

Feature Extraction for Breast Cancer Classification: A Comparative Study for Multiple Subsets of Features

Ass. Prof. Dr. Ismael Hadi Challob
ismaelhadi73@yahoo.com

Hanan Yasir Abd-Al Hassan
hananit1991@gmail.com

Department of Information Techniques
Technical Collage of Management \ Baghdad

Abstract— the most doctors spend a large part of their time looking at a benign tissue, which can easily be distinguished from cancer in most cases. This represents a waste of time and resources that could be better spent analyzing patients and to focus on cases where the disease is difficult to determine the classification or served with non-standard features. As a result, many researchers began to develop diagnostic methods of computer-aided through the application of image processing and computer vision techniques in an attempt to determine the spatial location of diseases such as breast cancer. This paper provides a preview of some work in progress on the computer system to support breast cancer diagnosis. For breast cancer diagnosis, the shape of the nuclei and the architectural pattern of the tissue are evaluated under high and low magnifications. In this study, the focus is on the development of classification prototype for the assessment of breast cancer images based on Fine Needle Aspiration. The parts of this study include: image segmentation process, features extraction process, then followed by classification process. The automatic system of malignancy classification was applied on a set of medical images. Three subsets of features (binary, color, and textural) features are used for comparison. Three classifiers (SVM, SOM, and KNN) are used to classify medical data for diagnosis. Color features and KNN classifier show the best accuracy among others.

Keywords—

FNA; Breast Cancer; Segmentation; Feature extraction; Classification.

1- INTRODUCTION

According to the International Agency for Research on Cancer, breast cancer is the most common cancer among women. In 2008, there were 1384155 diagnosed cases of breast cancer and 458503 deaths caused by the disease worldwide and according to the World Health Organization (WHO), breast cancer (BC) is one of the most deadly cancers diagnosed among middle-aged women aged from 40 to 60, and according to this organization there are 7.6 million deaths worldwide due to cancer each year, out of which 502,000 are caused by breast cancer alone. With such a high rate, breast cancer also is one of the most deadly cancers. For many years researchers have been trying to find the best way to treat breast cancer. Successful treatment is a key to reduce the high death rate.

Cancers in their early stages are vulnerable to treatment while cancers in their most advanced stages are usually almost impossible to treat [1]. The prognosis in breast cancer is strongly dependent on the disease development before any treatment is applied so the chance of recovery is a function of time of the detection of cancer. Modern medicine does not provide one hundred percent reliable, if possible cheap and at the same time non-invasive diagnostic methods for the diagnosis of breast pathology. As a result, in practice the important function acting in breast cancer diagnosis is the called (triple-test) [14], which is based on the summary of results of three medical examinations with different degrees of sensitivity and it allows to achieving high confidence of diagnosis.

The triple-test includes: Self-examination (palpation) of breast, mammography imaging of breast and FNB & FNA (Fine Needle Biopsy, Fine Needle Aspiration). The material (collected cells) is examined using microscope in order to confirm or exclude the presence of cancerous cells. The present approach requires a deep knowledge and experience of the cytologist responsible for diagnosis. In short, some pathologists can diagnose better than others. In order to make the decision independent of the arbitrary factor, morph metric analysis can be applied. Objective analysis of microscopic images of cells has been a goal of human pathology and cytology since the middle of the 19th century.

This work presents a method that allows distinguish malignant cells from the benign cells based on the analysis of cytological images of FNA material. The task

at hand is to classify a case (FNA of a patient) as benign or malignant. This is done by using morph metric, textural and topological features of nuclei isolated from microscopic images of the tumor. In previous work, the classification of the tumor is based on morph metric examination of cell nuclei.

In contrast to normal and benign nuclei, which are typically uniform in appearance, cancerous nuclei are characterized by irregular morphology that is reflected in several parameters. Morph metric measurements characterizing the size, cell grouping and color changes within the nuclei have been mainly used for feature extraction. It was decided not to use shape features because previous work showed that shape factors do not have good discriminative properties [19].

The quality of segmentation and feature extraction was tested by using the classifying algorithm .In this work three different classification methods was used to rate the feature subset : SVM , SOM , KNN . Infections as well as many other disorders In the literature one can find approaches to breast cancer classification [4,5,6,7,8,9,10,11,12,13,14,15,16]. The paper is divided into three sections. Section 1 gives an overview of breast cancer diagnosis techniques. Section 2 describes the process of acquisition of images used to breast cancer diagnosis. Section 3 deals with segmentation algorithm of image used to breast cancer diagnosis. Section 4 shows the experimental results obtained using the proposed approach. The last part of the work includes a conclusions and bibliography.

2- RELATED WORK

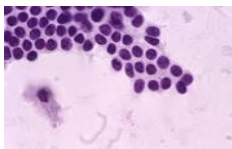
A. *A Historical Perspective In Medical Diagnosis Of Breast Cancer*

With the exception of surgical procedures, there are three main techniques normally used to assess breast mass: a physical examination, and mammography and suction cells fine-needle aspiration (FNA). It is called the tests the diagnosis triple assessment. Differ accuracy diagnosis of each one for the other, where it is estimated that out of his way accuracy 8 to 38 percent of breast cancers cannot be detected by palpation alone. When added to mammography screening physical, it has been detected up to 85% before the surgery. With the addition of suction cells fine-needle aspiration, it was diagnosed from 93 to 100% of cancers accurately [17, 18, 19].

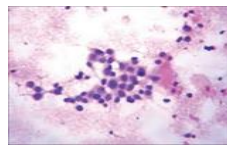
B. Origin and Acquisition of the Image:

Since the development of the first automated high-resolution whole-slide imaging (WSI) system by (Wetzel and Gilbertson in 1999), interest in using WSI for different applications in pathology practice has steadily grown. With advanced scanner technology (WSI) and cost effectiveness of the scanners and so could dissect tissue slices are now to be digitized and stored in the form of a digital image, with the availability and analyze much larger than the variables set along with imaging techniques and sophisticated analysis, it can be quickly replaced from the traditional model of the coroner microscopic examination with the coroner digital rely on a plate large flat screen to view and analyze digital tissue sections quickly. [20, 21].

Testing existing and new developed algorithms requires having databases at disposal, on which tests and benchmarks can be realized, especially in the domain of image analysis where in many problems domain knowledge need to be taken into account. Probably, the most commonly known data set of FNA images is Wisconsin Database of Breast Cancer (WDBC), which can be obtained from repository of machine learning database University of California. WDBC contains both raw images and visually extracted features, but the quality of images is rather poor and it does not fit for automatic feature assessment. In our study, a new data set that could be applied for completely automatic process of image analysis is designed [21].



A



B

Fig. 1: microscope image.

(A) Benign case, (B) malignant case

C. Computer-Aided Diagnosis procedure:

Breast cancer diagnosis is a very wide field of research studying only medical issues but also computer science issues. Breast cancer diagnosis is a multi-stage



process that involves different diagnostic examinations. Pattern classification is a well-known problem in the field of Artificial Intelligence concerned with the discrimination between classes of different objects. In this paper, the same techniques can be used in cancer diagnosis to assist doctors with their decisions.

This field of breast cancer examination is also an interest to many scientists, I will concentrate on some of the techniques used for classification and detection of the cancerous nuclei since it is very closely related to the research presented in this paper. Most automatic diagnostic systems are based on similar configuration of several steps. At the beginning images are adjusted in preprocessing phase. Then objects of interest are extracted from the images in segmentation step, which is the most challenging task. For separated objects morph metric and colorimetric features are calculated. Finally, objects are classified, e.g. as benign or malignant case [2, 3, 4, 5, 15, 16].

3-PROPOSED PROTOTYPE:

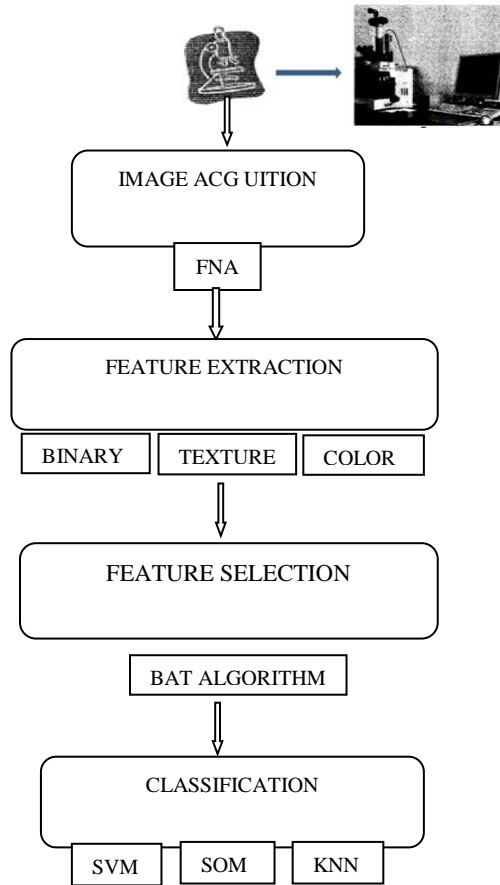


Fig. 2: The proposed prototype

3-1 collecting the images of Study :

The process of obtaining the first phase of the proposed system in the letters , and in this research were obtained digital images, a microscopic images (microscope image) as a special technique using WSI (Whole – Slide Image that contains a special camera to take pictures optical microscope technology that is displayed on the computer screen tied to a microscope using a special display for viewing images and software as in Figure follows[2] :

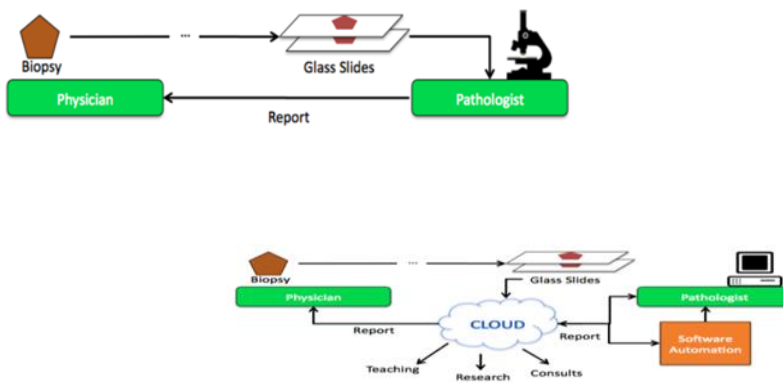


Fig. 3: The process of obtaining the images show [2]

3-2 Segmentation of Nuclei:

To determine whether the tumor is benign or malignant, cell nuclei need to be isolated from the background and from other objects on the image (e.g., red blood cells). Then from the nuclei certain features can be extracted and the malignancy determined. Because all subsequent analysis is based on the results obtained in the segmentation step, it is very important that the nuclei were properly

extracted. In the literature, many different approaches have been already proposed to extract cells or nuclei from microscope images [22, 23]. This task is usually done automatically or semi-automatically, using one of the well-known methods of image segmentation [4, 5, 15, 16, 17, and 22]. However, reliable nuclei segmentation is a challenging task. FNB images are particularly difficult due to the way they are prepared. The material is extracted by a needle and smeared on a glass. This may result in partial destruction of the tissue structure, and some time seven of nuclei. The cells are usually not uniformly distributed on the preparation. They often form 3-D structures, and they can possibly be in contact with and/or occluded by other cells. In this study, the classical region growing technique with (clustering segmentation method) is combined by using (Fuzzy C-mean) algorithm.

Fuzzy clustering is a power full unsupervised method for the analysis of data and construction of models, in many situations fuzzy clustering is more natural than hard clustering. Objects on the boundaries between several classes are not forced to fully belong to one of the classes, but rather are assigned membership degrees between 0 and 1 indicating their partial membership. Fuzzy C-mean is a method which allows one piece of data to belong to two or more clusters.

The general method is developed by Jim Bezdek in 1973, improved by Dunn in 1974, and Bezdek in 1981 is frequently used in pattern recognition. This algorithm works by assigning membership to each data point corresponding to each cluster center on the basis of distance between the cluster center and the data point. More the data is near to the cluster center more is its membership towards the particular cluster center. Clearly, summation of membership of each data point should be equal to one [23,24,25].

1-Initialize $U = [u_{ij}]$ matrix , u^0

2-At K – step : calculate the centers vectors $c^k = [C_j]$ with $u^{(k)}$

$$C_i = \frac{\sum_{j=1}^n u_{ij}^m x_j}{\sum_{j=1}^n u_{ij}^m}$$

3- Update $U^{(k)}$, $U^{(k+1)}$

4- Calculated: $d_{ij} = \sqrt{\sum_{i=1}^n (X_i - C_i)^2}$

$$u_{ij} = \frac{1}{\sum_{k=1}^c \frac{d_{ij}^{2/(m-1)}}{d_{kj}^{2/(m-1)}}}$$

5-If $\|U(k+1) - U(k)\| < E$ then STOP; other wise return To step 2.

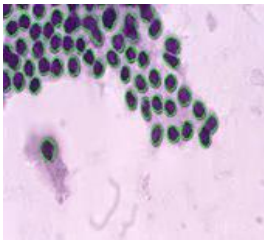
Here m is any red number greater than 1

u_{ij} is the degree of member ship of X_i in the cluster j ,

X_i is the i the of d- dimensional measured data ,

C_j is the d- dimension center of the cluster ,

The algorithm work by assigning membership to each data point corresponding to each cluster on the basis of distance between the cluster center and the data point. More the data is near to the cluster center. Clearly, summation of member ship of each data point should be equal to one after each iteration member ship and cluster centers are update according to the formula.



A

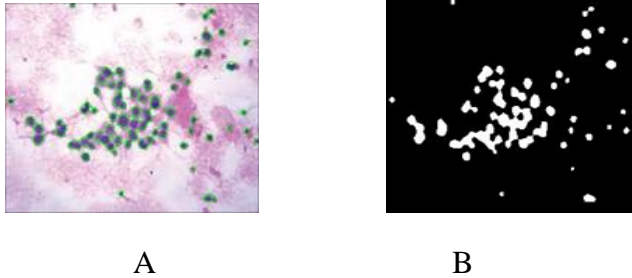


B

A) Detected circles imposed on the origin

B) Nuclei mask Image

Fig. 4: Manually segmented FNA images (benign case)



A) Detected circles imposed on the original

B) Nuclei mask image

Fig. 5: Manually segmented FNA images (malignant case)

3-3 Feature Extraction:

In contrast to normal and benign cells, which are typically uniform in appearance malignant cells are characterized by irregular morphology that is reflected in several parameters. Morph metric measurements characterizing the shape and size have been mainly used for feature extraction. The extracted features are [4]:(Binary features , Texture features , color features). After the isolation of nuclei from the images, as determined by the circles classified as correct in the previous step, features are extracted and used in the classification procedure. In our approach the features chosen reflect the observations of cytologists.

The key features associated with the diagnosis of breast cancer as used by specialists can be divided into three groups.

The first groups are features related to the size of the nuclei. A large variation of sizes in the image suggests malignancy of the tumor Small uniform nuclei argue for a benign case. This group is represented by the area and perimeter of a nucleus. Another important feature is the distribution of chromatin in the nuclei of healthy cells. Frequent occurrences of distinct lumps of chromatin may indicate presence of

cancer [16]. The second group of features represent this dependence with texture features based on gray-level co-occurrence matrix (GLCM) [25] and gray-level run-length matrix (GLRLM) [26,27], as well as the mean and variance of pixel values in each RGB channel [3]. The last group of features is related to the distribution of nuclei in the image. Healthy tissues usually form single-layered structures while cancerous cells tend to break up which increases the probability of encountering separated nuclei. To express this relation, features representing the distance to centroid of all nuclei are used, and the distance to c-nearest nuclei below is a detailed description of all the features:

A- Binary Features:

Features calculated based on the binary image. A set of nuclei in the image, $N = \{N_1, N_2 \dots N_n\}$ can be defined as a collection of all connected components and the nucleus N_i is a set of pixels that are contained in the extracted nuclei. The features are:

- 1- **Area:** The actual number of pixels of the nucleus.
- 2- **Perimeter:** The sum of distances between each adjoining pair of pixels around the border of the nucleus.
- 3- **Eccentricity:** The ratio of the distance between the foci of the ellipse and its major axis length.
- 4- **Major axis length:** the length of the major axis of the ellipse that has the same normalized second central moments as the region.
- 5- **Minor axis length:** the length of the minor axis of the ellipse that has the same normalized second central moments as the region.
- 6- **Distance from the centroid:** the Euclidean distance between the geometric center of the nucleus and mean of geometric centers of all the nuclei in the image.
- 7- **Average area:** Is calculated as the average number of nuclei pixel.
- 8- **Dispersion (D):** is defined as a variation of cluster areas (A_c).

9- **Convexity (Ci)**: Is calculated as the ratio of nucleus area and its convex hull which is the area of the minimal convex polygon that contains the nucleus.

10- **Number of groups (NG)**: calculate the number of groups in the image that weren't removed during the segmentation process.

B-Textural features:

Textural features are used to measure the texture information of the image. Here, the texture of the nucleus is taken into consideration. To extract textural features, a co-occurrence matrix is calculated, which provides us with information about the relation of pairs of pixels and corresponding grey level.

1- Gray-Level Co-Occurrence Matrix Feature:

The first four features are based on GLCM. The $N \times N$ matrix P , where N is the number of gray levels, is defined over an image to be the distribution of co-occurring values of pixels at a given offset. In other words each element of P specifies the number of times a pixel with gray-level value i occurs shifted by a given distance to a pixel with the value j [17]. In our case, the mean of four GLCMs is calculated and determined for offsets corresponding to 0° , 45° , 90° and 135° using eight gray-levels. The textural features are co-occurrence features calculated using GLCMs are

- Contrast: The intensity contrast between a pixel and its neighbor over the whole image.
- Correlation: The correlation of a pixel to its neighbor over the whole image.
- Energy: In literature also known as uniformity—the sum of squared elements in the GLCM.
- Homogeneity: The closeness of the distribution of elements in the GLCM to the GLCM diagonal.

2- Gray-Level Run-Length Matrix Features

The remaining eleven texture features are based on GLRLM. The $N \times M$ matrix p , where N is the number of gray levels and M is the maximum run length, is defined for a given image as the number of runs with pixels of gray level i and run length j [27]. Similarly to the GLCM, run length matrices is used for 0° , 45° , 90° and 135° angles using eight gray-levels. The textural features are run-length features calculated using GLRLMs are:

SRE: Short run emphasis.

LRE: Long run emphasis.

GLN: Gray-level non-uniformity.

RLN: Run length non-uniformity.

RP : Run percentage.

LGRE: Low gray-level run emphasis.

HGRE: High gray-level run emphasis.

SRLGE: Short run low gray-level emphasis.

SRHGE: Short run high gray-level emphasis.

LRLGE: Long run low gray-level emphasis.

LRHGE: Long run high gray-level emphasis.

C. Color Features:

Color images consist of three components each representing a primary color. Each of these components can be treated as a separate intensity image. To calculate color features any of the previously defined features and apply it to each color band can be used. This requires three times as many calculations as for the grey level image.

To overcome this problem, the spherical coordinate transform (SCT) can be applied to the RGB image which takes into consideration the relation between the RGB channels. The features are: (Red channel, Green channel, Blue channel, Gray Level of image) [3].

3-4 Classification Methods:

A set of 3 different classifiers was used to test the effectiveness of the features in diagnosing new samples. It was decided to use well known classification algorithms such as k-nearest neighbor (KNN), Support vector machine (SVM), and Self – Organization map (SOM) [4,5,6,7,8,9,10,11, 12,13, 14, 15,16,]. The idea behind using such number of classification techniques was to check how the method can influence the classification accuracy.

4- Experimental Results:

The features were tested for the classification efficiency, which is defined as a percentage of successfully recognized cases among all cases. For classification, (SVM, SOM, KNN) classifiers with value $k=10$ are used. There were 296 patients: 167 benign and 138 malignant. Each patient was represented by 1 image. The image belonging to the same patient was never at the same time in the training and testing set. The final diagnosis was obtained by a majority voting of the classification of individual images belonging to the patient.

The results show that there are three types' features providing important diagnostic information. They are (Binary, Color, Textural) features. Taking into account the fact the different subset can be optimal for different classifiers.

Table 1: Classification results after feature extraction

(All features)

Testing ratio	K= 10		
	SVM	KNN	SOM
10%	0.4448	0.5977	0.59172
30%	0.54324	0.63821	0.58434
50%	0.60082	0.6417	0.54132

Table 2: Classification results after feature extraction (Binary Features)

(Binary Features)

Testing ratio	K= 10		
	SVM	KNN	SOM
10%	0.4864	0.53224	0.5016
30%	0.53682	0.5258	0.53444
50%	0.5447	0.5458	0.55602

Table 3: Classification results after feature extraction (Color features)

(Color features)

Testing ratio	K= 10		
	SVM	KNN	SOM
10%	0.472	0.6225	0.48002
30%	0.53562	0.6425	0.54886
50%	0.54142	0.6625	0.53436

Table 4: Classification results after feature extraction (Texture features)

(Texture features)

Testing ratio	K= 10		
	SVM	KNN	SOM
10%	0.4192	0.5142	0.50468
30%	0.524	0.5342	0.53582
50%	0.52814	0.5542	0.57624

The results in Table 1 (Classification accuracy for all features) show that the best classification rate (50%) was obtained in KNN classifiers (using all three types of features: Binary, Color, and Textural and the weak result in classification rate (10%) were obtained in SVM using all three types of features. While the result in Table 2 shows that the best classification rate (50%) was obtained in (SOM) using subset of features (Binary) feature and the weak result in classification rate is (10%) obtained in (SVM) using subset of feature (Binary) feature. The result in Table 3 shows that the best classification rate (50%) was obtained for (KNN) using subset of features (Color) and the weak result in classification rate (10%) were obtained in (SVM) using subset of feature (Color) feature. Where the result in Table 4 show that the best classification rate (50%) was obtained in (SOM) using subset of features (Textural) and the weak result classification rate (10%) was obtained in (SVM) using subset of feature (Textural features).Color features show the best accuracy among other features.

5- Conclusion:

The aim of the work was to test whether three types of features (Binary, Color, Textural) might provide essential diagnostic information in automatic breast cancer diagnosis based on analysis of FNA images. In experiments, 296 cytological images from patients are used. In order to segment nuclei, a Fuzzy C-mean clustering is used where in feature extraction 10-Binary features, 4-Color features and (4-GLCM, 11-GLRLM) Textural features are used. Three classifiers (SVM, SOM, and KNN) are used for Classification process. The experiments show that KNN classifier gives the best results. Whereas Color-based features presents best accuracy among other Binary and Textural features. There are three challenges for the near future. First, the recognition rate should be improved by adding more sophisticated features that are not tested during current investigations. As a second challenge, the proposed approach must be applied for automatically segmented images. So, previously developed segmentation algorithms must be extended to deal properly with overlapped cells. Finally, the whole segmentation and classification system need to be applied for virtual slides generated by virtual scopes which are able to produce images. Such huge slides require a long analysis, respectively so it will be very helpful if automatic system can recognize suspected fragments of the slide and automatically present them in the first place.

REFERENCES

- [1] Bloom, H. J. G., & Richardson, W. W. (1957). Histological grading and prognosis in breast cancer: a study of 1409 cases of which 359 have been followed for 15 years. *British journal of cancer*, 11(3), 359.
- [2] Cruz-Roa, A., Basavanhally, A., González, F., Gilmore, H., Feldman, M., Ganesan, S., & Madabhushi, A. (2014, March). Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. In *SPIE medical imaging* (pp. 904103-904103).
- [3] Jeleri, L. (2009). Computerized cancer malignancy grading of fine needle aspirates (Doctoral dissertation, Concordia University Montreal, Quebec. Canada).
- [4] Filipczuk, P., Kowal, M., & Marciniak, A. (2010). Feature selection for breast cancer malignancy classification problem. *Journal of Medical Informatics & Technologies*, 15, 193-199.
- [5] Kowal, M., Filipczuk, P., Obuchowicz, A., & Korbicz, J. (2011). Computer-aided diagnosis of breast cancer using Gaussian mixture cytological image segmentation. *Journal of Medical Informatics & Technologies*, 17.
- [6] Marciniak, A., Obuchowicz, A., Monczak, R., & Kołodziński, M. (2005). Cytomorphometry of fine needle biopsy material from the cancer. In *Computer Recognition Systems* (pp. 603-609). Springer Berlin Heidelberg.
- [7] Doyle, S., Agner, S., Madabhushi, A., Feldman, M., & Tomaszewski, J. (2008, May). Automated grading of breast cancer histopathology using spectral clustering with textural and architectural image features. In *2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (pp. 496-499). IEEE.
- [8] Pérez, N., Guevara, M. A., & Silva, A. (2012). Evaluation of features selection methods for breast cancer classification. In *Icem15: 15th International Conference on Experimental Mechanics* (p. 10)
- [9] Chaiwun, B., Settakorn, J., Ya-in, C., Wisedmongkol, W., Rangdaeng, S., & Thorner, P. (2002). Effectiveness of fine-needle aspiration cytology of breast: Analysis of 2,375 cases from northern Thailand. *Diagnostic cytopathology*, 26(3), 201-205.
- [10] JELEŃ Ł., FEVENS T., KRZYŻAK A., Classification of breast cancer malignancy using cytological images of fine needle aspiration biopsies, *Int. J. Appl. Math and Comp. Sci.* Vol. 18, No. 1, 2008, pp. 75–83.
- [11] Kuenen-Boumeester, V., Menke-Pluymers, M., De Kanter, A. Y., Obdeijn, I. M., Urich, D., & Van Der Kwast, T. H. (2003). Ultrasound-guided fine needle aspiration cytology of axillary lymph nodes in breast cancer patients. A preoperative staging procedure. *European Journal of Cancer*, 39(2), 170-174.
- [12] Nithya, R., & Santhi, B. (2011). Classification of normal and abnormal patterns in digital mammograms for diagnosis of breast cancer. *International Journal of Computer Applications*, 28(6), 21-25.

- [13] Lavanya, D., & Rani, D. K. U. (2011). Analysis of feature selection with classification: Breast cancer datasets. *Indian Journal of Computer Science and Engineering (IJCSE)*, 2(5), 756-763.
- [14] . Wolberg, W. H., Street, W. N., & Mangasarian, O. L. (1993). Breast cytology diagnosis via digital image analysis. *Analytical and Quantitative Cytology and Histology*, 15(6), 396-404.
- [15] Filipczuk, P., Fevens, T., Krzyżak, A., & Monczak, R. (2013). Computer-aided breast cancer diagnosis based on the analysis of cytological images of fine needle biopsies. *IEEE transactions on medical imaging*, 32(12), 2169-2178.
- [16] Mohanty, A. K., Beberta, S., & Lenka, S. K. (2011). Classifying benign and malignant mass using GLCM and GLRLM based texture features from mammogram. *International Journal of Engineering Research and Applications*, 1(3), 687-693.
- [17] Dowlatshahi, K., Yaremko, M. L., Kluskens, L. F., & Jokich, P. M. (1991). Nonpalpable breast lesions: findings of stereotaxic needle-core biopsy and fine-needle aspiration cytology . *Radiology*, 181(3), 745-750.
- [18] Rothenberg, K. H. (1997). Breast Cancer, the Genetic Quick Fix, and the Jewish Community: Ethical, Legal and Social Challenges. *Health Matrix*, 7, 97.
- [19] Bloom, H. J. G., & Richardson, W. W. (1957). Histological grading and prognosis in breast cancer: a study of 1409 cases of which 359 have been followed for 15 years. *British journal of cancer*, 11(3), 359.
- [20] Ghaznavi, F., Evans, A., Madabhushi, A., & Feldman, M. (2013). Digital imaging in pathology: whole-slide imaging and beyond. *Annual Review of Pathology: Mechanisms of Disease*, 8, 331-359.
- [21] Peters, W. H., & Ranson, W. F. (1982). Digital imaging techniques in experimental stress analysis. *Optical engineering*, 21(3), 213427-213427.
- [22] Zhu, K. J., Su, S. H., & Li, J. L. (2005). Optimal number of clusters and the best partition in fuzzy C-mean. *Systems Engineering-theory & Practice*, 3, 52-61.
- [23] Rezaee, M. R., Lelieveldt, B. P., & Reiber, J. H. (1998). A new cluster validity index for the fuzzy c-mean. *Pattern recognition letters*, 19(3), 237-246.
- [24] Suganya, R., & Shanthi, R. (2012). Fuzzy c-means algorithm-a review. *International Journal of Scientific and Research Publications*, 2(11), 1.
- [25] Haralick, R. M., & Shanmugam, K. (1973). Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6), 610-621.
- [26] Castellano, G., Bonilha, L., Li, L. M., & Cendes, F. (2004). Texture analysis of medical images. *Clinical radiology*, 59(12), 1061-1069.
- [27] Tang, X. (1998). Texture information in run-length matrices. *IEEE transactions on image processing*, 7(11), 1602-1609.