International Journal on Emerging Technologies **10**(2): 403-407(2019)

ISSN No. (Print) : 0975-8364
ISSN No. (Online) : 2249-3255

An Efficient Mammogrammic Image Retrieval using Ring-based Classification

Sonia Jenifer Rayen[1] and R. Subhashini[2]
[1]Research Scholar, School of Computing, Sathyabama Institute of Science and Technology, India.
[2]Professor, School of Computing, Sathyabama Institute of Science and Technology, India.

(Corresponding author: Sonia Jenifer Rayen)
(Received 09 May 2019, Revised 25 July 2019 Accepted 31 July 2019)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: Content based Image Retrieval (CBIR) systems are used to retrieve useful contents from a massive amount of medical images. CBIR system is used to detect breast cancer in women who have an indication of the disease using mammography. In the existing system, the images are first retrieved and then classified. In the proposed method, the image retrieval time is less since it has to search the whole database for performing the retrieval. Moreover the retrieval rates obtained by the existing techniques are not satisfactory. Hence in this paper a novel Supervised Euclidean Distance - Ordered Area under Curve (ED-OACC) ranking method has been proposed for image retrieval process. The query image is partitioned into multiple rings. Initially the inner most ring is selected and calculated the distance based on the Region of Interest (ROI) for extracting the image features. The extracted features are ordered to select the relevant features using threshold value. The relevant features of all rings are consolidated into most suitable images which are extracted from the database. Multilayered feature extraction and ranking method performs faster retrieval process with more accuracy. The proposed image retrieval is compared with various classification algorithms, ranking algorithms and CBIR algorithms. The analysis is done by using precision, recall and F-measure parameter in order to attain more accuracy. The objective of the proposed method is to retrieve the relevant images from the database images in an efficient and reliable manner.

Keywords: CBIR, Digital Mammography, Image Ranking, Classification, Multimedia description, Machine Learning.

Abbreviations: CBIR-Content Based Image Retrieval

I. INTRODUCTION

Content Based Image Retrieval (CBIR) is a technique which is used to retrieve an image from the database for user query. There are huge amount of images are produced by different sources such as internet, social networking, sensors etc. These images are stored in a large volume of image database with low level and high level features. Traditionally the images are stored with annotation which will not provide solution, so an automatic image annotation has been introduced. A semantic image retrieval technique is used to get an exact image according to the user request.

Mammogram analysis is generally centered on the radiologist subjective judgment; also this procedure is difficult and inaccurate sometimes, that leads to missing lesion recognition during the routine check. Therefore, it's a requirement for a system to perform effective prediction and also retrieval of mammogram images

II. RELATED WORKS

Multi-Directional Search technique splits the user's query into multiple sub quires for retrieving identical images as clusters but it is not suitable for integrating semantic text with visual image features [1]. A multimodal content retrieval framework retrieves the images with heterogeneous multimedia sources with weighting scheme for achieving high retrieval accuracy [2]. Re-ranking method prepares rank list of all images based on indexing structure but it should be strengthen by using CBIR descriptors such as local features and motion patterns [3]. Recognizing Activities of Daily Living (RADL) method tracks and monitors a smart home with two components namely smart home management and ADL pattern monitoring which will be extended to all daily activities in the home infrastructure [4]. Incremental Annotations based image search with clustering (IAISC) which adopts the textual features and visual features for retrieving an exact image from the database. This technique suffers in two issues such as insignificant annotation degree and non relevant annotation [5].

New content-based image retrieval technique uses features like colour, texture and shape of the image and it can be extended to real time data from various sources [6]. An image based clinical trials have been analyzed and distributed based on the features which are collected from the healthcare database in order to reduce the error with in a limited time [7]. Dual directional multi-motif XOR patterns (DDMMXoRP) are used to extract and index the images by using V colour space of HSV colour plane in dual direction with better performance when compared to existing techniques [8]. Structural Support Vector Machine (S-SVM) model uses a ranking based prediction for local and global cues which achieves better consistency in semantic image segmentation [9].

A multi-dimensional inverted index method has been introduced for an ordered quantization among features descriptor with accuracy in the content based image retrieval [10]. A wearable single and dual-polarization antenna have been implemented for detecting breast cancer which leads performance problem. This can be achieved by integrating wearable arrays with bra-like prototype to microwave breast cancer detection [11]. Radiomics-Driven Conditional Random Field (RD-CRF) framework is used to extract voxel-level quantitative radiomics features and ensure the interconnected tissue characteristics of tumors [12]. Hashing-based approximate nearest neighbour search uses kernel-based method for extracting images by considering unlabeled image with kernel space. This method can be extended by implementing hashing in active learning methods for solving the CBIR problems with reduced searching time [13]. Unsupervised graph theoretic approach retrieves the image by using graph region based technique with modelling and evaluating similarities in the query image. An automatic feature learning and weighting algorithm is introduced to characterize the image regions with swift parallel algorithm [14]. Composite Anchor Graph Hashing with Iterative Quantization (CAGH-ITQ) is an unsupervised hashing algorithm integrates multimodal features and distance metrics as composite Anchor Graph by improving the efficiency, and also it minimizes the quantization error. This will be extended to retrieve the mammogram with hash code based supervised learning technique with relevant features [15, 24].

Robust Domain Adaptation with Low-Rank Reconstruction (RDALR) is used to categorize and to classify the information as outliers and noise between the samples in source domain and target domain space. This suffers a limited number of labels in target space, so semi supervised method has been introduced [16]. Linear Neighbourhood Propagation (LNP) is a problem of fitting identified samples to new target domains which leads to performance problem during classification process [17, 25].

Semi-supervised RDALR (SRDALR) technique is used to find the samples from partial source domain sample set and generates the cumulative sample set for further processing. It is not suitable for large sample set as well as different target sample space [18]. Label Propagation with Instance Weighting (LPIW) has multiple terms for predicting the labels from the source sample space which cannot be fit to the target sample space [19]. SRDALR+LPIW are used to provide the positive effects and high confidence with less number of samples. It suffers accuracy problem while handling large sample space [20].

III. PROPOSED METHOD

The flow diagram of the proposed method is shown in Fig. 1. Initially the query image is collected from the sources and performs pre-processing operation.

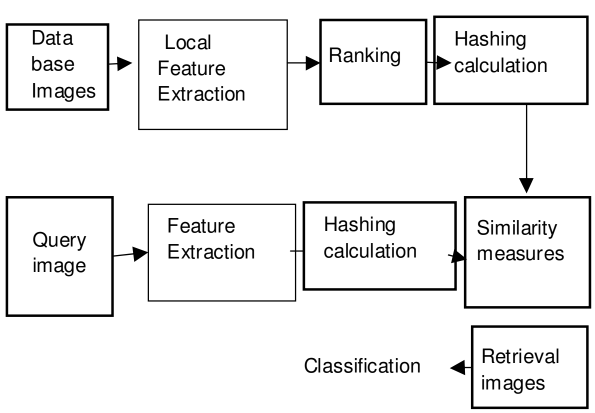

Fig. 1. Flow diagram.

The local visual features are extracted and classified in order to identify various layers of the image. The layer based supervised learning method is carried out in which the relevant features are used for subsequent process of the image retrieval. The classified features are ranked by using the ring structure. Multiple rings are identified and also extracted the global features from the ring. Every ring has features which are ordered and maintained in a separate rank list. There are N ring in the image with M rank list. The threshold has been fixed in global rank list for choosing best features. The hash operation is performed over ordered features for ensuring the changes in the image after performing the ranking process. The hash calculation of the query image and database images are compared. The relevant images are retrieved with high quality. The steps involved to retrieve the image using content based image retrieval Technique is described in the following pseudo code,

Step 1: Read the query image
Step 2: Perform Classification based on different visual features
Step 3: Perform the feature extraction for classified features
Step 4: Rank the classified features
Step 5: Perform Hashing for classified features
Step 6: Compare the Hash value of query image with database image
Step 7: Retrieve the relevant image

IV. RING BASED FEATURE EXTRACTION AND RANKING

The image is partitioned into multiple rings for identifying and extracting the reliable and relevant features for further retrieval process. Initially the innermost ring is selected and extracts the features as separate list. The ranking process has been carried out based on the extracted features and it selects the top most features based on the threshold. The ordered features are combined into a single list with suitable threshold value. The parameters used for proposed ranking process are Image Ring (IR) ={ $Ring_1$, $Ring_2$, ... $Ring_n$}, F = {f_1,f_2,...f_n}, R={R_1,R_2,...,R_n}, RN={RN_1,RN_2,...RN_n} and Cumulative Rank is CR. Fig. 2 shows that the process of ring based ranking method.

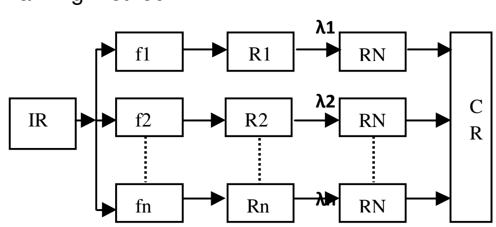

Fig. 2. Ring based ranking.

Where IR is an Image Ring, f1 is a image features, R1 is a ranking features, λ is a threshold value, RN is a New Ranking of Features, CR is a Cumulative Features. The Euclidean Distance (ED) of each ring is calculated by considering all samples with ranking process. This relevant feature in a particular ring is considered as Ring1 rank list. The second ring features are collected by performing the ED with already extracted relevant features. The second ring ranking process is carried out along with the Ring1 rank list. This ranking process gives a cumulative ranking of features. The same process repeats until the outer most ring is reached. The final rank list contains the cumulative rank of all relevant features in the image.

Ring1 ED calculation is represented as follows

$$L1 = |(x11, y11)| + |(x12, y12)|$$

$$L2 = \sqrt{|(x11, y11)|^2 + |(x21\ y21)|^2}$$

L3 = MAX (L1, L2)
Ring2 ED calculation is described as follows
L12 = |(x11,y11)+(x21,y21)| + |(x12,y12)+(x22,y22)|

$$L22 = \sqrt{|(x11, y11) + (x21\, y21)|^2 + |(x12\, y12) + (x22, y22)|^2}$$

L32 = MAX(L12, L22)
The same process repeats until the outer most ring is reached. The final cumulative ranking feature gives an efficient result during the retrieval process.

V. ANALYSIS OF RING BASED RANKING ALGORITHM

The rank list is maintained based on the ring in the query image is described in equation (1).

$$Ring\ List = \bigcup_{i=1}^{n} Rin > 0 \quad (1)$$

The local features are extracted as Feature List from the respective ring is expressed with suitable parameter in equation (2).

$$Feature\ List = \bigcup_{i=1}^{n} Ri - \bigcup_{i-1}^{m} fi\!:.nm > 0 \quad (2)$$

The feature ordering is performed based on the threshold applied over ring in order to extract relevant features is specified in equation (3).

$$Rank(R) = sort\ (\bigcap_{i=1}^{n} fin > 0) \quad (3)$$

The new feature list has been created with the suitable threshold λ and local rank list is described in the equation (4)

$$New\ Rank(RN) = s\lambda iXri\ \varepsilon\ R \quad (4)$$

The cumulative rank list is created based on ordered local features and global features with threshold is shown in equation (5)

$$Cumulative\ Ranking = \sum_{i-1}^{n} R\ Ni\, n > 0 \quad (5)$$

VI. SIGNIFICATION OF ASSESSMENTSCORE

The assessment code provides the report for the breast cancer in various perspectives. These codes indicate the normal and abnormal condition in the mammogram with negative and positive respectively. The assessment code is described in Table 1. The importance of assessment score is to find out breast cancer in all levels.

Table 1: Assessment score.

Code	Report
1	Negative
2	No Cancer (Benign)
3	Probable Benign
0	Need Additional Evaluation
4	Suspicious Cancer
5	Cancer (malignancy)

A. Attribute and its category
The attributes related to the breast cancer identification is based on the variables such as age, Assessment, cancer, density, mammogram comparison, density, family history, status of hormone therapy, prior mammogram, mammogram type and cancer type. Fig. 7 shows that the comparison of the variables related to the breast cancer detection.

VII. EXPERIMENTAL EVALUATION

A. Data Set
The database images are taken from Digital Database for Screening Mammography (DDSM) with various categories. Dataset contains 40000 records of different age group with various cancer diseases

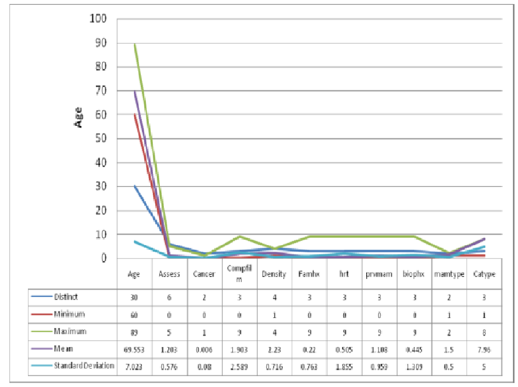

Fig. 3. Comparison of the variables.

VII. EXPERIMENTAL EVALUATION

A. Data Set

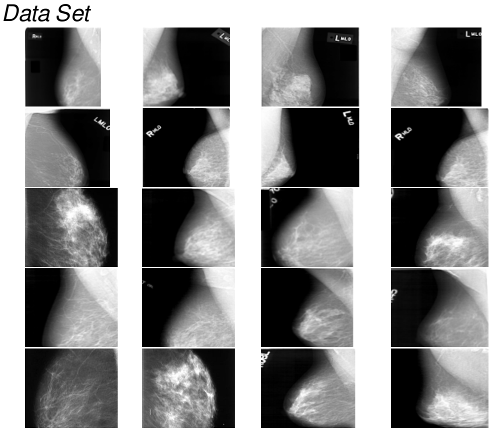

Fig. 4. Sample database mammogram images.

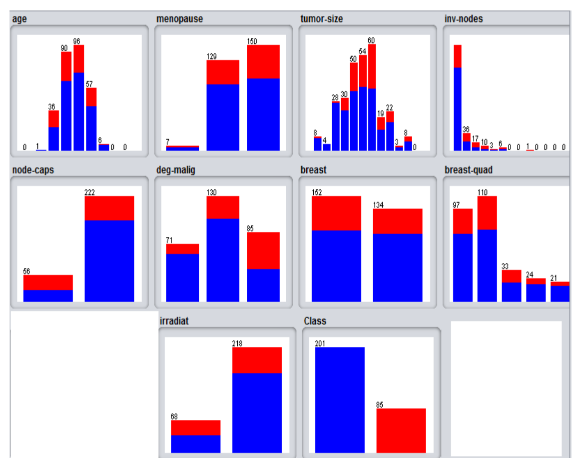

Fig. 5. Variable comparison of Dataset1.

In this dataset compare various features such as person age, menopause, tumor size, tumor nodes, node caps tumor catagery and class. It is shows in Fig. 5.

VIII. PERFORMANCE EVALUATION

The comparison of the precision and recall value is shown in Fig. 6. After the ranking process has been completed the most fitted features are collected and form a new feature list. The performance has measured by using Precision, Recall and F-measure values with proposed algorithm. The precision, recall and F-

Measure are specified in equation 4, equation 5 and equation 6 respectively.

$$recision = \frac{Number\ of\ Relavant\ Iages}{Total\ Number\ of\ Retrived\ Images} \quad (6)$$

$$Recall = \frac{Number\ of\ Relavant\ Iages}{Total\ Number\ of\ Retrived\ Images\ in\ the\ database} \quad (7)$$

$$F - Measure = \frac{2*Recall*Precision}{Precision*Recall} \quad (8)$$

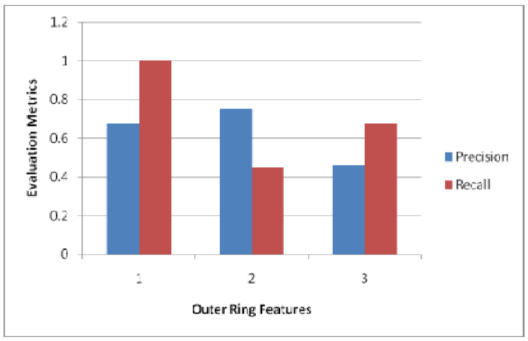

Fig. 6. Ring comparison of the precision and recall.

A. Comparison of Classification Algorithm

The proposed method has compared with various classifiers based on the precision and recall value. The proposed model has compared with SMO SDG, Navie Bayers and Bayer Net classification models. Precision and recall value is carried in whole classification process. The proposed method performs classification with less iteration because of maintaining ring based features which is shown in Fig. 7.

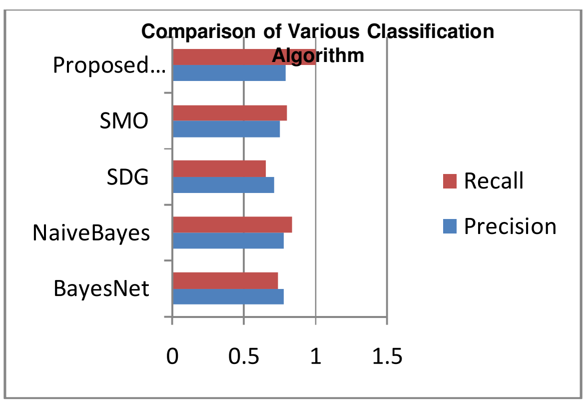

Fig. 7. Comparison of various Classification Algorithms.

B. Comparison of Ranking Algorithms

Euclidean Distance (ED) and Manhattan Distance (MD) are true order for rank measurement but it leads performance problem in multi layered samples. The proposed method integrates the ED and OACC in order to perform the efficient ranking process. The comparison of various ranking algorithm with proposed ED-OACC is shown in Fig. 8. This graph shows the precision and recall analysis. Various ranking algorithm has been compared with the proposed techniques. But the proposed precision and recall rate is high while comparing the existing methods.

C. Comparison of CBIR Algorithms

The proposed ring based classification of supervised technique provides a better performance in the retrieval process. The comparison of different Content based Image Retrieval an algorithm with proposed ED-OACC is shown in Fig. 9. An efficiency of the CBIR algorithm is calculated based on the parameters like precision, recall and accuracy rate.

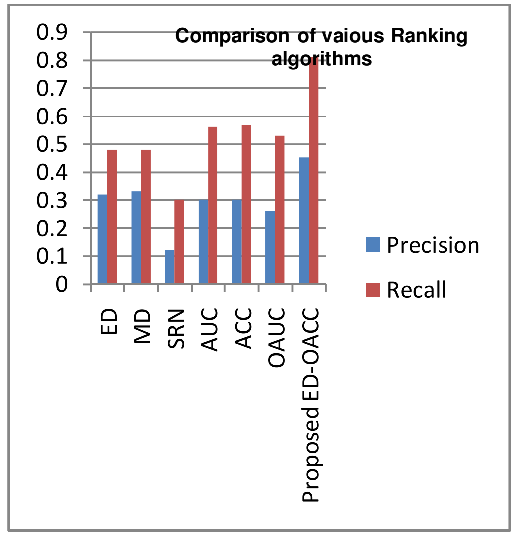

Fig. 8. Comparison of Ranking Algorithms.

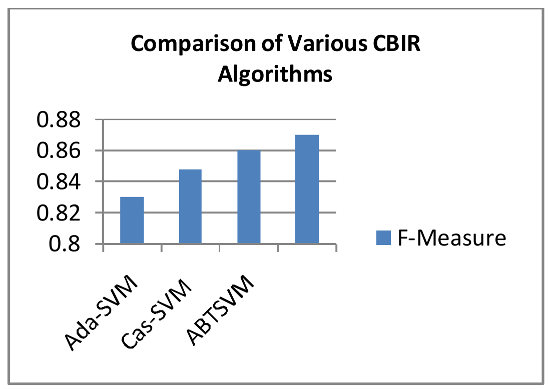

Fig. 9. Comparison of CBIR Algorithms.

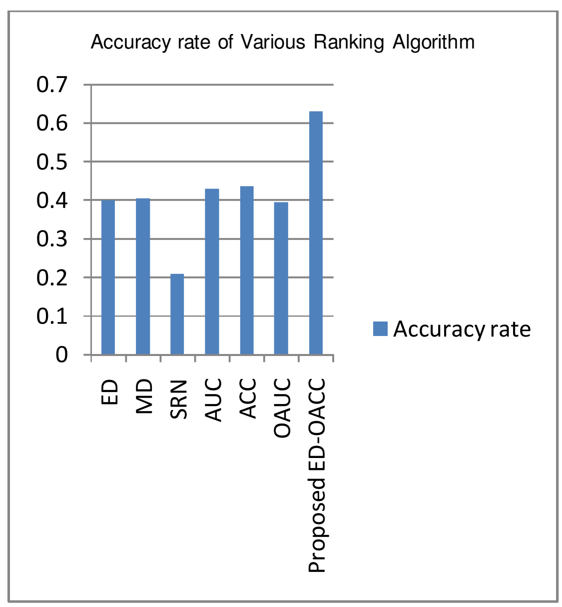

Fig. 10. Accuracy rate comparison of CBIR Algorithms.

XI. CONCLUSION

This paper proposed a new Content Based Image Retrieval (CBIR) method named Multi-Feature combine with SVM in applications. This type of image retrieval systems is used to retrieve useful information of medical images. In the existing methods, the image retrieval time is more. The proposed Multi-features based SVM technique provides a clear value. Also here the output retrieval rate is high compared to existing models. The proposed method uses a supervised classification method for extracting multiple levels of image features. The features are selected based on the rings available in the image. The ranking operation has performed in two levels namely local ring level features and global ring level features. Ada-SVM, Cas-SVM and ABTSVM classification methods are also compared with the proposed method. From the above analysis the proposed supervised rank based classification process achieves maximum efficiency and accuracy. In future the proposed method can be extended to data sciences for handling large amount of real time samples.

REFERENCE

[1]. Ning Yu, Kien A. Hua, Hao Cheng, (2012). A Multi Directional Search technique for image annotation propagation. *Journal of Visual Communication Image Retrival,* Vol. **23**, pp. 237–244.

[2]. D. Rafailidis, S. Manolopoulou, P. Daras, (2013). A unified framework for multimodal retrieval. *Pattern Recognition,* **46**, 3358–3370.

[3]. Daniel Carlos Guimarães Pedronette, Jurandy Almeida, (2014). Ricardo da S. Torres. A scalable re-ranking method for content-based image retrieval. *Information Sciences,* **265**, pp. 91–104.

[4]. Ihn-Han Bae, (2014). An ontology-based approach to ADL recognition in smart homes. *Future Generation Computer Systems,* **33**, pp. 32–41.

[5]. Fan Wu, Hao-Ting Pai, Yi-Feng Yan, Jeff Chuang, (2014). Clustering results of image searches by annotations and visual features. *Telematics and Informatics,* **31**, pp.477–491.

[6]. Kashif Iqbal, Michael O. Odetayo, Anne James, (2012). Content-based image retrieval approach for biometric security using colour, texture and shape features controlled by fuzzy heuristics. *Journal of Computer and System Sciences,* **78**, Pp.1258–1277.

[7]. Ximing Wang, Brent J Liu, Clarisa Martinez, Xuejun Zhang, Carolee J. Winstein, (2015). Development of a novel imaging informatics-based system with an intelligent workflow engine (IWEIS) to support imaging based clinical trials. *Computers in Biology and Medicine*, Pp. 1-9.

[8]. Santosh Kumar Vipparthi, Subrahmanyam Murala, Shyam Krishna Nagar, (2015). Dual directional multi-motif XOR patterns: A new feature descriptor for image indexing and retrieval. Optik, **126**, Pp.1467–1473.

[9]. Deli Pei, Zhenguo Li, Rongrong Ji, Fuchun Sun, (2014). Efficient semantic image segmentation with multi-class ranking prior. *Computer Vision and Image Understanding,* **120**, Pp.81–90.

[10]. Feng, D., Yang, J., & Liu, C. (2013). An efficient indexing method for content-based image retrieval. *Neurocomputing*, **106**, 103-114.

[11]. Hadi Bahramiabarghouei, Adam Santorelli and Leslie A. Rusch, (2015). Flexible 16 Antenna Array for Microwave Breast Cancer Detection. *IEEE Transactions On Biomedical Engineering*, Vol. **62**, NO. 10.

[12]. Chung, A.G., Khalvati, F., Shafiee, M.J., Haider, M.A., & Wong, A. (2015). Prostate cancer detection via a quantitative radiomics-driven conditional random field framework. *IEEE Access*, **3**, 2531-2541.

[13]. Begum Demir and Lorenzo Bruzzone, (2016). Hashing-Based Scalable Remote Sensing Image Search and Retrieval in Large Archives. *IEEE Transactions on Geoscience and Remote Sensing,* Vol. **54**, No. 2, Pp.892-904.

[14]. Bindita Chaudhuri, Begüm Demir, Lorenzo Bruzzone and Subhasis Chaudhuri, (2016). Region-Based Retrieval of Remote Sensing Images Using an Unsupervised Graph-Theoretic Approach. *IEEE Geoscience And Remote Sensing Letters*, Vol. **13**(7), Pp. 987-991.

[15]. Jingjing Liu, Shaoting Zhang,Wei Liu, Cheng Deng, Yuanjie Zheng and Dimitris N. Metaxas, Scalable (2016). Mammogram Retrieval Using Composite Anchor Graph Hashing with Iterative Quantization, DOI 10.1109/TCSVT.2016.2592329, Pp.1-11.

[16]. Jhuo, I.H., Liu, D., Lee, D.T., & Chang, S.F. (2012). Robust visual domain adaptation with low-rank reconstruction. In *2012 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2168-2175). IEEE.

[17]. Olga Zoidi; Nikos Nikolaidis; Ioannis Pitas, (2013). Exploiting clustering and stereo information in label propagation on facial images. *IEEE Symposium on Computational Intelligence in Biometrics and Identity Management (CIBIM), IEEE*, pp. 31–35.

[18]. Qian Shi; Bo Du; Liangpei Zhang (2015). Domain Adaptation for Remote Sensing Image Classification: A Low-Rank Reconstruction and Instance Weighting Label Propagation Inspired Algorithm. *IEEE Transactions on Geoscience and Remote,* Volume: **53**, Issue: 10, Pp.5677 – 5689.

[19]. Yuzhou Wu, Yu Sun, Xinle Liang, Ke Tang, Zixing Cai, (2015). Evolutionary semi-supervised ordinal regression using weighted kernel Fisher discriminant analysis. *IEEE Congress on Evolutionary Computation* (CEC), DOI: 10.1109/CEC.2015.7257300, IEEE, 2015.

[20]. Lilian Berton, Alneu de Andrade Lopes, (2014) "Graph Construction Based on Labeled Instances for Semi-supervised Learning. *2014 22nd International Conference on Pattern Recognition*, Pp. 2477– 2482.

[21]. Wei, L., Yang, Y., & Nishikawa, R.M. (2009). Microcalcification classification assisted by content-based image retrieval for breast cancer diagnosis. *Pattern recognition*, **42**(6), 1126-1132.

[22]. Seetharaman, K., & Sathiamoorthy, S. (2016). A unified learning framework for content based medical image retrieval using a statistical model. *Journal of King Saud University-Computer and Information Sciences*, **28**(1), 110-124.

[23]. Basha, S.M., Bagyalakshmi, K., Ramesh, C., Rahim, R., Manikandan, R. and Kumar, A. (2019). Comparative Study on Performance of Document Classification Using Supervised Machine Learning Algorithms: KNIME. *International Journal on Emerging Technologies*, **10**(1): 148-153

[24]. Balasaravanan, K., & Prakash, M. (2018). Detection of dengue disease using artificial neural network based classification technique. *Int. J. Eng. Technol.*, **7**(13), 13-15.

[25]. Annamalai, R., Srikanth, J., & Prakash, M. (2015). Accessing the data efficiently using prediction of dynamic data algorithm. *International Journal of Computer Applications*, **116**(22).

How to cite this article: Rayen, Sonia Jenifer and Subhashini, R. (2019). An Efficient Mammogrammic Image Retrieval using Ring-based Classification. *International Journal of Emerging Technologies*, **10**(2): 403–407.