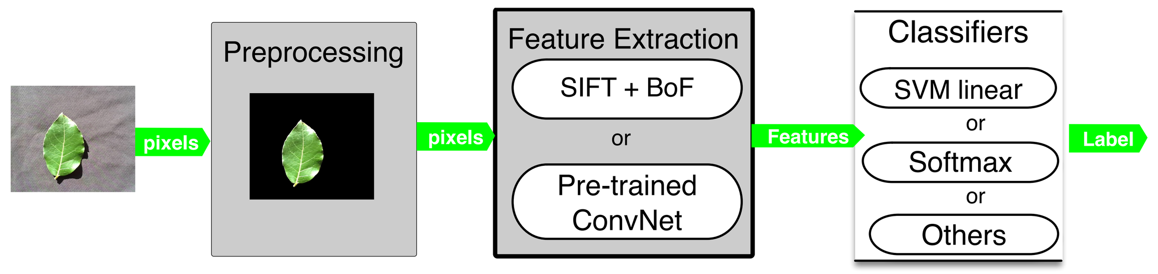
PLANT LEAF RECOGNITION

{ ALBERT LIU AND YANGMING WANG }@STANFORD.EDU

PROBLEM

Fine-grained leaf recognition with image processing has important application in weeds identification, species discovery, plant taxonomy, and so on. The subtle differences between species and the sheer large number of categories makes it hard to solve.

METHOD



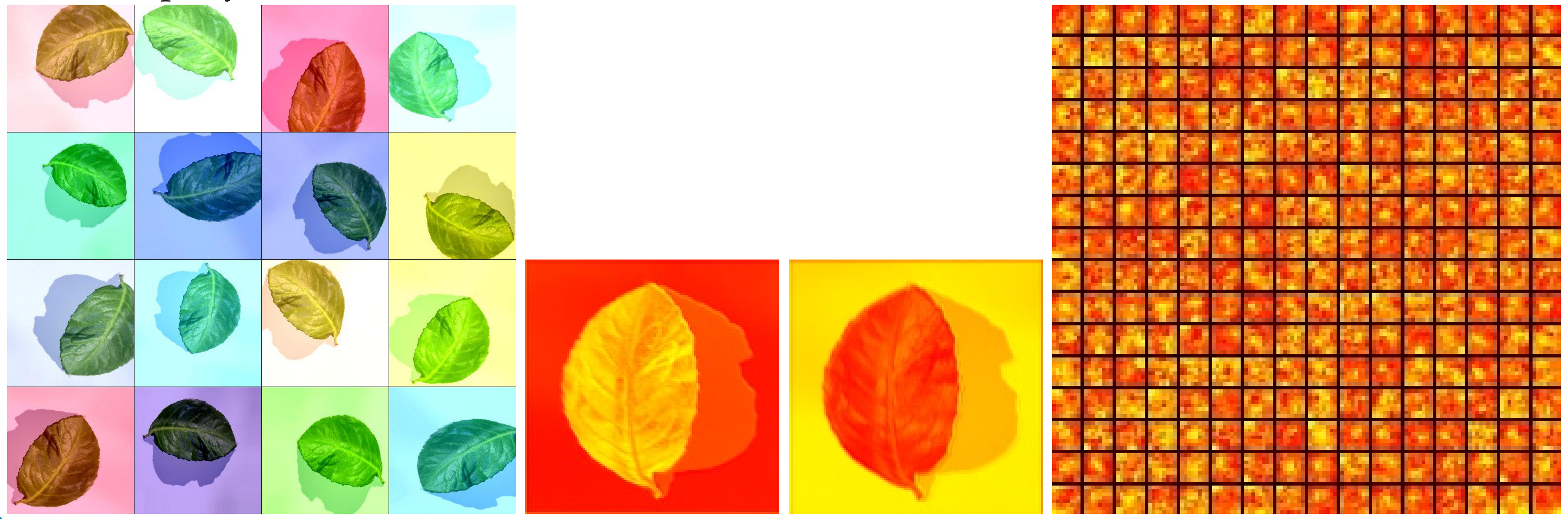
Firstly, during preprocessing, we apply Contrast Limited Adaptive Histogram Equalization to reduce lighting condition variation and raw images are resized to fit to the next layer. In traditional approach, we also use K-means to remove background color heuristically. For challenging dataset, we find the convex hull containing the largest N contours and then use GrabCut to segment leaf out of the background clutter. Next we extract features and we explore two options,

1. ConvNets. We take transfer learning approach. see right
2. SIFT + Bag of Features Key points are densely sampled and SIFT feature descriptors are retrieved from each raw images. Then

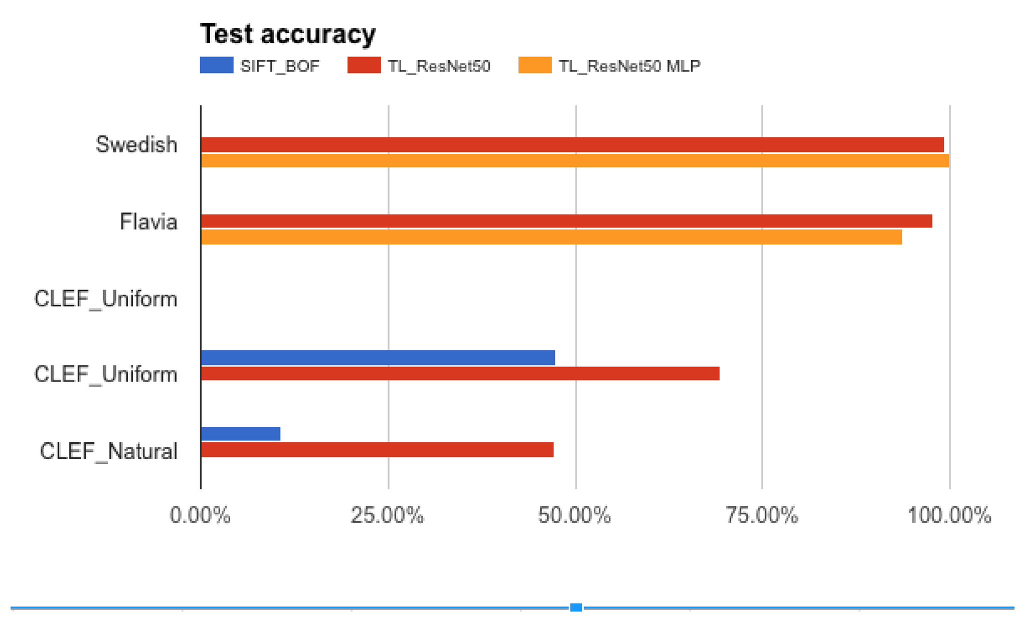
At last, we train a classifier, using the feature vectors from the above step and predict labels for our test data.

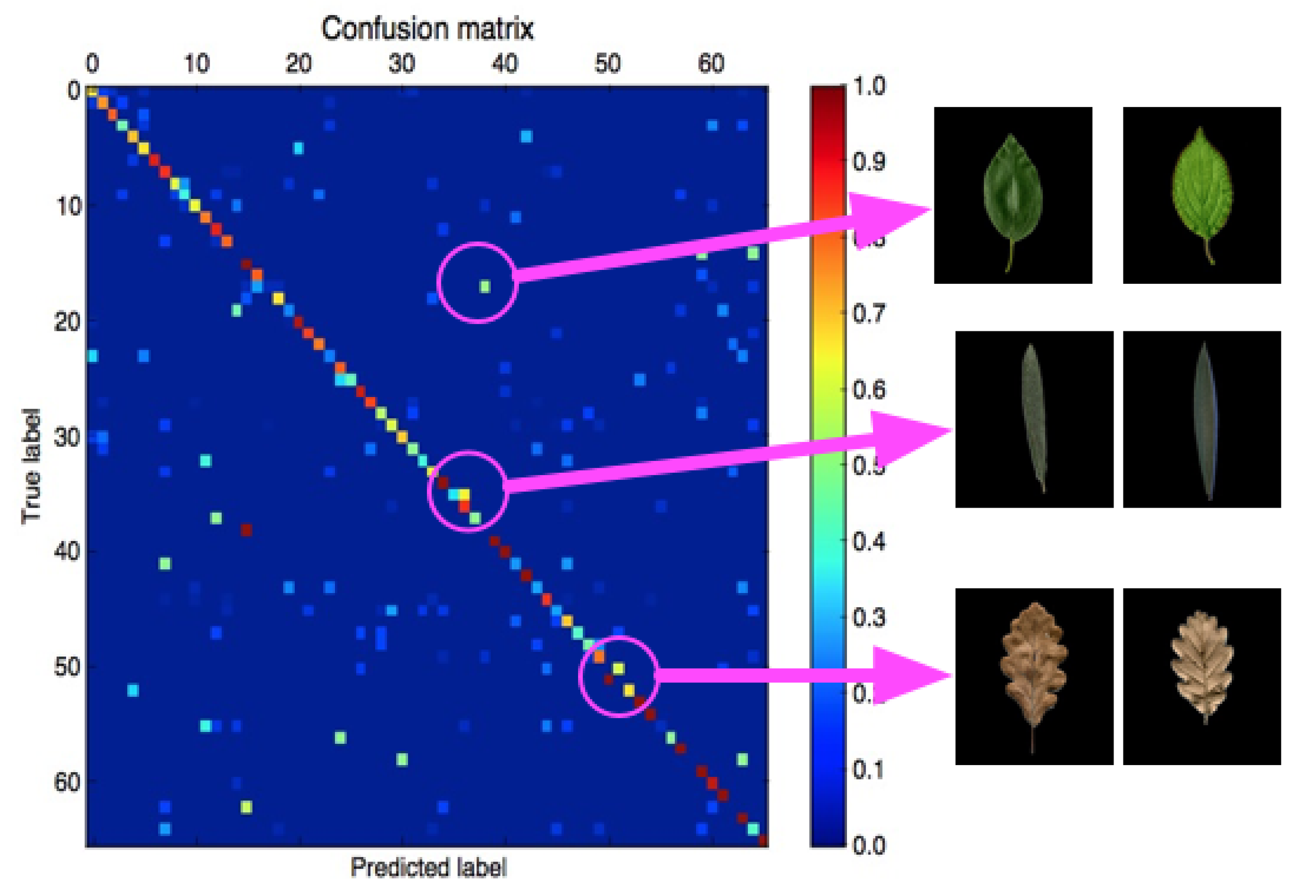
TRANSFER LEARNING - CONVNET

[Explain architecture and give more insights for ConvNet]
With the power of ConvNets with the constrains of time and computations. Specifically, we take a couple of ConvNets that are pretrained with ImageNet for ILSVRC object image classification task, remove one or two top layers and then treat the rest of the ConvNet as fixed feature extractor for the our data set.



RESULTS





DATASET

1. Clean images: little or no luminance or color variations), e.g. Swedish leaf dataset.
2. Challenging images: considerable variations on lighting conditions, viewpoints, background clutters and occlusions, no bounding box, e.g. ImageCLEF which is crowd sourced. (To effectively evaluate our approaches, we use subsets of this dataset.)

DISCUSSIONS

1. As expected, pre-trained CNN codes yields better accuracy than SIFT+BOF off the shelf. Particularly traditional method is more sensitive to background clutter.

2. Looking at the confusion matrix, we believe the causes for top misclassification

   - Very fine differences between species which are even hard for humans
   - Noisy and plausibly not-representative train data lead to overfitting,
   - The ConvNet models are pre-trained with different dataset and task

3. Image augmentation/normalization help can noise (background clutter, color/viewpoint variations).

FUTURE WORKS

1. Given enough time, we would approach more data (augmented or retrieve elsewhere) and fine-tune the ConvNet to solve the overfitting problem.

2. For the most challenging data set, i.e. Image CLEF natural images, we can segment the leaf out of background clutter first, which is an single object localization problem. We believe this would help as there are multiple objects and sometime leaf is not the most salient object for this dataset.