Classification results of the logo candidate regions for 8-class classification problem are shown in Table 4.

Table (4): Classification results of the shown logos in Fig. (10)

Class	Number of logos	True accept	False accept	False reject
1	15	8	14	7
2	15	11	3	4
3	15	12	6	3
4	15	6	42	9
5	15	7	8	8
6	15	14	3	1
7	15	11	3	4
Reject	2011	1932	36	79
Total	2116	2001	115	115

4. Conclusion

In this paper, we present a novel framework for logo detection and recognition from document images using texture-based document image segmentation and segm–ented candidate region classification. In this framework, the document image is segmented by a two-stage segme–ntation algorithm and then the segmented regions are classified by a final KNN classifier and two MLP classi–fiers in a hierarchical structure. A perfect set of texture and shape features are used to classify segmented regions and logo candidate regions. The logo candidates are recognized into pre-defined classes by a KNN classifier. The obtained results in the logo detection and recognition stages of the proposed framework show more than efficiency and effectiveness the previously presented works.

References

[1] K. Gao, Sh. Lin, Y. Zhang, Sh. Tang, D. Zhang, "Logo detection based on spatial-spectral saliency and partial spatial context", IEEE Int. Conf. on Mult. and Expo., pp.322-329, 2009.

[2] M. Goria, M. Magginia, S. Marinaib, J.Q. Shengc, G. Sodab, "Edge-backpropagation for noisy logo recognition", Pattern Recognition, Vol.36, No.3, pp.103-110, 2003.

[3] D. Doermann, E. Rivlin, I. Weiss, "Applying algebraic and differential invariants for logo recognition", Machine Vision and Application, Vol.9, No.2, pp.73-86, 1996.

[4] I.S. Hsieh, K.C. Fan, "Multiple classifiers for color flag and trademark image retrieval", IEEE Trans. on Image Proc., Vol.10, No.6, pp.938-950, 2001.

[5] J. Neumann, H. Samet, A. Soffer, "Integration of local and global shape analysis for logo classification", Pattern Recognition Letters, Vol.23, No.12, pp.1449-1457, 2002.

[6] P. Suda, C. Bridoux, B. Kammerer, G. Maderlechner "Logo and word matching using a general approach to signal registration", In Proc. Int'l Conf. Doc. Ana. and Rec., pp.61-65, 1997.

[7] K. Zyga, J. Schroeder, R. Price, "Logo recognition using retinal coding", In Proc. 38th Asilomar Conf. Signals, Sys. and Comp., Vol.2, pp.1549-1553, 2004.

[8] E. Francesconi, P. Frasconi, M. Gori, S. Marinai, J.Q. Sheng, G. Soda, A. Sperduti, "Logo recognition by recursive neural networks", Graph. Rec. Alg. and Sys., Vol.1389, pp.104-117, 1998.

[9] S. Seiden, M. Dillencourt, S. Irani, R. Borrey, T. Murphy, "Logo detection in document images", Proc. of the Inte. Conf. on Imag. Sci., Sys., and Tech., Las Vegas, Nevada, pp.446-449, 1997.

[10] T. Pham, "Unconstrained logo detection in document images", Pattern Recognition, Vol.36, No.12, pp.3023-3025, 2003.

[11] J. Chen, M.K. Leung ,Y. sheng Gao, "Noisy logo recognition using line segment Hausdorff distance", Pattern recognition, Vol.36, No.5, pp.943-955, 2003.

[12] G. Zhu, D. Doermann, "Automatic document logo detection", 9th Int. Conf. on Doc. Ana. and Rec., Vol.2, pp.864-868, 2007.

[13] Z. Ahmed, H. Fella, "Logos extraction on picture documents using shape and color density", The Int. IEEE Conf. on Indus. Elec., pp.2492-2496, 2008.

[14] H. Wang, Y. Chen, "Logo detection in document images based on boundary extension of feature rectangles", 10th Conf. of Doc. Ana. and Rec., pp.1335-1339, 2009.

[15] Y.S. Kim, W.Y. Kim, "Content-based trademark retrieval system using a visually salient feature", Image and Vision Computing, Vol.16, No.10, pp.931-939, 1998.

[16] P.Y. Yin, C.C. Yeh, "Content-based retrieval from trademark databases", Patt. Rec. Lett., Vol.23, No.9, pp.113-126, 2002.

[17] M.H. Hung, C.H. Hsieh, C.M. Kuo, "Similarity retrieval of shape images based on database classification", J. Vis. Commun. Image R., Vol.17, No.4, pp.970-985, 2006.

[18] H. Pourghassem, "Page layout analysis of the document image based on the region classification in a decision hierarchical structure", Jou. Elec. and Pow. Engi., Najafabad University, Vol.1, No.3, pp.37-44, Spring 2009.

[19] H. Pourghassem, H. Ghassemian, "Content-based medical image classification using a new hierarchical merging scheme", Jou. of Comp. Med. Imag. and Graph., Vol.22, No.8, pp.651-661, 2008.

[20] N. Otsu, "A threshold selection method from gray-level histograms", IEEE Trans. on Sys., Man., and Cyb., Vol.9, No.1, pp.62-66, 1979.

[21] A.K. Jain, P.W. Duin, J. Mao, "Statistical pattern recognition: A review", IEEE Trans. Pat. Anal. Mach. Intell., Vol.22, No.1, pp.4-36, 2000.

[22] H. Tamura, S. Mori, T. Yamawaki, "Texture features corresponding to visual perception", IEEE Trans. on Sys., Man., and Cyb., Vol.8, No.6, June 1978.

[23] R.M. Haralick, K. Shanmugan, I. Dinstein, "Textural features for image classification", IEEE Trans. on Sys. Man. and Cyb., Vol.3, No.6, pp.610-621, 1973.

[24] E. Persoon, K. Fu, "Shape discrimination using Fourier descriptors", IEEE Trans. Sys. Man. and Cyb., Vol.7, No.5, pp.170-179, 1977.

[25] A.L. Blum, P. Langley, "Selection of relevant features and examples in machine learning", Artificial Intelligence, Vol. 97, No.7, pp.245-271, 1997.

[26] W.Y. Kim, Y.S. Kim, "A new region-based shape descriptor", ISO/IEC MPEG99/M5472, TR15-01, Maui, Hawaii, Dec. 1999.

3.3.1. Feature Selection

List of the extracted features and their short form are obtained in Table 1. Image representations were generated in 6 different feature spaces (Table 2) using the following features: Tamura (including of contrast, coarseness and directionality) (T), co-occurrence (including of correlation, energy and homogeneity) (C), Fourier descriptor-complex representation (F), directional histogram (D) and tessellation-based spectral (S) features.

The final classifier (KNN) is evaluated with test data and various values of K. Classification results for 3-class classification problem (i.e., text, logo and pure picture classes) and different feature spaces with K=17 in KNN classifier are shown in Table 2.

All feature spaces in Table 2 consist of texture features (T and C), because these features have valuable distingui–shable information for region classification. Therefore, weights of these features are more than other features. The optimum feature spaces include of texture features and shape features. S, D and F features obtain higher classification accuracy, respectively. So, the weight of S is considered higher than weight of D and the weight of D is considered higher than weight of F feature. Results indicate that the (T,C,S,D,F) feature space provides higher classification accuracy than the other options investigated. Accuracy rate 98.43% is obtained with this feature space and (1.5,1.5,1.2,1.1,1) for its weight.

3.3.2. Classification Results

Segmented regions by two-stage segmentation algorithm are classified via multiple classifiers in a hierarchical structure. Fig. (8) shows classification results of segmented regions that are segmented by two-stage segmentation algorithm in previous section. Segmented regions that are classified as a text or logo region have bounded with solid blue or dash red lines, respectively.

3.4. Logo Recognition

After applying the hierarchical classification on the document image, the segmented regions are labeled with text, pure picture or logo classes. Then in logo recognition stage, the logo candidate regions are classified into pre-defined classes of logos by a KNN classifier. The used feature space in the logo recognition stage consists of Tamura, co-occurrence matrix, Fourier descriptor, tessellation-based spectral, directional histogram, ART, Zernike features. These features are applied in KNN classifier for the logo candidate region classification into *m* classes of pre-defined classes of logos and one class for rejection class. Based on complexity of the logo pattern, a set of the extracted features in Table 3 is formed until the best classification carried out. For example, to classify the shown logos in Fig. (9), feature vector as (C,F,S,D,A,Z) has formed by forward selection algorithm for a 7-class classification problem. The proposed algorithm obtains accuracy rate of 94.57% with feature space (C,F,S,D,A,Z) for 8-class classification problem (7 logo classes and one reject class). i.e., false accept of this classification problem is 5.43%. Fig. (10) shows sample images that have classified incorrectly in 7 classes. In other words, in Fig. (10), samples of false accept logos have shown.

Table (3): The extracted features.

Features	Feature length	Short form
Tamura (Contrast, Coarseness and Directionality)	3	T
Co-occurrence (Correlation, Energy and Homogeneity)	12	C
Fourier descriptor- complex representation	62	F
Tessellation-based spectral	16	S
Directional histogram	36	D
ART	35	A
Zernike	5	Z

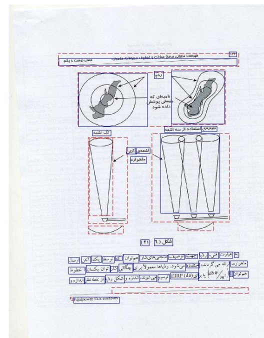


Fig. 8: Classification results document image by hierarchical classifier (bound in box of text and logo regions are determined with solid blue and dash red lines, respectively.).

Fig. 9: Image sample of logo database in 7 classes.

(a) (b)

Fig. 10: Image sample of false accept logos in each class. (a) Image sample of original logo, (b) image sample of false accept logos

3. Experimental Results

3.1. Document Image and logo Database

The used document image database is a collection of 977 images comprising three classes, document image (images only with text regions), pure picture (images without text regions) and combined images (images with text and picture regions). This database is used to evaluate the two-stage segmentation algorithm. The used logo database is a collection of 1980 images comprising international and Persian logos. To evaluate the proposed logo detection and recognition algorithm, these logos have inserted to document images manually. Image sample of the document image database and logo database are shown in Fig. (5) and Fig. (6), respectively.

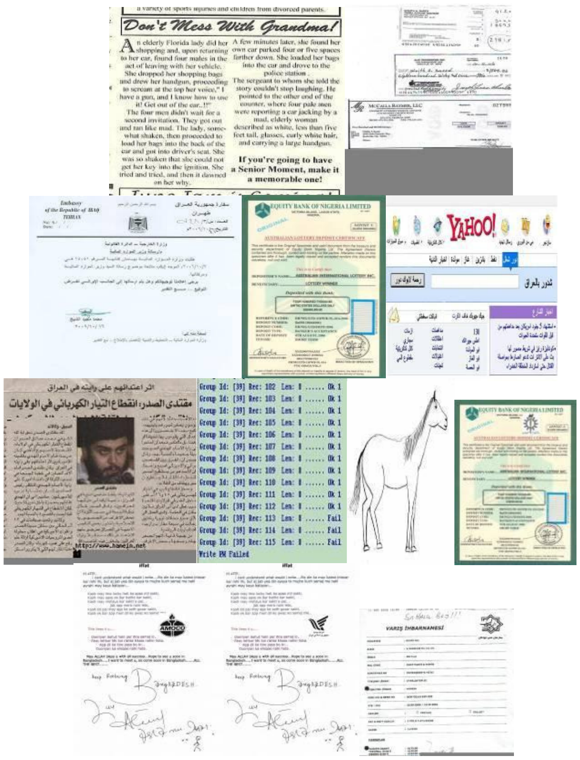
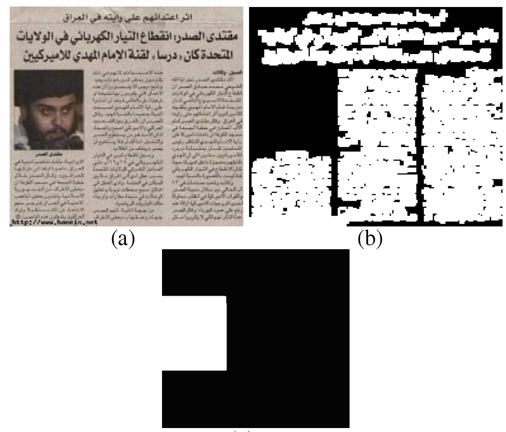


Fig. 5: Image sample of document images database.

Fig. 6: Image sample of logo database

3.2. Two-stage Segmentation

To evaluate two-stage segmentation algorithm, sample images are segmented by the wavelet and threshold-based segmentation algorithm. Fig. (7) shows image containing with text and picture regions. In these example, text regions are extracted from image in first stage of segmentation algorithm (Fig. (7.(b))) and picture region is segmented by threshold-based segmentation algorithm (Fig. (7.(c))). White regions in the images are segments and black regions are background of image.



(a) (b) (c)

Fig. 7: Applying two-stage segmentation on a scanned newspaper. (a). Scanned image, (b). Segmented regions with wavelet-based segmentation algorithm and (c). Segmented region with threshold-based segmentation algorithm on the rest of image from (b).

Table (1): The extracted features.

Features	Feature length	Short form
Tamura (Contrast, Coarseness and Directionality)	3	T
Co-occurrence (Correlation, Energy and Homogeneity)	12	C
Fourier descriptor- complex representation	62	F
Tessellation-based spectral	16	S
Directional histogram	36	D

Table (2): Classification results for KNN classifier with K=17 and 3-class (text, pure picture and logo classes).

Features(and their weights in feature vector)	Feature length	Accuracy%
T,C,F (1.5,1.5,1)	77	82.97
T,C,S (1.5,1.5,1)	31	89.35
T,C,D (1.5,1.5,1)	51	84.66
T,C,S,F (1.5,1.5,1.2,1)	93	92.15
T,C,S,D (1.5,1.5,1.2,1)	67	96.64
T,C,S,D,F (1.5,1.5,1.2,1.1,1)	129	98.43

3.3. Hierarchical classification

Segmented regions that are segmented by the two-stage segmentation algorithm, should be classified and labeled with text and picture (pure picture or logo) regions. There are two points in classification problems, first, the selection of optimum feature vector and second, the set of classifier parameters, for example, the number of neurons in middle layer of MLP classifier or the value of k in KNN classifier. To select the best set of extracted features is used from forward selection algorithm and to set the classifier parameters is used by trial and error way.

object is re-sampled to M samples by a uniform sampling function before performing the Fourier transform. The Fourier descriptor of the complex coordinate is:

$$f_Z = \left[\frac{F_{-(M/2-1)}}{|F_1|},\dots,\frac{|F_{-1}|}{|F_1|},\frac{|F_2|}{|F_1|},\dots,\frac{|F_{M/2}|}{|F_1|}\right] \quad (9)$$

where F_0 and F_1 are DC and the first non-zero frequency components used for normalizing the transform coefficients, respectively. In this paper, $M = 256$ is set. Therefore, a feature vector of the complex coordinate function has 254 elements.

2.2.2. Feature Selection

Our goal in feature selection is to find a minimum set of features that is obtained best discrimination between classes. We would like to find the features that most accurately discriminate among the classes and will yield the highest classification accuracy. Since the optimum set of features is unknown, usually two traditional forward selection and backward elimination algorithms are used [25]. Forward selection algorithm starts with an empty set of features and adds one feature at a time until the final feature set is reached. Backward elimination algorithm starts with a feature set containing all features and removes features. In this paper, forward selection algorithm is used and best set of features is selected.

2.3. Logo Recognition

After applying the two-stage segmentation algorithm and the hierarchical classification on the document image, the segmented regions are labeled with text, pure picture or logo classes. Then logo regions are classified into pre-defined classes of logos. Because, some non-logo regions may be labeled with logo regions by the final classifier in the previous stage, a class as reject class is added to logo classes in logo recognition component.

So, in the logo recognition stage, we encounter with a (m+1)-class classification problem that *m* is the number of logo classes. To solve this classification problem as other problems of classification, there are two problems, feature extraction of logo regions and classification of logos by extracted features. In this stage, therewith extracted features from regions in the previous stages, some new features are extracted from regions that improve classification accuracy of logo classes. These features are consisted of angular radial transform (ART) [26] and Zernike moments [15]. KNN classifier is used to classify segmented regions to a class of pre-defined logo classes or reject class.

To train this classifier with training samples, a new strategy in the training procedure is applied. To improve classification accuracy and also insensitive scale and rotation classification of logos, 14 new samples (rotated samples with rotation angles of $\pm 10^0$, $\pm 7^0$, $\pm 5^0$ and $\pm 2^0$ and resized samples with scale coefficients of 0.6, 0.7, 0.8, 0.9, 1.1 and 1.2) from each logo are generated and their feature vector extracted for training of KNN in the feature space. In test procedure, these 15 samples (14 new generated samples and original sample) are considered as 15 logo samples that should be classified into (m+1)-class by KNN classifier whereas for each input sample of 15 logo samples k nearest samples in the training dataset are identified and ultimately, maximum votes in (15×k) samples are selected as label of input logo. With this scheme, not only logo classification by KNN classifier becomes robust to resized and rotated unknown samples but also classification accuracy is improved.

2.3.1. Zernike moments

Zernike moments are defined inside the unit circle and the radial polynomial vector $R(\rho)$ is defined as [15]:

$$R_{mn}(\rho) = \sum_{s=0}^{\frac{n-|m|}{2}} (-1)^s \frac{(n-s)!}{s!(\frac{n+|m|}{2}-s)!(\frac{n-|m|}{2}-s)!}\rho^{n-2s} \quad (10)$$

$$R(\rho) = \{R_{mn}(\rho) \mid n = 0,1,2,..,\infty, |m| \le n, \text{ and } n-|m| \text{ is even}\} \quad (11)$$

Then the two-dimensional Zernike moment of an image, $I(\rho,\theta)$, in polar coordinate is defined as:

$$A = \frac{n+1}{\pi}\sum_{\rho}\sum_{\theta}[V(\rho,\theta)]^* I(\rho,\theta), \text{s.t.} \rho \le 1 \quad (12)$$

Here, $V(\rho,\theta)$ is a Zernike basis polynomial defined as:

$$V(\rho,\theta) = R(\rho)\exp(-jm\theta) \quad (13)$$

The magnitude of Zernike moment (ZMM) is defined as,

$$z = \|A\| \quad (14)$$

So, the **z** denotes the vector of z_{mn}. That is,

$$z = \{z_{mn} \mid n = 0,1,2,..,\infty, |m| \le n, \text{ and } n-|m| \text{ is even}\} \cdot \quad (15)$$

In our application, $n = 8$ is set. So Zernike moments feature has 5 elements.

2.3.2. Angular Radial Transform (ART)

ART is the 2-D complex transform defined on a unit disk that consists of the complete orthonormal sinusoidal basis functions in polar coordinates. The transformation is defined as [26]

$$F_{nm} = \langle V_{nm}(\rho,\theta), f(\rho,\theta)\rangle = \int_0^{2\pi}\int_0^1 V_{nm}^*(\rho,\theta), f(\rho,\theta)\rho\, d\rho\, d\theta \quad (16)$$

Here, F_{nm} is an ART coefficient of order n and m, $f(\rho,\theta)$ is an image function in polar coordinates, and $V_{nm}(\rho,\theta)$ is the ART basis function that are separable along the angular and radial directions, i.e.,

$$V_{nm}(\rho,\theta) = A_m(\theta)R_n(\rho) \quad (17)$$

The angular and radial basis functions are defined as follows:

$$A_m(\theta) = \frac{1}{2\pi}\exp(jm\theta) \quad (18)$$

$$R_n(\rho) = \begin{cases} 1 & n = 0 \\ 2\cos(\pi n\rho) & n \ne 0 \end{cases} \quad (19)$$

To describe a shape, all pixels constituting the shape are transformed with ART, and the transformed coefficients are formed into the ART descriptor. Twelve angular and three radial functions are used. By discarding the DC coefficient, 35 AC components form the descriptor vector.

only the number of samples in the neighborhood of unknown sample (k) should be set. Only texture features are used to train the first and second classifiers, but both texture and shape features are used to train the final classifier. Because, the first and second classifiers identify the segmented regions based on texture information but the final classifier classify the segmented regions based on texture and shape information.

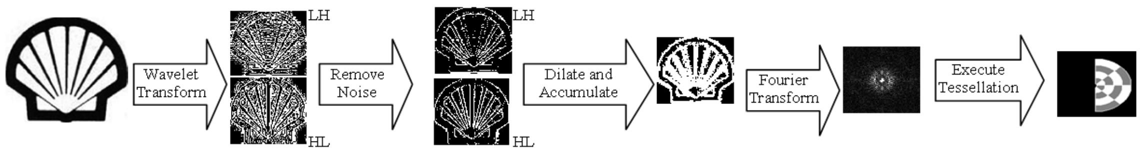


Fig. 4: Block diagram of the tessellation-based spectral feature extraction

2.2.1. Feature Extraction

In the hierarchical classifier, texture features such as Tamura features (including coarseness, contrast and directionality) [22], co-occurrence matrix features (including correlation, energy and homogeneity in 0^0, 45^0, 90^0 and 135^0 directions) [23], and shape features such as Fourier descriptor coefficient with representation complex coordinate function [18] and modified tessellation-based spectral and directional histogram features [19]. These features extract valuable information of both texture and shape contents of the segmented regions. Tessellation-based spectral feature has more than shape information whereas directional histogram feature has more than texture information. In the following, the tessellation-based spectral, co-occurrence matrix and Fourier descriptor coefficient features are described in detail.

2.2.1.1. The Tessellation-based Spectral Feature

Valuable information of edge and orientation (shape feature); also coarseness (texture feature) of objects in logo regions are captured to apply frequency analysis on the wavelet transform sub-bands. The block diagram of the tessellation-based spectral feature extraction is shown in Fig. (4). This feature is extracted in the following steps [19]:

1) To apply Meyer wavelet transform in one level on the medical image.
2) To remove noisy coefficients by applying a threshold.
3) To execute dilation morphological operators on the LH and HL sub-bands for amplifying edges in two horizontal and vertical directions, respectively.
4) To eliminate edges with the number of pixels less than a threshold.
5) To accumulate two LH and HL sub-bands into an accumulated image.
6) To execute the Fourier transform on the accumulated image.
7) To define a new tessellation scheme on the frequency spectrum based on an appropriate definition of frequency concept in the medical X-ray image classification.

In step 7, the tessellation scheme is defined as below:

$$S_{i,j} = \{(u,v) \mid R_i < r < R_{i+1}, \theta_j < \theta < \theta_{j+1}, 1 \le u \le N, 1 \le v \le M\} \quad (1)$$

where

$$r = \sqrt{(u-u_c)^2 + (v-v_c)^2} \quad (2)$$

$$R_i \in \{R_1, R_2, .., R_n\} \quad (3)$$

$$\theta = \tan^{-1}\left(\frac{v-v_c}{u-u_c}\right) \quad (4)$$

$$\theta_j \in \{-\pi/2, -\pi/4, 0, \pi/4, \pi/2\} \quad (5)$$

$N \times M$ is the image size, (u_c, v_c) is the center of frequency spectrum. R_i $(i = 1,2,..,n)$ is determined as follows:

$$Area_{l,m} = 2Area_{k,n} \quad \text{for all } m, n \text{ and } l = k+1 \quad (6)$$

where $Area_{i,j}$ is the number of pixels in the sector $S_{i,j}$. In our application, R_1=5 and R_5= (min (N, M))/3 is set. The Frequency spectrum of the real image is symmetric relation to the center of spectrum. So a half of spectrum information is redundancy and discarded. In this paper, the standard deviation within the sectors defines the feature vector. The standard deviation is defined as:

$$f_{i,j} = \sqrt{\sum_{Area_{i,j}} (I(u,v) - m_{ij})^2} \quad (7)$$

Where $m_{i,j}$ is the mean of the pixel values $S_{i,j}$. The extracted feature vector for an image is a $(n-1)$ by 4 (the number of arcs in the tessellation scheme) matrix. In this paper, $n = 5$ is set. So, the feature vector of the tessellation-based spectral feature has 16 elements.

2.2.1.2. Fourier Descriptors

Fourier descriptors describe the shape of an object with the Fourier transform of its boundary. Consider the contour of a 2D object as a closed sequence of successive boundary pixels (x_s, y_s), where $0 \le s \le N-1$ and N is the total number of pixels on the boundary. Then three types of contour representations, i.e., centroid distance, curvature, and complex coordinate function, can be defined [24]. In this paper, Fourier descriptor with complex coordinate function representation is used. The complex coordinate is obtained by simply representing the coordinates of the boundary pixels as complex numbers:

$$z(s) = (x_s - x_c) + j(y_s - y_c) \quad (8)$$

To ensure that the resulting shape features of all objects in a database have the same length, the boundary of each segmentation algorithm has two major advantages. Firstly, the text segments inside and adhesive to the picture segments are properly extracted in the first stage of the segmentation algorithm, secondly, the combined text-picture segments are segmented from the rest of the document image with more than accurate in the second stage of segmentation algorithm. Fig. (3) shows middle results of applying the wavelet-based segmentation algorithm on the sample image.

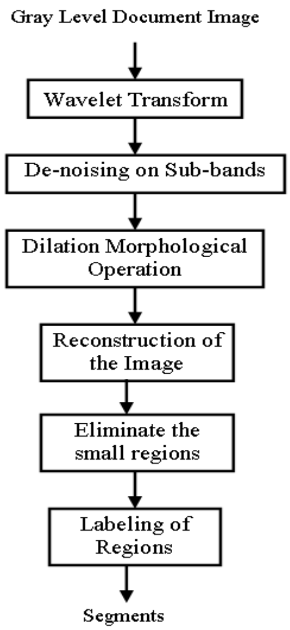


Fig. 2: Block diagram of the wavelet-based segmentation algorithm [18]

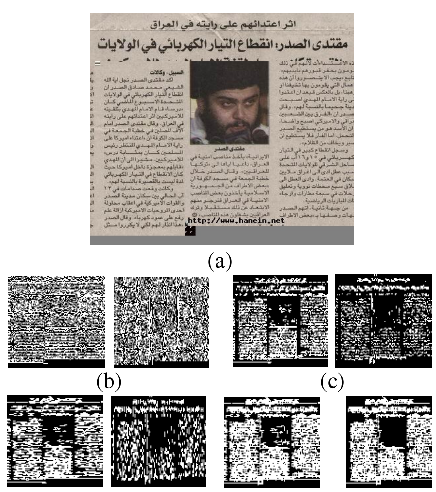


Fig. 3: Middle results of applying the wavelet-based segmentation algorithm on the sample image. (a). Sample image, (b). Horizontal (left image) and vertical (right image) sub-bands in level 2 of wavelet transform, (c). De-noised sub-bands (left and right images are horizontal and vertical sub-bands, respectively) of (b), (d). Dilated sub-bands (left and right images are horizontal and vertical sub-bands, respectively) of (c), (e). Accumulated image of horizontal and vertical sub-bands, (f). Filtered image of (e)

2.1.1. Threshold-based segmentation algorithm

The detected segments by the wavelet-based segmentation algorithm are removed from the document image and the rest of the image is segmented by the threshold-based segmentation algorithm. In the second segmentation algorithm, non-segmented regions from applying the first algorithm are divided to foreground and background sections by applying a threshold that is determined from gray-level histogram based on Otsu's method [20]. Then foreground sections are labeled and extracted from image as text or picture segments.

2.2. The Hierarchical Classification

There are two major reasons for combining multiple classifiers to solve a given classification problem [21]. Firstly, there are a number of different classifiers, each developed in a different context and for an entirely different representation/description of the same problem. Secondly, different classifiers trained on the same data may not only differ in their global performances, but they also may show strong local differences. Each classifier may have its own region in the feature space where it performs the best [21].

In our application, the text and picture classes have diverse samples and considerable overlap consequently they do not classify by a single classifier and limited features correctly. The hierarchical classifier is an appropriate approach to overcome this problem. In the hierarchical architecture, individual classifiers are combined into a structure, which is similar to that of a decision tree classifier. The tree nodes, however, may now be associated with complex classifiers demanding a large number of features [21]. The proposed hierarchical classifier is designed with three classifiers (Fig. (1)). The first and second classifiers are a multi-layer Perceptron (MLP) neural network and the final classifier is a k-nearest neighbor (KNN) classifier that all train with different features and samples, separately. MLP is a network with three layers (input, hidden and output layers) that trains with error back-propagation algorithm. The number of units in the input and output layers are equal to the number of feature vector elements and the number of classes, respectively. The number of units in the hidden layer has major effect on the classification performance that is determined by trial and error method.

The First and second classifiers classify segmented regions into text and picture classes that segment by the wavelet and threshold-based algorithm, respectively. In the first classifier, false accept for text class is close to zero.

Therefore, the text regions in this stage are removed from the document image and other regions are remained until classified. The more text regions in the document image are determined in the first classifier and seldom text region are remained in the document image. So, the second classifier is properly classified the remained regions into text and picture classes. Because, text regions can be as a logo or trademark,so, we reclassify the all text and picture regions into text, pure picture and logo classes by a KNN classifier. Logo regions are separated of the pure picture and text regions by the final classifier. In KNN classifier, examples to get a list of database trademarks ordered by similarity ranks. In [4], a region-based multiple classifier color image retrieval system was presented. In this approach, a region-growing technique for segmentation of the input image into logo candidate regions and three complementary region-based classifiers (color, shape and relational classifiers) were applied to logo recognition. In each classifier, a virtue probability representing the probability that an image is similar to the query image is defined. A set of virtue probabilities was calculated to define similarity measure in each classifier. In [17], a shape-based similarity retrieval system is developed based on database classification which exploits the contour and interior region of a shape efficiently. In this paper, angular radial transform (ART) region feature is employed to compare the query with the candidate sets according to the priority order.

In this paper, we present a novel framework for logo detection and recognition from document images using texture-based document image segmentation and segmented candidate region classification. In this framework, the document image is segmented by a two-stage segmentation algorithm that is performed as a cascade process on the document image [18]. The segmented regions via the first stage of segmentation algorithm (wavelet-based algorithm) are classified by a multi layer Perceptron (MLP) classifier. Classified regions as text region are removed from the document image until the second stage of segmentation algorithm (threshold-based algorithm) is performed to the rest of image. The segmented regions via threshold-based segmentation algorithm are classified into two classes, text and picture by a MLP classifier. Two-stage segmentation strategy provides better results due to proper segmentation of text region by the wavelet-based algorithm and complexity reduction of the image pattern in the threshold-based algorithm. Because logos can be formed from both text and shape elements, so the classified segments by two MLP classifiers are reclassified into text, pure picture and logo classes by a k-nearest neighbor (KNN) classifier as a final classifier. Ultimately, detected logos are classified into pre-defined classes by a KNN classifier with a new training scheme as logo recognition component. Texture and shape features are significantly used in the first two classifiers and final classifier, respectively. In logo recognition component, a set of shape and texture features that are previously presented in literature along with two features that are proposed in [19], are used.

This paper is organized as follows: Section 2 presents details of the proposed framework for logo recognition from the document image. In this section, two-stage segmentation algorithm, the hierarchical classification and feature extraction stages are described in detail. Experimental results are shown in Section 4. Section 5 provides a conclusion to the work.

2. The Proposed Framework for Logo Recognition from Document Images

Block diagram of the proposed framework for logo recognition is shown in Fig. (1). This framework consists of three major components, i.e. the two-stage segmen–tation, the hierarchical classification and the logo recognition. The two-stage segmentation is applied on the document image until it is segmented into small and large region [18]. The hierarchical classifier is then classified the detected segments of the document image into three classes, text, pure picture (personal photo, sign and table are labeled with pure picture) and logo. Ultimately, the extracted logo segments are labeled based on predefined classes. Each component of the proposed framework will be described in the following.

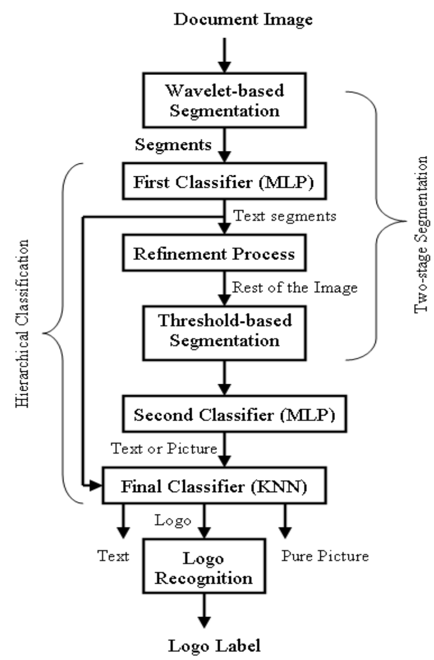


Fig. 1: Block diagram of the proposed framework for logo recognition

2.1. The two-stage segmentation algorithm

In this paper, the proposed segmentation algorithm in [18] is used to carry out with different strategies in two stages, the wavelet-based and threshold-based segmenta–tion algorithms. Block diagram of this algorithm is shown in Fig. (2). In the first stage, the text segments are mostly extracted from the document image by wavelet-based algorithm and the rest of image (including of the pure picture and logo regions, lines of tables and the rest of text segments) is segmented by the threshold-based segment–ation algorithm in the second stage. The extraction process of text regions (segments) from the document image that has detected by the first classifier is performed in the refinement component. The segmented picture regions via the wavelet-based segmentation algorithm are not removed from the document image until it is segmented again and reclassified by the threshold-based algorithm and the second classifier, respectively.

The regions which are classified by the first classifier as a picture segment have not been segmented perfectly. So, these regions remain in the document image for applying the next segmentation algorithm. The proposed two-stage

A Novel Framework for Logo Detection and Recognition from Document Images

Hossein Pourghassem[(1)] - Amir Salar Jafarpisheh[(2)]
(1) Assistant Prof. - Department of Electrical Engineering, Islamic Azad University, Najafabad Branch
(2) Ph.D Candidate - Department of Biomedical Engineering, Tehran University of Medical Sciences

Recive: Spring 2011 Accept: Spring 2012

Logo detection and recognition module is a vital requirement in official automation systems for document image archiving and retrieval applications. In this paper, we present a novel framework for logo detection and recognition based on sequential segmentation and classification strategy of document image. In this framework, using a two-stage segmentation algorithm (consisting of wavelet-based and threshold-based segmentation algorithms) and hierarchical classification by two multilayer Perceptron (MLP) classifiers and a k-nearest neighbor (KNN) classifier, a document image divides to text, pure picture and logo candidate regions. Ultimsately, in final decision, class of logo candidate region is determined based on pre-defined classes. In the hierarchical classification and logo recognition stages, the best feature space is selected by forward selection algorithm from a perfect set of texture and shape features. The proposed structure is evaluated on a variety and vast database consisting of the document and non-document images with Persian and international logos. The obtained results show efficiency of the proposed framework in the real and operational conditions.

Index Terms: Logo detection and recognition, document image, two-stage segmentation, hierarchical classification.

1. Introduction

With tremendous increase in multimedia databases, the demands for storing multimedia information (such as text, image, audio, and video) have increased. Along with the need for search and retrieval richer tools is unavoidable. In this domain, document image analysis and understanding have received a great deal of interests in the last few years for many diverse applications such as, digital library, Internet publishing and searching, on-line shopping and official automation systems. Along with logo detection and recognition is an important requirement in the document image analysis and shape matching domain as it enables us to identify the source of documents based on the organization where a document originates. For example, in official automation systems application, content of scanned official letters (such as a document image) should be recognized and classified by their logos. So, logos can act as a valuable means in identifying sources of documents [1].

The major previously research related to logo in document images have focused in logo recognition [2-8] and rare investigation have consisted of both logo detection and recognition [9-14]. In [11], a modified line segment Hausdorff distance has been proposed that incorporates structural and spatial information to compute dissimilarity between two sets of line segments rather than two sets of points. In this paper, logo is first generated to line segments and represented with feature vector. In [9], a logo detection system is presented based on segmentation the document image into smaller images using a top-down X-Y cut algorithm. In this paper, a total of sixteen features of the connected components in each segment are extracted and used by a rule-based classification scheme. In [12], an approach to detecting and extracting logos in document images using a multi-scale boosting strategy is presented. An initial two-class Fisher classifier at a coarse image scale on each connected component is used. Each detected logo candidate region is then classified at finer image scales by a cascade of simple classifiers [12]. In [10] a simple logo detection method has presented based on the assumption that the spatial density of foreground pixels in a logo region is greater than that in non-logo regions. A document image is first binarized into foreground and background pixels. Then, the spatial density within each fixed size window is computed and the region with the highest density is hypothesized as a logo region. In [15], a method for such a system based on the image content, using a shape feature was presented. Zernike moments of an image are used for a feature set. In this method, to recognize the detected logo, a similarly measure based on shape feature (Zernike moments) was defined. In [16], an automatic content-based logo retrieval method was proposed. The proposed method automatically selects appropriate features (such as area, deviation, symmetry, centralization, complexity and 2-level contour representation strings) based on feature selection principles to discriminate logo. In this method, the user can submit a query through logo