Data Mining Techniques to Deductthe Human Skin Diseases

[1]Mrs.K.SHUNMUGA PRIYA, [2]Dr.Mrs.V.SELVI

[1]Research Scholar, [2]Assistant Professor
[1]Department of Computer Science,
[1]Mother Teresa Women's University, Kodaikanal, Tamilnadu, India

Abstract**—***a large amount of Information can be extracted in the form of data that are hidden in images. The extraction process is possible by using data mining techniques. In this paper, we will use some system based on the decision tree for mining a data and Image processing. This system will be used for categorization of human skin diseases. we will can try the decision tree and digital image processing techniques to sense skin diseases using some built in features of digital image processing for a human skin. The major steps involved in this process are: pre-processing, features extraction and decision tree classifier. This system enhances the classification process to be more perfect. A dermatologist can make use of this accurate decision tree classification phase for classify with the affected skin images. This system is designed and can be implemented on MATLAB and is tested with the various images which are available in databases.*

Keywords—Artificial Intelligence, Data Mining, Image Processing, Skin Diseases, Decision Tree.

I. INTRODUCTION

Knowledge Discovery and Data Mining are rapidly developing areas of research that are at the intersection of several discipline, including figures, databases, AI, visualization, and high-performance and parallel computing. People in business, science, medicine, academia, and government collect such data sets, and a number of other commercial packages now offer general purpose Knowledge Discovery and data processing tools. An important information Discovery and Data Mining goal is to "turn information into knowledge." For example, knowledge acquire through such methods on a medical database could be published in a medical journal. Knowledge acquired from analyze a financial or marketing database could revise business practice and influence a management school's curriculum[1]. Valuable information are often hidden in images, however, few researches discuss data processing on them[2]. Further development of a computer-assisted diagnosis is associated with the use of new intelligent capabilities such as data processing, which permit discovering the relevant knowledge for image analysis and diagnosis from the database of image descriptions. The application of data mining will help to get some extra knowledge aboutspecific features of different classes and the way in which they are expressed in the image. The decision tree method has been used to categorize the medical images for diagnosis An additional advantage of data mining application for the decision of medical or other tasks is on the long-run the opportunity for creation of fully-automatic image analysis systems that could be very significant and helpful in the case of lack of knowledge for decision-making. Classification is one among the foremost common applications for data mining [3]. Classification is a crucial sort of knowledge extraction, and may help make key decision [4]. It corresponds to a task that happens frequently in lifestyle. For example, a hospital might want to classify medical patients into those that are at high, medium or low risk of acquiring a particular illness, an opinion polling company may wish to classify people interviewed into those that are likely to vote for every off variety of political parties or are undecided, or we might need to classify a student project as distinction, merit, pass or fail. Decision Tree has been a powerful and attractive tool in the field of classification, mainly because they produce easily interpretable and organized results. In general computationally efficient and capable of handling noisy data[4]. Image mining is quite just an expansion of knowledge mining to image domain. It is an interdisciplinary challenge that pulls upon proficiency in computer vision, digital image processing, image extraction, data mining, machine learning, databases, and artificial intelligence [5].

II. BACKGROUND

One of the problems that arise in any collection of data is the classification. The classification is important in many fields. One of the fields is image categorization (medical image classification). The classification process depends on the principal of knowledge mining, especially image mining, and may be done using one technique in data processing that's called decision tree. The decision tree is going to be wont to classify medical skin diseases images. This classification can done employing a system consists of the subsequent phases on images: pre-processing, features extraction and decision tree classifier. This system is extremely important and useful for physicians to detect skin diseasesof the human then to work out the acceptable medicine for that disease.

In this paper, we will propose a system to detect and recognize skin diseases in human. The system will differentiateamong normal skin and infected skin. The distinction is predicated on using data processing techniques, specifically the choice tree, and image processing to extract the important features utilized in classification process. This system are often employed by physicians to form the popularity process for skin diseases more accurate counting on a database of skin images.

III. RELATEDWORKS

At the present conditions of computerized skin diagnosis systems, there are some workarounds. Some Rights based cases are closed which are still under and research developments, and it's determined some restrictions and barriers in those and thus this solution tries to beat the issues that exist along with different take.

a) An automated system for deducting disease conditions of human skin

In this model, the condition of the skin condition is identified by evaluating disease images by using grey normalized symmetrical simultaneous occurrence stencils (GLCM) method. The proposed system is employed in an efficient and economical for the automated recognition of skin diseases. this method is beneficial for the skin to scale back the error with diagnosing. Another is that the first test for patients in rural areas, where the great doctors are missing. The system works with relational databases to the storage of implying the necessity for textual skin images. this method also can work for same variety of images directly over feature vectors [6].

b) Image-based diagnosis method

This system primarilyfocus on diagnose diseases of skin that are occurred by viruses and bacteria. This system used image of the various area and people images are taken then machine learning techniques and image processing applied to coach the pc to diagnose the skin problem. this can be an optional diagnosis method for these skin diseases and it's safe and no risks, side effects or inconveniences from the patient perspective. It also gives advantageous to doctors because it's fast and may be implemented in various ways (mobile phones, computers and digital cameras). And also it will be safely utilized by non-Specialized medical personnel. First, the patients were clinically analyzed by an expert (dermatologist/medical doctor), then laboratory tests were conducted to foresee and make sure the skin disease. The doctor then apprehended some images from the patients whose results showed that they'd a viral or bacterial infection [7].

c) Expert System for Diagnosis of Skin Diseases

This system is developed for diagnosing skin diseases which permit user to spot diseases of the human skin to supply advises or medical treatments during a very short fundamental quantity. The system uses technologies like image processing and data processing for the diagnosis of the disease of the skin. The image of disease is taken and it must be subjected to numerous processing for noise eliminating and enhancement of image. This image is straight away segmentation of images using threshold values. Finally data processing techniques are accustomed identify the skin problem and to suggest medical treatments or advice for users [8].

d) The Development of Online Pediatric Skin Diseases Diagnosis System

A system enables the user recognize skin diseases confronted of youngsters through the net and make user for advice or Treatments within the shortest period of your time. The is predicated on law and therefore the ahead was used a sequence heuristics engine for development from the system. With this technique, to help and allows the user to Recognition of Pediatric. Dermatology through the net and offer helpful Proposal the user[9] .

IV. ARCHITECTURE OF THE PROPOSED SOLUTION

An Expert System for the diagnosis of disease of the skin. In our situation we'd like resolved teams of the pattern opening in styles having a system to retrieve all the pictures to a selected level as styles is way more obvious to use no noisy and unwanted data then we extract specific features just like the colors characteristics will be accustomed create a model classification the realm. With this technique model eventually classification a predictable disease of the a brand new image from the skin condition. Building on over again on such diseases prophesied system would ask a user constitute and is predicated on the system and also the answer is decides to the kind of disease is that the case with used again within the data processing technique. Eventually refers or a medical the ideas that are supported the expected to as a results of a skin condition therapy system.

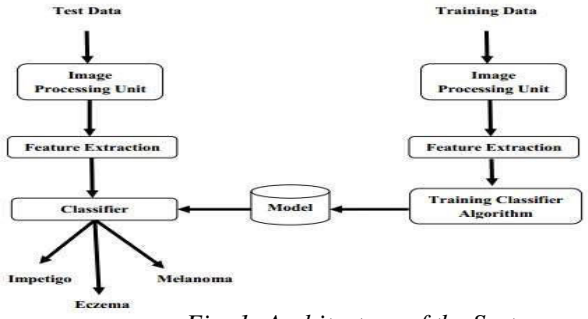

Fig. 1: Architecture of the System

This paper there's many searching in a picture data preparing and mining of understand what are simplest method of more accurately using the techniques and getting the most effective solutions for skin diseases to spot and diagnose. Search consists of three portions.

1. Processed image, and dividing and a bonus essence.
2. Model of classification and attribution disease.
3. The medical treatment or advice suggestions.

V. THE EXPERTSYSTEM

A. IMAGEPROCESSING

The image processing is that the main part within the design process at an expert system. within the beginning needed to define the affected the world of disease of the skin that a part of The image processing of this process must be it's been implemented. Imaging provide techniques and processes in creating images of physical body or samples for clinical purposes life science or for knowledge discovery. Digital image processing involves the screening of a neighborhood for processing and saving this region to a location (possibly a file) for processing [01]. The process of adjusting digital image in order to produce the result with more suitable for display or further image analysis for instance, we can remove noise or sharpen or brighten of digital images, making easier to spot key feature[00]. The detection of skin disorders and their evaluation is split into some basic steps[01]. The image processing and image filtering techniques. The mathematical concepts of convolution and therefore the kernel matrix are wont to apply filters to signals, to perform functions like extracting edges and reducing unwanted noise the Sobel operator and Gaussian smoothing filter are implemented in MATLAB to realize the functions previously mentioned, and are evaluated on test images. the consequences of adding Gaussian and 'salt and pepper' noise before filtering are thenpresented as an estimate to signals that occur in real applications. Pre-processing images refers to prior to applying other filters is shown to produce improved results when extracting edges from images with noise[13]. Gaussian filtering G is used to blura image and remove noise. In one dimensional the Gaussian function is given as follows:

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$$

Using algorithms in the image we can cut or remove the Background from the image. We separate a particular infected area from the image. It will be done with the help of retailinga part of imagefrom the picture. This feature extraction is to sendtheexact unitof datafor diagnosis.The extraction method is shows in Figure 2 Image Processing

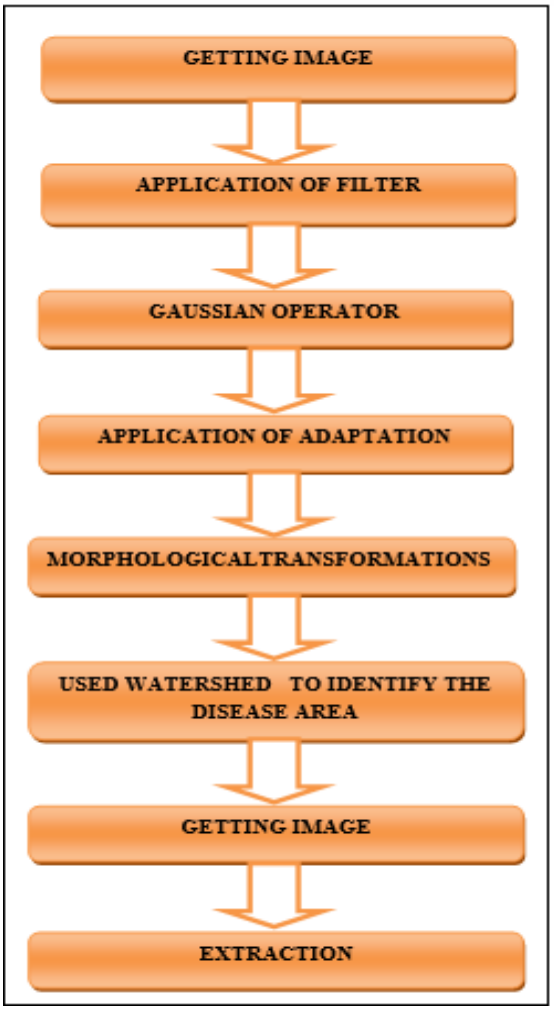

Figure 2 : Image Processing

B. DATA MINING

Data mining is that the extraction of hidden predictive information and unknown data, patterns, relationships and knowledge by exploring the massive data sets which are

difficult to seek out and detect with traditional statistical methods. data processing it's powerful technology whichwill discover most vital information from the info warehouse of the organizations. it's a really crucial step that collectively examine great deal of routinely data[14]. data processing is that the process of extracting meaningful information from large database. In Medical field the matter may arise within the era data processing has vital role to predict and diagnosis the disease in early stage with the utilization of machine learning tool[15]. data processing and KDD Process Furthermore before conducting a review and analysis work, we first need to understand what data processing is because the main area of the study declared that data processing came into existence within the middle of 1990"s and appeared as a strong tool that's suitable for fetching previously unknown at tern and useful information from huge dataset. Various studies highlighted that data processing techniques help the info holder to research and find out unsuspected relationships among their data which successively helpful for decisions making stated that a knowledge mining may be a technique that deals with the extraction of hidden predictive information from an outsized database. It uses sophisticated algorithms for the method of sorting through large amounts of knowledge sets and picking out relevant information. data processing the Analysis step of the Knowledge Discovery in Databases process, or KDD a comparatively young and interdisciplinary field of computing , is that the process of extracting Patterns from large data sets by combining methods from statistics and AI with management coined that the term Knowledge Discovery in Databases, or KDD for brief , refers to the broad process of finding knowledge in data, and emphasizes the "high-level" application of particular data processing methods. it's of interest to researchers in machine learning, pattern recognition, databases, statistics, AI , knowledge acquisition for expert systems and data visualization. Furthermore researcher had presented an overview of the steps of the KDD is that the last stage as shown in Figure 3[16].

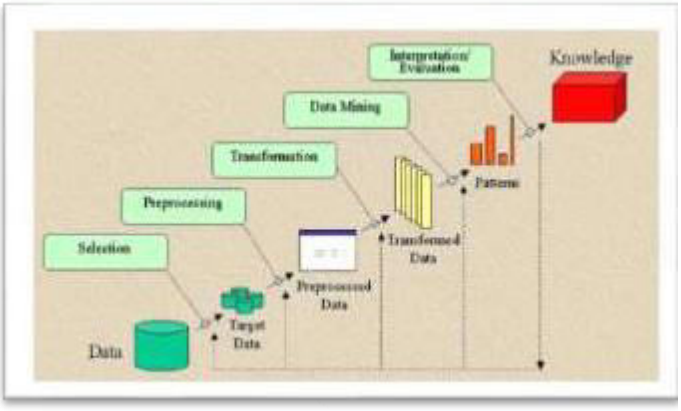

Fig. 3: Stages of data Discovery Process (KDD)

Data mining algorithms are needed in almost every step in KDD process starting from domain understanding to knowledge evaluation. it's necessary to spot and evaluate the foremost common data processing algorithms implemented in modern .Determining performance of knowledge mining solutions require much time and energy . data processing algorithms may concede better results for one sort of problems while others could also be suitable for various ones. the necessity is for algorithms with very high accuracy as diagnosis is taken into account as a big yet obscure task that must be administered precisely and efficiently[17]. The descriptive data extraction model is to seek out patterns within the data and identifies the correlation between traits actress within the data. In comparison of these intended to mining predictive input model largely to predict future results.

C. DECISION TREE

Decision tree produces identification or regression products all the way through a tree construction. It splits a dataset into smaller subsets while at an equivalent time connected decision tree is incrementally developed. the ultimate result's a choice tree with decision nodes and leaf nodes[18]. Decision tree may be a predictive data processing techniques often utilized in clinical medicine to simply visualize, and understand immune to noise in data. And is applicable in both regression and association data processing tasks[19]. Tree induction decision trees are wont to predict and/or classify tree are two phases, the training and implementation. During the training phase, the info set is partitioned iteratively[20]. In this, the target concept is represented within the form a tree, where the tree is made by using the principle of recursive partitioning. In this, attributes are selected as a partitioning attribute or as a node supported the knowledge gain criteria then the method continues repeatedly for each child node until all attributes are considered and a choice tree is made . Some pruning techniques may further be considered in order that the dimensions of the tree is reduced and therefore the overfitting is thereby avoided[21]. Be wont to create a model for predicting rating assessing disease of the skin expert system predicted the results may be a major task to create a model for predictive results within the system expert suggests classification algorithms to predict disease of the skin.

VI. CONCLUSION

The main of this paper is to focus is that the use of a proposed program and using image processing can predict and resolve enormous applications. The discovery of knowledge from large amounts of data considering both

the results of which is obtained by way of images and therefore the way the questionnaire addressed. That means, we'll have a system made up with many of the questions prepared by the system from the user. And the system to get help answer that was given to the above questions for the skin disease diagnosis. From this paper can manage information from a skin disease, symptoms, medical treatmenthandling and suggestions and generate a statement to demonstrate the report of the skin disease.

REFERENCES

[1] Q. Luo, "1. Luo, Q. on Knowledge Discovery and Data Mining WKDD 2008. (2008).on Knowledge Discovery and Data Mining WKDD 2008," no. January,2008.

[2] C. Science, "Image Processing and Image Mining using Decision Trees *," vol. 1003, pp. 989–1003, 2009.

[3] P. Smitha, L. Shaji, and M. G. Mini, "A Review of Medical Image Classification Techniques," *Int. Conf. VLSI, Commun. Instrum.*, no. Icvci, pp. 34–38, 2011.

[4] S. Nandgaonkar, R. Jagtap, P. Anarase, B. Khadake, and A. Betale, "Image mining of textual images using low-level image features," in *Proceedings - 2010 3rd IEEE International Conference on Computer Science and Information Technology, ICCSIT 2010*, 2010, vol. 9, pp.588–592.

[5] C. L. Devasena, "An Experiential Survey on Image Mining Tools , Techniques and Applications," vol. 3, no. 3, pp. 1155–1167,2011.

[6] N. Yadav, V. Kumar, and N. Director, "Skin Diseases Detection Models using Image Processing: A Survey Utpalshrivastava," *Int. J. Comput. Appl.*, vol. 137, no. 12, pp. 975–8887,2016.

[7] C. L. Aruta, C. R. Calaguas, J. K. Gameng, M. V. Prudentino, A. Anthony, and C. J. Lubaton, "Mobile-based Medical Assistance for Diagnosing Different Types of Skin Diseases Using Case-based Reasoning with Image Processing," vol. 3, pp. 115– 118,2015.

[8] A.A.L.C.Amarathunga,E.P.W.C.Ellawala,G. N. Abeysekara, and C. R. J. Amalraj, "Expert System For Diagnosis Of Skin Diseases," *Int. J. Sci. Technol. Res.*, vol. 4, no. 1, pp. 174–178,2015.

[9] M. M. Yusof, R. A. Aziz, and C. S. Fei, "The Development of Online Children Skin Diseases Diagnosis System," vol. 3, no. 2, pp. 231–234,2013.

[10] D. Okuboyejo, O. Olugbara, and S. Odunaike, "Automating Skin Disease Diagnosis Using Image Classification," *Proc. World Congr. Eng. Comput. Sci.*, vol. II, pp. 23–25,2013.

[11] S. Kumar, "Image Processing for Recognition of Skin Diseases," vol. 149, no. 3, pp. 37–40,2016.

[12] H. Motion, "Detection and Evaluation of Skin Disorders By One of Photogrammetric Image Analysis Methods," vol. XXXIX, no. September,pp.537–542, 2012.

[13] S. Kim, "Applications of Convolution in Image Processing with MATLAB," Univ. Washingt., pp. 1–19, 2013.

[14] S. Kaur and R. K. Bawa, "Future Trends of Data Mining in Predicting the Various Diseases in Medical Healthcare System," Int. J. Energy, Inf. Commun., vol. 6, no. 4, pp. 17–34, 2015.

[15] D. Sindhuja, "A Survey on Classification Techniques in Data Mining for Analyzing Liver Disease Disorder," vol. 5, no. 5, pp. 483–488, 2016.

[16] H. Leopord, W. KiprutoCheruiyot, and S. Kimani, "A Survey and Analysis on Classification and Regression Data Mining Techniques for Diseases Outbreak Prediction in Datasets," Int. J. Eng. Sci., pp. 2319–1813, 2016.

[17] P. Mahindrakar and M. Hanumanthappa, "Data Mining In Healthcare : A Survey of Techniques and Algorithms with Its Limitations and Challenges," Int. J. Eng. Res. Appl., vol. 3, no. 6, pp. 937–941, 2013.

[18] M. Preethi, "Analyzing Human Skin Texture using Machine Learning Approaches," Int. J. Comput. Appl., vol. 136, no. 1, pp. 975–8887, 2016.

[19] K. Danjuma and A. O. Osofisan, "Evaluation of Predictive Data Mining Algorithms in ErythematoSquamous Disease Diagnosis," no. Cvd, p. 10, 2015.

[20] T. D. Kale, "Informatics Application of Data Mining Techniques to Discover Cause of Under- Five Children Admission to Pediatric Ward : The Case of NigistEleni Mohammed Memorial Zonal Hospital," vol. 6, no. 1, pp. 1–14, 2015. [21]M. Kumarasamy, "An Improving Performance of Data Mining Using Intelligent Agent System Methods," vol. 6495, no. 12, pp. 104–110, 2016.