**CLASSIFICATION OF MAMMOGRAPHIC LESION FOR BREAST CANCER**

***Mrs. K. Kiruthika M.Tech***

***kkiruthika.cs@gmail.com***

***Assistant Professor***

***Department of Biomedical Engineering***

***Hindusthan College of Engineering and Technology, Coimbatore***

 ***Abstract- Breast Cancer is one of the most common cancers in women worldwide. Mammogram is the most widely accepted modality for screening breast cancer. Mammography is reliable method for detection of abnormality in the breast. However, the mammographic abnormalities are subtle in dense breast, making the detection a challenging task. Computer aided Diagnosis (CAD) system access breast images objectively as appeared to subjective analysis by radiologists. The performance of the classifier gets affected by the obscure boundaries of mass in dense breast. Global feature that capture the overall characteristic of images, will be erroneous due to incorrect segmentation in obscure mass cases. In our proposed system we utilized the local descriptors- Scale invariant feature transform (SIFT), Local Binary Pattern (LBP). Using these descriptors Classifier strategy model was performed for classifying normal, abnormal mass classification. In model based classifier we use MLP classifier along with Principal Component Analysis (PCA). These methods are tested on the benchmark dataset and acquired promising results***.

 *Keywords: Breast Cancer, Scale invariant feature transform (SIFT), Local Binary Pattern (LBP), and Principal Component Analysis.*

# INTRODUCTION

Worldwide, cancer remains a leading cause of mortality. The increasing global burden due to cancer can be attributed to the adoption of cancer-inducing lifestyle by the growing population of the world [1]. According to the World Health Organization (WHO) in 2012 has projected an increase in cancer incidence from 14.1 million to 21.6 million in 2030. The cell body maintains a cycle of regeneration processes. The balanced growth and death rate of the cells normally maintain the natural working mechanism of the body, Sometimes an abnormal situation occurs, where a few cells may start growing aberrantly which creates cancer [7][4]. In human body different types of cancer can be developed, among them breast cancer creates serious health concern. Women are more endangered to breast cancer than men because of anatomical changes in human body. There are different reason for breast cancer which includes breast density, age, family history, and obesity.

Breast cancer is the frequently diagnosed female cancer nearly affecting with age highest rate of 25.8 per one lakhs women and accounting 12.7 percent in mortality rate [1]. In 2020 it was suggested that the number would reach 1797900 [2]. Early stage of detection increases the recovery possibilities in breast cancer [3][6].

Female breast consists of ducts, lobules, fatty tissues and ducts. Milk is generated in lobules and carried towards nipple by ducts. Most of the BC will begin either in the breast tissue made up of glands or milk ducts. The remaining of the breast is made up of fatty, lymphatic, and lymphatic connective tissues [5]. When the cancer has begun, it also spread other parts of the body. Figure 1 shows the anatomy of female breast image.

BC tumor can be categorized into two broad scenarios.

* Benign (Noncancerous). Benign cases are treated as non cancerous, that is, non-life- threatening.
* Malignant (Cancerous). Malignant cancer begins from an abnormal cell growth and might rapidly spread or invade nearby tissue [7].

 Mammography is the screening and diagnostic technique, uses an X-ray image of the breast to examine and analyze the abnormal tissues [8]. Important indicators of breast cancer are masses



***Figure 1.1: Anatomy of female breast***

And micro calcification [9]. Usually, masses are identical from the surrounding tissue because it contains more variation in margin, size and shape [10]. So it is more challenging when detection of mammogram masses than the micro calcifications [11]. In dense breast, sensitivity of the mammogram is reduced by subtle nature of abnormalities and low contrasts [12] [13]. Breast Imaging Reporting and Data system (BIRADS) is a scheme that helps the patients to identify the class by following criteria, (i) BI-RADS I (0 to 25 % dense tissues, predominantly fat); (ii) BI-RADS II (26 to 50% dense tissues, fat with some fibro-glandular tissue); (iii) BI-RADS III (51 to 75% dense tissues, heterogeneously dense); and (iv) BI-RADS IV (above 75% dense tissues, extremely dense).



***Figure 1.2 shows the examples of breast in different BI- RADS classification.***

Despite its popularity, mammogram has a difficulty to making detection of abnormalities in dense breast. In mammography estimated that radiologist miss 10-30 % of cancer and also misinterpretation of benign as malignant may leads to biopsies and it make discomfort to the patients and It is proven that double reading can reduces the error, but in case of rural area its infeasible because of the shortage in radiologist.

Due to challenges and clinical significance of mammographic mass detection, [14] CAD sys- tem are invented to assist the radiologist and allows the doctors to analyze the medical images and finally detecting the abnormality in mammogram [15][16]. In classification based CAD system, region of interests (ROIs) has to be segmented, followed by global feature extraction and classification of normal and abnormal tissues [17].

Content Based Image Retrieval (CBIR) provide more convenient access to the images; while using conventional system for image retrieval, it provides clinically significance similar images by text based search. The existing system makes textual annotation problem when using attributes as a keyword [22]. In mammography cases, CBIR could provide visually similar mammograms based on the content based manner, query images are compared with the medical database images, based on the visual similarity and most relevant images are retrieved [23] [24]. Such approach have more advantages over the classifier based systems, they detect the unusual masses with more relevant database images, [47] and also it’s eliminates the segmentation and provides clinical evidence to assist the radiologist [25].

The widely used method that employs local descriptors for CBIR is the bag-of-features (BoFs) model/bag-of-words (BoWs) [28][29][30]. In First step multiple feature vectors are extracted from the image, for obtain a BoW. The feature space is defined by these vectors, where each image is represented as a set of feature vectors, after that codebooks are learned based on the set of representative feature [26]. This learning process is known as codebook/dictionary learning. Then for the each image constant length histogram has been extracted and encode the each image based on the learned dictionary quantize the feature. And finally histogram space is defined by using histogram and it can be used for retrieval tasks [27] [46].

# LITERATURE SURVEY

Menglin Jiang et al (2015), [31] had proposed an approach for diagnosis and retrieval of mammographic masses. In this method, the vocabulary tree stores all the quantized features of preceding diagnosed mammograms and the nodes of the vocabulary tree are refined. The existence of the mass in the query is determined by retrieved images. Low spatial-temporal cost of vocabulary tree improves the excellent scalability. Experiments are performed on digital database for screening mammography contains 11533 images, and as a result, vocabulary tree with adaptive Weight gives the classification Accuracy of 90.8 %.

Jingjing Liu et al (2017), [32] had proposed a extension of original Anchor Graph Hashing (AGH) and suggested a new unsupervised hashing algorithm, known as composite AGH with iterative quantization (C-AGH-ITQ). C-AGH-ITQ enables real-time searching by compressing mammograms into compact binary codes and in Hamming space. This algorithm, integrates the multiple features and different distance metric that are performed on the AGH. Introduction of orthogonal rotation matrix reduces the quantization error which helps in improving the effectiveness of the hash code. Experiments are conducted on the Digital Database for Screening Mammography images that contains 11533 ROIs. C-AGH-ITQ suggests 84 percentage of retrieval precision and 93 percentage of classification accuracy has been achieved by KNN prediction.

Devang Kulshreshtha etal (2017), had proposed a computer-based mammogram retrieval sys- tem [33]. Local binary pattern and k-mean clustering are used for feature extraction. In this approach based on the visual similarity of mammograms, it generates the clusters for query image features find the closest cluster that are matched with cluster representatives. Then mammographic ROIs are retrieved from closest cluster by computing Euclidean distance. During matching, the query images are searched depending on cluster size and they are not compared with database images. Finally, the approach provides a good response time and retrieval performance.

Vibhav Prakash Singh et al(2016), [34] proposed a CAD system , which automatically detects mammographic abnormalities. The method includes pre-processing, cropping of mammograms and suppression of Gaussian noise. GLCM based texture features from different distances of neighboring and angles are extracted. Furthermore, most relevant features are also examined using AdaBoost feature selection method. Finally, normal and abnormal mammograms are classified using Random forest (RF) classifier with accuracy of 89.02%.

Haixia Li et al (2017), [35] proposed a method in mammography images, translate 2D contour of breast mass into 1D signature, then obtain the new contour descriptor from the subsection by performing segmentation of 1D signature and features are extracted locally. Root mean square slope is a new contour descriptor, which ensures the roughness of the contour. Finally, classification is performed by using different classifiers such as SVM, KNN and ANN to detect benign and malignant mass and get the classification accuracy of 99.66% by using root measure square feature with SVM classifier.

Caio Eduardo et al (2017), [49] proposed the methods analyze the local feature of the Scale- Invariant Feature Transform (SIFT), Speed Up Robust Feature (SURF), Oriented Fast and Rotated BRIEF (ORB) and Local Binary Pattern (LBP) to discriminating the patterns of benign and malignant of masses in mammography image. These features are represented and dimensionality is reduced by using Bag of feature (BoF) model. Finally these features are classified using SVM, Adaboost and Random forest classifier. The system obtains 99.65% of accuracy for benign and malignant mass classification.

Taye et al (2018), [50] proposed a method to classify the mammogram as normal and abnormal, and from mammography image extract CNN feature and dimensionality of the feature is reduced by applying PCA. The experiments are done on Digital Database for Screening Mammography and Mammography Image Analysis Society (MIAS) datasets with the accuracy of 98.80% and 98.75% respectively.

1. **METHODOLOGY**
2. ***Dataset***

Digital Database for Screening Mammography is the database used for mammographic image analysis. The database consists of collection of patients images, which includes both normal and benign cases [45].

1. ***Feature Extraction***

The aim of this process is to extract and identify local descriptors from the mammographic ROI images. These descriptors contribute to the discrimination of lesions in breast masses. There are three set of local feature which includes SIFT, SURF and LBP are extracted separately to the mammographic ROI.

1. ***Scale In-variant feature transform (SIFT)***

 SIFT is a feature detection algorithm to detect local features in images [36] [37] SIFT descrip- tors are extracted through following steps. The first step involves finding local extrema by detection of scale invariant key points [38]. The second step is determining of scale and exact location in each key point with help of model fitting. The key points that are poorly localized with reduced contrast present on an edge are eliminated.

The next step focus on calculation of a gradient orientation histogram for histogram peak and key points of entire region at the selected scale and are taken account of the dominant orientation of key points. The last step is dividing the surrounding region into 4 x4 sub regions, following for each sub region which computes gradient orientations with 8 bin histogram of relative dominant orientation histogram and finally 128-D feature vector is formed by concatenating all the 16 histograms. The abovementioned method is designed such that the extracted SIFT descriptors are invariant to scale, rotation, translation. SIFT is discriminative in nature. i.e., a single feature vector can be accurately matched from a large database of features [39] [40].

1. ***Codebook learning and BoW Representation***

Initially for obtaining Bag of words, multiple feature vectors are extracted from the ROIs. By using set of representative feature codebook are learned. Afterwards for the each image constant length histogram are extracted based on the learned dictionary quantize the feature. Then histogram space is defined by using the histograms and this can be used for the retrieval task [15]. The following are the mathematical procedure to obtain BoW representation.

Each image consists of multiple feature vector (local descriptors) denoted as a. To develop the codebook, collection of feature vectors are clustered into number of clusters denoted by, $N\_{k}$ and use the corresponding centroids known as codeword denoted as $w\_{i}^{cc}$ and then distance has been computed between the each feature vector and the corresponding codeword.

$b\_{i}$= ⃦ a-$w\_{i }^{cc}$⃦…………………………………….. (1)

By using classical softmin function the distance are approximated,

$c\_{i}$ = exp ($\frac{-b\_{i}}{T})/\sum\_{j}^{}exp⁡(\frac{b\_{j}}{T})$………………..…… (2)

 Stabilization process is controlled by parameter T. Then literately calculates, for multiple feature vectors of the same images and calculate the different values for the different feature vector p,which is denoted by, $c\_{i}^{P}$. Finally integrates the response of single image over the all p points and obtain the final representation.

Feature

Extraction

Feature

Extraction

Database

Images

Testing

Database

Images

$d\_{i}$=$\sum\_{p=1}^{p}c\_{i}^{p}$ ……………………… (3)

1. ***Local Binary Pattern (LBP)***

SIFT

SIFT

LBP is applied to the region of mammographic lesions to analyze the local texture. LBP is a powerful method and has ability to describe the local spatial structure of an image, which includes non parameterized operator. Differentiation of texture feature is defined by,

LBP

LBP

LBP ($x\_{c},y\_{c}$) = $\sum\_{n=0}^{n-1}s(i\_{n}-i\_{c})2^{n}$………… (4)

Where, n denotes number of pixels related to the center one, $i\_{c}$ corresponds to gray level of center pixel, $i\_{n}$ n corresponds to gray level value of neighboring pixel, and S(x) is a function denotes results 1 if x ≥ 0 and 0, otherwise.

Normal/

Abnormal

Classifier

Model

Further, based on the results binary matrix is formed, and the neighborhood pixel values of the binary matrix are multiply with weights of the weight matrix. And LBP texture values are calculated by summing the pixel values.

$d\_{i}$=$\sum\_{p=1}^{p}c\_{i}^{p}$…………….. (5)

1. ***Model Based Classification***

The proposed system is explain in figure 3.1 explores classification of mammographic masses using classifier model. In our system feature level fusion can been employed followed by the classifier model

****

***Figure 3.1: Block Diagram of Model Based Classifier***

1. ***Feature Fusion***

 Feature fusion is a process of combining the specific extracted features which are stored in a dictionary to obtain a single feature file. Feature of SIFT and LBP are extracted from the mammographic ROI, then these two feature are concatenated serially and makes the feature representation, further it is given to the feature reduction techniques.

1. ***Principal component analysis (PCA)***

Feature reduction plays an important role in process of classification. Generally feature vector are high dimension. The main objective of the feature reduction is to reduce dimensionality of the feature; we use principal component analysis for feature reduction. Principal component analysis is a method to reduce the dimensionality of the features. To develop principal component PCA convert number of correlated variable into uncorrelated variables then by using feature vector first covariance matrix is calculated, then compute the eigenvector and eigen values and sorted accordingly to decreasing eigenvalue. For effective discriminate class bigger variance of data distribution are used than smaller variance.

1. ***Classifier model***

In the step of classification we adopt, multilayer perceptron neural networks (MLP) for mammographic mass classification. MLP consists of an input layer, one or more hidden layer followed by the output layer. Weights are assigned to the different layers and calculate the output by compares with it target output. Finally the error signal are back propagates to adjust the connection weight correspondingly. Multi layer neural network architecture is shown in Figure 3.2

****

***Figure 3.2: Multi layer architecture [47]***

**IV RESULTS AND DISCUSSION**

 The proposed approach has been evaluated on the benchmark DDSM database considering 2600 images are used in the training phase and 1700 images are used in the testing phase which includes both normal and abnormal images. In feature extraction step, local features are extracted from Mammographic ROIs, Figure 4.1 shows the original mammographic ROI. In SIFT feature extraction we localize the key points along with the descriptors, the dimension of the each feature vector is 128 and we extract the LBP features for original mammographic images, with the radius of 1 and 8 points are used to calculate the LBP mask and dimension of the feature vectors is 59, then we normalize the LBP histogram.





1. ***(b)***

***Figure 4.1: (a) Original mammographic ROI***

1. ***Mammographic ROI with SIFT key points***

 For learning the codebook, cluster the set of extracted feature vectors based on the optimal k value and use corresponding centroids to form a codebook. These centriods are used to quantize the feature vector, and to obtain the similarity between each feature vector and each codeword. The parameter T controls the quantization process.

 SIFT and LBP features are fused serially, and for the different PCA component the validation and testing accuracy are calculated. Table 4.1 shows the comparative analysis of different component with respect to different hidden size. Finally for .95 PCA components, achieves the better accuracy.



***Table 4.1: Accuracy of PCA with different components***

In Figure 4.2 shows the validation accuracy of PCA with different component, and based on the validation report we set the parameter accordingly Table 4.1 shows the component of .95 with dimension of 276 gives the better accuracy.



***Figure 4.2 Validation accuracy of (a) 40 component (b) 60 component (c) 80 component (d) 90 component (e) 95 component (f) comparison of different component***

Figure 4.3 shows the ROC curve for different PCA component, result shows .95 components provide good performance that other component.

***Table 4.2: Comparative analysis of different features***

******

***Figure 4.3: ROC curve for different PCA component***

 Table 4.2 shows the comparative analysis of different feature, from the results of performance metric we notice that SIFT LBP provide better classification accuracy than other feature. And SIFT LBP obtains 0.91 precision values which is higher than other feature precision values. Among the three feature, fusion of SIFT and LBP followed by applying SIFT LBP outperformed and provide effective mass characterization.



***Table 4.3: Density wise classification report***

 Table 4.6 shows the density wise classification the result shows the proposed fused features yielded reasonably good results for heterogeneously dense cases with an overall accuracy of 92.8%.

**V CONCLUSION**

 Mammographic mass classification is challenging task, due to the varying size of masses, shape and margin and difficulty to distinguish between normal and abnormal cases. Classifier performance is affected by the obscure boundaries in dense breast and incorrect segmentation by extracting global feature. In our proposed method SIFT, SURF and LBP based local descriptors were extracted followed by retrieval based classifier and model based classifier. Model based classifier utilizing fused SIFT and LBP evaluated on benchmark DDSM database with overall accuracy of 92.8%.

**REFERENCES**

[1] Datta, K., Choudhuri, M., Guha, S and Biswas, J.” Breast cancer scenario in a regional cancer entre in Eastern India over eight years-still a major public health problem.” Asian Pacific Journal of Cancer Prevention 13(3) (2012): 809-813.

[2] Malvia, S., Bagadi, S. A., Dubey, U. S., and Saxena, S. ”Epidemiology of breast cancer in Indian women.” AsiaPacific Journal of Clinical Oncology, 13(4) (2017): 289-295.

[3] Phadke, A. C., and Rege, P. P. ”’Fusion of local and global features for classification of abnormality in mammograms.” Sadhana 41(4)(2016): 385-395.

[4] American Cancer Society. Breast Cancer Facts and Figures 2017-2018. Atlanta: American Cancer Society Inc. 2017.

[5] Eberl Margaret M, Fox Chester H, Edge Stephen B, Carter Cathleen A and Mahoney Martin C.” BI-RADS classification for management of abnormal mammograms.” J. Am. Board Fam. Med. 19(219) 2006: 161164

[6] Tsochatzidis, L., Zagoris, K., Arikidis, N., Karahaliou, A., Costaridou, L., andPratikakis, I. ”Computer-aided diagnosis of mammographic masses based on a supervised content-based image retrieval approach.” Pattern Recognition 71 2017: 106-117.

[7] Nahid, A. A., and Kong, Y. ” Involvement of Machine Learning for Breast Cancer Image Classification: A Survey.” Computational and mathematical methods in medicine 2017.

[8] Li, Y., Chen, H., Cao, L., and Ma, J. ”A survey of computer-aided detection of breast cancer with mammography.” J.” Health Med. Inform 7(4) 2016.

[9] H.-D. Cheng, X.-J. Shi, R. Min, L.-M. Hu, X.-P. Cai, and H.-N. Du, ”Approaches for automated detection and classification of masses in mammograms, ” Pattern Recog vol. 39, no. 4, pp. 646668, 2006

[10] A. Oliver, J. Freixenet, J. Mart, E. Perez, J. Pont, E. R. E. Denton, and R. Zwiggelaar,”A review of automatic mass detection and segmentation in mammographic images,” Med.Image Anal vol. 14, no. 2, pp. 87110, 2010

[11] R. M. Rangayyan, F. J. Ayres, and J. E. Leo Desautels, A review of computer aided diagnosis of breast cancer: Toward the detection of subtle signs, J. Franklin Inst., vol. 344, pp. 312348, 2007.

[12] Chun-Chu Jen and Shyr-Shen Yu, ”Automatic detection of abnormal mammograms in mammographic images.” In: Expert Systems with Applications 42.6 (2015), pp. 3048- 3055.

[13] Sneha Simon, R Lavanya, Devi Vijayan, ”PSO based Density Classifier for Mammograms”, The 16th International Conference on Biomedical Engineering pp 62-66

 [14] Singh, V. P., Gupta, A., Singh, S., and Srivastava, R. ” An efficient content based image retrieval for normal and abnormal mammograms.” In Electrical Computer and Electronics (UPCON), pp. 1-6 ,2015

[15] Singh, V. P., Srivastava, A., Kulshreshtha, D., Chaudhary, A., and Srivastava, R. ” Mammogram classification using selected GLCM features and random forest classifier.” International Journal of Computer Science and Information Security, 14(6) 2016: 82.

[16] Tang, Jinshan, Rangaraj M. Rangayyan, Jun Xu, Issam El Naqa, and Yongyi Yang. ”Computer-aided detection and diagnosis of breast cancer with mammography: recent advances.” IEEE Transactions on Information Technology in Biomedicine 13, no. 2 2009: 236- 251.

[17] Ganesan, Karthikeyan, U. Rajendra Acharya, Chua Kuang Chua, Lim Choo Min, K. Thomas Abraham, and Kwan-Hoong Ng. ”Computer-aided breast cancer detection using mammograms: a review. ” IEEE Reviews in biomedical engineering 6 2013: 77-98.

[18] Jalalian, Afsaneh, Syamsiah BT Mashohor, Hajjah Rozi Mahmud, M. Iqbal B. Saripan, Abdul Rahman B. Ramli, and Babak Karasfi. ”Computer-aided detection diagnosis of breast cancer in mammography and ultrasound: a review.” Clinical imaging 37, no. 3 2013: 420-426.

[19] Warren Burhenne, Linda J., Susan A. Wood, Carl J. D’Orsi, Stephen A. Feig, Daniel B. Kopans, Kathryn F. O’Shaughnessy, Edward A. Sickles, Laszlo Tabar, Carl J. Vyborny, and Ronald A. Castellino. ”Potential contribution of computer-aided detection to the sensitivity of screening mammography.” Radiology 215, no. 2 2000: 554-562.

[20] Rui, Yong, Thomas S. Huang, and Shih-Fu Chang. ”Image retrieval: Current techniques, promising directions, and open issues.” Journal of visual communication and image representation 10, no. 1 1999: 39-62.

[21] Smeulders, Arnold WM, Marcel Worring, Simone Santini, Amarnath Gupta, and Ramesh Jain. ”Content-based image retrieval at the end of the early years.” IEEE Transactions on Pattern Analysis and Machine Intelligence 12 2000: 1349-1380.

[22] Long, Fuhui, Hongjiang Zhang, and David Dagan Feng. ”Fundamentals of content-based image retrieval.” In Multimedia Information Retrieval and Management, pp. 1-26. Springer, Berlin, Heidelberg, 2003.

[23] Liu, Ying, Dengsheng Zhang, Guojun Lu, and Wei-Ying Ma. ”A survey of content-based image retrieval with high-level semantics.” Pattern recognition 40, no. 1 2007: 262-282.

[24] Datta, Ritendra, Dhiraj Joshi, Jia Li, and James Z. Wang. ”Image retrieval: Ideas, influences, and trends of the new age.” ACM Computing Surveys (Csur) 40, no. 2 2008: 5.

[25] Muller, Henning, Nicolas Michoux, David Bandon, and Antoine Geissbuhler. ”A review of content-based image retrieval systems in medical applicationsclinical benefits and future directions.” International journal of medical informatics 73, no. 1 2004: 1-23.

[26] Passalis, Nikolaos, and Anastasios Tefas. ”Learning neural bag-of-features for large-scale image retrieval.” IEEE Transactions on Systems, Man, and Cybernetics: Systems 47, no. 10 2017: 2641-2652.

[27] Passalis, Nikolaos, and Anastasios Tefas. ”Neural bag-of-features learning.” Pattern Recognition 64 2017: 277-294.

[28] S. Lazebnik, C. Schmid, J. Ponce, ”Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories”, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2, pp. 2169-2178, 2006.

[29] Y.-G. Jiang, C.-W. Ngo, J. Yang, ”Towards optimal bag-of-features for object categorization and semantic video retrieval”, Proc. ACM Int. Conf. Image Video Retrieval, pp. 494-501, 2007.

[30] S. Lazebnik, C. Schmid, J. Ponce, ”Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories”, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2, pp. 2169-2178, 2006.

[31] Menglin Jiang, Shaoting Zhang, Member, IEEE, Hongsheng Li, Member, IEEE,and Dimitris N. Metaxas, Senior Member, IEEE, ”Computer Aided Diagnosis of mammographic Masses Using Scalable Image retrieval”, IEEE Transactions on Biomedical Engineering, Vol. 62, No. 2, February 2015

[32] Jingjing Liu, Shaoting Zhang, Senior Member, IEEE, Wei Liu, Member, IEEE, Cheng Deng, Member, IEEE, ”Scalable Mammogram Retrieval Using Composite Anchor Graph Hashing With Iterative Quantization”, Systems For Video Technology, November 2017

[33] Devang Kulshreshtha, Vibhav Prakash Singh, Content-Based Mammogram Retrieval Using k-means Clustering and Local Binary Pattern 2nd International Conference on Image, Vision and Computing, 2017

[34] Singh, V. P., Srivastava, A., Kulshreshtha, D., Chaudhary, A., and Srivastava, R. ”Mammogram classification using selected GLCM features and random forest classifier.” International Journal of Computer Science and Information Security, 14(6) 2016: 82.

 [35] Li, Haixia, Xianjing Meng, Tingwen Wang, Yuchun Tang, and Yilong Yin. ”Breast masses in mammography classification with local contour features.” Biomedical engineering online 16, no. 1 2017: 44.

[36] B. Andr e, T. Vercauteren, A. M. Buchner, M. B. Wallace, and N. Ayache, Learning semantic and visual similarity for endomicroscopy video retrieval, IEEE Trans. Med. Imag., vol. 31, no. 6, pp. 12761288, Jun. 2012.

[37] J. Liu, S. Zhang, W. Liu, X. Zhang, and D. N. Metaxas, Scalable mammogram retrieval using anchor graph hashing, in Proc. IEEE Int. Symp. Biomed. Imaging 2014, pp. 898901

 [38] K. Mikolajczyk and C. Schmid, A performance evaluation of local descriptors, IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 10, pp. 16151630, Oct. 2005.

[39] X. Wang, M. Yang, T. Cour, S. Zhu, K. Yu, and T. X. Han, Contextual weighting for vocabulary tree based image retrieval, in Proc. IEEE Int. Conf. Comput. Vis., 2011, pp. 209216.

[40] D. G. Lowe, Distinctive image features from scale-invariant keypoints, Int. J. Comput. Vis., vol. 60, no. 2, pp. 91110, 2004.

[41] N. Passalis, A. Tefas, Entropy optimized feature-based Bag-of-Words Representation for Information Retrieval, IEEE Transactions on Knowledge and Data Engineering, 2016.

[42] Jiu, Mingyuan, Christian Wolf, Christophe Garcia, and Atilla Baskurt. ”Supervised learning and codebook optimization for bag-of-words models.” Cognitive Computation 4, no. 4 (2012): 409-419.

[43] Csurka G, Dance C, Fan LX, Willamowski J, Bray C.” Visual categorization with bags of keypoints.” In: Proceedings of ECCV international workshop on statistical learning in computer vision. 2004. p. 122. [44] Bishop CM. ”Neural networks for pattern recognition.” Oxford: Oxford university press; 1994. p. 14045

[45]http://marathon.csee.usf.edu/Mammography/Database.html

[46] Liu J, Yang Y, Shah M. ”Learning semantic visual vocabularies using diffusion distance.”In: CVPR. 2009. p. 461-68

[47] H. Alto, R. M. Rangayyan, and J. E. L. Desautels, Content-based retrieval and analysis ofmammographicmasses, J. Electron. Imag., vol. 14, no. 2, pp. 023 0161023 01617, 2005

[48] Oral, Canan, and Hatice Sezgin. ”Effects of dimension reduction in mammograms classification.” In 2013 8th International Conference on Electrical and Electronics Engineering (ELECO), pp. 630-633. IEEE, 2013.

[49] Matos, C.E.F., Souza, J.C., Diniz, J.O.B., Junior, G.B., de Paiva, A.C., de Almeida, J.D.S., da Rocha, S.V. and Silva, A.C., 2018. Diagnosis of breast tissue in mammography images based local feature descriptors. Multimedia Tools and Applications, pp.1-26.

[50] Debelee, Taye Girma, Mohammadreza Amirian, Achim Ibenthal, Gnther Palm, and Friedhelm Schwenker. ”Classification of Mammograms Using Convolutional Neural Network Based Feature Extraction.” In International Conference on Information and Communication Technology for Develoment for Africa, pp. 89-98. Springer, Cham, 2017.