Open-set Plant Identification Using an Ensemble of Deep Convolutional Neural Networks

Mostafa Mehdipour Ghazi[1], Berrin Yanikoglu[1], and Erchan Aptoula[2]

[1] Faculty of Engineering and Natural Sciences, Sabanci University, Istanbul, Turkey
[2] Institute of Information Technologies, Gebze Technical University, Kocaeli, Turkey
{mehdipour,berrin}@sabanciuniv.edu
eaptoula@gtu.edu.tr

Abstract. Open-set recognition, a challenging problem in computer vision, is concerned with identification or verification tasks where queries may belong to unknown classes. This work describes a fine-grained plant identification system consisting of an ensemble of deep convolutional neural networks within an open-set identification framework. Two well-known deep learning architectures of VGGNet and GoogLeNet, pre-trained on the object recognition dataset of ILSVRC 2012, are fine-tuned using the plant dataset of LifeCLEF 2015. Moreover, GoogLeNet is fine-tuned using plant and non-plant images for rejecting samples from non-plant classes. Our systems have been evaluated on the test dataset of PlantCLEF 2016 by the campaign organizers and our best proposed model has achieved an official score of 0.738 in terms of the mean average precision, while the best official score is 0.742.

Keywords: open-set recognition, plant identification, deep learning, convolutional neural networks

1 Introduction

Automated plant identification is a fine-grained image classification problem concerned with small inter-class and large intra-class variations. As with many other problems, recent research in this area has concentrated on deep learning schemes in which significant improvements have been obtained compared to traditional methods [1,2,3,4]. Raw data is fed into these networks in multiple levels, allowing the system to automatically discover high-level features for plant species recognition. However, due to their high computational complexity for training from scratch, the existing pre-trained deep networks are fine-tuned for plant identification purposes [2,3,5].

The plant identification challenge of the Conference and Labs of the Evaluation Forum (CLEF) [6,7,8,9,10] is one of the most well-known annual events that benchmark content-based image retrieval from structured plant databases including photographs of leaves, branches, stems, flowers, and fruits. The latest annotated plant dataset is provided by the LifeCLEF 2015 campaign with over 100,000 pictures of herbs, trees, and ferns belonging to 1,000 species collected from Western Europe.

Open-set recognition is a challenging task in computer vision which deals with identification or verification problems where samples from unknown classes could be presented to the system as well [11,12]. To create a similar scenario, the PlantCLEF 2016 campaign has provided a test image query which is very different from the previous years' CLEF datasets in nature [13,14]. This dataset is composed of unknown plant species and non-plant objects to create an open-world plant dataset. Therefore, the task challenges are not limited to automatically recognizing the known plant species, but also rejecting unknown plants and objects as well.

Our team participated in the PlantCLEF 2016 challenge under the name of *SabanciUGebzeTU* achieved the second rank with a very small difference from the first-ranked team. In our proposed systems, we fine-tune the pre-trained deep convolutional neural networks of GoogLeNet [15] and VGGNet [16] for plant identification using the LifeCLEF 2015 plant task datasets. We augment this data using different image transforms such as rotation, translation, reflection, and scaling to overcome overfitting while training and to improve the performance during testing the system on highly noisy test data. The overall system is then composed of these two networks using score-level averaging. To enable rejections, we train another deep learning system to separate plants from non-plants. Specifically, we fine-tuned GoogLeNet using the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) 2012 dataset as negative examples–after removing the potted plant category.

The rest of this paper is organized as follows. Section 2 describes the proposed methods based on fine-tuning of the GoogLeNet and VGGNet models for plant identification and open-set rejection, data augmentation, and classifier fusion. Section 3 is dedicated to the description of the utilized dataset and presentation of designed experiments and their results. The paper concludes in Section 4 with the summary and discussion of the utilized methods and obtained results.

2 Proposed Method

Our proposed method for automated plant identification is based on fine-tuning and fusing of two successful deep learning models, i.e. GoogLeNet [15] and VGGNet[16]. These models are, respectively, the first-ranked and second-ranked architectures of the ILSVRC 2014–both trained on the ILSVRC 2012 dataset with 1.2 million labeled images of 1,000 object classes.

GoogLeNet [15] contains 57 convolutional layers, 14 pooling layers, and one fully-connected layer while VGGNet [16] involves 16 convolutional layers, five pooling layers, and three fully-connected layers. Both networks take color image patches of size 224×224 pixels as the input and connect a linear layer with Softmax activation in the output.

We used the plant task datasets of LifeCLEF 2015 for fine-tuning the pre-trained models. The training parameters consist of a weight decay of 0.0002, a base learning rate of 0.001, and a batch size of 20. Data augmentation is applied to decrease the chance of overfitting during training and to improve performance

while testing. For this purpose, we randomly extract K square patches from each original image around its center. The original image is also rotated by $\pm R$ degrees and the largest square image is cropped from the center of each rotated image. All the extracted patches as well as the original image ($K + 3$ images) are then scaled to 256×256 pixels and the mean image is subtracted from them. Finally, five patches of size 224×224 pixels are extracted from four corners and the center of each image. These patches are then also reflected horizontally, resulting in $10 \times (K + 3)$ patches per input image in total.

Score-level averaging is applied to combine the prediction scores assigned to each class for all the augmented patches within a single network. Finally, the obtained scores from every deep network classifier are combined using the same averaging rule. The *observation id* that indicates different images of the same unique plant is not used in our systems; hence the recognition results are based only on single images.

For the unknown-class rejection task within the unseen data, we separately fine-tuned the GoogLeNet model for a binary classification problem, i.e. plants vs. non-plants. The training samples for this purpose consist of plant images from LifeCLEF 2015 and non-plant images from ILSVRC 2012. Using the obtained results from the binary classifier, we reject those images that are classified as non-plant *and* receive a low confidence score in the main plant identification system.

3 Experiments and Results

To train and validate our systems, we used the plant dataset of LifeCLEF 2015 [17] consisting of 113,204 images of different plant organs belonging to 1,000 species of trees, herbs, and ferns. We randomly divided the *training portion* of the dataset into two subsets for training and validation, with 70,904 and 20,854 images, respectively. The *test portion* of the dataset consists of a separate set of 21,446 images; however the ground-truth for the test dataset was only recently released and thus used in a limited fashion in our system. In the remainder of this paper, we will call these three subsets train, validation, and test subsets, respectively.

The PlantCLEF 2016 test dataset includes 8,000 samples submitted by the users of the mobile application Pl@ntNet [13]. This dataset is highly noisy and essentially different from the plant task datasets of LifeCLEF 2015 since it contains pictures of unknown plants and objects. We use the data augmentation approach explained in Section 2 and set $K = 5$ and $R = 10$ to train and test 80 patches per input image. In some experiments, we set $K = 9$ and $R = \{10, 20\}$ to train and test 140 patches per image.

To fine-tune GoogLeNet for the open-set unknown-class rejection task, K is set to zero and no rotation is applied so as to augment data with only 10 patches per image. The training data for this problem is obtained by combining the earlier training subset of LifeCLEF 2015 for the plant class samples and an equal number of non-plant object samples from ILSVRC 2012, giving about 140,000

samples in total. In addition, the validation set is obtained from combining the earlier validation subset of LifeCLEF 2015 and an equal number of non-plant object samples of ILSVRC 2012, resulting in about 40,000 samples in total. Images used in the training and validation subsets contained distinct samples.

We implemented the GoogLeNet and VGGNet deep models using the Caffe deep learning framework [18] with pre-trained weights obtained from Caffe Model Zoo provided by the Berkeley Vision and Learning Center (BVLC). In the rest of this section, we explain our conducted experiments, their validation results, and the prepared runs.

It is worth mentioning that we have not used additional plant images for training any of the systems. Furthermore, all systems are fully automatic except for Run 3 where we manually remove 90 images in order to measure our performance in the case of a closed set problem (no query from unknown classes).

Run 1. In this run, we first fine-tuned the pre-trained models of GoogLeNet and VGGNet using the train subset of LifeCLEF 2015 that we augmented with 140 and 80 patches per image till 600,000 and 500,000 iterations, respectively. Next, we fine-tuned the pre-trained GoogLeNet with almost all of the data (train+test+validation/2) augmented with 80 patches per image until 200,000 iterations. We fused all the obtained scores for augmented image patches from different networks' classifiers. With this system, we achieved an accuracy of 79.80% on the given validation set (validation/2).

To reject test samples of unknown classes, we fine-tuned the pre-trained GoogLeNet with the obtained plant/non-plant dataset until 100,000 iterations and achieved an accuracy of almost 100% on the given validation set. Next, we tested the combined deep model on the non-plant validation set; from the obtained plant identification prediction scores it was inferred that these scores have a uniform-like distribution. We set the rejection threshold to $T = 0.4$ based on the top-1 score for maximal rejection of non-plant samples and minimum rejection of plant data. In the test stage, we utilized our combined system for score prediction as well as rejection of samples whose top-1 score was less than 0.4. We rejected 480 images from the PlantCLEF 2016 test dataset in this manner and obtained our first run.

Run 2. All the implemented steps for achieving this run were the same as those performed for Run 1. The only difference was that no image rejection was performed in obtaining this run.

Run 3. All the steps utilized in preparing this run were common with Run 1 except for the rejection method. In this experiment, we manually reviewed the test images and set aside 90 non-plant images. This was done to measure the system performance with only known classes; however, as this score is later provided by campaign organizers, this system is not very relevant.

Test Results. We submitted the classification results of the aforementioned systems on the official test set of the PlantCLEF 2016. The utilized metric for evaluation was the mean average precision. In other words, all prediction scores were extracted for each class and sorted in the descending order to compute the average precision for that class. Finally, the mean was computed across all classes. The released results by the challenge organizers [14] are shown in Figure 1. As the officials results indicate, we obtained the second place in the open-set plant identification competition of PlantCLEF 2016. Moreover, the unofficial results indicate that we achieved the top score in the image-based plant identification.

From the official results, we can conclude that the automatic rejection method (Run 1) was the top performer among our three systems followed by the manual rejection scheme (Run 3) and the method with no rejection (Run 2).

Comparing automatic rejection (Run 1) and no rejection (Run 2) systems, we see that the error in automatic rejection is very small when considering only known classes (0.806 vs 0.807), while the gains in the open-set and invasive species problems are slightly more significant (0.738 vs 0.736 and 0.704 vs 0.683, respectively).

4   Conclusions

In this paper, we reviewed the details of our proposed systems used in the open-set plant identification problem of PlantCLEF 2016 campaign. We fused two powerful deep learning methods of VGGNet and GoogLeNet after fine-tuning them with the augmented plant task datasets of LifeCLEF 2015. Meanwhile, we used plant and non-plant images to fine-tune GoogLeNet for a binary classification to perform unknown-class rejection. Our systems were officially evaluated on the test dataset of PlantCLEF 2016 and our best proposed model achieved a mean average precision of 0.738.

Our main focus in PlantCLEF 2016 was to see how we could improve the competitive results obtained in LifeCLEF 2015 [10]. To this end, we studied the best ways to fine-tune deep learning models; in particular, we experimented with the iteration size, batch size, and data augmentation. We observed that the accuracy steadily improved with the increased number of iterations and with data augmentation.

References

1. Chen, Q., Abedini, M., Garnavi, R., Liang, X.: IBM research Australia at LifeCLEF 2014: Plant identification task. In: CLEF (Working Notes). (2014)

2. Choi, S.: Plant identification with deep convolutional neural network: SNUMedinfo at LifeCLEF plant identification task 2015. In: CLEF (Working Notes). (2015)

3. Champ, J., Lorieul, T., Servajean, M., Joly, A.: A comparative study of fine-grained classification methods in the context of the LifeCLEF plant identification challenge 2015. In: CLEF (Working Notes). (2015)

4. Ghazi, M.M., Yanikoglu, B., Atoula, E., Muslu, O., Ozdemir, M.C.: Sabanci-Okan system in LifeCLEF 2015 plant identification competition. In: CLEF (Working Notes). (2015)

5. Ge, Z., McCool, C., Sanderson, C., Corke, P.: Content specific feature learning for fine-grained plant classification. In: CLEF (Working Notes). (2015)

6. Goëau, H., Bonnet, P., Joly, A., Boujemaa, N., Barthelemy, D., Molino, J.F., Birnbaum, P., Mouysset, E., Picard, M.: The CLEF 2011 plant images classification task. In: CLEF (Notebook Papers/Labs/Workshop). (2011)

7. Goëau, H., Bonnet, P., Joly, A., Yahiaoui, I., Barthelemy, D., Boujemaa, N., Molino, J.F.: The ImageCLEF 2012 plant identification task. In: CLEF (Online Working Notes/Labs/Workshop). (2012)

8. Goëau, H., Bonnet, P., Joly, A., Bakic, V., Barthelemy, D., Boujemaa, N., Molino, J.F.: The ImageCLEF 2013 plant identification task. In: CLEF (Working Notes). (2013)

9. Goëau, H., Joly, A., Bonnet, P., Selmi, S., Molino, J.F., Barthelemy, D., Boujemaa, N.: LifeCLEF plant identification task 2014. In: CLEF (Working Notes). (2014)

10. Goëau, H., Bonnet, P., Joly, A.: LifeCLEF plant identification task 2015. In: CLEF (Working Notes). (2015)

11. Scheirer, W.J., Jain, L.P., Boult, T.E.: Probability models for open set recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence **36**(11) (2014) 2317–2324

12. Bendale, A., Boult, T.: Towards open world recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2015) 1893–1902

13. Joly, A., Goëau, H., Glotin, H., Spampinato, C., Bonnet, P., Vellinga, W.P., Champ, J., Planqué, R., Palazzo, S., Müller, H.: LifeCLEF 2016: multimedia life species identification challenges. In: Proceedings of CLEF 2016. (2016)

14. Goëau, H., Bonnet, P., Joly, A.: Plant identification in an open-world (LifeCLEF 2016). In: CLEF (Working Notes). (2016)

15. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: IEEE Conference on Computer Vision and Pattern Recognition. (2015)

16. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. Computing Research Repository (CoRR) (2014) arXiv: 1409.1556.

17. Joly, A., Goëau, H., Spampinato, C., Bonnet, P., Vellinga, W.P., Planqué, R., Rauber, A., Palazzo, S., Fisher, B., Müller, H.: LifeCLEF 2015: multimedia life species identification challenges. In: Experimental IR Meets Multilinguality, Multimodality, and Interaction. (2015) 462–483

18. Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T.: Caffe: Convolutional architecture for fast feature embedding. In: Proceedings of the 22nd ACM International Conference on Multimedia. (2014) 675–678
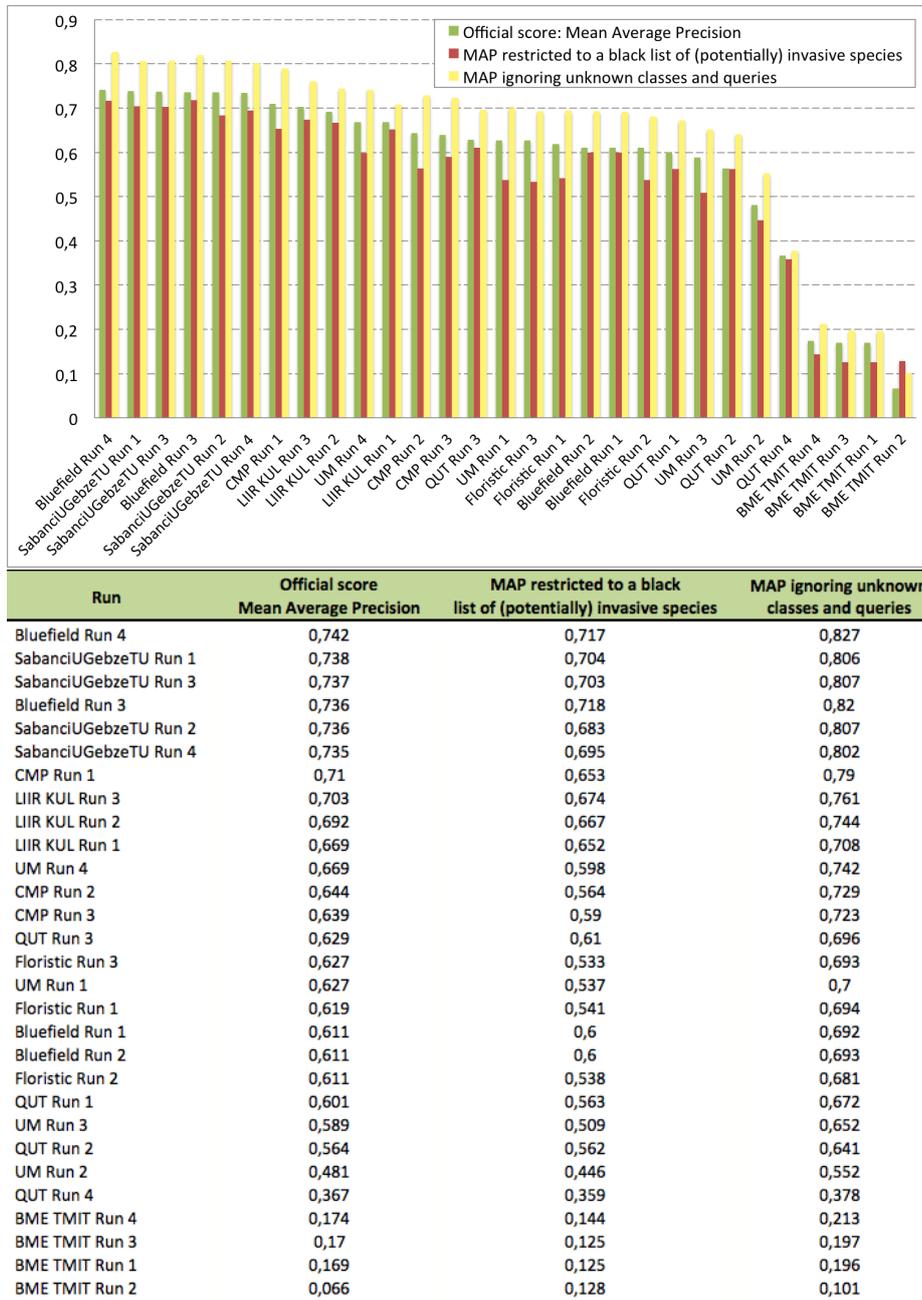
Run	Official score Mean Average Precision	MAP restricted to a black list of (potentially) invasive species	MAP ignoring unknown classes and queries
Bluefield Run 4	0,742	0,717	0,827
SabanciUGebzeTU Run 1	0,738	0,704	0,806
SabanciUGebzeTU Run 3	0,737	0,703	0,807
Bluefield Run 3	0,736	0,718	0,82
SabanciUGebzeTU Run 2	0,736	0,683	0,807
SabanciUGebzeTU Run 4	0,735	0,695	0,802
CMP Run 1	0,71	0,653	0,79
LIIR KUL Run 3	0,703	0,674	0,761
LIIR KUL Run 2	0,692	0,667	0,744
LIIR KUL Run 1	0,669	0,652	0,708
UM Run 4	0,669	0,598	0,742
CMP Run 2	0,644	0,564	0,729
CMP Run 3	0,639	0,59	0,723
QUT Run 3	0,629	0,61	0,696
Floristic Run 3	0,627	0,533	0,693
UM Run 1	0,627	0,537	0,7
Floristic Run 1	0,619	0,541	0,694
Bluefield Run 1	0,611	0,6	0,692
Bluefield Run 2	0,611	0,6	0,693
Floristic Run 2	0,611	0,538	0,681
QUT Run 1	0,601	0,563	0,672
UM Run 3	0,589	0,509	0,652
QUT Run 2	0,564	0,562	0,641
UM Run 2	0,481	0,446	0,552
QUT Run 4	0,367	0,359	0,378
BME TMIT Run 4	0,174	0,144	0,213
BME TMIT Run 3	0,17	0,125	0,197
BME TMIT Run 1	0,169	0,125	0,196
BME TMIT Run 2	0,066	0,128	0,101

Fig. 1. The official released results of PlantCLEF 2016