IDENTIFYING MEDICINAL PLANT LEAVES USING TEXTURES AND OPTIMAL COLOUR SPACES CHANNEL

C H Arun[1], and D Christopher Durairaj[2]

[1]Department of Computer Science, Nesamony Memorial Christian College, Marthandam-629165, India
[2]Department of Computer Science, Virudhunagar Hindu Nadars Senthikumara Nadar College, Virudhunagar-626001, India

E-mail: dass.arun@gmail.com

Abstract

This paper present an automated medicinal plant leaf identification system. The Colour Texture analysis of the leaves is done using the statistical, the Grey Tone Spatial Dependency Matrix (GTSDM) and the Local Binary Pattern (LBP) based features with 20 different color spaces (RGB, XYZ, CMY, YIQ, YUV, YCbCr, YES, U*V*W*, L*a*b*, L*u*v, lms, $l\alpha\beta$, $I_1I_2I_3$, HSV, HIS, IHLS, HIS, TSL, LSLM, and KLT). Classification of the medicinal plant is carried out with 70% of the dataset in training set and 30% in the test set. The classification performance is analysed with Stochastic Gradient Descent (SGD), k Nearest Neighbour(kNN), Support Vector Machines based on Radial basis function kernel(SVM-RBF), Linear Discriminant Analysis(LDA) and Quadratic Discriminant Analysis(QDA) classifiers. Results of classification on a dataset of 250 leaf images belonging to five different species of plants show the identification rate of 98.7 %. The results certainly show better identification due to the use of YUV, L*a*b* and HSV colour spaces.

Keywords: *colour spaces, texture features, plant identification, pattern recognition*

Abstrak

Makalah ini menyajikan sebuah tanaman obat sistem identifikasi daun otomatis. Analisis Warna Tekstur dari daun dilakukan dengan menggunakan statistik, Grey Tone Spatial Dependency Matrix (GTSDM) dan Pola Binary lokal (LBP) fitur berbasis dengan 20 ruang warna yang berbeda (RGB, XYZ, CMY, YIQ, YUV, YCbCr, YES, u * V * W *, L * a * b *, L * u * v, LMS, $l\alpha\beta$, $I_1I_2I_3$, HSV, HIS, IHLS, HIS, TSL, LSLM, dan KLT). Klasifikasi tanaman obat dilakukan dengan 70% dari dataset di set pelatihan dan 30% dalam tes set. Kinerja klasifikasi dianalisis dengan Stochastic Gradient Descent (SGD), k Tetangga terdekat (kNN), Dukungan Mesin Vector berdasarkan Radial fungsi dasar kernel (SVM-RBF), Linear Discriminant Analysis (LDA) dan kuadrat Analisis Diskriminan (QDa) pengklasifikasi. Hasil klasifikasi pada dataset dari 250 gambar daun milik lima spesies yang berbeda dari tanaman menunjukkan tingkat identifikasi 98,7%. Hasil tentu menunjukkan identifikasi yang lebih baik karena penggunaan YUV, L * a * b * dan ruang warna HSV.

Kata Kunci: *ruang warna, fitur tekstur, identifikasi tanaman, pengenalan pola*

1. Introduction

In the western ghats region of Asia, especially in Kanyakumari district, the southern most district of India, traditional home herbal remedies are readily available in most of the homes. These herbal plants are used in treating common sicknesses like common cold, diarrhoea and headache. The herbal practitioners, herbal medical industries and the even the common lay people have a good knowledge about using these herbal plants. They collect the needed herbal plants from the available sources and then use the plant leaves for the preparation of the medicine.

World has seen a steady interest and usage of traditional herbal remedies and herbal produces [1]. World Health Organization (WHO) reports that about 65-80% of the world's population in developing countries depend on plants for their primary healthcare due to poverty and lack of access to modern medicine [2].

Leaf is one of the major identifying features of a plant. Leaves are far being the only discriminant visual key between species but, due to the shape and size, they have the advantage to be easily observed, captured and described [3]. Though all the medicinal plant parts like roots, flowers and barks are used for medicinal purposes, the most common part with more medicinal value are the fresh leaves. Identification of the

respective medicinal plant leaves are normally done using plant taxonomic field guides by the Botanists. The Siddha medicinal practitioner (the traditional herbal physician) who mostly rely on medicinal plant leaves, and the households, depend on their knowledge gained and their experience in identifying the plant leaves. Though, this has been the regular practice for generations, modernisation has it is impact and most of the younger generations do not have enough knowledge in identifying the respective medicinal plant leaves. Wrong identification of these leaves can lead to more damage while treating the patients. As there is a need for precise and quick identification of the plant leaves, an automatic identification system would prove to be an effective solution.

Five medicinal plant leaves are considered in the present work and some of their medicinal uses are given below. A sample of the medicinal plant leaf dataset from different plant leaves is given in the Figure 1. (0) Desmodium gyrans - Antidote for snake poison, effective for heart diseases, rheumatic complaints, diabetes and skin ailments. (1) Butea monosperma - promotes diuresis & antihelminthic, treats leucorrhoea & diabetes. (2) Malpighia glabra - help lower blood sugar, increases collagen and elastin production, treats diarrhoea, dysentery, and liver problems. (3) Helicteres isora - help treat intestinal complaints, colic pains and flatulence. (4) Gymnema sylvestre - suppresses the sensation of sweet, anti-diabetic.

Automatic plant leaf identification is a difficult problem because there is often high intra-species variability, and low interspecies variation [4]. But many promising approaches have emerged. Image based plant leaf identification is done using morphological characteristics(shape, colour, texture, margin [4-7]). Automatic identification of medicinal plant leaves are done using texture, colour, statistical and Local Binary Pattern(LBP) features [8-10]. The identified plant leaves are used in food processing, medical, botanical gardening and cosmetic industry [11].

Colours are represented using different colour spaces. There is no particular colour space that is proving itself best for all the colour images. The use of colour increases the performances of the standard grey level texture analysis techniques [12,13]. To exploit the advantage of the colour component of the medicinal plant leaves which are mostly green in colour, identifying medicinal plant leaves using colour shall be considered in this research work. Though colour is not stable during the growth period of the leaves, colour space usage will certainly be providing more information in better identification. Selection of
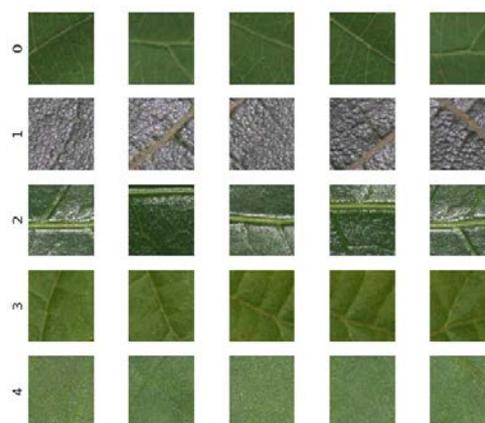

Figure 1. Leaf images of Desmodium gyrans(row 0), Butea monosperma(row 1), Malpighia glabra(row 2), Helicteres isora(row 3)and Gymnema sylvestre(row 4)

the suitable colour space for the particular identification depends on the specific image [14]. Different colour component based image processing produces reliable and accurate results [15] and improves classification performance [16]. Several approaches used in automatic identification of plant leaves use HSV colour space, which results in better performance [17].

The decorrelated colour space L*a*b* is said to perform best in colour transfer algorithms which yields better quantification results in foods with curved surfaces [18]. Y CbCr colour model outperforms other specified models RGB, YIQ, YCbCr, HSV and HSI in terms of objective quality assessment for Colour Image Fusion [19], YUV colour space is used to solve parameter optimisation of blind colour image fusion [20]. The hybrid colour space RCrQ possesses complementary characteristics and enhances discriminative power for face recognition [21].

The objective of the present work is to compute the colour texture features using various colour spaces based on greylevel, the Grey Tone Spatial Dependency Matrix (GTSDM) and the LBP operators and to find out the suitable colour space in identifying the medicinal plant leaf automatically.

Suitable image classification approaches can be used for the identification process. The rest of the paper is organized as follows: Methodology of the work is discussed in section 2 which includes the details of texture feature extraction, colour transform, classification and the description of the dataset. Design and implementation of the work are discussed in section 2. The results and discussion are given in section 3, which is followed by conclusion in section 4.

Algorithm 1:

Overview

1: Get image L from the leaf dataset.

2: Apply the colour transform in L producing the transformed image T.

3: Extract the features f_i from the transformed image T (where i = 1 to 12).

4: Store the feature values f_i in the database.

5: Repeat the steps 1 to 4 for all training images.

6: Read a test image I from the dataset.

7: Extract the features t_i from the test image I (where i= 1 to 12)

8: Apply the classifiers using the features f_i and t_i.

9: Display the classified output.

10: Calculate the classification accuracy.

11: Repeat the stages 6 to 10 for all testing images.

end

Figure 2. Block diagram for Classification

2. Methods

The general approach adopted for identifying medisinal plant leaves are explained here. The different stages like texture feature extraction, colour transform and classification algorithms are presented in detail. Primary dataset is elaborated. The Algorithm 1 shows the overview of procedure to obtain the classified medicinal plant leaf.

Texture Feature Extraction

Texture features in identifying medicinal plant leaves can be computed using first-order statistical moments (mean, variance and skewness) and second-order statistical moment [Grey Tone Spatial Dependency Matrix-GTSDM (energy, homogeneity, dissimilarity, entropy, correlation and contrast) and Local Binary Pattern-LBP (mean and standard deviation)] operators[9]. First-order grey level statistical features are the simplest texture descriptors. GTSDM or GLCM are the second-order statistical features which are more sensitive and more intuitive than first-order features. Local Binary Pattern (LBP) operators are known for its invariance to local gray scale variations, monotonic photometric changes and its high descriptive power.

Colour Transform

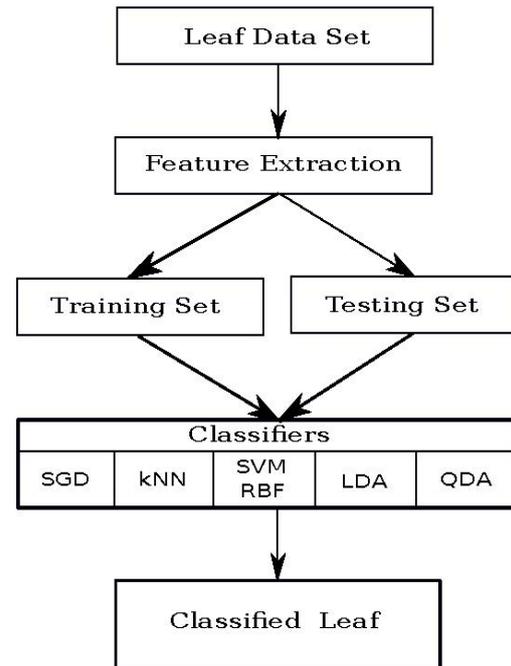The colour space suitable for medicinal plant leaf identification has to be found using experimental testing, as in most cases, the best established colour space is not suiting to all types of dataset [15]. Out of the various existing colour spaces, twenty colour spaces are considered. The colour spaces considered for our experimental testing are categorized into five colour space families [22]. The transformation equations for all the considered colour spaces are given in their respective citations. The colour spaces are, primary colour spaces: RGB, XYZ [23] and CMY [24]; luminance-Chrominance spaces: YIQ[23], YUV [23], Y CbCr [21], YES [25], U*V *W* [26], L*a*b* [27], L*u*v* [27], lms [28] and l [28]; independent axis space: $I_1I_2I_3$ [29]; the perceptual spaces: HSV [18], HIS [30], IHLS [31], HIS [32] and TSL [33] and the other colour spaces: LSLM [27] and KLT [34].

Classification

Texture based image classification involves deciding the most pertinent texture category of the observed image [35]. Classification here would mean the machine language classification of the different classes and not the linnaean taxonomic system of plant classification [36]. When the prior knowledge of the established classes are available and the texture features are extracted, the given image could be classified to the appropriate class.

The texture class i consisting of a set of n images can be represented as

$$T_i = \{t_{i,1}, t_{i,2}, \dots, t_{i,n}\} \qquad (1)$$

Where $t_{i,j}$ is the member image.

Various classifiers used in this work are Stochastic Gradient Descent (SGD) [37,38], k Nearest Neighbour (kNN) [39,40], Support Vector Machines based on Radial basis function kernel(RBF) [41], Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) [42]. SGD is an approach to discriminative learning of linear classifiers under convex loss functions. kNN computes the distance between a test point and all points in the training set. SVM are supervisor learning methods which are similar to SGDs and these are effective in high dimensional feature spaces. Linear Discriminant Analysis (LDA) is a classifier with a linear decision boundary, generated by fitting class conditional densities to the data. Quadratic Discriminant Analysis (QDA) is a classifier with a quadratic decision boundary, generated by fitting class conditional densities to the data. Parameter tuning is of importance while dealing with certain classifiers. The k value in the kNN classifier ranges from 1 to a maximum of the square root of the number of instances [43]. Choosing a lower k value will have a greater influence of noise on the result while a higher k value will increase the computation time. Moreover the value of k is chosen to be a odd value to lessen the computation time [44]. The parameters of SVM with a Gaussian radial basis function (RBF) kernel are C and γ the parameter C controls the influence of each individual support vector. A low bias and high variance is obtained from larger C and vice versa. The parameter handles non-linear classification. A higher gamma value will give a high bias, low variance and vice versa. The standard and optimal method to choose the optimal parameters C and γ is a Grid Search (GS) [45].

The classifier places all of the N training images in a vector space with N dimensions, based on the features extracted from the image. When a new uncategorized test image is placed in that vector space, the best suited image is found using the pairwise-distance function. The pairwise-distance between two images are calculated as given in the equation (2) where d1 and d2 are the two images. When the distance between d1 and d2 are the smallest, the most suitable image is obtained.

$$distance(d1, d2) = 1/L \qquad (2)$$

Where L is the number of matching images.

The test image is estimated to belong to a specific category, after generation of its features, when the majority of its neighbors are also in the same category.

Algorithm 2:
Functional

1: Let L be the leaf image, where $L = LIj$ and $j = 1, 2 \cdots p$

2: Form the transformed image T using the colour transform for the leaf image L, where $T = CTj$ and $j = 1, 2 \cdots p$

3: Calculate all the following features for each channel of the transformed matrix T.

    a) Grey features - mean($\mu 0$), variance($\mu 2$), standard deviation(σ) and skewness($\mu 3$)

    b) Grey Tone Spatial Dependency matrix(GTSDM) - energy(ζ), homogeneity(η), dissimilarity(ξ), entropy(E), correlation(ρ) and contrast(λ)

    c) Local Binary Pattern(LBP) - mean(μL) and standard deviation(σL).

4: Insert calculated features in the feature vector, $f1$ for each channels of the transformed matrix T.

5: Repeat the stages 1 to 4 for all the training set images

6: Read the test image I from the dataset

7: Form the transformed image K for the test image I, where $K = CTj$ and $j = 1, 2 \cdots p$

8: Calculate all the features listed in step 3 for each channel of the transformed image K.

9: Insert the calculated feature in the feature vector $f2$ for the test image.

10: Apply the classifiers SGD, kNN, SVM-RBF, LDA and QDA separately using the feature set $f1$ and $f2$.

11: Repeat the stages 6 to 10 and display the classified output

12: Calculate the classification accuracy

end

Datasets

The medicinal plant leaves are collected around the Western Ghats region of Kanyakumari forests. The digital images are obtained using the Canon EOS 1000D digital camera on the abaxial portions of the medicinal plant leaves in the uncompressed JPEG format with the dimension, 3420 x 4320 x 3, in a closed environment to maintain constant illumination. As the leaf characteristics, especially in terms of the colour, vary widely from its tender stage to the mature stage, the proposed algorithm is restricted for images of mature leaves of a
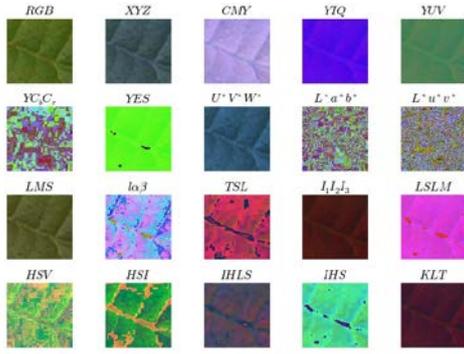
Figure 3. A leaf sample image in all considered Colour Spaces

TABLE 1
FIRST ORDER STATISTICAL SAMPLE
FEATURE VALUES OF VARIOUS CLASSES

Mean	Variant	Std.Dev	Skewness	Class
133.1596	545.1827	23.3491	0.47	0
58.1753	525.6315	22.9267	3.1372	1
53.3689	713.9928	26.7206	-0.469	2
68.3939	452.3641	21.2689	0.0202	3
60.5693	221.0215	14.8668	0.2602	4

TABLE 3
CLASSIFICATION RESULT BASED ON DIFFERENT
COLOUR SPACES

Colour Space	SGD	kNN	SVM RBF	LDA	QDA
RGB-B	94.7	96.0	96.0	93.3	96.0
XYZ-Y	94.7	93.3	94.7	90.7	96.0
CMY-Y	94.7	96.0	96.0	93.3	96.0
YIQ-Y	94.7	94.7	94.7	90.7	96.0
YUV-U	96.0	93.3.	93.3	92.0	98.7
YCbCr-Cb	92.0	92.0	94.7	93.3	96.0
YES-E	94.7	86.7	92.0	89.3	97.3
U*V*W*-V*	94.7	94.7	94.7	90.7	96.0
L*a*b*-a*	97.3	93.3	98.7	93.3	94.7
L*u*v*-L*	96.0	88.0	93.3	85.3	94.7
LMS-S	94.7	96.0	96.0	93.3	96.0
$l\alpha\beta$-l	94.7	92.0	93.3	89.3	97.3
$I_1I_2I_3$-I_1	94.7	93.3	96.0	90.7	96.0
HSV-H	96.0	97.3	98.7	92.0	96.0
HIS-I	94.7	93.3	96.0	90.7	96.0
IHLS-H	92.0	92.0	96.0	93.3	96.0
IHS-I	94.7	93.3	96.0	90.7	96.0
TSL-T	92.0	90.7	96.0	90.7	96.0
LSLM-LM	92.0	92.0	89.3	88.0	97.3
KLT-K	94.7	93.3	96.0	90.7	96.0

plant. The working dataset is formed by dividing the images into a size of 50 x 50 x 3 and forming a set total of 250 images, 50 images each from 5 medicinal plant leaves. The training and the testing sets are formed from the image datasets.

Design and Implementation

The implementation of the proposed algorithm is done using Python as the programming language in Linux platform along with the open-source libraries namely numpy, scipy, scikit-learn, colorsys and grapefruit. The various stages of the algorithm developed in identifying the medicinal plant leaves using colour textures are as given in Algorithm 2 and are explained here.

The medicinal leaf image from our dataset is available as an m x n x p colour image L with p colour channels in the form of equation (3)

$$LI_k = [c_{ij}] \qquad (3)$$

where i = 1,2,…m, j = 1, 2,… n, k = 1,2…p. p represents the number colour channels of the image.

The different colour space transformation, CTk in equation(4), is calculated from the colour channels LIk using the corresponding colour transform equations given in the respective citations of section 2. A sample of all the considered colour space leaf images are given in the Figure 3

$$CT_k = [t_{ij}] \qquad (4)$$

where i = 1, 2,… m,j = 1, 2… n, k = 1, 2… p. p represents the number colour channels of the image.

The statistical features using mean ($\mu 0$), variance ($\mu 2$), standard deviation (σ) and skewness ($\mu 3$) are calculated directly from the transformation CTk: The co-occurrence matrix and the haralick features energy (ζ), homogeneity (η), dissimilarity (ξ), entropy (ϵ), correlation (ρ) and contrast (λ), are calculated from the selected channel of the transformation CTk [35]. The feature values are used for classification.

The image CTk is used to calculate the local binary pattern operator feature using 8 neighbors and a unit radius for computation of the pattern. From the computed LBP histogram image, the mean of the histogram (μL) and standard deviation (σL) are calculated and used as a feature.

The general procedure of classification shown in Figure 2 is followed with the dataset. After extracting the features $\mu 0$, $\mu 2$, σ, $\mu 3$, ζ, η, ξ, ϵ, λ, ρ, μL and σL for the data set, which can be divided into two sets, the classification is done in two stages: training and testing. The 70% of dataset is considered as training samples and the remaining 30% as testing images. The medicinal plant leaves can be classified using the trained

TABLE 2
SECOND ORDER GTSDM, LBP SAMPLE FEATURE VALUES OF VARIOUS CLASSES

| GTSDM Energy	GTSDM Homogeneity	GTSDM Dissimilarity	GTSDM Entropy	GTSDM Correlation	GTSDM Contrast	LBP Mean	LBP SD	Class
0.077	0.3624	2.2302	5.4448	0.8678	8.7101	119.0731	84.315	0
0.2847	0.6988	0.9385	3.3051	0.8725	4.7482	108.5133	91.9709	1
0.2775	0.7737	0.4689	2.89	0.8868	0.5507	113.5907	88.2488	2
0.2312	0.6958	0.6588	3.2908	0.8587	0.9144	113.0331	86.4206	3
0.4263	0.88	0.2412	2.086	0.8727	0.2476	97.5045	90.3729	4
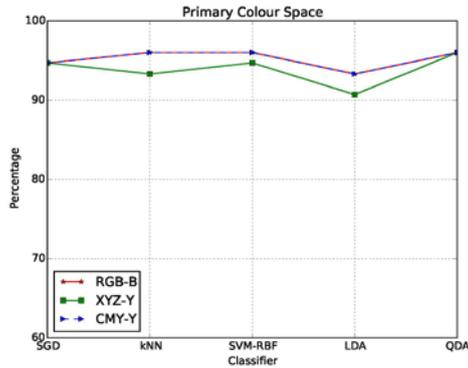
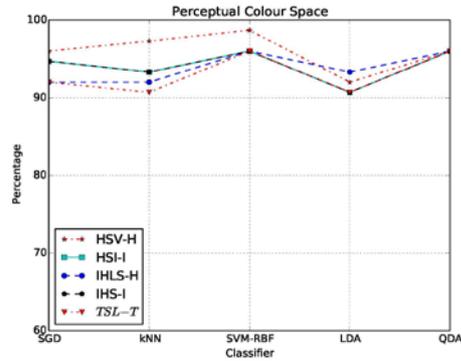
Figure 4. Primary Colour Spaces
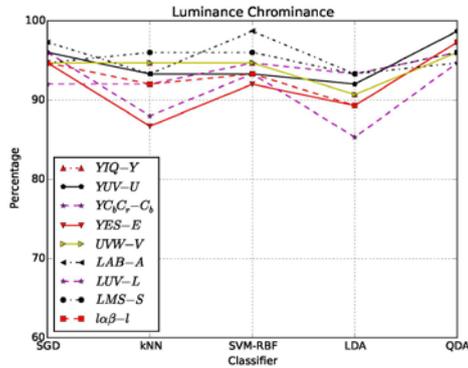

Figure 6. Perceptual Colour Spaces

Figure 5. Luminance Chrominance Colour Spaces
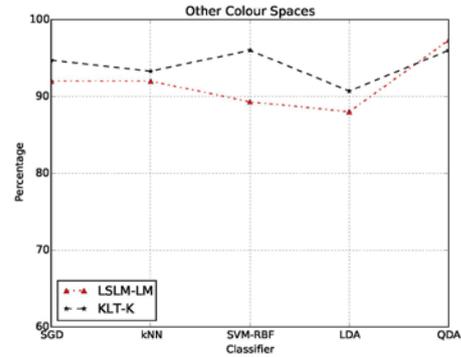

Figure 7. Other Colour Spaces

samples. The classification is performed by the different classifiers SGD, kNN, SVM-RBF, LDA and QDA. The parameter k for the kNN classifier in our work is chosen to be 3 which is an odd k value and is between the specified range. The C and γ parameters in SVM-RBF classifier are selected using Grid Search function as C = 1 and γ = 2. The parameters for SGD classifier considered are α = 0.001 and the number of iterations = 100.

3. Results and Analysis

The different colour spaces play a key role in this process of classification. The leaf images taken from the dataset are transformed into different colour spaces. For each colour space is considered with different colour channels. The feature values are computed for all colour channels. Sample feature values of various classes are given in the Table 1. The classification is done using the feature values of each colour channel.

Performance Analysis

The recognition percentage obtained for different colour spaces varies on the colour channels. The best classification result for different colour spaces shown here is considered based on the best classification result channel. The detailed classification result obtained for the colour channel of various colour spaces are shown in Table 3.
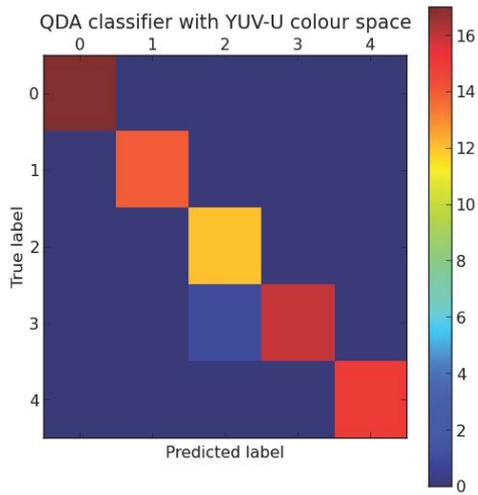
Figure 8. Confusion Matrix for QDA classifier with YUV-U colour space

TABLE 4
CLASSIFICATION REPORT ON THE QDA
CLASSIFIER FOR YUV-U COLOUR SPACE

Leaf	Precision	Recall	f1-score	Support
Leaf 0	1.00	1.00	1.00	17
Leaf 1	1.00	1.00	1.00	14
Leaf 2	0.92	1.00	0.96	12
Leaf 3	1.00	0.94	0.97	17
Leaf 4	1.00	1.00	1.00	15
Avg/total	0.99	0.99	0.99	75

Primary Colour Spaces: It is observed that the blue channel produced good result of 96% than the red and green channels of RGB colour space for kNN, SVM-RBF and QDA classifiers. In the XYZ colour space, the Y channel gives 96%, which is the best accuracy with X, Y and Z channels for QDA classifier. The CMY colour space performs best with 96% on the Y channel for the three classifiers kNN, SVM-RBF and QDA in yellow channel. The consolidated primary family colour spaces performances are given in Figure 4 The RGB-B and CMY-Y are both giving the best performance of 96 % with kNN, SVM-RBF and QDA.

Luminance Chrominance Colour Spaces: The YIQ colour space presents the best recognition rate of 96% for QDA classifier in Y channel. For YUV colour space, U channel gives very good performance than all other colour space and channels with 98.7% for QDA classifier. The Y CbCr colour space performs the result of 96% in QDA classifier with Cb channel. In YES colour space E channel presents the 97.3% recognition rate for QDA classifer. The U*V*W* colour space produced 96% in V* channel for QDA classifier. The colour space L*a*b* very well performed

TABLE 5
COMPARISON WITH OTHER METHODS

Work	Method	Dataset (classes)	Accuracy
[47]	Fourier Moments	32 (flavia)	62.00%
[48]	1-NN	60	82.33%
[47]	PNN-PCNN	32 (flavia)	91.00%
[9]	I+GTDSDM+LBP	5	94.70%
[46]	Neighborhood	30	95.83%
[47]	SVM-BDT	32 (flavia)	96.00%
Proposed	I+GTDSDM+LBP+ 20 colour spaces	5	98.70%

with 98.7% in SVM-RBF classifier. The channel a* projected the highest recognition rate of 98.7% in L*a*b*. The L*u*v* colour space manages 96% in L* channel for the SGD classifier. The S channel in LMS colour space proves with the performance of 96% in three classifiers. The L*u*v* colour space provides 97.3% in L* channel for QDA classifier. The consolidated Luminance Chrominance colour spaces performances are given in Figure 5. The YUV-U with QDA, L*a*b*-a* with SVM-RBF gives the best performance of 98.7%.

Independent Colour Space: The $I_1I_2I_3$ colour space gives 96% in I_1 channel for SVM-RBF and QDA classifiers.

Perceptual Colour Spaces: The HSV colour space is the best performing colour space for all classifiers except LDA. For SVM-RBF classifier, HSV-H presents the maximum performance of 98.7%. The colour spaces HSI-I, IHLS-H, IHS-I and TSL-T are best at 96% with SVM-RBF and QDA. The performances of Perceptual family of colour spaces are given in Fig 6. The HSV-H yields the best performance of 98.7% with SVM-RBF.

Other Colour Spaces: LSLM-LM presents the maximum performance of 97.3%. The LSLM-LM colour space performs best at 97.3% with QDA. The KLT-K colour space performs best at 96.0% with SVM-RBF and QDA.The consolidated performances are given in Figure 7.

When the classifier performance is considered, the QDA performs the best in all colour spaces. The SVM-RBF comes next to QDA in the classification rate. For L*a*b* and HSV, the SVM-RBF produced the best result of 98.7%. For YUV, the QDA produced the best result of 98.7%.

Precision Recall

The precision, recall, f1-support and support are calculated for one of the best performing classifiers QDA, which gives a performance of 98.7% in identification of the leaf and is given in the

Table 4. The confusion matrix for the classifier QDA with the YUV-U channel is given in Figure

Comparison with Related Works

The proposed work is done with the texture features and the colour transforms. The leaf shape, venation are not considered here. The work presented here produced best results than the previous [9] work done with the same database. The comparative results of the different methods are listed out in Table 5. The computations were done with gray scale operations. The neighborhood method [46] operated with 30 plants gives 95.83% accuracy. The Fourier moments [47] with Flavia database gives 62.00%. The PNN-PCNN [47] method applied with Flavia database produces 91.00% accuracy. The SVM-BDT [47] applied with Flavia database presents 96.00% performance. The 82.33% accuracy is found for 1-NN method [48] with 60 species. The texture feature method [9] for 5 species of medicinal plants produced the performance accurac of 94.7%. The proposed method applied with the individual components of the 20 colour spaces considered, yielded the better accuracy of 98.7%.

4. Conclusion

In this paper, the method to identify the medicinal plants automatically with the help of colour textures calculating the statistical, GTSDM and LBP texture features is proposed. Though a single shade of colour for the considered mature plant leaf is not stable during the growth period of the leaves, usage of colour space is providing more information that helps for better identification. The limitation of this method in identification is that this method works only with the matured leaves of the plant and that it demands more computational time. The colour space YUV-U with QDA, L*a*b*-a* with SVM-RBF, and HSV-H with SVM-RBF give the finest performance of 98.7% among the twenty colour spaces considered and hence it is concluded that the suitable colour space for identifying medicinal plant leaves are YUV with U channel, L*a*b* with a* channel and HSV with H channel. The QDA classifier performs best in most of the different colour spaces considered.

References

[1] "Homeopathic and herbal remedies-us-report." Mintel, April 2011.

[2] J. B. Calixto, "Twenty-five years of research on medicinal plants in latin america: a personal view," Journal of ethnopharmacology, vol. 100, no. 1, pp. 131–134, 2005.

[3] A. Joly, H. Goe¨au, P. Bonnet, V. Bakic´, J. Barbe, S. Selmi, I. Yahiaoui, J. Carre´, E. Mouysset, J.-F. Molino et al., "Interactive plant identification based on social image data," Ecological Informatics, vol. 23, pp. 22–34, 2014.

[4] T. Beghin, J. S. Cope, P. Remagnino, and S. Barman, "Shape and texture based plant leaf classification," in International Conference on Advanced Concepts for Intelligent Vision Systems. Springer, 2010, pp. 345–353.

[5] H. Kebapci, B. A. Yanikoglu, and G. B. U¨nal, "Plant image retrieval using color, shape and texture features," Comput. J., vol. 54, no. 9, pp. 1475–1490, 2011.

[6] J. S. Cope, P. Remagnino, S. Barman, and P. Wilkin, "Plant texture classification using gabor co-occurrences," in International Symposium on Visual Computing. Springer, 2010, pp. 669–677.

[7] J. S. Cope, D. P. A. Corney, J. Y. Clark, P. Remagnino, and P. Wilkin, "Plant species identification using digital morphometrics: A review," Expert Syst. Appl., vol. 39, no. 8, pp. 7562–7573, 2012.

[8] Y. Herdiyeni and N. Wahyuni, "Mobile application for indonesian medicinal plants identification using fuzzy local binary pattern and fuzzy color histogram," in Advanced Computer Science and Information Systems (ICACSIS), 2012 International Conference on, Dec 2012, pp. 301–306.

[9] C. H. Arun, W. R. Sam Emmanuel, and D. Christopher Durairaj, "Texture feature extraction for identification of medicinal plants and comparison of different classifiers," International Journal of Computer Applications, vol. 62, no. 12, pp. 1–9, January 2013, published by Foundation of Computer Science, New York, USA.

[10] Y. Herdiyeni, E. Nurfadhilah, E. A. Zuhud, E. K. Damayanti, K. Arai, and H. Okumura, "Article: A computer aided system for tropical leaf medicinal plant identification," International Journal on Advanced Science, Engineering and Information Technology, vol. 3, no. 1, pp. 23–27, 05 2013.

[11] B. Satyabama, "Content based leaf image retrieval (cblir) using shape, color and texture features," Indian Journal of Computer Science and Engineering (IJCSE), vol. 2, no. 2, pp. 202–211, Apr - May 2011.

[12] A. Drimbarean and P. F. Whelan, "Colour texture analysis: A compara- tive study," Vision System Laboratory, School of

Electronic Engineering, Dublin City University, 2000.

[13] P. F. Whelan and O. Ghita, "Colour texture analysis," 2008.

[14] K. Y. Lin, J. H. Wu, and L. H. Xu, "A Survey On Color Image Segmentation Techniques," Journal of Image And Graphics, vol. 10, no. 1, pp. 1–10, 2005.

[15] E. Reinhard and T. Pouli, "Colour spaces for colour transfer," in Proceedings of the Third International Conference on Computational Color Imaging, ser. CCIW'11. Berlin, Heidelberg: Springer-Verlag, 2011, pp. 1–15.

[16] A. Drimbarean and P. F. Whelan, "Experiments in colour texture analysis," Pattern Recogn. Lett., vol. 22, no. 10, pp. 1161–1167, Aug. 2001.

[17] A. Ehsanirad and Y. Sharath Kumar, "Leaf recognition for plant clas- sification using glcm and pca methods," Oriental Journal of Computer Science and Technology, vol. 3, no. 1, pp. 31–36, 2010.

[18] F. Mendoza, P. Dejmek, and J. Aguilera, "Calibrated color measurements of fruits and vegetables using image analysis," Postharvest Biology and Technology, vol. 41, pp. 285–295, 2006.

[19] W. Rattanapitak and S. Udomhunsakul, "Comparative efficiency of color models for multi-focus color image fusion," Hong Kong, 2010.

[20] Y. Niu and L. Shen, "The optimal multi-objective optimization using pso in blind color image fusion," in 2007 International conference on multimedia and ubiquitous engineering (MUE'07). IEEE, 2007, pp. 970–975.

[21] Z. Liu and C. Liu, "Fusion of color, local spatial and global frequency information for face recognition." Pattern Recognition, vol. 43, no. 8, pp. 2882–2890, 2010.

[22] L. Busin, J. Shi, N. Vandenbroucke, and L. Macaire, "Color space selection for color image segmentation by spectral clustering," in Signal and Image Processing Applications (ICSIPA), 2009 IEEE International Conference on, Nov 2009, pp. 262–267.

[23] P. Shih and C. Liu, "Comparative assessment of content-based face image retrieval in different color spaces." IJPRAI, vol. 19, no. 7, pp. 873–893, 2005.

[24] P. Bourke, "Converting between rgb and cmy, yiq, yuv," Texture Colour, pp. 1–7, Feb 1994.

[25] E. Saber and A. M. Tekalp, "Frontal-view face detection and facial fea- ture extraction using color, shape and symmetry based cost functions," Pattern Recognition Letters, vol. 19, no. 8, pp. 669 – 680, 1998.

[26] J. Ladd and J. Pinney, "Empirical relationships with the munsell value scale," Proceedings of the Institute of Radio Engineers, vol. 43, no. 9, p. 1137, September 1955.

[27] P. Colantoni et al., "Color space transformations," see http://www.radugar-yazan. ru/files/doc/colorspacetrans-form95. pdf, 2004.

[28] C.-M. Wang and Y.-H. Huang, "A novel color transfer algorithm for image sequences," Journal of Information Science and Engineering, vol. 20, no. 6, pp. 1039–1056, 2004.

[29] Y. Ohta, Knowledge-based interpretation of outdoor natural color scenes. Morgan Kaufmann, 1985, vol. 4.

[30] W. ladys law Skarbek and A. Koschan, "Colour image segmentation a survey," IEEE Transactions on circuits and systems for Video Technol- ogy, vol. 14, 1994.

[31] J. Angulo and J. Serra, "Mathematical morphology in color spaces ap- plied to the analysis of cartographic images," Proceedings of GEOPRO, vol. 3, pp. 59–66, 2003.

[32] T.-M. Tu, S.-C. Su, H.-C. Shyu, and P. S. Huang, "A new look at ihs-like image fusion methods," Information Fusion, vol. 2, no. 3, pp. 177–186, 2001.

[33] D. Butler, S. Sridharan, and V. Chandran, "Chromatic colour spaces for skin detection using gmms," in Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on, vol. 4. IEEE, 2002, pp. IV–3620.

[34] R. Kountchev and R. Kountcheva, "Image color space transform with enhanced klt," in New Advances in Intelligent Decision Technologies, ser. Studies in Computational Intelligence, K. Nakamatsu, G. Phillips-Wren, L. Jain, and R. Howlett, Eds. Springer Berlin Heidelberg, 2009, vol. 199, pp. 171–182.

[35] R. M. Haralick, K. S. Shanmugam, and I. Dinstein, "Textural features for image classification." IEEE Transactions on Systems, Man, and Cybernetics, vol. 3, no. 6, pp. 610–621, 1973.

[36] J. Comstock, An introduction to the study of botany: including a treatise on vegetable physiology, and descriptions of the most common plants in the middle and northern states, ser. Harvard science and math textbooks preservation microfilm project. Robinson, Pratt & Co., 1837.

[37] Y. Tsuruoka, J. Tsujii, and S. Ananiadou, "Stochastic gradient descent training for l1-

regularized log-linear models with cumulative penalty," in Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 1 - Volume 1, ser. ACL '09. Stroudsburg, PA, USA: Association for Computational Linguistics, 2009, pp. 477–485.

[38] Y. Lin, F. Lv, S. Zhu, M. Yang, T. Cour, K. Yu, L. Cao, and T. Huang, "Large-scale image classification: fast feature extraction and svm train- ing," in Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. IEEE, 2011, pp. 1689–1696.

[39] A. Bosch, A. Zisserman, and X. Mun˜oz, "Scene classification using a hybrid generative/discriminative approach," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, no. 4, pp. 712–727, 2008.

[40] G. Amato and F. Falchi, "kNN based image classification relying on local feature similarity," in Proceedings of the Third International Conference on SImilarity Search and APplications. ACM, 2010, pp. 101–108.

[41] Y.-W. Chang, C.-J. Hsieh, K.-W. Chang, M. Ringgaard, and C.-J. Lin, "Training and testing low-degree polynomial data mappings via linear svm," Journal of Machine Learning Research, vol. 11, pp. 1471–1490, 2010.

[42] T. Hastie, R. Tibshirani, and J. H. Friedman, The elements of statistical learning: data mining, inference, and prediction: with 200 full-color illustrations. New York: Springer-Verlag, 2008.

[43] R. O. Duda, P. E. Hart, and D. G. Stork, Pattern classification. John Wiley & Sons, 2012.

[44] A. B. Hassanat, M. A. Abbadi, G. A. Altarawneh, and A. A. Alhasanat, "Solving the problem of the k parameter in the knn classifier using an ensemble learning approach," arXiv preprint arXiv:1409.0919, 2014.

[45] M. Habshah, "A hybrid technique for selecting support vector regression parameters based on a practical selection method and grid search procedure," 2016.

[46] L. Jiming, "A new plant leaf classification method based on neigh- borhood rough set," Advances in information Sciences and Service Sciences(AISS), vol. 4, no. 1, pp. 116–123, 1 2012.

[47] S. G. Krishna Singh, Indra Gupta, "SVM-BDT PNN and fourier moment technique for classification of leaf shape," International Journal of Signal Processing, Image Processing and Pattern Recognition, vol. 3, no. 4, pp. 67–78, December 2010.

[48] C.-L. Lee and S.-Y. Chen, "Classification of leaf images," International Journal of Imaging Systems and Technology, vol. 16, no. 1, pp. 15–23, 2006.