Scientific Journal Impact Factor (SJIF): 1.711
International Journal of Modern Trends in Engineering and Research
ISSN (ONLINE) : 2349 - 9745 ISSN (PRINT) : 2393 - 8161

A Mobile Application for Quick Classification of Plant Leaf Based on Color and Shape

Mohammed Zaid Shahid Akhtar[1], Dr. Sunil A. Patekar[2], Mandar G. Sohani[3]
[1,2,3] *Computer Engineering, Vidyalankar Institute of Technology*

Abstract— This paper presents an effective way to develop an android application that gives user the ability to identify plant species based on photographs of the plant's leaf taken with mobile phone. The algorithm acquires a color image of the leaf using mobile phone's camera and perform the classification by comparing the acquired image with the stored database image.

The method proposed in the paper performs preprocessing to obtain the binary contour of the leaf, then morphological features of the leaf are obtained to generate feature vector of the leaf, then classifies the species based on the combination of computed feature vector. The system is first trained against several known species and then used to classify unknown species.

The paper also presents a way to improve current system by providing classification based on shape and color of the plant's leaf. To classify the leaf based on shape, the binary contour of the plant's leaf is obtained using Otsu's method and sobel operator. The classification based on color is achieved using the dominant color method. To compare the morphological features, Euclidean distance method is used.

Keywords— android; photographs; morphological feature; classification; preprocessing; shape; dominant; color; Euclidean distance, Otsu's, sobel.

I. INTRODUCTION

Identifying an unknown species previously required the consultation of a heavy field guide, where user would make detail observation of plant's features and navigate a complex decision tree. The process was non-trivial even for seasoned botanist, and carrying the guides into the field was often impractical – particularly for hikers and other casual users [9].

Today, with the help of smart phones, which is computationally strong, equipped with large storage capacity and high resolution camera it is possible to not only replace but improve the database storage and decision tree but also to create automatic leaf classification applications.

Plants are basically classified according to shapes and structures of their flowers or genitals [8]. However it is difficult to input shapes and structures of flowers or genitals because they have complex structures. On the other hand, leaves of plants are planer and input of their shapes is easy [8].

Compared with other methods, such as cell and molecule biology methods, classification based on leaf image is the first choice for plant leaf classification. There has been several attempts in the plant leaf classification field using different methods and techniques, where all of them was computer processed simulation, for example using probabilistic neural network [7], using move median centers (MMC) hypersphere classifier and morphological features [6] etc. However the proposed work in the paper is an android based mobile application for plant species classification.

Most of the previous shaped based algorithms focus on the binary contour of the plant's leaf. The algorithm starts with the image acquisition of the plant's leaf, where the image is a color image.

Next is the preprocessing step, where the image is converted to mutable bitmap image and then into binary image using Otsu's method. Boundary of the leaf is obtained using Sobel operator. Several methods are available to obtain the binary contour of a leaf, e.g. centroid-contour distance (CCD) curve. The dominant color of the leaf is also obtained in this process to be later used in the color matching phase. After preprocessing, numerous morphological features like aspect ratio, rectangularity, sphericity etc. are extracted from the image to generate leaf's feature vector. Finally, to determine the query leaf's correct match, the query leaf's feature vector is computed and its difference from each leaf class represents its feature similarity vector. The feature similarity vector is then normalized and weighted, and their Euclidean distance is calculated. The order of the resulting vector of class distances represents the order of similarity [9].

II. LITERATURE SURVEY

The application of digital image processing techniques to the problem of automatic plant leaf classification began two decades ago and it has been since proceeding in earnest [9]. In 1989, Petry and Kuhbauch were the first to extract digital morphological features for use with identification model [1]. The technology found its some of the earliest applications in industrial agriculture, where the desire was to separate crop species from weed species, allowing for decreased use of pesticides [2].

J. Hemming and T. Rath developed a computer vision based weed identification system, in which special attention was given to the open field experiments. Eight different morphological features and three color features were calculated for each single object to build a joint feature space [3]. A membership function based on fuzzy logic approach was formed and used for the classification [3].

Another approach was an autonomous agricultural mobile robot for mechanical weed control in outdoor environments. The robot employs two vision systems: one gray-level vision system that is able to recognize the row structure formed by the crops and to guide the robot along the rows and a second, color based vision system that is able to identify a single crop among the crops [4].

For shape based leaf image retrieval Z. Wang, Z. Chi and D. Feng proposed an efficient two-stage approach for leaf image retrieval by using simple shape features including centroid-contour distance (CCD) curve, eccentricity and angle code histogram (ACH). In the first stage, the images that are dissimilar with the query image will be first filtered out by using eccentricity to reduce the search space, and fine retrieval will follow by using all three sets of features in the reduced search space in the second stage [5].

In 2007, Stephen Gang Wu, Forrest Shen Bao, Xu, Wang, Chang proposed a probabilistic neural network (PNN) with image and data processing techniques to implement a general purpose automated leaf recognition for plant classification [7].

III. PROPOSED SYSTEM

The proposed system works as follows. An image of the leaf interested is acquired through the mobile phone camera and is preprocessed to obtain the binary image of the leaf's contour. For several samples of known species, morphological features are extracted from the binary contour image. This information is used to train the algorithm by determining the median value of each feature for a given species. With more information samples, the algorithm becomes for effective at identifying a given species.

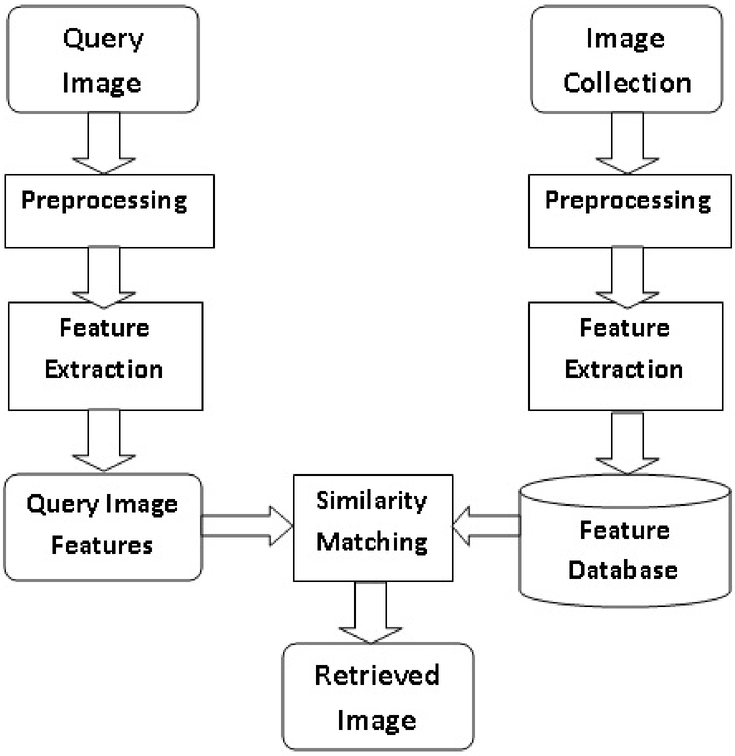

Figure 1. Proposed System

The proposed system consists of the following phases:

1. Query Image Acquisition
2. Preprocessing
3. Morphological Feature Extraction
4. Training and Matching

3.1. Query image acquisition

Currently, every mobile phone running the Android operating system is equipped with a high-quality digital camera with support for auto-focus and flash photography [10]. Input to the system is the image of a leaf of unknown species captured through the user's mobile phone camera. A certain degree of uniformity is assumed in the acquired images, i.e. the picture will be taken at a reasonable distance, in decent lighting, roughly normal to the surface, and against a background which provides sufficient contrast (white background). Application also provides the feature to rotate the captured image to the desired angle and to crop the image [Fig. 2.a] [Fig. 2.b]

International Journal of Modern Trends in Engineering and Research (IJMTER)
Volume 02, Issue 10, [October – 2015] ISSN (Online):2349–9745 ; ISSN (Print):2393-8161

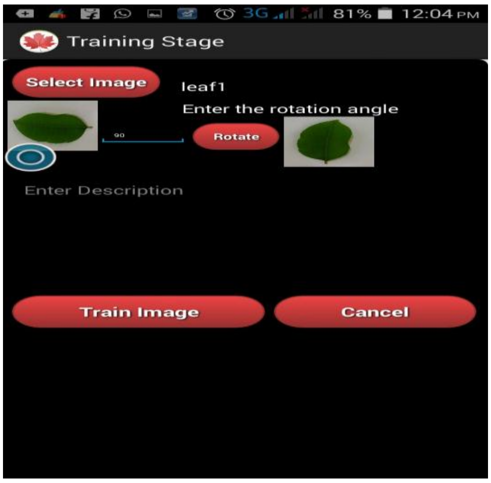

Figure 2.a. Image rotation

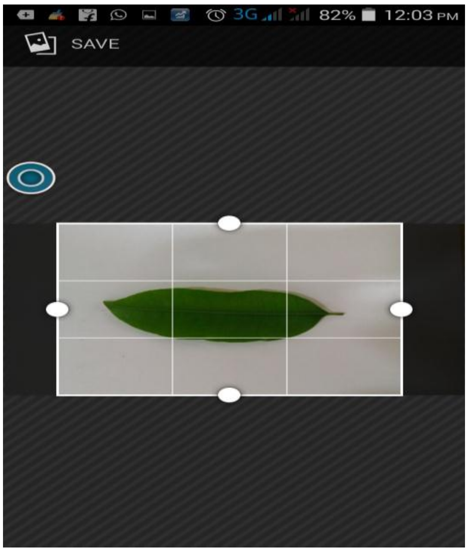

Figure. 2.b. Image crop

3.2. Preprocessing

To find the outline contour of the leaf, the acquired image is first preprocessed. In this process the dominant color of the leaf is obtained, to be used later in the leaf color matching phase. The acquired image is converted to a mutable bitmap image and is converted in to binary image using Otsu's method. Following this, the boundary of the image is obtained using sobel operator (kernel matrix of 3x3) [Fig. 3].

Figure 3. Sample leaf before and after preprocessing

3.3. Morphological feature extraction

Following review of literature with various combinations of digital morphological features, we decided to include the following features in our final algorithm: aspect ratio, rectangularity, form factor, sphericity and eccentricity. Each of the features is described below.

3.3.1. Aspect ratio (AR). The aspect ratio is the ratio between the maximum length D_{max} and the minimum length D_{min} of the minimum bounding rectangle (MBR) [Fig 4]

$$\mathbf{AR = D_{max} / D_{min}} \quad (1)$$

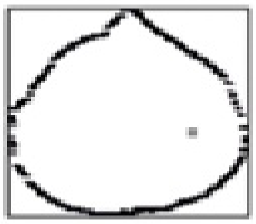

Figure 4. MBR

3.3.2. Rectangularity (R).): Rectangularity is defined as the ratio between the region-of-interest (ROI) area and the MBR area.

$$\mathbf{R = A_{ROI} / (D_{max} / D_{min})} \quad (2)$$

3.3.3. Form factor (F). Form factor is a well-known shape description characteristic given by

$$\mathbf{FF = 4\pi A_{ROI} / P^2_{ROI}} \quad (3)$$

3.3.4. Sphericity (S). Sphericity is the ratio of the radius of the incircle of the ROI (r_i) and the radius of the excircle of the ROI (r_c).

$$\mathbf{S = r_i / r_c} \quad (4)$$

3.3.5. Eccentricity (E). Eccentricity is the ratio of the length of the main inertia axis of the ROI (E_A) and the length of the minor inertia axis of the ROI (E_B).

$$\mathbf{E = E_A / E_B} \quad (5)$$

3.4. Training and matching

Five samples of each class were acquired for training the system. They were segmented from the background and their contours were extracted using Otsu's method and sobel operator respectively. Five scores for each of the features listed above were averaged to determine the feature vector for the class.

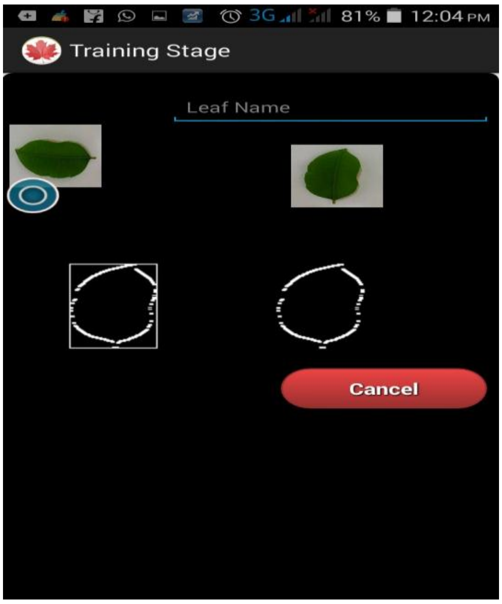

Figure 5. Training stage

Matching the query leaf image with the stored classes consists of 3 processes: matching by shape, color and morphological features. As explained above, the dominant color of the query leaf image will get extracted in the preprocessing phase and the image gets converted into a binary image. The binary contour of the query leaf image is obtained by performing sobel operator [Fig. 3]. The binary contour of the query leaf image is used to perform matching by shape. The dominant color of the query leaf image is then compared with the color of each classes stored in the system. At last, to determine the query leaf's correct match, its feature vector is computed by extracting the morphological features, in the similar manner as that of the training leaf's and its difference from each leaf class is calculated based on Euclidean distance method using the formula.

$$\text{Euclidean distance} = \sqrt{(a_2 - a_1)^2 + (b_2 - b_1)^2 \ldots + (n_2 - n_1)^2} \quad (6)$$

IV. ANDROID IMPLEMENTATION

The describe classification algorithm was implemented on Eclipse platform using the programming language JAVA. Algorithm was initially tested on the inbuilt emulator of eclipse and then later tested on a handset. Since Eclipse doesn't support many higher level functions on morphological features e.g. convex hull, moments of inertia etc. like MATLAB, they were not implemented in this algorithm. Pre-calculated training data for the trained classes are stored in the system and classification calculations are performed locally on the handset.

V. EXPERIMENTAL RESULTS

For the Android application field test, we tried to test 5 leafs from different classes. To keep statistical independence, a new set of test leafs were taken [Fig. 9]. Experimental results have been classified into 3 different tables, Table 1, Table 2, and Table 3 for shape, color and feature matching respectively.

Looking at Table 1 and 2, classification based on shape and color had high rates of successful classification for class 1, 2, 3 and 4, where class 5 had faired result. From Table 3, classification based on morphological features had successful matching with class 1 and 2, faired matching with classes 4 and 5 and rather inconsistent matching with class 3. The inconsistent behavior may be due to the shape similarity of the class, in which case larger number of training samples may provide improvement. Additionally the training samples were captured without the light of the mobile phones, in which case, if flash light is used to capture testing images, result for color based classification becomes inconsistent.

	Class 1	**Class 2**	**Class 3**	**Class 4**	**Class 5**
Class 1	**86 %**	83 %	82 %	82 %	64 %
Class 2	84 %	**85 %**	82 %	81 %	65 %
Class 3	84 %	82 %	**85 %**	81 %	77 %
Class 4	81 %	80 %	80 %	**82 %**	69 %
Class 5	84 %	83 %	84 %	80 %	**81 %**

Table 1. Result for Leaf class match (Shape)

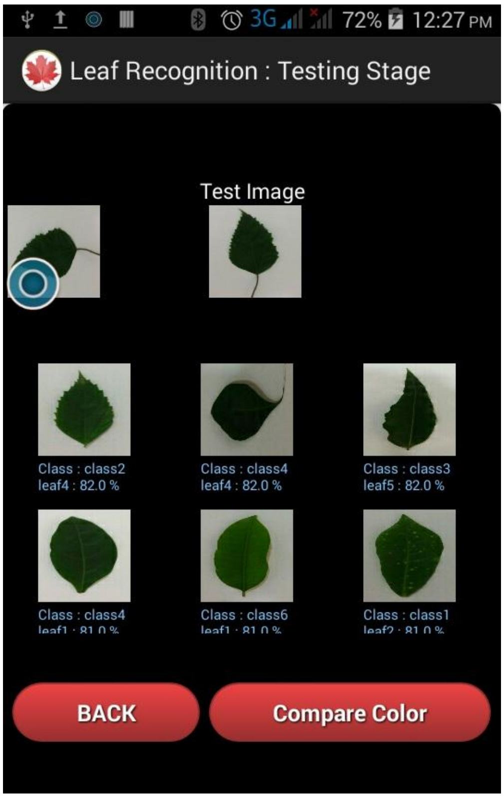

Figure 6. Shape compare

	Class 1	**Class 2**	**Class 3**	**Class 4**	**Class 5**
Class 1	**97 %**	85 %	93 %	89 %	61 %
Class 2	88 %	**97 %**	92 %	82 %	86 %
Class 3	73 %	87 %	**94 %**	53 %	75 %
Class 4	98 %	95 %	96 %	**99 %**	83 %
Class 5	87%	86 %	88 %	75 %	**91 %**

Table 2. Result for Leaf class match (Color)

International Journal of Modern Trends in Engineering and Research (IJMTER)
Volume 02, Issue 10, [October – 2015] ISSN (Online):2349–9745 ; ISSN (Print):2393-8161

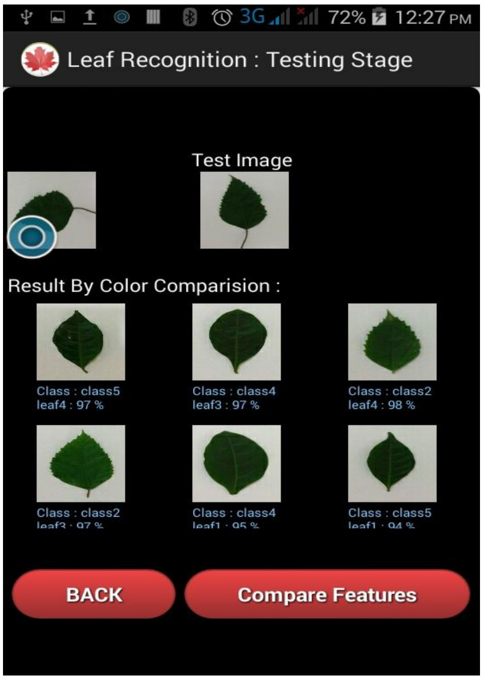

Figure 7. Color compare

	Class 1	Class 2	Class 3	Class 4	Class 5
Class 1	**127.93**	156.89	223.23	233.37	270.12
Class 2	60.27	**55.67**	68.46	66.24	71.99
Class 3	59.20	63.65	**97.25**	98.82	115.35
Class 4	57.43	54.50	73.85	**72.59**	81.78
Class 5	59.62	55.50	70.19	68.22	**74.63**

Table 3. Result for Leaf class match (Morphological features)

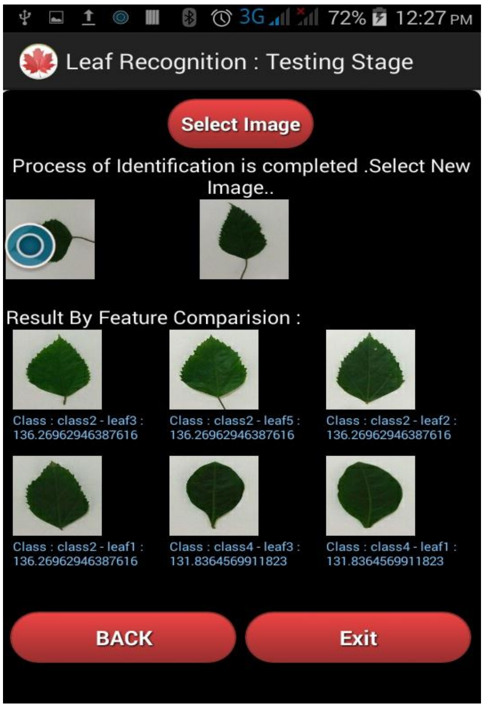

Figure 8. Feature compare

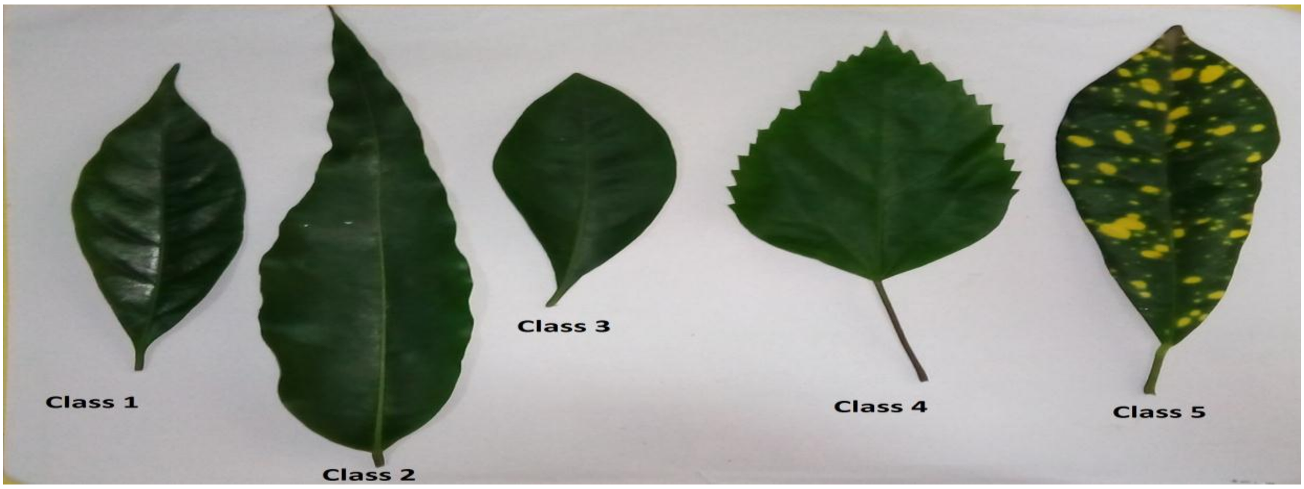

Figure 9. Leaf Classes

VI. CONCLUSION

The system offers significantly fair classification on the sample classes. However, to improve the accuracy of the algorithm, much large number of classes and samples for each is required, which leads to the limitation of storage and processing speed of the handset. To address this limitation, the application could be used in the cloud based environment where storage and classification is done over cloud application.

The application fairly classifies the sample classes, performance of the algorithm may degrade when attempting to classify samples from plant species that are related or similar in shape. The result of the color matching phase gets affected if the image for testing has been captured using flash light, since the training samples had been captured in natural day light without the use of flash light of handset. Lastly, the morphological classification performs average if compared with other methodology. However, with a large number of samples the performance can be improved.

The described mobile Android mobile application makes use of three classification phase, shape, color and morphological feature classification, to classify the plant species. Sobel operator, dominant color and Euclidean distance method is used to classify based on shape, color and morphological features respectively. The method proved quite robust and efficient under normal circumstances. With the increased number of training samples and classes, the application will be able to classify more plant species.

REFERENCES

[1] W. Petry and W. Kuhbauch, “Automated discrimination of weed species with quantitative image analysis”, Journal of Agronomy & Crop Science, vol. 163, pp. 345-351, 1989.
[2] R. D. Tillet, “Image analysis for agriculture processes: a review of potential opportunities”, Journal of Agricultural Engineering Research, vol. 12, pp. 247-258, 1991
[3] J. Hemming and T. Rath, “Computer-vision based weed identification under field conditions using controlled lightning”, Journal of Agricultural Engineering Research, vol. 78-3, pp. 233-243, March 2001.
[4] B. Astrand and A. Baeveldt, “An agricultural mobile robot with visionbased perception for mechanical weed control”, Autonomous Robots, vol. 13-1, pp. 21-35, July 2002.
[5] Z. Wang, Z. Chi and D. Feng, “Shape based leaf image retrieval”, IEEE Proceedings: Vision, Image, and Signal Processing, vol. 150-1, pp. 34-43, February 2003.
[6] J. Du, X. Wang and G. Zhang, “Leaf shape based plant species recognition”, Applied Mathematics and Computation, vol. 185-2, pp. 883-893, February 2007.
[7] S. Wu, F. Bao, E. Xu, Y. Wang and Q. Xiang, “A leaf recognition algorithm for plant classification using probabilistic neural network”, in Proceedings of 2007 IEEE International Symposium on Signal Processing and Information Technology, Giza, December 2007.

Volume 02, Issue 10, [October – 2015] **ISSN (Online):2349–9745 ; ISSN (Print):2393-8161**

[8] C. Im, H. Nishida and T. L. Kunil, "Recognizing plant species by leaf shapes – a case study of the acer family", in Proceedings of 1998 IEEE International Conference on Pattern Recognition, Brisbane, August 1998.
[9] David Knight, James Painter, Matthew Potter, "Automated Plant Leaf Classification for a Mobile Field Guide".
[10] Rahul Mehta, Nishchol Mishra, Sanjeev Sharma, "Color-Texture based Image Retrieval System", Journal of Computer Applications, vol. 24-No.5, June 2011.

