IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Identification Of Skin Disease Using Machine Learning And Python

[1]Ramya, [2]MullaiVendhan, [3] Dr.Ananthalakshmi

[1]PG Student, [2] Working professional, [3]Assistant Professor
[1]Electronics and Communication,
[1]Meenakshi College of Engineering, Chennai, Tamilnadu

Abstract: It is a difficult element for physicians of diagnosing problems even rising technology, in skin ailment symptoms. Most humans are uncovered to serious pores and skin ailments that require them to go to hospitals and go via quite a few distinctive high priced clinical examinations which take up to days. The proposed work can clear up the above hassle to an approachable extent, thru the graph of software by means of Python Machine Learning. An approach is based totally on vectors and pixels classification of the images, and is proposed to discover the 3 quite a number kind of pores and skin diseases: The ailments are particularly Melanoma, seborrheic keratosis and Nevus. It is a manner to become aware of the kind of sickness in simply a few seconds, making the prognosis extra quickly and realistic. These undertaking goals to classify the specific illnesses based totally on photographs given as input. The challenge is in simple terms primarily based on a python software program platform. Here the proposed approach is ply of Machine Learning with the tensor goes with the flow for educating raw data and CNN Algorithm to classify the 3 sorts of pores and skin ailments in Python application. These kind of illnesses like Melanoma, seborrheic keratosis and Nevus can be recognized and indicates the output as the identity of the ailment in python Software as output.

***Index Terms* - CNN, Python, Machine Learning, vector classification.**

I. INTRODUCTION

To ensure whether the given set of images is Melanoma, seborrheic keratosis, and Nevus using image processing methodology. Many skin diseases cause itchiness, dry skin, or rashes. Treatment can reduce symptoms and may even keep them at bay for months at a time. Skin diseases include all conditions that irritate, clog or damage your skin, as well as skin cancer. It's important to check your skin for any changes, including new or non-healing spots or changes in moles. If detected and treated early, the majority of skin cancers are curable. The deepest layer of the epidermis, located just above the dermis, contains cells called melanocytes. Melanocytes produce the skin's pigment or color. Melanoma, seborrheic keratosis, and Nevus begin when healthy melanocytes change and grow out of control, forming a cancerous tumor. Malignant refers to the ability of a cancerous tumor to develop and metastasize to other areas of the body. Sometimes, Melanoma, seborrheic keratosis, and Nevus develop from a normal mole a person already has on their skin. When this happens, the mole will undergo changes that usually can be seen, such as changes in shape, size, color, or the border of the mole. Melanoma, seborrheic keratosis, and Nevus can develop anywhere on the body, including the head and neck, the skin under the fingernails, the genitals, and even the bottom of the feet or palms of the hands. Melanoma, seborrheic keratosis, and Nevus may not be colored like a mole. It may have no color or be slightly red, which is called melanotic melanoma.

When found early, Melanoma, seborrheic keratosis, and Nevus can often be cured with surgery. However, melanoma is one of the most dangerous forms of skin cancer. This may grow deep into the skin; this is called invasive melanoma. It can also invade blood vessels and spread to lymph nodes and distant parts of the body; this is called metastatic melanoma. Seborrheic keratosis are both non-contagious and nontoxic. They don't require treatment, but if they get irritated by clothes or you don't like the way they appear, you could choose to have them removed. Nevi are often identified clinically by dermatoscopy or by observation with the unaided eye. For separating melanocytic nevi from melanoma, more sophisticated imaging procedures are available, such as computerized dermatoscopy and image analysis. The type of nevus and the level of diagnostic ambiguity affect how it is managed. Since certain nevi are known to be benign, they may just need to be periodically checked. Others may need further testing, including a biopsy, for histological analysis (looking at a sample of skin under a microscope to detect unique cellular properties).

II. Literature Review

Local gradient-based descriptors can be drawn to extract individual insulator caps and then using elliptical descriptors to detect the insulator caps and utilizing local outlier factor (LOF) algorithm to determine the possible defects of insulators. Traditional threshold segmentation method was adopted by the Literature to extract insulators from the background. After that, faults can be recognized according to the ratio of pixels in an area. Literature firstly segments the aerial photo of power transmissionsystem with K-means clustering to acquire the foreground connected regions of the insulators; then, it locates explosion based on the characteristics of insulators after segmentation. The above are few image processing methods using threshold segmentation of images of power transmission line and fault recognition on the segmentation based on the explosion features of insulator. Due to the difficulty of extraction the insulator from complex background through manual method, the accuracy of the recognition is critically low and the robustness is weak. The main method used by this paper were to set up the network of detection, segmentation and recognition based on deep learning, training each network through samples collected by drones and ultimately achieving the intelligent of detection and recognition. With deep learning, the subjectivity in manual extraction for features and threshold selection can be avoided.

Each electrical insulator must go through a number of insulator tests to assure the required performance and to prevent unintended insulator failure. We shall attempt to comprehend the many reasons why insulators fail before doing insulator testing. Insulator testing ensures the electrical insulator's quality since the insulator's quality determines the likelihood that the insulation will fail.

Table 1. ***Literature survey tabulation***

Title	Author	Summary
Melanoma skin cancer detection using deep learning and classical machine learning techniques	T. Shanthi a , R.S. Sabeenianb , R. Anandc	The challenges in automating the process includes the variation of skin tones, location of the disease, specifications of the image acquisition aliquando id per
Detection and recognition for fault insulator based on deep learning	Yongli Wang, Jiao Wang, Feng Gao, Panfeng Hu, Li Xu, Jian Zhang,Yiliang Yu, Jun Xue, Jianqing Li	Insulator extracting and segmentation by deep learning enhances robustness improves the accuracy.
Type of Skin Disease Identification by Machine Learning Using Python	Dr. R. Muthalagu 1, M. Mohammed Musheer 2, K. Nandhini 3, N. Saranraj4	The project is purely based on python software platform. The images are collected from various publicly available databases like DermWeb, Dermnet etc. Here the proposed method is the use of Machine Learning with tensor flow for training the dataset and the SVM Algorithm to classify the five types of skin diseases in Python software.

III. CNN ALGORITHM

Deep Learning has established itself as a very potent tool over the last few decades due to its capacity for handling massive volumes of data. Hidden layer technology is much more popular than conventional methods, particularly for pattern recognition. The most well-liked deep neural network is Convolutional Neural Networks. General outline of how CNN Algorithm works is shown

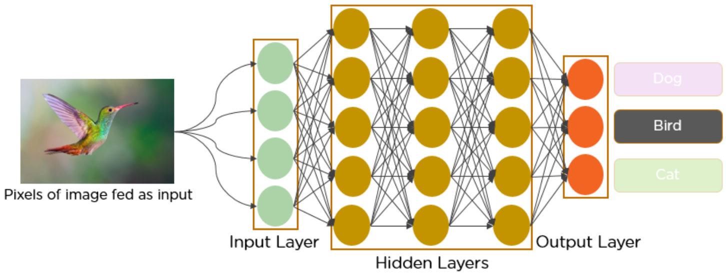
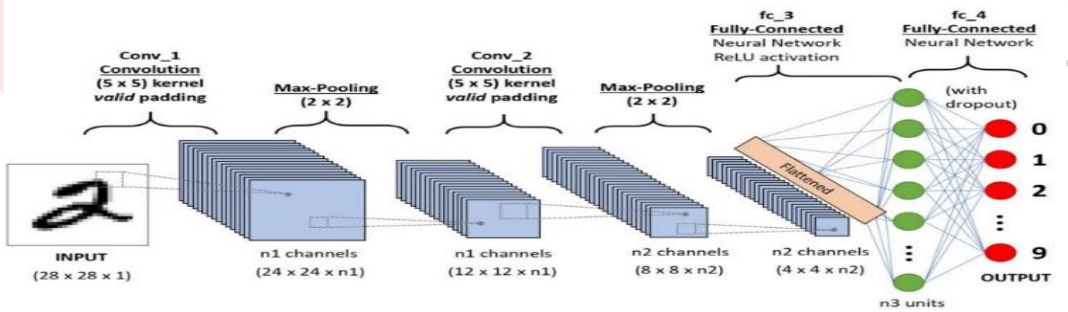


Figure 1. General Outline of CNN

CNNs are regularized versions of multilayer perceptron's. Multilayer perceptron's often refer to networks that are completely linked, meaning that every neuron in one layer is coupled to every neuron in the layer above it. Due to their "complete connectedness," these networks are vulnerable to data overfitting. Regularization or overfitting prevention methods frequently include punishing training parameters (such as weight decay) or cutting connectivity (skipped connections, dropouts, etc.). By utilizing the hierarchical structure in the data and assembling patterns of increasing complexity using smaller and simpler patterns imprinted in their filters, CNNs adopt a novel strategy for regularization. Consequently, in terms of connection and CNNs take a different approach towards regularization: they take advantage of the hierarchical pattern in data and assemble patterns of increasing complexity using smaller and simpler patterns embossed in their filters. Therefore, on a scale of connectivity and complexity, CNNs are on the lower extreme.

IV. BACKGROUND OF CNN

CNN's were first developed and used around the 1980s. The most that a CNN could do at that time was recognize handwritten digits. It was mostly used in the postal sectors to read zip codes, pin codes, etc. The important thing to remember about any deep learning model is that it requires a large amount of data to train and also requires a lot of computing resources. Because of this significant disadvantage at the time, CNNs were restricted to the postal industry and were unable to penetrate the machine Learning field.



.

Figure 2. Layers of CNN

In 2012 Alex Krizhevsky realized that it was time to bring back the branch of deep learning that uses multi-layered neural networks. The availability of large sets of data, to be more specific ImageNet datasets with millions of labelled images and an abundance of computing resources enabled researchers to revive CNNs.

3.1Inception V3

On the ImageNet dataset, it has been demonstrated that the picture recognition model Inception v3 can achieve more than 78.1 percent accuracy. The model is the result of several concepts that have been established by various scholars throughout the years. On the original paper, it is based on: "Rethinking the Inception Architecture for Computer Vision" by Szegedy, et. al. Convolutions, average pooling, max pooling, concatenations, dropouts, and fully linked layers are some of the symmetric and asymmetric building components that make up the model itself. The model makes considerable use of batch normalisation, which is also applied to the activation inputs. Using Softmax, the loss is calculated. The screenshot below displays the model's high-level diagram: The Model of Inception More details on the Inception architecture may be found. The README:

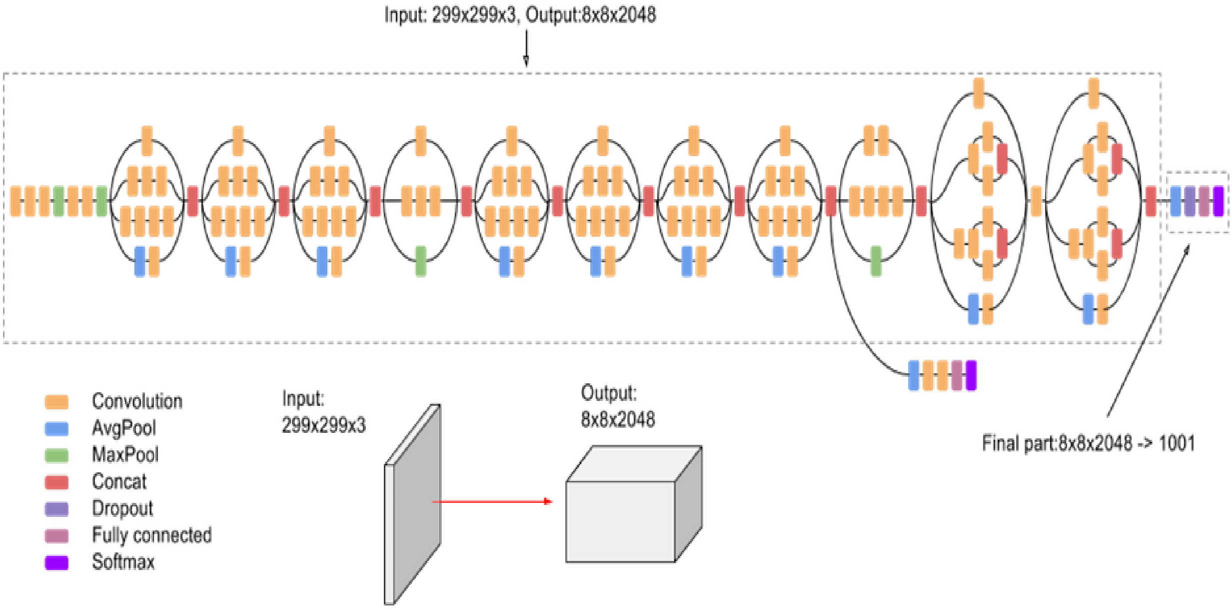


Figure 3 A high-level diagram of CNN model

Convolutional neural network model Inception-v3 has 48 layers and was pre-trained. It is a network that has been trained using a variant of the ImageNet database's more than a million photos. It is the third iteration of Google's Inception CNN model, which was first developed for the ImageNet Recognition Challenge. This pre-trained network can categorize photos into 1000 different item categories, including various animals, a keyboard, a mouse, and a pencil. The network has therefore acquired rich feature representations for a variety of pictures. The size of the network's picture input is 299 by 299 pixels. In the first section, the model collects generic features from the input photos, and in the second, it classifies the images based on those features.

3.2 Confusion Matrix

A given collection of data is classified into classes through the process of classification. In machine learning (ML), the issue is framed, the data is collected and cleaned, some essential feature variables are added (if any), the model is trained, its performance is measured, it is improved using a cost function, and finally it is ready for deployment. But how do we gauge its effectiveness? Are there any specific features to consider? Comparing the actual values to the expected values is a simple and general solution. But that doesn't make the problem go away. Let's try to analyse the issue using the well-known MNIST dataset. Examining the confusion matrix is a significantly more effective technique to judge a classifier's performance. The basic concept is to record the frequency with which occurrences of class A are categorised as class B. For instance, you would check the 5th row and 3rd column of the confusion matrix to find out how many times the classifier confused photos of 5s with 3s.

A Confusion matrix is an *N x N matrix* used for evaluating the performance of a classification model, where N is the number of *target classes*. The matrix compares the actual target values with those predicted by the machine learning model.

The following 4 are the basic termithaty which will help us in determining the metrics we are looking for.

- True Positives (TP): when the actual value is Positive and the predicted is also positive.
- True negatives (TN): when the actual value is Negative and the prediction is also negative.
- False positives (FP): When the actual is negative but the prediction is Positive. Also known as the Type 1 error
- False negatives (FN): When the actual is Positive but the prediction is Negative. Also known as the Type 2 error

For a binary classification problem, we would have a 2 x 2 matrix as shown below with 4 values:

		ACTUAL VALUES	
		Positive	Negative
PREDICTED VALUES	Positive	TP	FP
	Negative	FN	TN

Table No. 2 Confusion Matrix for the Binary Classification

- The target variable has two values: **Positive** or **Negative**
- The **columns** represent the **actual values** of the target variable

3.3 Sensitivity and specification

So our model gets about 0.72 probability of a positive test given that the patient has the disease (bottom right of the confusion matrix), that's often called sensitivity. Sensitivity is a statistical measure that is widely used in medicine that is given by the following formula

$$\begin{aligned}\text{sensitivity} &= \frac{\text{number of true positives}}{\text{number of true positives} + \text{number of false negatives}}\\ &= \frac{\text{number of true positives}}{\text{total number of sick individuals in population}}\\ &= \text{probability of a positive test given that the patient has the disease}\end{aligned}$$

The other metric is specificity, you can read it in the top left of the confusion matrix, we got about 63%. It is basically the probability of a negative test given that the patient is well:

$$\begin{aligned}\text{specificity} &= \frac{\text{number of true negatives}}{\text{number of true negatives} + \text{number of false positives}}\\ &= \frac{\text{number of true negatives}}{\text{total number of well individuals in population}}\\ &= \text{probability of a negative test given that the patient is well}\end{aligned}$$

In our example, out of all patients that has a benign, we predicted 63% of them as benign.s

IV. RESULTS

4.1 Results

Python coding is executed and the visualization of output is clearly given in below images. Considering all input images of skin lesions, we have analyzed the parameters of feature extraction are measured.

4.2 Conclusion

The heading of the Acknowledgment section and the References section must not be numbered. Skin disease is the deadliest form of skin cancer when we missed to diagnosis. It can be a serious form of cancer if it is not diagnosed at an early stage. Thus, it becomes vital to use supportive imaging techniques that have been shown to improve and facilitate the diagnosis process. These techniques are build based on strategies invented by physicians to capture the melanoma at an early stage. We have introduced a hybrid method for Melanoma, seborrheic keratosis and Nevus skin cancer detection that can be used to examine any suspicious lesion.

4.1 Limitations and Future Studies

In Future this project can be implemented through Video Data's. That is instead of using input images we can use input data's as Video clip.

V. ACKNOWLEDGMENT

We extend our sincere thanks and deep sense of gratitude to our project guide Mrs. Ananthalakshmi M.E., Ph.D., Assistant professor for providing us with the necessary inputs and guidance to complete this project within the stipulated time.

REFERENCES

1. Ali, A. 2001. M. S. Mim, M. Das and M. A. Kiber, "Feature Based Skin Disease Estimation Using Image Processing for Teledermatology," 2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2), Rajshahi, 2018, pp. 1-5, doi: 10.1109/IC4ME2.2018.8465487.
2. Jinen Daghrir 1,2, Lotfi Tlig 2, Moez Bouchouicha 3, Mounir Sayadi Melanoma skin cancer detection using deep learning and classical machine learning techniques, 2020, pp: 16
3. Alenezi, Nawal, (2019). A Method Of Skin Disease Det ection Using Image Processing And Machine Learning Procedia Computer Science. 163. 85-92. 10.1016/j.procs.2019.12.090.
4. Ashwin R Jadhav, Arun G Ghontale, and Vimal K Shrivastava. Segmentation and border detection of melanoma lesions using convolutional neural network and svm. In Computational Intelligence: Theories, Applications and Future Directions-Volume I, pages 97–108. Springer, 2019.
5. A. Anand, T. Shanthi, R.S. Sabeenian, S. Veni, Real time noisy dataset implementation of optical character identification using CNN, Int. J. Intell. Enterp. Inderscience Enterprises Ltd. 7 (1/2/3) (2020) 67–80.
6. Jeffrey Glaister, Student Member, IEEE, Alexander Wong, Member, IEEE,and David A. Clausi, Senior Member, IEEE Segmentation of Skin Lesions From Digital Images Using Joint Statistical Texture Distinctiveness,April 2014.
7. Chanki Yu, Sejung Yang, Wonoh Kim, Jinwoong Jung, Kee-Yang Chung, Sang Wook Lee, and Byungho Oh. Acral melanoma detection using a convolutional neural network for dermoscopy images. PloS one, 13(3), 2018.
8. Zaher Khazaei, F Ghorat, AM Jarrahi, HA Adineh, M Sohrabivafa, and E Goodarzi. Global incidence and mortality of skin cancer by histological subtype and its relationship with the human development index (hdi); an ecology study in 2018. World Cancer Research Journal, 6:13, 2019.
9. Dirk Schadendorf, Alexander CJ van Akkooi, Carola Berking, Klaus G Griewank, Ralf Gutzmer, Axel Hauschild, Andreas Stang, Alexander Roesch, and Selma Ugurel. Melanoma. The Lancet, 392(10151):971– 984, 2018.
10. S. Kalaiarasi , H. Kumar and S. Patra, "Dermatological Disease Detection using Image Processing and Neural Networks," International Journal of Computer Science and Mobile Applications, Vol.6 Issue. 4, April- 2018, pp. 109-118..