VNU Journal of Science: Comp. Science & Com. Eng, Vol. 34, No. 2 (2018) 1-15

VNU Journal of Science:
Computer Science and Communication Engineering

Journal homepage: http://www.jcsce.vnu.edu.vn/index.php/jcsce

Score-based Fusion Schemes for Plant Identification from Multi-organ Images

Nguyen Thi Thanh Nhan[1,3,*], Do Thanh Binh[1,2], Nguyen Huy Hoang[1,2], Vu Hai[1], Tran Thi Thanh Hai[1], Thi-Lan Le[1]

[1]International Research Institute MICA, HUST - CNRS/UMI-2954 - GRENOBLE INP, Hanoi, Vietnam
[2]School of Information and Communication Technology, HUST, Hanoi, Vietnam
[3]University of Information and Communication Technology, Thai Nguyen University, Thai Nguyen, Vietnam

Abstract

This paper describes some fusion techniques for achieving high accuracy species identification from images of different plant organs. Given a series of different image organs such as branch, entire, flower, or leaf, we firstly extract confidence scores for each single organ using a deep convolutional neural network. Then, various late fusion approaches including conventional transformation-based approaches (sum rule, max rule, product rule), a classification-based approach (support vector machine), and our proposed hybrid fusion model are deployed to determine the identity of the plant of interest. For single organ identification, two schemes are proposed. The first scheme uses one Convolutional neural network (CNN) for each organ while the second one trains one CNN for all organs. Two famous CNNs (AlexNet and Resnet) are chosen in this paper. We evaluate the performances of the proposed method in a large number of images of 50 species which are collected from two primary resources: PlantCLEF 2015 dataset and Internet resources. The experiment exhibits the dominant results of the fusion techniques compared with those of individual organs. At rank-1, the highest species identification accuracy of a single organ is 75.6% for flower images, whereas by applying fusion technique for leaf and flower, the accuracy reaches to 92.6%. We also compare the fusion strategies with the multi-column deep convolutional neural networks (MCDCNN) [1]. The proposed hybrid fusion scheme outperforms MCDCNN in all combinations. It obtains from + 3.0% to + 13.8% improvement in rank-1 over MCDCNN method. The evaluation datasets as well as the source codes are publicly available.

Received 28 March 2018, Revised 18 September 2018, Accepted 13 December 2018

Keywords: Plant identification, Convolutional neural network, Deep learning, Fusion.

1. Introduction

Plant identification plays an important role in our daily life. Nowadays, automatic vision-based machines for the plant identification usually utilizes image(s) from individual plant organs such as leaf [2-4], flower [5], branch [6]. Recently, this topic has obtained a considerable attention of scientists in the fields of multimedia retrieval, computer vision, and

* Corresponding author. Email: nttnhan@ictu.edu.vn
https://doi.org/10.25073/2588-1086/vnucsce.201

pattern recognition. In recent competitions for the plant identification (e.g., PlantCLEF 2014, 2015, 2016 and 2017), deep learning technique has emerged as an effective tool. However, with a large number of species, the single organ identification accuracy is still limited. In addition, complex backgrounds and the appearance of multiple organs in one image increase the difficulty of this task. The performance issues of the classifiers, but using images from individual plant organ also has some practical and botanical limitations. For instance, the appearance of leaves can be easily changed by temperature, weather condition. Some leaves of specific species are often too young or too much depends on periods of the year. The appearance of flowers is more stable and less variant with such changes. However, some organs are not visible throughout the year such as fruit, flower, or even leaf. Following the point of view of botanists and biological experts, images from single organ do not have enough information for the identification task due to the large inter-class similarity and large intra-class variation. They also comment that there are many practical situations where separating species can be very difficult by just observing leaves, while it is indisputably easier with flowers. Recently, more researches have been dedicated to plant identification from images of multi-organs especially with the release of a large dataset of multi-organs images of PlantCLEF since 2013 [6-10]. Pl@ntnet is the first tool that identifies plants based on multi-organ [11]. It first performs plant identification from an image of each organ and then combines the identification results of multi-organs to create the final identification result. To leverage the role of organs, each type of organ has different weight. For example, flowers have higher weights than leaves because flowers have better distinguishing characteristics than leaves. In this tool, the weights for each organ are empirically optimized. Studies [10-14] have shown that the plant identification based on multiple organs outperforms that of single organ.

In [15] and relevant works [14], for single organ plant identification, we proposed to use deep CNN that could achieve the higher performance than conventional hand-designed feature approaches. However, it is noticed that the performances of a CNN strongly depend on image varieties within each species in the training dataset. The performances of the plant identification task could be increased when the number of images for each species is large enough. Especially, a large number of images of each plant organ with same species is required in the context of the multi-organ combination. Therefore, we take into account collecting the images of different organs of same species for the context of the multi-organ combination. Then, three fusion techniques that are transformation-based fusion approaches, classification-based fusion approaches [16], and our own proposed robust hybrid fusion (RHF) are evaluated. Four most common types of organs that are leaf, flower, branch and entire are used in the evaluation. Each pair of organs is combined and examined with these fusion approaches.

Our work focuses on score-based fusion schemes for determining the name of species based on images of different organs. In the previous work [15], a method for plant identification from multi-organs images is proposed. As a consequence, the experimental results in [15] confirmed that fusion approach is a potential solution to increase the accuracy rate for identifying plant species. This paper is an extended version [15] with the following new contributions. First, in this paper, for single organ plant identification, with the aim of answering the question: “Is it possible to learn one sole network for all types of organs?”, we define and evaluate two schemes: (1) one CNN for each organ and (2) one CNN for all organs. The first scheme allows to make explicit fusion for each organ while the second does not require to know the type of the organ and consumes less computation resources. Second, besides AlexNet used in [15], in this work, we employ another network architecture (ResNet) for single organ plant identification. Several experiments have been carried with the aim of evaluating the performance of two proposed schemes and CNNs (AlexNet and ResNet) for single plant identification as well as multiple organ plant identification through the proposed fusion schemes. The experimental results show that the proposed method obtains from +3.0% to +13.8% improvement in rank-1 over the MCDCNN method [1]. Finally, we public the codes and evaluation datasets that are used in this paper.

This paper is organized as follows: Section 2 surveys relevant works of the plant identification and the fusion approaches. The overall framework is presented in Section 3. The single organ identification using a convolutional neural network is described in Section 4. In Section 5, we present the combination of multi-organ images with various fusion schemes. Section 6 shows the experimental results. The conclusions and discussions are given in Section 7.

2. Related work

2.1. Single organ plant identification

Since the last decade, the plant identification tasks mainly utilize images from leaves on a simple background [17-21] because leaves usually exist in a whole year and are easily collected. However, leaves often do not have enough information to identify a plant species. The plant identification task has recently been expanded with images from different organs [1, 22] such as leaf, flower, fruit, stem, and entire on a complex background so that the identification accuracy is better. The performances of the recent approaches are listed in a technical report of the LifeCLEF 2015 [6]. Readers can also refer to a recent comprehensive survey on plant species identification using computer vision techniques in [23].

There are two main approaches to the plant identification task. The first one uses hand-designed feature [17, 24, 25] where the automatic vision-based machines applied a variety of generic feature extraction and classification techniques. The common features [23] are morphological, shape-based, color, textures, while the Support Vector Machines (SVM) and Random Forest (RF) are common classifiers. These approaches are steady but achieve low performances when facing a large number of species such as 500 species in PlantCLEF 2014, 1000 species in PlantCLEF 2015/2016 datasets [6] and 10000 species in PlantClef2017 [10]. The second one employs the deep learning techniques. Convolutional neural networks (e.g., AlexNet, VGGNet, GoogLeNet and ResNet) obtained state-of-the-art results in many computer vision tasks [26, 27]. The teams utilizing deep learning techniques are top winners in PlantCLEF competition. In PlantCLEF 2014 [28], the winner used AlexNet from scratch to classify 500 plant species. Continuing this success, many research groups have used the deep learning approaches for the plant identification [6, 29]. In PlantClef 2015 [6], the CNN is mostly used by GoogLeNet. GoogLeNet, VGGNet, CaffeNet, AlexNet, ResNet, Inception v4 and Inception-ResNet are used by most teams in the PlantCLEF 2016/2017 competition [9, 10], including the winning team. Applying some CNNs, then classifier ensembles tend to yield better results than applying one CNN [10, 29], this is a new trend for plant identification. In [30], a CNN is used to learn unsupervised feature representations for 44 different plant species collected at the Royal Botanic Gardens, Kew, England. [14] carried out and analyzed a comparative evaluation between hand-designed features and deep learning approaches. They show that CNN-based approaches are significantly better than the hand-designed schemes.

2.2. Multi-organ plant identification

The fact that the state-of-the-art results of the plant identification using a single organ are still far from practical requirements. Currently, the best rank-1 plant identification accuracy is approximately 75% by using flower images. In our empirical evaluation, this performance is significantly reduced when the number of species is increased. The classifiers utilizing the image(s) from individual organs face a challenge that is the small variation among species, and a large variation within a species. Therefore, some recent studies proposed the combinations of multiple organs of plants [1, 22].

There are two main approaches for plant identification from multi-organs. The first approach tries to secure the final performance by focusing on improving the performance of single-organ plant identification while the second one attempts to develop fusion schemes.

The works belonging to the first approach simply apply average function to get the final plant identification from those obtained for different organs [6, 29, 31]. It is worth to note that the average is equivalent to Sum rule and in the experiment section, we will show that this fusion technique is not suitable for plant identification as it does not take into account the role of the plants' organs.

Concerning the second approach, most works apply late fusion at score level for identifying the plant species from the identification results of different organs. The score level fusion can be categorized into three groups: transformation-based approaches, classification-based approaches, and density-based approaches [16]. In transformation-based approaches, the matching or confidence scores are normalized first. Then they are fused by using various rules such as max rule, product rule, or sum rule, to calculate a final score. The output decision is marked based on that final score. [14] used the sum rule to combine identification results from leaf and flower images and got the better result than those of single organ. In classification-based approaches, multiple scores are treated as feature vectors and a classifier, such as Support Vector Machine and Random Forest, is constructed to discriminate each category. The signed distance from the decision boundary is usually regarded as the fused score. The last group, density-based approaches guarantee the optimal fusion as long as the probability density function of the score given for each class is correctly computed. However, such kind of approaches are suitable only for verification issue, but not for identification task.

In this paper, we examine various fusion techniques to answer the questions that which ones achieve the best performances and which pair of organs could achieve the best identification accuracy.

3. Overall framework

In this paper, we focus the second approach for plant identification from multi-organs. In our study, we apply the state-of-the-art methods for plant identification from single organ and focus our contributions on fusion schemes. The proposed framework that consists of two main steps: single organ plant identification and multi-organ plant identification is illustrated in Fig. 1 and Fig. 2. Concerning plant identification from image of single organ, we apply CNN as it has been proved to be effective in previous studies [9]. When applying deep learning for plant identification from image of single organ, one question is naturally raised: Do we need to train a proper CNN for each organ? To answer this question, we propose two schemes as illustrated in Fig. 1: (1) one proper CNN for each organ and (2) one CNN for all organs. The first scheme allows making explicit fusion for each organ while the second does not require to know the type of organ and consumes less computation resources. It is worth to note that in these two schemes, any network can be applied. In this paper, we choose two networks that are AlexNet and ResNet. We obtain confident scores at the output of each single organ plant identifier. For identifying plants using multi-organ images, we propose different late fusion techniques that are classified into transformation-based, classification-based and hybrid fusion schemes. In the section 4 and section 5, we will explain in detail the network architecture used for single organ plant identification as well as the fusion approaches.

4. Single organ identification using deep convolutional neural networks

Plant identification from images of single organ aims to determine the name of species based on images taken from one sole organ of plants. It is worth to note that most works have been dedicated to the single organ plant identification where leaf and flower [32] are two most widely used organ images. Previous studies have shown that deep learning has outperformed hand-crafted features for the single plant identification [10]. In this paper, we take into account the fusion schemes based on the single organ plant identification. In particularly, we employ two well-known CNN networks that are AlexNet and ResNet.We investigate the performance of these networks for the single organ plant identification with two schemes: one CNN for each organ and one CNN for all organs.

AlexNet, which is developed by Alex Krizhevsky, Ilya Sutskever, and Geoff Hinton [27], is the first CNN that has become the most popular nowadays. It succeeds in the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) dataset [33] with roughly 1.2 million labeled images of 1,000 different categories. The AlexNet's architecture is shown in Fig. 3. It has approximately 650,000 neurons and 60 million parameters. There are five convolutional layers (C1 to C5), two normalization layers, three max-pooling layers, three fully-connected layers (FC6, FC7, and FC8), and a linear layer with a Softmax classification in the output. The main reason is that AlexNet runs quite fast on common PC or workstation and achieves comparative results compared with some recent CNNs such as GoogLeNet, VGGNet.

The second network is Residual Network named ResNet. It is the Convolutional neural network of Microsoft team that won ILSRVC 2015 classification task [34]. ResNet-50 is one of the versions provided in experiments, it is a 50 layer Residual Network. There are other variants like ResNet101 and ResNet152 also [34]. ResNet introduces the new terminology is residual learning. The difference between ResNet and others networks is that it aims at leaning some residuals rather than learning features at the end of its layers. Residual can be seen as subtraction of feature learned from layer input. Shortcut connection from input of n^{th} layer to $(n+x)^{th}$ layer is used for ResNet. This kind of network is more efficient and results in better accuracy.

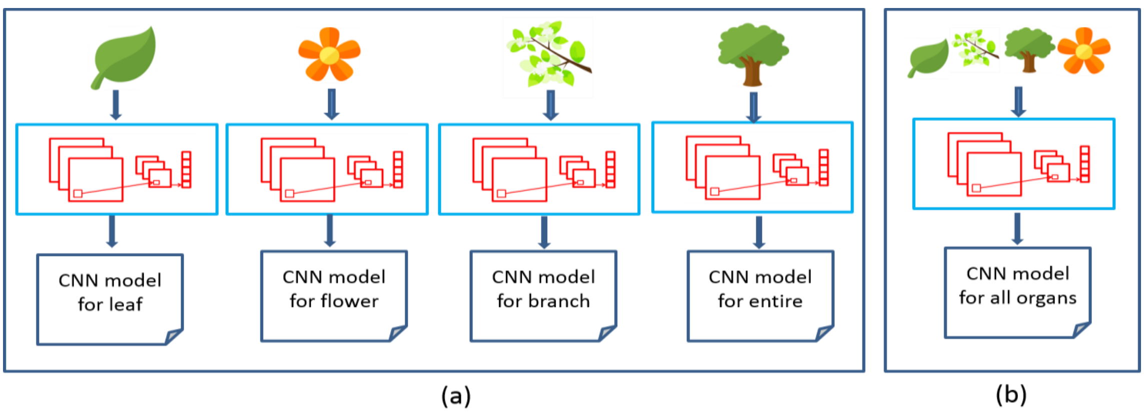

Fig. 1. Single organ plant identification.
a) Scheme 1: One CNN for each organ; b) Scheme 2: One CNN for all organs.

In this study, AlexNet and ResNet are deployed on computer with 2.20 GHz CPU, 16GB RAM and GeForce GTX 1080 Ti GPU. We fine-tuned AlexNet, ResNet-50 with the pre-trained parameters of it in the ImageNet dataset. The output is 50 classes instead of 1000 classes as the default. We optimized the model for this particular task of plant identification, some of the optimization parameters are used in AlexNet are follows: learning rate=0.01, batch size=50, weight decay=0.0005, dropout=0.5, number of epochs=200. In ResNet we use some optimization parameters: learning rate=0.001, batch size=64, weight decay=0.0001, number of epochs=200.

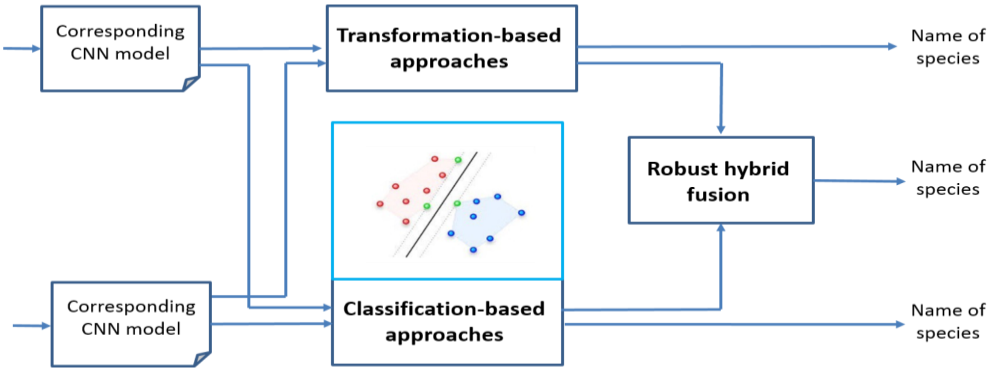

Fig. 2. Multi-organ plant identification.

In the test phase, the output matching/confidence scores obtained for an image is an C-dimensional vector $[s_1 s_2 ... s_C]$ where C is the number of species, s_i is the confidence score to i^{th} plant species, $s_i \in R$, $0 \le s_i \le 1$. The larger s_i is, the greater the probability that the image is taken from the species i^{th} is.

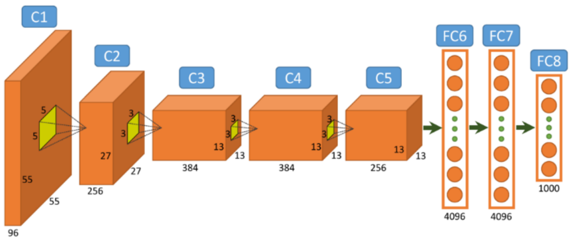

Fig. 3. AlexNet architecture taken from [27].

5. The proposed fusion strategies

5.1. Transformation-based approaches

We combine the identification results from N images of two organs as the following rules. Given the query-images $q = \{I_1, I_2, .., I_N\}$ of a pair of organs, let us define some notations: C is the number of species, $s_i(I_k)$ is the confidence score to i^{th} plant species when using image I_k as a query from a single organ plant identification, where $1 \le i \le C$, $1 \le k \le N$. In our experimental, we choose $N = 2$. The input query q is assigned to class c according to the following rules:

Max rule is one of the most common transformation-based approaches. Maximal score is selected as the final confidence score. In this case, we assign the input query q to class c such that:

$$c = \arg\max_i \max_{k=1..N} s_i(I_k) \tag{1}$$

Sum rule is also the representative of the transformation-based approaches. Summation of the multiple scores provides a single fused score. The sum rule assigns the input query to class c such that:

$$c = \arg\max_i \sum_{k=1}^{N} s_i(I_k) \tag{2}$$

Product rule is based on the assumption of statistical independence of the representations. This assumption is reasonable because observations (e.g., leaf, flower, entire) of a certain species are mutually independent. This allows us using images from multi-organ in order to make a product rule for the plant identification task. The input query is assigned to class c such that:

$$c = \arg\max_{i} \prod_{k=1}^{N} s_i(I_k) \quad (3)$$

5.2. Classification-based approaches

The score-based level fusion can be formed as a classification-based approach. Once the multiple confidence scores are concatenated into a single feature vector, we can build a binary or multiple classifier for it. In this study, we adopt works in [16] which deploys a classification-based approach for fusing multiple human gait features. The plant identification task is formed as a one-versus-all classification. We define a positive/negative sample as a pair of scores at the true/false position of species. Positive and negative samples are chosen as shown in the Fig. 5. An SVM classifier is trained by using positive and negative training samples in the score space.

The distribution of positive and negative samples, which are obtained from confidence scores of branch and leaf images, is shown in Fig. 4. In the test phase, after pushing a pair of organs into the CNN model, we have a pair of score vectors correspondingly. We split it into C pairs where C is the number of species. Then we push each pair into the SVM classifier and we keep it if it is a positive sample. The species of the positive sample, which has the maximum distance to the decision bound, is the label of the pair of organs.

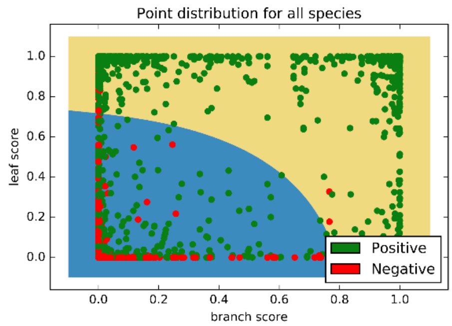

Fig. 4. Distributions of negative and positive samples based on the branch and leaf scores.

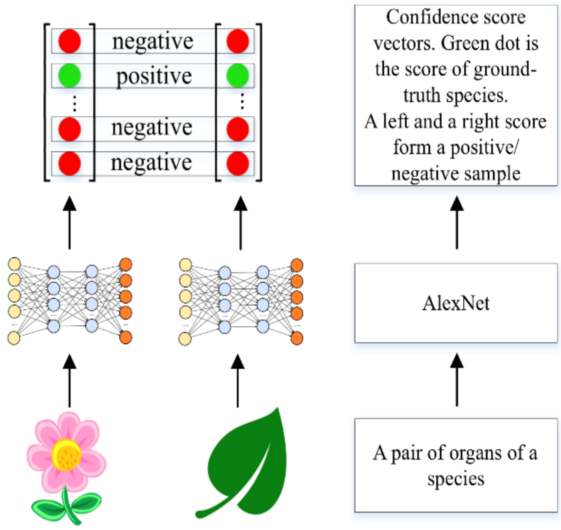

Fig. 5. Explaination for positive and negative samples.

5.3. The proposed robust hybrid fusion

The above classification-based approach can lose distribution characteristics for each species because all positive and negative samples of all species are merged and represented in a metric space only. Therefore, we build each species an SVM model based on its positive and negative samples. For example, Fig. 6 shows a score distribution of a specific species. When we input a pair of organs to our model, we will know the probability that it belongs to each species by these SVM classifiers. Then we combine this probability with the confidence score of each organ. As far as we know, q is the query of a pair of two image organs, and $s_i(I_k)$ is i^{th} species confidence score for image I_k. Let us denote the probability p_i that q is a positive sample of the i^{th} species SVM model. The robust hybrid fusion model is formed as independence observations:

$$c = \arg\max_{i} p_i.(\prod_{k=1}^{N} s_i(I_k)) \quad (4)$$

This model is an integration between a product rule and a classification-based approach. We expect that the positive probability of point q affects the fusion result. If the positive probability of point q is high, the probability of point q belonging to i^{th} species is high, too.

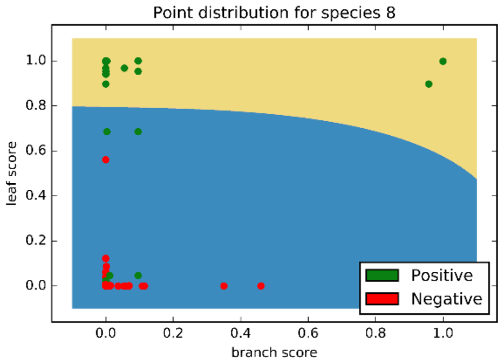

Fig. 6. Distributions of negative and positive samples based on the branch and leaf scores for species id 8.

6. Experimental

6.1. Collecting the database

The proposed fusion strategies are evaluated with four types of organs including leaf, flower, entire and branch. For deploying a CNN successfully, it always requires a large training data. Moreover, for deploying multi-organ plant identification, we must be ensured with different organs of same species. The fact that even with a large PlantCLEF 2015 dataset, there are only 12.5% observations that have at least two organs [1].

In this study, we deploy the following scheme to enrich the experimental dataset of the plant species. Firstly, we extract the most common species (the species with the largest number of images) from PlantCLEF 2015 dataset [6] which is collected fromWest Europe with more than one hundred thousand pictures of 1000 plant species. As a result, we collect 50 species which consist of the largest number of observations. [35] shows that as the number of training images per class increases, the accuracy on the test set will increase, so in this work we used Bulk Image Downloader, which is a powerful tool for collecting images from Internet resources, to collect more data using species' name. The searching results are manually screened later with the help of botanists. The details of our final evaluation dataset are shown in Table 1. The average of images for each organ of each species after enrichment is larger than 50. This is larger than the original PlantCLEF 2015 dataset.

The collected dataset is separated into three parts with the ratio 5:3:2 respectively. The first part is the training data of CNN for single organ identification, as explained in Section 4. We used the third part of the dataset to evaluate the performances of CNN and late fusion methods. For the fusing based on classification approaches, to deploy an SVM classifier, the results from the second part of the dataset returning from CNN was used as training dataset of the SVM model. In order to balance the number of positive and negative sample, we randomly collect the negative points instead of taking all of those. The proposed hybrid fusion scheme utilizes the testing schemes of the product rule and the classification-based approaches.

6.2. Evaluation measurement

To evaluate the performances of the proposed fusion approaches, we use the identification accuracy rate that is defined as follows:

$$Accuracy = \frac{T}{N} \qquad (5)$$

where T is the number of true predictions, N is the number of queries. A query is correctly identified if its actual species is in the k first species returned from the retrieved list. We compute the accuracy at rank-1 and rank-5 in our experiments.

6.3. Experimental results

6.3.1. Evaluation of two schemes for single organ plant identification

We compare the performance of two schemes used for single organ plant identification that are (1) Scheme 1: A CNN (AlexNet or ResNet) for each organ and (2) Scheme 2: A CNN (AlexNet or ResNet) for all organs. The results obtained for the two proposed schemes with two networks are shown in Table. 2, Table. 3. We can observe that ResNet obtained better results than that of AlexNet in both schemes and for most organs except Entire in Scheme 1. It is interesting to see that Scheme 1 is suitable for high discriminative and salient organs such as leaf and flower while Scheme 2 is a good choice for others organs such as branch and entire. The results of branch and entire identification in Scheme 2 are improved because some images of flower and leaf might contain the branch and entire information. The advantage of using scheme 2 for single organ identification is that it does not require to define the type of organ. In the section 6.3.2 and section 6.3.3, the multi-organ plant identification results of the two proposed schemes with two networks will be reported.

Table 1. The collected dataset of 50 species with four organs

	Flower	**Leaf**	**Entire**	**Branch**	**Total**
CNN Training	1650	1930	825	1388	5793
SVM Input	986	1164	495	833	3478
Testing	673	776	341	553	2343
Total	3309	3870	1661	2774	11614
Species number = 50					

Table 2. Single organ plant identification accuracies with two schemes: (1) An AlexNet for each organ; (2) An AlexNet for all organs. The best result is in bold

Organ	Scheme 1		Scheme 2	
	An AlexNet for each organ		An AlexNet for all organs	
	Rank-1 (%)	Rank-5 (%)	Rank-1 (%)	Rank-5 (%)
Leaf (Le)	**66.2**	**89.8**	63.8	87.0
Flower (Fl)	**73.0**	**90.8**	72.2	90.4
Branch (Br)	43.2	70.4	**47.4**	**72.6**
Entire (En)	32.4	64.0	**33.8**	**61.0**

Table 3. Single organ plant identification accuracies with two schemes: (1) A ResNet for each organ; (2) A ResNet for all organs. The best result is in bold

Organ	Scheme 1		Scheme 2	
	A ResNet for each organ		A ResNet for all organs	
	Rank-1 (%)	Rank-5 (%)	Rank-1 (%)	Rank-5 (%)
Leaf (Le)	**73.4**	88.0	70.6	**90.2**
Flower (Fl)	**75.6**	92.6	75.4	**92.8**
Branch (Br)	48.6	73.0	**54.6**	**80.2**
Entire (En)	32.4	63.2	**39.0**	**65.0**

6.3.2. Evaluation of fusion schemes for multiple organ plant identification

Table 4 and Table 5 show the performance obtained when combining a pair of organs for plant identification. The experimental results show that almost the fusion techniques highly improve the accuracy rate compared with utilizing images from one sole organ (see Table 2 and Table 3). In the case, applying scheme 1 for single organ plant identification, for the AlexNet, the best performance for single organ is 73.0% for flower images, whereas by applying the proposed RHF, the accuracy rate of a combination between leaf-flower images dramatically increases by 16.8% to 89.8%. When applying ResNet, the combination of leaf and flower (Le-Fl) improves +17% over the single organ. Not only the leaf-flower scenario but in all six pairs of multi-organs combination, the product rule and its variant RHF also retain the highest performances. Almost the other fusion performances are also higher than those of single organ. Fig. 7 demonstrates that using multiple organs gives a correct identification result even the results of each organ is incorrect.

We continue evaluating the performance of the proposed fusion schemes using Cumulative Match Characteristic curve (CMC), as shown in Fig. 8, Fig. 9, Fig. 10, Fig. 11. It measures the plant identification performances at various ranks. The better performance, the higher CMC is achieved. The higher CMCs are obtained with the most of the fusion schemes. The best CMC is achieved by a combination of Flower-Leaf with the RHF fusion.

To further evaluate advantages of the proposed fusion schemes, we attempt to find out the rank-k so that the identification accuracy reaches 99%. In this evaluation scenario, the fusion performances are better than those of single organ. The detailed results are given in Table. 6 and Table. 7. The RHF and product rule continue showing the significant performance compared with the results of other techniques. With leaf-flower combination, it can reach the accuracy 99% at rank-7 for product rule, or rank-9 for RHF in case of using AlexNet for single organ plant identification. ResNet allows to obtain the same accuracy at rank-4 in both product rule and RHF. It is much lower than the best case of using images from a single organ, where rank-29 is required.

Pistacia lentiscus L.	9.0e-01
Robinia pseudoacacia L.	1.0e-01
Fraxinus excelsior L.	1.8e-03
Fraxinus angustifolia Vahl	2.5e-04
Gleditsia triacanthos L.	4.5e-06

Daucus carota L.	7.4e-01
Robinia pseudoacacia L.	2.5e-01
Viburnum opulus L.	1.8e-03
Crataegus monogyna Jacq.	2.5e-04
Sambucus nigra L.	6.6e-09

Robinia pseudoacacia L.	1.0
Daucus carota L.	4.4e-07
Sambucus nigra L.	2.1e-08
Pistacia lentiscus L.	1.5e-08
Crataegus monogyna Jacq.	6.6e-09

Fig.7. Comparison of identification results using leaf, flower, and both leaf and flower images. The first column are query images. The second column shows top 5 species returned by the classifier. The third column is the. corresponding confidence score for each species. The name of species is Robinia pseudoacacia L.

Table 4. Obtained accuracy at rank-1 when combining each pair of organs with different fusion schemes in case of using AlexNet. The best result is in bold

		Scheme 1 for single organ identification					Scheme 2 for single organ identification				
Accuracy (%)		**Max rule**	**Sum rule**	**Product rule**	**SVM**	**RHF**	**Max rule**	**Sum rule**	**Product rule**	**SVM**	**RHF**
En-Le	R1	66.2	67.2	75.6	74.0	**76.6**	66.8	67.2	77.4	71.4	**78.6**
	R5	88.6	88.8	93.2	81.8	**94.6**	88.4	88.2	93.6	80.2	**94.4**
En-Fl	R1	73.8	74.4	78.8	77.2	**81.2**	73.84	73.6	78.8	76.24	**80.4**
	R5	92.6	92.8	94.2	84.2	**94.4**	88.8	89.2	94.8	83.6	**95.6**
Le-Fl	R1	81.6	82.0	88.6	86.2	**89.8**	78.8	81.2	89.6	83.2	**89.6**
	R5	96.8	96.8	98.2	90.4	**98.4**	95.6	96.0	99.2	88.8	**99.2**
Br-Le	R1	70.2	71.0	76.8	73.8	**78.4**	66.4	68.2	78.2	73.6	**78.2**
	R5	89.6	90.0	93.4	79.6	**93.8**	92.0	93.0	95.6	81.6	**96.0**
Br-Fl	R1	74.2	75.4	80.8	79.0	**81.4**	70.2	70.6	80.6	76.6	**81.4**
	R5	90.8	91.4	95.2	83.0	**95.4**	90.4	90.6	95.4	84.6	**95.6**
Br-En	R1	51.6	52.2	58.0	58.0	**58.6**	52.4	52.8	60.6	60.6	**61.6**
	R5	76.8	77.6	83.6	81.4	**83.8**	78.2	78.6	83.6	83.4	**84.9**

Table 5. Obtained accuracy at rank-1 when combining each pair of organs with different fusion schemes in case of using ResNet. The best result is in bold

		Scheme 1 for single organ identification					Scheme 2 for single organ identification				
Accuracy (%)		**Max rule**	**Sum rule**	**Product rule**	**SVM**	**RHF**	**Max rule**	**Sum rule**	**Product rule**	**SVM**	**RHF**
En-Le	R1	70.4	72.2	75.2	73.2	**78.0**	73.6	75.4	80.8	73.2	**80.8**
	R5	91.8	92.6	92.8	90.6	**93.2**	94.2	94.4	94.8	90.6	**95.2**
En-Fl	R1	73.8	75.4	80.0	76.4	**83.2**	74.6	76.0	80.2	76.4	**83.2**
	R5	93.2	93.6	95.0	89.2	**95.4**	94.4	95.0	**95.8**	89.2	95.2
Le-Fl	R1	90.0	91.4	92.4	91.4	**92.6**	85.8	87.6	89.2	91.4	**92.6**
	R5	98.0	98.8	99.0	96.0	**99.2**	98.4	98.4	99.0	96.0	**99.2**
Br-Le	R1	77.8	79.2	82.0	79.4	**83.2**	79.8	81.4	**83.6**	79.4	83.2
	R5	91.8	92.2	94.0	90.4	**94.6**	94.4	94.4	**96.4**	90.4	94.6

Br-Fl	R1	80.0	81.0	84.4	82.0	**86.4**	78.8	80.4	85.6	81.0	**86.0**
	R5	93.6	94.4	97.6	91.4	**97.8**	95.6	96.0	96.2	91.4	**97.6**
Br-En	R1	52.4	54.4	**62.2**	55.0	60.6	60.4	66.2	69.0	55.0	**69.0**
	R5	82.0	83.4	86.6	80.4	**87.4**	84.8	85.6	**89.6**	80.4	87.6

Table 6. Rank number (k) where 99% accuracy rate is achieved in case of using Alexnet. The best result is in bold

	Scheme 1 for single organ identification						Scheme 2 for single organ identification					
	En-Le	En-Fl	Le-Fl	Br-Le	Br-Fl	Br-En	En-Le	En-Fl	Le-Fl	Br-Le	Br-Fl	Br-En
Organ 1	42	42	27	46	46	46	48	48	30	47	47	47
Organ 2	27	29	29	27	29	42	30	25	25	30	25	48
Sum rule	17	24	10	21	25	**25**	30	30	16	24	21	33
Max rule	19	24	10	23	25	26	30	33	16	23	21	32
Product rule	16	20	7	22	**18**	**25**	**17**	20	**9**	**12**	14	**25**
SVM	50	50	50	50	50	50	50	50	50	50	50	50
RHF	**14**	**19**	9	**19**	**18**	**25**	**17**	**19**	19	17	**12**	**25**
MCDCNN[1]	24	29	12	20	**18**	33	24	29	12	20	18	33

Table 7. Rank number (k) where 99% accuracy rate is achieved in case of using ResNet. The best result is in bold

	Scheme 1 for single organ identification						Scheme 2 for single organ identification					
	En-Le	En-Fl	Le-Fl	Br-Le	Br-Fl	Br-En	En-Le	En-Fl	Le-Fl	Br-Le	Br-Fl	Br-En
Organ 1	44	44	29	38	38	38	41	41	30	37	37	37
Organ 2	29	15	15	29	15	44	30	29	29	30	29	41
Sum rule	**16**	14	5	19	13	**23**	22	33	**8**	19	20	**27**
Max rule	17	15	7	19	14	25	22	33	9	20	19	29
Product rule	18	**11**	**4**	**16**	10	26	**18**	33	10	**14**	19	29
SVM	50	50	50	50	50	50	50	50	50	50	50	50
RHF	19	13	**4**	**16**	**8**	26	25	**28**	12	22	**17**	30

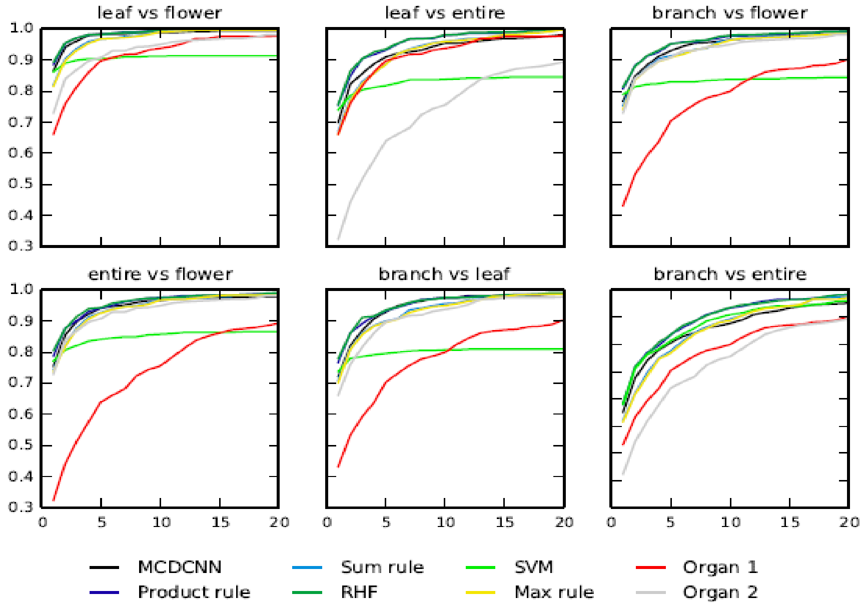

Fig. 8. Cumulative Match Characteristic curve obtained by the proposed method with AlexNet (Scheme 1 for single organ identification).

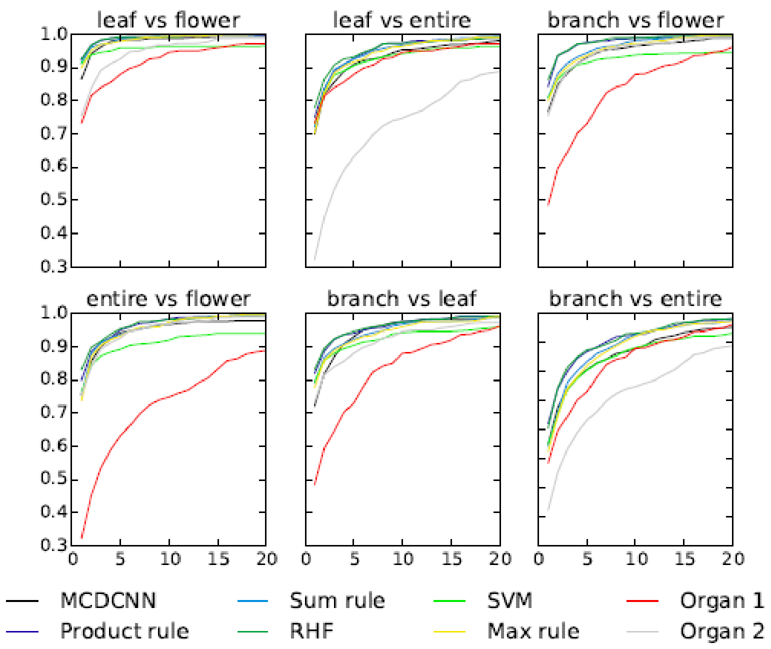

Fig. 9. Cumulative Match Characteristic curve obtained by the proposed method with ResNet (Scheme 1 for single organ identification).

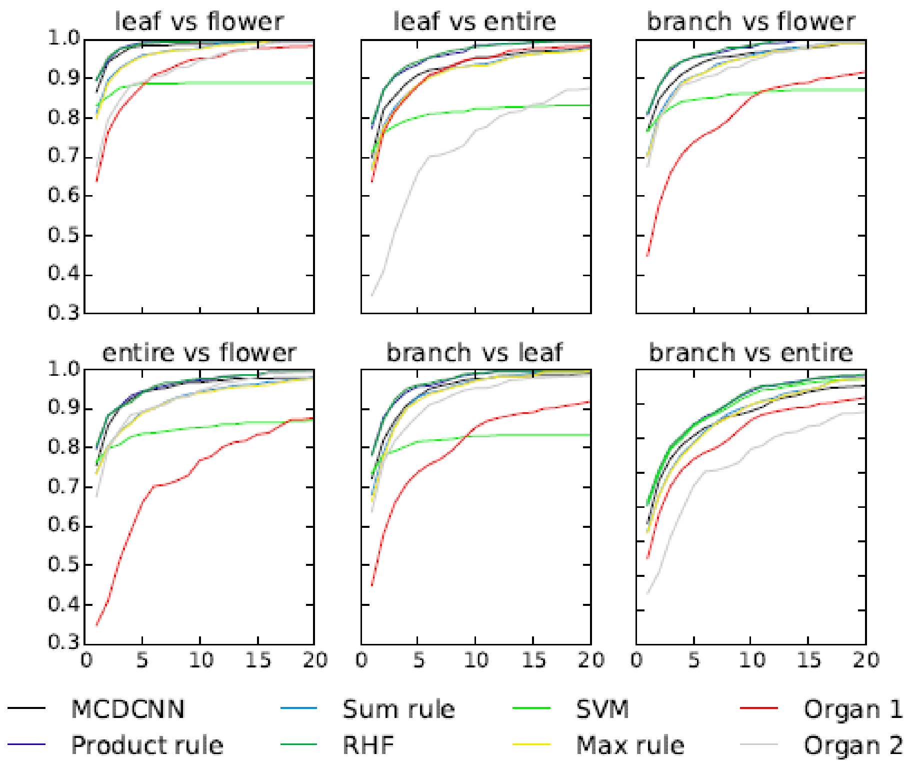

Fig. 10. Cumulative Match Characteristic curve obtained by the propsed method with AlexNet (Scheme 2 for single organ identification).

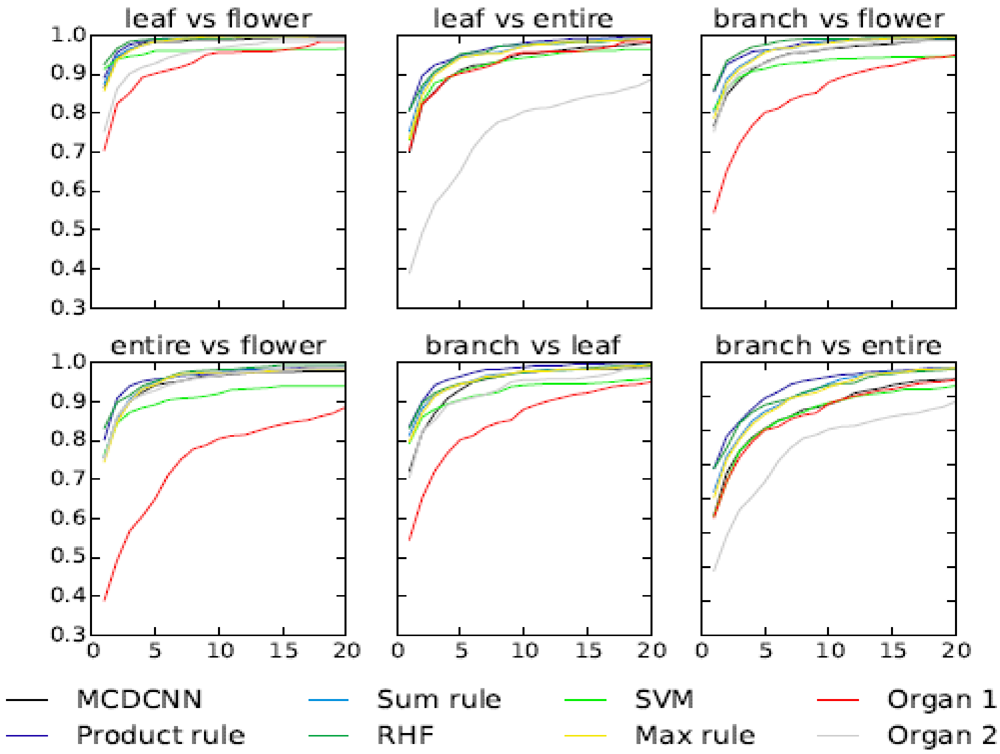

Fig. 11. Cumulative Match Characteristic curve obtained by the proposed method with ResNet (Scheme 2 for single organ identification).

6.3.3. Comparison to MCDCNN

To show the effectiveness of the proposed fusion scheme, we compare its performance with the performance of MCDCNN [1]. Since the implementation of MCDCNN is unavailable, we implement this network as described in [1]. MCDCNN has four columns where each column is a CNN with the pre-trained parameters from the ImageNet dataset and it stands for an organ (i.e. flower, leaf, branch, and entire). The CNN chosen in MCDCNN is AlexNet. MCDCNN combines all features at the last layer of each column to a feature vector before using a fully connected layer for species classification. On the preparation of training data for the model, with each pair of organs, we produce all pairs of images, where each organ has one image from the training set. For example, following the statistics of the dataset in

Table.1, there are $1,650 \times 1,930$ = 3,184,500 pair of flower and leaf images. Concerning the structure, the proposed method and that proposed in [1] shares one common point that is the Scheme 1 for single organ plant identification. However, two methods differ in the most important part: fusion technique for determining the result of plant identification from those obtained with single organs. In the proposed method, the fusion is done on confidence score after the images passed through the CNN model while in the MCDNN [1] the combining is done at Fully Connected Layers. Furthermore, besides Scheme 1 and AlexNet architecture, in the proposed method, we investigate the performance of both schemes (Scheme 1 and Scheme 2) with two networks (AlexNet and ResNet). The obtained results on the same dataset in Table 8 show that the proposed method outperforms MCDCNN in all combinations. The improvement is up to 10% for the combination of branch and leaf. The source codes and evaluation datasets are available at[1].

7. Conclusions

This paper examined two schemes for single organ plant identification as well as

[1] http://mica.edu.vn/perso/Le-Thi-Lan/jcsce.html

several fusion schemes for the plant identification task using multi-organ images. The experiments show that the fusion techniques increase the performances dramatically. Also, the robust hybrid fusion model presents the best result in all evaluations. It obtains from + 3.0% to + 13.8% improvement in rank-1 over MCDCNN method. In future work, we attempt to investigate a method to identify species for observations with an unfixed number of organs and number of images in each organ.

Acknowledgments

The authors thank Collaborative Research Program for Common Regional Issue (CRC) funded by ASEAN University Network (Aun-Seed/Net), under the grant reference HUST/CRC/1501 and thank Assoc. Prof Van-Sam Hoang, an expert in botany from Vietnam University of Forestry for helping us in validating the working dataset.

References

[1] A. He, X. Tian, Multi-organ plant identification with multi-column deep convolutional neural networks, in: 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2016, pp. 002020-002025. doi:10.1109/SMC.2016.7844537.

[2] S. G. Wu, F. S. Bao, E. Y. Xu, Y.-X. Wang, Y.-F. Chang, Q.-L. Xiang, A leaf recognition algorithm for plant classification using probabilistic neural network, in: Signal Processing and Information Technology, 2007 IEEE International Symposium on, IEEE, 2007, pp. 11-16.

[3] X. Gu, J.-X. Du, X.-F. Wang, Leaf recognition based on the combination of wavelet transform and gaussian interpolation, Advances In Intelligent Computing (2005) 253-262.

[4] J. Chaki, R. Parekh, Plant leaf recognition using shape based features and neural network classifiers, International Journal of Advanced Computer Science and Applications 2 (10).

[5] K. Lunau, Innate flower recognition in bumblebees (bombus terrestris, b. lucorum; apidae): optical signals from stamens as landing reaction releasers, Ethology 88 (3) (1991) 203-214.

[6] H. Goëau, P. Bonnet, A. Joly, LifeCLEF Plant Identification Task 2015, in: CEUR-WS (Ed.), CLEF: Conference and Labs of the Evaluation forum, Vol. 1391 of CLEF2015 Working notes, Toulouse, France, 2015. URL https://hal.inria.fr/hal-01182795

[7] H. Goëau, P. Bonnet, A. Joly, V. Bakic, D. Barth´el´emy, N. Boujemaa, J.-F. Molino, The imageclef 2013 plant identification task, in: CLEF, 2013.

[8] H. Goëau, A. Joly, P. Bonnet, S. Selmi, J.-F. Molino, D. Barth´el´emy, N. Boujemaa, Lifeclef plant identification task 2014, in: CLEF2014 Working Notes. Working Notes for CLEF 2014 Conference, She_eld, UK, September 15-18, 2014, CEUR-WS, 2014, pp. 598-615.

[9] H. Goëau, P. Bonnet, A. Joly, Plant identification in an open-world (lifeclef 2016), CLEF working notes 2016 (2016) 428-439.

[10] H. Goëau, P. Bonnet, A. Joly, Plant identification based on noisy web data: the amazing performance of deep learning (lifeclef 2017), CEUR Workshop Proceedings, 2017.

[11] A. A-ouard, H. Goëau, P. Bonnet, J.-C. Lombardo, A. Joly, Pl@ntnet app in the era of deep learning, in: ICLR 2017 Workshop Track-5th International Conference on Learning Representations, 2017.

[12] H. Zhu, X. Huang, S. Zhang, P. C. Yuen, Plant identification via multipath sparse coding, Multimedia Tools and Applications 76 (3) (2017) 4599-4615.

[13] S. H. Lee, Y. L. Chang, C. S. Chan, Lifeclef 2017 plant identification challenge: Classifying plants using generic-organ correlation features, Working Notes of CLEF 2017.

[14] T. T.-N. Nguyen, T.-L. Le, H. Vu, H.-H. Nguyen, V.-S. Hoang, A combination of deep learning and hand-designed feature for plant identification based on leaf and flower images, in: Advanced Topics in Intelligent Information and Database Systems, Springer, 2017, pp. 223-233.

[15] T. B. Do, H. H. Nguyen, T. T. N. Nguyen, H. Vu, T. T. H. Tran, T. L. Le, Plant identification using score-based fusion of multi-organ images, in: 2017 9th International Conference on Knowledge and Systems Engineering (KSE), 2017, pp. 191-196. doi:10.1109/KSE.2017.8119457.

[16] Y. Makihara, D. Muramatsu, H. Iwama, Y. Yagi, On combining gait features, in: Automatic Face

and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops on, IEEE, 2013, pp. 1-8.
[17] J. S. Cope, D. Corney, J. Y. Clark, P. Remagnino, P. Wilkin, Plant species identification using digital morphometrics: A review, Expert Systems with Applications 39 (8) (2012) 7562-7573.
[18] T. Beghin, J. Cope, P. Remagnino, S. Barman, Shape and texture based plant leaf classification, in: Advanced Concepts for Intelligent Vision Systems, Springer, 2010, pp. 345-353.
[19] A. N. Hussein, Content-based image retrieval system for plant leaf database using texture, Ph.D. thesis, Universiti Putra Malaysia (2011).
[20] E. Aptoula, B. Yanikoglu, Morphological features for leaf based plant recognition, in: Image Processing (ICIP), 2013 20th IEEE International Conference on, IEEE, 2013, pp. 1496-1499.
[21] B. Yanikoglu, E. Aptoula, C. Tirkaz, Automatic plant identification from photographs, Machine vision and applications 25 (6) (2014) 1369-1383.
[22] H. Goeau, P. Bonnet, J. Barbe, Multi-organ plant identification, in: 2012 ACM international workshop on Multimedia analysis for ecological data, 2012, pp. 41-44.
[23] J. Waldchen, P. Mader, Plant species identification using computer vision techniques: A systematic literature review, Archives of Computational Methods in Engineering (2017) 1-37.
[24] N. Kumar, P. Belhumeur, A. Biswas, D. Jacobs, W. Kress, I. Lopez, J. Soares, Leafsnap: A computer vision system for automatic plant species identification, Computer Vision–ECCV 2012 (2012) 502-516.
[25] T.-L. Le, N.-D. Duong, V.-T. Nguyen, H. Vu, V.-N. Hoang, T. T.-N. Nguyen, Complex background leaf-based plant identification method based on interactive segmentation and kernel descriptor, in: Proceedings of the 2nd International Workshop on Environmental Multimedia Retrieval, ACM, 2015, pp. 3-8.
[26] H.-J. Yoo, Deep convolution neural networks in computer vision, IEIE Transactions on Smart Processing & Computing 4 (1) (2015) 35-43.
[27] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in neural information processing systems, 2012, pp. 1097-1105.
[28] Q. Chen, M. Abedini, R. Garnavi, X. Liang, Ibm research australia at lifeclef2014: Plant identification task., in: CLEF (Working Notes), 2014, pp. 693-704.
[29] M. M. Ghazi, B. Yanikoglu, E. Aptoula, Plant identification using deep neural networks via optimization of transfer learning parameters, Neurocomputing 235 228-235.
[30] S. H. Lee, C. S. Chan, P. Wilkin, P. Remagnino, Deep-plant: Plant identification with convolutional neural networks, in: Image Processing (ICIP), 2015 IEEE International Conference on, IEEE, 2015, pp. 452–456.
[31] G. Cerutti, L. Tougne, C. Sacca, T. Joliveau, P.-O. Mazagol, D. Coquin, A. Vacavant, Late information fusion for multi-modality plant species identification, in: Working notes for Conference and Labs of the Evaluation Forum, 2013.
[32] J. W¨aldchen, P. M¨ader, Plant species identification using computer vision techniques: A systematic literature review, Archives of Computational Methods in Engineering 25 (2) (2018) 507-543. doi:10.1007/s11831-016-9206-z. URL https://doi.org/10.1007/s11831-016-9206-z
[33] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al., Imagenet large scale visual recognition challenge, International Journal of Computer Vision 115 (3) (2015) 211-252.
[34] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770-778.
[35] G. Van Horn, O. Mac Aodha, Y. Song, Y. Cui, C. Sun, A. Shepard, H. Adam, P. Perona, S. Belongie, The inaturalist species classification and detection dataset (2018) 8769-8778.