A Review on Classification Methods for Plants Leaves Recognition

Khaled Suwais[1]

Faculty of Computer Studies
Arab Open University (AOU)
Riyadh, Saudi Arabia

Khattab Alheeti[2], Duaa Al_Dosary[3]

College of Computer Sciences and Information Technology
University of Anbar
Anbar, Iraq

Abstract—**Plants leaves recognition is an important scientific field that is concerned of recognizing leaves using image processing techniques. Several methods are presented using different algorithms to achieve the highest possible accuracy. This paper provides an analytical survey of various methods used in image processing for the recognition of plants through their leaves. These methods help in extracting useful information for botanists to utilize the medicinal properties of these leaves, or for any other agricultural and environmental purposes. We also provide insights and a complete review of different techniques used by researchers that consider different features and classifiers. These features and classifiers are studied in term of their capabilities in enhancing the accuracy ratios of the classification methods. Our analysis shows that both of the Support Victor Machines (SVM) and the Convolutional Neural Network (CNN) are positively dominant among other methods in term of accuracy.**

Keywords—Leaf recognition; feature extraction; leaf features; classifiers; image processing

I. INTRODUCTION

An important role introduced by plants to maintain the ecological balance of the earth by providing us with breathing, shelter, fuel and medicine. Pattern recognition and image processing techniques are exploited by using plant images to build plant lists for the conservation and preservation of existing classes of the plant [1]. Leaves are considered convenient for the recognition and classification of different plant species because they are capable to present flat and two-dimensional surfaces with various characteristics like texture, colour, and shape. Many biological and environmental factors affect leaves to be damaged. So, many characteristics of a damaged leaf will be not useful to provide identifying signals. Therefore, a recognition system that depends on such characteristics may lead to unreliable and inconsistent outcomes.

Plant species recognition and classification method by conventional artificial processes are present time consuming, due to the depending on specific botanical information used by common persons [1]. Many research topics based on the automatic classification of the plant species are important. Some effective algorithms in computer science such as pattern recognition, image processing and machine learning and some technologies such, mobile devices and digital cameras, present the idea of automated classification for plant species by extracting different characteristics from the images of a plant leaf. With the development of machine learning, image processing, mobile devices, computer software, and hardware [2], it is possible to present an efficient and quick automated system to manage, recognize and understand a plant species [3].

In the area of plant taxonomy, leaf analysis has an essential role used to analyze, recognize and understand plant recognition and leaf patterns. The automatic plant recognition based on some features and characteristics, including leaf texture, leaf shape, leaf colour, and, other geometric features has been exploited. These characteristics are dependent on the recognition of the plant species. One of the essential challenges for plant recognition/classification is the diversity of leaf shapes [4]. The colour feature is more dependent to classify and identify plant species because leaf colour is can be changed according to the environment in different seasons. The texture features are more based on the information assured from its vein and venation. Recently, leaf venation patterns are considered an important factor to identify plant species with few techniques to extract leaf vein structure. Many methods depended on automatic or manual leaf venation extraction from leaf patterns. Furthermore, there have been few efforts to correlate and evaluate leaf venation and leaf spectral signatures [5].

In general, texture, shape, and colour features for each kind of plant leaf utilized to recognize plant species [4]. Therefore, most of the existing systems and methods of plant species recognition depend on these features of leaf image with its ability to be valid and reliable for years.

In this paper, different methods used in the plant recognition and classification field are discussed. The implementation and performance of various methods of plant recognition is important for the advancement of these technologies in supporting environment. Hence these methods are reviewed and analyzed. The presented methods have advantages and disadvantages for the recognition and identification of leaf patterns. The remainder of this paper organized as follows: Section 2 presents and discusses various earlier works. Section 3 presents the advanced methods used in leaf recognition. In Section 4, difficulties and directions related to the earlier proposed methods of leaf recognition are discussed. Conclusions are presented in Section 5.

II. LITERATURE REVIEW

In general, there is a general step for leaf recognition, including capturing leaf's images, applies pre-processing method on the captured image, extract feature and classify leaf. Fig. 1 illustrates the flowchart of the major steps carried out in the process of leaf recognition.

A. Images Capturing

In various studies, a scanner or digital camera is used for acquiring leaf images. In [6], the authors used a Samsung camera (DV300F SAMSUNG zoom Lens 5X 16.1 megapixels) to capture images of on-branch green apples, apricot, nectarine, sour cherry, peach, and amber-coloured plums. A digital camera (SONY W730) is used in [7] to capture the green apple targets. The Microsoft Kinect 2.0 camera is chosen in [8] to capture juicy peach images for colour, depth, and point cloud features. While the authors in [9] used an MX808 camera to collect green pepper plant images to create a new dataset. The Canon 660D digital camera used to collect 8911 images of rice leaf disease as a dataset used in the paper [10].

To collect 2D images for apple fruit counting and diameter, the authors in [11] used a thermal camera for accurate results. Also, a thermal camera is used in [12] to collect 2D images of oranges for recognition. Because of the limitation presented with 2D images related to incomplete information, 3D images are considered in many types of research. A laser scanner used in [13], [14], [15] to scan 3D images. Alternatively, an RGB-D camera is used in [16], [17], [18] to present a complete and significantly 3D scan.

B. Images Pre-Processing Methods

An important concept in the leaf recognition system is the pre-processing phase. This phase includes the following steps: image re-orientation, image cropping, convert the image to a grayscale image than to a binary image, remove noise, stretch contrast, and threshold inversion [19]. Various preprocessing techniques are developed based on efficient machine learning methods. How leaf images' features are extracted, and the outcomes of pre-processing phase are important aspect of visual-based machine learning. The study in [20], suggested that to extract leaf features, the leaf image is divided into 2/4 parts, instead of the whole leaf extraction. Vein, colour, Fourier descriptors (FD) exploited in the presented image processing techniques. To achieve a sufficient rate of accuracy, Gray-Level Co-occurrence Matrix (GLCM) methods and the Flavia leaf dataset are used to present 99.1% accuracy. In [21], presented a study and analysis of different methods used various image pre-processing techniques. Simple Linear Iterative Clustering (SLIC) used in one of the studied methods, which uses on super-pixel for grouping them with a defined value through many iterations of the closed neighbour to determine a data vector with a similar value. In [22], the Guided Active Contour (GAC) method is developed. In this method, the snake segmentation technique is used to enhance the polygonal framework for the elongated leaf shape.

For extraction of segments from the data, a hierarchical model based on the Kurtz algorithm is proposed [23]. The proposed approach suggests extracting the interesting parts from data. The data is arranged from the lowest to the highest resolution as clusters as a tree. The first cluster represents the colour features of coarse image patches. The Binary Partition Tree (BPT) used to arrange the individual patches in a hierarchical manner. This method shows that the precision of the system reached up to 85.1%. In [24], a pre-processing technique is used in the proposed system for recognition of soybean and weed leaf. The data used include the images captured by the 2G-R-B camera where the erosion algorithm utilized to remove images distortion. Moment invariant is used to identify scale, invariability, rotation, and translation of soybean leaf image. The image pre-processing technique used can improve the classification rate to 90.5%.

C. Feature Extraction Methods

Some important characteristics such as colour, size, and shape are used for leaf recognition. The segmented image can be a source of information for feature extraction and could assist in the proper classification of the anomaly. Some statistical measures used for textural features extraction such as Color Co-occurrence Matrix (CCM), Spatial Grey Level Dependence Matrix (SGLDM), Grey Level Co-occurrence Matrix (GLCM), Local Binary Patterns (LBP). Various existing systems and methods of plant recognition depend on the colour, size, shape, and texture of the leaf image.

The study in [25] leaves in plants have holes or diseases that could cause reduction of leaves, and thus cause segmentation. First, point searched by pixel scanning and arranged as foreground/background. When the pixel is categorized as foreground, this process cuts off and the next line is scan. Every individual pixel passed with this process for identification. The result of this model was provided with an average error of 3.00 for five leaves. The study in [26] proposed a system for applying feature extraction by utilizing a method known as area labelling. The pre-processing phase is applied for image processing to provide binary image output. Next, the output binary image is offered to area labelling for identified region production. In this work, when the pointer defines a pixel with the value '1' then the eight-connecting area algorithm is used to acquire more search for the eight-connecting area by the kernel. The features of the leaf image are reflected when the pixels are marked and contend for features extraction.
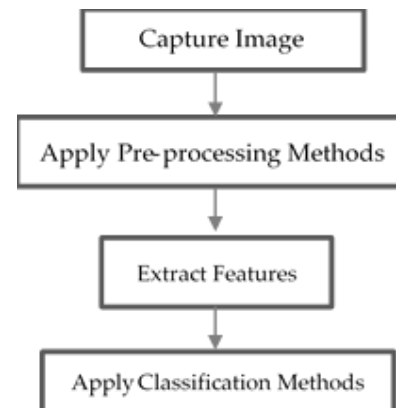

Fig. 1. The Basic Steps of Leaf Recognition Method.

In [27], Gopal et al. attended to present the medicinal images for classification goals depending on the colour features extraction. First, a digital scanner used for providing input image attached for the preprocessing phase. Later, for feature extraction, the image is pushed into the program to gain the colour feature according to its Fourier descriptor. In the training phase, 100 leaf images are used and 50 leaf images are used in the testing phase. The results show that the efficiency of the method is 92%. Feature extraction represents an important role in providing accurate precision and accuracy in leaf recognition/classification system built on utilizing machine learning mechanism. This belongs to the fact that the predetermined feature in the network affects the architecture of machine learning. Different mechanisms used in different approaches to solving different problems so that there are various feature extraction methods to be utilized.

III. Leaf Recognition and Classification Methods

Various related researches proposed for leaf plant recognition and classification is discussed in this section. In [28], the Local Binary Patterns (LBP) method is used to propose an alternative method for plant leaves classification. The proposed method uses the extracted texture features from plant leaves to recognize plant leaves. LBP, the R and G colour of images. In addition, the method efficiency against Gaussian, pepper, and salt are evaluated. Next, the Extreme Learning Machine (ELM) method is used to classify and test the acquired features from the proposed system. In this system, Swedish, Flavia, Foliage, and ICL datasets are used. The obtained results are compared to prove that the proposed method can identify noiseless from noisy images. The accuracy results achieved is claimed to be (98.94%) Flavia, (99.46%) Swedish, (83.71%) ICL and (92.92%) Foliage datasets.

An automatic and accurate segmentation method is proposed in [29]. The authors have used an efficient encoding method for the feature depth information extraction. Later, Mask R-CNN is deployed to train the used RGB-D data. For more efficiency, the features of the data are fused in the Feature Pyramid Network (FPN) structure. Next, Density-Based Spatial Clustering of Applications with Noise (DBSCAN) provided to segment a single leaf from overlapping leaves in the explored scope using the detected leaf areas and depth data. The experimental results are compared to prove that the proposed system automatically detects leaves with an accuracy of around 89.3%. In [30], the authors used the dataset of apple leaf image that employed six apple leaf diseases to provide 2462 images for method evaluation. The proposed method is compared with the traditional multi-classification method based on cross-entropy loss function for results evaluation. The traditional multi-classification method achieves an accuracy of 92.29%, while the proposed method in [30] presents better accuracy with 93.51%, 93.31%, and 93.71% on the test set, respectively.

In [31], Jaya Algorithm with the optimized deep neural network used to propose a system for paddy leaf diseases identification. The leaves image of the rice plant is taken normally from the field, brown spot, blast, and sheath rot diseases. In the pre-processing phase, the RGB images are converted into HSV images and binary images are extracted to split the non-diseased and diseased samples. For the segmentation of non-diseased portion, diseased portion, and background a clustering method is utilized. Jaya Optimization Algorithm (DNN_JOA) with Optimized Deep Neural Network is used in the Classification of diseases phase. The results of the work prove that the proposed method achieved an accuracy of 90.6%.

The authors in [32] presented a classification method of plant's leaves based on Multiscale Triangle Descriptor (MTD) and Local Pattern Histogram Fourier (LBP-HF). The two methods are employed to characterize shape and texture, respectively. Based on their experiments, the recognition accuracy ratio is found to be 99.1%, 98.4%, 95.6% when applied on Flavia, Swedish and MEW2012 datasets, respectively. However, the method has some limitations. The features of the leaves need to be designed manually as no automated process of learning is applied. In [33], an alternative recognition method is presented based on Generalized Procrustes Analysis (GPA). The method uses contour (shape) features for classification. The core of the method depends on performing some computation to calculate the distance between a set of contour points and the center of the contour upon applying some alignments. The results show that the recognition accuracy rate is 84.4% and 98.4% on Leafsnap and Flavia datasets, respectively.

A recognition method based on Multiscale Sliding Chord Matching (MSCM) is presented in [34]. The method aims to recognize soybean cultivar by joint leaf patterns. The MSCM strategy is implemented to extract shape features. The experiment over 6000 sample images shows that the accuracy ratio is 72.4%. The analysis shows that such a low ratio results from several reasons. The leaves of the soybean plan have different visual cues for soybean cultivar identification. In addition, the joint leaf pattern is not integrated with the descriptors of leaves from different parts of soybean plants. There are many other classification models found in various researches. These methods include Support Vector Machine, Artificial Neural Network, Convolutional Neural Network, K-Nearest Neighbors, and Probabilistic Neural Network.

A. Support Vector Machine (SVM)

In SVM is an essential machine-learning technique for data learning and solving classification and identification problems. The study in [35] proposed utilizing leaf contour and centroid for proposing the leaf image recognition systems. The proposed method aimed to use image processing techniques as well as SVM utilized as a classifier. Flavia dataset utilized to take 70 patterns with their shape and geometrical features. Their results prove that the highest achievement accuracy of 97.7%. In [36], the authors provide a comparative analysis for leaf recognition and classification. SVM used as the classifiers in this system and a shape detector utilized to extract 14 leaf features. In the training dataset, the Flavia database used to provide sixteen different plant species. The results show that the highest accuracy of 90.9% by exploiting SVM.

Araujo et al. [37] used SVM and neural network as classifiers of leaf image classification. These classifiers used

for training four different features, the histogram of gradients (HOG), namely local binary pattern (LBP), Zernike Moments (ZM), and speed of robust features (SURF). The results show that using multiple classifiers of the system overcame the performance of monolithic methodologies and the best results reported. A significant improvement proved to be effective to detect plants by using SVM as a classifier for an environment with heavy overlapping and interferences cases [38]. In this experiment, the authors exploited 300 leaf images of three plant species for identification. A marker-controlled watershed segmentation was used to capture and segment the images. The system achieves 86.7% accuracy for identification. The accuracy can be improved by adding more features as well as the dataset used for the experiments. SVM suffer from different limitations, such as the complexity of its structure, and the slowness of training and testing. On the other hand, SVM is considered robust and has high potentials for generalization.

B. Artificial Neural Network (ANN)

The proposed system of leaf pattern recognition in [39] exhibits that using ANN as a classifier is reliable. There was a study presents 98.6% of accuracy for recognition which can be increased when more dataset used [39]. In [40], using ANN as a classifier to recognize and identify the medicinal plant leaves can improve the results. The ANN classifier was used to train the extracted colour, shape, and texture of leaf images. The results show that the system presents an accuracy of 94.4% using 63 leaf images. The accuracy of the extracted leaf venation improved in [41] by about 10% when selecting the ANN as classifier combined with thresholding. The results show that the accuracy improved to 97.3% by combining ANN with thresholding. ANN can recognize the relationships between dependent/independent variable, and support simplistic statistical testing. As for the limitations, ANN requires a high computational load and a high tendency of data overfitting.

C. Convolutional Neural Network (CNN)

In [42], CNN is used to establish a cotton growth recognition algorithm. Confusion matrix and recognition efficiency exploited for the optimization process where a CNN model is established, and its precision was proved by modifying training /test sets based on the concept of the k-fold test. The results show that this method is suitable for the recognition task and can achieve good results in the term of high precision, low cost, and real-time. The method proposed in [43] presents an automated system for medicinal plant classification using CNN. A 3-layer CNN is employed to extract high-level features for classification. The method is supported by a data augmentation technique for higher efficiency. The experimental results show that the recognition accuracy rate of the method is around 71.3%.

To solve the disease similarity problem, an efficient method is proposed in [44]. Two types of diseases happening in the same leaf and the influence of external light lead to this problem. In the beginning, they gained a cucumber leaf disease dataset, then they build a classification model by using the EfficientNet method for the above four types. Finally, they used CNN-based EfficientNet-B4 to demonstrate a two-classification model of cucumber similar diseases. The obtained results prove that their proposed method has a considerable effect on the similar diseases of cucumber classification of accuracy around 96%. In [10], the authors used CNNs to extract the rice leaf disease image features. Later, for classification and prediction of the specific disease SVM method is applied. In their work, the cross-validation method was the optimal parameter of SVM. The results show that the average accuracy of the proposed recognition model was 96.8% based on utilizing deep learning and SVM techniques. The experiment is applied over a dataset prepared by the authors as per the details stated in Table I.

In [45], a deep convolutional neural network used to build an automatic classification and recognition framework of various paddy crop stress as biotic/ abiotic using the field images. The dataset used includes 12 different stress categories of healthy/normal with 30,000 field images of five different paddy crop varieties. The results show that the proposed model can achieve an average accuracy of 92.89%. In image recognition tasks, CNNs are used as feature extractors and classifiers to introduce the better performance. In CNN's, Multiple features are extracted simultaneously as well as they are robust to noise. These advantages made CNN an interesting classifier in many types of research. In [46], the authors aimed to identify leaf diseases based on the traditional CNN by integrating of inception structure and a pooling layer. In this model, the number of parameters reduced and the identification accuracy improved by up to 91.7%. Similarly, the model in [47] used CNN classifier for maize leaf disease detection. This method can classify diseases according to three types. For plant disease classification and recognition, CNN is proven to be an effective manner. The method in [48] integrates deep learning with CNN for classification. The results show that even reducing the number of parameters, would not affect the recognition accuracy.

CNN considered a faster recognition process as it extracts and recognizes the features concurrently. CNN is accurate for plant classification due to the numerous sets of data trained by users before it is considered to be capable enough for application. CNN shows that the accuracy of leaf classification achieved up to 94% [49]. The integration of deep learning knowledge with CNN provided an efficient model for feature extraction to recognize and identify vein samples from the presented image [50].

D. K-Nearest Neighbors (KNNs)

In the recognition and classification methods, the accuracy of identification increased when the number of images for testing is increased. The study in [51] shows that Principal Component Analysis (PCA) algorithm and Cosine k-Nearest Neighbors (KNN) classifier is improved compared to SVM and Patternnet neural network. KNN classifier provides 83.5% of accuracy [19]. Such low accuracy is relatively weak to be agreeable even the process of feature extraction is quick and simple. KNN classifier is not capable to handle samples distortion and could cause inaccuracy in the classification process. A method proposed for this classifier with a specific colour histogram increases the accuracy up to 87.3% [19].

TABLE I.   PREVIOUS LEAF CLASSIFICATION METHODS

Ref.	Published year	Dataset	Classifier	Extracted Features	Average Accuracy Rates
[28]	2019	Flavia *dataset* Swedish *dataset* ICL *dataset* Foliage *dataset*	LBP	Color Texture	Flavia = 98.94% Swedish = 99.46% ICL = 83.71% Foliage = 92.92%
[29]	2020	7988 *images*	Mask R-CNN	Color	89.03%
[31]	2019	650 *images*	Deep Neural Network	Color Texture	90.57%
[32]	2021	Flavia *dataset* Swedish *dataset* MEW2012 *dataset*	MTD + LBP-HF	Texture Shape	Flavia = 99.10% Swedish = 98.40% MEW2012 = 95.60%
[33]	2018	Leafsnap *dataset* Flavia *dataset*	GPA	Shape	Leafsnap = 84.40% Flavia = 98.40%
[34]	2020	6000 *images*	MSCM	Shape	72.40%
[35]	2017	Flavia *dataset*	SVM	Shape	97.70%
[36]	2018	Flavia *dataset*	SVM	Shape	90.90%
[38]	2015	300 *images*	SVM	Shape	86.70%
[37]	2017	ImageCLEF 2011 *dataset* ImageCLEF 2012 *dataset*	SVM + Neural Network	Texture Shape	ImageCLEF 2011 = 86.20% ImageCLEF 2012 = 64.10%
[10]	2020	8911 *images*	SVM + CNN	Shape Color	96.80%
[45]	2020	6000 *images*	CNN	Shape	92.89%
[30]	2020	2462 *images*	CNN	Color Texture Shape	92.29%
[44]	2020	2816 *images*	CNN	Color Texture Shape	96.00%
[42]	2020	1443 *images*	CNN	Texture	93.27%
[46]	2019	6108 *images*	CNN	Color Texture	91.70%
[47]	2019	54306 *images*	CNN	Texture	92.85%
[48]	2019	ImageNet *dataset* PlantVillage *dataset*	CNN	Color	97.14%
[49]	2017	Flavia *dataset*	CNN	Shape	99.70%
[43]	2020	3570 *images*	CNN	Shape Vein	71.30%
[39]	2006	180 *images*	ANN	Shape Vein	94.40%
[40]	2013	63 *images*	ANN	Shape Color Texture	94.40%
[41]	2007	2940 *images*	ANN	Color Vein	97.33%
[51]	2019	ImageCLEF 2012 *dataset* Leafsnap *dataset* Flavia *dataset*	KNNs	Texture	ImageCLEF 2012 = 88.80% Leafsnap = 74.50% Flavia = 98.70%
[52]	2016	Flavia *dataset*	KNNs	Shape	94.37%
[53]	2010	1200 *images*	PNN	Shape	91.41%
[54]	2008	900 *images*	PNN	Shape Texture	93.70%
[55]	2014	Flavia *dataset* Swedish *dataset*	PNN	Shape	Flavia = 82.01% Swedish = 80.01%
[56]	2012	2448 *images*	PNN	Texture Color	74.51%.
[57]	2007	1800 *images*	PNN	Texture Color Shape	90.00%

The authors in [52] produced an improvement in leaf classification based on utilizing KNN classifier with edge and shape features. Flavia dataset exploited to provide 32 plant species to be tested. The results show that the presented method improves the average classification accuracy to 94.4%.

E. Probabilistic Neural Network (PNN)

In the recognition and classification methods, PNN is utilized as a classifier due to many advantages, including high resistance of distortion, flexibility to modify data, and the specimen can be classified into multiple outputs. In this section, we study the efficiency of PNN in classifying leaves.

The work in [53] presents an algorithm for plant species classification of leaf image based on PNN. The points of the leaf's shape are extracted from the background and a binary image is produced accordingly. After that, the leaf is aligned horizontally with its base point on the left of the image. Several morphological features, such as eccentricity, area, perimeter, major axis, minor axis, equivalent diameter, convex area and extent, are extracted. The network was trained with 1200 simple leaves from 30 different plant species with an accuracy rate of 91.41%. The authors in [54] address the issue of low recognition rate in plant identification since the objects broad and the classification features are not synthetic. To resolve this issue, PNN is presented for a rapid recognition method that is applied over thirty kinds of broad-leaved trees. The shape and texture features of broad-leaved trees combine, composing a synthetic feature vector of broad leaves to realize the computer automatic classification towards broad-leaved plants. The use of PNN has achieved an average recognition rate of 93.70%.

An alternative PNN-based leaf classification method is proposed in [55]. Upon converting the RGB image to its binary image representation, the binary image is passed to a canny operator to recognize the edges of the image. Sampling is then used to compute the centroid distance of these points and the distance of sampling points from the axis of the least inertia line. A probabilistic neural network has been used as a classifier. The results show that the average accuracy rates of the method on Flavia and Swedish datasets are 82.1% and 80.1%, respectively. In [56], the researchers present a mobile application for identifying Indonesian medicinal plants. The application uses both Fuzzy Local Binary Pattern (FLBP) and Fuzzy Color Histogram (FCH) methods for extracting leaf image texture and colour, respectively. For fusion of FLBP and FCH, the Product Decision Rules (PDR) method is applied. As for the classifier, PNN is utilized to classify medicinal plant species. The accuracy of this work is claimed to be around 74.51%. PNN appear to be an effective classifier for the automated leaf recognition method proposed in [57]. The method relies on the use of image and data processing techniques, and applied over 1800 leaf images. The method managed to extract 12 leaf features organized into 5 basic variables which compromise the PNN input vector. The PNN is trained by 1800 leaves to classify 32 kinds of plants. The accuracy is found to be reasonable around 90%. However,

aside from the advantages of PNN mentioned above, PNN is considered as a complicated network layout, and it requires long time on training. In addition, PNN has a tendency for overfitting with too many traits. Table I summarizes the key facts and finding of our analysis.

IV. Discussion and Analysis

In the early presented plants leaves species recognition systems, several issues related to providing better classification results are addressed. Our analysis of existing classification methods focuses on different issues, including the commonly used features and classifiers and their impact on classification accuracy, what datasets are used for testing, and research trends on leaf classification methods.

Researchers have used several features in their methods, including (colour (C), shape (S), texture (T) and vein (V)). We have also found some researches combine multiple features to enhance the accuracy ratio. Most of the researches ($\approx 41\%$ of existing methods) focus on shapes features in their classification methods. Analysis of the accuracy ratio of these methods shows that combining multiple features in the classification method helps in enhancing the accuracy rations of leaves classifications. Our analysis also reveals that there is a lack of studies on methods that use vein as a feature of classification, as only $\approx 6\%$ of existing studies tickle such feature in their methods. However, considering the vein features shows promising results when combined with shape, colour or texture features. Fig. 2(a) shows the percentage of studies discuss each type of features, while Fig. 2(b) reflects the accuracy ratios achieved by these features according to the existing classification methods.

As for classifiers, various techniques are found in the state of the art. We found that there is a greater focus on CNN-based classifiers. Several methods show enhanced performance when combining CNN with other classifiers, such as SVM and LBP. Most of the accuracy ratio shows that CNN-based methods outperform other classifiers. On the other hand, there is a growing interest in ANN classifiers as it shows high accuracy ratios. In the three existing studies on ANN classifiers, results show that the accuracy ratio ranges between 94.4 and 97.3. Such high accuracy should give ANN classifier more interest for researchers in developing new classification methods. Fig. 3 presents the accuracy ratio achieved by different classifiers.
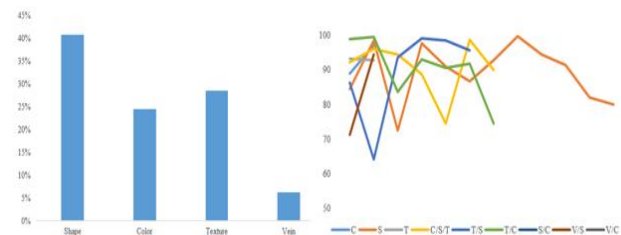

Fig. 2. (Left) Appearance Ratio of Features in Existing Methods, (Right) Accuracy Ratio of different Features.

In term of testing environments, our analysis shows that the majority of researchers (≈ 48%) has developed their datasets for their testing. Using well-known standard datasets such as Flavia and Swedish appeared in less than 30% of the studies. In this regard, researchers should focus on updating and considering standard datasets to enhance the scientific judgments on proposed classification methods. Fig. 4 shows the utilization of different datasets for testing leaves classification methods.

As for the classification method, we noticed that the current researches are oriented toward three main areas. These areas are CNN, SVM and PNN. We found that CNN occupies ≈ 31% of existing classification methods, while each of PNN and SVM found in ≈ 16% of methods. Fig. 5 illustrates the frequencies of different classification methods used in the state of arts.
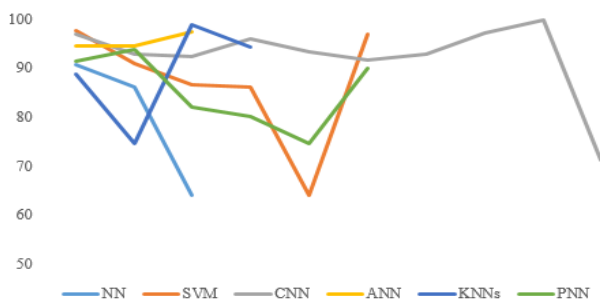

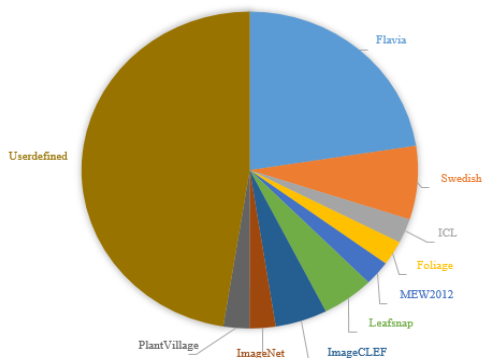Fig. 3.   Accuracy Ratios Achieved by different Classifiers.

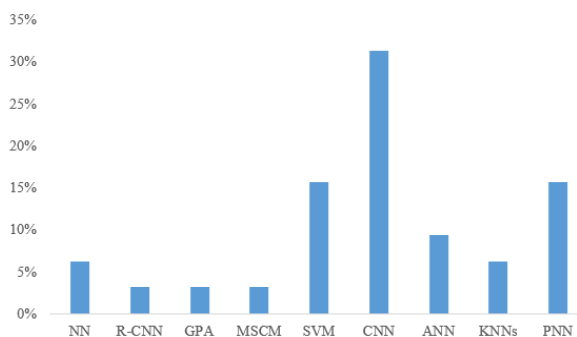Fig. 4.   Datasets used for Testing Leaves Classification Methods.

Fig. 5.   Classification Methods used for Leaves Classification and Recognition.

V.   CONCLUSION

In this research, we have made an effort to study and analyze the latest researches in the field of leaves classification and recognition. We have provided helpful insight on the process of leaves classification using different features of leaves. These features are discussed and analyzed thoroughly, and their efficiency in enhancing the recognition and classification process is presented. In addition, various classifiers and classification methods are studies. Our unique analysis has discussed and analyzed different factors that might affect the accuracy of the classification process. These factors include features, classifiers, and testing datasets. We found that combining multiple features have a positive impact on the classification process. However, greater efforts should be made by researchers to examine and investigate the best combination of features. For instance, none of the researches has combined the vein feature with the colour, shape and texture in one method. We found that CNN classifiers groups the attention of researchers, while SVM classifiers are found more attractive in recent researches. As SVM classifiers look interesting in recent years, further investigations are needed to study the relation between accuracy and the best leaves' features that should be used in SVM-based classification methods. As for the testing datasets, we found that more efforts should be made on unifying these datasets for integrity purposes. The majority of researchers have tested their methods based on some user-defined datasets, which makes the comparisons between proposed methods inaccurate.

REFERENCES

[1]   Y. Shao, "Supervised global-locality preserving projection for plant leaf recognition," Computers and electronics in agriculture, vol. 158, no. October 2018, pp. 102–108, 2019.

[2]   J. Zhang, D. Huang and T. Lok, "A novel adaptive sequential niche technique for multimodal function optimization," Neurocomputing, vol. 69, pp. 2396–2401, 2006.

[3]   M. Seeland, M. Rzanny and N. Alaqraa, "Plant species classification using flower images—A comparative study of local feature representations," PLoS one, vol. 12, no. 2, 2017.

[4]   Q. Zhao, H. Ma. and M. Cheung, "An efficient android-based plant leaf identification system," Neurocomputing, vol. 151, pp. 1112–1119, 2015.

[5]   S. Green, A.Walton, S. Little, C. Price, S. Wing et al., "Reading the leaves: a comparison of leaf rank and automated areole measurement for quantifying aspects of leaf venation," Applications in plant sciences, 2, 2014.

[6]   S. Iman, and H. Khosravi, "Expert systems with applications a deep neural network approach towards real-time on-branch fruit recognition for precision horticulture," Expert systems with applications, vol. 159, p. 113594, 2020.

[7]   S. Sun, Q. Wu, L. Jiao, Y. Long, D. He et al., "Recognition of green apples based on fuzzy set theory and manifold ranking algorithm," International journal for light and electron optics, 2018.

[8]   G. Wu, B. Li, Q. Zhu, M. Huang and Y. Guo, "Using color and 3d geometry features to segment fruit point cloud and improve fruit recognition accuracy," Computers and electronics in agriculture, vol. 174, no. January, p. 105475, 2020.

[9]   W. Ji, X. Gao, B. Xu, G. Chen and D. Zhao, "Target recognition method of green pepper harvesting robot based on manifold ranking," Computers and electronics in agriculture, vol. 177, no. July, 2020.

[10] F. Jiang, Y. Lu, Y. Chen, D. Cai and G. Li, "Image recognition of four rice leaf diseases based on deep learning and support vector machine," Computers and electronics in agriculture, vol. 179, no. October, p. 105824, 2020.

[11] D. Stajnko, M. Lakota and M. Hočevar, "Estimation of number and diameter of apple fruits in an orchard during the growing season by thermal imaging," Computers and electronics in agriculture, vol. 1, pp. 31–42, 2004.

[12] D. Bulanon, T. Burks and V. Alchanatis, "Image fusion of visible and thermal images for fruit detection," Biosystems engineering, vol. 113, pp. 12–22, 2009.

[13] P. Eizentals and K. Oka, "3D pose estimation of green pepper fruit for automated harvesting," Computers and electronics in agriculture, vol. 128, pp. 127–140, 2016.

[14] J. Mack, C. Lenz, J. Teutrine and V. Steinhage, "High-precision 3D detection and reconstruction of grapes from laser range data for efficient phenotyping based on supervised learning," Computers and electronics in agriculture, vol. 135, pp. 300–311, 2017.

[15] S. Paulus, J. Dupuis, A. Mahlein and H. Kuhlmann, "Surface feature based classification of plant organs from 3D laserscanned point clouds for plant phenotyping," BMC bioinformatics, vol. 14, p. 238, 2013.

[16] E. Barnea, R. Mairon and O. Ben-Shahar, "Colour-agnostic shape-based 3D fruit detection for harvesting robots," Biosystems engineering, vol. 146, pp. 57–70, 2016.

[17] R. Perez, F. Cheein and J. Rosell-Polo, "Flexible system of multiple RGB-D sensors for measuring and classifying fruits in agri-food Industry," Computers and electronics in agriculture, vol. 139, pp. 231–242, 2017.

[18] T. Yongting and Z. Jun, "Automatic apple recognition based on the fusion of color and 3D feature for robotic fruit picking," Computers and electronics in agriculture, vol. 142, pp. 388–396, 2017.

[19] T. Munisami, M. Ramsurn, S. Kishnah and S. Pudaruth, "Plant leaf recognition using shape features and colour histogram with k-nearest neighbour classifiers," Procedia computer science, vol. 58, pp. 740 – 747, 2015.

[20] M. Turkoglu and D. Hanbay, "Recognition of plant leaves: An approach with hybrid features produced by dividing leaf images into two and four parts," Applied mathematics and computation, vol. 352, pp. 1–14, 2019.

[21] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua et al., "SLIC superpixels compared to state-of-the-art superpixel methods," IEEE transactions on pattern analysis and machine intelligence, vol. 34, pp. 2274–2281, 2012.

[22] G. Cerutti, L. Tougne, J. Mille, A. Vacavant and D. Coquin, "Understanding leaves in natural images—a model-based approach for tree species identification," Computer vision and image understanding, vol. 117, pp. 1482–1501, 2013.

[23] C. Kurtz, N. Passat, P. Gançarski and A. Puissant, "Extraction of complex patterns from multiresolution remote sensing images: a hierarchical top-down methodology," Pattern recognition, vol. 45, pp. 685–706, 2012.

[24] Z. Bo, W. Hua, L. Jun, M. Hua and Z. Chao, "Research on weed recognition method based on invariant moments," 11 world congress on intelligent control and automation, pp. 2167–2169, 2014.

[25] C. Lu, H. Ren, Y. Zhang and Y. Shen, "Leaf area measurement based on image processing," 9th international conference on measuring technology and mechatronics automation, vol. 2, pp. 580–582, 2010.

[26] V. Shivling, A. Singla, C. Ghanshyam, P. Kapur and S. Gupta, "Plant leaf imaging technique for agronomy," 2011 International conference on image information processing, pp. 1–5, 2011.

[27] A. Gobal, S. Reddy and V. Gayatri, "Classification of selected medicinal plants leaf using image processing," 2012 International conference on machine vision and image processing, pp. 5–8, 2012.

[28] M. Turkoglu and D. Hanbay, "Leaf-based plant species recognition based on improved local binary pattern and extreme learning machine," Physical A: statistical mechanics and its applications, vol. 527, p. 121297, 2019.

[29] X. Liu, C. Hu and P. Li, "Automatic segmentation of overlapped poplar seedling leaves combining mask R-CNN and DBSCAN," Computers and electronics in agriculture, vol. 178, no. August, p. 105753, 2020.

[30] Y. Zhong and M. Zhao, "Research on deep learning in apple leaf disease recognition," Computers and electronics in agriculture, vol. 168, no. October 2019, p. 105146, 2020.

[31] S. Ramesh and D. Vydeki, "Recognition and classification of paddy leaf diseases using optimized deep neural network with Jaya algorithm," Information processing in agriculture, vol. 7, no. 2, pp. 249–260, 2020.

[32] C. Yang, "Plant leaf recognition by integrating shape and texture features", Pattern recognition, vol.112, p. 107809, 2021.

[33] S. Choudhury, J. Yu and A. Samal, "Leaf recognition using contour unwrapping and apex alignment with tuned random subspace method", Biosystems engineering, vol. 170, pp. 72-84, 2018.

[34] B. Wang, Y. Gao, X. Yuan and S. Xiong, "From species to cultivar: Soybean cultivar recognition using joint leaf image patterns by multiscale sliding chord matching", Biosystems engineering, vol. 194, pp. 99-111, 2020.

[35] A. Khmag, S. Al-Haddad and N. Kamarudin, "Recognition system for leaf images based on its leaf contour and centroid," 2017 IEEE 15th student conference on research and development, pp. 467–472, 2017.

[36] V. Srivastava and A. Khunteta, "Comparative analysis of leaf classification and recognition by different SVM classifiers," 2018 International conference on inventive research in computing applications, pp. 626–631, 2018.

[37] V. Araujo, A. Britto, A. Brun, A. Koerichm and R. Palate, "Multiple classifier system for plant leaf recognition," 2017 IEEE international conference on systems, man and cybernetics, pp. 1880–1885.

[38] R. Nesaratnam and C. Murugan, "Identifying leaf in a natural image using morphological characters," International conference on innovations in information, embedded and communication systems, 2015.

[39] Q. Wu, C. Zhou and C. Wang, "Feature extraction and automatic recognition of plant leaf using artificial neural network," Research on computing science, vol. 20, pp. 3–10, 2007.

[40] R. Janani and A. Gopal, "Identification of selected medicinal plant leaves using image features and ANN," 2013 International conference on advanced electronic systems, pp. 238–242, 2013.

[41] H. Fu and Z. Chi, "Combined thresholding and neural network approach for vein pattern extraction from leaf images," IEE Proc. - Vision, image signal process, vol. 153, pp. 881–892, 2007.

[42] S. Wang, Y. Li, J. Yuan, L. Song and X. Liu, "Recognition of cotton growth period for precise spraying based on convolution neural network," Information processing in agriculture, pp. 1–13, 2020.

[43] R. Akter and M. Hosen, "CNN-based leaf image classification for Bangladeshi medicinal plant recognition", Emerging technology in computing, communication and electronics, pp. 1-6, doi: 10.1109/ETCCE51779.2020.9350900.

[44] P. Zhang, L. Yang and D. Li, "EfficientNet-B4-ranger: a novel method for greenhouse cucumber disease recognition under natural complex environment," Computers and electronics in agriculture, vol. 176, no. July, p. 105652, 2020.

[45] B. Anami, N. Malvade and S. Palaiah, "Deep learning approach for recognition and classification of yield affecting paddy crop stresses using field images," Artificial intelligence in agriculture, vol. 4, pp. 12–20, 2020.

[46] B. Hang, D. Zhang, P. Chen and J. Zhang, "Classification of plant leaf diseases based on improved convolutional neural network," Sensors, vol. 19, p. 4161, 2019.

[47] M. Sibiya and M. Sumbwanyambe, "A Computational procedure for the recognition and classification of maize leaf diseases out of healthy leaves using convolutional neural networks," AgriEngineering, vol. 1, pp. 119–131, 2019.

[48] Y. Toda and F. Okura, "How convolutional neural networks diagnose plant disease," Plant phenomics, vol. 2019, Article ID 9237136, doi: 10.34133/2019/9237136.

[49] W. Jeon and S. Rhee, "Plant leaf recognition using a convolution neural network," Korean institute of intelligent systems, vol. 17, pp. 26–34, 2017.

[50] G. Guillermo, L. Uzal, M. Larese and P. Granitto, "Deep learning for plant identification using vein morphological patterns," Computers and electronics in agriculture, vol. 127, pp. 418–424, 2016.

[51] F. Kheirkhah and H. Asghari, "Plant leaf classification using GIST texture features," IET computer vision, vol. 13, p. 369, 2018.

[52] P. Kumar, K. Rao, A. Raju, and D. Kumar, "Leaf classification based on shape and edge feature with k-NN classifier," 2016 2nd Int. Conf. Contemp. Comput. Informatics, pp. 548–552, 2016.

[53] J. Hossain, and M. Amin, "Leaf shape identification based plant biometrics," Proc. 2010 13th Int. Conf. Comput. Inf. Technol. ICCIT 2010, no. Iccit, pp. 458–463, 2010.

[54] L. Huang, and P. He, "Machine recognition for broad-leaved trees based on synthetic features of leaves Using Probabilistic Neural Network," Proc. - Int. Conf. Comput. Sci. Softw. Eng. CSSE 2008, vol. 4, pp. 871–877, 2008.

[55] K. Mahdikhanlou, and H. Ebrahimnezhad, "Plant leaf classification using centroid distance and axis of least inertia method," 2014 22nd Iran. Conf. Electr. Eng., pp. 1690–1694, 2014.

[56] Y. Herdiyeni, and N. Wahyuni, "Mobile Application for Indonesian Medicinal Plants Identification using Fuzzy Local Binary Pattern and Fuzzy Color Histogram," 2012 Int. Conf. Adv. Comput. Sci. Inf. Syst., pp. 978–979, 2012.

[57] S. Wu, F. Bao, E. Xu, Y. Wang, Y. Chang, and Q. Xiang, "A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network," IEEE Int. Symp. Signal Process. Inf. Technol., 2007.