International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Impact Factor (2012): 3.358

Identification of Ovarian Mass: A Survey

Hemita Pathak[1], Vrushali Kulkarni[2], Sarika Bobde[3]

[1]PG Student, Dept of Computer Engineering, MIT, Pune, India.

[2]Associate Professor, Dept of Computer Engineering, MIT, Pune, India.

[3]Assistant Professor, Dept of Computer Engineering, MIT, Pune, India

Abstract: ***Ovarian Cancer is leading cause of cancer deaths in women today. Early detection of the cancer can reduce mortality rate. Studies have shown that radiologists can miss the detection of a significant proportion of abnormalities in addition to having high rates of false positives. To detect malignancy, methods of pattern recognition and image processing are used. Pattern recognition in image processing requires the extraction of features from regions of the image and the processing of these features with a pattern recognition algorithm. In recent age, cases of ovarian cancer are increasing day by day, so diagnosis of ovarian cancer should be appropriate and up to the mark. Ultrasound imaging is widely used for diagnosis over the other imaging modalities like Positron Emission Tomography (PET), Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) due to its noninvasive nature, portable, accurate, harmless to the human beings and capability of forming real time imaging. In this paper we have presented some of the techniques for diagnosis ovarian malignancy. Generally there are main three phases to detect the malignancy. First is pre-processing in that need to smooth the image, second is feature extraction from the image and third is classification of the features.***

Keywords: Ultrasound images, Pre-processing, Feature extraction, Classification

1. Introduction

Ultrasound imaging application in medicine and other fields is enormous. It has several advantages over other medical imaging modalities. The use of ultrasound in diagnosis is well established because of its noninvasive nature, low cost, capability of forming real time imaging.

As ovarian cancer is silent killer for women there are no symptoms of cancer and mostly it detect in the final stage. In the ovarian mass identification it plays key role from the ultrasound image it can be diagnose that cyst is malignant or benign and it can also be diagnose abnormality of the ovary. If the diagnosis of the ultrasound images is accurate than only few patients have to go for further biopsy. And it can be great help for the doctors as they can detect the disease in early stage and can be saved patient also.

From decades many researchers are trying to find efficient methods for efficient detection of malignant cyst. Up till now there are many methods have found out. Machine Learning's techniques are widely used for feature extraction and classification. In earlier time statistical method was used for detection but now a days machine learning's techniques are used because its accuracy. Generalized method for detection of ovarian cyst or malignancy is first to remove spackle or extraneous noise from the image. After removing noise features are extracted from image and at final stage classification is performs for extracted features. Algorithm like GLCM (gray level co-occurrence matrix), GLRLM (gray level run length matrix) are mostly used for the feature extraction. As feature extraction is also crucial in detection of the malignancy. Many classification algorithms like Decision Tree, SVM (support vector machine), Genetic Algorithm, Neural Network, etc. are used for the efficient classification of the extracted feature.

In this paper we have showed different methods to detect the malignancy. The subsequent discussions of this paper have been organized as follows:

Procedure for removal of the noise and to smooth the images pre-processing techniques are discussed in section 2. After pre-processing feature extraction from the images are discussed in section 3. In section 4 classification techniques are discussed for the obtained features. Section 5 concludes the paper and future work.

2. Approaches For Pre-Processing

Pre-processing of the images commonly involves removing low-frequency background noise, normalizing the intensity of the individual particles images, removing reflection and masking portions of images.

2.1 Approach-1

Golam et al [1] proposed technique for the preprocessing of the images, have two steps: merging and t-testing. Merging: Merging refers to creating a single data matrix from a number of sample files. This matrix is used for the preprocessing such as alignment, handling of missing values and refining of the dataSet. This merging is useful for efficient implementation of the algorithm.

T-test: T-test will be containing matrix a having cancer data set and data set by having non cancer data set. Now T-test will discard non-significant feature of the data set.

2.2 Approach-2

Usha et al [2] proposed technique for the preprocessing of the ultrasound images. Steps for the technique are: R-plane, ROI (Region of Interest), Spackle Filter and morphological operations.

2.2.1 R-plan

In this algorithm 24 bit RGB image is required but only R-

Licensed Under Creative Commons Attribution CC BY

plane is considered. And in every image position of ovary should be in center. So images containing ovary is extracted.

2.2.2 Spackle filter

To remove the noise from the ultrasound images spackle reducing anisotropic diffusion filter is used to denoising ultrasound images.

2.2.3 Morphological operations

Unwanted border and background of the images and are removed so that ovary can be clearly visible. Morphological operations erosion and dilation is performed. If ovary attached to the border then apply erosion thrice followed by clearing border, region filling and dilation thrice. With this operation ovary is segmented successfully from its background. By applying these methods can remove unwanted noise from ultrasound images.

3. Approaches For Feature Extraction

Feature extraction is spatial form of dimension reduction. Texture is one of the most important characteristics of an image. It is used to describe the local spatial variations in image brightness which is related to image properties such as coarseness, and regularity. This is achieved by performing numerical manipulation of digitized images to get quantitative measurements. Texture analysis is used to find a unique way of representing the underlying characteristics of textures and represent them in some simpler but unique form, so then they can be used to accurately and robustly classify and segment objects.

Normally texture analysis can be grouped into four categories: model-based, statistical-based, structural-based, and transform-based methods. Model-based methods are based on the concept of predicting pixel values based on a mathematical model. Statistical methods describe the image using pure numerical analysis of pixel intensity values. Structural approaches seek to understand the hierarchal structure of the image. Transform approaches generally perform some kind of modification to the image [3].

3.1 GLCM (Gray Level Co-occurrence Matrix)

Rajendra et al [4] proposed method for extraction of feature from given image.

3.1.1 Fractal dimension

It identifies irregularity in the pixel intensities of the image. Differential box counting with sequential algorithm is used to calculate FD [5]. Gray scale image is fed and grid size is of power 2 for better computation. Maximum and minimum intensities for each (2 x 2) box are obtained to sum their difference, which gives the *M* and *r* by *s/M* where *M=min(R, C)*, *s* is the scale factor, *R* and *C* are the number of rows and columns, respectively. When the grid size gets doubled, *R* and C reduce to half of their original value and above procedure is repeated iteratively until *max(R, C)* is greater than 2. Linear regression model is used to fit the line from plot *log (Nr)* vs. *log (1/r)* and the slope gives the *FD.*

3.1.2 GLCM (Gray Level Co-occurrence Matrix)

The elements of the GLCM C_d (i, j) are made up of the relative number of times the gray level pair (a, b) occurs when pixels are separated by the distance (a, b) = (1, 0)[20].

3.1.3 HOS (Higher Order Spectra)

Higher order statistics denote higher order moments (order greater than two) and non-linear combinations of higher order moments, called the higher order cumulants. They help to extract information on the phase and nonlinearities present in the signal [6].

From above methods features extracted are Normalized bispectrum entropy, normalized bispectral squared entropy, and normalized bispectral cubed entropy for every one degree of Radon Transform between 0 to 180 degrees. Thus, the total number of extracted features would be 724.

Md. Sohail et al [7] proposed direction for the feature extraction. GLCM (Gray Level Co-occurrence Matrix) based feature extraction*:* It provides second order method for extracting feature. Two parameters are taken in account d= inter distance difference and θ= orientation.

$$C_{i,j} = \frac{P_{i,j}}{\sum_{i,j=1}^{G} P_{i,j}} \quad (1)$$

$C_{i,j}$ = Co-occurrence probability between i and j
$P_{i,j}$ = number of occurrence of grey level i and j within window
G= quantized number of grey levels

To obtain features from the images four co-occurrence matrices from each image are calculated where θ= {0, 45, 90,135} degree and d=1 pixel. 14 statistical texture descriptors are calculated from each co-occurrence matrices proposed by Haralick et al [8]. Therefore, total 56 texture feature were extracted from each image.

3.2 GLRLM (Grey Level Run Length Matrix)

Md. Sohail et al [9] proposed method for feature extraction from images.

GLRLM (Grey Level Run Length Matrix): Method for extracting higher order statistical texture features [10].

G= number of gray levels, R= longest run, N= number or pixels in the image. It is a two dimensional matrix of (G×R) elements in which each element ($P_{i,j}$/ θ) gives the total number of occurrences of runs having length j of gray level i in a given direction θ.

Local relative GLRLM- based Texture Feature Extraction: Proposed method work on the principle of "Local Matching" [11]. For features divide images into uniform grids of three level partitioning.

Three levels partitioning for extracting local relative GLRLM-based features: For first level partitioning divide image into 4^k non-overlapping upper layer partitioning blocks(fig a). In the second level partitioning $\frac{4^k}{4}$ overlapping

Licensed Under Creative Commons Attribution CC BY

partitioning blocks are created. Therefore, $4^k \times \frac{4^k}{4} = 5\times4^{k-1}$ partition blocks are created during first and second level partitioning (fig b). Third level partitioning process divides each of this $5\times4^{k-1}$ upper layer blocks into 4 equal parts

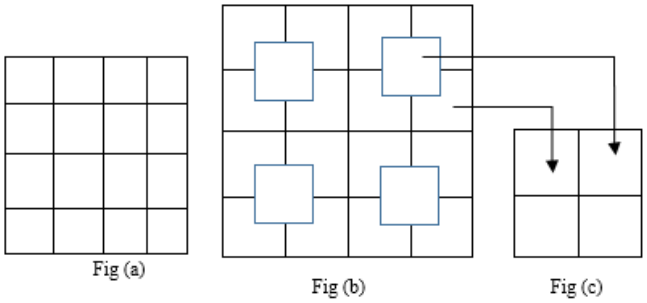

Figure 1: GLRLM Partitioning

Generating altogether 5×4^k lower layer partitioning sub-block (fig c)

3.3 Selection of optimal subset of image descriptors using MOGA (Multi Objective Genetic Algorithm)

Md. Sohail et al [12] proposed technique to find optimal descriptor from ultrasound image. In this papers they describe GLCM [7] features obtained by four co-occurrence matrices by using G= 256 and θ= {0, 45, 90, 135} and finally 19 descriptors were calculated. In GLRLM [9] feature extracted from four matrices by using G=64 and θ= {0, 45, 90, 135} and 11 descriptors were calculated.

Selection of optimal subset of image descriptors using MOGA (Multi Objective Genetic Algorithm): Method for selecting optimal feature descriptor through MOGA is demonstrated in figure 2.

3.3.1 Entropy based feature ranking

It removes irrelevant features, so it will reduce the entropy. It ranks features in descending order of the relevance. Therefore, for M dimensional feature vector, the method is repeated M times to find the rank of M features. The entropy measure of a data set of M instances (N feature vector) is calculated as:

$$E = -\sum_{i=1}^{N}\sum_{j}^{N}(S_{ij} \times \log S_{ij} + (1 - S_{ij}) \times \log(1 - S_{ij})) \quad (2)$$

S_{ij} = similarity measure normalized to [0, 1]

Feature Ranking with T-statistics*:* Let us consider that the data set consists of N instances

(Feature vector): $X_i = (X_1^i, X_2^i, \ldots, X_M^i)$; where, $1 \leq i \leq N$, and M is the dimension of each feature vector.

Ranking score T (X_j) is obtained by [13]

$$T(Xj) = \frac{|\mu_1^j - \mu_2^j|}{\sqrt{\frac{(\sigma_1^j)^2}{M^1} + \frac{(\sigma_2^j)^2}{M^2}}} \quad (3)$$

The final score of an attribute (feature) is obtained by taking average of the scores calculated for that particular attribute considering all the classes. When making selection, features with the highest scores are considered as the most discriminatory features.

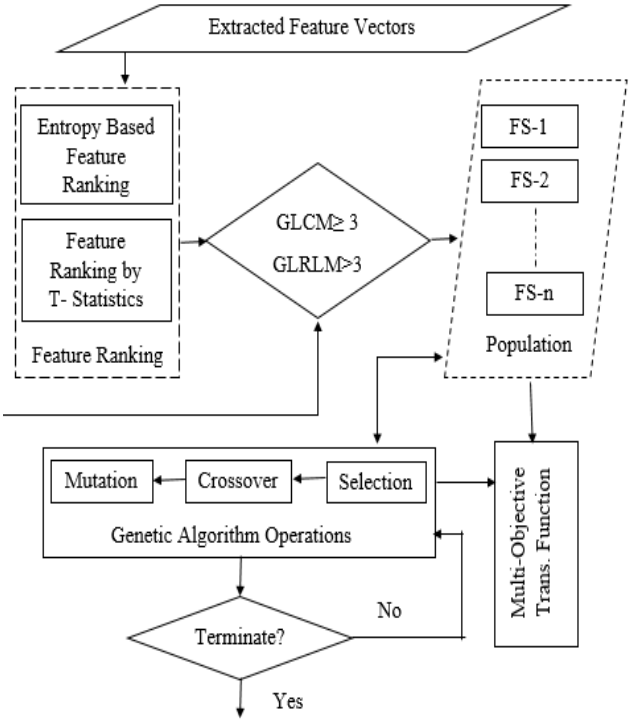

Figure 2: Multi Objective Genetic Algorithm

3.3.2 Fitness calculation

A fitness value is associated with each chromosome that represents the degree of fitness of the solution.

$$f(v) = F(g(v)) = \alpha S_b(g(v)) - \beta S_w(g(v)) \quad (4)$$

v = binary encoded chromosome

g(v) = mapping function that maps binary encoded chromosome to the original dataset.

Operation of Genetic Algorithm Selection*:* The selection operation determines which parent chromosome may participate in producing offspring for the next generation.

Crossover: A standard n point crossover operator, operating on two individuals with coding subsets of size m1 and, tends to yield offspring with complexity of approximately $\frac{(m1+m2)}{2}$.

Mutation: mutations are performed by flipping randomly one or more bits in a single parent chromosome to generate an offspring.

4. Approaches for Classification

Classification is the problem of identifying to which of a set of categories, a new observation belongs, on the basis of a training set of data containing observations (or instances) whose category membership is known. An example would be assigning a given email into "spam" or "non-spam" classes or assigning a diagnosis to a given patient as described by observed characteristics of the patient.

Licensed Under Creative Commons Attribution CC BY

3.4 Decision Tree

Decision tree Learned from class label tuples. Decision tree are hierarchically arranged questions and answers that lead to classification. In decision tree, each internal node denotes a test on an attribute, each branch represents an outcome of test, and each leaf node holds a class label, for an unknown tuple X, path is traced from root to leaf node for prediction of class. Rajednra et al [4] proposed method for classification.

In the case of Decision Trees (DT), the input features are used to construct a tree, and then a set of rules for the different classes are derived from the tree [14]. These rules are used for determining the class of an incoming new image.
DT classifier will give output with high accuracy of 95.1% and predicts class of the patient.

3.5 Support Vector Machine

Golam et al [1] proposed method to classify texture feature.

Approach -1: Support vector machine [15] is supervised machine learning technique. In this method data are mapped into a higher dimensional input space and construct an optimal hyperplane for separating these data. The best separation is obtained by the hyperplane that has the largest distance to the nearest training data points of any class samples. SVM maps the original input space by using a kernel function such as linear kernel, quadratic kernel, Gaussian radial basis function kernel, polynomial kernel and multilayer perceptron kernel. Kernel function $k(\chi_i;\chi_j)$ computes the inner product of two vectors in the feature space with the mapping function:

$$k(\chi_i, \chi_j) = \phi(\chi_i)\,\phi(\chi_j) = z_i \cdot z_j \quad (5)$$

Three types of commonly used kernel functions are:

Linear Kernel $k(\chi_i; \chi_j) = \chi_i, \chi_j$ (6)

Polynomial kernel $k(\chi_i;\chi_j) = \chi_i,\chi_j$ (7)

Gaussian kernel $k(\chi_i; \chi_j) = \left(\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)$ (8)

Linear kernel is used for this implementation.
It gives average 99% accuracy when 2-fold to 10-fold classifier is used.

Approach -2: Md. Sohail et al [7] proposed method for classification. Support Vector Machine (SVM) belongs to the class of maximum margin classifiers. They perform pattern recognition between two classes by finding a decision surface that has the maximum distance to the closest points in the training set known as support vectors [16]. Assuming linearly separable data, the goal of maximum margin classification is to separate the two classes by a hyperplane such that the distance to the support vectors is maximized. This hyperplane is known as the *Optimal Separating Hyperplane* and is expressed as:

$$f(x) = \sum_{i=1}^{L} a_i y_i x_i x + b \quad (9)$$

α_i and b are solutions of a quadratic programming problem. Each data point χ is separated from hyperplane as:

$$d(x) = \frac{\sum_{i=1}^{L} a_i y_i\, x_i \cdot x + b}{\left\|\sum_{i=1}^{L} a_i y_i x_i \cdot x\right\|} \quad (10)$$

d= classification result of χ

Each point χ in the input space is mapped to a point

$z= \phi(\chi)$ of a higher dimensional space, called the *feature space*, where the data are separated by a hyperplane. The key property in this construction is that the mapping $\phi(.)$ is subject to the condition that the dot product of two points in the feature space $\phi(\chi)$. $\Phi(y)$ can be rewritten as a kernel function k (χ, y). Using the kernel function, the decision surface is defined by:

$$f(x) = \sum_{i=1}^{L} y_i a_i k(x, x_i) + b \quad (11)$$

Finally, multi-class classification was performed by arranging $\frac{q(q-1)}{2}$ 2-class SVMs in "pair-wise" top down tree structured approach proposed in [17]. Here, q represents the number of classes in the dataset.

From the experiment it can be said that for average accuracy of SVM-RBF- 86.90% significantly outperforms SVM-Polynomial-82.95%, SVM-Sigmoid- 84.81%, Neural Network- 79.56%, *k*-NN- 82.45%.

5. Conclusion and Future Research

As describe in earlier sections diagnosis of ovarian mass through ultrasound images is useful and efficient way of detection. And we presented some techniques which are widely used for the detection. Efficient feature extraction is also a key part if the diagnosis as from the optimum feature it can conclude that weather cysts is malignant of benign, and from previously presented methods combination of GLCM and GLRLM can perform better and can give better result. Classification is also important because algorithms which we stated above will classify those extracted feature which will lead to the efficient diagnosis. From above experimental result it can be conclude that SVM, SVM-RBF can be decent choice for the classification.

Though much research in medical diagnosis field is going on and many new techniques are getting introduced. But we can increase accuracy, specificity, true positives and decrease false positives etc. We can find new techniques for optimal feature selection. We can also make computer aided diagnosis system for better diagnosis which will help doctors and patients as well. And instead on manual ROI can also find some solution for automated segmentation for ultrasound images of ovary.

References

[1] Golam Morshed Maruf, Sabrina Rashid *"A novel self-adaptive algorithm for cancer classification based on*

Licensed Under Creative Commons Attribution CC BY

feature reduction of SELDI-TOF data using wavelet decomposition" ICCIT 2011.

[2] Usha B S, Sandya S "*Measurement of ovarian size and parameters"* INDICON 2013

[3] Mohamed, S.S. and SalamaM.M."*Computer Aided diagnosis for Prostate cancer using Support Vector Machine*" Publication: Proc., medical imaging conference 2005, California, SPIE Vol. 5744, pp. 899-907.

[4] U Rajendra Acharya, VinithaSree S, Luca Saba, Filippo Molinari, Stefano Guerriero, Jasjit S Suri " *Ovarian Tumor Characterization and Classification: A class of GyneScan™ Systems"*International Conference of the IEEE EMBS San Diego, California USA, , 2012

[5] M. K. Biswas, T. Ghose, S. Guha, and P. K. Biswas, "*Fractal dimension estimation for texture images: A parallel approach,*" *Pattern Recogn Letters*, vol. 19, pp. 309-313, 1998.

[6] C. Nikias and A. Petropulu," *Higher-Order Spectral Analysis"*. Englewood Cliffs, NJ: Prentice-Hall, 1997.

[7] Abu Sayeed Md. Sohail, Prabir Bhattacharya, Sudhir P.

[8] Mudur, and Srinivasan Krishnamurthy "*retrieval and classification of ultrasound images of ovarian cysts Combining texture features and histogram moments"* ISBI 2010

[9] R.M. Haralick, K. Shanmugan, and I.H. Dinstein, "*Textural Features forImage Classification,"*IEEE Trans. Systems, Man and Cybernetics, vol. 3, no. 6, pp. 610–621, May 1973

[10] Abu Sayeed Md. Sohail, Prabir Bhattacharya, Sudhir P. Mudur, and Srinivasan Krishnamurthy *"local relative GLRLM-based texture feature extraction for Classifying ultrasound medical images"* IEEE CCECE 2011M.M.

[11] Galloway, "*Texture Analysis Using Gray Level Run Lengths,"* Computer Graphics Image Processing, vol. 4, pp. 172–179, June 1975

[12] X. Pan, and Q.Q. Ruan, "*Palmprint Recognition Using Gabor- Based Local Invariant Features*," Neurocomputing, vol. 72, no. 7-9, pp. 2040 - 2045, 2009.

[13] Abu Sayeed Md. Sohail, Prabir Bhattacharya, Sudhir P. Mudur, and Srinivasan Krishnamurthy "*selection of optimal texture descriptors for retrieving Ultrasound medical images"*ISBI 2011

[14] H. Liu, J. Li, and L. Wong, "*A Comparative Study on Feature Selection and Classification Methods Using Gene Expression Profiles and Proteomic Patterns*," *Genome Informatics*, vol. 13, pp. 51-60, 2002.

[15] D. T. Larose," *Decision Trees. In: Discovering Knowledge in Data: An introduction to data mining."* New Jersey, USA: Wiley Interscience,pp. 108-126, 2004

[16] V. N. Vapnik, "*Statistical learning theory.*" John Wiley andSons, New York; 1998.

[17] V.N. Vapnik, "*The Nature of Statistical Learning Theory*", 2nd edn, Springer-Verlag, New York, 2000

[18] J. Platt, N. Cristianini, and J. Shawe-Taylor, "Large margin dags formulticlass classification," *Advances in Neural Information ProcessingSystems,* vol. 12, pp. 547–553, 20

Author Profile

Hemita Pathak pursuing Masters of Engineering from Maharashtra Institute of Technology, Pune. Studied Bachelors of Engineering from VNSGU, Surat. Currently working on Machine Learning and Bioinformatics domain.

Prof. Vrushali Kulkarni working as Assoiciate Professor in Maharashtra Institute of Technology, Pune. She has 21 years of teaching experience and 15 national/international journal/conference publications. Her research interest is in area of Data Mining and Machine Learning.

Prof. Sarika Bobde working as Assistant Professor in Maharashtra Institute of Technology,Pune. She has 15 years of teaching experince. Her area of interest is Data structures and Database Managemnt Systems.

Licensed Under Creative Commons Attribution CC BY