# Plant identification with deep convolutional neural network: SNUMedinfo at LifeCLEF plant identification task 2015

Sungbin Choi

Department of Biomedical Engineering, Seoul National University, Republic of Korea

wakeup06@empas.com

**Abstract.** This paper describes our participation at the LifeCLEF Plant identification task 2015. Given various images of plant parts such as leaf, flower or stem, this task is about identification of plant species given multi-image observation query. We utilized GoogLeNet for individual image classification, and combined image classification results for plant identification per observation. Our approach achieved best performance in this task.

**Keywords:** Image classification, Deep convolutional neural network, GoogLeNet, Borda-fuse

## 1 Introduction

In this paper, we describe the participation of the SNUMedinfo team at the LifeCLEF Plant Identification task 2015. Each query is composed of multi-image observation, which represents individual plant observed the same day by a same person. Each observation has multiple image, taken from various parts of plant such as leaf, stem or flower. So this task is about identification of plant species given multi-image observation query. For a detailed introduction of the task, please see the overview paper of this task (1).

In recent years, deep Convolutional Neural Network (CNN) has improved automatic image classification performance dramatically (2). In this study, we experimented with GoogLeNet (3) which has shown effective performance in recent ImageNet Challenge (4). Although LifeCLEF Plant identification task is about more fine-grained image classification compared to ImageNet's general object category classification, finetuning CNN pretrained on ImageNet dataset was very effective in performance. Our experimental methods are detailed in the next section.

## 2 Methods

We applied CNN for individual image classification (Section 2.1). Then image classification results are combined to produce observation classification (Section 2.2).

### 2.1 Image classification using deep convolutional neural network

**Finetuning from GoogLeNet**

We utilized GoogLeNet for individual plant image classification. GoogLeNet incorporates Inception module with the intention of increasing network depth with computational efficiency.

We randomly divided observations in LifeCLEF Plant identification training set into five-fold. Images from one fold is used as validation set, and images from other four fold is used as training set.

Training CNN for plant identification started from GoogLeNet pretrained on ImageNet dataset. We finetuned CNN on plant identification training set (initial learning rate 0.001; batch_size:120; number of iteration:100,000). Only horizontal mirroring (left-right flipping of image) and image random cropping (cropping 224 x 224 image out of 256 x 256 input image) is used for data augmentation.

We trained five separate CNNs[1]. CNN output score is used to produce ranked list of relevant plant species. Five ranked list is combined into single ranking using Borda-fuse method (5).

### 2.2 Observation classification by combining image classification result

Each query observation is composed of multiple image. We combined image classification result from Section 2.1 using two different rank aggregation method.

(1) Borda-fuse method
(2) Majority voting based method

## 3 Results

We submitted four different runs. Details of runs are summarized in the following table.

**Table 1.** Different setting of submitted runs

| | **Image classification** | **Observation classification** |
|---|---|---|
| **SNUMedinfo1** | Only 1 CNN is used[2] | Borda-fuse |
| **SNUMedinfo2** | Only 1 CNN is used | Majority voting based method |
| **SNUMedinfo3** | 5 CNNs are used | Borda-fuse |
| **SNUMedinfo4** | 5 CNNs are used | Majority voting based method |

1 We arbitrarily determined number of CNN classifier for experiment as five. In this study, we tried to assess the effects on performance when more CNNs are trained and their results are combined.

2 Among five trained CNNs, only one CNN is used for classification.

Primary evaluation metric for this task was average classification score. Inverse of the rank of the correct species are scored between 0 and 1, and then it is macro-averaged over distinct user who has taken photos of observation query images.
Evaluation results on test set is described in following table.

**Table 2.** Evaluation results of submitted runs

| | **Image classification score** | **Observation classification score** |
|---|---|---|
| **SNUMedinfo1** | 0.594 | 0.604 |
| **SNUMedinfo2** | 0.594 | 0.611 |
| **SNUMedinfo3** | 0.652 | 0.663 |
| **SNUMedinfo4** | 0.652 | 0.667 |

Performance was clearly better when five CNNs are combined for image classification (SNUMedinfo3 and SNUMedinfo4), compared to when only one CNN is used (SNUMedinfo1 and SNUMedinfo2). This is observed from both per image classification score and per observation classification score.
With regard to the rank aggregation methods used in observation classification, majority-voting based method showed slightly better performance compared to the Borda-fuse method, but the difference was negligible.

# 4 Discussion

## 4.1 CNN finetuning from other task model

In Chen et al.'s experiments (6) in last year, CNN trained without finetuning from other external dataset showed inferior performance, compared to their advanced feature encoding method (7) based on SIFT and Color Moments features. But when CNN is finetuned from ImageNet pretrained GoogLeNet, it was very effective, even though plant identification is targeted for finer-grained image classification task between different plant species compared to the ImageNet's general object category classification.

## 4.2 Combining CNN output

From table 2, we could observe that training multiple CNN and combining their outputs improve classification performance. As also experimented in (8), training and combining multiple CNN output method is considered to be effective to cope with CNN's variance.

## 4.3 Training plant part-specific CNN

In this task, each image is tagged with plant part name (e.g., stem, flower). We also tried dividing training set images according to the tagged part and training CNN per each part separately. But in our preliminary experiments, these part-specific image trained CNNs mostly showed no performance gain (similar or slightly worse performance, compared to when no part-specific training is used). So we chose not to use tagged plant part information for CNN training.

## 5 Conclusion

In LifeCLEF Plant identification task 2015', we applied GoogLeNet pretrained on ImageNet dataset for training by finetuning on the plant training set. Although task is more finer-grained image category classification compared to the ImageNet, and the number of plant species has doubled compared to the last year's plant task (9), classification performance was very effective. Also, training multiple CNNs and combining their output improved classification performance further. In our future study, we will explore other CNN architectural design options and different classification result combination methodologies.

## 6 References


1. Cappellato L, Ferro, N., Jones, G., and San Juan, E. CLEF 2015 Labs and Workshops. CEUR Workshop Proceedings (CEUR-WS.org); 2015.
2. Krizhevsky A, Sutskever I, Hinton GE, editors. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems; 2012.
3. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al. Going deeper with convolutions. arXiv preprint arXiv:14094842. 2014.
4. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, et al. Imagenet large scale visual recognition challenge. arXiv preprint arXiv:14090575. 2014.
5. Aslam JA, Montague M. Models for metasearch. Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval; New Orleans, Louisiana, USA. 384007: ACM; 2001. p. 276-84.
6. Chen Q, Abedini M, Garnavi R, Liang X, editors. Ibm research australia at lifeclef2014: Plant identification task. Working notes of CLEF 2014 conference; 2014.
7. Perronnin F, Dance C, editors. Fisher kernels on visual vocabularies for image categorization. Computer Vision and Pattern Recognition, 2007 CVPR'07 IEEE Conference on; 2007: IEEE.
8. Cireşan D, Giusti A, Gambardella L, Schmidhuber J. Mitosis Detection in Breast Cancer Histology Images with Deep Neural Networks. In: Mori K, Sakuma I, Sato Y, Barillot C, Navab N, editors. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013. Lecture Notes in Computer Science. 8150: Springer Berlin Heidelberg; 2013. p. 411-8.
9. Goëau H, Joly A, Bonnet P, Selmi S, Molino J-F, Barthélémy D, et al., editors. Lifeclef plant identification task 2014. CLEF2014 Working Notes Working Notes for CLEF 2014 Conference, Sheffield, UK, September 15-18, 2014; 2014: CEUR-WS.