International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Exploring Statistical Models in Dermatological Disorders Identification

M. Kavitha[1], Dr. P. Lakshmi Prasanna[2]

Submitted: 26/12/2023 **Revised**: 06/02/2024 **Accepted**: 15/02/2024

Abstract: Amid the numerous diseases to mankind, skin diseases are one such diseases which are usually caused by virus, bacteria, fungus or other organisms. There are chances are spreading of the diseases and therefore timely analysis and identification is at most important so as to minimize further complication. Advances in laser and photonic-based medical technology have helped to identify skin diseases more accurately and rapidly. Characterization plays an important role in helping to classify skin diseases. This work contributes to skin disease research based on HOG feature-based extraction and modelling the output using statistical methods. In this article we have considered Bivariate Gaussian Mixture Model (BGMM). The results derived are tested against benchmark metrics.

Keywords: *Skin Disease, Statistical Model, HOG, Quality metrics, BGMM.*

1. Introduction

The human skin accomplishes various roles such as the synthesis of Vitamin D, protecting the internal organs, preventing loss and protecting the body from damage caused by environmental pollution. In this skin, there are three layers: epidermis, dermis and hypodermis. Among these layers, epidermis layers are considered to be the external layer, which is very thin and its thickness varies from 0.05mm to 0.15mm. This is the layer which protects the human against harmful chemical and UV radiations. It also provides mechanical resistance. The second layer or the middle layer called the dermis is effective in implementing its role against protects the body from stress and anxiety. This middle layer is further divided into papillary dermis and reticular dermis. The last layer, which is also considered as the bottom layers or hypodermis is a fatty layer, it protects the body from injuries and also protects the muscles. The nerve system and red vessels branch out to connect the hypodermis to the rest of the body.

Once the virus attacks the body the external layer is highly influenced and thereby creates infections, which penetrates into the other layer causing very high damage to the individual. Many of the disease at the primary stage have its influence on the skin and they appear s rashes or identified as rough patches and red spots. These deformities if left unnoticed will spread into the body and result into even the most serious disease cancerous. Therefore, identification of these disease at the root level plays a significant role from controlling its development and narrow downing its likelihood chances of leading towards mortality. There are many methodologies that help out in defining the type of diseases and slow down its penetration. However, due to many conditions like atmospheric, environmental, virus and other infections. These microorganisms that leading to generating of the diseases is very hard to control, also most of the diseases that root out from the skin as a median to share common features. Therefore, identifying these features appropriately and classifying them into the proper medical clusters is a challenging task. Among the various methodologies in the literature most significance works are presented using machine learning techniques together with deep learning approaches. However, the challenging task raised needs a large training set for the modeling the disease and labelling them into the appropriate groups. To overcome these characters, in this article we propose statistical methodologies. The main advantage of using this statistical approach is that, appropriate modeling or labelling can be assumed based on the features identified, also these models perform more accurately in low or medium range training data availability. The total epochs of the model are built on the feature extraction. To obtain the effective features, in this article we have considered the HOG features. The main advantage is that it tries to identify the relevant images pertaining to a particular disease which is common between the individual immaterial of size, shape and the orientation. Also, in the proposed approach these features are given to the model BGMM. The main advantage of the dual features is that it maximizes the estimation of the parameters more predominantly. The parameters of the proposed model are updated so that a component likelihood of the model can e bitterly approximated. The rest of the article is explained below; Part 2 of the article provides a brief review of the literature. In section 3 the data under consideration is presented. Part 4 of the article deals with basic information based on HOG features. The BGMM is defined in section 5 of the article. Part 6 of the article is dedicated towards the experimentation and the results obtained together with

[1] *KL University, Vijayawada – 522302, Andhra Pradesh, INDIA*
ORCID ID : 0009-0006-1441-2710
[2] *KL University, Vijayawada – 522302, Andhra Pradesh, INDIA*
ORCID ID : 0000-0001-9346-811X
** Corresponding Author Email: kavithareddy2414@gmail.com*

evaluation metrics are presented in the corresponding section 7. The paper concludes with summarization in the section 8

2. Review of Literature

Several Academics have projected diverse procedures to notice the skin diseases. Here we briefly present some of the methodologies reported in the literature.

Preprocessing is often recommended as a way to improve stratification outcomes, but it decelerates down the process. To accumulate, many of the currently used policies rely heavily on preprocessing methods and defining the required factors. adapting to the different appearances of the input image which can be puzzling. Heterogeneity of dermato-scopic images often prevents accurate classification of melanoma. Various features include the presence of unwanted objects, including hairs, capillaries, colored spots, rulers, ink marks, vignettes, noise, uneven illumination, and bright images instigated by dermato-scopic image assortment methods. Many of the currently used segmentation techniques trust profoundly on numerous pre-processing stages to elude the end results of unwanted artifacts that can compromise the precise stratification of skin lesions. Correct segmentation of melanoma in dermatoscopic images may suffer significantly from occlusion caused by unwanted artifacts [1]. Destruction of many artifacts as a result of this problem, occlusion techniques have been developed in manually segmented images. The thresholding [2, 3], morphology [4, 5], filtering [6] and Dull Razor-based artifact removal [9, 7, 8] practices are debated underneath.

Dermoscopy images have truncated concentration and non-uniform illumination, which are also normally improved with prior artificial image enhancement. These enhancement techniques include contrast matching [6, 10], filtering [10, 11, 12], adaptive histogram equalization [5, 13], and constrained adaptive histogram equalization (CLAHE) [14, 15, 16]. Reconstructed from the beginning, CLAHE was popularized as an progressive enhancement technique used to produce medical images. In addition, studies have used histogram preprocessing [17], average interpretations [18], deep learning [19], multilevel corrosion [20], adaptive gamma correction [21], and Z-score transformation [22] and the Broken Vessel filter [23]. To reduce residual noise, pre-segmentation and post-processing techniques are used, typically feature extraction and image enhancement methods.

Most of the modern classification methods use deep learning based methods, especially CNN. Much of CNN's innovation comes from its ability to learn larger and more complex groups than traditional training models and algorithms. The latest model is the structure of the CNN segment, the Full Sensor Network (FCN), U-Net, SegNet, and DeepLab [22] are some examples, among them. As the ability to learn from heterogeneous data improves, researchers have recently used a deep neural network to classify melanoma. The ability of U-Net designs to handle the natural usage of the most demanding networks. The network hop network [24] can recover the pieces of information lost in the clustering process [78]. Wei et al. [25] showed that the Att-DenseUnet network is based on the combination of the dense U-Net network and the use of tools in the U-Net, to achieve better results in melanoma. Ibtehaz etal [26] an in-depth study of the U-Net model led to the design of the U-Net3 architecture.

3. Dataset

This data includes images of 23 types of skin diseases downloaded from http://www.dermnet.com/dermatology-pictures-skin-disease-pictures. The total number of images is approximately 19,500, of which approximately 15,500 are sorted in the training set and the rest are sorted in the testing set. The images are in JPEG format, consisting of 3 channels, i.e. RGB. Resolution varies by image and category, but generally these are not very high-resolution images.

These categories include acne, melanoma, eczema, seborrheic keratosis, ringworm, blistering disease, poison ivy, psoriasis, vascular tumors, etc. The largest online dermatology resource aims to provide online medical education.

4. Feature Extraction

Due to their ability to deal with variables and local spatial changes as well as simplicity in implementation, HOG methods are used [31]. In general, HOG occurs naturally by simulating how the brain processes visual information and thoughts. This has proven to be a useful method for predicting visibility and local behaviour using force gradients or edge orders in their distribution. In the process of using the features of HOG, these features are compared by finding the parameters of the intensity of the edge in the region of the wound. To achieve this goal, two logical units are defined locally, namely the cell and the block. Normally the cell size is 8×8 pixels and each block has 2×2 cells (i.e. 16×16 pixels) for each HOG display. Since the HOG detector is based on the window coding principle by default, the HOG blocks are usually arranged so that each cell can contribute to the final signal more than once. Adjacent boundary blocks are spanned by eight pixels horizontally and vertically. The Main advantage of HOG is, if we are able to find a similar image which of reduced or increased in size, based on the features, we can still identify the image more accuracy.

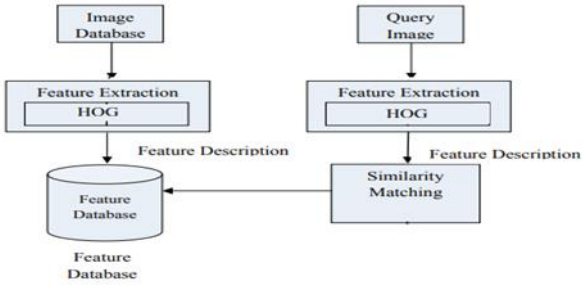


Fig. 1. Hog Architecture.

The process involved in calculate gradient histogram includes; Preprocessing (resizing). Compute gradient images, Compute the histogram of gradients in 8×8 cells, perform block normalization, Form the HOG characteristic vector, calculating the gradient magnitude.

$$\sqrt{(x-d)2+(y-d)2}$$

And then estimate the Gradient using the formula,

$$\tan^{-1}((x\text{-}d) / (y\text{-}d))$$

4.1 Algorithm for HOG

The HOG feature vector is generated by summing the gradient calculations of each pixel.

a. Create a histogram for each block using the gradient value.

b. Calculation of normalization of histograms.

5. Truncated Gaussian Mixture Model

In order to model the features and extract the diseased skin together with the type of disease, in this article, we have used a statistical modeling based on Truncated Gaussian distribution.

The main advantage of building a model-based approach using statistical parameters is that it helps to understand the finer details of the pixels more effectively based on the parameters and hence better understating of the diseases is possible.

- In model-based image segmentation algorithm, the entire image is to be viewed as a collection of image regions.
- Hence finite mixture models are utilized to characterize the pixels inside the image.

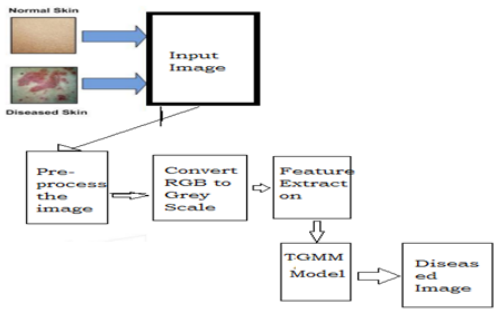


Fig. 2. Hog Architecture

The main advantage beyond the consideration of this model is that most of the diseases from the scanning, which are considered for analysis are mostly having the pixel ranges within the finite region and also these pixels are mostly truncated towards the right-side regions, therefore, in this article a Truncated Gaussian mixture model is used.

The Probability density function of the Truncated Gaussian Mixture model is given by

- The Probability Density function (P.D.F) is given by,

$$f(Z)=\frac{1}{\sqrt{2\pi\sigma}}\ e^{-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2} \quad 1$$

Where x is the pixel value of the diseased skin, μ , π and σ are the mean weighted probabilities and standard deviations respectively.

$$g(Z)=\begin{cases} -\infty\le Z<\infty \\ \dfrac{f(z)}{\int_{Z_M}^{Z_L} f(z)dz} & Z_L \le Z \le Z_M \\ Z_M \le Z<\infty \end{cases} \quad \text{----- 2}$$

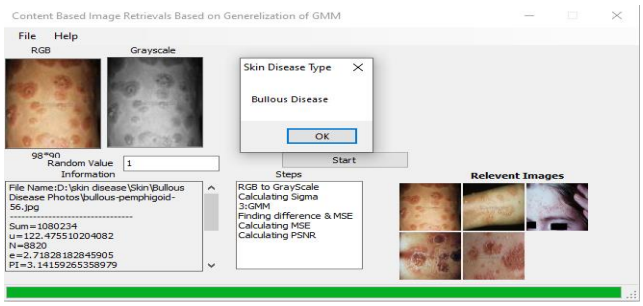
$$g(z)=\frac{f(z)}{B-A},\ z_L<z<z_M \quad \text{----- 3}$$

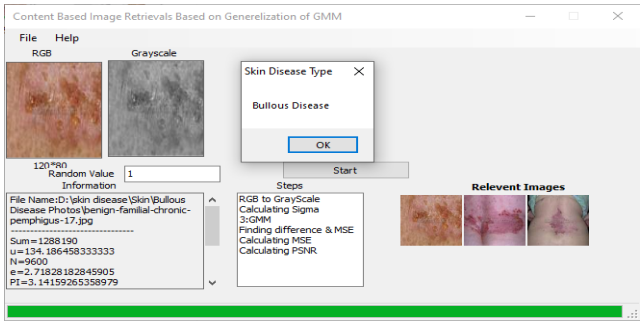
The truncating points are chosen as the minimal intensity value of the pixels and the maximum intensity value of the pixels, that is the entire image is now normalized between the limits A to B, where A denotes minimum pixel intensity, B denotes the maximum pixel intensity. And are given by the formulas.

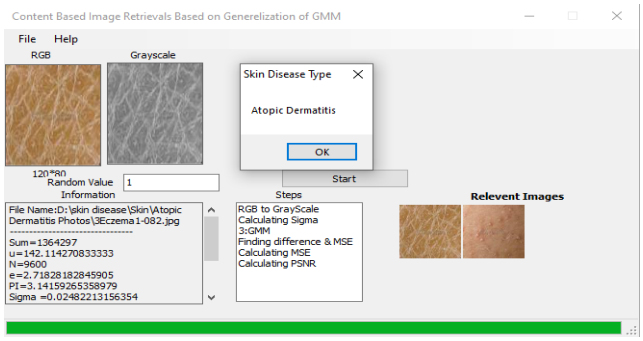
$$A=\int_{-\infty}^{Z_L}\frac{e^{\left(\frac{-1}{2}\left(\frac{z-\mu}{\sigma}\right)^2\right)}}{\sqrt{2\pi\sigma}}dz \ \text{ and } \ B=\int_{-\infty}^{Z_M}\frac{e^{\left(\frac{-1}{2}\left(\frac{z-\mu}{\sigma}\right)^2\right)}}{\sqrt{2\pi\sigma}}dz \quad \text{----- 4}$$

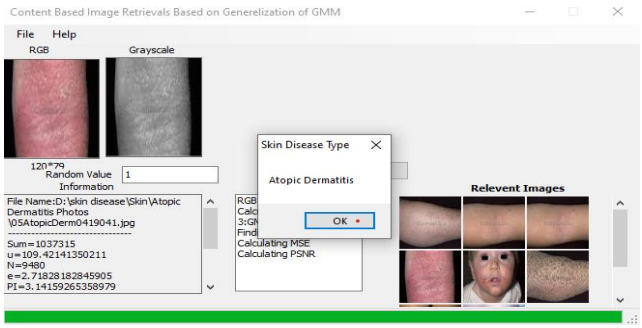
6. Experimentation and Results

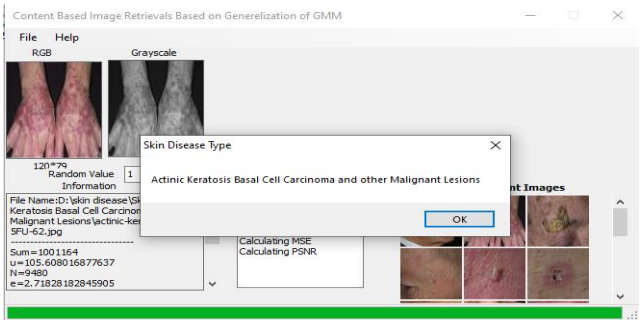
The experimentation is carried out by using different diseases from the above-mentioned dataset and the experimentation is carried out in JAVA environment, the results derived are presented in the following figures. The results are also evaluated using metrics like Precision, Recall and F-1 score, Accuracy, MSE, PSNR, FPR, FNR, PWC and the results derived are presented in the following table-1











Evaluation Metrics of different methods on DATASET		
Metrics\Methods	**GMM**	**TRUNCATED GMM**
PRECISION	0.0129	0.0167
RECALL	0.123	0.0437
ACCURACY	0.9374	0.9731
F-SCORE	0.0182	0.0385
MSE	0.0144	0.02
RMSE	0.1202	0.1415
FPR	0.0626	0.0269
FNR	0.723	0.942
PSNR	66.5661	75.917
PWC	6.2635	2.6926

Table 1. Accuracy of the Proposed model based on different metrics.

7. Conclusions

In this article, a methodology is presented for skin disease identification using Truncated Gaussian Mixture model, The experimentation is carried out using the standard dataset of skin diseases and the results are evaluate using various bench mark metrics like Precision, Recall, F-Score, MSE, PSNR, RMSE, FNR, FPR, PWC. Investigations are carried out based on HoG parameters (contrast, correlation, energy and homogeneity) showed that different types of dermatological diseases have unique peak structure can be easily predicted. The above metrics show cased the efficacy of the proposed model.

References

[1] Li, W.; Raj, A.N.J.; Tjahjadi, T.; Zhuang, Z. Digital hair removal by deep learning for skin lesion segmentation. Pattern Recognit. 2021, 117, 107994.

[2] Guarracino, M.R.; Maddalena, L. SDI+: A novel algorithm for segmenting dermoscopic images. IEEE J. Biomed. Health Inform. 2019, 23, 481–488.

[3] Jahanifar, M.; Tajeddin, N.Z.; Gooya, A.; Asl, B.M. Segmentation of lesions in dermoscopy images using saliency map and contour propagation. arXiv 2017, arXiv:1703.00087v2.

[4] Nida, N.; Irtaza, A.; Javed, A.; Yousaf, M.H.; Mahmood,

M.T. Melanoma lesion detection and segmentation using deep region based convolutional neural network and fuzzy C-means clustering. Int. J. Med. Inform. 2019, 124, 37–48.

[5] Pennisi, A.; Bloisi, D.D.; Nardi, D.; Giampetruzzi, A.R.; Mondino, C.; Facchiano, A. Skin lesion image segmentation using Delaunay Triangulation for melanoma detection. Comput. Med. Imaging Graph. 2016, 52, 89–103.

[6] Guarracino, M.R.; Maddalena, L. SDI+: A novel algorithm for segmenting dermoscopic images. IEEE J. Biomed. Health Inform. 2019, 23, 481–488.

[7] Ünver, H.M.; Ayan, E. Skin lesion segmentation in dermoscopic images with combination of YOLO and grabcut algorithm. Diagnostics 2019, 9, 72.

[8] Hu, K.; Liu, S.; Zhang, Y.; Cao, C.; Xiao, F.; Huang, W.; Gao, X. Automatic segmentation of dermoscopy images using saliency combined with adaptive thresholding based on wavelet transform. Multimed. Tools Appl. 2020, 79, 14625–14642.

[9] Garnavi, R.; Aldeen, M.; Celebi, M.E.; Varigos, G.; Finch, S. Border detection in dermoscopy images using hybrid thresholding on optimized color channels. Comput. Med. Imaging Graph. 2011, 35, 105– 115.

[10] Jahanifar, M.; Tajeddin, N.Z.; Gooya, A.; Asl, B.M. Segmentation of lesions in dermoscopy images using saliency map and contour propagation. arXiv 2017, arXiv:1703.00087v2.

[11] Joseph, S., & Olugbara, O. O. (2022). Preprocessing effects on performance of skin lesion saliency segmentation. Diagnostics, 12(2), 344.

[12] Samsudin, S. S., Arof, H., Harun, S. W., Abdul Wahab, A. W., & Idris, M. Y. I. (2022). Skin lesion classification using multi- resolution empirical mode decomposition and local binary pattern. Plos one, 17(9), e0274896.

[13] Sarker, M. M. K., Rashwan, H. A., Akram, F., Singh, V. K., Banu, S. F., Chowdhury, F. U., ... & AbdelNasser, M. (2021). SLSNet: Skin lesion segmentation using a lightweight generative adversarial network. Expert Systems with Applications, 183, 115433.

[14] Kimori, Y. (2022). A Morphological Image Preprocessing Method Based on the Geometrical Shape of Lesions to Improve the Lesion Recognition Performance of Convolutional Neural Networks. IEEE Access, 10, 70919-70936.

[15] Hossen, E. (2021). Design a Convolutional Neural Network for Melanocytic Skin Lesion Identification (Doctoral dissertation, Texas A&M University-Commerce).

[16] Abduljaleel, H. K. (2019). Deep CNN based skin lesion image denoising and segmentation using active contour method. Engineering and Technology Journal, 37(11 Part A).

[17] Guo, Y., Ashour, A. S., & Smarandache, F. (2018). A novel skin lesion detection approach using neutrosophic clustering and adaptive region growing in dermoscopy images. Symmetry, 10(4), 119.

[18] Devi, S. S., Laskar, R. H., & Singh, N. H. (2020). Fuzzy C-means clustering with histogram based cluster selection for skin lesion segmentation using non-dermoscopic images.

[19] J. Long, E. Shelhamer, and T. Darrell, ‘‘Fully convolutional networks for semantic segmentation,’’ in Proc. IEEE Conf. Comput Vis. Pattern Recognit. (CVPR), Jun. 2015, pp. 3431–3440.

[20] O. Ronneberger, P. Fischer, and T. Brox, ‘‘U-Net: Convolutional networks for biomedical image segmentation,’’ in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. Cham, Switzerland: Springer, 2015, pp. 234–241.

[21] V. Badrinarayanan, A. Kendall, and R. Cipolla, ‘‘SegNet: A deep convolutional encoder-decoder architecture for image segmentation,’’ IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 12, pp. 2481–2495, Dec. 2017.

[22] L. C. Chen, G. Papandreou, and I. Kokkinos, ‘‘DeepLab: Semantic image segmentation with deep convolutional nets, Atrous convolution, and fully connected CRFs,’’ IEEE Trans. Pattern Anal. Mach. Intell., vol. 40, no. 4, pp. 834–848, Jun. 2016.

[23] Y. Xie, J. Zhang, Y. Xia, and C. Shen, ‘‘A mutual bootstrapping model for automated skin lesion segmentation and classification,’’ IEEE Trans. Med. Imag., vol. 39, no. 7, pp. 2482–2493, Dec. 2020. /

[24] Afza, F., Khan, M. A., Sharif, M., & Rehman, A. (2019). Microscopic skin laceration segmentation and classification: A framework of statistical normal distribution and optimal feature selection. Microscopy research and technique, 82(9), 1471-1488.

[25] Kuzmina, I., Diebele, I., Jakovels, D., Spigulis, J., Valeine, L., Kapostinsh, J., & Berzina, A. (2011). Towards noncontact skin melanoma selection by multispectral imaging analysis. Journal of Biomedical optics, 16(6), 060502.

[26] Querleux, B., Guillot, G., Ginéfri, J. C., Poirier-Quinot, M., & Darrasse, L. (2019). Quantitative Magnetic Resonance Imaging of the Skin: In Vitro and In Vivo Applications. Imaging Technologies and Transdermal Delivery in Skin Disorders, 341-369.