International Journal of Latest Engineering and Management Research (IJLEMR)
ISSN: 2455-4847
www.ijlemr.com || Volume 03 - Issue 03 || March 2018 || PP. 38-42

A Study on Biomedical image Classification using Data Mining Methods

P. Thamilselvan[1], Dr. J. G. R. Sathiaseelan[2]
[1]*Research Scholar,*
Department of Computer Science,
Bishop Heber College (Autonomous)
Tiruchirappalli, TN, India
[2]*Associate Professor and Head,*
Department of Computer Science,
Bishop Heber College (Autonomous),
Tiruchirappalli, TN, India

Abstract: Image mining is an important region in data mining. In which the changes in medical image innovation have extraordinarily expanded for image considers. This paper proposes contemplate on some transcendent information mining algorithms for image classification. Genetic algorithm, Decision tree and k-means algorithms are taken for this specialized study to locate the biomedical image classifications. We examine the utilization of various information mining strategies, for example, decision tree, k-means algorithm and genetic algorithms. The result of this examination is to discover the execution algorithms in view of the classification exactness in biomedical images. In addition, the trial examine show the utilization of bio medical image classification.

Key Words: Data Mining; Image Classification; Biomedical Images; Classification Accuracy

I. INTRODUCTION

The great incidence of biomedical images in developed countries has basically increased. This paper presents some experimental study in biomedical image classification. We studied different data mining methods such as decision tree, genetic method and k-means method. The biomedical image standard allows to storing texture descriptions along with images [1]. The morphology of the ductal network can potentially provide valuable information to developing biomedical images and assist in diagnosis anomalies [2]. The mixture models is too distinct the exudates the structures from contextual followed by and back propagation technique edge detection technique [3-4]. The size of sub biomedical images is assigned by radius it is fixed to range among 45-55 pixels [5].

The fundamental value of the X-ray remains the same today, as it was over 100 years ago – different structures (bone, cartilage, tissue, tumor, metal, etc.) can be identified based on their ability to block the X-ray/Röntgen beam. The initial uses of *in-vivo* imaging were to diagnose broken bones and locate foreign objects, such as, bullets, inside a patient‘s body. As imaging techniques and resolutions improved, physicians began to use these methods to locate medical abnormalities both for planning surgery and for diagnosing the disease. It is natural to ask whether the improved resolution of medical imaging has clinical value. Consider the use of CTs to diagnose lung cancer. The expected 5-year survival rate for all patients with a diagnosis of lung cancer is merely 15%. In the United States, lung cancer is the leading cause of cancer death for both men and women and costs almost $10 billion to treat annually. Genetic algorithms are modeled on biological processes in which parents pass character traits to their offspring. The biomedical images classify the digital mammograms in two ways such as abnormal and normal. The normal case mention that healthy patient and the abnormal state mention that benign cases. The biomedical images represents the mammograms images, cancerous tumors and digital mammograms.

This paper presents study of some predominant data mining algorithms such as Decision tree, genetic algorithm and k-means method to find biomedical image classification accuracy.

II. ALGORITHM USED

In section II, the various algorithms pertaining to image classification are discussed. In section III highlights the comparison these algorithms. In section IV research findings are presented. Section V covers the conclusion part of this research paper.

A. Decision Tree (DT)

Vasileios et al. [6] proposed a decision tree method for signifying and classify tree structures in biomedical images. The tree structure is regularly encountered in biomedical image contexts. The author have use tree encoding techniques to obtain a symbolic representation. The main problem of classifying biomedical images is reduced to string classification. In biomedical image analysis the feature image classification is associated with functions and it can be used to assist medical diagnosis problems. The concerned authors have proposed multistep method for classifying tree structure in biomedical images. This approach uses tree encoding schemes to obtain a representation of tree structures problems. The main goal of this method is to develop effective descriptors of tree structure that can be used for classification. To find classification accuracy this method is applied in clinical x-ray images. This technique is independent on image modality and the clinical relevance applications therefore this method applied in any type of biomedical images. This technique shows 86 % accuracy in biomedical image classification.

B. Genetic Algorithm (GA)

Bhuvaneswari et al. [7] cultivated a method genetic algorithm for automatic classification of lung diseases in CT lug images to find classification accuracy. This method defines new image feature extraction method for ordering the lung CT images. It contains three type of stages like pre-processing, feature extraction method and classification method. This method tested in more than 400 biomedical lung image datasets to find the performance of this method. The lung disease can CT images can be affected by impurity an exposure at the medications, workplace and various disorders. X-ray chest radiography and CT images are two common and main anatomic imaging that are routinely used in the diagnosis of various lung disease. The proposed method works effectively for detection of lung diseases images with high specificity, sensitivity and classification accuracy. This algorithm shows 91.53 % accuracy in image classification by using CT lung diseases images.

C. K-Means

Rajput et al. [8] implemented the k-means method to find classification accuracy in biomedical images. This proposed method the retinal color images processed to classify and identify the better results. The computational complexity of k-means method over the classification methods the authors decided to use k-means classification in the proposed work. This technique is a main classification method to generate hard classification accuracy. The biomedical images classified by namely hard exudates, hemorrhages and soft exudates. The exudates features is identified by using k-means classification method. This proposed method shows 97 % accuracy in biomedical image classification.

III. COMPARISON BETWEEN DIFFERENT CLASSIFIERS

Table I shows the comparison of genetic algorithms, decision tree and k-means algorithms based on different criteria as given below.

TABLE I. COMPARISON OF CLASSIFIERS

Table1	**Comparison of Methods**		
S. No	***Algorithm***	***Limitation***	***Usage***
1	DT	Sometime Accuracy will be Loss.	It is proper for real world problems
2	GA	Difficult to find Classification Accuracy	This Technique efficiently works well for lung diseases
3	K-Means	.Algorithms fails in Large datasets .This method fails in portion of optic disk images	Depends on Many Parameters

IV. RESEARCH FINDINGS

A. Biomedical Image Mining

Biomedical is one of the best method in image classification. In this study, we intend to use larger mammographic database and getting valuable information from the images. In separately, classification method based on data mining technique becomes more accurate in large datasets. In this paper, we studied three technique for biomedical image classification.

B. Data Sources

To find the biomedical image classification, the concerned authors have used in their methods from different biomedical image datasets such as retinal images, X-ray Galactograms and CT lung images. The table II describes the different datasets implemented in data mining algorithm.

TABLE II. DESCRIPTION DATASETS

Table2	**Dataset Description**	
S. No	***Algorithm***	***Biomedical Dataset***
1	Decision Tree	X-ray galactography
2	Genetic Algorithm	Computed Tomography lung images
3	K-Means	Color retinal images

C. Image classification

Image Classification is a process of finding document class in each type of data. In this study, the different classification techniques are used such as K-Means method, genetic algorithm and decision tree to find the biomedical image classification accuracy.

D. Process to classification accuracy

- ✓ **Image Acquisition**

Image Acquisition is a process of create digital images. It is also include compression, processing, display and printing in digital image process.

- ✓ **Image Enhancement**

Image enhancement is used to develop the eminence of images.

- ✓ **Preprocessing**

The preprocessing technique is used to filter unwanted noisy data from biomedical images. This method will remove the unwanted noises and produce the enhanced images.

- ✓ **Feature Extraction**

The feature extraction method is used to decrease the unique dataset by calculating some certain features. Each type of sub images and pixel coefficient values are removed. To analyze the classification of texture in the biomedical images can be completed by using numerical feature extraction method.

- ✓ **Feature Selection**

Feature selection technique is a process of selecting an a bridged relevant features that expands classification accuracy by finding for greatest features from subset images. This method also involves the selection of subset from to total dataset based on given criterion.

E. Performance of Techniques

Decision tree approach classifying and characterizing like tree structure in biomedical images. This method is used to tree encoding schemes to obtain string representation of tree structure. The goal of decision tree method is to develop effective classification accuracy that can be used for analyze the performance of this technique. The concerned author analysis the biomedical x-ray images in order to find the normal cases images. This method independent of the biomedical imaging modality and medical applications. So it can also applied to any other method like k-means and genetic algorithms. The proposed decision tree method demonstrate the effectiveness of classification and characterization in biomedical images. The concerned author have considered 54 x-ray images to find performance of decision tree method. Those images acquired from more than 30 women

with ages ranging from 25 to 70 years. In each type of test the author retrieve the nearby neighbor trees based on the similarity space. Table III illustrates the classification accuracy obtained by the decision tree, genetic algorithms and k-means method when using biomedical images.

TABLE III. PERFORMANCE STUDY OF ALGORITHMS

Table3	**Methods**		
S. No	***Methods***	***Correctly Classified Instances***	***Incorrectly Classified Instances***
1	Decision Tree	86%	14%
2	Genetic Algorithms	91.53%	8.47
3	K-Means Method	93%	7%

The table III contains classification accuracy of k-means method, decision tree method and genetic algorithms. The decision tree method gives 86 % accuracy in image classification by using biomedical x-ray images. Genetic algorithm shows 91.53 % accuracy in image classification by using biomedical computed tomography lung images. The k-means technique shoes 93 % accuracy in classification by using biomedical color retinal images. The highest classification accuracy is 93 % obtained by k-means method.

Genetic algorithm is used to increase capability through calculation. The solution of this problem is represented in the form of chromosome called genes that hold set of values for the optimization variables. Once the original population is generated, the fitness function which notify the goodness of chromosome in optimization task. Based on the fitness function, gene are nominated and some genetic operations such as crossover and mutation it is applied in the selected chromosome. These genes develop best individuals till it spreads the global optimum solution. This genetic algorithm is instigated on lung diseases CT images. The input dataset of proposed method consist of more than 400 images.

The k-means classification method is a superior circumstance of the classification methods. The proposed k-means technique experimented on more than 100 biomedical images. The morphological process is used to removing the blood vessel network to getting good results. The proposed technique has successfully classified the biomedical images especially in exudates for retinal images. This type of classification helps to physicians in diagnosing retinal diseases. This k-means technique is implemented by CIELAB to identify and classify the better classification accuracy. The figure 1 spectacles the classification performance of studied algorithms.

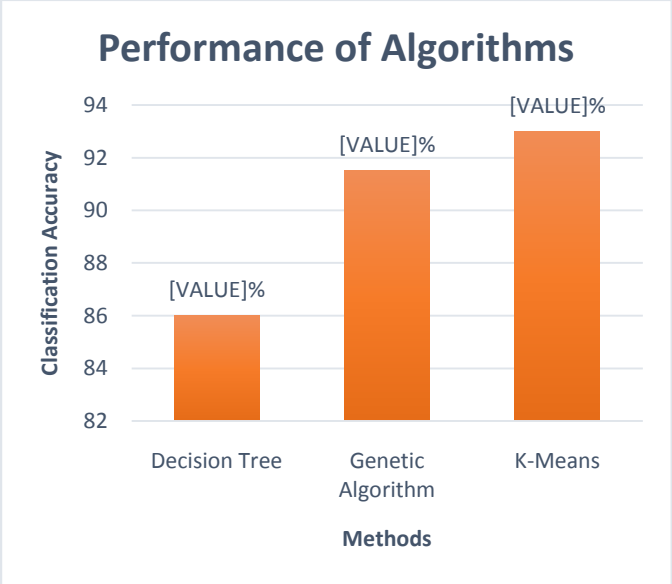

Fig. 1. Performance of methods in image classification

V. CONCLUSION

The improvements of biomedical image technology have prominently improved for the biomedical imaging study. Image classification is a process of finding expedient information from hefty amount of data.

Image classification is one the essential technique in image mining. In this paper, we studied three algorithms for image classification in biomedical images. The decision tree method shows 86 % accuracy, genetic algorithm shows 91.53 % accuracy and k-means method shows 93 % accuracy in different biomedical images such as retinal images, x-ray images and CT images. From this k-means algorithm shows highest accuracy when compare with other two methods. So we suggest that k-means method is one of the predominant method in biomedical image classification. In future, we planned to reduce processing time of k-means method because k-means method takes the long processing time to find the classification accuracy.

REFERENCES

[1] Graham R, Perriss R, Scarsbrook A. Dicom demystified: a review of digital file formats and their use in radiological practice. Clin Radiol 2005:1133–40.

[2] Kalender WA. Computed tomography: fundamentals, system technology, image quality, applications. 3rd ed. Erlangen: Publicis Publishing; 2011.

[3] M. Garcia, C. I. Sanchez, M. I. Lopez, D. Abasolo and R. Hornero, "Neural network based detection of hard exudates in retinal images," Computer Methods and Programs in Biomedicine, vol. 93, pp. 9-19, 2009.

[4] AshagowdaKaregowda, AsfiyaNasiha, M.A. Jayaram and A.S. Manjunath proposed an "Exudates detection in retinal image using back propagation neural network" International Journal of Computer Applications (0975 – 8887) Volume 25– No.3, July 2011.

[5] Jaspreet Kaur, Dr H. P. Sinha proposed "Automated Localization of Optic Disc and Macula from fundus Images" , International Journal of Advanced Research in Computer Science and Software Engineering, ISSN – 2277128X, Volume 2, issue 4, April 2012.

[6] Vasileios Megalooikonomou, Michael Barnathan, Despina Kontos, Predrag R. Bakic, Andrew D. A. Maidment "A Representation and Classification Scheme for Tree-Like Structures in Medical Images: Analyzing the Branching Pattern of Ductal Trees in X-ray Galactograms" IEEE Transactions On Medical Imaging, Vol. 28, No. 4, April 2009.

[7] C. Bhuvaneswari, P. Aruna, D. Loganathan "A new fusion model for classification of the lung diseases using genetic algorithm" Elsevier, Egyptian Informatics Journal, pp 1-9, 2014.

[8] Dr. G. G. Rajput, Preethi N. Patil, "Detection and classification of exudates using k-means clustering in color retinal images" IEEE Fifth International Conference on Signals and Image Processing, ISBN 978-0-7695-5100-5, 2014.