RESEARCHERID
THOMSON REUTERS
ISSN: 2277-9655
[ICEMESM-18] Impact Factor: 5.164
IC™ Value: 3.00 CODEN: IJESS7

IJESRT
INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

GLOBAL CLASSIFICATION OF DERMATITIS DISEASE WITH K-MEANS CLUSTERING IMAGE SEGMENTATION METHODS

Prafulla N. Aerkewar[1] & Dr. G. H. Agrawal[2]
[1] Ph.D.Research scholar, B.D.College of Engg., Sevagram, Dist. Wardha,RTMNU,India
[2] Professor, Department of EE, K.D.K.College of Engineering,RTMNU, Nagpur,India

DOI: 10.5281/zenodo.1229020

ABSTRACT

The objective of this paper to presents a global technique for classification of different dermatitis disease lesions using the process of k-Means clustering image segmentation method. The word global is used such that the all dermatitis disease having skin lesion on body are classified in to four category using k-means image segmentation and nntool of Matlab. Through the image segmentation technique and nntool can be analyze and study the segmentation properties of skin lesions occurs in dermatitis disease. A skin lesion is a superficial growth or patch of the skin that does not resemble the area surrounding it. It have also been proposed that which are suitable for the processing of various images for different types of patches for various skin diseases. The skin lesion in different dermatitis diseases are different in appearance and have different properties if they looks similar in some circumstances. The main objective to classify the lesions of different dermatitis diseases based on its twelve parameters like contrast, Energy, Homogeneity etc where it would be able to classify the similar patch in to different disease. The focus is on Leprosy, Vitiligo, Psoriasis and other like birth mark or burn or boil skin patch which can be classify using K-means clustering methods and nntool of Matlab.. This review the original ideas and concepts of the above methods, because we believe this information collected and analyzed parameter are helpful to classify disease using Image Segmentation Technique

Keywords: Dermatitis, Lesion, Patches, Disease, Leprosy, Vitiligo, Psoriasis, Entropy, Autocorrelation.

I. INTRODUCTION

In day to day life, any person, rich or poor may cause patches on body due to many diseases such as Leprosy, Scar tissue, Contact dermatitis, Vitiligo, Psoriasis, Ring Worm, Eczema which may have the main issue that the skin is affected and may have the similar appearing symptoms like patches on some part of body or all over the body. As the patches on the body are looks similar but may cause by different reason. So due to similar appearance it may cause misdiagnosis and wrong treatment may be given. Most of the time, it is complicated when such a thing happens complication like reaction of a treatment may affect a life of victim.

So to avoid such misdiagnosis, image segmentation using k-means suggest the methodology for perfect diagnosis of the dermatological disease and helps in the analysis of such a disease, which can analyze the prior situation also. There are different image segmentation procedure that can be apply to build up such a powerful electronics tool which may create revolution in field of medical as well as engineering field. This effective technique work on analyzing the different parameter of skin lesion image, nntool and fuzzy logic of pervious database particular skin disease patch image.

Before starting the actual work of research in this I have studied the some available technology and present methodology related to the general image processing and image segmentation algorithm. In the procedure of analysis of skin related object, every entity being analyze as a digital image from which skin patch can be extracted and have to proceed for segmentation. So following is the process for analyzing skin images using this segmentation techniques. Out of many segmentation method K-means clustering extract all the necessary parameter that is required for diagnosis.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

II. METHODOLOGY

Database of images

The database of Images is collected via personal contacts with a patients or finding person having patch or patches on body. Also visited private skin care clinic and taken guidance regarding dermatitis disease.Total 70 people are contacted via visiting Maharogi seva Samiti, Dattapur and some private clinic. Also personally contacted some people. 14 samples of each category have been taken for analysis. Out of 70 people 30 people does have any patch on body so they come under other category. Out of remaining 40 people 30 found the patient of leprosy. Among 30 patient 14 sample are tested using algorithm and detected Leprosy correctly. Some Image samples are taken from patches image taken from internet. Sample images of Vitiligo and psoriasis also detected correctly. For analysis . Above database is segmentated using k-Means clustering method. This method uses the means clustering to cluster the dermatological lesions. By using the cluster values we classify the types of dermatological ulcers. Fuzzy K-Nearest neighbor classification is used to classify the types of dermatological ulcers. Each sample should be classified similarly to its surrounding samples, therefore, a unknown sample could be predicated by considering the classification of its nearest neighbor samples. Given a sample set values, a fuzzy *M* class partition of these vectors specify the membership degrees of each sample corresponding to each class.

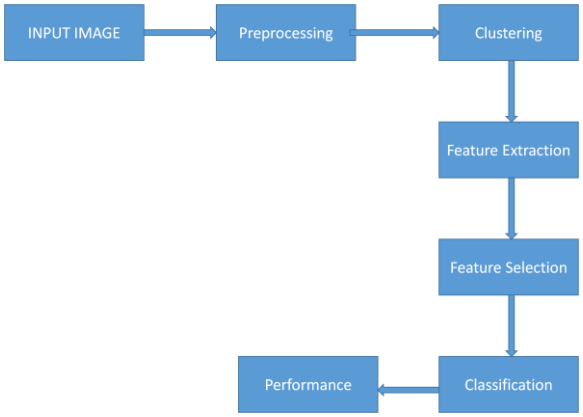


Fig 1. System architecture

Feature Extraction and Feature Selection:-

The collected database of skin lesions are process using k-means clustering and median filter and denoise image is generated. The denoised image forwarede for feature extraction and many feature like energy, Entropy, Autoco-orelation, homogeneity are extracted from database. The same 12 features are extracted from the test image to be diagnosis. Following are the 12 parameter are use for train and test images for feature extraction using below procedure.

Contrast: *Contrast* is a measure of the local variations present in an image. If there is a large amount of variation in an image the **P[i,j]**'s will be concentrated away from the main diagonal and contrast will be high (***typically k=2, n=1***).

$$C(k,n) = \sum_i \sum_j (i-j)^k P_d[i,j]^n$$

Homogeneity:- A homogeneous image will result in a *co-occurrence matrix* with a combination of high and low P[i,j]'s.

Where the ***range of gray levels*** is small the **P[i,j]** will tend to be clustered around the main diagonal.

$$C_h = \sum_i \sum_j \frac{P_d[i,j]}{1+|i-j|}$$

Entropy: Entropy is a measure of information content. It measures the randomness of intensity distribution.Such a matrix corresponds to an image in which there are no preferred gray level pairs for the

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

distance vector d. Entropy is highest when all entries in P[i,j] are of similar magnitude, and small when the entries in P[i,j] are unequal**.**

$$C_e = -\sum_i \sum_j P_d[i, j] \ln P_d[i, j]$$

Correlation: Correlation is a measure of image linearity.

$$C_c = \frac{\sum_i \sum_j [ijP_d[i, j]] - \mu_i \mu_j}{\sigma_i \sigma_j}$$

$$\mu_i = \sum iP_d[i, j], \qquad \sigma_i^2 = \sum i^2 P_d[i, j] - \mu_i^2$$

Correlation will be high if an image contains a considerable amount of linear structure.

Energy: One approach to generating texture features is to use local kernels to detect various types of texture. After the convolution with the specified kernel, the *texture energy measure (TEM)* is computed by summing the absolute values in a local neighborhood: If ***n*** kernels are applied, the result is an ***n***-dimensional feature vector at each pixel of the image being analyzed.

$$L_e = \sum_{i=1}^{m} \sum_{j=1}^{n} |C(i, j)|$$

Maximum Probability: This is simply the largest entry in the matrix, and corresponds to the strongest response. This could be the maximum in any of the matrices or the maximum overall.

$$C_m = \max_{i,j} P_d[i, j]$$

Cluster Shade

$$SHADE = \sum_{i=0}^{2G-2} (i - 2\mu)^3 H_s(i \mid \Delta x, \Delta y)$$

where

$$\mu = \frac{1}{2} \sum_{i=0}^{2G-2} iH_s(i \mid \Delta x, \Delta y)$$

Local Homogeneity, Inverse Difference Moment (IDM) :

$$IDM = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{1}{1 + (i - j)^2} P(i, j)$$

IDM is also influenced by the homogeneity of the image. Because of the weighting factor IDM will get small contributions from inhomogeneous areas. The result is a low IDM value for inhomogeneous images, and a relatively higher value for homogeneous images.

Sum of Squares, Variance :

$$VARIANCE = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (i - \mu)^2 P(i, j)$$

This feature puts relatively high weights on the elements that differ from the average value of P(i, j).

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

Cluster Prominence :

$$PROM = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\{i+j-\mu_x-\mu_y\}^4 \times P(i,j)$$

Dissimilarity :

$$\sum_{i,j=1}^{G} C_{ij}|i-j|$$

Autocorrelation :

Other statistical approaches include an autocorrelation function, which has been used for analyzing the regularity. This function evaluates the linear spatial relationships between primitives. The set of autocorrelation coefficients shown below are used as texture features:

$$C(p,q) = \frac{MN}{(M-p)(N-q)} \frac{\sum_{i=1}^{M-pN-q}\sum_{j=1} f(i,j)f(i+p,j+q)}{\sum_{i=1}^{M}\sum_{j=1}^{N} f^2(i,j)}$$

where *p, q* is the positional difference in the *i, j* direction, and *M, N* are image dimensions.

C. Algorithm:-

Algorithm is design as per the following module that perform possible analysis for applied digital image of dermatitis skin patch by calculating above discussed features.

Load Image:

In this module we initially load the test image which is to be classified. It can be done by using the uigetfile and imread functions in the matlab. To display the input image we use imshow function. After the loading the input image we will convert into gray scale image using rgb2gray function.

Preprocessing:

After the gray scale conversion we apply median filter to that image.
The Median filter is a nonlinear digital filtering technique, often used to remove noise. Such noise reduction is a typical pre-processing step to improve the results of later processing. Median filtering is very widely used in digital image processing because, under certain conditions, it preserves edges while removing noise.

Feature Extraction:

The feature values are extracted from the filtered test image . We extract theGLCM feature for the input image. The Gray Level Co occurrence Matrix (GLCM) method is a way of extracting second order statistical texture features. A GLCM is a matrix where the number of rows and columns is equal to the number of gray levels in the image.

Texture Feature Extraction:

When the input data to an algorithm is too large to be processed and it is suspected to be notoriously redundant (much data, but not much information) then the input data will be transformed into a reduced representation set of features (also named features vector). Transforming the input data into the set of features is called *feature extraction*. The features provide the characteristics of the input type to the classifier by considering the description of the relevant properties of the image into a feature space. If the features extracted are carefully chosen, it is expected that they will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input. Feature extraction involves simplifying the amount of resources required to describe a large set of data accurately. When performing analysis of complex data one of the major problems stems from the number of variables involved.

GLCM

A gray level co-occurrence matrix (GLCM) contains information about the positions of pixels having similar gray level values.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

A co-occurrence matrix is a two-dimensional array, **P**, in which both the rows and the columns represent a set of possible image values.

A GLCM **Pd**[i,j] is defined by first specifying a displacement vector **d**=(dx,dy) and counting all pairs of pixels separated by **d** having gray levels i and j.

The **GLCM** is defined by:
where $\mathbf{n_{ij}}$ is the number of occurrences of the pixel values **(i,j)** lying at distance **d** in the image.
The co-occurrence matrix $\mathbf{P_d}$ has dimension **n× n**, where n is the number of gray levels in the image.

III. RESULT AND DISCUSSION

The developed advanced image segmentation algorithm for analysis and detection of Dermatitis disease is useful for the classification of dermatitis disease using k-Means clustering method of segmentation and can classify disease into 4 category based on the database stored in nntool and using fuzzy logic it matched the parameter of test image to the database image. This algorithm extract the 12 feature 12 database and a test image of patch of disease and compare them. On the basis best matched 5 feature it classify to one of the four category. In this paper analysis of vitiligo test image is given. This algorithm first remove the noise of image and then perform k-means clustering image segmentation to extract features like energy, contrast, homogeneity etc and match with 12 featured of stored database of all images. Best 4 matched feature help to analyze and detect disease. The result of best selected 5 features of 56 database images and feature of test image is shown in table1.
Figure and Tables:-

Table1:- **Feature Selection from database Images**

	SF1	SF2	SF3	SF4	SF5
V1	0.9838	0.9914	-46.0167	1.4087	0.0337
V2	0.0788	1.6656	0.9611	-5.0706	0.9607
V3	0.0489	1.5177	-3.2019	0.9758	0.974
V4	0.9443	0.9708	1.9275	0.1129	-12.4592
L1	0.0326	-0.0143	0.9842	0.9731	1.3122
L2	1.6561	0.9764	0.9906	0.0492	11.6469
L3	0.9655	2.2651	0.9296	0.4037	-22.8034
L4	0.1246	0.9421	0.9665	2.0448	6.6486
PS1	0.102	0.8506	0.3318	1.7808	0.6369
PS2	0.799	0.9828	0.4865	0.0358	-0.2757
PS3	-0.843	0.9418	1.7357	0.1202	0.9346
O1	0.1312	1.1413	0.9595	0.8693	0.0841
O2	0.957	0.9535	1.7885	-1.8702	0.0938
O3	0.9379	-6.475	0.1256	0.9121	1.4638

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

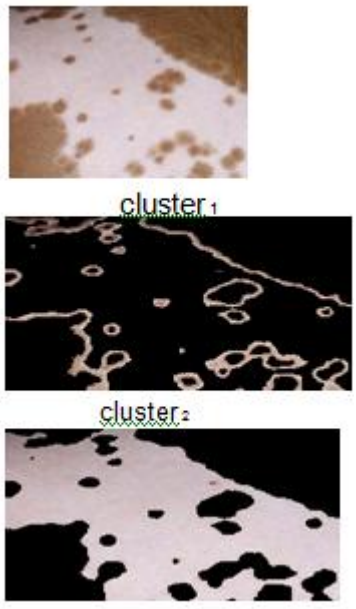


Fig 1:Test Image of Vitiligo and Segmentation

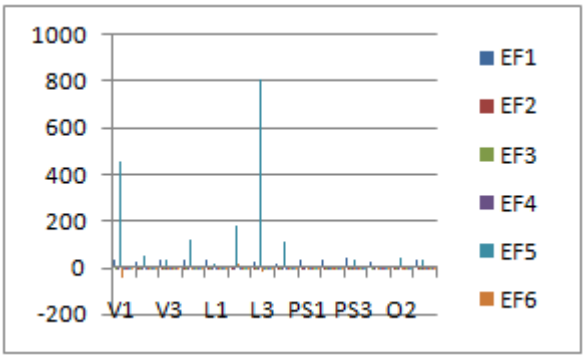


Fig A:-Extracted Features of Database Diseases

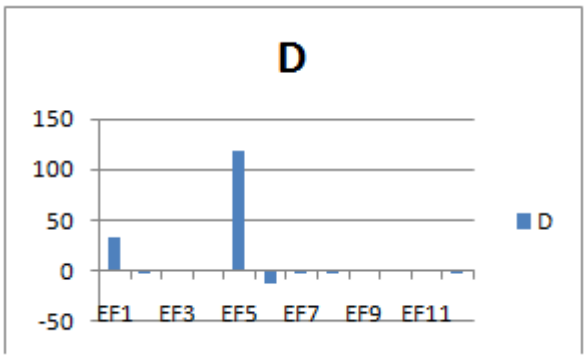
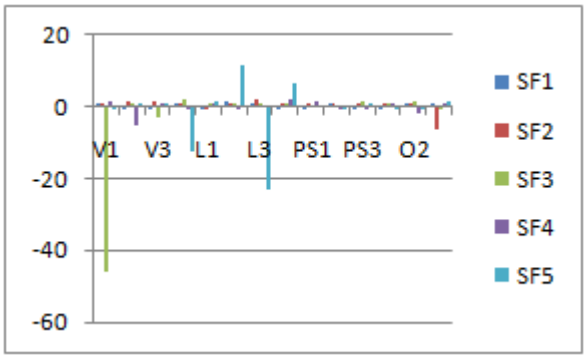


Fig B:-Extracted features of Vitiligo

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7



FigC:-Selected features of database Diseases

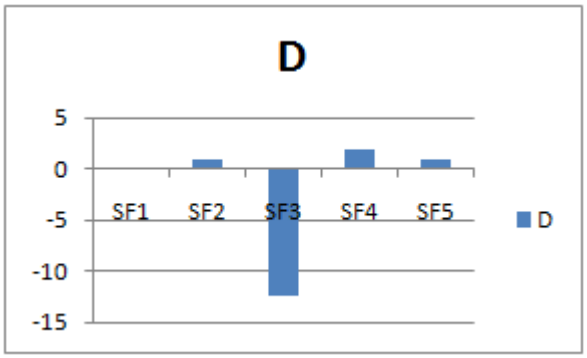


Fig D:- Selected features of Viltiligo

Feature Selection from database Images

Table 2:-Feature Selection of Test Image compared with database

20.45952	0.048813	0.953154	10.611	1.237093

IV. CONCLUSION

The advanced image segmentation algorithm is useful to classify the input test image of any disease into one of the four category like Leprosy, Vitiligo, Psoriasis or Other. It first perform the segmentation using K-means segmentation method to 12 extract features of test image and then using feature selection module of fuzzy logic it classify in to appropriate disease.

V. REFERENCES

[1] M. J. Callam, C. V. Ruckley, D. R. Harper, and J. J. Dale, "Chronic ulceration of the leg: Extent of the problem and provision of care," Brit. Med. J. (Clin. Res. ed.), vol. 290, no. 6485, pp. 1855–1856, 1985.
[2] S. Herrick, P. Sloan, M. McGurk, L. Freak, C. McCollum, and M. Ferguson, "Sequential changes in histologic pattern and extracellular matrix deposition during the healing of chronic venous ulcers," Amer. J. Pathol., vol. 141, no. 5, pp. 1085–1095, 1992.
[3] M. Witte and A. Barbul, "General principles of wound healing," Surg. Clin. North Amer., vol. 77, no. 3, pp. 509–528, 1997.
[4] K. Burnand, I. Whimster, and A. Naidoo, Pericapillary fibrin in the ulcerbearing skin of the leg: The cause of lipodermatosclerosis and venous ulceration," Brit. Med. J. (Clin. Res. ed.), vol. 285, no. 6348, pp. 1071– 1072, 1982.
[5] M. Rezvani, M. Robbins, J. Hopewell, and E. Whitehouse, "Modification of late dermal necrosis in the pig by treatment with multi-wavelength light," Brit. J. Radiol., vol. 66, no. 782, pp. 145–149, 1993.
[6] T. Pavicic and H. Korting, "Xerosis and callus formation as a key to the diabetic foot syndrome: Dermatologic view of the problem and its

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

[7] Silvio M. Pereira, Marco A. C. Frade, Rangaraj M. Rangayyan, Fellow, IEEE, and Paulo M. Azevedo-Marques, Member, IEEE "Classification of Color Images of Dermatological Ulcers", IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, VOL. 17, NO. 1, JANUARY 2013
[8] Payel Ghosh and Melanie Mitchell, "Segmentation of medical images using a genetic algorithm," in GECCO, Mike Cattolico, Ed. 2006, pp. 1171–1178, ACM.