DOI: 10.21917/ijivp.2016.0179

CLASSIFICATION OF CERVICAL CANCER CELLS IN PAP SMEAR SCREENING TEST

S. Athinarayanan[1] and M.V. Srinath[2]

[1]*Department of Centre for Information Technology and Engineering, Manonmaniam Sundaranar University, India*
E-mail: aathithee@gmail.com

[2]*Department of Computer Application, Sengamala Thayaar Educational Trust Women's College, India*
E-mail: drmvsri@gmail.com

Abstract

Cervical cancer is second topmost cancers among women but also, it was a curable one. Regular smear test can discover the sign of precancerous cell and treated the patient according to the result. However sometimes the detection errors can be occurred by smear thickness, cell overlapping or by un-wanted particles in the smear and cytotechnologists faulty diagnosis. Therefore the reason automatic cancer detection was developed. This was help to increase cancer cell mindfulness, diagnosis accuracy with low cost. This detection process consists of some techniques of the image preprocessing that is segmentation and effective texture feature extraction with SVM classification. Then the Final Classification Results of this proposed technique was compared to the previous classification techniques of KNN and ANN and the result would be very useful to cytotechnologists for their further analysis.

Keywords:

Cancer, Cervical Cancer, Classification

1. INTRODUCTION

Cervical cancer is communal cancer among woman in universal but at the same period it is one of the furthermost preventable and treatable cancer. Mostly cervical cancer starts with pre-cancerous changes and develops very slowly. So, up to 92% of this cancers may be prevented because cell changes are detected and treated early by using Smear Test [1]. Typically the Smear Test based on two types. That is Conventional and Thin Prep. Both of this Two tests brushing tool (Cytobrush and Broomlike device) [2]. But almost developed countries, they use only thinprep method test. Then the color of a cell indicates its age which is not a significant information for cancer identification [3], therefore all sample images of smear will be converted to grey scale images for proper evaluation. Cervical cancer is usually affected by Squamous and Glandulas cells but mostly it was affected by the first type and is called Squamous cell carcinoma [4].

A lot of research papers attentive on segmentation, classification, or both have been published over past 35 years. Though, the majority of them is trying to solve a very exact problems while working with very limited datasets. Segmentation typically focus on localisation of nucleus and does not converse with overlapping cytoplasm. Proposed classification methods usually deal with already separated cervix cells, and it doesn't consider the overlap [5]. Segmentation The greatest popular and communal selections for the segmentation task and it was based on automatic thresholding, morphological and contours process. Bam and Lovell segmented the nucleus using an active contour model that was assessed by using dynamic programming to novelty the edging with the minimum cost inside a limited space around the darkest point in the image [6].

Marina detected the locations of nuclei centroids detected the locations of nuclei centroids in smear images by using the image gradient, eliminated the centroids that were too close to each other and used a support vector machine (SVM) classifier for the final selection of points using color values in square neighbor hoods. In [7], However, the state of squamous cell dysplasia can be described more specifically, see Fig.1, which has been considered in [8] and [9] cervical cells are categorized into normal, LSIL and HSIL classes but they didn't notable types of normal cells. Then this LSIL and HSIL Classes are distinguished as abnormal cell that is (Mild, Moderate, Severe and CIS types of the dysplasia).

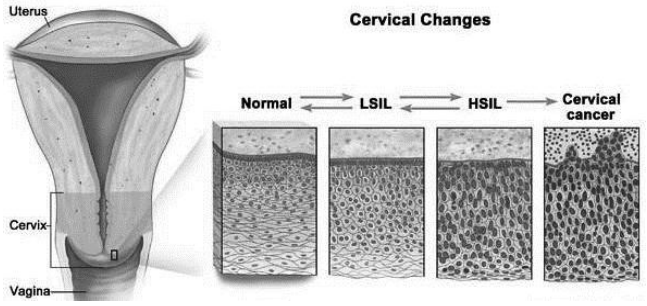


Fig.1. Squamous Cell Dysplasia

There are two or three examinations have been finished by powers to understand if a cell in Pap smear pictures is hazardous or not. They were done in perspective of two or three issues, for occasion, nonattendance of a specialist cytotechnologist, a huge number slide for reliably, lacking lab work environments, techniques persuades the chance to be repetitive and amazingly fragile to human goofs. Modified cervical exposure have been well known by bosses all together with manage the issues [10,11,12]. Recollecting the completed goal to total in the midst of destructive and non-hazardous cells, there are couple of techniques have been proposed by scientists. Jere [13] has shown Outline Affirmation on 2D Cervical Cytological Propelled Pictures. In this paper tries to research outlines of cervix cells considering its morphological segments, in term of size, shape, and shading. Empowers of this paper, the parameters that will be utilized for their computerized attestation structure were cytoplasm to focus degree, shading force, and wavelet attributes to empower changed obvious affirmation of assortment from the standard in the cervix cells. Calculation classifiers utilized by her study are NB and SVM classifier. A brief time allotment later, Tana [14] proposed another strategy which is game-plan cervical improvement utilizing Fourier change. The segments that used to detach in the midst of standard and the unusual cells are figured

in context of the mean, change and entropy picked up from the rehash pieces along the circle clear drew in at the purpose of joining of the compass and the rehash areas along the developed line having an edge θ. Figuring the segments were finished by utilizing some numerical recipes and as a part of along these lines can give a result whether the Pap smear pictures are unsafe or not.

Furthermore the Disease identification methods compress the distinctive order strategies. At first the subtle elements of Straight Classifiers, routinely utilized direct classifiers for cervical malignancy discovery and grouping are straight discriminate investigation [15] and logistic relapse [16]. The significant outline of Straight Discriminative Investigation is to find the direct blend of the elements which best separate two or more classes of the information. Counterfeit Neural Systems, This arrangement systems are the collection of numerical shape that mirror the properties of organic sensory system and the rule of versatile natural learning [17]. In the glade of cervical malignancy discovery and characterization, three number of sorts of manufactured neural systems ANN are routinely utilized that is Back-Proliferation neural system, self-arranging map and various leveled ANN [18] [19]. At that point the thought behind Bayesian neural system is to cast the charge of preparing a system as a predicament of surmising, which is understood by utilizing hypothesis of Bayes' [20]. At that point the Bayesian neural system is further finest and intense than customary neural systems, essentially when the preparation information set is little. The SVM Classifier preparing problem [21] take into consideration discovering misclassification of boisterous. In [22] [23] [24], SVM [25] was connected to arrange the typical or irregular of the classes effectively.

2. PROPOSED METHOD

The figure of the proposed system diagram is shown in Fig.2. In first the Pap smear test images are converted from color into grayscale images. Then Adaptive Path Smooth Filter has been applied in order to remove a small noise while preserving edge sharpness. The process of this Filter is replacing a pixel value with a weighted sum of the native pixels on the minimum-cost path, this algorithm reduces noise without blurring edges. It determines the weighted sum by the lowest-cost pixel at the boundary of a local neighborhood back to the original pixel.

2.1 SEGMENTATION

Segmentation process based on two steps, the first step is segmentation of cell that is background removal and second one segmentation of cytoplasm and the nucleus. This Segmentation Process is based on Modified Otsu Thresholding Algorithm. The steps of the algorithm is given below.

1. Finding histogram and probabilities of all intensity level
2. Fixed up initial $\omega_i(0)$ and $\mu_i(0)$
3. Footstep through all probable thresholds $t = 1\ldots$ maximum intensity
4. Modernize ω_i and μ_i.
5. Calculate $\sigma_b^2(t)$.
6. Preferred threshold related to the maximum $\sigma_b^2(t)$.

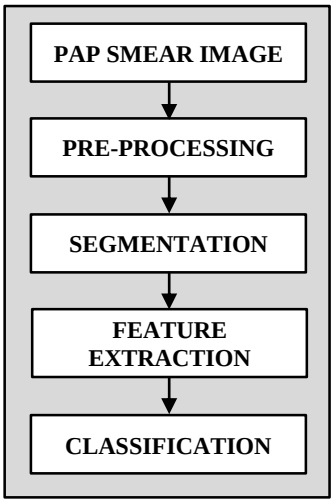


Fig.2. Proposed System

2.2 FEATURE EXTRACTION

2.2.1 Texture Feature Extraction based on GLCM:

Gray-level-based features: features based on the change between the gray-level in the pixel and a statistical value descriptive of its surrounds. It contains the second-order statistical information of neighboring pixels of an image. It is estimated of a probability density function (PDF) of gray level duos in an image [26]. It can be expressed in the following Eq.(1),

$$P_\mu(i,j) \qquad (i,j = 0,1,2,...N-1) \tag{1}$$

where, i, j indicate the gray level of two pixels, N is the gray image dimensions, μ is the position relation of two pixels. Different values of μ decides the distance and direction of two pixels. Normally Distance (D) is 1, 2 and Direction(θ) is 0^0, 45^0, 90^0, 135^0 are used for calculation.

Texture features can be extracted from gray level images using GLCM Matrix. In our proposed method, five features namely, energy, correlation, entropy, contrast and homogeneity are experiments. These features are extracted from the segmented Pap Smear images and analyzed using different directions and spaces.

Energy states the replication of pixel pairs in an image as given in Eq.(2),

$$k_1 = \sum_{i=0}^{N-1}\sum_{j=0}^{k-1} P_\mu^2(i,j). \tag{2}$$

Local variations present in the image is measured by Contrast. If the contrast value is high means the image has large variations as given in Eq.(3).

$$k_2 = \sum_{i=0}^{N-1} t^2 \left\{ \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} P_\mu(i,j) \right\}. \tag{3}$$

Correlation is a measure linear dependency of gray values in a matrices. It is a 2D histogram in which distinct pixel pairs are allotted to each other on the base of a specific, predefined displacement vector as given in Eq.(4).

$$k_3 = \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} \frac{(i,j)\,p(i,j) - \mu_1\mu_2}{\sigma_1^2\sigma_2^2}. \tag{4}$$

where, μ_1, μ_2, σ_1, σ_2 are mean and standard deviation values accrued in the *x* and *y* directions individually.

Entropy is a measure of non-uniformity in an image based on the likelihood of Co-occurrence values, it also indicates the complication of the image as given in Eq.(5).

$$k_4 = -\sum_{i=0}^{k-1}\sum_{j=0}^{k-1} P_\mu(i,j)\log\left(P_\mu(i,j)\right). \quad (5)$$

Homogeneity is inversely proportional to contrast at constant energy, whereas it is inversely proportional to energy as given in Eq.(6).

$$k_5 = \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} \frac{P_\mu(i,j)}{1+(1-j)^2}, i \neq j\ . \quad (6)$$

2.3 CLASSIFICATION

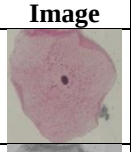
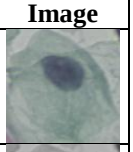
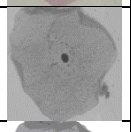
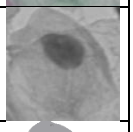
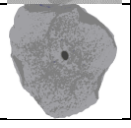
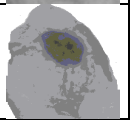
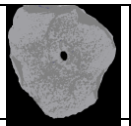
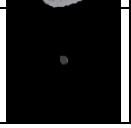
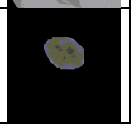
The Classification for Pap smear Analysis aims at classifying the Pap smear cells whether it is affected or not by using the extracted features of gray level co-occurrence matrix technique. Then in this paper using Support Vector Machine, KNN and ANN Classification techniques. Then the final classification results, SVM is better Classification method compared to other methods of KNN and ANN. These Classification performance are analyzed by using in terms of sensitivity, specificity and accuracy.

3. EXPERIMENTAL RESULTS & DISCUSSI ON

3.1 EXPERIMENTAL IMAGE DATA SET AND RESULTS

The cytology images are attained through a controlling microscope by the skilled Cyto technicians. All images were captured with a resolution of 0.201 μm/pixel from the public database of cervical cancer, Herlev University Hospital, Denmark. In this Experimental Result Section as per the three process of this paper which was referred in our previous section (2.1-2.3), dataset images 40 Normal and 60 Abnormal Images could be processed. The Sample of the two images (one normal and one abnormal) Result was given in Table.1.

Table.1. Experimental Results

Image Details	Normal Image	Abnormal Image
Input Image		
Gray Image		
Cell Segmented Image		
Cytoplasm Segmented Image		
Nuclei Segmented Image		

3.2 PERFORMANCE ANALYSIS

Sensitivity and specificity measure the statistical performance of a binary classification test, also known in statistics or classification functions, sensitivity measures the proportion of actual positives which are correctly identified. Specificity measures the proportion of negatives which are correctly identified.

TP = True Positive is correctly classified the cell as cancer cell.

FN = False Negative is incorrectly classified the cell as no cancer.

FP = False Positive is incorrectly classified the cell as cancer.

TN = True Negative is correctly classified the cell as no cancer.

$$Sensitivity = TP/(TP + FN) \quad (7)$$

$$Sensitivity = TN/(TN + FP) \quad (8)$$

$$Accuracy = (TN + TP)/(TN + TP + FN + FP) \quad (9)$$

3.3 COMPARISON ANALYSIS

We have compared our proposed cervical cancer classification system, against the neighbor technique (KNN) and neural network (ANN) techniques. The performance analysis has been made by plotting the graphs of estimation metrics such as sensitivity, specificity and the accuracy are shown in Table.2.

Table.2. Experimental Result with proposed system

Evaluation metrics		Texture Features with SVM(Proposed)	Texture Features with KNN	Texture Features with ANN
Input Image Data Set	TP	72	59	56
	TN	14	11	9
	FP	9	18	22
	FN	5	12	13
	Sensitivity	94	83	81
	Specificity	60	38	29
	Accuracy	86	70	65

By analyzing the performance of the classification a graph will be plotted according to the value of the Table.2. Then based on the results plotted in a graph, the proposed SVM Classification is a best method, when it compare the other methods KNN and ANN. The Resultant Graph was shown in Fig.3, Fig.4 and Fig.5.

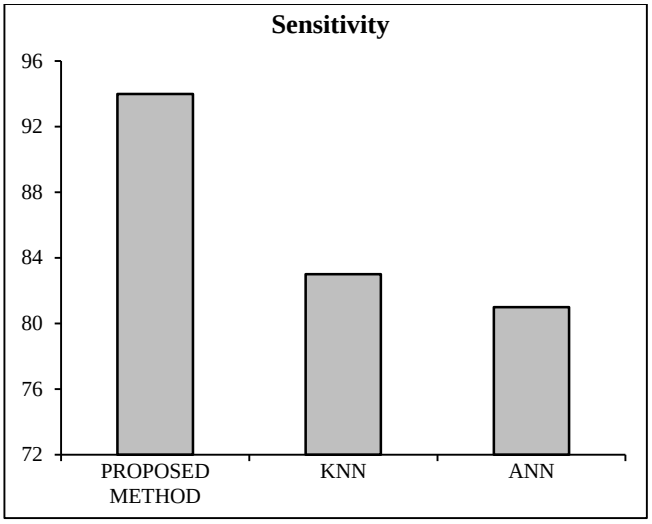


Fig.3. Comparison result analysis by using Sensitivity parameter for SVM, KNN and ANN

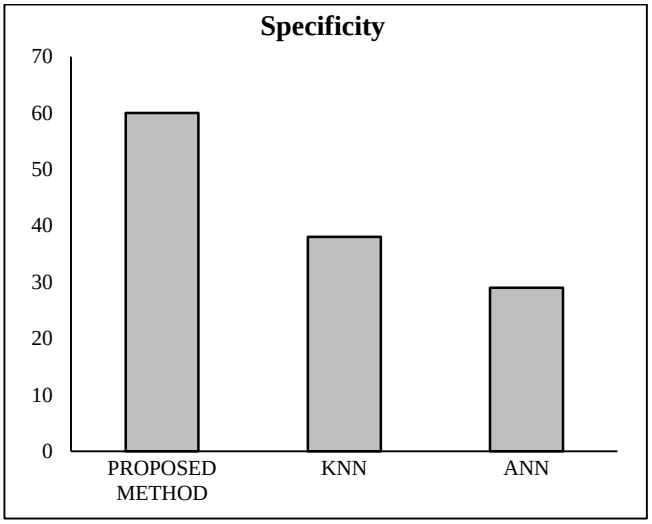


Fig.4. Comparison result analysis by using Specificity parameter for SVM, KNN and ANN

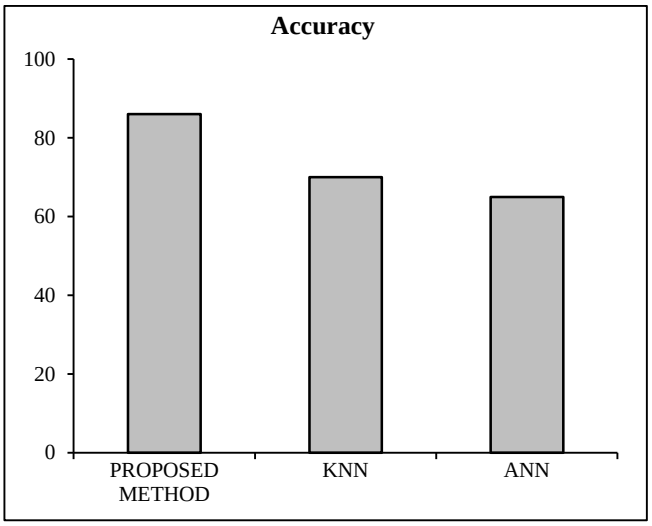


Fig.5. Comparison result analysis by using Accuracy parameter for SVM, KNN and ANN

4. CONCLUSION

At the ending session of this paper, we developed an automatic cervical cancer detection system for finding whether the given pap smear image contains the cell is normal or abnormal. Then the decision making system was premeditated with the Texture features and Support Vector Machine. The proposed system involves feature extraction and classification. The advantage of the system is to support the physician to make the final decision without hesitancy. Conferring to the experimental results, the proposed method is effective for the classification of Pap Smear Cell image into normal and abnormal. For comparative analysis, our proposed approach is compared with other classification such as KNN and ANN. The accuracy level (86%) for our proposed method proved that the proposed algorithm graph is good at detecting the cancer in the experimental images.

REFERENCES

[1] D. Cobzas, N. Birkbeck, M. Schmidt, M. Jagersand and A. Murtha, "3D Variational Brain Tumor Segmentation using a High Dimensional Feature Set", *Proceedings of 11th International Conference on Computer Vision*, pp. 1-8, 2007.

[2] M.E. Plissiti, C. Nikou and A. Charchanti, "Automated Detection of Cell Nuclei in Pap Smear Images using Morphological Reconstruction and Clustering", *IEEE Transactions on Information Technology in Biomedicine*, Vol. 15, No. 2, pp. 233-241, 2011.

[3] American Cancer Society, http://www.cancer.org/cancer/cervicalcancer/index.

[4] Haematoxylin Eosin (H&E) staining, http://protocolsonline.com/histology/dyes-and-stains/haematoxylin-eosin-he-staining/

[5] W. David Stinson, "Pathology of an Abnormal Pap Smear", http://wdavidstinsonmd.com/Pap%20test.htm

[6] Yu Pend, et al., "Detection of Nuclei Clusters from Cervical Cancer Microscopic Imagery Using C4.5", *Proceedings of 2nd International Conference on Computer Engineering and Technology*, Vol. 3, pp. 593-597, 2010.

[7] N. Mustafa, N.A. Mat-Isa, U.K. Ngah, M.Y. Mashor and K.Z. Zamli, "Linear Contrast Enhancement Processing on Preselected Cervical Cell of Pap Smear Images", *Technical Journal School of Electrical & Electronic Engineering*, Vol. 10, pp. 30-34, 2004.

[8] A. Jayachandran and R. Dhanasekaran, "Brain Tumor Detection using Fuzzy Support Vector Machine Classification based on a Texton Co-occurrence Matrix", *Journal of Imaging Science and Technology*, Vol. 57, No. 1, pp. 10507-1-10507-7, 2013.

[9] S. Mukhopadhyay and B. Chanda, "Multiscale Morphological Segmentation of Gray Scale Images", *IEEE Transactions on Image Processing*, Vol. 12, No. 5, pp. 533-549, 2003.

[10] J. Suryatenggara, B.K. Ane, M. Pandjaitan and W. Steinberg, "Pattern Recognition on 2D Cervical Cytological Digital Images for Early Detection of Cervix Cancer",

World Congress on Nature and Biologically Inspired Computing, pp. 257-262, 2009.

[11] J.S.-J. Lee, J.-N. Hwang, D.T. Davis and A.C. Nelson, "Integration of Neural Networks and Decision Tree Classifiers for Automated Cytology Screening", *Proceedings of International Joint Conference on Neural Networks*, Vol. 1, pp. 257-262, 1991.

[12] N. Lassouaoui, L. Hamami and N. Nouali, "Morphological Description of Cervical Cell Images for the Pathological Recognition", *International Journal of Medical, Health, Biomedical, Bioengineering and Pharmaceutical Engineering*, Vol. 1, No. 5, pp. 313-316, 2007.

[13] D.C. Wilbur and Bibbo, "*Automation in Cervical Cytology*", Chapter 34, Comprehensive Cytopathology, 2008.

[14] Y. Srinivasan, S. Yang, B. Nutter, S. Mitra, B. Philips and R. Long, "Challenges in Automated Detection of Cervical Intraepithelial Neoplasia", *Proceedings of SPIE, Medical Imaging 2007: Computer-Aided Diagnosis*, Vol. 6514, 2007.

[15] F.H. White and K. Gohari, Alterations in the Volume of the Intercellular Space between Epithelial-cells of the Hamster Cheek-Pouch: Quantitative Studies of Normal and Carcinogen-treated Tissues", Journal of Oral Pathology & Medicine, Vol. 13, No. 3, pp. 244-54, 1984.

[16] F.B. Sorensen, P. Bichel and A. Jakobsen, "Stereological Estimates of Nuclear Volume in Squamous-cell Carcinoma of the Uterine Cervix and its Precursors", *Virchows Archive A: Pathological Anatomy and Histopathology*, Vol. 418, No. 3, pp. 225-233, 1991.

[17] Rahmadwati, Golshah Naghdy, Montse Ross, Catherine Todd and E. Norachmawati, "Classification Cervical Cancer using Histology Images", *Proceedings of Second International Conference on Computer Engineering and Applications*, Vol. 1, pp. 515-519, 2010.

[18] Thanatip Chankong, Nipon Theera-Umpon and Sansanee Auephanwiriyakul, "Cervical Cell Classification using Fourier Transform", *Proceedings of 13th International Conference on Biomedical Engineering*, Vol. 23, pp. 476-480, 2009.

[19] Yanxi Liu, Tong Zhao and Jiayong Zhang, "Learning Multispectral Texture Features for Cervical Cancer Detection", *Proceedings of IEEE International Symposium on Biomedical Imaging: Macro to Nano*, pp. 169-172, 2002.

[20] Eric Njoroge, Stephen R. Alty, Mahbub R. Ghani and Maha Akatib, "Classification of Cervical Cancer Cells using FTIR Data", *Proceedings of the 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 5338-5341, 2006.

[21] Xiangjun Shi, H.D. Cheng, Liming Hu, Wen Ju and Jiaweitian, "Detection and Classification of Masses in Cervical Ultrasound Images", *Digital Signal Processing*, Vol. 20, No. 3, pp. 824-836, 2010.

[22] Peter A. Lachenbruch, "*Discriminant Analysis*", Haffner Press, 1975.

[23] 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 Mammogram", *Pattern Recognition*, Vol. 39, No. 4, pp. 646-668, 2006.

[24] J.A. Noble and D. Boukerroui, "Ultrasound Image Segmentation: a Survey", *IEEE Transactions on Medical Imaging*, Vol. 25, No. 8, pp. 987-1010, 2006.

[25] Yu Len Hunang, Kao LunWang and Dar Ren Chen, "Diagnosis of Cervical Tumors with Ultrasonic Texture Analysis using Support Vector Machines", *Neural Computing and Applications*, Vol. 15, No. 2, pp. 164-169, 2006.

[26] Christopher Bishop, "*Pattern Recognition and Machine Learning*", 1st Edition, Springer, 2006.

[27] Jae H. Song, Santosh S. Venkatesh, Emily F. Contant, Ted W. Cary, Peter H. Arger, and Chandra M. Sehgal, "Artificial Neural Network to Aid Differentiation of Malignant and Benign Cervical Masses by Ultrasound Imaging", Proceedings of SPIE, Medical Imaging 2005: Ultrasonic Imaging and Signal Processing, Vol. 5750, pp. 148-152, 2005.

[28] J.G. Daugman, "Complete Discrete 2-D Gabor Transforms by Neural Networks for Image Analysis and compression", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. 36, No. 7, pp. 1169-1179, 1998.