Dariusz Król · Ngoc Thanh Nguyen
Kiyoaki Shirai
Editors

# Advanced Topics in Intelligent Information and Database Systems

Springer

*Editors*
Dariusz Król
Faculty of Computer Science and Management
Wrocław University of Science and Technology
Wrocław
Poland

Kiyoaki Shirai
School of Advanced Science and Technology
Japan Advanced Institute of Science and Technology
Nomi, Ishikawa
Japan

Ngoc Thanh Nguyen
Faculty of Computer Science and Management
Wrocław University of Science and Technology
Wrocław
Poland

and

Division of Knowledge and System Engineering for ICT
Faculty of Information Technology
Ton Duc Thang University
Ho Chi Minh City
Vietnam

ISSN 1860-949X ISSN 1860-9503 (electronic)
Studies in Computational Intelligence
ISBN 978-3-319-56659-7 ISBN 978-3-319-56660-3 (eBook)
DOI 10.1007/978-3-319-56660-3

Library of Congress Control Number: 2017935835



Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

# A Combination of Deep Learning and Hand-Designed Feature for Plant Identification Based on Leaf and Flower Images

**Thi Thanh-Nhan Nguyen, Thi-Lan Le, Hai Vu, Huy-Hoang Nguyen and Van-Sam Hoang**

**Abstract** This paper proposes a combination of deep learning and hand-designed feature for plant identification based on leaf and flower images. The contributions of this paper are two-fold. First, for each organ image, we have performed a comparative evaluation of deep learning and hand-designed feature for plant identification. Two approaches for deep learning and hand-designed feature that are convolutional neuron network (CNN) and kernel descriptor (KDES) are chosen in our experiments. Second, based on the results of the first contribution, we propose a method for plant identification by late fusing the identification results of leaf and flower. Experimental results on ImageClef 2015 dataset show that hand designed feature outperforms deep learning for well-constrained cases (leaf captured on simple background). However, deep learning shows its robustness in natural situations. Moreover, the combination of leaf and flower images improves significantly the identification when comparing leaf-based plant identification.

**Keywords** CNN · KDES · Plant identification

## 1 Introduction

Plant identification aims at determining the name of species based on observations. This task is time consuming and difficult even for the botanist experts. Recently, the advanced research in computer vision community allows building automatic plant

T. Thanh-Nhan Nguyen (✉) · T.-L. Le · H. Vu · H.-H. Nguyen
International Research Institute MICA, HUST-CNRS/UMI-2954-GRENOBLE INP, Hanoi University of Science and Technology, Hanoi, Vietnam
e-mail: nttnhan@ictu.edu.vn

T. Thanh-Nhan Nguyen
Thai Nguyen University of Information and Communication Technology, Thai Nguyen, Vietnam

V.-S. Hoang
Vietnam Forestry University, Hanoi, Vietnam



D. Król et al. (eds.), *Advanced Topics in Intelligent Information and Database Systems*, Studies in Computational Intelligence 710,
DOI 10.1007/978-3-319-56660-3_20

identification based on images analysis. However, a large number of automatic plant identification methods still work with only one sole organ of the plants. Among different organs of the plant, leaf and flower are the most widely used. Leaf is usually flat and easy to collect almost the whole year while flower has a high distinguishing capacity. However, using one sole organ for plant identification is not always relevant because one organ cannot fully reflect all information of a plant due to the large inter-class similarity and large intra-class variation.

Recently, there are two new trends in plant identification. First, for each organ-based plant identification, instead of using hand-designed feature, learning feature has been investigated more and more [1, 2]. Second, the plant identification moves from one sole organ to multi-organ [3]. This is motivated by a real scenario where one user tries to identify a plant by observing its different organs. This also can reflect the process of plant identification of botanists. Observing simultaneously several organs allows the botanists to disambiguate species that could be confused when using only one organ.

This paper has two main contributions. First, for each organ image, we have performed a comparative evaluation of deep learning and hand-designed feature for plant identification. Two approaches for deep learning and hand-designed feature that are convolutional neuron network (CNN) and kernel descriptor (KDES) are chosen in our experiments. Second, based on the results of the first contribution, we propose a method for plant identification by late fusing the identification result of leaf and flower.

## 2 Related Work

There are two main approaches for plant identification. The first one uses hand-designed feature while the second one employs the deep learning strategy. A number of hand-crafted (or hand-designed) features have been used for plant identification, such as edge features [4], KDES [5] and shape properties [6]. The hand-designed feature approach requires knowledge about the application. Moreover, approaches using hand-crafted features also require pre-processing techniques. For the leaf-based images, LeafSnap is a noticeable application that achieved more than 80% accuracy rate. For the flower-based images, by using a dataset of 120 species, [7] obtains the recognition rate of 76.3%. Recently, convolutional neural network obtains state-of-the-art results in many computer vision applications [8, 9]. Typically CNNs are AlexNet, VGG, GoogLeNet and ResNet. Some studies try to apply CNN such as [1] to learn unsupervised feature representations for 44 different plant species collected at the Royal Botanic Gardens, Kew, England. However, it still lacks a detail analysis and comparative evaluation of hand-designed feature and deep learning for plant identification.

Recently, the plant identification has been expanded with multi-organ/multi-images such as leaf-scan, leaf, flower, fruit so that the obtained results are better. The common way to do it is to combine recognition results by late-fusion techniques.

Such kind of approaches are listed in technical report of the LifeCLEF competitions 2015 [3]. However, to the best of our knowledge, there is no plant identification technique combining two most important organs which are leaves and flowers, to achieve better performance for the plant identification.

# 3 Proposed Method

## *3.1 Overview*

The overview of plant identification based on of multi-organ images is illustrated in Fig. 1. The system consists of three main components: Preprocessing, Single organ-based plant identification and Multiple organ-based plant identification. In our work, we focus on two most important organs of the plant that are leaf and flower.

- Preprocessing: depending on the characteristic of the image, we propose corresponding preprocessing techniques.
- Single organ-based plant identification: In this component, we propose to use two approaches: hand-designed feature and learning feature. For hand-designed feature, we employ KDES with multi-class SVM (Support Vector Machine) while for learning feature approach, we propose to use GoogLeNet because of its remarkable result for object recognition.
- Multiple organ-based plant identification: we perform late fusion with Sum rule technique to combine results of single organ-based plant identification.

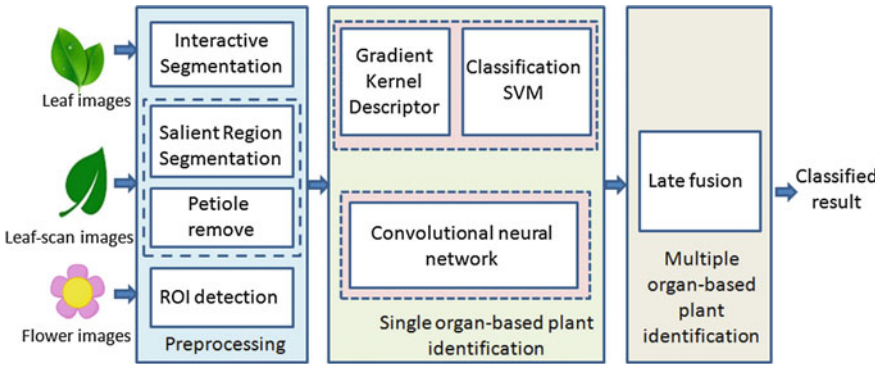


**Fig. 1** Overview of multiple organ-based plant identification

## 3.2 Preprocessing

The images of object of interest (e.g. leaf and flower) can be captured in different conditions. In our work, we use ImageClef2015 dataset [3]. In this dataset, leaf images are divided into two categories: leaf (leaf image captured on natural background) and leaf-scan (leaf image captured on simple background) while flower images are normally captured in natural background.

For leaf images, an interactive segmentation method [5] based on Watershed algorithm is first applied to segment leaf from background regions. Then, the main direction of the leaves are normalized based on moment calculations.

With leaf-scan images, a saliency extraction method and a common segmentation technique (e.g., mean-shift algorithm) [10] are employed. A segmented region is selected based on a condition that its saliency value is large enough. The connected-region techniques then are applied to connect the selected regions. Since leaf-scan images usually contain simple background, we can obtain relatively good segmentation results for leaf-scan images.

Concerning flower images, as flower images are normally taken in complex background, it is not easy to have stable and good segmentation results. Therefore, saliency-segmentation procedure that is similar for leaf-scan image is chosen for determining ROI (Region Of Interest) from flower images. Flower ROI is determined by top-left and bottom-right points on boundary of the connected-regions. Detail information on these techniques can be found in [11]. Some examples of leaf, leaf-scan and flower images after applying the preprocessing techniques are illustrated in Fig. 2.

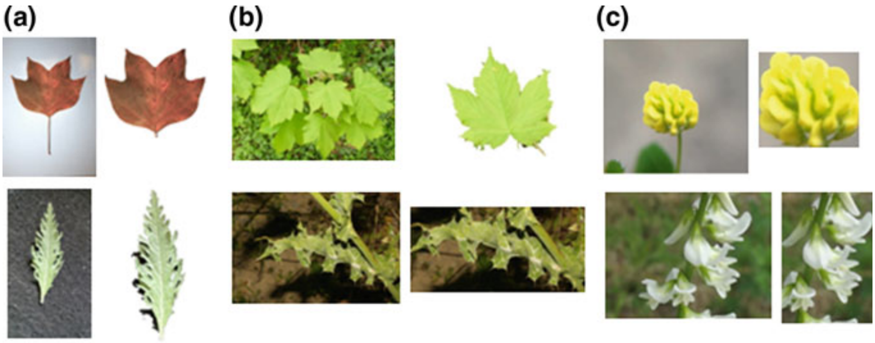


**Fig. 2** Examples of leaf-scan, leaf and flower images after applying corresponding pre-processing techniques

### 3.3 Hand-Designed Feature for Single Organ Identification

Kernel descriptor (KDES) [12] has been applied in our previous work for leaf-based plant identification [5]. In this paper, we compare the performance of this feature with an effective deep neural network that is convolutional neural network. KDES is extracted through 3 steps: pixel-level feature, patch-level feature and image-level feature extraction.

**Pixel-level features extraction**

For each pixel z, its gradient vector consists of two components: the magnitude m(z) and the orientation $\theta(z)$ where the orientation vector is defined as $\tilde{\theta}(z) = [\sin(\theta(z))\ \cos(\theta(z))]$.

**Patch-level features extraction**

The patch-level feature is extracted through two steps. The first step aims at generating a set of patches from image while the second one allows computing patch feature. Derived from match kernel representing the similarity of two patches, we can extract feature vector for the patch using approximate patch-level feature map, given a designed patch level match kernel function. The gradient match kernel, defined in Eq. 1, is formed from three kernels that are gradient magnitude kernel $k_{\tilde{m}}$, orientation kernel $k_o$ and position kernel $k_p$. These kernels are defined in [12].

$$K_{gradient}(P, Q) = \sum_{z \in P} \sum_{z' \in Q} k_{\tilde{m}}(z, z') k_o(\tilde{\theta}(z), \tilde{\theta}(z')) k_p(z, z') \tag{1}$$

where P and Q are patches of two different images that we need to measure the similarity. $z$ and $z'$ denote the position of a pixel in the image patch P and Q respectively. $\tilde{\theta}(z)$ and $\tilde{\theta}(z')$ are gradient orientations at pixel $z$ and $z'$ in the patch P and Q respectively. Then, the approximative feature over image patch P is constructed as:

$$\bar{F}_{gradient}(P) = \sum_{z \in P} \tilde{m}(z) \phi_o(\tilde{\omega}(z)) \otimes \phi_p(z) \tag{2}$$

where $\otimes$ the Kronecker product, $\tilde{m}(z)$ is normalized $m(z)$, $\phi_o(\tilde{\omega}(z))$ and $\phi_p(z)$ are approximate feature maps for the kernel $k_o$ and $k_p$, respectively. The basic idea of representation based on kernel methods is to compute approximate explicit feature map for kernel function.

**Image-level features extraction**

After extracting patch-level feature, the feature of whole image will be computed as follows. First, a spatial pyramid is built by dividing the image into cells at several layers. Then, the feature map is defined as:

$$\bar{\phi}_S(X) = \frac{1}{|X|} \sum_{x \in X} \phi(x) \tag{3}$$

where C is a cell that has a set of patch-level features X = $\{x_1 \ldots x_p\}$ and $\phi(x)$ is approximative feature maps for the kernel k(x, y).

After computing feature at image level, we apply multi-class SVM as classifier.

### 3.4 Deep Learning for Single Organ Identification

GoogLeNet of Szegedy et al. that won the classification and object recognition challenges in the ILSVRC 2014 [13] used a new variant of convolutional neural network called "Inception" for classification with the intention of increasing network depth with computational efficiency. Figure 3 shows a schematic view of GoogLeNet network. GoogLeNet is a very deep neural network model with 22 layers when counting only layers with parameters (or 27 layers if we also count pooling). In this work, we fine-tune GoogLeNet with plant flower and leaf images with the following parameters: initial learning rate: 0.001; batch size: 5; number of iteration: 50,000. GoogLeNet output score is employed to produce the ranked list of relevant plant species.

### 3.5 Result Fusion

We combine identification results from leaf and flower image as follows:

$$score(q, species) = \sum_{i=1}^{N} score(I_i, species) \quad (4)$$

where q is query observation, N is number of images of this observation, *score* $(I_i, species)$ is obtained similarity score when using image $I_i$ of the observation. Note that the number of images of each observation can be different.

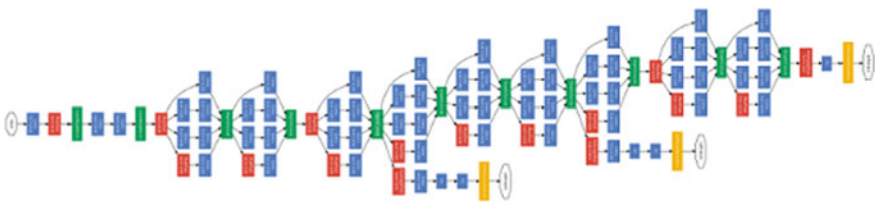

**Fig. 3** A schematic view of GoogLeNet network (adapted from [13])

# 4 Experimental Results

## 4.1 Dataset

In our experiment, we use a large dataset of plant named PlantCLEF 2015 [3]. This dataset has more than one hundred thousand images belonging to 41,794 observations of 1000 plants species living in West Europe. One observation is a set of images depicting the same individual plant, observed by the same person in the same day with the same device. For each observation, images of seven different organs including leaf, leaf-scan (leaf image captured on simple background), flower, fruit, stem, entire plant and branch are captured. In this paper, we use leaf including leaf and leaf-scan and flower images. Table 1 gives detail information of training and testing set. Some examples of the flower and leaf images are illustrated in Fig. 4.

**Table 1** Training and testing sets provides by PlantCLEF2015

| | Leaf | Leafscan | Flower | Total |
|---|---|---|---|---|
| Train | 8,885 | 13,198 | 27,975 | 50,058 |
| Test | 2,690 | 221 | 8,327 | 11,238 |
| All | 11,575 | 13,419 | 36,302 | 61,296 |

**Fig. 4** Examples of flower and leaf with simple and complicated background images on PlantCLEF 2015

### 4.2 Evaluation Measures

We evaluate our proposed approach at two different levels: single-organ based plant identification named image level and multi-organs plant identification named observation level. For this, we compute the accuracy at rank k at image and at observation level as follows: $Accuracy = \frac{T}{N}$ where T and N are the number of correct recognition and the number of queries respectively. One image or one observation is considered as correct recognition if the relevant plant belongs to the k first plants of the retrieved list. In our experiments, accuracy at rank 1 (k = 1) and rank 10 (k = 10) are used.

### 4.3 Results and Discussions

Our system is implemented in C++, Matlab with the use of three libraries that are OpenCV, KDES and Caffe.

**Comparative evaluation for single organ identification** In order to compare the performance of KDES and CNN, we have performed two experiments. The aim of the first experiment is to evaluate KDES and GoogLeNet on the preprocessed images. Accuracies obtained at Rank 1 and Rank 10 are shown in Table 2. With leaf and flower images, result of GoogLeNet is much better than KDES. However, with leaf-scan, KDES outperforms GoogLeNet. This result allows recommending using hand-designed feature if the working images are images of leaf captured on a well-constraint condition. Otherwise, deep learning is a good choice.

The effect of preprocessing step for deep learning is analyzed in the second experiment. For this, we compare accuracy of GoogleNet of raw and preprocessed database. Table 3 shows results of GoogLeNet with and without applying preprocessing technique. For leaf and leafscan, the result on raw images is lower than that on preprocessed images. This means that preprocessing is important step to improve identification accuracy. However, with flower images, the obtained results on raw images and preprocessed images are relatively similar. This shows the capacity of deep learning in recognizing objects of nature scene.

**Table 2** Comparison of hand-designed feature (KDES) and deep learning (GoogleNet) for plant identification at image level on preprocessed images

| | Leaf | | LeafScan | | Flower | |
|---|---|---|---|---|---|---|
| | Rank 1 | Rank 10 | Rank 1 | Rank 10 | Rank 1 | Rank 10 |
| KDES | 24.26 | 46.36 | **78.28** | **92.76** | 10.95 | 24.62 |
| GoogLeNet | **47.30** | **72.70** | 69.78 | 86.22 | **66.60** | **90.23** |

**Table 3** Accuracy of plant identification obtained with GoogleNet at image level on images with and without preprocessing

| | Leaf | | Leafscan | | Flower | |
|---|---|---|---|---|---|---|
| | Rank 1 | Rank 10 | Rank 1 | Rank 10 | Rank 1 | Rank 10 |
| Without preprocessing | 35.39 | 62.79 | 67.56 | 83.11 | 67.45 | 90.82 |
| With preprocessing | **47.30** | **72.70** | **69.78** | **86.22** | 66.60 | 90.23 |

**Table 4** Accuracy for plant identification based on only leaf, and combining leaf and flower

| | Accuracy (%) | |
|---|---|---|
| | Rank 1 | Rank 10 |
| Only leaf | 33.28 | 68.81 |
| Only flower | 61.67 | 86.76 |
| Fusion of leaf and flower | **64.46** | **90.77** |

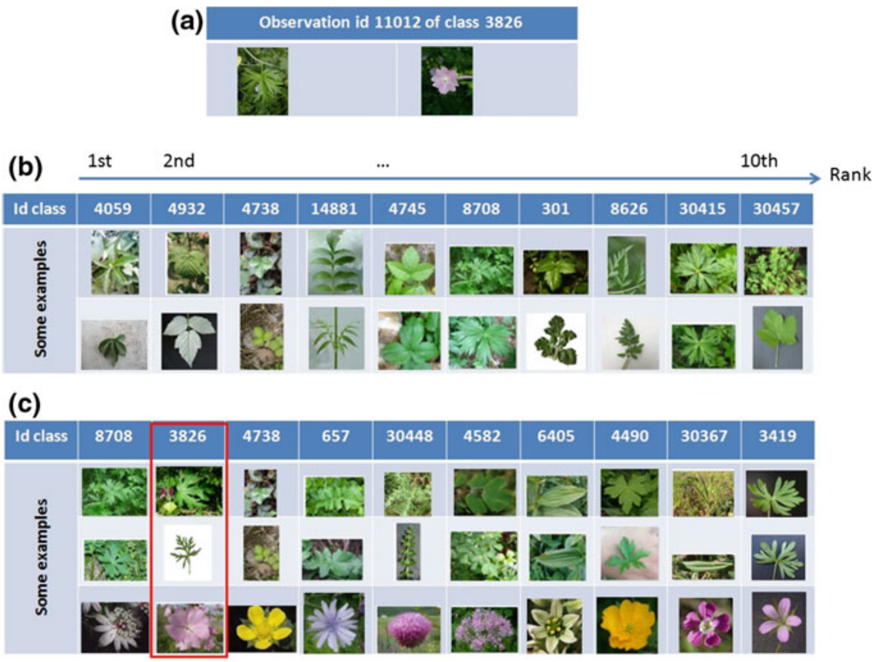


**Fig. 5** Example of multiple organ-based plant identification **a** Input: observation 11012 of class 3826 having two images: one image of leaf and one of flower, **b** 10 first retrieved results when using only leaf, **c** 10 first retrieved results when combining leaf and flower images. The correct species is bounded by a *red* rectangle

**Plant identification result by combining leaf and flower** From testing dataset, we extracted 574 observations that contain images of both flower and leaf in order to evaluate multiple organs plant identification. The obtained results are shown in Table 4. We can observe that when combining flower and leaf images, the accuracy increases significantly for both Rank 1 and Rank 10. The result also shows that flower is a distinguishing organ for plant identification. However, this organ is not always available. An example of retrieval results for observation 11012 is illustrated in Fig. 5. This observation has two images: one of leaf and one of flower (see Fig. 5a). Figure 5b, c show the 10 first retrieved results sorted in descending order of confidence score when using only leaf and fusing leaf and flower respectively. We may observe that due to the large inter-class similarity of leaf images, this observation is not correct identified. However, when adding information of flower, the correct plant is returned at second rank.

## 5 Conclusions

In this paper, we have performed a comparative evaluation of two approaches for plant identification. The obtained results show that for the well-constrained case, hand-designed feature outperforms deep learning. However, in weak-constrained situation, deep learning seems to show its robustness. Then, based on the evaluation, we have proposed a plant identification method combining leaf and flower images. The obtained results are promising. However, the fusion technique is still simple and does not take into account the role of each organ. In the future, we will focus on developing more robust combination of these organs.

**Acknowledgements** 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.

## References

1. Lee, S.H., Chan, C.S., Wilkin, P., Remagnino, P.: Deep-plant: plant identification with convolutional neural networks. In: 2015 IEEE International Conference on Image Processing (ICIP), pp. 452–456, IEEE (2015)
2. Choi, S.: Plant identification with deep convolutional neural network: Snumedinfo at lifeclef plant identification task 2015. In: Working Notes of CLEF 2015 Conference (2015)
3. Goëau, H., Bonnet, P., Joly, A.: lifeclef plant identification task 2015. In: CEUR-WS, ed. CLEF: Conference and Labs of the Evaluation forum. Volume 1391 of CLEF2015 Working Notes, Toulouse, France, Sept (2015)
4. Kumar, N., Belhumeur, P.N., Biswas, A., Jacobs, D.W., Kress, W.J., Lopez, I.C., Soares, J.V.: Leafsnap: a computer vision system for automatic plant species identification. In: Computer Vision–ECCV 2012, pp. 502–516. Springer (2012)
5. Le, T.L., Duong, N.D., Nguyen, V.T., Vu, H., Hoang, V.N., Nguyen, T.T.N.: 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, pp. 3–8. ACM (2015)
6. Cope, J.S., Corney, D., Clark, J.Y., Remagnino, P., Wilkin, P.: Plant species identification using digital morphometrics: a review. Expert Syst. Appl. **39**(8), 7562–7573 (2012)
7. Nilsback, M.E., Zisserman, A.: A visual vocabulary for flower classification. In: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), vol. 2, pp. 1447–1454. IEEE (2006)
8. Yoo, H.J.: Deep convolution neural networks in computer vision. IEIE Trans. Smart Process. Comput. **4**(1), 35–43 (2015)
9. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. Adv. Neural Inf. Process. Syst. 1097–1105 (2012)
10. Achanta, R., Hemami, S., Estrada, F., Susstrunk, S.: Frequency-tuned salient region detection. In: IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2009, pp. 1597–1604. IEEE (2009)
11. Le, T.L., Dng, D., Vu, H., Nguyen, T.N.: Mica at lifeclef 2015: multi-organ plant identification. In: Working Notes of CLEF 2015 Conference (2015)
12. Bo, L., Ren, X., Fox, D.: Kernel descriptors for visual recognition. In: Adv. Neural Inf. Process. Syst. 244–252 (2010)
13. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–9 (2015)