Classification of Skin Diseases Using Deep Learning

[1]Rashid Ansari, [2]Sanjay Kadam, [3]Samadhan Sonavane
[1]PG Scholar, [3]Assistant Professor
[1]Sandip University, Nashik, India, [2]CDAC, Pune, India, [3]Sandip University, Nashik, India

Abstract: Skin diseases conventionally detect by observing primary features. Skin diseases possess a lot of resemblances. Detection of skin diseases becomes problematic and vary from a dermatologist to dermatologist as a number of the feature of skin diseases increases. So, it is important to develop a computer-aided diagnosis system which can detect skin diseases without any problem. In this paper, three skin diseases such as tinea corporis, tinea cruris, and tinea faceii are classified by machine learning and deep learning method. Primary, pre-processing is performed to remove the unnecessary region from skin images and extract the region of interest. Next, HOG and GLCM algorithms are applied to extract feature. Lastly, NuSVC classifier is used to determine skin diseases. The results obtained shows that HOG features with SVM classifier achieved an accuracy of 78 % which is better than other techniques. In a deep learning system, pre-trained Xception model is fine-tuned by retraining whole model with lower learning rate. This model achieved an accuracy of 82%.

Keywords: image processing, machine learning, deep learning, dermatologist, classification

I. INTRODUCTION

A classification problem is when the target variable belongs to some particular category. A classification model endeavors to reach some determination from watch esteems. Given at least one data sources, a classification model will attempt to foresee the estimation of at least one results. Classification of skin diseases is a difficult task because of some skin diseases have similar symptoms and looks relevant to each other.

Deep learning has appeared most remarkable in the field of machine learning over the last few years and the image classification performance has adequately increased [16]. Deep learning model consists of various layers. Deep learning is popular because of the transformation of hardware which could deal with a lot of information. Specifically, the convolutional neural systems made the most astounding progress in the field of image classification and are generally connected to a large portion of the cutting edge deep learning techniques.

Each sort of skin infection is having certain distinctive features. Based on these attributes, characterization is performed. The surface is a significant component that recognizes the article present in a picture. The texture is characterized by the spatial circulation of pixels in the area of an image.

Skin diseases image classification also showed high performance when utilizing deep learning techniques. Feature extraction and learning from the input data is performed automatically in deep learning techniques though the machine learning techniques have particular techniques to perform feature extraction.

In this paper, classification is performed using machine learning and deep learning techniques. In machine learning techniques, feature extraction is performed using a Histogram Of Oriented Gradient, and Gray-Level Co-Occurrence Matrix (GLCM) [10], dimensionality reduction is performed using principal component analysis (PCA) [13] and classification is performed using classification algorithms (Support Vector Machine, Decision Tree and Random Forest) [13]. Out of which HOG features with Support Vector Machine perform better as compared to other techniques. Whereas, in deep learning techniques we've use of transfer learning Xception model which is an extension of inception v3 model [15].

II. RELATED WORKS

M. S. Manekar et al. [9] proposed a system which automatically segments different skin cancers. Images are pre-processed by applying RGB to L*a*b conversion and contrast enhancement and image segmentation are performed using C-means clustering and watershed algorithm. The feature is extracted using GLCM and Image Quality Assessment (IQA) followed by SVM classifier. C-means algorithm generated a superior result (Acc. 98%) as compared to the watershed algorithm (Acc. 92%).

E. Kazmierczak et al. [8] prosed system which is semi-supervised scaling segmentation algorithm and can segment scaling right away from erythema and skin images. Features are extracted through Gabor texture analysis and scaling contrast map. Their method has been more successful than SVM and Markov Random Field.

R. Maurya et al. [7] proposed a system for skin cancers classification of four types. In these, system skin lesions features extracted through GLCM and classification is performed using Support Vector Machine. They got an accuracy of 81.43%.

Md. Nazrul Islam et al. [6] used a maximum entropy thresholding method for image segmentation and GLCM for feature extraction. The proposed framework comprises of a feedforward multilayer network with backpropagation as a preparation model accomplishing a precision of 80%.

Esteva et al. [5] build a Deep CNN model which have the ability to achieve fine-grained classification. There model show correspondent result with all tested experts. Their model performs better when trained on finer disease partition as compared to multiple class. CNN learned internal features using t-SNE (t-distributed Stochastic Neighbor Embedding) and achieved an accuracy of 72.1%.

Hanging Zhou et al. [4] also used CNN for multi-class classification and achieved an accuracy of 65.8%(six diseases) and 90% (two diseases).

Zhang X et al. [3] build a model based on CNN. First, they preprocessed skin images by removing noise and enhance those images. Convolutional Neural Network is used to extract features and perform classification with the help of softmax classifier.

Y. Hasija et al. [2] proposed a system which is a combination of support vector machine and convolution neural network join with image processing tools to achieve accuracy up to 95.3%.

Xavier Giro-i-Nieto et al. [1] develop a system with three methods which are CNN from scratch, CNN as feature extractor and Fine-tuning with CNN (VGGNet). The proposed solution is built around VGGNet architecture.Results M1-66.00%,M2-68.67%,M3-81.33

III. MATERIALS AND METHODS

In this paper, the classification of skin diseases is performed using machine learning and deep learning models.

1. Classification using machine learning

In machine learning, we have to follow following steps to achieved classification.

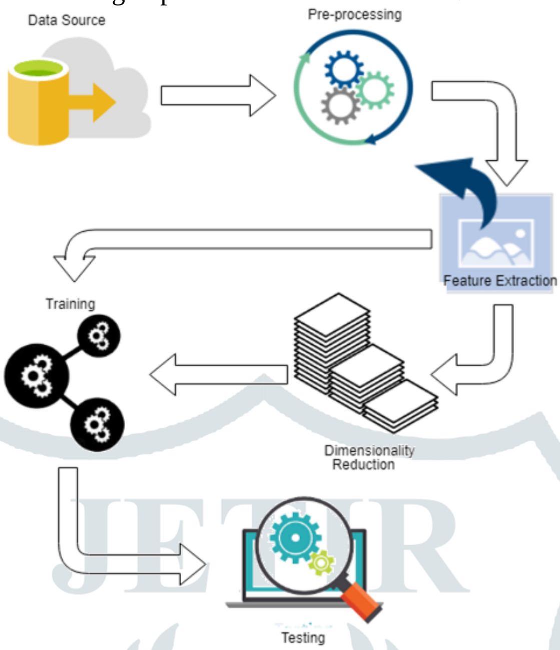

Fig 3.1 Machine Learning System

a) Dataset

Dataset is comprised of 317 skin diseases images of three diseases such as tinea corporis, tinea cruris, and tinea faceii. Data is collected from various online dermatologist repository like dermnet, dermis and data is collected from the hospital.

b) Pre-processing

In this step, blurry and corrupted images are removed manually. Images are crop manually to eliminate the unnecessary background and to extract the region of interest. Further, the image may have varying size, so before further processing, images should have standardized shape and size.

c) Feature Extraction

This is an important part of this system. In which feature extracted from image helps to identify the whole image. The feature contains a piece of necessary and useful information which helps to discriminate various images. Algorithms used for feature extraction are:

- **Histogram of Oriented Gradient (HOG):**

In this feature extraction algorithm, dissemination of the order of gradients is used as features [10, 11]. The gradient is highly desirable due to the magnitude of the gradient is considerably high around corners and edges. Below are the steps which help to find HOG feature descriptor:

a) Fixed aspect ratio is considered while analyzing patches.
b) Horizontal and vertical gradients are calculated to find HOG descriptor.
c) Direction and magnitude of the gradient are also calculated.
d) Image is separated into 8x8 cells and for each 8x8 cells histogram of gradients is calculated. Each cells contains 9 bins corresponding to angles 0, 20, 40.. 160 [10].
e) To make a gradient of image is independent of lighting variation image is normalized [10].

- **Gray-Level Co-Occurrence Matrix:**

GLCM is a matrix of different combinations of pixel brightness values (gray level) in an image [10, 12]. GLCM texture thinks about the relation between the reference pixel and the neighbor pixel in an image. GLCM depicts the frequency of occurrence of pixels with specific values in a specified spatial relationship in an image, which is then followed by extracting statistical measures from this matrix, like homogeneity, correlation, symmetry, and energy, etc [10, 12].

d) Dimensionality Reduction

A total number of features extracted from the above algorithms is very large. Working on these features are computation intensive and difficult to deal with it. This problem can be solved by using dimensionality reduction techniques which help to diminish the number of arbitrary features analysis by recovering a lot of key factors. In this study, we've used principal component analysis for this task [13].

e) Classification

Classification can be done on many factors. Factors represent a feature. This feature used to characterize the input image from the available categories. There are various techniques for classification. In this study, we've used SVM.

SVM: Support vector machine is a binary as well as a multiclass classifier which used one vs one or one vs rest techniques to predict multiple classes. SVM separated classes using hyperplane. Implementation of SVM is easy and with less computation can generate excellent output. Benefits of using SVM is it beneficial for handling higher dimensional data and is memory efficient [13].

2. Classification using deep learning

Deep learning models figure out how to perform classification directly from images. Sometimes this model can produce results better than human [14].

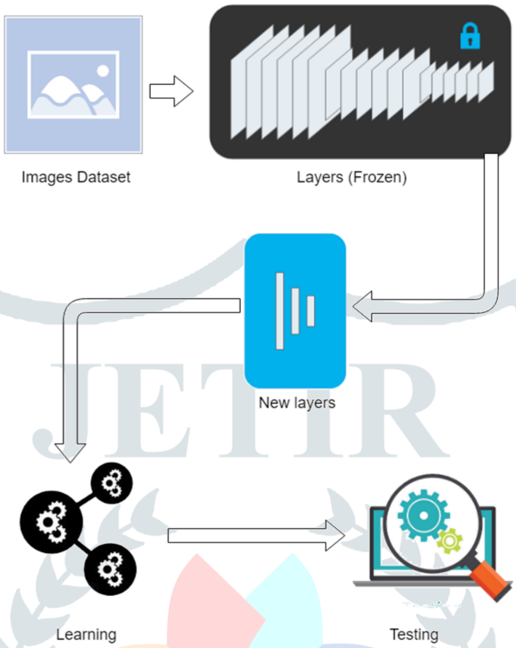

Fig 3.2 Deep Learning System

In this paper, pre-trained model Xception is used. Xception is an extension of Inception. It has depthwise separable convolution which shows better performance as compared to inception-v3. This model is 1st Runner Up in ILSVRC 2015 for both ImageNet ILSVRC and JFT datasets [15].

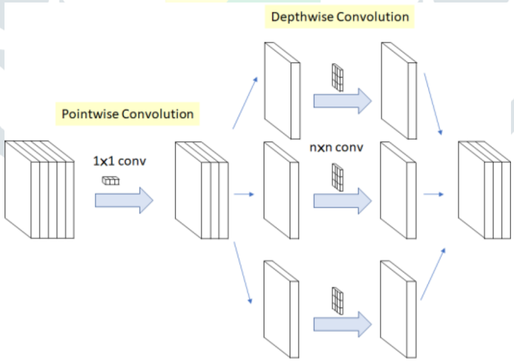

Fig 3.3 Modified Depthwise Separable

From depthwise convolution pursue by a pointwise convolution is modified to pointwise convolution pursued by a depthwise convolution in xception is shown in Fig 3.3 [15]. This alteration suggests that 1×1 convolutional is performed before any n×n spatial convolutions which inspired from inception v3 [14, 15].

In this system, we create a custom exception model. We utilized weights from IMAGENET dataset.

In this system, we build custom xception model. This model is trained over an imagenet dataset so, we've utilized the weight of this model. Subsequently, we've frozen the body layer and train only the top classifier. Afterward, we train the whole model without freezing any layer. Whole model train with a lower learning rate.

IV. EXPERIMENT AND RESULTS

a. Using machine learning techniques

For the datasets collected, we perform the following process to classify skin diseases. Firstly, images were manually cropped to remove unnecessary background from the images and keep only the affected region. Then, features are extracted from these images by using a Histogram of Oriented Gradient and GLCM algorithm with a fixed image size of 250*250 pixels. These, extracted is safe in an h5 file locally. Later, these extracted features are used by Nu-SVC classifier to predict the class of skin diseases.

Parameters settings used in this classifier are, the kernel is set to 'sigmoid', 0.001 is the tolerance of stopping criteria, verbose is set to True and rest of parameters values are set to default. Results are shown in Table 4.1.

Table 4.1 Machine Learning Model Accuracy

Feature algo/ Classifier	No. of Features	SVM
HOG	62500	78.48
GLCM	65536	70.54

There is a lot of features is generated from the above algorithm which takes a lot of time while training. To reduced training time we've used PCA for feature reduction and results is shown in Table 4.2.

Table 4.2: Machine Learning Model Accuracy with feature reduction

Feature algo/ Classifier	No. of Features	Reduced Features	SVM
HOG	62500	106	78.48
GLCM	65536	99	70.54

b. Using deep learning techniques

To classify skin diseases through lesion images of corresponding skin diseases we train a Neural Network. Once our model gets a train, we test it on the test dataset and save the model in an h5 file in JSON form. The new unseen image is provided to our model for classification. To train our model we used Imagenet weights and build a custom Xception model. First, we train only the top classifier for 5 epochs. Then, re-train the whole model for 20 epochs with a lower learning rate. The result is shown in table 4.3.

Table 4.3: Classification Report

Labels	Precision	Recall	F1-score
TineaCorporis(0)	0.85	0.85	0.85
TineaCruris(1)	0.79	0.79	0.79
TineaFaceii(2)	0.83	0.83	0.83
Average	0.82	0.82	0.82

Fig 4.2: Confusion Matrix

V. CONCLUSION AND DISCUSSION

The study has been done to determine three typical skin diseases which are tinea corporis, tinea cruris, and tinea faceii. Machine learning models and deep learning model is used to analyze the input lesion images belong to which class. Image contains a lot of extraneous information which can be reduced by cropping an image to remove unnecessary region and extract the region of interest. Pursuant to that, the diverse feature extraction algorithm is applied to extract meaningful features from images. Also, to classify skin diseases various classification algorithm is used. From the results, it is observed that feature extraction through HOG with support vector machine classifier perform well as compared to other techniques. Xception pre-trained model is utilized for classification purpose. This model is fine-tuned by retraining the whole model with a lower learning rate. From the result, it is observed that deep learning techniques perform better than other conventional techniques. By executing deep learning techniques we can achieve high precision value even for numbers of skin diseases.

REFERENCES

[1]. A. Romero Lopez, X. Giro-i-Nieto, J. Burdick and O. Marques, "Skin lesion classification from dermoscopic images using deep learning techniques," 2017 13th IASTED International Conference on Biomedical Engineering (BioMed), Innsbruck, Austria, 2017, pp. 49-54.

[2]. Y. Hasija, N. Garg and S. Sourav, "Automated detection of dermatological disorders through image-processing and machine learning," 2017 International Conference on Intelligent Sustainable Systems (ICISS), Palladam, 2017, pp. 1047-1051.

[3]. Zhang X, Wang S, Liu J, Tao C. Towards improving diagnosis of skin diseases by combining deep neural network and human knowledge. BMC Med Inform Decis Mak. 2018;18(Suppl 2):59. Published 2018 Jul 23.

[4]. H. Zhou, F. Xie, Z. Jiang, J. Liu, S. Wang, and C. Zhu, "Multi-classification of skin diseases for dermoscopy images using deep learning," 2017 IEEE International Conference on Imaging Systems and Techniques (IST), Beijing, 2017, pp. 1-5.

[5]. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017 Feb;542(7639):115.

[6]. M. N. Islam, J. Gallardo-Alvarado, M. Abu, N. A. Salman, S. P. Rengan and S. Said, "Skin disease recognition using texture analysis," 2017 IEEE 8th Control and System Graduate Research Colloquium (ICSGRC), Shah Alam, 2017, pp. 144-148.

[7]. R. Maurya, Surya Kant Singh, A. K. Maurya and A. Kumar, "GLCM and Multi-Class Support vector machine based automated skin cancer classification," 2014 International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, 2014, pp. 444-447.

[8]. J. Lu, E. Kazmierczak, J. H. Manton and R. Sinclair, "Automatic Segmentation of Scaling in 2-D Psoriasis Skin Images," in IEEE Transactions on Medical Imaging, vol. 32, no. 4, pp. 719-730, April 2013.

[9]. Manerkar MS, Snekhalatha U, Harsh S, Saxena J, Sarma SP, Anburajan M. Automated skin disease segmentation and classification using multi-class SVM classifier.

[10]. "Classification of Skin Diseases Using Machine Learning Techniques", International Journal of Emerging Technologies and Innovative Research, ISSN:2349-5162, Vol.6, Issue 5, page no.362-369, May-2019.

[11]. N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 2005, pp. 886-893 vol. 1.

[12]. Robert M Haralick; K Shanmugam; Its'hak Dinstein (1973). "Textural Features for Image Classification" (PDF). IEEE Transactions on Systems, Man, and Cybernetics. SMC-3 (6): 610–621.

[13]. James, Gareth, et al. An introduction to statistical learning. Vol. 112. New York: springer, 2013.

[14]. Géron, Aurélien. *Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems*. " O'Reilly Media, Inc.", 2017.

[15]. Chollet, François. "Xception: Deep learning with depthwise separable convolutions." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1251-1258. 2017.

[16]. Brinker, Titus Josef, et al. "Skin cancer classification using convolutional neural networks: systematic review." *Journal of medical Internet research* 20.10 (2018): e11936.