Improving Feature Selection Capabilities in Skin Disease Detection System

Vedanti Chintawar, Jignyasa Sanghavi

Abstract: ***Feature extraction is the process of description of the input imagery into a fixed set of values. For a good feature extraction algorithm these values are sufficient in order to describe the entire properties of the image under test. There are many kind of features which can be extracted from the image, these features vary from color features, shape features, to morphological and texture features. Feature extraction is usually application dependent, and allows the application designers to incorporate various kinds of descriptors for the image under test. An optimum feature extraction system is the one which can accurately and uniquely identify each image separately via uniqueness in the feature vector. In this paper, we analyze various feature extraction techniques and identify the best features suited for the application of skin disease detection systems, and also provide some acute observations on how these techniques can be improved to further optimize the accuracy of identification.***

Index Terms: Accuracy, color, feature, morphological, shape, texture, uniquely.

I. INTRODUCTION

For characterization of pictures, the area highlights of the image are utilized to perceive the particular pictures. These highlights are catogoriesed on the distinctive key piece of picture data like shading power, edges of the articles present in picture, surface, etc [1]. The capability of highlight extraction strategy redesigns the further getting ready of an image in a manner of speaking. These highlights can be used in picture organizing, plan affirmation and recuperation. These applications require the diminished and appropriate information to achieve abnormal state of exactness. A data picture posses significant incredible and dreary information. The route toward trading this information to decrease set of highlight (or highlight vector) is known as the element assurance. Picture Analysis is a methodology in which highlight of the photos are removed and break down for further handling. It is unique in relation to other picture handling tasks like reclamation, coding and improvement. Picture examination includes the recognition, division, extraction and order strategies [2]. Feature extraction strategy is utilized to remove the features by keeping however much data as could reasonably be expected from extensive arrangement of information of picture. Productivity and viability of feature determination and extraction are extreme test these days. Various techniques are utilized to extricate features like shading, surface and shape as feature vector. The methods for feature extractions are characterized are appeared in Fig. 1.

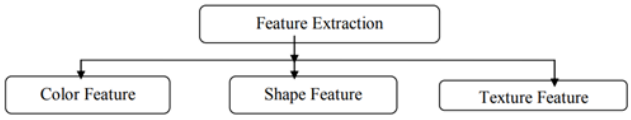

Fig 1:Types of feature extraction methods

II. LITERATURE REVIEW

There are different techniques in writing that have been proposed for the feature extraction. It is a sort of information decrease procedure. The point is to decrease the informational index of features which is profitable data present in a picture. [3]. Information present in a picture are mind boggling and high dimensional, it is an important advance to separate the instructive feature from a picture for article acknowledgment and division [4]. Other than cutting down the computational cost, highlight extraction is in like manner a strategies for controlling the affirmed chide of dimensionality. In picture examination, all highlights of the photos are separated in such a way, to the point that secure the class notice ability well [5, 6]. Thresholding is one of the exercises executed as guide errand toward concentrate the low measurement include. Shading, edges and corners are other low measurement include for request of picture. In beginning of picture examination distinctive shading spaces are used to perceive the photos. There are differing shading spaces, for instance, RGV, LUV, HSV and HMMD [7]. Shading histograms, Color Coherence vector, shading minutes based and shading correlogram are used for the extraction of highlights in pictures. These strategies rely upon expelling the mean, skewness and standard deviation of intensity of the image pixel [7]. Shading histogram was productive and speedier in recognizing shading apportionment includes in some irregular pictures meeting central necessities. However, it was unsuccessful in planning colossal course of action of pictures and no satisfies the going with criteria (Consistency, Accuracy) [8]. Out of these technique Color Moment is least intricate, littler and healthy system to evacuate the component [9]. As result

Revised Manuscript Received on May 24, 2019.

Vedanti Chintawar, M.tech Scholar, Nagpur University, Ramdeobaba college of engineering & management, Nagpur, India.
Jignyasa Sanghavi, Assistant professor, Nagpur University, Ramdeobaba college of engineering & management, Nagpur, India.

showed up in shading space extraction were not met at the craving. Picture examination limited with recognizing verification of edges and corners. Edge acknowledgment which conveys the line drawing used for low measurement highlight which portray the condition of things [10]. The idea of edge extraction include is exceptionally subject to helping conditions, a comparative power and the closeness of bustle. A corner locator figuring called FAST (Features from Accelerated Segment Test) in perspective on the SUSAN (Smallest Unvalued Segment Assimilating Nucleus) [11]. With FAST, the disclosure of corners was composed over edges as they ensured that corners are a champion among the most characteristic sorts of highlights that exhibit a strong two dimensional power change, and are along these lines particularly perceived from the neighboring centers [12]. Surface based element extraction can be designated spatial and extraordinary surface subject to their diverse focal points to use in the image getting ready. Spatial surface is definitely not hard to use and can be remove information from any shape. These element are very sensitive to fuss and mutilations. Supernatural surface is lively and requires less figuring. For compelling element ridiculous surface require square region with satisfactory size [3]. Gabor channel is commonly used to isolate the surface component for picture request. Gabor channel or wavelets portray an image by gaining the center repeat and presentation parameter. An element vector is made by getting the imperativeness at a specific repeat and bearing [4]. Shape highlight extraction systems can be requested into two social affairs as Contour based and locale based methodologies. Structure based framework register shape include just from the utmost and region based system removes highlight from the entire region. These strategies incorporate two kind of approach. First is steady strategy which does not disengage shape into subpart. It uses as far as possible to gathered the component vector. Second is Discrete (Global) Approach separates as far as possible into sub part and procedure the multi-dimensional element vector. The Shape descriptor incorporates figuring area, circularity, flightiness, genuine center point presentation, and turning imperativeness [8]. Typical systems for cutoff crumbling rely upon polygonal gauge, recurring pattern deterioration and twist fitting. In zone based frameworks, all of the pixels inside a shape region are considered to get the shape depiction, instead of simply use limit information as in structure base methodologies. Essential region based strategies utilize minute descriptors to depict shapes. Other territory based systems consolidate organize methodology, shape matrix, raised structure and media turn. Like structure based strategies, zone based methodology are progressively effective as whole shape area is considered for descriptor where every pixel of shape is considered [10]. Zhang and Lu proposed the Generic Fourier descriptor (GFD) to annihilation of multidimensional examination of a shape. The GFD is acquired by applying a 2-D Fourier change on a polarraster investigated shape picture. Neural frameworks are extraordinarily promising methodology for highlight extraction in light of pivotal parallel part of estimation. There exists wide extent of ANNs that are adequately arranged to perform dimensionality lessening of the data to make new game plan of littler and material component vector. ANN acknowledges the data as pixel and concentrates the component through the layers of framework reliant on learning. [7] presented a remarkable element extraction neural framework model of a one dimensional fundamental fragment examination (PCA) which was extended to various estimations [8, 11]. Baldi and Hornik [9] showed that three-layer auto-associator frameworks performed better contrasting with PCA. Direct PCA can't deal capably with non-straight data subspace. In subsequent [10, 11], auto-associator frameworks with multilayers were shown extraordinary execution in non-straight dimensionality decline including preeminent surfaces [12]. It is in like manner possible to use a mix of direct subspaces to evaluated a non-straight subspace [13]. There exists an approach by social affair the relative component into little course of action of highlight pack to decrease high-dimensional data. Clustering secure the overabundance highlight of the data while arranging it.

III. CATEGORIZATION OF FEATURE EXTRACTION TECHNIQUES

Shape based features comprises of edges, corners, districts of intrigue and edges. Edges are focuses where there is a limit (or an edge) between two picture locales. When all is said in done, an edge can be of practically subjective shape, and may incorporate intersections. Practically speaking, edges are typically characterized as sets of focuses in the picture which have a solid slope greatness. Besides, some basic calculations will at that point chain high angle guides together toward structure a progressively total portrayal of an edge. These calculations typically place a few limitations on the properties of an edge, for example, shape, smoothness, and slope esteem. Locally, edges have a one-dimensional structure.

The terms corners and intrigue focuses are utilized to some degree reciprocally and allude to point-like features in a picture, which have a nearby two dimensional structure. The name "Corner" emerged since early calculations originally performed edge location, and after that broke down the edges to discover fast alters in course (corners). These calculations were then grown with the goal that express edge identification was never again required, for example by searching for elevated amounts of ebb and flow in the picture angle. It was then seen that the alleged corners were likewise being identified on parts of the picture which were not corners in the customary sense (for example a little splendid spot on a dull foundation might be distinguished). These focuses are every now and again known as intrigue focuses, yet the expression "corner" is utilized by custom.

Masses give a corresponding depiction of picture structures as far as locales, rather than corners that are more point-like. By the by, mass descriptors may regularly contain a favored point (a neighborhood limit of an administrator reaction or a focal point of gravity)

which implies that many mass indicators may likewise be viewed as intrigue point administrators. Mass finders can distinguish territories in a picture which are too smooth to be in any way identified by a corner detector. Consider contracting a picture and after that performing corner recognition. The indicator will react to focuses which are sharp in the shrunk picture, however might be smooth in the first picture. It is now that the contrast between a corner finder and a mass indicator turns out to be to some degree obscure. To a substantial degree, this refinement can be helped by including a proper idea of scale. By the by, because of their reaction properties to various kinds of picture structures at various scales, the LoG and DoH mass indicators are likewise referenced in the article on corner location.

For lengthened items, the thought of edges is a characteristic instrument. An edge descriptor registered from a dark dimension picture can be viewed as a speculation of an average pivot. From a pragmatic perspective, an edge can be thought of as a one-dimensional bend that speaks to a hub of symmetry, and moreover has a quality of nearby edge width related with each edge point. Sadly, be that as it may, it is algorithmically harder to separate edge features from general classes of dark dimension pictures than edge-, corner-or mass features. All things considered, edge descriptors are habitually utilized for street extraction in airborne pictures and for extricating veins in medicinal pictures. Low-level features can be extricated specifically from the first pictures, though abnormal state feature extraction relies upon low dimension features

A shading piece (or shading feature) is a segment of a production, (for example, a paper or magazine) that centers predominantly around impressions or portrayals of the topic. It for the most part underlines the elucidating perspectives. The shading histogram is invariant to pivot of the picture on the view hub, and changes in little advances when turned generally or scaled. The development of the shading histogram is an immediate procedure, including examining the picture, allotting shading esteems to the goals of the histogram, and building the histogram utilizing shading parts as records. The histogram calculation has O(x,y) multifaceted nature for pictures of size x × y . The unpredictability for a solitary picture coordinate is direct, O(n), where n speaks to the quantity of various hues, or goals of the histogram. The shading histogram measure is fundamentally littler than the picture itself, expecting shading quantization. In this manner it is the best decision for shading feature extraction.

A picture surface is a lot of measurements determined in picture preparing intended to evaluate the apparent surface of a picture. Picture surface gives us data about the spatial course of action of shading or forces in a picture or chose district of a picture. Picture surfaces can be falsely made or found in normal scenes caught in a picture. Picture surfaces are one way that can be utilized to help in division or characterization of pictures. For increasingly exact division the most valuable features are spatial recurrence and a normal dim dimension. To dissect a picture surface in PC designs, there are two different ways to approach the issue: Structured Approach and Statistical Approach.

For picture content depiction, shape is an essential visual feature and one of the crude feature. Shape content depiction can't be characterized precisely in light of the fact that estimating the likeness between shapes is troublesome. For shape feature extraction, edge coordinating is imperative. Edge coordinating: Uses edge discovery systems, for example, the Canny edge recognition, to discover edges. Changes in lighting and shading as a rule don't have much impact on picture edges.

Many researchers have used different algorithms to perform the task of feature extraction in order to improve overall system accuracy. Some models use only shape features, only color features or only texture features, while some others use a hybrid combination of them. In our analysis and conclusion section, we found that the classification accuracy of hybrid features is very high as compared to the accuracy of individual features.

IV. PROPOSED WORK

In our work we are using the following blocks for the system,

- Hair removal
- Segmentation using OTSU
- Feature extraction

This is partial implementation of our proposed work, but in this implementation we have extracted the hybrid features based on the analysis done by us in this text. These features are further used for classification, which we are working upon. The hair removal block applies edge detection on the input image, and then removes all these edges from the image, due to the fact that hairs are generally contained in edges of the image. Furthermore, this operation is repeated for several iterations, so that we can get a better pre-processed result.

The pre-processed image is given to an OTSU segmentation block, which produces an OTSU mask. This mask is applied on the image in order to get the segmented image, which is further used for feature extraction. The feature extraction unit finds out hybrid features, namely Circularity, High Luminance Scale, FAST comers, Solidity, Shape Skewness and Border Skewness, which describe the image completely. The following results show the outputs of segmentation of our technique.

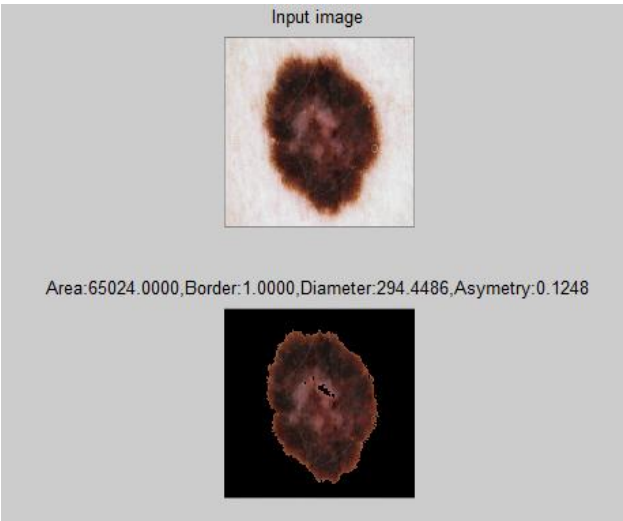

Fig2: Cancerous mole segmented image

Fig 3: Non-cancerous mole detected image

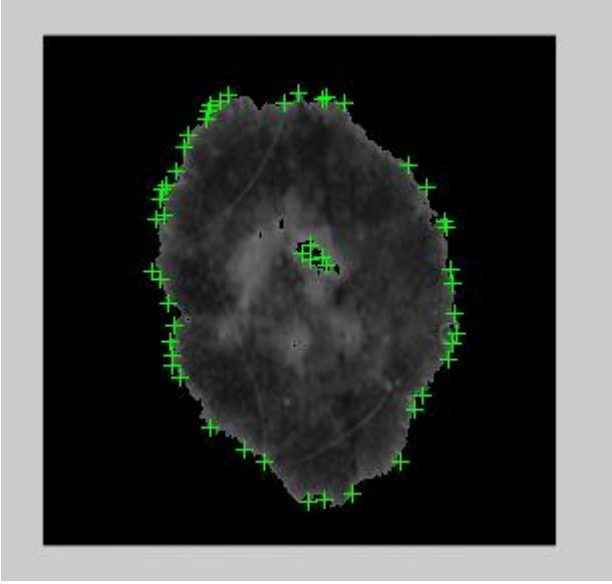

Fig 4: FAST Corners for affected image

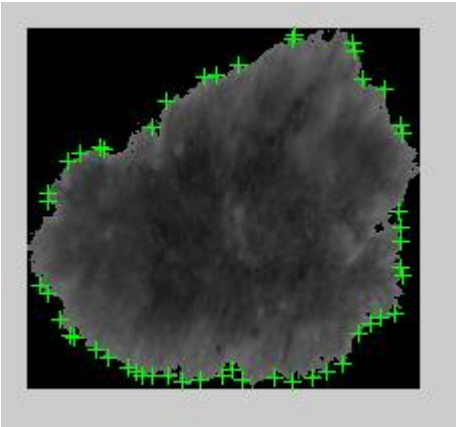

Fig 5: FAST corners for non-affected image

From the above results we can be assured that the segmentation and feature extraction techniques are giving good variant features across both affected and non-affected images.

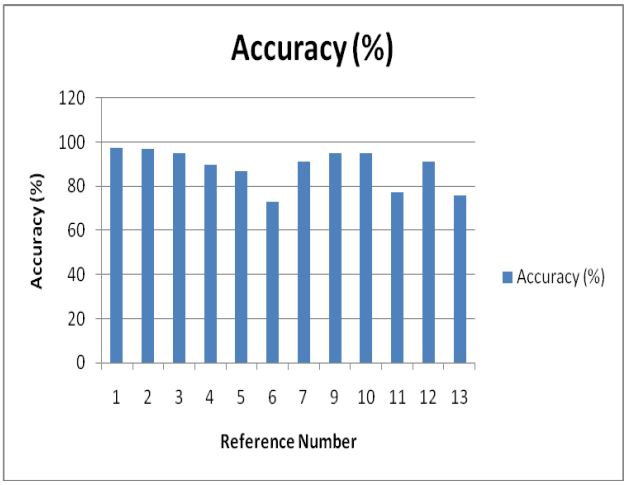

Fig 6: Accuracy comparison.

Ref no.	Feature name	Classifier	Accuracy(%)
1.	Shape (Asymmetry, Border, Color and Diameter)	Neural networks	97.51
2.	Texture + color	SVM	97
3.	GLCM	SVM	95
4.	Shape (Asymmetry, Border, Color and Diameter)	Threshold based	
5.	Hybrid Features (Circularity, High Luminance Scale, FAST comers, Solidity, Shape Skewness and Border Skewness)	SVM	86.67

6.	Convolutional neural network (CNN)	CNN	73
7.	Lesion Feature Network	CNN	91.4
9.	Orientation Features	SVM	95
10.	Fully convolutional residual network	CNN	94.9
11.	Siamese Convolutional Neural Networks	CNN	77.45
12.	Image Net VGG	CNN	91
13.	Color Histogram + Sparse Coding + LBP	Deep residual network	76

I. Descriptive comparison of accuracy.

From the above analysis we concluded that the accuracy of hybrid features which includes shape, colors and textures along with neural networks, support vector machines or convolutional neural networks is better as compared to other methods. Thus, researchers must use these pre-defined methods in order to design the base systems for skin disease detection.

V. FUTURE WORK

The current work produces high accuracy of more than 95% when features of color, texture and shape are combined, but the delay and storage needed for such a complex feature set evaluation is very high, so researchers can apply machine learning and artificial intelligence algorithms to these combined feature sets and develop a feature selection algorithm which produces the best features for giving highest accuracy with minimum delay, so that the system can be used for real time applications.

REFERENCES

[1] Wilson F. Cueva, F. Muñoz, G. Vásquez., G. Delgado, "Detection of skin cancer "Melanoma" through Computer Vision", 2017 IEEE XXIV International Conference on Electronics, Electrical Engineering and Computing (INTERCON), IEEE 2017.

[2] Farzam Kharaji Nezhadian, Saeid Rashidi,"Melanoma skin cancer detection using color and new texture features",2017 Artificial Intelligence and Signal Processing (AISP), IEEE 2017.

[3] Uzma Bano Ansari,Tanuja Sarode,"Skin Cancer Detection Using Image Processing", International Research Journal of Engineering and Technology (IRJET), Volume: 04,Issue: 04, Apr-2017.

[4] Shivangi Jain, Vandana jagtap, Nitin Pise,"Computer aided Melanoma skin cancer detection using Image Processing",International Conference on Intelligent Computing, Communication & Convergence (ICCC-2015), Elsevier - 2015.

[5] Suleiman Mustafa, Akio Kimura,"A SVM-based diagnosis of melanoma using only useful image features", 2018 International Workshop on Advanced Image Technology (IWAIT), IEEE 2018.

[6] Andre Esteva, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau, and Sebastian Thrun,"Dermatologist-level classification of skin cancer with deep neural networks", Vol 542, p-115-127, Springer Nature Feb-2017

[7] Yuexiang Li, Linlin Shen,"Skin Lesion Analysis towards Melanoma Detection Using Deep Learning Network", p-1-16 Sensors 2018.

[8] Yading Yuan, Ming Chao, Yeh-Chi Lo, "Automatic Skin Lesion Segmentation Using Deep Fully Convolutional Networks with Jaccard Distance", IEEE Transactions on Medical Imaging, Volume: 36, Issue: 9, Sept. 2017, IEEE 2017.

[9] Supriya Joseph, Janu R Panicker,"Skin Lesion Analysis System for Melanoma Detection with an Effective Hair Segmentation Method", IEEE International Conference on Information Science (ICIS), IEEE Aug-2016.

[10] Lequan Yu, Hao Chen, Qi Dou, Jing Qin, Pheng-Ann Heng, "Automated Melanoma Recognition in Dermoscopy Images via Very Deep Residual Networks", IEEE Transactions on Medical Imaging, Volume: 36, Issue: 4, April 2017.

[11] Yu-An Chung, Wei-Hung Weng, "Learning Deep Representations of Medical Images using Siamese CNNs with Application to Content-Based Image Retrieval",31st Conference on Neural Information Processing Systems (NIPS 2017).

[12] Haofu Liao,"A Deep Learning Approach to Universal Skin Disease Classification", Graduate Problem Seminar - Project Report, University of Rochester, 2015.

[13] N. C. F. Codella, Q.B. Nguyen, S. Pankanti, D. A. Gutman, B. Helba, A. C. Halpern, J. R. Smith,"Deep learning ensembles for melanoma recognition in dermoscopy images", IBM Journal of Research and Development, Volume: 61, Issue: 4/5, July-Sept. 2017.

AUTHORS PROFILE

Vedanti R. chintawar has done Bachelor of Engineering in Computer technology from Gondwana University , India in 2017 and Currently Pursuing P.G in Computer Science and Engineering from Ramdeobaba College of Engineering and Management , Nagpur , India and Research Paper work focuses on Image Processing

Jignyasa Sanghavi is Assistant Professor at Shri Ramdeobaba College of Engineering and Management, Nagpur, India. She received B.E. degree and M.Tech degree in Computer Science and Engineering from Rashtrasant Tukodoji Maharaj Nagpur University. She has teaching experience of 8.5 years in Computer Science and Engineering department. She is member of ACM. Her research interest includes Image Processing, Parallel Computing, Databases and Algorithms. She has published 10 research papers in various conferences, workshops and International Journals.