International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 03 Issue: 03 | Mar-2016 www.irjet.net p-ISSN: 2395-0072

Review of Image Processing and Machine Learning Techniques for Eye Disease Detection and Classification

Mr. Langade Umesh[1], Ms. Malkar Mrunalini[2], Dr. Swati Shinde[3]

[1]*Student(BEIT),PCCoE College of Engineering , Pune*
[2]*Student(BEIT),PCCoE College of Engineering , Pune*
[3]*Professor, Dept. of Information Technology, PCCoE College, Pune .*

Abstract - *The ever increasing amounts of patients data in the form of medical images, imposes new challenges to clinical routine such as diagnosis, treatment and monitoring. Image mining is the process of searching and discovering valuable information knowledge of data. It is applied on image processing and machine learning. Image processing is having significance for disease detection on medical images. These disease recognition and classification are specific to human organ and image type. With help of image processing and machine learning techniques it is possible to automate and/or assist physicians in clinical diagnosis. This paper describes the application of various image processing and machine learning techniques for detection of eye diseases.*

Key Words: **Medical Image Mining, Image Processing, NB, KNN, SVM, AUC, DCT, HMM and PCA approaches.**

1. INTRODUCTION

Lots of people in rural and semi-urban areas get suffered from eye diseases such as Diabetic Retinopathy, Glaucoma; Age based Macular Degradation and etc. Here using some methods and technique takes symptoms and take image of disease eye into consideration and will detect and classify. Using this we can minimize the need of the doctor and it will also notify the patient about his disease and its solution. In this paper we are trying to understand various eye disease detection and its classification using some image processing and machine learning techniques. In this paper we cover image processing techniques such as noise suppression, sharpening, contrast enhancement, image segmentation, etc. and as well as machine learning technique such as NB, KNN, SVM, AUC, HMM, etc.

2. MEDICAL IMAGE PROCESSING

To perform the medical image processing and disease detection, a sequence of image processing operations are required to improve quality of acquired image and to perform the detection. These processing stages are :

Enhancement: Medical images are often affected by noise due to interference and other factors that affect the imaging processes. Image enhancement is the improvement of image quality to increase the perception of information in images for medical specialists. This enhancement id achieved using following methods which are listed below:

a. Noise suppression
b. Sharpening
c. Contrast Enhancement
d. Image Segmentation
e. Feature extraction
f. Statistical analysis
g. Classification based on a classifier

These steps are helps in improving the quality of the image and algorithms used in these methods are depends upon that condition or situation.

Image Processing: Various image processing techniques used in automated recent diagnosis and analysis of various eye disease are Enhancement, Registration, Image Fusion, image Segmentation, Feature extraction[6], pattern matching, classification, Statistical measurements and analysis[2].

Image Recognition: The goal of image recognition is the classification or structural description of images. Image classification involves feature detection property measurement; image description involves, in addition, segmentation and relational structure extraction [7]. Some significant ideas in each of these areas are reviewed in the following paragraphs. Historically, the techniques used have usually been developed on heuristic grounds, but there is increasing interest in deriving optimum techniques based on models for the classes of images to be analyzed.

3. MACHINE LEARNING

Naive Bayes: The NB classifier[1] has been widely and successfully applied for research on medical data[1]. NB classifier is one of the highly effective and efficient classification algorithms, through comparison of NB with other popular classifiers such as Logistic regression, nearest neighbour, Decision Tree, Neural Network and Rule Based on medical data sets. The Classifiers are compared depending on the area under the Receiver Operating Characteristics (ROC)[1] curve[3]. kononenko(2001) considered NS as a benchmark algorithm that in any medical domain has to be tried before any other advanced method. Compared to other classifiers, Naive Bayes is simple, computationally efficient, requires relatively little data for training, need not to have lot of parameters and it is naturally robust to unavailable and noisy data.

KNN: K - Nearest Neighbor[4] is a kind of instance-based learning, where the function is only locally approximated and all computation is referred until classification[1]. This technique is called lazy learning because, it does not need any training or minimal training phase. All the training data is needed only during the testing phase and this technique uses all the training data so that if we have a large data set then we need special method to work on part of data which is the algorithmic approach[1]. Although classification is the primary application of KNN, we can also use it for density estimation also. The k-nearest neighbor algorithm is one of the the simplest algorithm of all machine learning algorithms.[4] KNN classification[4] was formulated from the requirement to perform several analysis when reliable parametric estimates of probability densities are not known or difficult to determine.

SVM (Support Vector Machine): In machine learning support vector machines (SVMs also known as Support Vector Networks) are supervised learning models with correlated learning algorithms that learns data and determines patterns, used for regression and classification analysis[4]. Given a set of training examples, each marked as referring to one category for one of two categories, an SVM training algorithm creates a model that divides new examples into one category or the other devising it a non-probabilistic binary linear classifier[1]. An SVM model is a representation of the example as points in space assigned so that examples of the different categories are divided [1]. In addition to performing linear classification, SVMs can expeditiously perform a nonlinear classification using the trick called the kernel trick, implicitly mapping their into high-dimensional feature spaces.

HMM (Hidden Markov Model): We describe an embedded HMM [1] - based approach for face recognition and detection that uses an effective set of observation vectors gained from the 2D-DCT coefficients. The embedded HMM can sculpture the two dimensional data finer than the one-dimensional HMM and is computationally
less difficult than the two-dimensional HMM [1]. This model is well suited for face images since it exploits important facial characteristics, structure of "states" inside each of these "super states".

DCT: In the field of images processing and recognition, discrete cosine transform (DCT)[1] and linear discrimination are two widely used techniques. Based on them, we present a new face and palmprint recognition approach in this paper. It first uses a two-dimensional separability judgement to select the DCT [1] frequency bands with favourable linear separability. Selected bands, it extracts the linear discriminative features by an improved Fisherface method and perform the classification by the nearest neighbour classifier [1]. We detailed analyse theoretical advantages of our approach in feature extraction. It can significantly improve the recognition rates for face and palmprint data and effectively reduce the dimension of feature space.

PCA: A new technique coined two-dimensional component analysis is done for image representation. As opposed to PCA, 2DPCA [1] is based on image matrices rather than 1D factor so; the image metrics does not need to be transformed into a factor prior to for feature extraction. Instead an image covariance metrics is constructed directly using the original image matrices and its eigenvector are derived for image feature extraction [1]. To test 2DPCA and evaluate the performance, a series of experiments were performed on their face image databases: ORL, AR and Yale face databases. The Experimental result indicates that the extraction of image features is computationally very efficient using 2DPCA than PCA.

AUC: The AUC[1] is the a part of performing matric of logistic regression is a widely for used evaluation matric for binary classification problems, like predicting a disease is there or not.

4. LITERATURE SURVEY

1. A Study on Retinal Disease Classification and Filtration Approaches. Parul1, Neetu Sharma2 [1]

Image processing is having is significance for disease on medical images. Theses disease recognition and classification are specific to human organ and image type. One of such disease class includes detection of retinal disease such as glaucoma detection or diabetic detection.

The paper has defined study on disease recognition approaches such as SVM , DCT ,HMM, and PCA approaches. This paper also defines the image processing operation applied to filter the medical image and to perform disease area segmentation. To perform image processing and disease detection, a series of image processing operations are required to improve the quality of acquired image and to perform the detection.

2. Review of Image Processing Technique for Glaucoma Detection , Preeti , Jyotika Pruthi [1]

The review paper describes the application of various image processing techniques for automatic detection of glaucoma. Glaucoma is a neurodegenerative disorder of optic nerve, which causes partial loss of vision. Large number of people suffers from eye disease relies upon examining retinal fundus image fusion, image segmentation, feature extraction, image enhancement, morphology, pattern matching, image classification, analysis and statistical measurements.

3. A Novel Approach for Classifying Medical Images Using Data Mining Techniques. Alamelu Mangai, Jagadish Nayak and V. Santhosh Kumar [1]

In this paper, a novel approach for automatic classification of fundus images is proposed. The method uses image and data pre-processing techniques to improve the performance of machine learning classifiers. Further a discretization method is proposed to improve the accuracy of the classifiers. Experiments were done on retinal fundus images using the proposed method on three classifiers Naive Bayes NB, k nearest neighbour KNN and support vector machine SVM. Results in terms of accuracy of classification and area under ROC curve AUC show that NB outperform the other classifiers as per the proposed method.

4. An Improved k Nearest Neighbor Classifier Using Interestingness Measures for Medical Image Mining. J. Alamelu Mangai, Satej Wagle, and V. Santhosh Kumar [1]

In this research a medical image classification framework using data mining techniques is proposed. It involves the feature extraction, feature selection, feature discretization and classification. In the classification phase, the performance of the traditional kNN k-nearest neighbor classifier is improved using a feature weighting scheme and a distance weighted voting instead of simple majority voting. Feature weights are calculated using the interestingness measures used in association rule mining. Experiments on the retinal fundus images show that the proposed framework improves the classification accuracy of traditional kNN from 78.57% to 92.85%.

5. Detection of Retinal Hemorrhage in Fundus Images by Classifying the Splat Features Using SVM. Inbarathi.R, Karthikeyan.R [1]

The objective of our proposed work is to detect retinal hemorrhage for automatic screening of DR using Support Vector Machine (SVM) classifier. To detect retinal hemorrhage, retinal fundus images are taken from Messidor dataset. After pre-processing, retinal images using pixel of same color and intensity, the image is

partitioned into non-overlapping area that covers the entire image. Splat and GLCM feature are extracted to improve the classification accuracy. In order to classify the given input images, different classes must be represented using relevant and significant features with the help of selection method that is processed by filter and wrapper approaches. Then hemorrhage affected retina is detected by SVM classifier. Finally classification accuracy is compared with K-Nearest Neighbor (KNN) classifier.

5. PROPOSED SYSTEM

The proposed system will present the design of an expert system that aims to provide the patient and diagnosis of the eye diseases. The eye has always been viewed as a tunnel to the internal workings of the human body structure. Symptoms of the eye, like many of the times there is a swollen eye or red eyes, or redness of eyes. These symptoms can be easily identified by looking at the eyes.

Here we are designing an expert system that will take the image of the patient's eye. The patient will also select or insert the other symptoms of the disease that are experienced by him. Many of the times it has been observed that, the two different diseases can have the same image for the disease but the symptoms of the disease may vary. So, the system will take all these symptoms and image of an eye into consideration and will generate the appropriate result that will notify the patient about his disease. This system can

minimize the need of the doctor or it can be used where the availability of the doctor is less.

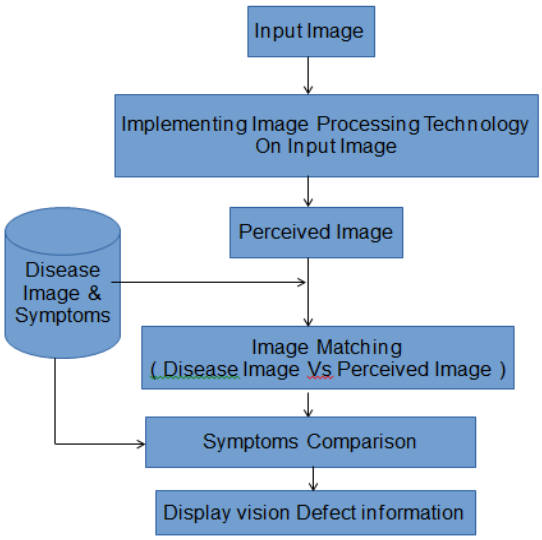

Figure-1: System architecture

An implementation of the image enhancement algorithm for the accuracy of the image of the patient's eye. Preparation of the dataset which store various images of the diseases and it will be used while testing the real data. An implementation of the image comparison algorithm which will be using the image of the eye and the dataset that consisting of the various images of diseased eye. Feature extraction from the image comparison and testing for the symptoms of the disease entered by the user.

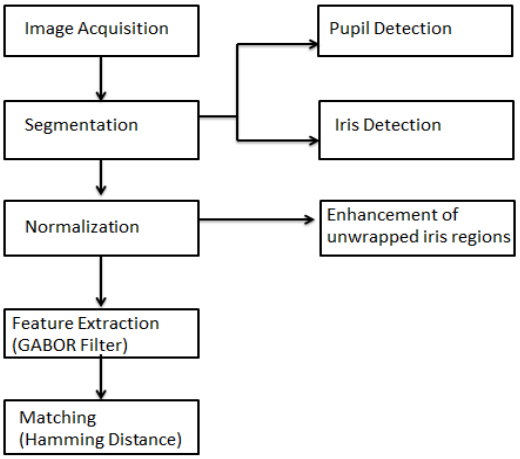

Fig -2: Image processing techniques [11]

1 Image Acquisition:

Fig-3: Image Acquisition

Image acquisition is the very first step in the iris recognition system. The size and colour of iris of every person is different therefore it is very difficult to recognize. The acquisition process [8] [11] produces different results for the same persons due to the different lighting effect, different positioning and different separation of distance.

2. Image Segmentation[11]:

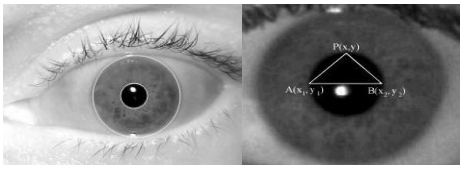

Fig-4 : Image Segmentation

The Image segmentation is the process of consume all the different parts of the eye like pupil diameter[7], eyelashes, eyelid, sclera, retina part of eye, inner and outer part of the eye and removes all irrelevant details to increase the efficiency and same time on recognition process. Inner boundary and outer boundary of typical iris can be taken as circles. The two circles are usually not to be co-centric.

3. Image Normalization:

Normalization refers to preparing a segmentation of input image for the feature extraction process. Due to the variation of the illumination and the associated elastic deformations in the iris texture the size of pupil may change and may interface with the results of pattern matching[11]. For proposed system we are going to use Daugman's normalized model [9]. In this model the process will produce the iris regions, which has the same dimensional of the same captured image of iris under the different lighting effects.

4. Feature Extraction:

The input image features are extracted by circular symmetric filter method and grabber filter method [5]. This method describes the relationship between low frequency information and high frequency information. It improves the efficiency and correctness of the eye disease recognition

system with the help of image pre-processing and feature representation. The inner and outer boundary limits of an input iris localized by filtering, edge detection and Hough transform[11].

5. Matching

This will be the last step for proposed system. In this step the encoding process will extract the feature from iris image and used for the matching process[11]. The encoding process will encode the pattern of disease image into 3002 bit iris code. After the encoding process the Hamming Distance method will be used to matching process, this method gives the measure in two bit patterns that how many bits are same. The purpose of hamming distance reduce the errors motive by false accept and false reject rate.

6. CONCLUSIONS

The paper has defined a survey on review of medical image processing and machine learning techniques for detecting and classifying eye disease images for disease recognition. Proposed system will use all the image processing techniques and algorithms mentioned in the paper. Eye disease detection & recognition can be achieved by proposed system with the use of Image Processing and Data Mining techniques.

REFERENCES

[1]A Study on Retinal Disease Classification and Filtration Approaches. Parul1, Neetu Sharma2.

[2] Review of Image Processing Technique for Glaucoma Detection, Preeti, Jyotika Pruthi

[3] A Novel Approach for Classifying Medical Images Using Data Mining Techniques. Alamelu Mangai, Jagadish Nayak and V. Santhosh Kumar

[4] An Improved k Nearest Neighbor Classifier Using Interestingness Measures for Medical Image Mining. J. Alamelu Mangai, Satej Wagle, and V. Santhosh Kumar

[5] Detection of Retinal Hemorrhage in Fundus Images by Classifying the Splat Features Using SVM. Inbarathi.R, Karthikeyan.R

[6] Pattern Recognition and Image Processing King-Sun Fu, Fellow, IEEE, and Azriel Rosenfeld, Fellow, IEEE

[7] Proposed Iris Recognition Algorithm through Image Acquisition technique, Sachin Gupta Assistant Professor, Department of Computer Science and Engineering Maharishi Markandeshwar Group of Institutions, Karnal, Haryana, Ashish Gagneja Associate Professor, Department of Electronics and Communication Engineering. Maharishi Markandeshwar Group of Institutions,Karnal, Haryana

[8] Image Understanding for Iris Biometrics: A Survey Kevin W. Bowyer, Karen Hollingsworth, and Patrick J. Flynn Department of Computer Science and Engineering University of Notre Dame Notre Dame, Indiana 46556

[9] Iris localization using Daugman's algorithm, Oad Percy Ahmad Waqas

[10] A Review on Feature Extraction Techniques of Iris, R. B. Patil, R. R. Deshmukh Department of Computer Science & IT, Dr. Babasaheb Ambedkar Marathwada University,Aurangabad

[11] Generation of Iris Template for recognition of Iris in Efficient Manner Ruchika Gupta1 Harish Saini2 CSE, Kurukshetra H.No-1805 Sector-15 H.B.C Sonipat, Haryana (India).