Plant Leaf Recognition

Albert Liu, albertliu@stanford.edu Yangming Huang, yangming@standford

I. Introduction

The number of plant species is estimated to be over 220000 in the world [1]. Automatic plant species recognition with image processing has gained increasing interests recently. The main application are crop/weeds identification, plan biology research and species tracking [3]. We discuss a implementation that is suitable to identify leaf based on camera picture captured from natural background.

II. Related Work

III. Approach

A. Overview

We break down the problem into two pieces. Firstly, we want to locate the leaf in the image which is an object detection problem. This becomes more important if the leaf is not the single salient object in the image. We use typical method which is to scan the image and determine where the match may occur by a constructed binary classifier. The variations of leaves can this harder. Then, optionally we can apply background removal on the detected image. Finally, we identify the leaf to one of the varied classes, which is a category recognition problem. For which we explore traditional method Bag of Feature method and CNN.

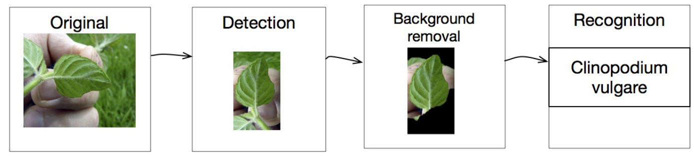

Fig. 1. Pipeline of the system

The subsequent sections are organized as following

IV. DataSet

We found these datasets.

1) UCI leaf dataset [4]: 40 species with 5 to 16 samples per specie
2) Kaggle leaf dataset[3]: 99 species with 16 samples per specie
3) Swedish leaf dataset [5]: 15 species with 952 samples (roughly 60 samples per specie)
4) Flavia leaf dataset [6]: 33 species with roughly 60 samples per specie
5) Pl@ntNet [7]: the most challenging dataset as images are collected through crowd sourced application. 71 species with 5436 images.

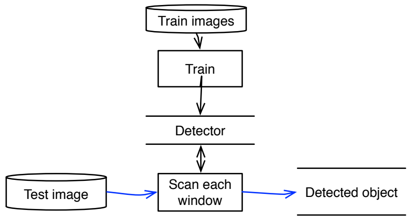

Fig. 2. Leaf Detection

A. Object Detection with HOG

V. Object Detection & Segmentation

A. Object Detection with CNN

B. Object Segmentation

VI. Object recognition

A. Bag of Features (BoF) + local feature descriptors

Due to the simplicity and performance, this well established approach was taken at first. Interest points are detected from the raw images and then local invariant feature descriptors are collected, which are clustered to form the visual vocabulary/codebook. Afterwards, each raw image can be represented with histograms of visual words, i.e. term vectors. We have prototyped a BoF system, based on the OpenCV package. Here is the illustration of the system 3.

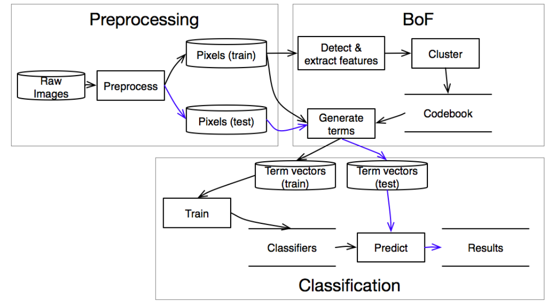

Fig. 3. System Design for BoF

During preprocessing, raw images are converted to gray-scaled images and resized to reduce computation complexity. We extract SIFT descriptors from the pixels after detecting the key points. Limiting the width of image to 128 pixels, we have roughly 200 SIFT descriptors per image. Then all the descriptors are clustered to build visual words via K-Means. Due to computation complexity, we pick randomly 100 training images to build the visual vocabulary for the initial run.

B. CNN

VII. Results

VIII. Conclusion

Acknowledgment

The authors would like to thank...

References

[1] Charles Mallah, James Cope, James Orwell. Plant Leaf Classification Using Probabilistic Integration of Shape, Texture and Margin Features. Signal Processing, Pattern Recognition and Applications, in press. 2013

[2] Itheri Yahiaoui, Nicolas Herve, and Nozha Boujemaa. Shape-based image re-trieval in botanical collections, Lecture Notes in Computer Science including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinfor matics, vol. 4261 LNCS, pp 357-364, 2006.

[3] Pedro F. B. Silva, Andre R.S. Marcal, Rubim M. Almeida da Silva. Evaluation of Features for Leaf Discrimination. 2013. Springer Lecture Notes in Computer Science, Vol. 7950, 197-204.

[4] https://archive.ics.uci.edu/ml/datasets/Leaf

[5] Oskar J. O. Sderkvist. Computer vision classifcation of leaves from swedish trees. Master's Thesis, Linkoping University, 2001.

[6] Stephen Gang Wu, Forrest Sheng Bao, Eric You Xu, Yu-Xuan Wang, Yi-Fan Chang and Chiao-Liang Shiang, A Leaf Recognition Algorithm for Plant classification Using Probabilistic Neural Network, IEEE 7th International Symposium on Signal Processing and Information Technology, Dec. 2007, Cario, Egypt

[7] http://www.imageclef.org/2011/Plants