

International Journal of Engineering Sciences & Research Technology

(A Peer Reviewed Online Journal)
Impact Factor: 5.164



Chief Editor
Dr. J.B. Helonde

Executive Editor
Mr. Somil Mayur Shah

ABSTRACT

It is a very huge process to predict a disease based on the visual diagnosis of cell type with accuracy especially when multiple features are associated. If we get the information about the dead skin which is not visible by bare eyes. One of the major problems coming in the medical field is that doctors are not able to detect that infected part which is not visible by bare eyes and therefore they only operate the visible infected part of the skin and this may cause a major problem like hazardous disease in the future. Skin disease classification system is developed and the relationship of the skin disease image across different type of neural network is established. The collected medical images are feed into the system, and using different image processing schemes image properties are enhanced. Useful information can be take out from these medical images and pass to the classification system for training and testing using PYCHARM image processing toolbox for detection of dead skin.

KEYWORDS: Digital image processing, PYCHARM, enhancement, neural network.

1. INTRODUCTION

Despite humans evolving through generations aren't immune to diseases. And as some skin diseases look similar to rashes or insect bites are difficult to be diagnosed. Also there aren't many dermatologists to treat each and every one. Along with it being too costly for poor people, and are thus ignored. The symptoms of skin diseases being small pimples, skin bulges of various colors and sizes, these are sometimes itchy too and puss can come out through these things. There are many homemade remedies for common diseases like cold, fever, but as skin diseases aren't common and lack of knowledge about these can lead to an outbreak. In this paper we are going to learn about three different skin diseases Dermatitis, Herpes, Psoriasis and how to diagnose them.

Skin diseases:

Herpes is a virus (Herpes simplex virus) causing contagious sores, mostly occurs around the Mouth or on the genitals. Symptoms include blisters, ulcer and pain while urinating, cold sores and vaginal discharge. There is no cure for this. However there are medicines that can prevent or shorten outbreaks.

Dermatitis is nothing but skin inflammation but is not contagious. There are many types of dermatitis and depending on them the causes vary. Symptoms are rashes, blisters, redness, itchy skin and swelling. There are homemade remedies and medical treatments, and even some dermatitis clear up on their own after three to four weeks.

Psoriasis is a very common condition in which skin cells build up and form scales and itchy, dry patches. It is a chronic disease that often comes and goes and is not contagious. Common symptoms include red patches of skin covered with thick, silvery scales, itching, burning or soreness, thickened, pitted or ridged nails. There are many medical and therapeutic treatments to get over this

2. MATERIALS AND METHODS

The proposed solution is a prototype with a data base of three common skin diseases, using which a patient can self-diagnose and get some earlier knowledge of their skin disease before consulting a dermatologist. This prototype can be used in mobile hospitals in rural areas. These days everybody is connected through mobile phones. Thus, this prototype can be accessed even in the most remote locations in the country. The proposed prototype provides a non-invasive method of skin disease detection where the patient provides a picture of the infected area as an input to the prototype and any further analysis is done on this input image. No pricking or prodding of the skin is required. Image processing is used to detect skin diseases in humans. It describes the current methods employed for detecting skin diseases, proposes a digital method to detect skin diseases and states the benefits of this method. It also includes a detailed description of the transforms used to implement the proposed method. The proposed system consists of two processes: Training and Testing. Training process analyses and performs image processing on the training data base where as the testing process analyses the input image provided by the patient. Before any analysis on the input image is done, image pre-processing is done so that tall images are consistent in desired characteristics. Resolution matching is done on the image so that they are all the same size (128 x 128). 128 x 128 is chosen so that the processing time is less.

3. SYSTEM REQUIREMENTS:

PyCharm is software used for python programming is a numerical getting ready state and fourth age programming tongue. Made as a result of Math Works, its licenses compose of controls, plotting of cut off points and information, use of user interfaces, calculations with programs written in python language.

Desktop and development environment

The below fig shows the course of action of mechanical musters and work places which helps us to use Pycharm records and limits. In which image can be processed, as we know that python has libraries to work directly on project here we pre-installed libraries such as OpenCV, Matplotlib and CV2.

OpenCV (Open Source Computer Vision) is a library of programming functions mainly aimed at real-time computer vision. In simple language it is library used for Image Processing. It is mainly used to do all the operation related to Images.

Python OpenCV cv2. OpenCV-Python is a library of **Python** bindings designed to solve computer vision problems. **cv2.imread()** method loads an image from the specified file. If the image cannot be read (because of missing file, improper permissions, unsupported or invalid format) then this method returns an empty matrix.

Matplotlib is a plotting library for **Python**. It is used along with NumPy to provide an environment that is an effective open source alternative for MatLab. It can also be used with graphics toolkits like PyQt and wxPython.

NumPy is a general-purpose **array**-processing package. It provides a high-performance multidimensional **array** object, and tools for working with these arrays. It is the fundamental package for scientific computing with Python.

4. ALGORITHM

KNN algorithm is one of the simplest classification algorithms and it is one of the most used learning algorithms. KNN is a non-parametric, lazy learning algorithm. Its purpose is to use a database in which the data points are separated into several classes to predict the classification of a new sample point.

The KNN algorithm

1. Load the data
2. Initialize K to your chosen number of neighbors
3. For each example in the data
 - I. 3.1 Calculate the distance between the query example and the current example from the data.
 - II. 3.2 Add the distance and the index of the example to an ordered collection

4. Sort the ordered collection of distances and indices from smallest to largest (in ascending order) by the distances
5. Pick the first K entries from the sorted collection
6. Get the labels of the selected K entries
7. If regression, return the mean of the K labels
8. If classification, return the mode of the K labels

5. HOW IT WORKS?

Let's take a simple case to understand this algorithm. Following is a spread of red circles (RC) and green squares (GS)

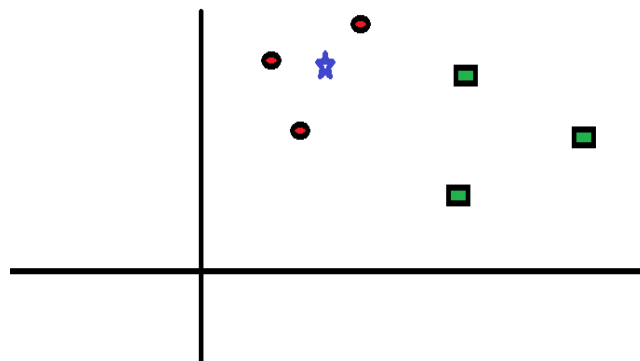


Fig.1 Sample image of KNN algorithm

You intend to find out the class of the blue star (BS) BS can either be RC or GS and nothing else. The "K" is KNN algorithm is the nearest neighbors we wish to take vote from. Let's say $K = 3$. Hence, we will now make a circle with BS as center just as big as to enclose only three data points on the plane. Refer to following diagram for more details:

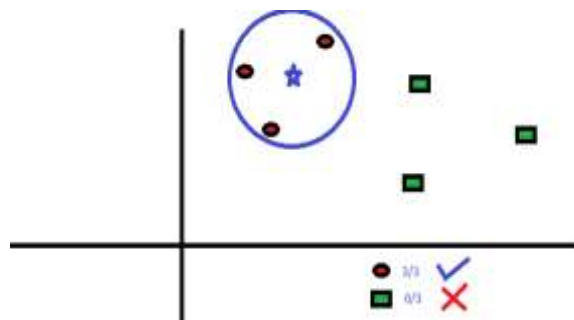


Fig.2 Resultant image of KNN algorithm

The three closest points to BS is all RC. Hence, with good confidence level we can say that the BS should belong to the class RC. Here, the choice became very obvious as all three votes from the closest neighbor went to RC. The choice of the parameter K is very crucial in this algorithm. Next we will understand what are the factors to be considered to conclude the best K.

6. HOW DO WE CHOOSE THE FACTOR K?

First let us try to understand what exactly K influences in the algorithm. If we see the last example, given that all the 6 training observation remain constant, with a given K value we can make boundaries of each class. These boundaries will segregate RC from GS. The same way, let's try to see the effect of value "K" on the class boundaries. Following are the different boundaries separating the two classes with different values of K.

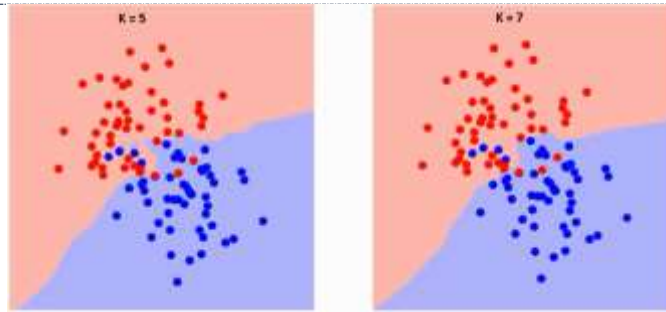


Fig3. Sample images to calculate the K values in KNN algorithm

If you watch carefully, you can see that the boundary becomes smoother with increasing value of K. With K increasing to infinity it finally becomes all blue or all red depending on the total majority. The training error rate and the validation error rate are two parameters we need to access on different K-value. Following is the curve for the training error rate with varying value of K.

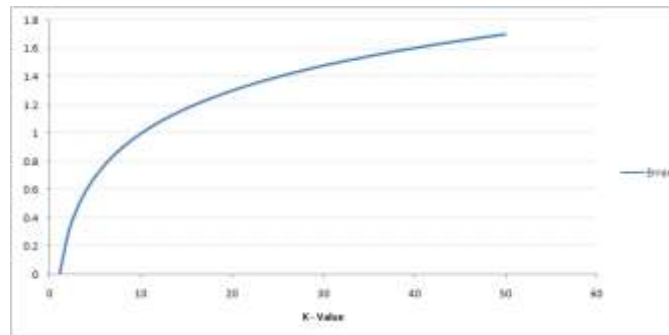


Fig4(a). Graphical representation of KNN algorithm for different values for K

As you can see, the error rate at K=1 is always zero for the training sample. This is because the closest point to any training data point is itself. Hence the prediction is always accurate with K=1. If validation error curve would have been similar, our choice of K would have been 1. Following is the validation error curve with varying value of K:

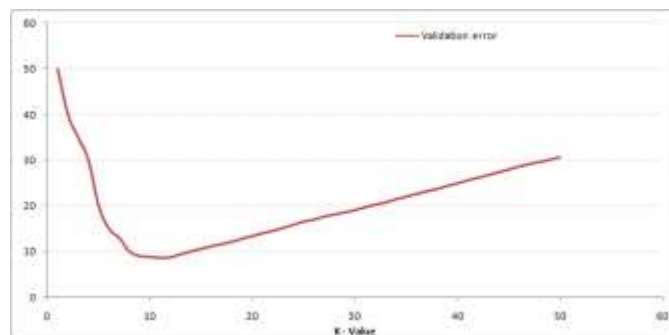
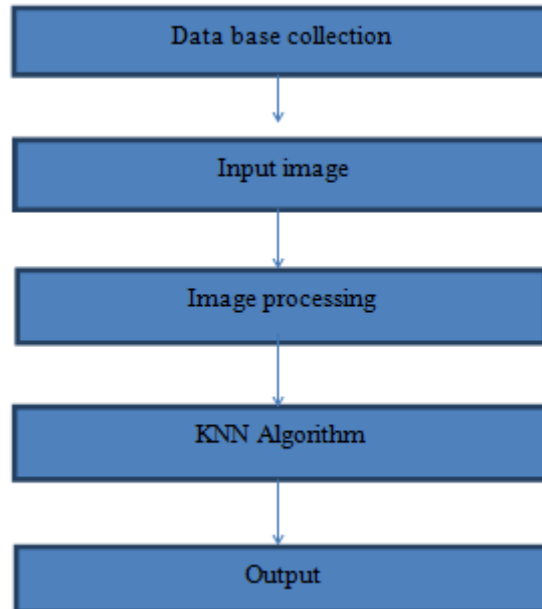


Fig4(b). Graphical representation of KNN algorithm for different values for K

This makes the story more clearer. At K=1, we were over fitting the boundaries. Hence, error rate initially decreases and reaches a minimal. After the minima point, it then increases with increasing K. To get the optimal value of K, you can segregate the training and validation from the initial dataset. Now plot the validation error curve to get the optimal value of K. This value of K should be used for all predictions.

7. FLOW CHART



8. CONCLUSION

A Computer based skin disease detection system is proposed. The diagnosing methodology uses Digital Image Processing Techniques for the classification of infected skin. Image will be processed in the pycharm software with the help of opencv, numpy and cv2. And comparing user image with the data base hence we identify the disease.

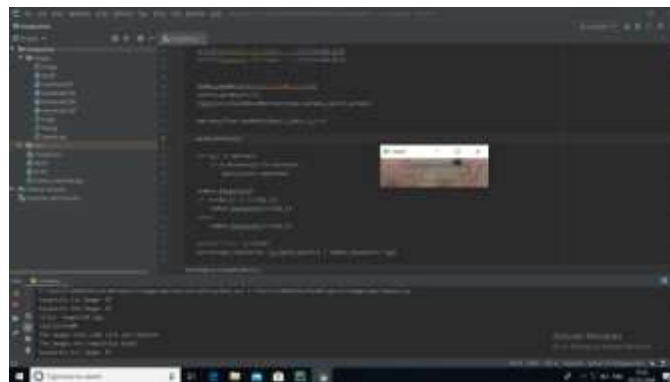


Fig 5. Resultant image of this paper(sample image)

REFERENCES

- [1] T.H. Lau and A.A. Jumaily, "Automatically Early Detection of Skin Cancer." 978-0-7695-3879-2/09 IEEE Image analysis system to detect skin disease
- [2] A.Bhardwaj and J.S. Bhatia, "An Image Segmentation Method for Early Detection and Analysis of Melanoma." 2279-0853, p-ISSN: 2279-0861. Volume 13.IOSR 2013.
- [3] Suneel Kumar and Ajit Singh. Image Processing for Recognition of Skin Diseases. International Journal of Computer Applications 149(3):37-40, September 2016
- [4] Adnan Firoze, Hong Yan, M. Ashraf ul Amin, M. Golam Kibria, and M. Shamsul Arifin, "Dermatological Disease Diagnosis using Colour-skin Images", Proceedings of the 2012 International Conference on Machine Learning and Cybernetics Xian, July, 2012, pp.1675-1680.



-
- [5] Shervan F. E, Mohammad S, Farshad T, An Innovative Skin Detection Approach Using Color Based Image Retrieval Technique, *Int. J. Multimed. Its Appl.* 4 (2012) 9. doi:10.5121/ijma.2012.4305.
- [6] Muhammad Z. A, AsgharMj, Sheikh S, Shakeel A, Diagnosis of Skin Diseases using Online Expert System, *International Journal of Computer Science and Information Security*, June 2011
- [7] <https://www.ncbi.nlm.nih.gov/books/NBK212/>

