Artificial Intelligence in Medicine 50 (2010) 33–41

Contents lists available at ScienceDirect

Artificial Intelligence in Medicine

journal homepage: www.elsevier.com/locate/aiim

Fusion of fuzzy statistical distributions for classification of thyroid ultrasound patterns

Dimitris K. Iakovidis [a,*], Eystratios G. Keramidas [b], Dimitris Maroulis [b]

[a] Department of Informatics and Computer Technology, Technological Educational Institute of Lamia, GR 35100 Lamia, Greece
[b] Department of Informatics and Telecommunications, University of Athens, Panepistimiopolis, GR-15784 Athens, Greece

ARTICLE INFO

Article history:
Received 24 September 2008
Received in revised form 21 September 2009
Accepted 29 March 2010

Keywords:
Fuzzy feature extraction
Support vector classification
Ultrasound imaging
Thyroid nodules

ABSTRACT

Objective: This paper proposes a novel approach for thyroid ultrasound pattern representation. Considering that texture and echogenicity are correlated with thyroid malignancy, the proposed approach encodes these sonographic features via a noise-resistant representation. This representation is suitable for the discrimination of nodules of high malignancy risk from normal thyroid parenchyma.
Materials and methods: The material used in this study includes a total of 250 thyroid ultrasound patterns obtained from 75 patients in Greece. The patterns are represented by fused vectors of fuzzy features. Ultrasound texture is represented by fuzzy local binary patterns, whereas echogenicity is represented by fuzzy intensity histograms. The encoded thyroid ultrasound patterns are discriminated by support vector classifiers.
Results: The proposed approach was comprehensively evaluated using receiver operating characteristics (ROCs). The results show that the proposed fusion scheme outperforms previous thyroid ultrasound pattern representation methods proposed in the literature. The best classification accuracy was obtained with a polynomial kernel support vector machine, and reached 97.5% as estimated by the area under the ROC curve.
Conclusions: The fusion of fuzzy local binary patterns and fuzzy grey-level histogram features is more effective than the state of the art approaches for the representation of thyroid ultrasound patterns and can be effectively utilized for the detection of nodules of high malignancy risk in the context of an intelligent medical system.

1. Introduction

Ultrasound imaging presents a valuable modality that has come to play an increasingly important role in the diagnostic evaluation of soft tissues such as eyes, heart, thyroid, and other internal organs. Recent advances in ultrasound technology lead to high frequency transducers (5–15 MHz) which provide both deep ultrasound penetration and highdefinition images, achieving spatial resolution between 0.7 and 1.0 mm. It also combines a set of unique virtues such as non-invasive, low cost, quick, convenient, and with no side effects real time imaging.

The thyroid gland is located in the anteroinferior part of the neck covered by a thin layer of muscle and dermal tissue. Due to the superficial location of the thyroid gland, ultrasound technology has become the most widely employed imaging method for the diagnosis and follow-up of thyroid disorders such as nodules, tumors and cysts [1–3]. Many thyroid diseases can present clinically with one or more thyroid nodules. Although in some cases such nodules can be detected through palpation, the majority of them are clinically occult (smaller than 1.5 cm), but can be readily detected by high resolution sonography. Two of the most useful sonographic features recognised by the radiologic community for detection and malignancy risk assessment of thyroid nodules are echogenicity and texture [4].

However, an inherent characteristic of ultrasound imaging is the presence of speckle noise [5]. Speckle is a random, deterministic, interference pattern in an image formed by coherent radiation of a medium containing many sub-resolution scatterers. Speckle has a negative impact on ultrasound imaging, since it tends to reduce the image effective resolution and contrast, inducing a degree of uncertainty which can be consequently propagated to medical diagnosis. Bamber and Daft [6] have showed that lesion detectability decreases approximately by a factor of eight due to the presence of speckle noise. Therefore, a methodology for the representation of thyroid ultrasound patterns should not ignore the noise-originated uncertainty.

Several endeavours have been undertaken to improve interpretation of thyroid ultrasound images through quantitative

* Corresponding author. Tel.: +30 6932 252591; fax: +30 22310 33945.
E-mail address: dimitris.iakovidis@ieee.org (D.K. Iakovidis).

0933-3657/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.artmed.2010.04.004

criteria [7–15]. These approaches take no special consideration of the noise-originated uncertainty. In order to obtain an uncertainty-aware representation of thyroid ultrasound patterns we propose a noise-resistant coding of both texture and echogenicity, based on fusion of a fuzzy distribution of local binary patterns, referred to as fuzzy LBP (FLBP) features, and ultrasound echogenicity represented by the fuzzy grey-level histogram (FGLH).

In this paper the performance of the proposed fusion scheme is thoroughly investigated for the classification of nodular and normal thyroid ultrasound patterns. The classification task has been assigned to SVMs because of their robust performance even in high dimensional feature spaces [16]. The experimental evaluation is based on reliable performance metrics estimated from receiver operating characteristics (ROCs) and involve comparisons with several thyroid ultrasound pattern representation approaches.

The rest of this paper comprises four sections. Section 2 review today's literature in interpretation of ultrasound thyroid images. Section 3 describes the material used in this study and the proposed methodology. Section 4, refers to the experimental evaluation and the classification results obtained. A discussion along with a summary of conclusions is provided in the last section.

2. Related work

There have been various attempts towards less subjective techniques for the evaluation of thyroid ultrasound images. Some of the earliest approaches were based on the use of grey-level histograms (GLHs). In [7] GLHs have been used for efficient supervised and unsupervised classification of normal and abnormal thyroid tissue. In [8] GLHs have been effectively used for the discrimination of tissue with Hasimoto disease from normal thyroid tissue. Morifuji [9] proposed an optimal combination of statistical features extracted from GLH for malignancy determination of thyroid nodules. These features include: the difference of mean, the standard deviation, the skewness, and the krutosis. In another study [10], the 10% percentile of GLH has been utilized for the quantification and classification of echographic findings in the thyroid gland (Table 1).

The aforementioned studies show that GLH carries substantial information for the characterization of thyroid tissue. However, two ultrasound image regions may have the same histogram but still different textures [17], since the GLH does not encode any information related to the spatial distribution of image pixels. On those grounds latter studies have incorporated second or higher order statistical features for texture analysis of ultrasound thyroid images. In [11], both Haralick's co-occurrence features (CM) [18] and Muzzolini's spatial features [19] have been evaluated for the discrimination of chronic inflamed from normal thyroid tissue. Two feature selection schemes were employed; one based on the compactness and the separability, and another based on the classification error. Both schemes resulted in the same optimal set of features.

Table 1
Summary of previous feature extraction approaches for thyroid ultrasound image analysis, ordered by year of publication.

Category of features	Feature extraction approach	Year	Ref
Echogenicity	GLH	1985	[7]
	GLH	1986	[8]
	GLH mean, standard deviation, skewness and krutosis	1989	[9]
	GLH 10%	1989	[10]
Texture–echogenicity	CM–M	2003	[11]
	CM-MGL	2005	[12]
Texture	CM	2006	[13]
	Radon	2007	[14]
	LBP	2007	[15]

In [12], three types of textural features were utilized for malignancy risk assessment of thyroid nodules. Initially a set of features were calculated from GLHs, CMs and run-length matrices (RL) [20], but only the mean grey-level value of GLHs and the sum variance from the co-occurrence matrix (CM-MGL) were selected as most appropriate via exhaustive search of feature combinations based on the minimum classification error as a criterion.

In [13] only co-occurrence matrix features have been utilized for the characterization of thyroid tissue in ultrasound images. Another study [14] proposes the use of Radon Transform features for the discrimination between normal, low malignancy risk and high malignancy risk nodular tissue. This effort has been based on the hypothesis that tissues in thyroid ultrasound images may be differentiated by directionality patterns, which can be encoded in the Radon domain.

A recent study [15] provided promising results for thyroid nodule detection with a two phase scheme based on local binary pattern (LBP) features [21]. In the first phase of this scheme the thyroid lobe boundaries are detected. These boundaries define a region from which LBP features are extracted in the second phase. The LBP feature vectors are then classified by a k-nearest neighbour classifier.

This literature review shows that the previous approaches use GLH and/or various textural descriptors for the representation of thyroid ultrasound patterns; however, none of them takes any special consideration of the noise-originated uncertainty in the ultrasound images. In order to obtain an uncertainty-aware representation of thyroid ultrasound patterns we propose a noise-resistant coding of both texture and echogenicity, based on fusion of fuzzy statistical distributions. Ultrasound texture is represented by a fuzzy distribution of local binary patterns, referred to as fuzzy LBP (FLBP) features, and ultrasound echogenicity is represented by the fuzzy grey-level histogram (FGLH), which is well known for its insensitivity to noise [22]. Some initial experimental studies indicate the suitability of the FLBP features for ultrasound texture representation [23,24], and it's resistance in the presence of additive noise for the representation of natural textures from reference databases [25,26].

3. Materials and methods

3.1. Thyroid ultrasound patterns

The material used in this study is a set of anonymized B-mode thyroid ultrasound images accompanied with ground truth information provided by the Euromedica Medical Center of Athens in Greece with the approval of its ethical committee. All ultrasound examinations were performed with a digital ultrasound system Philips HDI 5000. A broadband curved array transducer with a frequency range of 2.0–5.0 MHz was used. All the ultrasound images were stored in digital imaging and communications in medicine (DICOM) format. The settings of the ultrasound scanner that affect image attributes (e.g. overall time-gain, near-field and far-field gain compensation) were kept constant throughout the entire study. The ultrasound images were digitized at 8-bit grey-levels with an effective resolution of 470×480 pixels.

In total, 75 patients who had ultrasonographic examinations of their thyroid gland were enrolled in this study. Each of the 200 longitudinal ultrasound images obtained from these patients contains one or more hypoechoic nodules classified as Grade 3 or Grade 4 [27], by two out of three expert evaluators (Fig. 1(a and b)).

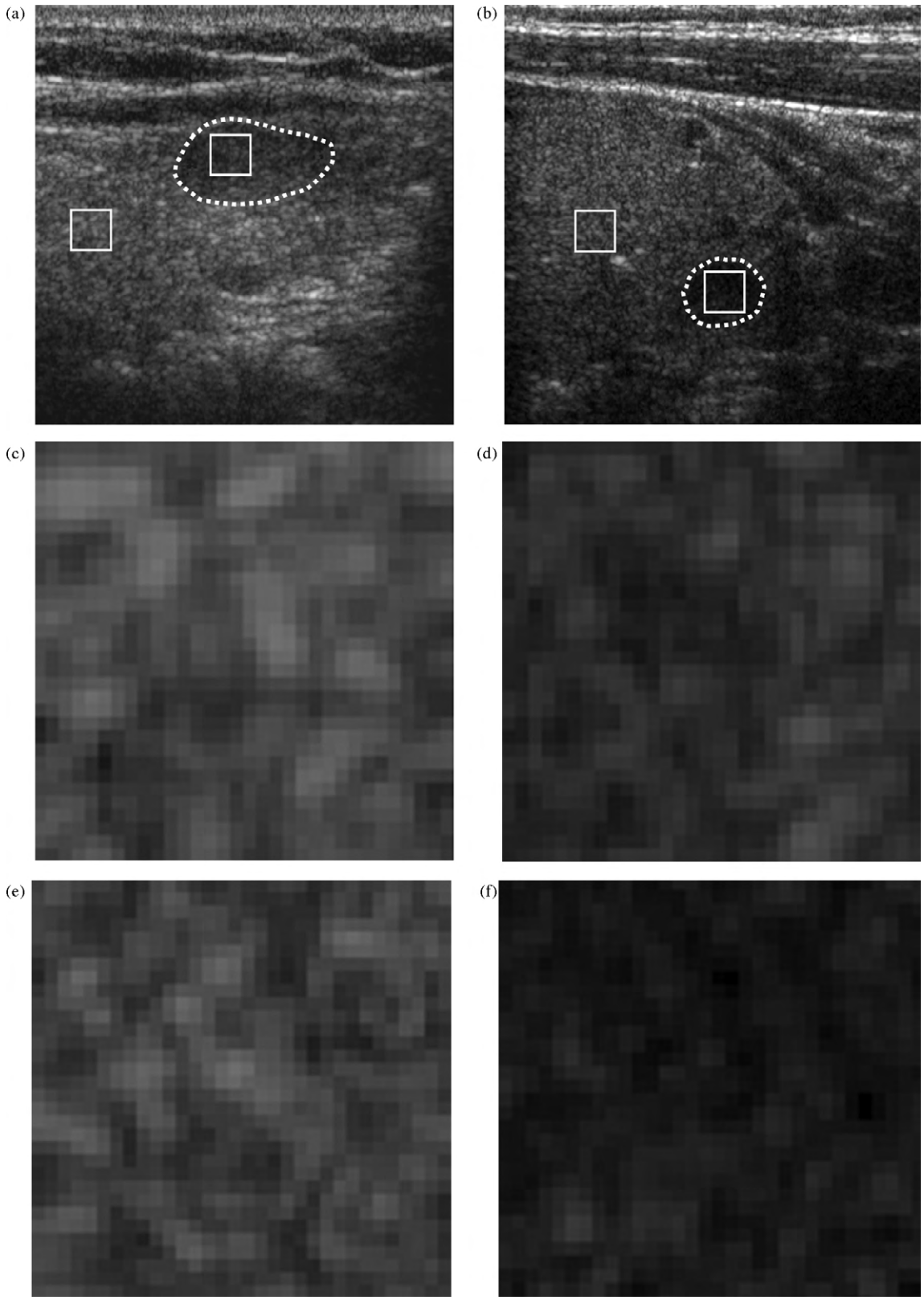


Fig. 1. (a, b) Representative thyroid ultrasound images with one visible nodule delineated and two square blocks selected from normal and nodular thyroid tissue; (c, e) square block sampled from normal thyroid parenchyma; (d, f) square block sampled from nodular tissue.

Our study focuses only on the discrimination of these types of nodules because they are associated with a significantly higher malignancy risk than isoechoic and hyperechoic nodules.

The ground truth boundaries of the nodules (Fig. 1(a and b)) were obtained by following the rule that a pixel belongs to the nodule when it is included in at least two out of the three delineations drawn by the experts [28].

From each image a balanced set of non-overlapping 32 × 32-pixel square block samples from normal (Fig. 1(c and e)) and nodular (Fig. 1(d and f)) tissue of the thyroid gland was selected.

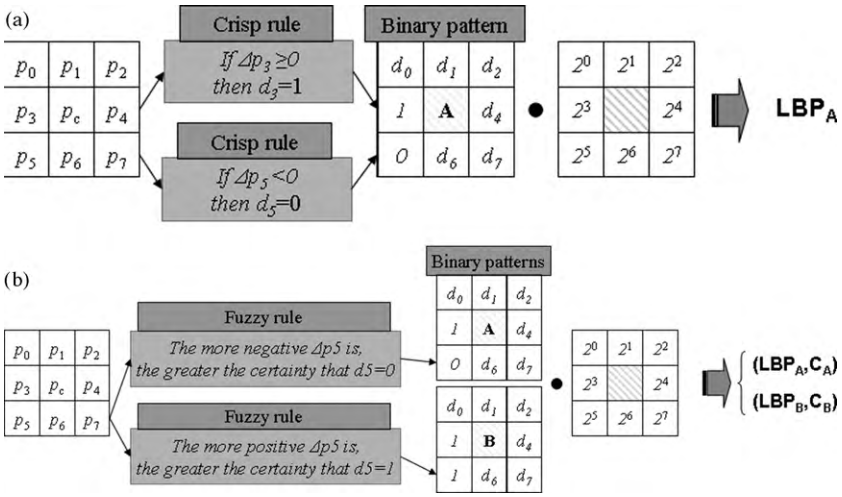


Fig. 2. (a) Example of LBP computation scheme and (b) example of the FLBP computation scheme. For both schemes $\Delta p_3 \geq 0$, $\Delta p_5 < 0$, for every $i \neq 5$, $|\Delta p_i| > F$, and for $i = 5$, $|\Delta p_i| < F$.

The number of blocks sampled from each nodule was determined by the maximum number of blocks that could fit within its boundaries, whereas an equicardinal set of representative samples from the normal thyroid parenchyma was selected by the experts. This sample selection process produced a total of 250 thyroid ultrasound patterns constituting a solid dataset for experimentation.

3.2. Pattern representation

3.2.1. Local binary patterns

The local binary pattern (LBP) [21] is a popular approach to texture representation that has provided excellent results in many studies [29–32]. Derived from a general definition of texture, the LBP is based on the concept of texture primitives, often called texels or textons [17]. This approach provides a theoretically and computationally simple and efficient methodology for texture analysis.

To represent the formations of a textured image, the LBP approach, models 3×3 textons as illustrated in Fig. 2(a). A 3×3 texton consists of a set of nine elements, $P = \{p_c, p_0, p_1, \ldots, p_7\}$, where p_c represents the grey-level value of the central pixel and p_i $(0 \leq i \leq 7)$ represent the grey-level values of the peripheral pixels. Each texton then, can be characterized by a set of binary values d_i $(0 \leq i \leq 7)$, where

$$d_i = \begin{cases} 1 & \text{if } \Delta p_i \geq 0 \\ 0 & \text{if } \Delta p_i < 0 \end{cases} \quad (1)$$

and $\Delta p_i = p_i - p_c$.

For each 3×3 texton a unique LBP code can be derived by these binary values, as follows:

$$\text{LBP} = \sum_{i=0}^{7} d_i \cdot 2^i \quad (2)$$

Every pixel in an image generates such an LBP code. A single LBP code represents the local microtexture information around a pixel by a single integer code LBP $\in$ [0,255]. Then a histogram is created to collect up the occurrences of different LBP codes from all pixels in the image. This histogram forms the LBP feature vector, which characterizes image texture.

The LBP approach is based on Boolean logic through hard thresholding of peripheral pixels. This characteristic makes the LBP texture representation sensitive to small variations of the pixel intensities such as those usually caused by noise. In the following, the LBP approach will also be referred to as crisp LBP so as to differentiate it from the fuzzy LBP approach.

3.2.2. Fuzzy local binary patterns

In order to enhance the LBP approach so as to cope with the uncertainty introduced by the speckle noise in ultrasound images, fuzzy logic modeling has been considered. Fuzzy logic has certain major advantages over traditional Boolean logic when it comes to real world applications such as texture representation of real images. Introduced by Zadeh [33], fuzzy theory excels in producing exact results from imprecise data. The main difference between the fuzzy and the classic logic is that statements are no longer 0 or 1, but assume any real value between 0 and 1, that allows more human-like interpretation and reasoning.

The incorporation of fuzzy logic into the LBP approach includes the transformation of the input variables to respective fuzzy variables, according to a set of fuzzy rules. We define two linguistic rules to describe the relation between the intensity values of the peripheral pixels p_i and the center pixel p_c of a 3×3 neighbourhood in a more human-like fashion, as follows:

Rule 0: The more negative Δp_i is, the greater the certainty that d_i is 0.

Rule 1: The more positive Δp_i is, the greater the certainty that d_i is 1.

Based on these two rules two membership functions $\mu_0(\,)$ and $\mu_1(\,)$ can be defined. Let function $\mu_0(\,)$ define the degree to which Δp_i is negative, i.e. the degree to which d_i is 0. As a membership

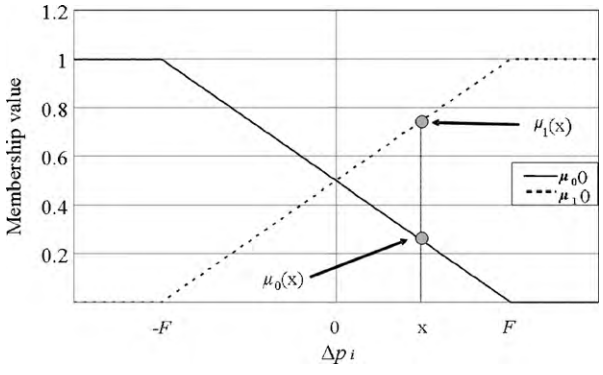


Fig. 3. Membership functions $\mu_0(\,)$ and $\mu_1(\,)$ as a function of Δp_i.

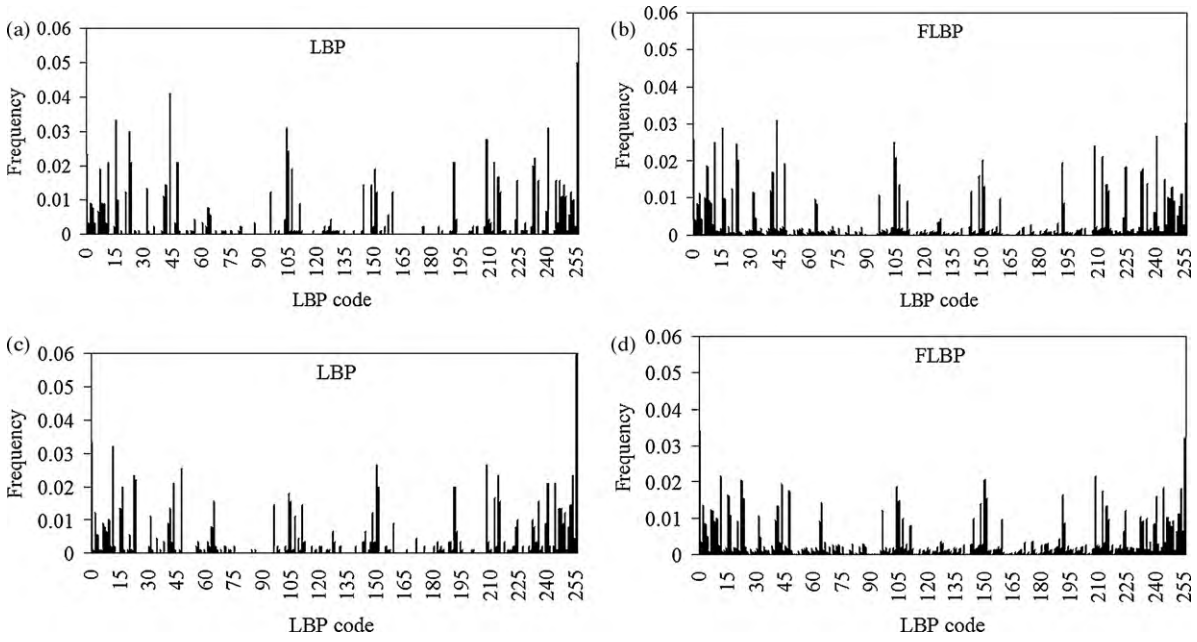


Fig. 4. Histograms obtained by LBP and FLBP approaches applied on images presented on Fig. 1(c and d). (a) LBP histogram from normal thyroid tissue. (b) FLBP histogram from normal thyroid tissue for $F = 5$. (c) LBP histogram from nodular thyroid tissue and (d) FLBP histogram from nodular thyroid tissue for $F = 5$.

function $\mu_0(\ )$ we consider a decreasing function (Fig. 3) defined as follows:

$$\mu_0(i) = \begin{cases} 0 & \text{if} \quad \Delta p_i \geq F \\ \dfrac{F - \Delta p_i}{2 \cdot F} & \text{if} \quad -F < \Delta p_i < F \\ 1 & \text{if} \quad \Delta p_i \leq -F \end{cases} \tag{3}$$

Similarly $\mu_1(\ )$ is an increasing function that defines the degree to which Δp_i is positive, i.e. the degree to which d_i is 1. As membership function $\mu_1(\ )$ we consider the following (Fig. 3):

$$\mu_1(i) = \begin{cases} 1 & \text{if} \quad \Delta p_i \geq F \\ \dfrac{F + \Delta p_i}{2 \cdot F} & \text{if} \quad -F < \Delta p_i < F \\ 0 & \text{if} \quad \Delta p_i \leq -F \end{cases} \tag{4}$$

For both $\mu_0(\ )$ and $\mu_1(\ )$, $F \in [0,255]$ represents a parameter that controls the degree of fuzziness. Contrary to what stated above for the original LBP operator, where a single LBP code characterize a 3×3 texton, in the proposed FLBP approach, a texton can be characterized by more than one LBP codes. Fig. 2(b) presents an example of the FLBP approach, where two LBP codes (A and B) characterize a 3×3 texton. The degree to which each LBP code characterize a texton (C_A, C_B), depends on the values of the membership functions $\mu_0(\ )$ and $\mu_1(\ )$ that correspond to each peripheral pixel. For a 3×3 texton, the contribution C_{LBP} of each LBP code in the FLBP histogram is defined as:

$$C_{LBP} = \prod_{i=0}^{7} \mu_{d_i}(i) \tag{5}$$

where $d_i \in \{0,1\}$ and the LBP code can be obtained from Eq. (2). Each 3×3 texton contributes to more than one bins of the FLBP histogram and always the total contribution of a single neighbourhood is

$$\sum_{LBP=0}^{255} C_{LBP} = 1 \tag{6}$$

Fig. 4(a and c) illustrates two LBP histograms calculated from the normal and nodular thyroid tissue samples illustrated in Fig. 1(c and d). It can be observed that in these histograms 125 and 114 out of 255 bins have zero value, respectively. This results in a small set of significant peaks that can be identified for each histogram. The corresponding FLBP histograms are illustrated in Fig. 4(b and d). These histograms do not have bins with zero values and there are more spikes, though limited in magnitude. This indicates that FLBP histograms are more informative than LBP histograms. Considering that entropy is defined as [34]

$$\Phi = -\sum_{LPB=0}^{255} p_{LBP} \cdot \log(p_{LBP}) \tag{7}$$

where p_{LBP} is the membership degree of the LBP-th pattern, the more diversified the histogram signal, the higher the entropy, and the more the actual information. If all the bins have equal probability, the maximum entropy will be reached. Apparently, the entropy of the FLBP histograms is always greater than the entropy of the crisp LBP histograms.

3.3. Fuzzy grey-level histograms

The echogenicity of B-mode thyroid ultrasound images carries substantial information on the pathology of the examined tissue and can be represented by grey-level histograms (GLH) [35].

Given an ultrasound image region, its normalized histogram represents the empirical probability density function of pixel values and can be defined as $H(g) = h_g/n$ where h_g is the number of pixels with grey-level g, and $N = \sum_{g=0}^{G-1} h_g$.

To cope with the uncertainty introduced in the pixel intensities in the presence of noise, we consider the use of fuzzy histograms as a noise-resistant representation of thyroid ultrasound image regions [36]. The definition of a fuzzy grey-level histogram (FGLH) requires that a membership function $\mu_g(p_i)$ is specified for each grey-level $g \in [0,G]$, where G is the maximum number of grey-levels in the ultrasound image. Each membership function defines the degree of membership of image pixel p_i to histogram bin g. The exact type and shape of membership functions $\mu_g(\ )$ can greatly vary, and in general it depends on the specific problem [37]. The triangular function is commonly used as a membership function

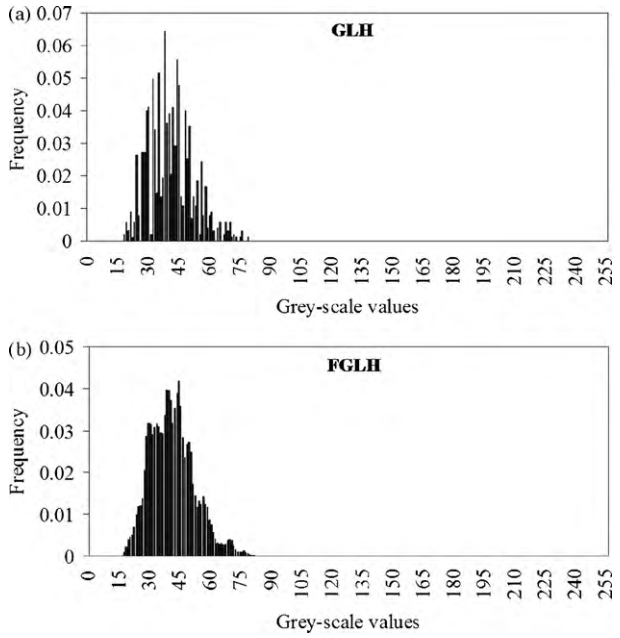


Fig. 5. Histogram and fuzzy histogram ($F = 5$) for image presented on Fig. 1(d).

and can be defined as

$$\mu_g(p_i) = \begin{cases} \dfrac{F - |g_i - g|}{F^2} & |g_i - g| < F \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

where g_i is the grey-level value of pixel p_i and $F \in [0,G]$ is the fuzzification parameter. Other, higher order membership functions such as the quadratic or cubic functions could also be used [37].

Then, the normalized fuzzy histogram of an image region can be defined as (Fig. 5):

$$H(g) = \frac{1}{N}\sum_{i=1}^{N}\mu_g(p_i) \quad (9)$$

4. Results

Comprehensive classification experiments were conducted for the evaluation of proposed thyroid pattern representation approach using linear, 3rd-degree polynomial ($l = 3$) and Gaussian SVMs [38]. The SVM parameters considered were in the range of 2^{-5} to 2^{15} for the cost parameter c and 2^{-15} to 2^{3} for the parameter γ as these have been proposed in [39].

The classification performance was investigated using ROC analysis. The ROC curve presents a convenient way of visualizing a classifiers performance in order to select a suitable operation point, or decision threshold. The area under ROC curve (AUC) offers a reliable single figure measure of the classification performance [40,41]. It represents the probability of correctly distinguishing between nodular and normal thyroid ultrasound patterns. In all the experiments the AUC was estimated by 10-fold cross-validation [42].

The results are organized in two sections. In the first section we evaluate the FLBP features along with various textural features previously proposed for thyroid ultrasound pattern representation. Thereinafter in the second section, we present the results obtained with the proposed feature fusion approach and compare them with the ones obtained with state of the art fusion approaches.

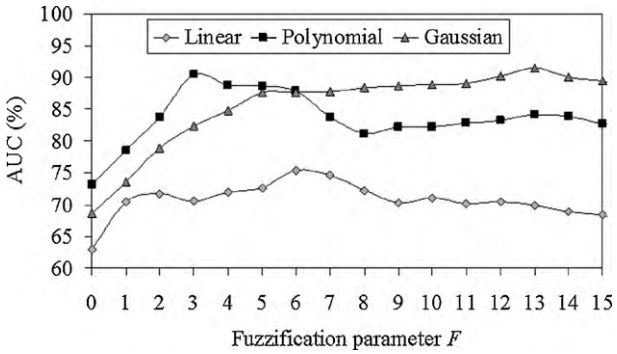


Fig. 6. AUC classification results for LBP ($F = 0$) and FLBP with $F \in [1,15]$.

4.1. Evaluation of textural features

The classification performance of the FLBP features was investigated for $F = 1,2\ldots,15$. The results obtained with various SVM kernel functions are illustrated in Fig. 6. For $F = 0$ the depicted performance corresponds to the one obtained with the crisp LBP features. It can be noticed that the maximum AUC obtained with the crisp LBP approach is 73.1% with polynomial kernel SVM. For each kernel SVM, the AUC obtained with the FLBP features, for all the values of F parameter tested, is higher than the corresponding AUC obtained with LBP features. The maximum AUC of the FLBP features reaches 91.4% for $F = 13$ with the Gaussian kernel SVM. The performance improvement offered by the FLBP approach is clearly depicted by the corresponding ROC curves illustrated in Fig. 7.

Three more thyroid texture representation approaches proposed in the literature were implemented and included in the experimental evaluation presented in this study. The different approaches are:

a. The co-occurrence matrix (CM) feature representation suggested in [11,43].
b. The Radon (R) domain features proposed in [14].
c. Muzzolini's (M) spatial features initially proposed in [19] and applied on thyroid ultrasonography in [43],

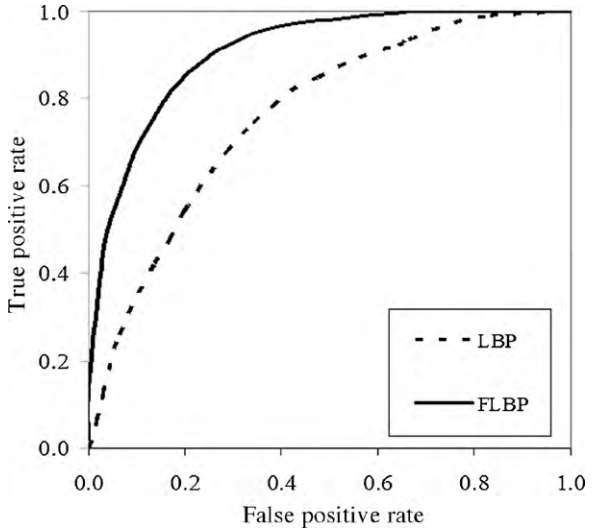
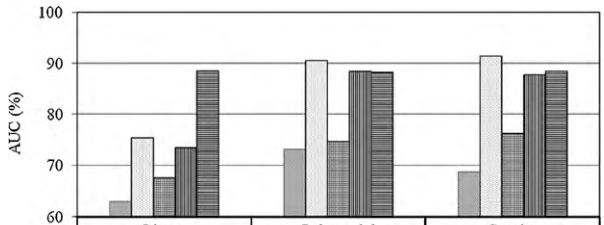
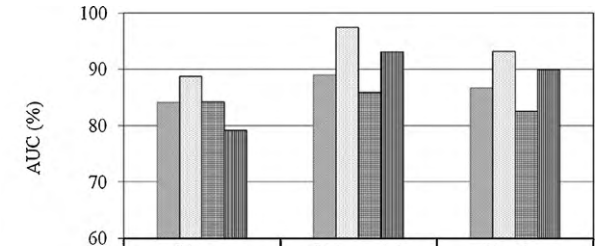


Fig. 7. ROC curves obtained with LBP feature vectors and polynomial kernel SVM (AUC = 73.1%) and with FLBP feature vectors ($F = 13$) and Gaussian kernel SVM (AUC = 91.4%).



	Linear	Polynomial	Gausian
LBP	63.0	73.1	68.7
FLBP	75.4	90.5	91.4
CM	67.6	74.7	76.2
R	73.4	88.4	87.7
M	88.6	88.2	88.4

Fig. 8. AUC accuracies obtained by five thyroid texture representation approaches.



	Linear	Polynomial	Gausian
LBP-GLH	84.1	89.0	86.7
FLBP-FGLH	88.8	97.5	93.2
CM-MGL	84.2	85.9	82.6
CM-M	79.2	93.1	90.0

Fig. 9. AUC accuracies obtained by four feature fusion approaches to thyroid pattern representation.

The best AUCs obtained with these approaches on the available dataset using various SVM classifiers are illustrated in Fig. 8. It can be noticed that the FLBP features (for $F = 13$) provide the best discrimination between the nodular and the normal thyroid ultrasound patterns with the Gaussian SVMs. The discrimination capability of the Radon and of the spatial features is slightly lower.

Comparing the results obtained with the different classification schemes it can be derived that only the spatial features are approximately linearly separable. The lowest AUC was obtained with the crisp LBP approach. This can be attributed to the fact that it does not involve any noise suppressive mechanism such as integral operations or fuzzy modeling in feature computation.

4.2. Evaluation of feature fusion approaches

By introducing fuzzy luminance information into the fuzzy texture representation obtained by the FLBP approach we aim to enhance the discrimination of the nodular from the normal thyroid patterns, since both ultrasound texture and echogenicity provide substantial cues to the clinical assessment of thyroid nodules [2]. To validate this theoretical argumentation we proceeded to extensive experiments investigating the classification performance of the proposed approach in comparison with the following fusion approaches:

a. Fusion of crisp LBP and GLH.
b. The mean value of the local grey-level histogram and the sum variance estimated from the co-occurrence matrix (CM-MGL), as proposed in [12].

The fusion of Muzzolini's spatial features and grey-level co-occurrence matrix features (CM–M) proposed in [11].

The crisp and the fuzzy grey-level histograms considered were of 16, 32, 64, 128 and 256 bins. The histograms with less than 256 bins were obtained by uniform grey-level quantization.

The results of the experimental evaluation are summarized in Fig. 9. The proposed approach outperformed all the other approaches by achieving an AUC of 97.5% with the 128-bin FGLH, with $F = 13$ and the polynomial SVM. The best classification performance achieved with the fusion of the crisp LBP and GLH is significantly lower, reaching only 89.0%. The large difference in these two thyroid pattern representation approaches indicates the significance of fuzzy modeling in the representation of ultrasound patterns.

The fusion methodology CM–M provided the second best AUC (93.1%) with the polynomial SVM. However, this is still much lower than the maximum AUC obtained by the proposed approach. The fusion of the simple features CM-MGL provided the lowest overall classification performance.

A more detailed view of the results for various decision thresholds is presented in Fig. 10. The depicted ROC curves validate that the proposed approach achieves the best tradeoff between the true positive and false positive rates.

For the evaluation of the statistical significance of the pairwise differences between the AUCs obtained by 10-fold cross-validation, a 1-way ANOVA [42] was applied. The results obtained for the case of the polynomial kernel SVMs are summarized in Table 2. Pairs 1–3 show that the proposed feature fusion approach performs significantly better than the other three fusion approaches investigated, with $p\text{-value} \leq 2.5 \times 10^{-2}$ and $F\text{-ratio} \geq 5.9$.

In order to provide a visual cue on the discrimination of the thyroid ultrasound patterns by the proposed feature fusion approach we randomly selected a balanced set of 50 samples from the available dataset and used Sammon's mapping [44] to project the extracted feature vectors into a three-dimensional space. The resulting space is illustrated in Fig. 11, where it can be observed that the vectors belonging to the different classes are indeed separable but in a non-linear way.

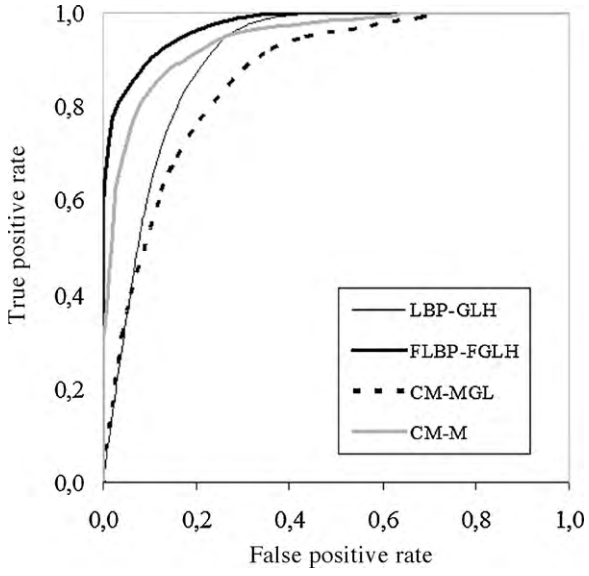


Fig. 10. ROC curves obtained by 3rd-degree polynomial kernel SVMs and four feature fusion approaches to thyroid pattern representation.

Table 2
1-way ANOVA results.

Pair	Approach 1	Approach 2	p-Value	F-ratio	Confidence interval 95%
1	FLBP–FGLH	LBP–GLH	2.5×10^{-2}	5.90	[−0.222, −0.016]
2	FLBP–FGLH	CM-MGL	8.3×10^{-3}	8.77	[0.034, 0.203]
3	FLBP–FGLH	CM–M	2.2×10^{-2}	6.48	[−0.084, −0.008]
4	LBP–GLH	CM–M	1.5×10^{-1}	2.17	[0.177, 0.031]
5	LBP–GLH	CM-MGL	9×10^{-2}	3.19	[−0.012, 0.159]
6	CM-MGL	CM–M	1.8×10^{-2}	5.12	[−0.128, −0.020]

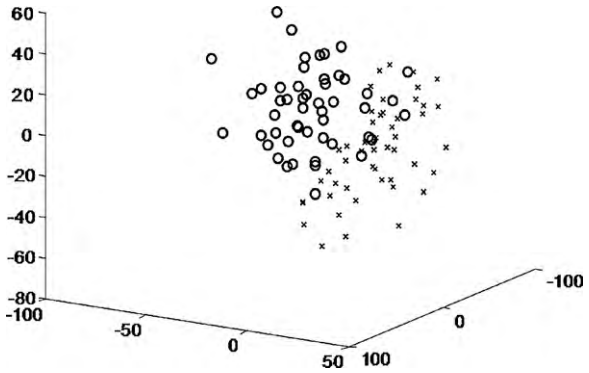


Fig. 11. Visualization of the proposed fused fuzzy feature space. The diagram is plotted with Sammon's mapping of the multidimensional vectors produced by the proposed feature fusion approach into a three-dimensional space.

5. Discussion and conclusions

The present work proposed a novel thyroid ultrasound pattern representation approach for the discrimination of high malignancy risk nodules from normal thyroid parenchyma. Sonographic features that characterize high risk thyroid nodules include hypoechogenicity due to high density of cells and microcalcifications appearing as tiny, punctuate hyperechoic foci. On those grounds we proposed a novel approach that co-evaluates both ultrasound texture and echogenicity. Based on that ultrasound images are inherently characterized by speckle noise we considered fuzzy logic as a means to cope with noise-originated uncertainty. The proposed approach encodes ultrasound texture into FLBP distributions, and echogenicity into FGLH.

Supervised classification analysis based on support vector machines was applied. Comprehensive classification experiments were conducted to investigate the discriminativity between nodular and normal thyroid ultrasound patterns using various feature extraction methodologies and various kernel functions. The ROC and the corresponding AUC have been estimated as performance evaluation criteria more reliable than classification accuracy.

The conclusions derived from this study can be summarized as follows:

- Thyroid nodules of high malignancy risk can be discriminated from normal thyroid parenchyma using textural ultrasound image features.
- The FLBP feature extraction method provides better descriptors of thyroid ultrasound texture than previous methods.
- The classification performance obtained with the FLBP descriptors is significantly enhanced by the proposed fusion of FGLH into the FLBP feature vector. This approach leads to the best discrimination of the high malignancy risk nodules from the normal thyroid parenchyma, as compared to state of the art methods.
- The proposed fusion approach results in high dimensional feature vectors, however their derivation does not involve complex computations. This suggests that the vector classification should be treated by algorithms that are not easily affected by the "curse of dimensionality", such as SVMs.
- Support vector classification experiments testing various kernel functions indicate that the feature space produced by the proposed fusion features approach is non-linearly separable. The polynomial function was indicated as more suitable for the approximation of the decision surface separating the nodular from the normal thyroid ultrasound patterns.

The results of this study indicate that the proposed FLBP–FGLH approach exhibits resistance to ultrasound speckle since it performs better than the fused crisp equivalent LBP–GLH approach. Moreover, preliminary experiments [25] on Gaussian noise-degraded textures have showed that the FLBP approach outperforms the original LBP for various noise levels and that the advantage of the FLBP over the LBP approach becomes more evident for lower signal to noise ratios. These results amplify our expectations about the resistance of the proposed approach to different levels of speckle noise. However the investigation of this issue requires extensive experimentation with images degraded with various levels of artificial speckle noise and could be the topic of a future study.

The features developed in this research promise to improve the management of thyroid cancer by providing a second, more objective, assessment of the thyroid nodules during ultrasound examinations. Future work includes research on the quantification of the malignancy risk of thyroid nodules and integration of the developed methodology into an integrated intelligent medical system.

Acknowledgments

We would like to thank EUROMEDICA S.A., Greece for the provision of the medical images. We would also like to thank N. Dimitropoulos, M.D. and G. Legakis M.D. for their continuous support and advice. This work was supported by the Greek General Secretariat of Research and Technology (25%), the European Social Fund (75%), and private sector, under the framework of Measure 8.3 of E.P. Antagonistikotita – 3rd European Support Framework – PENED 2003 (grant no. 03-ED-662).

References

[1] Bushberg J, Seibert A, Leidholdt E, Boone J. The essential physics of medical imaging, 2nd ed., Philadelphia: Lippincott Williams & Wilkins; 2001.
[2] Rumack CM, Wilson SR, Charboneau JW, Johnson JA. Diagnostic ultrasound, 3rd ed., Netherlands: Mosby; 2004.
[3] Chikui T, Okamura K, Tokumori K, Nakamura S, Shimizu M, Koga M, et al. Quantitative analyses of sonographic images of the parotid gland in patients with Sjögren's syndrome. Ultrasound in Medicine and Biology 2006;32(5): 617–22.
[4] Papini E, Guglielmi R, Bianchini A, Crescenzi A, Taccogna S, Nardi F, et al. Risk of malignancy in nonpalpable thyroid nodules: predictive value of ultrasound and color-Doppler features. Journal of Clinical Endocrinology and Metabolism 2002;87(5):1941–6.

[5] Jasjit S, Wilson D, Laxminarayan S. Handbook of biomedical image analysis. New York: Springer-Verlag; 2005.
[6] Bamber JC, Daft C. Adaptive filtering for reduction of speckle in ultrasonic pulse-echo images. Ultrasonics 1986;24(1):41–4.
[7] Mailloux G, Bertrand M, Stampfler R, Ethier S. Local histogram information content of ultrasound B-mode echographic texture. Ultrasound in Medicine and Biology 1985;11(5):743–50.
[8] Mailloux G, Bertrand M, Stampfler R, Ethier S. Computer analysis of echographic textures in Hashimoto disease of the thyroid. Journal of Clinical Ultrasound 1986;14(7):521–7.
[9] Morifuji H. Analysis of ultrasound B-mode histogram in thyroid tumors. Nippon Geka Gakkai Zasshi 1989;90(2):210–21.
[10] Hirning T, Zuna I, Schlaps D. Quantification and classification of echographic findings the thyroid gland by computerized b-mode texture analysis. European Journal of Radiology 1989;9(4):244–7.
[11] Smutek D, Sara R, Sucharda P, Tjahjadi T, Svec M. Image texture analysis of sonograms in chronic inflammations of thyroid gland. Ultrasound in Medicine and Biology 2003;29(13):1531–43.
[12] Tsantis S, Cavouras D, Kalatzis I, Piliouras N, Dimitropoulos N, Nikiforidis G. Development of a support vector machine-based image analysis system for assessing the thyroid nodule malignancy risk. Ultrasound in Medicine and Biology 2005;31(11):1451–9.
[13] Skouroliakou C, Lyra M, Antoniou A, Vlahos L. Quantitative image analysis in sonograms of the thyroid gland. Nuclear Instruments and Methods in Physics 2006;569(2):606–9.
[14] Savelonas MA, Iakovidis DK, Dimitropoulos N, Maroulis D. Computational characterization of thyroid tissue in the radon domain. In: IEEE international symposium on computer-based medical systems; 2007.p. 189–92.
[15] Keramidas EG, Iakovidis D, Maroulis D, Karkanis SA. Efficient and effective ultrasound image analysis scheme for thyroid nodule detection lecture notes in computer science, 4633. Berlin/Heidelberg: Springer; 2007. p. 1052–1060.
[16] Vapnik V. Estimation of dependencies based on empirical data. New York: Springer-Verlag; 1982.
[17] Julesz B. Textons, the elements of texture perception, and their interactions. Nature 1981;290(5802):91.
[18] Haralick RM, Dinstein I, Shanmugamm K. Textural features for image classification, systems, man and cybernetics. IEEE Transactions 1973;3(6):610–21.
[19] Muzzolini R, Yang YH, Pierson R. Texture characterization using robust statistics. Pattern Recognition 1994;27(1):119–34.
[20] Galloway MM. Texture analysis using gray level run lengths. Computer Graphics Image Processing 1975;4:172–9.
[21] Ojala T, Pietikäinen M, Harwood D. A comparative study of texture measures with classification based on featured distribution. Pattern Recognition 1996;29(1):51–9.
[22] Jawahar CV, Ray AK. Incorporation of grey-level imprecision in representation and processing of digital images. Pattern Recognition Letters 1996;17(7):541–6.
[23] Keramidas EG, Iakovidis DK, Maroulis D, Dimitropoulos N. Thyroid texture representationvia noise resistant image features. In: IEEE international symposium on computer-based medical systems; 2008.p. 560–5.
[24] E.G. Keramidas, Ultrasound image processing and analysis framework, MSc thesis, University of Athens, Greece, 2007.
[25] Keramidas EG, Iakovidis D, Maroulis D. Noise robust statistical feature distributions for texture analysis. In: European signal processing conference; 2008.
[26] Ahonen T, Pietikäinen M. Soft histograms for local binary patterns. In: Finnish signal processing symposium; 2007.
[27] Tomimori EK, Camaro CY, Bisi H, Medeiros-Neto G. Combined ultrasonographic and cytological studies in the diagnosis of thyroid nodules. Biochimie 1999;81(5):447–52.
[28] Kaus MR, Warfield SK, Jolesz FA, Kikinis R. Segmentation of meningiomas and low grade gliomas in MRI. In: International conference on medical image computing and computer-assisted intervention, vol. 1679; 2006. p. 1–10.
[29] Iakovidis DK, Maroulis DE, Karkanis SA. An intelligent system for automatic detection of gastrointestinal adenomas in video endoscopy. Computers in Biology and Medicine 2006;36(10):1084–103.
[30] Ruiz A, Sertel O, Ujaldon M, Catalyurek U, Saltz J, Gurcan M. Pathological image analysis using the GPU: stroma classification for neuroblastoma. In: IEEE international conference on bioinformatics and biomedicine; 2007.p. 78–88.
[31] Hadizadeh H, Baradaran Shokouhi S. Random texture defect detection using 1-D hidden Markov models based on local binary patterns transactions on information and systems, 7. Fremont: Oxford University Press; 2008 . p. 1937–1945.
[32] Grabner H, Nguyen TT, Gruber B, Bischof H. On-line boosting-based car detection from aerial images. Journal of Photogrammetry and Remote Sensing 2008;63(3):382–3.
[33] Zadeh LA. Fuzzy sets. Information and Control 1965;8:338–53.
[34] De Luca A, Termini S. Entropy measures in fuzzy set theory. Systems and Control Encyclopedia 1988;1467–73.
[35] Schiemann U, Gellner R, Riemann B, Schierbaum G, Menzel J, Domschke W, et al. Standardized grey scale ultrasonography in Graves' disease: correlation to autoimmune activity. European Journal of Endocrinology 1999;141(4):332–6.
[36] Jawahar CV, Ray AK. Fuzzy statistics of digital images. IEEE Signal Processing Letters 1996;3(8):225–7.
[37] Yager RR, Filev DP. Essentials of fuzzy modeling and control. New York: Wiley Interscience; 1994.
[38] Vapnik V. Statistical learning theory. New York: Wiley; 1998.
[39] Chang CC, Lin CJ. LIBSVM: a library for support vector machines; 2001, Software available at http://www.csie.ntu.edu.tw/~cjlin/papers/libsvm.ps.gz.
[40] Ling CX, Huang J, Zhang H. AUC: a better measure than accuracy in comparing learning algorithms. In: Canadian conference on artificial intelligence, vol. 2671; 2003.p. 329–41.
[41] Huang J, Ling CX. Using AUC and accuracy in evaluating learning algorithms. IEEE Transactions on Knowledge and Data Engineering 2005;17(3):299–310.
[42] Theodoridis S, Koutroumbas K. Pattern recognition, 3rd ed., Academic Press; 2006.
[43] Smutek D, Sara R, Sucharda P, Tesar L. Different types of image texture features in ultrasound of patients with lymphocytic thyroiditis. In: ACM international conference proceeding series, vol. 49; 2003.p. 100–2.
[44] Sammon JW. A nonlinear mapping for data structure analysis. IEEE Transactions on Computers 1969;18(5):401–9.