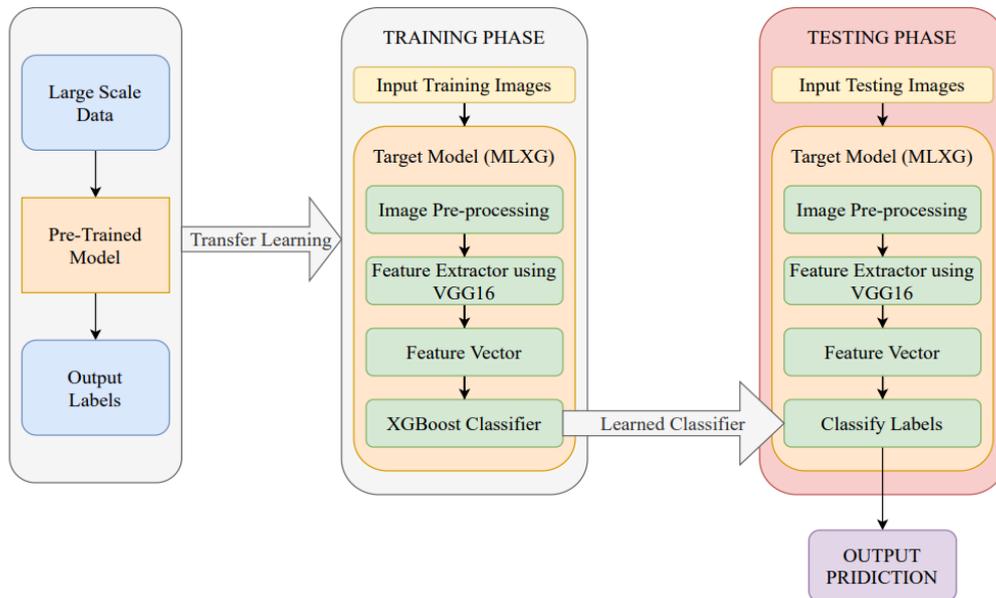


Figure 10: Overview of the Proposed model

a. Configuration

We used the initial weights trained on the ImageNet dataset to train the VGG16 model on the proposed system. For fine tuning, image resize and augmentation are implemented to stabilize the dataset. The image will be converted into BGR. All images used in dataset are resized with the targeted values 224 * 224 pixels. Further, the image will be converted into an n-dimensional array using NumPy. The pre-trained model will extract the features from the images. Label Encoder has been used to encode the labels of the images. As we are working with pre-trained weights so we have made loaded layers as non – trainable. So, the trainable parameter will be zero once the feature has

been extracted by VGG16. The features are extracted from the image and the image is reshaped into its original form. XGBoost has been used as a classifier algorithm because the optimization of XGBoost is faster than any other algorithm and it will classify skin diseases.

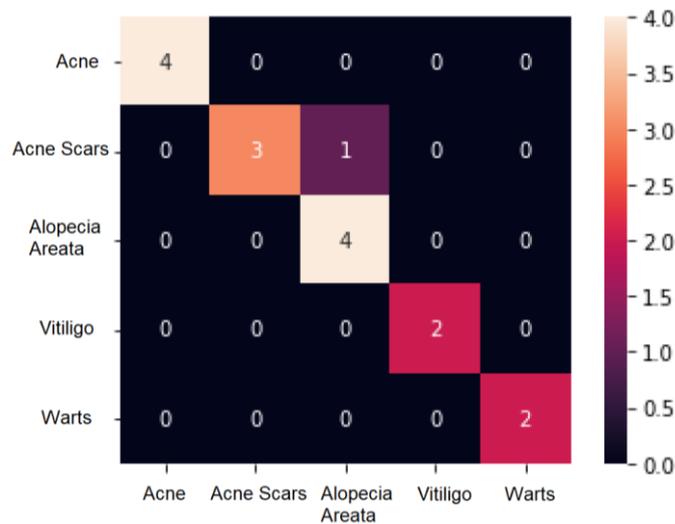


Figure 11: Confusion Matrix of model MLXG

b. Results

To display the effectiveness of human pores and skin sickness detection, it's far crucial that we compare our method to test and to recognize the overall performance of the gadget to test if the underlying troubles are solved or not. Common evaluation for the diagnose are Precision, Accuracy, F1-measure, and Recall. Training and testing of the final proposed MLXG model was done on around 1000 images of skin disease. Fig 5 illustrates the extraction of features in different layers using the VGG16 model. The values for each are detected from minimum value till maximum value shown in Table 3.

Figure 11 shows the confusion matrix for skin disease detection. A confusion matrix is a matrix commonly used to give an account about the execution and efficiency of a classification model (or "classifier") on a test data set for which the true values are known. The confusion matrix is between test labels and prediction labels. Fig 12 illustrates the confusion matrix where TP is an acronym for True Positive, FP for False Positive, FN for False Negative and TN for True Negative.

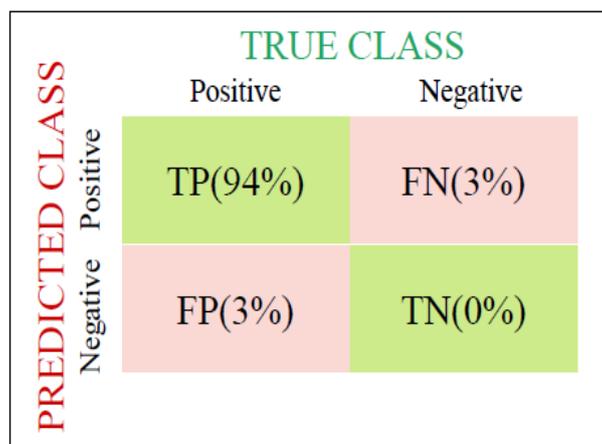


Figure 12: Confusion matrix between test labels and prediction labels

The accuracy achieved using the MLXG model is 94% as shown in figure 9. The accuracy is calculated using the definition as True Positive (TP) when the cases are predicted to have a skin disease; true Negative (TN) when the cases predicted not to have a skin disease. False Positive (FP) when the cases predicted to have a skin disease but actually have different disease not the one predicted; false Negative (FN) when the cases predicted not to have a skin disease but actually they have the skin disease. Figures 13, 14 and 15 show the output of the image and the predicted output from the machine learning algorithm and the actual label of the image is to be given.

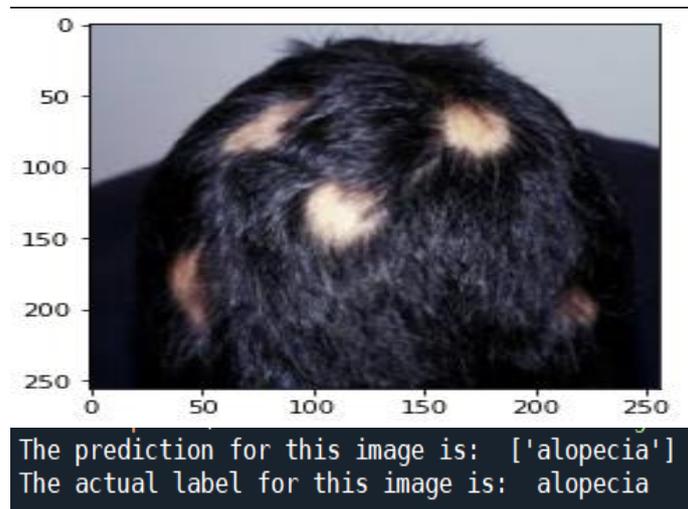


Figure 13: Output using MLXG model



Figure 14: Output using MLXG model

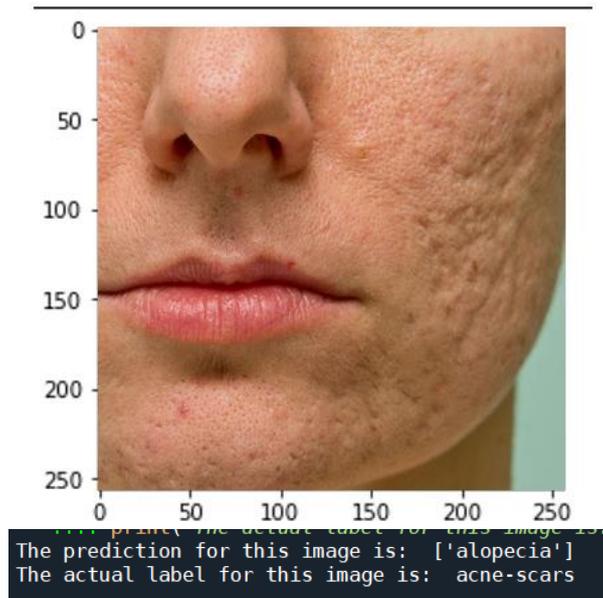


Figure 15: Output using MLXG model

5. Conclusion and Future Scope

This article introduced an improved classification system for detecting human skin disorders using techniques of machine learning and this paper presented focuses on exploring a unique approach to human skin disease detection using the same. As far as we know, our approach has not been evaluated in the disease-related studies selected for the detection of human skin diseases. This proposed method achieves 94% accuracy and great performance with the combination of feature extractor VGG16 which is a part of deep learning and machine learning classifier XGBoost for classifying the problems. The proposed system can correctly identify acne, acne scars, alopecia areata, vitiligo and warts. The MLXG method employs a deep learning technique by learning features of an image through every single pixel using a pre-trained VGG16 network. The classification is then performed by running the XGBoost classifier on the basis of extracted feature vectors. We can reach the conclusion that this technique can be potentially used for the classification of diagnostic images when developing MLXG expert systems. An average of 94% accuracy was achieved in the detection of human skin diseases.

This paper attempts a new study of the combination of VGG16 and XGBoost. The proposed model is built using a limited set of images in the dataset. We can test this model on big datasets as it doesn't use more space and the time taken to give the output is faster than other models. The skin diseases selected are unique and the prediction given by the model is 100% accurate and can be used by the dermatologist. In modern times, the efficient, proper and right use of technology is required. This project will help develop the technological infrastructure of our nation. It will also support the Digital India Campaign which was launched by our honourable Prime Minister Mr. Narendra Damodardas Modi.

6. References

- [1] "SKIN DISEASE DETECTION USING COMPUTER VISION AND MACHINE LEARNING TECHNIQUE." https://ejmcm.com/article_2063.html (accessed Jan. 13, 2022).
- [2] "Skin Disease detection based on different Segmentation Techniques | IEEE Conference Publication | IEEE Xplore." <https://ieeexplore.ieee.org/document/8862403> (accessed Jan. 13, 2022).
- [3] "IRJET- Skin Disease Identification using Image Processing and Machine Learning Techniques by IRJET Journal - Issuu." <https://issuu.com/irjet/docs/irjet-v7i3265> (accessed Jan. 13, 2022).

- [4] N. Alenezi, "A Method of Skin Disease Detection Using Image Processing and Machine Learning," *Procedia Comput. Sci.*, vol. 163, pp. 85–92, Jan. 2019, doi: 10.1016/j.procs.2019.12.090.
- [5] "Disease Recognition Method Based on Image Color Skin and Texture Features." <https://www.hindawi.com/journals/cmmm/2018/8145713/> (accessed Jan. 13, 2022).
- [6] "Deep convolutional neural network for face skin diseases identification | IEEE Conference Publication | IEEE Xplore." <https://ieeexplore.ieee.org/document/8940336> (accessed Jan. 13, 2022).
- [7] K. Melbin and Y. J. V. Raj, "An Enhanced Model for Skin Disease Detection using Dragonfly Optimization based Deep Neural Network," in 2019 Third International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Dec. 2019, pp. 346–351. doi: 10.1109/I-SMAC47947.2019.9032458.
- [8] Ma. C. R. Navarro, E. Bustillos, and D. P. Y. Barfeh, "Skin Disease Analysis using Digital Image processing," in 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), Dec. 2019, pp. 311–316. doi: 10.1109/ICCIKE47802.2019.9004267.
- [9] Y. Wang, J. Cai, D. C. Louie, H. Lui, T. K. Lee, and Z. Jane Wang, "Classifying Melanoma and Seborrheic Keratosis Automatically with Polarization Speckle Imaging," in 2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP), Nov. 2019, pp. 1–4. doi: 10.1109/GlobalSIP45357.2019.8969331.
- [10] T. A. Rimi, N. Sultana, and Md. F. Ahmed Foysal, "Derm-NN: Skin Diseases Detection Using Convolutional Neural Network," in 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), May 2020, pp. 1205–1209. doi: 10.1109/ICICCS48265.2020.9120925.
- [11] A. Haddad and S. A. Hameed, "Image Analysis Model For Skin Disease Detection: Framework," in 2018 7th International Conference on Computer and Communication Engineering (ICCCE), Sep. 2018, pp. 1–4. doi: 10.1109/ICCCE.2018.8539270.
- [12] V. Bevilacqua, A. Brunetti, A. Guerriero, G. F. Trotta, M. Telegrafo, and M. Moschetta, "A performance comparison between shallow and deeper neural networks supervised classification of tomosynthesis breast lesions images," *Cogn. Syst. Res.*, vol. 53, pp. 3–19, Jan. 2019, doi: 10.1016/j.cogsys.2018.04.011.
- [13] X. Y. Liew, N. Hameed, and J. Clos, "An investigation of XGBoost-based algorithm for breast cancer classification," *Mach. Learn. Appl.*, vol. 6, p. 100154, Dec. 2021, doi: 10.1016/j.mlwa.2021.100154.
- [14] "VGGNet-16 Architecture: A Complete Guide | Kaggle." <https://www.kaggle.com/blurredmachine/vggnet-16-architecture-a-complete-guide> (accessed Jan. 13, 2022).
- [15] J. Parashar, Sumiti, and M. Rai, "Breast cancer images classification by clustering of ROI and mapping of features by CNN with XGBOOST learning," *Mater. Today Proc.*, Nov. 2020, doi: 10.1016/j.matpr.2020.09.650.
- [16] "XGBoost | Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining." <https://dl.acm.org/doi/10.1145/2939672.2939785> (accessed Jan. 13, 2022).
- [17] "Simple guide to confusion matrix terminology." <https://www.dataschool.io/simple-guide-to-confusion-matrix-terminology/> (accessed Jan. 13, 2022).
- [18] C. Flohr and R. Hay, "Putting the burden of skin diseases on the global map," *Br. J. Dermatol.*, vol. 184, no. 2, pp. 189–190, 2021, doi: 10.1111/bjd.19704.
- [19] L. Tizek, M. c. Schielein, F. Seifert, T. Biedermann, A. Böhner, and A. Zink, "Skin diseases are more common than we think: screening results of an unreferred population at the Munich Oktoberfest," *J. Eur. Acad. Dermatol. Venereol.*, vol. 33, no. 7, pp. 1421–1428, 2019, doi: 10.1111/jdv.15494.