

Plant leaf classification

Proposal

Team members

Albert Liu/SUID : 06167442
Yangming Huang/SUID : 06168334

albertpl@stanford.edu
yangming@stanford.edu

1 Introduction

Automatic plant species recognition with image processing has gained increasing interests recently. The main application are crop/weeds identification, plant biology research and species tracking [2]. The number of plant species is over 220000 and leaf images are considered the most available and effective attributes for such identification. This problem can be seen as a multi-class classification with relatively small training samples (raw image of leaf).

The main goal of this project is to apply various algorithms and techniques we learned in this class to solve this problem. First step is to use KNN on the pre-extracted shape, margin, texture features to establish baseline. This is because most of existing works [1] [2] suggest KNN and the aforementioned features are readily available from the data sets. Secondly we will approach with a set of models, including Linear SVM, Quadratic SVM, Linear Discriminative Analysis and others. Then we compare the results of different algorithms. Lastly we will explore to extract customer features from the raw images and apply feature extraction/selection techniques.

The data set

1. UCI [3] 40 species with 5 to 16 samples per species
2. kaggle [4] 99 species with 16 samples per species

The feature extraction techniques are discussed in [2].

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

- [3] <https://archive.ics.uci.edu/ml/datasets/Leaf>
- [4] <https://www.kaggle.com/c/leaf-classification/data>